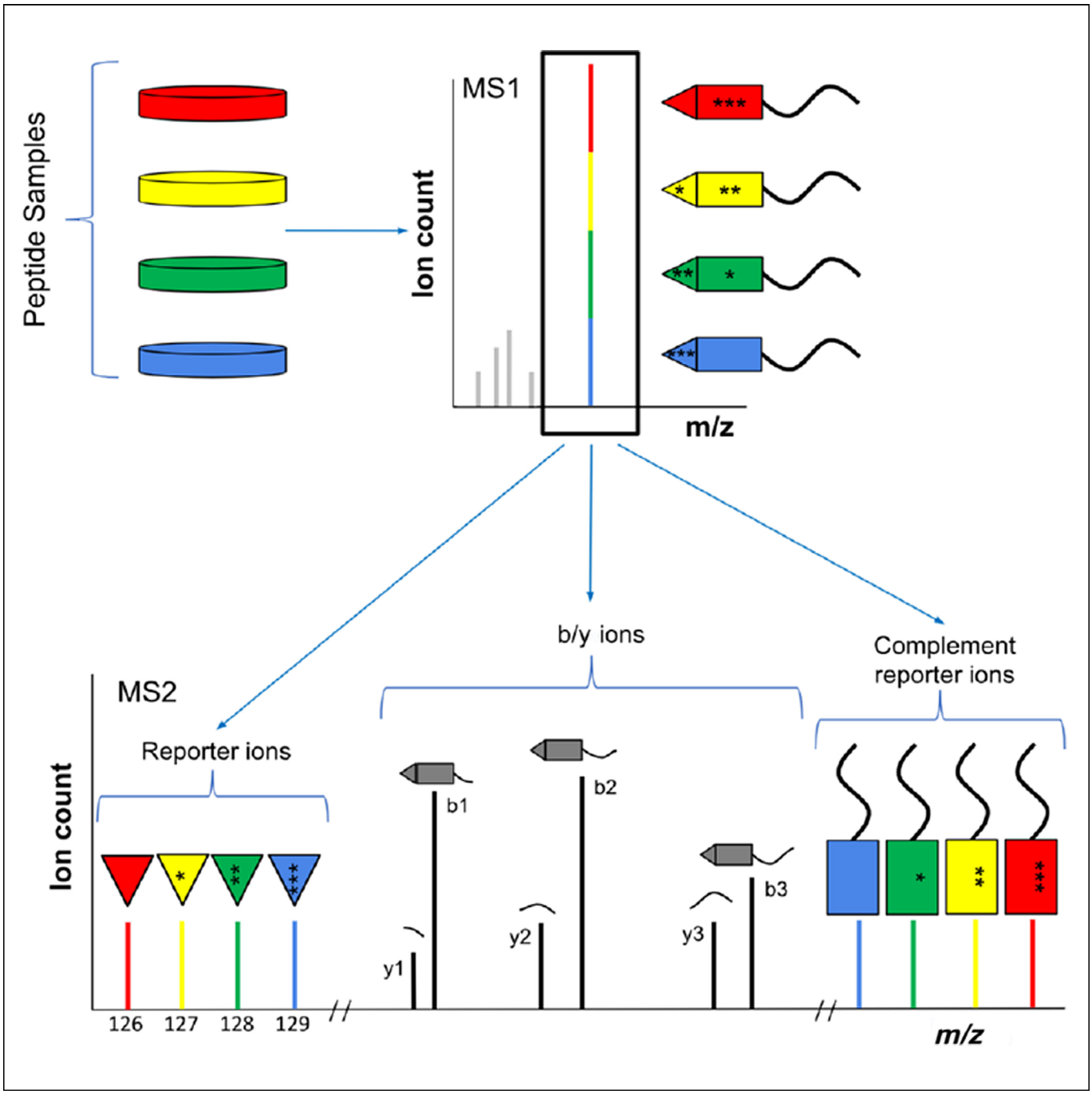
ChemPubSoc Europe
DOI: 10.1002/cbic.201800650
CHEMBIOCHEM Reviews

A Review on Quantitative Multiplexed Proteomics

Nishant Pappireddi,[a, b] Lance Martin,[a, b] and Martin Wühr*[a, b]



Over the last few decades, mass spectrometry-based proteomics has become an increasingly powerful tool that is now able to routinely detect and quantify thousands of proteins. A major advance for global protein quantification was the introduction of isobaric tags, which, in a single experiment, enabled the global quantification of proteins across multiple samples. Herein, these methods are referred to as multiplexed proteomics. The principles, advantages, and drawbacks of various multiplexed proteomics techniques are discussed and compared with alternative approaches. We also discuss how the emerging combination of multiplexing with targeted proteomics might enable the reliable and high-quality quantification of very low abundance proteins across multiple conditions. Lastly, we suggest that fusing multiplexed proteomics with data-independent acquisition approaches might enable the comparison of hundreds of different samples without missing values, while maintaining the superb measurement precision and accuracy obtainable with isobaric tag quantification.

1. Background on Protein Identification and Quantification in Mass Spectrometry-Based Shotgun Proteomics

Throughout this paper, we refer to proteomics methods that use isobaric tags to analyze multiple protein samples as multiplexed proteomics. Multiplexed proteomics builds on decades of technological development in proteomics prior to isobaric tags. To put multiplexed experiments in context, we begin this review with an overview of protein identification methods and alternative quantitative approaches. We then cover multiplexed proteomics techniques, which involve the use of isobaric labeling. Finally, we discuss how merging multiplexed proteomics with other quantification strategies might help to overcome current technical limitations.

Throughout this review, we only discuss bottom-up proteomics, in which proteins are first digested into peptides and the peptides are analyzed with a mass spectrometer. An entire field is devoted to the analysis of intact proteins through mass spectrometry (MS), known as top-down proteomics. For these approaches, we refer the reader to excellent reviews published elsewhere.[1]

1.1. Peptide identification in shotgun proteomics

Proteomic analysis is typically performed by means of liquid chromatography–mass spectrometry (LC-MS).[2] For shotgun proteomics, protein samples derived from cell or tissue lysate are digested into peptides with proteases, such as trypsin (Figure 1A).[3] Trypsin cleaves, with fairly high specificity, protein peptide bonds at the C termini of arginine and lysine residues.[4] To better probe the complexity of the proteolytic mixture, tryptic peptides can be fractionated into multiple samples based on properties such as charge, size, polarity, or hydrophobicity.[5] The peptides in each sample are then separated by means of LC. The peptides elute through a thin opening from the column directly in front of the mass spectrometer. Voltage between the column opening and the inlet of the mass spectrometer leads to a process called electrospray ionization (ESI). The eluting droplets undergo evaporation, concentrating positively charged peptides until coulombic repulsion overwhelms surface tension and the droplets explode, resulting in charged peptides in the gas phase.[6]

The efficiency of the ionization process can differ by orders of magnitude for different peptides.[7] Additionally, the efficiency of protein conversion into analyzable peptides can also vary drastically due to different digestion efficiencies and/or peptide solubility. Therefore, the number of ions inside the mass spectrometer is not a direct readout of how many proteins were originally in the sample. Because of this problem, MS is an inherently nonquantitative method and significant additional efforts are required to obtain quantitative information.

Peptide molecules ionize before entering the mass spectrometer, where researchers can detect or filter them based on their mass-to-charge (m/z) values. Figure 1B shows the chromatogram of the most abundant ion species collected during a typical experiment lasting about 2 h. We call a spectrum of all the intact peptides eluting at a given time an MS1 spectrum (Figure 1C). The height of each signal reflects the number of detected ions.[8] The human genome encodes approximately 20000 proteins, resulting in approximately 10^6 possible tryptic peptides. With the resolving power of current mass analyzers, it is not possible to identify peptides solely based on their intact masses. However, it is possible to fragment peptide ions in a mass spectrometer at the weakest bonds (usually the peptide bond between amino acids) by colliding them with inert gases. The resulting fragment ions are analyzed in an MS2 or MS/MS spectrum, which can be used to identify the amino acid sequence and to detect post-translational modifications, such as phosphorylation (Figure 1D).[9] By convention, the fragment ions containing the N terminus of the peptide are called b ions, and the fragments containing the C terminus are called y ions.[10]

If all of the b or y ions were formed and detected, then the differences in m/z values would allow the amino acid sequence of the peptide to be determined de novo because each genetically encoded amino acid has a different molecular weight (except leucine and isoleucine). However, the situation is typi-

[a] N. Pappireddi, Dr. L. Martin, Prof. M. Wühr
Department of Molecular Biology
Princeton University, Icahn Laboratory, #246
Washington Road, Princeton, NJ 08544 (USA)
E-mail: wuhr@princeton.edu

[b] N. Pappireddi, Dr. L. Martin, Prof. M. Wühr
The Lewis-Sigler Institute for Integrative Genomics, Princeton University
Princeton, NJ 08544 (USA)

Supporting information and the ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201800650.

This article is part of the young researchers' issue ChemBioTalents. To view the complete issue, visit http://chembiochem.org/chembiotalents

cally not this ideal, and de novo peptide sequencing is not practical for most spectra obtained from actual experiments. Instead, the analyzed samples typically come from organisms for which we know all possible protein sequences, and hence, all possible peptides. Rather than having to sequence peptides de novo from spectra, we typically only need to find the most likely match to known amino acid sequences. To this end, theoretical MS2 spectra are created for these peptides, based on all b and y ions that can result from fragmentation. By comparing the MS1 mass and the corresponding observed MS2 spectrum to the theoretical spectra of possible peptides, the best match can be found, resulting in the peptide being assigned to the spectrum.[11] Multiple search algorithms are available to automatically perform this analysis.[11,12] Moreover, various machine-learning strategies have been developed to confidently assign spectra to peptides and proteins.[13]

Nishant Pappireddi is a PhD student in the molecular biology program at Princeton University, working in the Wühr Lab. He graduated from UC Berkeley with a BA in molecular & cell biology and mathematics. His research aims to determine the molecular mechanisms affecting protein partitioning between the nucleus and cytoplasm of the *Xenopus* oocyte. To perform these studies, he uses quantitative multiplexed proteomics techniques.

Lance Martin obtained his PhD in biochemistry and macromolecular biophysics from Caltech. With Pamela Bjorkman, he solved the crystal structure of the neonatal Fc receptor in complex with a heterodimeric form of the Fc protein. He joined Pehr Harbury's lab at Stanford Biochemistry for his postdoc, working on the synthesis of chemical crosslinking reagents. After his postdoc, he founded the company Martin-Protean. Recently, he joined the Wühr Lab to work on quantitative proteomics.

Martin Wühr studied biochemistry at TU Munich. During his PhD at Harvard in the laboratory of Tim Mitchison, he studied how the cell division machinery adapted to drastically changing cell sizes in early vertebrate development. During his postdoc with Steven Gygi and Marc Kirschner, he invented the complement reporter ion approach for multiplexed proteomics and used proteomics to study nucleocytoplasmic partitioning. In 2016, he started his lab at Princeton University. The Wühr Lab aims to advance proteomics technology and applies quantitative proteomics to study cellular organization.

1.2. Absolute and relative quantification in proteomics experiments

Given the intrinsic quantification limitations of MS in quantitative proteomics, we distinguish between absolute quantification (determining the absolute concentration of a protein in a sample) and relative quantification (determining the relative ratio of the amounts of a given protein in different samples). As discussed earlier, the efficiencies of turning protein concentrations into MS signals are nonuniform and currently unpredictable. The signal in the mass spectrometer is therefore only an indirect readout for the abundance of a peptide in solution.

For the relative quantification of proteins between two or more samples, their peptides must first be relatively quantified from each sample and the data from multiple peptides integrated to obtain a ratio for the overall protein.[14] The signal size corresponding to a peptide is proportional to the number of peptide ions detected by the instrument. Because the ionization efficiencies of different peptides are different, it is not possible to directly compare the MS signal of different peptides to determine their abundance in a sample. However, it is possible to compare signals of the same peptide, with the same ionization efficiency, in different samples, which is the basis for relative quantification.[15] All of the methods described below use this as the basis for relative quantification of peptides, and ultimately proteins.

Absolute quantification in proteomics is usually an extension of relative quantification methods that quantify relative to an added spiked-in standard, with known absolute concentration. Due to the high costs of such standards, these experiments are typically limited to studies of a smaller subset of the proteome.[16] Although not directly correlated, the total ion signal of a protein seems to be related to its in vivo abundance through a power law.[17] By using an internal standard, the absolute protein abundance for all proteins detected in a sample can therefore be inferred, but this comes with a wide approximately twofold error.[17b]

1.3. MS1-based, label-free quantification

Currently, the most widely used form of quantitative proteomics is based on quantifying the MS1 signal of peptides obtained from tryptic digestion. Because there is no attempt to covalently modify the peptides, this version of quantitative proteomics is often referred to as label-free. Label-free proteomics involves running different samples consecutively (Figure 2A).[18] The MS1 signal for a given peptide is integrated over time from all MS1 spectra in which it can be observed (Figure 2B, C). The obtained area under the curve is a measure of the total number of ions for a given peptide. Although this area is not a good readout for the absolute amount of peptide

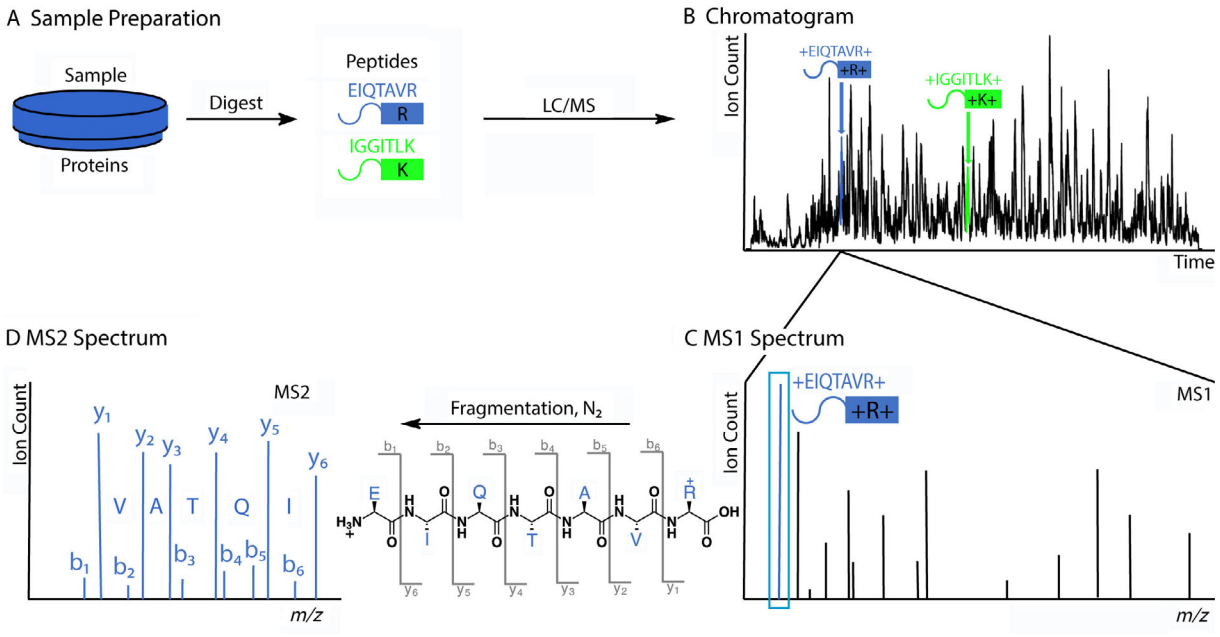


Figure 1. Outline of peptide identification with shotgun proteomics. A) A sample of proteins is digested by trypsin, which cleaves peptide bonds at the C terminus of lysine and arginine residues. The example peptide, EIQTAVR, which we follow throughout this figure, is shown in blue. To reduce complexity, the peptides are separated by LC, ionized by electrospray, and injected into the mass spectrometer. B) The chromatogram of the most abundant signal at each retention time. The blue and green peptides elute at different retention times. C) At any given time, for example, when the blue peptide elutes, multiple different peptides coelute. The mass spectrometer can typically distinguish them by their mass to charge ratio (*m/z*). The mass spectrum of the intact peptides is called the MS1 spectrum. The signal corresponding to the peptide EIQTAVR is highlighted in blue. D) In complex mixtures, mass alone is not enough for peptide identification. Inside the mass spectrometer, a signal corresponding to a peptide is isolated and fragmented by collision with an inert gas. The *m/z* values of the fragment ions, derived from the blue peptide, are recorded in the MS2 spectrum. By convention, peptide fragments containing the N terminus are called b ions, whereas fragments with the C terminus are called y ions. The characteristic masses of the fragment ions, together with the precursor mass from the MS1 spectrum, are typically sufficient to identify a peptide unambiguously.

in the sample, the corresponding ion counts (i.e., the area under the curve) for the same peptide in a different sample can be used for relative quantification.[15,19] These peptide ratios are then combined to give a relative ratio of proteins. This can be achieved in a variety of ways, such as by using the mean or median of (all or the top *n*) peptide ratios, taking a weighted average of peptide ratios based on signal intensity, calculating the ratio of total peptide ion counts, or by using linear regression to fit a line through the signal intensities for each peptide.[20] If there are more than two samples, pairwise protein ratios can be calculated by using any of these methods, and a least-squares analysis can be used to interpolate relative protein amounts in each sample.[19] MS2 spectra are required for peptide identification, but their signal is typically not used for quantification in these label-free approaches.

Compared with quantitation methods involving tags or labels, a label-free method avoids additional expense and sample preparation steps. Furthermore, label-free quantification is feasible on hundreds or even thousands of samples. On the other hand, there are some major limitations to a label-free approach. A major limitation is the requirement for multiple runs, which reduces throughput. Another drawback is the comparatively poor measurement precision; the median protein coefficients of variation (CVs) between replicates are typically about 20%.[19] Many less abundant proteins typically exhibit even larger variability, although this also tends to be a problem for other quantification methods. This comparatively poor reproducibility comes from each sample being run separately and data acquisition varying when MS1 and MS2 spectra are obtained. Another major limitation of the label-free approach is that, even in replicates, a significant fraction of peptides will not be detected in every sample. This is known as the missing value problem.[21] One can decide to concentrate on the proteins identified in every sample, but this will quickly reduce the number of quantified proteins to only the most abundant proteins in all samples. However, some statistical methods have been developed to tackle the missing value problem, and multiple papers have discussed the effective imputation of missing values.[21,22] Cox and co-workers have developed a system of tools, known as MaxLFQ, which imputes missing values by matching retention times and *m/z* values between different samples.[19] The continuous improvement in MS technology enables increasingly faster collection of MS2 spectra, which cover an increasing number of signals from the MS1 spectrum.[23] These advances might help to overcome the missing value problem in label-free quantification.

1.4. MS1-based quantification with heavy-isotope labeling

In contrast to label-free quantification, multiple methods label peptides with heavy, non-radioactive, isotopes. Peptides can be either labeled in vivo, for example, by the addition of heavy

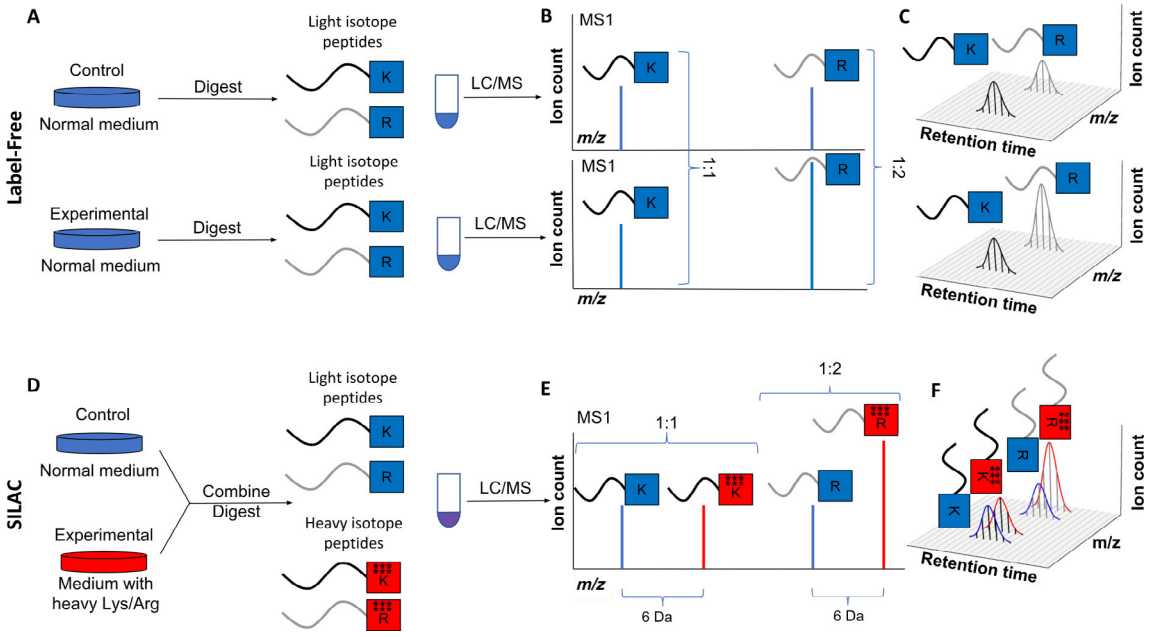


Figure 2. Outline of label-free and SILAC quantification. A)–C) Principles of label-free quantification. A) In label-free quantification, multiple protein samples are digested with trypsin (which cleaves after K or R). The resulting peptides are separated by means of LC and ionized before entering the mass spectrometer. Shown throughout are two peptides. The one ending in K has equal concentration in the two analyzed samples, and the one ending in R is concentrated twofold higher in the experimental sample relative to the control. B) The MS1 spectrum records the number of ions for various *m/z* values of the intact peptide eluting at a given time. C) The elution of a peptide takes about 20 s. During this time, about 10 MS1 spectra are collected, each showing the peptide at potentially different intensities. The integration of this intensity over time approximates the total number of ions ionizing in the mass spectrometer. D)–F) Principles of MS1-based quantification through heavy-isotope labeling (e.g., SILAC). D) In SILAC, cell samples are grown either in media with amino acids with naturally occurring isotopes (light) or in media in which amino acids (K and R) contain heavy isotopes (here 6). Importantly, the heavy isotopes do not alter the chemical properties of the peptides. Cells are lysed and combined. The proteins are digested together, and the resulting peptides are simultaneously separated by LC and ionized before entering the mass spectrometer. E) Peptides in the heavy sample are shifted to the right on the MS1 spectrum relative to those from the light sample. Ratios between signals within one spectrum can thus be used for relative quantification. F) To utilize all available information, typically, the ion intensity is integrated over the entire elution profile.

amino acids to tissue culture medium (stable isotope labeling with amino acids in cell culture or "SILAC"),[24] or in vitro, for example, by performing chemical modifications after proteolytic digestion.[25] Heavy isotopes, with the exception of deuterium, have essentially identical chemistry and elution patterns as those of their light equivalents, but the mass spectrometer can easily distinguish between different *m/z* ratios. The main advantage of this approach is that samples can be labeled separately with different isotopes and then combined before injection into the mass spectrometer. The samples can therefore be coanalyzed (Figure 2D, E) and relative quantification occurs within a single experiment rather than between runs (Figure 2E, F). This inherently leads to much higher reproducibility (i.e., higher measurement precision) and avoids the missing value problem of label-free approaches, if the number of samples does not exceed the maximum number of labeling combinations.[26] If there is no signal for a peptide, it is known to be below the detection limit, rather than not being picked up by chance, which can be the case for label-free approaches.

The major limitation of using the MS1 signal to quantify isotopically labeled peptides is that the complexity of the MS1 spectrum increases with the number of samples because each sample is isotopically labeled with a different mass. In practice, this limits the number of samples that can be compared in a single experiment to two or three.[26] A recent clever extension of SILAC can avoid this limitation by using labels, the masses of which only differ by a few mDa.[27] However, these experiments require current mass analyzers to be engineered to exceptionally high standards, which hinders wider application of the technology. Another limitation of MS1-based quantification is that the number of ions that can be accumulated in the most commonly used high-resolution analyzer, the Orbitrap, is limited. The number of ions for low-abundance peptides can therefore be very small if some very high abundance peptides coelute at the same time in the MS1 spectrum, resulting in less precise quantification due to poor ion statistics. This limitation has been somewhat alleviated by ion-mobility separation or BoxCar.[28]

1.5. Data-independent acquisition (DIA)

One feature common to the standard implementations of label-free and many other quantitative proteomics methods is the data-dependent acquisition (DDA) of MS2 spectra (Figure 3A–D). Based on the MS1 spectrum, the instrument successively chooses the largest signals for acquisition of MS2 spectra and peptide identification (Figure 3A).[29] Intuitively, this makes sense because the goal is to spend the limited number of MS2 spectra on the signals in the MS1 spectrum that can most likely be successfully identified and quantified (Figure 3B–E).

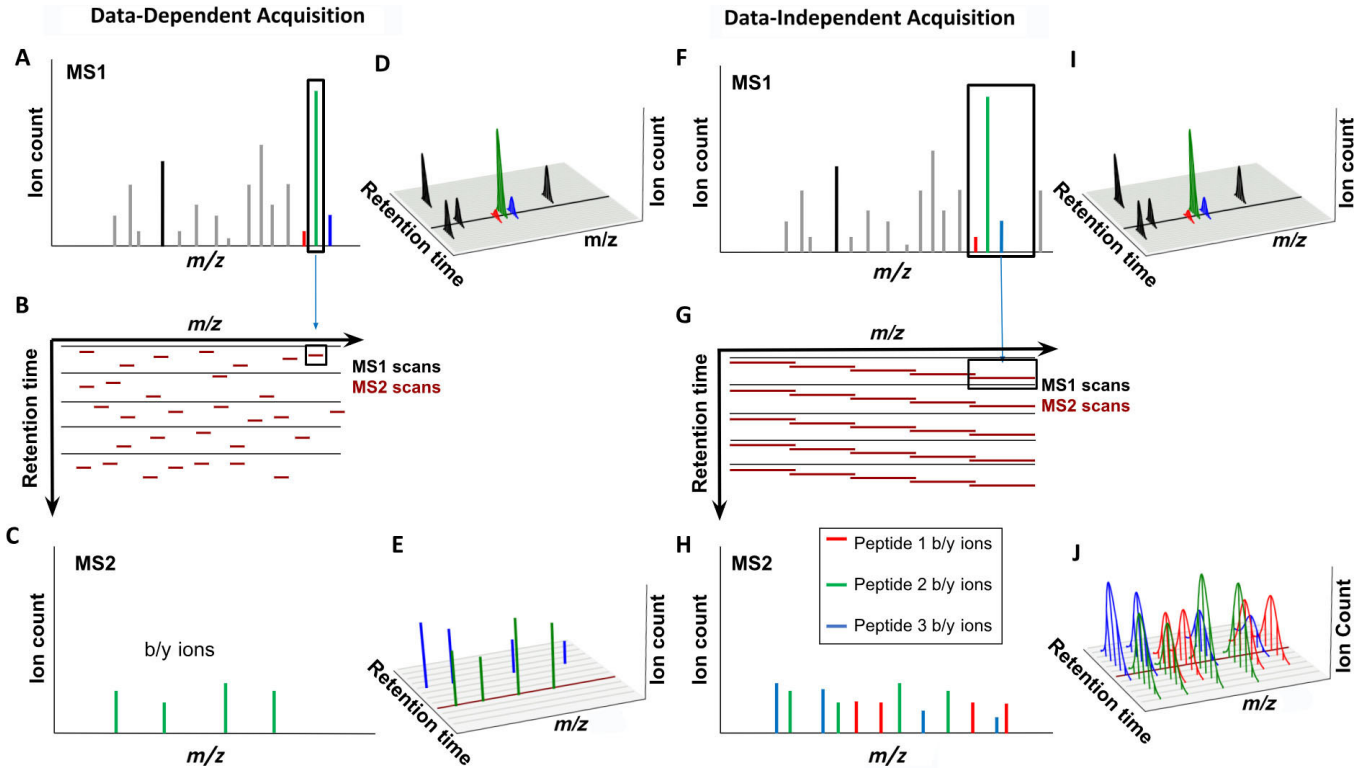


Figure 3. Comparison of data-dependent-aquisition (DDA) and data-independent-aquistion (DIA) approaches. A)–E) DDA. A) The goal of DDA is to identify as many peptides as possible, one at a time. The highest signals in the MS1 spectrum are selected for isolation, with an isolation window of about 1 Th. Peptides in this window are isolated and fragmented for readout in the MS2 spectra. B) Shown is the sequence of MS1 spectra (black) and the data-dependent MS2 isolation windows (dark red), centered on the highest abundant signals. Each MS1 spectrum is followed by multiple MS2 spectra; with current instrumentation and duty cycles of about 2 s, this would be about 30 MS2 spectra following each MS1 spectrum. C) The MS2 spectrum resulting after isolation and fragmentation consists mainly of b and *y* ions from the target peptide, which allows for comparatively simple peptide identification. D) For quantification, for example, with label-free approaches, the peptides in the MS1 spectrum are continuously monitored through the signal intensities in the MS1 spectra. Shown are the retention profiles for various peptides—the area under this curve is typically used for peptide quantification. The black line represents the single MS1 scan shown in B). E) Shown here are signals for the b and y ions for the green and blue peptides. A peptide is typically only isolated once for MS2 analysis, the signal height cannot be used for quantification. Not all signals in the MS1 will trigger the collection of an MS2 spectrum (signals for the red peptide are missing). Additionally, low-abundance signals might be below the detection limit in the MS1 spectrum, and thus, cannot trigger MS2 spectra. The dark red line represents the single MS2 scan shown in C). F)–J) DIA. F) The goal of DIA is to continuously collect fragment ion intensities for all eluting peptides. To make this approach compatible with current MS speed, significantly wider isolation windows (≈10 Th) compared with those of the DDA approach (≈1 Th) are required. All ions within this comparatively wide isolation window are isolated and simultaneously fragmented. G) Shown is the schedule of MS1 spectra (black) and the isolation windows of MS2 spectra (red). H) The simultaneous isolation and fragmentation of multiple peptides results in a complex MS2 spectrum consisting of b and y ions from nummerous isolated peptides. I) Similarly to DDA, MS1 intensities of peptides are collected and can be used for quantification. The black line indicates the time for the MS1 spectrum in G). J) Unlike in the DDA equivalent, ion intensity information for b and *y* ions is available throughout the entire elution profile for each peptide. This makes it possible to use fragment ion intensities for quantification. Because the entire *m/z* space is continuously covered, information for more peptides than with the DDA approach is available. Here, the abundance of the red peptide can be quantified. The dark red line represents the single MS2 scan shown in H).

However, there are usually more signals available than the mass spectrometer can isolate for fragmentation and which signals are chosen is an inherently stochastic process. Which MS2 spectra are acquired and at what time during the elution profile will differ from run to run, even if the exact same sample is reanalyzed.

DIA was envisioned to overcome this limitation by continuously and methodically collecting MS2 spectra covering the entire MS1 spectrum, so that for each *m/z* value information at the MS1 and MS2 levels are available (Figure 3 F–J).[30] Current instruments are not fast enough to collect enough MS2 spectra with the typical approximately 1 Th (Dalton/elementary charge) isolation window. Therefore, wider windows are chosen to reduce the number of MS2 spectra needed to cover the total *m/z* range, and so the resulting MS2 spectra typically contain fragment ion series from multiple precursors (Figure 3 F). Another alternative is to simultaneously isolate multiple small MS2 windows.[31] This results in a very complex series of MS2 spectra that are more difficult to analyze than those obtained from DDA methods (Figure 3 H). The Aebersold group introduced an approach, known as SWATH-MS, to analyze these complex spectra, by using prior knowledge of the chromatographic and MS behavior of the peptides.[30b] This approach was recently reviewed by Ludwig et al., who described improvements to DIA and how SWATH-MS could be used to analyze both total cell lysates and protein samples enriched for post-translational modifications.[32] The recent drastic improvements of DIA measurements are mostly due to computa-

tional advances.[33] The major advantage of DIA is its coverage: every peptide is fragmented multiple times. Thus, DIA does not have as severe a missing value problem as that of label-free DDA approaches. DIA seems particularly attractive for the comparison of many samples. For quantification by DIA, either MS1 or MS2 spectra can be chosen, although MS2 quantification is predominantly used (Figure 3I, J).[30b,c] Although the DDA and DIA acquisition strategies are currently mutually exclusive, the rapid advance of instrument speed will likely result in the two different approaches merging.[23] Once the instrument is fast enough to continuously cover the entire precursor space with the small (≈1 Th) windows commonly used for DDA strategies, the DDA and DIA methods may start to become identical, in terms of window size and breadth of coverage.

2. Multiplexed Proteomics with Isobaric Labeling

The methods discussed so far have key limitations. Label-free quantification provides comparatively poor measurement precision. Additionally, missing values of peptides that are only identified in some samples are hard to interpret, even qualitatively. Although MS1-based isotope labeling offers exquisite quantification for more abundant peptides, it suffers from a lack of multiplexing capability because, as the number of samples increases, so does the complexity of the MS1 spectrum. DIA mostly overcomes the missing value problem of a label-free approach, but samples are still analyzed one at a time, which limits measurement precision and requires significant instrument time.

2.1. Principles of quantitative multiplexed proteomics

Multiplexed proteomics, based on isobaric mass tags, promises to overcome, or at least mitigate, these limitations.[34] The most commonly used isobaric tags are the tandem mass tag (TMT)[35] and the isobaric tag for relative and absolute quantitation (iTRAQ),[36] which are both commercially available, but there are other isobaric tags, as described later. Isobaric tags are reagents used to covalently modify peptides, by using the heavy-isotope distribution in the tag to encode different conditions, and are generally added after digestion. Unlike isotopic labeling methods, such as SILAC, that were discussed earlier (Figure 2D–F), each variant of an isobaric tag set has an identical total mass. The only difference is how the heavy isotopes are distributed along the tag. Because the tags contain a site that fragments in the MS2 spectrum, reporter ions with different masses result, depending from which sample the peptide originated. In addition to the reactive group, which reacts with the peptide, each tag contains a reporter group with a differential number of heavy isotopes. To keep the total mass of the tag constant, the number of heavy isotopes on the mass balancer group is adjusted accordingly (Figure 4A). Identical peptides from different samples elute at the same time, and therefore, appear as a single signal in the MS1 spectrum. This is a major advantage because the complexity of the MS1 spectra does not increase significantly with the number of samples. This is in contrast to SILAC-like experiments, in which even comparing replicates will double the number of signals in the MS1 spectra. Therefore, the number of conditions that can be compared with isobaric tag experiments in a single experiment is higher (currently up to 11) than that with SILAC-like methods. Quantification occurs after isolation and fragmentation of these labeled peptides in the MS2 spectrum (Figure 4B). Usually, the amount of energy added for fragmentation is only sufficient for one bond to break. This could either be a peptide bond on the backbone or the intended breakage point in the isobaric tag. Each tag has several heavy isotopes that are distributed differently relative to this fragile bond. Upon breakage, the isobaric tag produces low *m/z* reporter ions that contain different masses, depending on the conditions from which

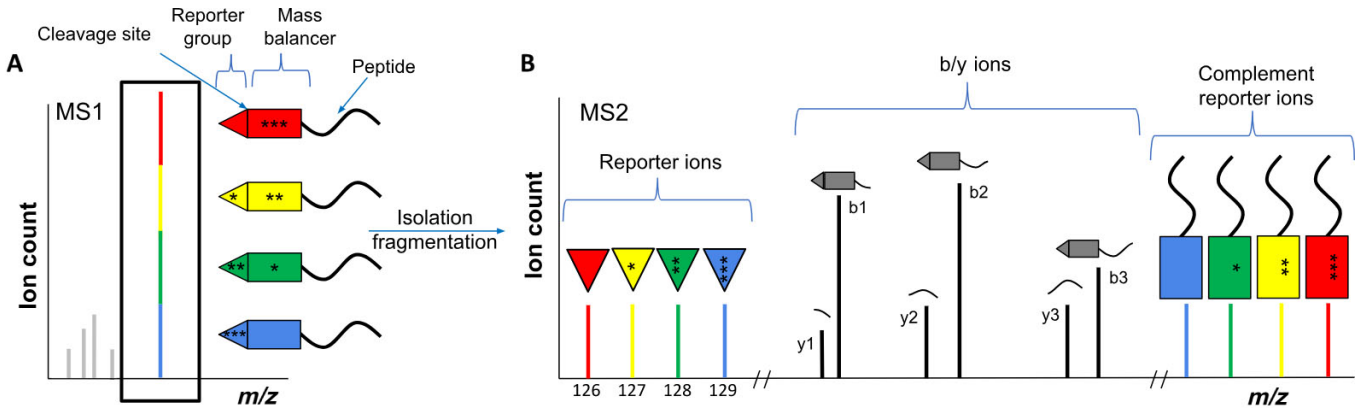


Figure 4. Outline of multiplexed proteomics with isobaric tags. A) Isobaric tags have the same total mass, but differing distributions of heavy isotopes between the reporter group and mass balancer. Heavy isotopes are shown as asterisks. Peptides from four different samples are labeled with tags of the same mass, resulting in a single MS1 signal, which can be isolated. With more tags (conditions), the complexity of the MS1 spectrum does not increase. This makes isobaric tags compatible with higher multiplexing (currently up to 11) compared with, for example, SILAC (Figure 2). B) After a peptide is isolated based on the MS1 spectrum, fragmentation will either cleave off the reporter ions, or lead to fragmentation of the peptide backbone. The reporter ions show different masses in the MS2 spectrum and can be used for relative quantification. Similarly, the intact peptide with the balancing groups, the complement reporter ions, can also be used for quantification. The b and *y* ions are used for peptide identification (Figure 1).

they originate, and can therefore be used for relative quantification (Figure 4B). Additionally, breakage of the isobaric tag leads to the formation of complement reporter ions, which contain the balancing part of the isobaric tag and the intact peptide.[37,38] The balancing group of the isobaric tag also encodes the experimental conditions, and the complement reporter ions can therefore be similarly used for quantification (Figure 4B).[38] We discuss later the key advantages and disadvantages of the utilization of different reporter ions for quantification. The complement reporter ions were observed, for example, by the group of Mechtler, but were not initially used for quantification.[39] Instead, they removed these signals to increase the peptide identification success rate. The use of complement reporter ions for quantification is similar to an approach reported by Yan et al., who labeled peptides differentially on their N and C termini with heavy isotopes to generate isobaric peptides.[40] They used the fragment ions for quantification.

Although the labeling step after protein digestion could introduce some variability, and although there is a limit to the number of samples that can be labeled by an isobaric tag system, there are many advantages of isobaric tags that compensate for these deficiencies. The ability to analyze many samples at once mostly circumvents the missing value problem. If no signal is detected for a peptide under particular conditions, it can be inferred that the peptide is indeed much less abundant than it is under other conditions. It is still possible for peptides to be excluded from MS2 fragmentation, but, since all samples have peptides eluting under the same signal, all labeled versions of the same peptide with either be isolated together or not at all. Another major advantage is the inherent high reproducibility between samples due to the samples being combined after labeling and co-analyzed. Relative to MS1 quantification methods, such as SILAC, data quality is even further improved because each analysis heavily enriches the peptides of interest in the MS2 spectrum, in which quantification occurs, resulting in useful peptide ion statistics, even for low-abundance peptides.[41] Multiplexed proteomics therefore demonstrates very high reproducibility, with CVs of about 5%, and very few peptides with CVs above 10%.[42] The last major advantage, at least compared with label-free quantification, is that throughput is markedly increased because multiple samples can be co-analyzed in one run. This limits expense and makes it compatible with the analysis of prefractionated samples.[43] For equivalent amounts of machine time, this results in significantly more proteins that can be quantified relative to label-free approaches.[42,44] Together, these benefits make multiplexed proteomics a very attractive option for relative quantification.

2.2. Main problem of multiplexed proteomics: Interference/ratio distortion

In the previous section, we discussed the principles and promises of multiplexed proteomics. However, in its initial implementation, the method came with a major measurement artifact: ratio distortion. In Figure 4, we pretended that it was possible to specifically isolate one peptide of interest. For technical reasons, however, the smallest possible isolation window currently achievable with a mass filter (e.g., a quadrupole) is approximately 0.5 Th.[45] In complex mixtures, such as tryptic digests derived from cell lysates, whenever a peptide is isolated in the MS1 spectrum for MS2 analysis, other peptides with similar *m/z* values will nearly always be co-isolated (Figure 5A).[46] Because both the target peptide and contaminating peptides carry the same reporter groups, after MS2 fragmentation, the reporter ion signal for that particular isolation will be a combination of reporter ions stemming from the peptide of interest and from all other contaminating peptide ions (Figure 5B). Nearly all measurements are therefore distorted, often to a significant extent. In general, these contaminating ions tend to bias the relative ratios between different conditions towards a 1:1 ratio.[46c] This distortion tends to be more significant for low-abundance peptides, for which the interfering signal is relatively greater. However, it is also possible that a 1:1 peptide is distorted by a changing contaminant, resulting in unsubstantiated measurements of changes.[47] Despite these problems, many groups successfully use multiplexed methods that are vulnerable to interference.[48] For some studies, qualitative knowledge of which proteins change is sufficient. However, if one is interested in the quantitative change of protein expression levels, addressing interference is essential.[46b,c,49]

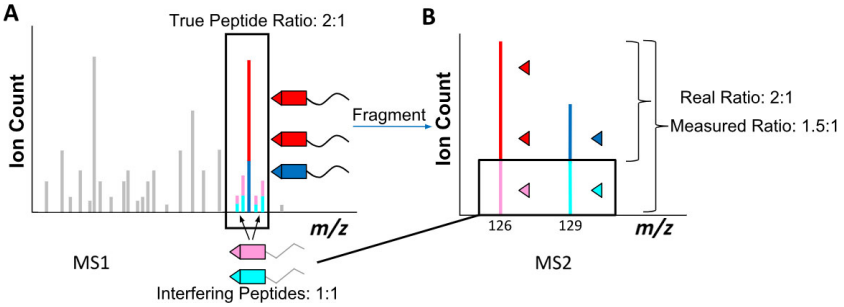


Figure 5. The problem of multiplexed proteomics: ratio distortion. A) Even for the smallest technically possible isolation window centered on a peptide of interest (red and dark blue), in a real experiment, other peptides with similar *m/z* and retention time will be co-isolated (pink and light blue). These interfering peptides will also be labeled with identical isobaric tags. B) Upon co-isolation and co-fragmentation, in the MS2 spectrum, the low *m/z* reporter ions are identical, regardless of origin, and distort the quantification. Most background peptides tend to not change, showing a 1:1 ratio between control and experiment. The observed ratio for a peptide of interest, which changes twofold between control and experiment, will typically be compressed towards a 1:1 ratio.

Recently, multiple statistical methods have been suggested to bioinformatically correct for this distortion.[50] Nevertheless, the best sample quantification can currently be obtained by applying experimental remedies for this major problem, which we discuss in the following section.

2.3. Overcoming interference with further gas-phase purification (QuantMode, MS3)

One of the earliest methods to reduce interference from contaminating ions is an approach known as QuantMode.[46b] This method reduces the charge of all peptides by one. After isolation of the new desired *m/z* window, interfering peptides with similar *m/z* but different charge from that of the targeted peptide are removed. QuantMode was thus able to significantly reduce interfering ions, which resulted in more accurate quantification. The main drawback of the method is that interfering ions of the same charge as that of the target ion can still be co-isolated. Additionally, the proton-transfer process, which alters the ion charge, is comparatively slow, resulting in fewer collected spectra and a shallower assaying of the sample.

Currently, the most widely used approach to counteract ratio distortion involves a further fragmentation and isolation step to produce an MS3 spectrum.[46c] An MS3 spectrum results from the isolation of ions in the MS2 spectrum and their further fragmentation (Figure 6 A–C). This filters out the interfering peptides, which allows the target peptide to be quantified more accurately. The original version of the MS3 method only isolated a single isobaric-tag-labeled fragment ion from each MS2 spectrum, which greatly reduced the sensitivity of the quantification.[46c] This drawback was overcome by a more advanced method called MultiNotch MS3.[47] The use of isolation waveforms with multiple frequency notches enables the simultaneous precursor selection (SPS) in the linear ion trap.[51]

With this approach, multiple fragment ions from each MS2 spectrum are simultaneously isolated, which results in greater sensitivity (Figure 6 A–C). Thermo Fisher Scientific commercialized this approach on the Orbitrap Fusion and Lumos mass spectrometers. As a result, MultiNotch MS3 is now widely used and currently considered to be state-of-the-art; it is able to detect changes of about 10% in protein abundance with high confidence.[52]

Despite this, there are a number of limitations to the MS3 approach. Perhaps the most significant disadvantage is the requirement for additional MS scans. This results in a loss of ions and comparatively slow cycle times. Furthermore, the MS3-based methods require instrumentation that is more complex and expensive. Finally, even MultiNotch MS3 fails to completely remove interference, especially for peptides with low abun-

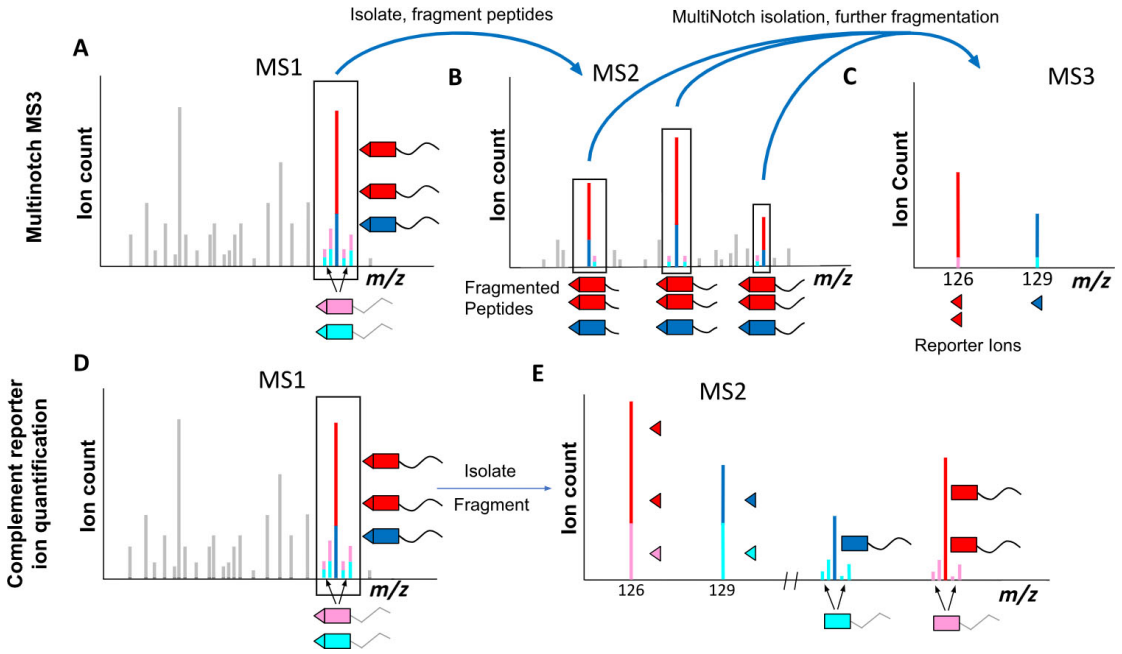


Figure 6. Strategies to overcome ratio distortion. A)–C) Overview of the MultiNotch MS3 approach. A) The MultiNotch MS3 method acquires an MS2 spectrum, similar to the standard approach, by isolating a target peptide (red and dark blue), along with interfering peptides (pink and light blue). This spectrum is used for peptide identification. B) Instead of quantifying the reporter ions in the MS2 spectrum, the highest abundant signals, which typically are b and y ions from the peptide of interest, are simultaneously isolated and further fragmented for an MS3 spectrum. C) Reporter ions in the MS3 spectrum are used for quantification. Additional gas-phase purification typically leads to removal of most interfering signals. Although not perfect, the measured ratios are typically significantly more accurate than those obtained with a standard MS2 approach. D)–E) Overview of the complement reporter ion quantification strategy. D) For the complement reporter ion method, a standard MS2 spectrum is acquired, which will co-isolate and co-fragment the peptide of interest and interfering species. E) The low *m/z* reporter ions show interference, as discussed in Figure 5. However, this method involves analyzing the complement reporter ions in the MS2 spectrum, in which the peptide is still attached to the mass balancer group. Because the target peptide and interfering peptides typically have slightly different masses, this allows them to be distinguished with a high resolution mass analyzer, such as an Orbitrap. This results in significantly more accurate quantification compared with that of MS2 and even the MS3 approach. Although interference still occasionally leads to ratio distortion, to the best of our knowledge, this method currently generates the most accurate data.

dance because interfering ions in the MS2 spectrum are still co-isolated in the MS3 spectrum (Figure 6C).[42,52a] It is likely that MultiNotch MS3 data quality can be further improved by setting the notches in a peptide specific manner. For shotgun approaches, this would require the ability to identify MS2 spectra immediately after their acquisition and before the corresponding MS3 scan.[53]

Another approach to reduce interference is ion mobility spectrometry.[28b,54] This separation method, which is orthogonal to LC and *m/z* analyzers, promises to suppress ratio distortion because interfering ions will be separated from the peptide of interest.

2.4. Overcoming interference with the complement reporter ion-based approach (TMTc)

An alternative method to overcome the ratio distortion problem is based on the complement fragment ions in the MS2 spectrum.[38] If an isobaric tag (e.g., TMT) breaks, it produces low *m/z* reporter ions, but also the intact peptide with the balancing group of the isobaric tag still attached (Figure 4B). Due to their complementary nature, these were named complement reporter ions, or TMTc ions if the experiment was performed with TMT tags.

TMTc ions containing the same peptide differ in mass, depending on the experimental conditions, similar to low *m/z* reporter ions. These TMTc ions can therefore also be used for multiplexed quantification (Figure 6D, E). The key advantage of using TMTc over low *m/z* reporter ions for quantification is that any interfering peptides typically will have slightly different masses to those of the target peptide. The ability to distinguish different TMTc masses in the Orbitrap is about 100-fold higher than that of the lowest feasible resolving power of quadrupole ion isolation. TMTc is therefore much more robust to interfering ions than that of the standard MS2 approach. TMTc is even able to outperform MS3-based methods in terms of measurement accuracy.[42] Furthermore, compared with Quantmode or MultiNotch MS3, TMTc does not require an additional fragmentation step, which saves time and, in principle, increases sensitivity. Because no higher order scans are required, the complement reporter ion approach can be performed on comparatively simple instruments, such as quadrupole Orbitraps or QTOFs.

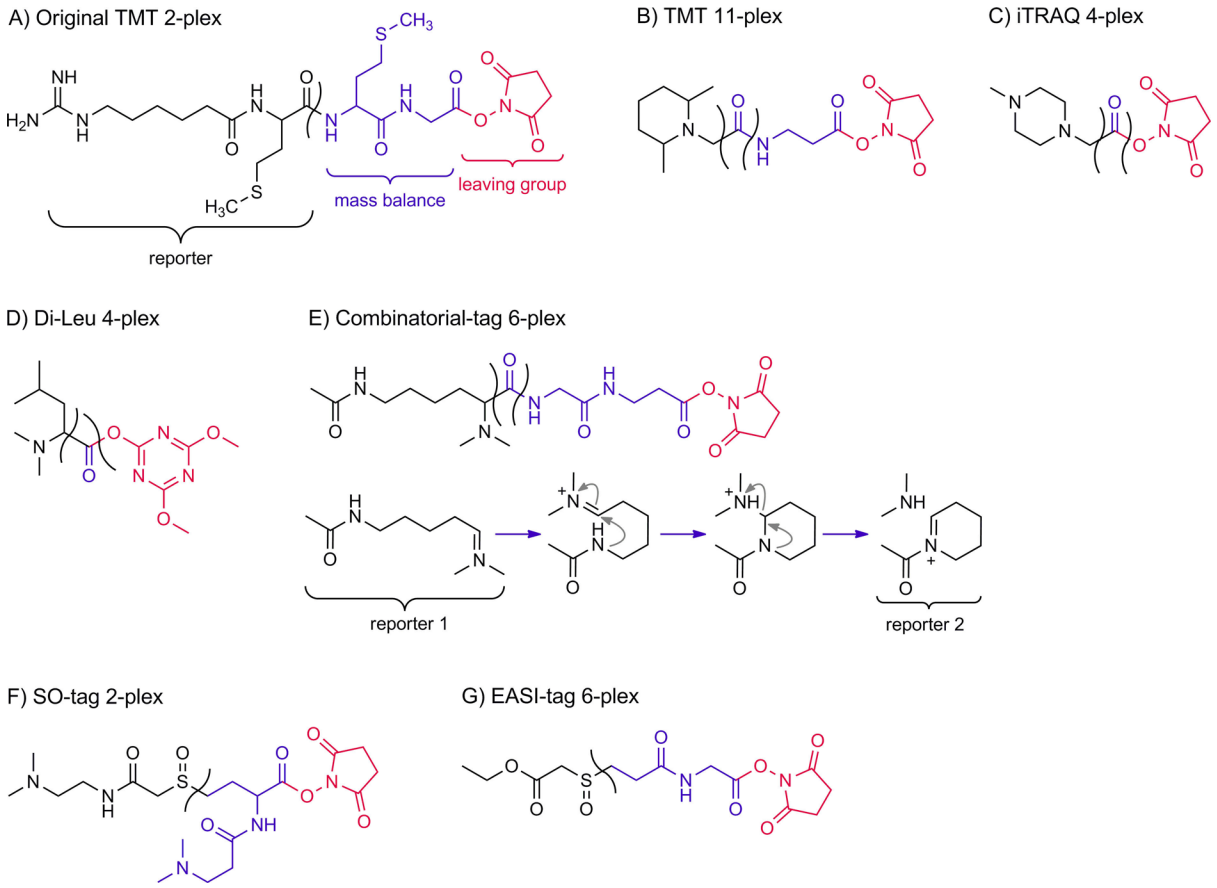
Figure 6 heavily simplifies the actual picture by portraying a single signal corresponding to each peptide. In reality, peptides elute as an isotopic envelope of multiple peaks, spaced apart by 1 Th due to the natural frequency of ^{13}C, ^{15}N, ^{18}O, and other heavy isotopes in biological molecules. If the entire isotopic envelope of a peptide is isolated, the complement reporter ion cluster has to be deconvolved from the isotopic envelope of the precursor peptide.[38] This deconvolution process results in a loss of quantitative precision. To combat this shortcoming, we have developed a refinement of the TMTc deconvolution method known as TMTc+.[42] The TMTc+ method uses a narrow isolation window. In the extreme case, in which only one precursor signal is chosen, deconvolution becomes similar to the simplified cases represented in Figure 6, and only isotopic impurities have to be accounted for. The resulting data comes with a drastic improvement of measurement precision, while still preserving superb measurement accuracy.

Despite the promises of the complement reporter ion approach, several key limitations remain: at high *m/z* values, mass spectrometers cannot distinguish the extra neutron in heavy nitrogen or carbon in commercial isobaric tags with currently feasible resolving power. Thus, although with MS2 or MS3 approaches, up to 11 conditions can currently be compared, only 5 TMT channels are currently distinguishable with the complement reporter ion strategy (Figure S1 in the Supporting Information). This lowered multiplexing capacity is a major drawback to the TMTc method, but future isobaric tags should be able to address this limitation. Furthermore, emerging super-resolution approaches are, at least in principle, able to further increase multiplexing capacity by providing the resolving power to distinguish the extra neutron in different elements, even at high *m/z* regions.[55] Another major hurdle is the poor formation of the complement reporter ions. Commercially available tags were not intended for this purpose and the complement reporter ions formed inefficiently. Recently, two tags (the SO-tag and EASI-tag; further discussed below)[37,56] were designed specifically for the formation of the complement reporter ions. Although the complement forms efficiently, it comes with the drawback of making identification of peptides difficult because breakages of both the tag and the peptide backbone can occur, leading to many additional signals, which are not classical b or y ions and are not recognized by standard search algorithms. The advantageous combination of measurement sensitivity, precision, and accuracy of the complement reporter ion approach (specifically TMTc+) makes it our current method of choice for most experiments in our laboratory. Although TMTc-based methods have not been widely used outside of our laboratory, we think that they provide a viable alternative to the more prevalent MS3-based methods. Nevertheless, many shortcomings remain, and considerable extra efforts will be required to exploit the full potential of the complement-reporter-ion method.

2.5. Overview of different isobaric tags

All isobaric tags contain a functional group that enables covalent attachment to peptides. Typically, this group reacts with primary amines on the N terminus of a peptide or with lysine side chains. However, some tags react with carbonyl or sulfhydryl groups.[57] In addition, all isobaric tags contain a reporter group and a mass balancer group (Scheme 1A). The total number of heavy isotopes in the tag is constant, making them isobaric in the MS1 spectrum. However, the distribution of heavy isotopes between the reporter and balancing group differs for different conditions (Figure S1).

Isobaric tags were first introduced by Thompson et al. for the relative quantification of peptides.[34] The original tag was a 2-plex called tandem mass tag (TMT; Scheme 1A). Although this original TMT tag was used to prove an important new concept, the structure was comparatively bulky, which led to additional unintentional fragmentation patterns, poor ionization



Scheme 1. Overview of isobaric tags. The black part of the structure indicates the reporter part; the balancing group is blue; and the leaving group, which is removed after the tag reacts with the peptides, is red. A) The original 2-plex TMT from Thompson et al.[34] B) The current commercial TMT, which can encode up to 11 different conditions (see Figure S1 for the heavy-isotope distribution). C) In the iTRAQ structure, the oxygen can be either ^{16}O or ^{18}O. D) The DiLeu tag is a 4-plex tag developed by Xiang et al.[60] E) Braun et al. developed combinatorial tags.[62] These tags generate multiple reporter ions, which allow for a high multiplexing capacity for a given number of heavy isotopes. After fragmentation at the cleavage site shown, reporter 1 forms. However, this further fragments into reporter 2. F) Stadlmeier et al. developed the sulfoxide-based tag, which was optimized for complement reporter ion formation due to fragmentation of the sulfoxide bond at lower energies.[37] The two tertiary amines result in higher charge states of peptides after ionization and further facilitate fragmentation. G) The EASI-tag developed by Virreira Winter et al. similarly fragments comparatively easily.[56] The "reporter" part of the EASI-tag is a neutral loss. Therefore, quantification with the EASI-tag is only possible through the complement reporter ions.

properties, and poor identification success rates. Members of Proteome Sciences further developed the TMT into the significantly smaller version currently available commercially from Pierce (Scheme 1 B).[35] This TMT tag is currently able to encode up to 11 different conditions (Figure S1). To obtain this high multiplexing capacity, a mass analyzer is required that can distinguish between the additional neutron masses in ^{13}C versus ^{14}N, which differ by 6 mDa (Figure S1).[58]

An alternative commercially available tag from AB Sciex is iTRAQ (Scheme 1 C). The structure can encode up to four different conditions.[36] An 8-plex iTRAQ is also commercially available, but, to the best of our knowledge, its structure has not been published.[59] Pichler et al. found that 4-plex iTRAQ had a higher peptide identification rate than those of 8-plex iTRAQ or 6-plex TMT.[39] They concluded that the 8-plex iTRAQ might suffer due to the appearance of fragment ions from the larger tag in the MS2 spectrum, which they suggested hinders peptide identification. This indicates that isobaric tags should be designed to be as small as possible, while allowing sufficient multiplexing capacity.

An alternative to the commercial tags are the *N*,*N*-dimethyl leucine (DiLeu) tags (Scheme 1 D).[60] These tags contain a reporter group consisting of a dimethylamine connected to a leucine side chain, and a mass balancing group consisting of the CO atoms of the carboxyl group. These DiLeu tags were originally 4-plex. Using deuterium isotopes as labels, Frost et al. reported an upgraded version of DiLeu that increased its multiplexing capacity to 12.[61] Nevertheless, deuterium-labeled peptides typically show different elution profiles to those of unlabeled peptides. For MS1-based quantification, this can be acceptable because the entire elution profile can be integrated (Figure 2 F).[24] However, for multiplexed proteomics, typically only one MS2/MS3 spectrum is acquired per peptide. Differential elution profiles in different channels could therefore lead to serious quantification artifacts.

A clever set of isobaric tags was showcased by Braun et al. (Scheme 1 E).[62] Known as combinatorial isobaric mass tags (CMTs), the fragmentation of these molecules results in multiple reporter ions. Because of this, their multiplexing capacities are larger than those of conventional tags of comparable size

and number of heavy isotopes. Nevertheless, quantification depends on a deconvolution approach, which comes at the cost of measurement precision, particularly if interfering ions alter the true peptide ratios. The report on CMT demonstrates a 6-plex version by using two different reporter ions. Furthermore, by taking into account a third reporter ion that was also detected, the chemical structure of the tag allows for 28-plex tags, if utilizing the 6 mDa spacing between heavy carbon and nitrogen. These high values could be achieved with just five heavy isotopes on each tag.

Motivated by the inefficient formation of the complement reporter ions (Figure 6B) from the commercial isobaric tags, Stadlmeier et al. developed a sulfoxide-based tag, known as SO-tag (Scheme 1F).[37] In this tag, the reporter and balancer groups are linked by a sulfoxide group. This allows the tag to be fragmented at low energies, which increases the yield of complement reporter ions. Indeed, the SO-tag is much more favorable to fragmentation, relative to the peptide backbone, and so typically many signals are available for quantification. An interesting idea put forward in this publication was quantification by using complement b and y ions. These are the fragment ions that develop if both the isobaric tag and peptide backbone break. Similar to complement reporter ions that result from only breakage of the isobaric tag, these ions also encode different sample conditions.

Another sulfoxide-based tag was recently developed by Virreira Winter et al., known as the easily abstractable sulfoxide-based isobaric (EASI) tag (Scheme 1G).[56] The EASI-tag also contains an asymmetric sulfoxide bond that is cleaved at relatively low energy. An interesting novelty of the EASI-tag is that the low *m/z* reporter ion equivalent is a neutral loss, which only makes quantification possible through the complement reporter ions.

Both SO-tag and EASI-tag seem to suffer from comparatively poor success rates in identifying peptide spectra (Figure 1D). This is because the tags fragment more easily than the peptide backbone. Typically, the b and y ions additionally lose the low *m/z* reporter part of the isobaric tag, which results in signals that standard search algorithms do not consider for identification. Improved search algorithms that consider these ions might mitigate this problem. However, the spectra are much more complex, and it is not clear how much adapted search algorithms will be able to overcome this major limitation, which results in comparatively few quantified peptides and ultimately proteins. We believe the most promising way forward for complement reporter ion quantification is the development of new chemical structures that balance the formation of complement reporter ions with the ability to reliably and efficiently identify peptides.

3. Emerging Multiplexed Proteomics Technologies

Multiplexed proteomics in its current form is highly attractive and well suited for many studies. However, significant shortcomings remain. Among them are the difficulties in detecting low-abundance proteins. These are often some of the most interesting proteins, such as transcription factors or signaling molecules. To overcome these limitations, we discuss the emerging fusion of targeted proteomics with multiplexing technologies to reliably reach low-abundance proteins. Another major limitation of multiplexed proteomics is the maximal multiplexing capacity. The current limit, with TMT tags, is 11-plex, but, in many studies, it is desirable to compare hundreds or even thousands of different samples. In principle, these can be split into several 11-plex experiments, but then, similar to label-free approaches, some peptides will only be analyzed in a subset of the experiments. The interpretation of these missing values is difficult. Additionally, quantification between 11-plexes is challenging. Here, we suggest the fusion of the complement reporter ion quantification strategy with DIA approaches to enable the comparison of hundreds of samples with few missing values and high measurement quality.

3.1. Targeted multiplexed proteomics

So far, the methods we have described involve analyzing protein samples globally, with the aim of identifying and quantifying as many proteins as possible. However, it is also possible to sacrifice global coverage and focus the limited ion injection times on peptides of a few ($\approx$100) proteins of particular interest. Such approaches are called targeted proteomics.[63] By predefining the data acquisition towards specific ions that elute at specific times, targeting, at least in principle, enables detection and quantification of low abundant peptides. Although this approach requires a significant amount of setup, it can be used to analyze low-abundance peptides that would be missed by a shotgun approach.

The simple combination of targeted proteomics with an isobaric MS2 approach is not attractive. The interference problem is especially problematic for low-abundance peptides and quantification would be very unreliable and likely severely distorted (Figure 5). However, the reduction of ratio distortion by using MS3-based methods made targeted multiplexing of proteolyzed cell lysates more feasible (Figure 6A). As a result, Erickson et al. developed a targeted multiplexing method known as TOMAHAQ.[64] In this method, samples were spiked with trigger peptides labeled with TMT0, which was the standard TMT structure without any heavy isotopes. This resulted in the trigger peptides eluting simultaneously at a known *m/z* offset away from the target sample peptides in the MS1 spectrum, which were labeled with standard TMT 10-plex labels. The spiked-in trigger peptides were sufficiently abundant to be consistently observed in the MS1 scan. The instrument was programmed to isolate and fragment the sample peptides at the known mass offset, even if there were no detectable signals from the target peptides at that *m/z*. The notches for the MS3 scan were preprogrammed to fit the peptide of interest and to be specific for fragment ions containing an intact TMT tag. Furthermore, this approach was refined by selecting in real time from the MS2 spectra only those b and y ions that had minimal interfering ions. This allows the researchers to obtain accurate quantifications of even dilute peptides, which

suffer from significant interference, even upon using the standard SPS-MS3 method.

Although this MS3-based method has proved critical to targeted multiplexed proteomics, the complement reporter ion approach is, in our opinion, particularly attractive for this purpose (Figure 6B). The complement reporter ion strategy is superbly able to distinguish signal from interfering background noise, and the lack of an additional gas-phase isolation makes it, at least in principle, more sensitive. Thus, we believe that, for a targeted approach, the use of complement b and y ions could be particularly attractive because this would provide an additional layer of distinction (apart from the precursor mass) and allow the separation of signals, even for isobaric peptides with nearly identical elution times.

3.2. Fusing multiplexed proteomics with DIA

Currently, multiplexed proteomics is very attractive for comparing up to 11 samples, which is the maximum number that can be analyzed in a single experiment. However, many studies require the comparison of hundreds or even thousands of different conditions. In such scenarios, multiplexed experiments suffer from the missing value problem, similar to label-free approaches. Additionally, it is hard to compare proteins quantified in one 11-plex with another 11-plex. Often, so-called bridging channels are used, which are analyzed in all experiments.[65] However, it is possible that these do not contain a subset of proteins or contain it at a significantly different concentration than that of other analyzed samples. In such cases, quantification between two 11-plexes relies on the ratios of unreliable ratios. Currently, the best remedy for quantitative comparison of many samples might be DIA approaches. However, these come with comparatively low measurement precision and inefficient instrument time usage. A fusion of DIA with isobaric labeling approaches might be able to combine the advantages of both methods.

Because DIA involves fragmenting all of the MS1 ions in a certain mass range (Figure 3F–G), it is incompatible with MS2 or MS3 methods, which use the low *m/z* reporter ions (Figures 5 and 6A–C). However, for multiplexed methods based on the complement reporter ion approach, the reporter ion signal is precursor specific (Figure 6D, E). Complement ions from different peptides would have different masses and would be distinguishable (Figure 7). In a proof of principle, we demonstrated the quantification of two different peptides in one spectrum.[38] Particularly attractive might be the use of b or y ions that additionally have a broken isobaric tag, thereby forming complement fragment ions, which can also be used for quantification.[37] Regardless, several hurdles have to be overcome to make this approach feasible. Simultaneously fragmenting all of the precursor ions within an *m/z* range will give a very complex MS2 spectrum, both due to the large number of isolated peptide species and due to the differing masses of the complement reporter ions, which will likely make the analysis quite challenging. Although interference will lead to some ratio distortion, the combination of multiple quantification events over multiple spectra might provide sufficient data to overcome such challenges. Additionally, the very wide isolation windows will require deconvolution of the isotopic envelopes. Isobaric tags with multiple Dalton spacing might make this approach more feasible with high measurement precision. Despite these challenges, a successful fusion of DIA with multiplexing could be a highly attractive method.

4. Summary and Outlook

Multiplexed proteomics in its current form is highly attractive and often the best suited quantitative proteomics option for many studies. The higher throughput enabled by multiplexing has made it possible to analyze hundreds of samples with reasonable depth.[66] Over the last several years, remarkable technological progress has been made, particularly in addressing

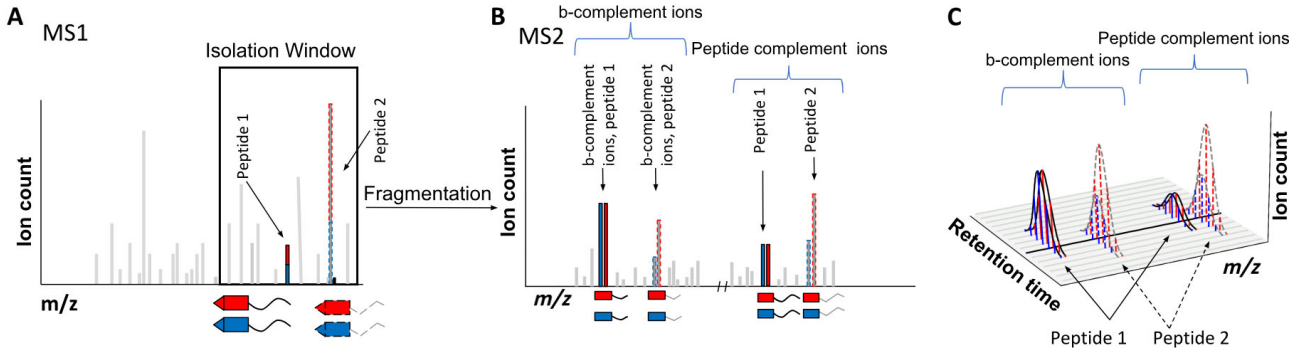


Figure 7. Proposed fusion of DIA with multiplexed proteomics. A) Peptides would be labeled with isobaric tags, similar to a normal multiplexing experiment. For sake of simplicity, we only show two conditions. Similar to a normal DIA workflow, all signals within a certain wide *m/z* window in an MS1 scan are co-isolated. This window contains multiple peptides, which will all be simultaneously isolated and fragmented into an MS2 spectrum. For simplicity, only two peptides are shown in detail and are depicted with solid and dashed outlines. B) In the MS2 spectrum, the low *m/z* reporter ions cannot be used for quantification because the reporter ions of all co-isolated peptides will be identical. However, simultaneous quantification is possible through the peptide complement reporter ions. Additionally, complement b and y ions that result from backbone breakage and loss of the reporter groups can be used for peptide-specific quantification. C) The continuous monitoring of peptide complement reporter ions and b- and y-fragment complement reporter ions allow the relative quantification of multiplexed abundances, even between various runs. Additionally, the number of missing values in samples larger than the multiplexing capacity of a single tag should be drastically reduced.

ratio distortion, which is the major shortcoming of multiplexed proteomics. Currently, isobaric tag-based multiplexed proteomics can accurately, precisely, and sensitively quantify thousands of proteins simultaneously across up to 11 samples. With the resulting data, changes of less than about 10% can be detected with high confidence. Despite these advantages, major limitations remain. One major remaining hurdle is the reliable quantification of low-abundance proteins. Emerging methods for targeted multiplexing promise to overcome the problem of quantifying low-abundance proteins across multiple conditions. Another major remaining challenge is how to quantify protein abundances among hundreds of samples, while limiting missing values. Herein, we have suggested that the fusion of DIA with the complement reporter ion approach might be able to unite the best of both worlds.

Acknowledgements

We would like to thank members of the Wühr Lab for helpful suggestions and discussions. We thank Michael Stadlmeier for comments on the manuscript. N.P. was supported by NIH grant T32GM007388. This work was supported by NIH grant 1R35GM128813 and the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award numbers DE-SC0018420 and DE-SC0018260.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: analytical methods · isotopic labeling · mass spectrometry · proteins · proteomics

[1] a) A. D. Catherman, O. S. Skinner, N. L. Kelleher, *Biochem. Biophys. Res. Commun.* **2014**, *445*, 683 – 693; b) T. K. Toby, L. Fornelli, N. L. Kelleher, *Annu. Rev. Anal. Chem.* **2016**, *9*, 499 – 519.

[2] R. E. Ardrey, *Liquid Chromatography – Mass Spectrometry: An Introduction*, Wiley, New York, **2003**.

[3] R. L. Gundry, M. Y. White, C. I. Murray, L. A. Kane, Q. Fu, B. A. Stanley, J. E. Van Eyk, *Curr. Protoc. Mol. Biol.* **2009**, *10*, 1025.

[4] J. V. Olsen, S. E. Ong, M. Mann, *Mol. Cell. Proteomics* **2004**, *3*, 608 – 614.

[5] a) B. Manadas, V. M. Mendes, J. English, M. J. Dunn, *Expert Rev. Proteomics* **2010**, *7*, 655 – 663; b) H. J. Lee, H. J. Kim, D. C. Liebler, *J. Proteome Res.* **2016**, *15*, 2346 – 2354.

[6] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science* **1989**, *246*, 64 – 71.

[7] J. Muntel, S. A. Boswell, S. Tang, S. Ahmed, I. Wapinski, G. Foley, H. Steen, M. Springer, *Mol. Cell. Proteomics* **2015**, *14*, 430 – 440.

[8] D. Chelius, P. V. Bondarenko, *J. Proteome Res.* **2002**, *1*, 317 – 323.

[9] D. F. Hunt, J. R. Yates, 3rd, J. Shabanowitz, S. Winston, C. R. Hauer, *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 6233 – 6237.

[10] P. Roepstorff, J. Fohlman, *Biomed. Mass Spectrom.* **1984**, *11*, 601.

[11] J. K. Eng, A. L. McCormack, J. R. Yates, *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976 – 989.

[12] a) D. N. Perkins, D. J. Pappin, D. M. Creasy, J. S. Cottrell, *Electrophoresis* **1999**, *20*, 3551 – 3567; b) J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, M. Mann, *J. Proteome Res.* **2011**, *10*, 1794 – 1805; c) J. K. Eng, T. A. Jahan, M. R. Hoopmann, *Proteomics* **2013**, *13*, 22 – 24.

[13] A. I. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold, *Anal. Chem.* **2003**, *75*, 4646 – 4658.

[14] a) M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, B. Kuster, *Anal. Bioanal. Chem.* **2007**, *389*, 1017 – 1031; b) M. Bantscheff, S. Lemeer, M. M. Savitski, B. Kuster, *Anal. Bioanal. Chem.* **2012**, *404*, 939 – 965.

[15] a) P. V. Bondarenko, D. Chelius, T. A. Shaler, *Anal. Chem.* **2002**, *74*, 4741 – 4749; b) W. Wang, H. Zhou, H. Lin, S. Roy, T. A. Shaler, L. R. Hill, S. Norton, P. Kumar, M. Anderle, C. H. Becker, *Anal. Chem.* **2003**, *75*, 4818 – 4826.

[16] S. A. Gerber, J. Rush, O. Stemman, M. W. Kirschner, S. P. Gygi, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6940 – 6945.

[17] a) C. Ludwig, M. Claassen, A. Schmidt, R. Aebersold, *Mol. Cell. Proteomics* **2012**, *11*, M111 013987; b) M. Wühr, R. M. Freeman, Jr., M. Presler, M. E. Horb, L. Peshkin, S. Gygi, M. W. Kirschner, *Curr. Biol.* **2014**, *24*, 1467 – 1475.

[18] a) S. E. Ong, M. Mann, *Nat. Chem. Biol.* **2005**, *1*, 252 – 262; b) J. S. Andersen, C. J. Wilkinson, T. Mayor, P. Mortensen, E. A. Nigg, M. Mann, *Nature* **2003**, *426*, 570 – 574.

[19] J. Cox, M. Y. Hein, C. A. Luber, I. Paron, N. Nagaraj, M. Mann, *Mol. Cell. Proteomics* **2014**, *13*, 2513 – 2526.

[20] a) L. Hultin-Rosenberg, J. Forshed, R. M. Branca, J. Lehtio, H. J. Johansson, *Mol. Cell. Proteomics* **2013**, *12*, 2021 – 2031; b) B. Carrillo, C. Yanofsky, S. Laboissiere, R. Nadon, R. E. Kearney, *Bioinformatics* **2010**, *26*, 98 – 103.

[21] B. J. Webb-Robertson, H. K. Wiberg, M. M. Matzke, J. N. Brown, J. Wang, J. E. McDermott, R. D. Smith, K. D. Rodland, T. O. Metz, J. G. Pounds, K. M. Waters, *J. Proteome Res.* **2015**, *14*, 1993 – 2001.

[22] a) C. Lazar, L. Gatto, M. Ferro, C. Bruley, T. Burger, *J. Proteome Res.* **2016**, *15*, 1116 – 1125; b) R. Wei, J. Wang, M. Su, E. Jia, S. Chen, T. Chen, Y. Ni, *Sci. Rep.* **2018**, *8*, 663.

[23] a) F. Meier, A. D. Brunner, S. Koch, H. Koch, M. Lubeck, M. Krause, N. Goedecke, J. Decker, T. Kosinski, M. A. Park, N. Bache, O. Hoerning, J. Cox, O. Rather, M. Mann, *Mol. Cell. Proteomics* **2018**, *17*, 2534 – 2545; b) G. Espadas, E. Borras, C. Chiva, E. Sabido, *Proteomics* **2017**, *17*, 1600416.

[24] S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, M. Mann, *Mol. Cell. Proteomics* **2002**, *1*, 376 – 386.

[25] a) S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, R. Aebersold, *Nat. Biotechnol.* **1999**, *17*, 994 – 999; b) A. C. Tolonen, W. Haas, *J. Vis. Exp.* **2014**, *89*, e51416.

[26] Z. Li, R. M. Adams, K. Chourey, G. B. Hurst, R. L. Hettich, C. Pan, *J. Proteome Res.* **2012**, *11*, 1582 – 1590.

[27] a) A. E. Merrill, A. S. Hebert, M. E. MacGilvray, C. M. Rose, D. J. Bailey, J. C. Bradley, W. W. Wood, M. El Masri, M. S. Westphall, A. P. Gasch, J. J. Coon, *Mol. Cell. Proteomics* **2014**, *13*, 2503 – 2512; b) K. A. Overmyer, S. Tyanova, A. S. Hebert, M. S. Westphall, J. Cox, J. J. Coon, *Nat. Protoc.* **2018**, *13*, 293 – 306; c) A. S. Hebert, A. E. Merrill, D. J. Bailey, A. J. Still, M. S. Westphall, E. R. Strieter, D. J. Pagliarini, J. J. Coon, *Nat. Methods* **2013**, *10*, 332 – 334.

[28] a) E. C. Yi, M. Marelli, H. Lee, S. O. Purvine, R. Aebersold, J. D. Aitchison, D. R. Goodlett, *Electrophoresis* **2002**, *23*, 3205 – 3216; b) S. Pfammatter, E. Bonneil, F. P. McManus, S. Prasad, D. J. Bailey, M. Belford, J. J. Dunyach, P. Thibault, *Mol. Cell. Proteomics* **2018**, *17*, 2051 – 2067; c) F. Meier, P. E. Geyer, S. Virreira Winter, J. Cox, M. Mann, *Nat. Methods* **2018**, *15*, 440 – 448.

[29] D. C. Stahl, K. M. Swiderek, M. T. Davis, T. D. Lee, *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 532 – 540.

[30] a) J. D. Venable, M. Q. Dong, J. Wohlschlegel, A. Dillin, J. R. Yates, *Nat. Methods* **2004**, *1*, 39 – 45; b) L. C. Gillet, P. Navarro, S. Tate, H. Rost, N. Selevsek, L. Reiter, R. Bonner, R. Aebersold, *Mol. Cell. Proteomics* **2012**, *11*, O111 016717; c) M. J. Rardin, B. Schilling, L. Y. Cheng, B. X. MacLean, D. J. Sorensen, A. K. Sahu, M. J. MacCoss, O. Vitek, B. W. Gibson, *Mol. Cell. Proteomics* **2015**, *14*, 2405 – 2419.

[31] J. D. Egertson, A. Kuehn, G. E. Merrihew, N. W. Bateman, B. X. MacLean, Y. S. Ting, J. D. Canterbury, D. M. Marsh, M. Kellmann, V. Zabrouskov, C. C. Wu, M. J. MacCoss, *Nat. Methods* **2013**, *10*, 744 – 746.

[32] C. Ludwig, L. Gillet, G. Rosenberger, S. Amon, B. C. Collins, R. Aebersold, *Mol. Syst. Biol.* **2018**, *14*, e8126.

[33] a) B. C. Searle, L. K. Pino, J. D. Egertson, Y. S. Ting, R. T. Lawrence, B. X. MacLean, J. Villen, M. J. MacCoss, *Nat. Commun.* **2018**, *9*, 5128; b) C. C. Tsou, D. Avtonomov, B. Larsen, M. Tucholska, H. Choi, A. C. Gingras, A. I. Nesvizhskii, *Nat. Methods* **2015**, *12*, 258 – 264; c) H. L. Rost, G. Rosenberger, P. Navarro, L. Gillet, S. M. Miladinovic, O. T. Schubert, W. Wolski,

B. C. Collins, J. Malmstrom, L. Malmstrom, R. Aebersold, *Nat. Biotechnol.* **2014**, *32*, 219 – 223; d) R. Bruderer, O. M. Bernhardt, T. Gandhi, S. M. Miladinovic, L. Y. Cheng, S. Messner, T. Ehrenberger, V. Zanotelli, Y. Butscheid, C. Escher, O. Vitek, O. Rinner, L. Reiter, *Mol. Cell. Proteomics* **2015**, *14*, 1400 – 1410.

[34] A. Thompson, J. Schafer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann, R. Johnstone, A. K. Mohammed, C. Hamon, *Anal. Chem.* **2003**, *75*, 1895 – 1904.

[35] L. Dayon, A. Hainard, V. Licker, N. Turck, K. Kuhn, D. F. Hochstrasser, P. R. Burkhard, J. C. Sanchez, *Anal. Chem.* **2008**, *80*, 2921 – 2931.

[36] P. L. Ross, Y. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, D. J. Pappin, *Mol. Cell. Proteomics* **2004**, *3*, 1154 – 1169.

[37] M. Stadlmeier, J. Bogena, M. Wallner, M. Wühr, T. Carell, *Angew. Chem. Int. Ed.* **2018**, *57*, 2958 – 2962; *Angew. Chem.* **2018**, *130*, 3008 – 3013.

[38] M. Wühr, W. Haas, G. C. McAlister, L. Peshkin, R. Rad, M. W. Kirschner, S. P. Gygi, *Anal. Chem.* **2012**, *84*, 9214 – 9221.

[39] P. Pichler, T. Kocher, J. Holzmann, M. Mazanek, T. Taus, G. Ammerer, K. Mechtler, *Anal. Chem.* **2010**, *82*, 6549 – 6558.

[40] W. Yan, J. Luo, M. Robinson, J. Eng, R. Aebersold, J. Ranish, *Mol. Cell. Proteomics* **2011**, *10*, M110 005611.

[41] P. Mertins, N. D. Udeshi, K. R. Clauser, D. R. Mani, J. Patel, S. E. Ong, J. D. Jaffe, S. A. Carr, *Mol. Cell. Proteomics* **2012**, *11*, M111 014423.

[42] M. Sonnett, E. Yeung, M. Wühr, *Anal. Chem.* **2018**, *90*, 5032 – 5039.

[43] a) B. Ruprecht, J. Zecha, D. P. Zolg, B. Kuster, *Methods Mol. Biol.* **2017**, *1550*, 83 – 98; b) A. Edwards, W. Haas, *Methods Mol. Biol.* **2016**, *1394*, 1 – 13.

[44] J. D. O'Connell, J. A. Paulo, J. J. O'Brien, S. P. Gygi, *J. Proteome Res.* **2018**, *17*, 1934 – 1942.

[45] R. A. Scheltema, J. P. Hauschild, O. Lange, D. Hornburg, E. Denisov, E. Damoc, A. Kuehn, A. Makarov, M. Mann, *Mol. Cell. Proteomics* **2014**, *13*, 3698 – 3708.

[46] a) N. A. Karp, W. Huber, P. G. Sadowski, P. D. Charles, S. V. Hester, K. S. Lilley, *Mol. Cell. Proteomics* **2010**, *9*, 1885 – 1897; b) C. D. Wenger, M. V. Lee, A. S. Hebert, G. C. McAlister, D. H. Phanstiel, M. S. Westphall, J. J. Coon, *Nat. Methods* **2011**, *8*, 933 – 935; c) L. Ting, R. Rad, S. P. Gygi, W. Haas, *Nat. Methods* **2011**, *8*, 937 – 940.

[47] G. C. McAlister, D. P. Nusinow, M. P. Jedrychowski, M. Wühr, E. L. Huttlin, B. K. Erickson, R. Rad, W. Haas, S. P. Gygi, *Anal. Chem.* **2014**, *86*, 7150 – 7158.

[48] a) M. M. Savitski, F. B. Reinhard, H. Franken, T. Werner, M. F. Savitski, D. Eberhard, D. Martinez Molina, R. Jafari, R. B. Dovega, S. Klaeger, B. Kuster, P. Nordlund, M. Bantscheff, G. Drewes, *Science* **2014**, *346*, 1255784; b) H. Keshishian, M. W. Burgess, H. Specht, L. Wallace, K. R. Clauser, M. A. Gillette, S. A. Carr, *Nat. Protoc.* **2017**, *12*, 1683 – 1701.

[49] S. Y. Ow, M. Salim, J. Noirel, C. Evans, I. Rehman, P. C. Wright, *J. Proteome Res.* **2009**, *8*, 5347 – 5355.

[50] a) J. J. O'Brien, J. D. O'Connell, J. A. Paulo, S. Thakurta, C. M. Rose, M. P. Weekes, E. L. Huttlin, S. P. Gygi, *J. Proteome Res.* **2018**, *17*, 590 – 599; b) B. Bai, H. Tan, V. R. Pagala, A. A. High, V. P. Ichhaporia, L. Hendershot, J. Peng, *Methods Enzymol.* **2017**, *585*, 377 – 395; c) E. Ahrné, T. Glatter, C. Vigano, C. Schubert, E. A. Nigg, A. Schmidt, *J. Proteome Res.* **2016**, *15*, 2537 – 2547.

[51] a) A. G. Marshall, T. C. L. Wang, T. L. Ricca, *J. Am. Chem. Soc.* **1985**, *107*, 7893 – 7897; b) F. W. McLafferty, D. B. Stauffer, S. Y. Loh, E. R. Williams, *Anal. Chem.* **1987**, *59*, 2212 – 2213; c) L. Chen, A. G. Marshall, *Int. J. Mass Spectrom. Ion Processes* **1987**, *79*, 115 – 125; d) J. C. Schwartz, M. W. Senko, J. E. Syka, *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 659 – 669.

[52] a) M. Wühr, T. Guttler, L. Peshkin, G. C. McAlister, M. Sonnett, K. Ishihara, A. C. Groen, M. Presler, B. K. Erickson, T. J. Mitchison, M. W. Kirschner, S. P. Gygi, *Curr. Biol.* **2015**, *25*, 2663 – 2671; b) L. Peshkin, L. Ryazanova, M. Wühr, *bioRxiv* **2019**, DOI: https://dx.doi.org/10.1101/210476.

[53] J. Graumann, R. A. Scheltema, Y. Zhang, J. Cox, M. Mann, *Mol. Cell. Proteomics* **2012**, *11*, M111.013185.

[54] K. E. Swearingen, R. L. Moritz, *Expert Rev. Proteomics* **2012**, *9*, 505 – 517.

[55] A. N. Kozhinov, Y. O. Tsybin, *Anal. Chem.* **2012**, *84*, 2850 – 2856.

[56] S. Virreira Winter, F. Meier, C. Wichmann, J. Cox, M. Mann, F. Meissner, *Nat. Methods* **2018**, *15*, 527 – 530.

[57] a) A. Palmese, C. De Rosa, G. Chiappetta, G. Marino, A. Amoresano, *Anal. Bioanal. Chem.* **2012**, *404*, 1631 – 1635; b) K. T. Pan, Y. Y. Chen, T. H. Pu, Y. S. Chao, C. Y. Yang, R. D. Bomgarden, J. C. Rogers, T. C. Meng, K. H. Khoo, *Antioxid. Redox Signaling* **2014**, *20*, 1365 – 1381.

[58] a) T. Werner, I. Becher, G. Sweetman, C. Doce, M. M. Savitski, M. Bantscheff, *Anal. Chem.* **2012**, *84*, 7188 – 7194; b) G. C. McAlister, E. L. Huttlin, W. Haas, L. Ting, M. P. Jedrychowski, J. C. Rogers, K. Kuhn, I. Pike, R. A. Grothe, J. D. Blethrow, S. P. Gygi, *Anal. Chem.* **2012**, *84*, 7469 – 7478.

[59] A. Pierce, R. D. Unwin, C. A. Evans, S. Griffiths, L. Carney, L. Zhang, E. Jaworska, C. F. Lee, D. Blinco, M. J. Okoniewski, C. J. Miller, D. A. Bitton, E. Spooncer, A. D. Whetton, *Mol. Cell. Proteomics* **2008**, *7*, 853 – 863.

[60] F. Xiang, H. Ye, R. Chen, Q. Fu, L. Li, *Anal. Chem.* **2010**, *82*, 2817 – 2825.

[61] D. C. Frost, T. Greer, L. Li, *Anal. Chem.* **2015**, *87*, 1646 – 1654.

[62] C. R. Braun, G. H. Bird, M. Wühr, B. K. Erickson, R. Rad, L. D. Walensky, S. P. Gygi, W. Haas, *Anal. Chem.* **2015**, *87*, 9855 – 9863.

[63] a) V. Lange, P. Picotti, B. Domon, R. Aebersold, *Mol. Syst. Biol.* **2008**, *4*, 222; b) T. A. Addona, S. E. Abbatiello, B. Schilling, S. J. Skates, D. R. Mani, D. M. Bunk, C. H. Spiegelman, L. J. Zimmerman, A. J. Ham, H. Keshishian, S. C. Hall, S. Allen, R. K. Blackman, C. H. Borchers, C. Buck, H. L. Cardasis, M. P. Cusack, N. G. Dodder, B. W. Gibson, J. M. Held, T. Hiltke, A. Jackson, E. B. Johansen, C. R. Kinsinger, J. Li, M. Mesri, T. A. Neubert, R. K. Niles, T. C. Pulsipher, D. Ransohoff, H. Rodriguez, P. A. Rudnick, D. Smith, D. L. Tabb, T. J. Tegeler, A. M. Variyath, L. J. Vega-Montoto, A. Wahlander, S. Waldemarson, M. Wang, J. R. Whiteaker, L. Zhao, N. L. Anderson, S. J. Fisher, D. C. Liebler, A. G. Paulovich, F. E. Regnier, P. Tempst, S. A. Carr, *Nat. Biotechnol.* **2009**, *27*, 633 – 641; c) A. C. Peterson, J. D. Russell, D. J. Bailey, M. S. Westphall, J. J. Coon, *Mol. Cell. Proteomics* **2012**, *11*, 1475 – 1488.

[64] B. K. Erickson, C. M. Rose, C. R. Braun, A. R. Erickson, J. Knott, G. C. McAlister, M. Wühr, J. A. Paulo, R. A. Everley, S. P. Gygi, *Mol. Cell* **2017**, *65*, 361 – 370.

[65] a) J. A. Paulo, F. E. McAllister, R. A. Everley, S. A. Beausoleil, A. S. Banks, S. P. Gygi, *Proteomics* **2015**, *15*, 462 – 473; b) J. A. Paulo, S. P. Gygi, *Proteomics* **2015**, *15*, 474 – 486.

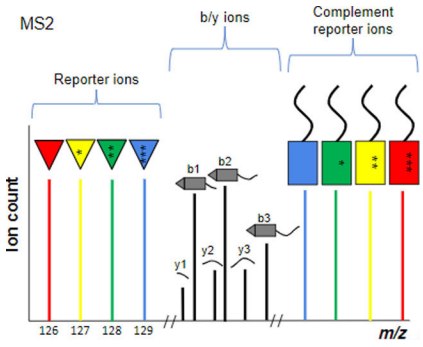
[66] a) J. M. Chick, S. C. Munger, P. Simecek, E. L. Huttlin, K. Choi, D. M. Gatti, N. Raghupathy, K. L. Svenson, G. A. Churchill, S. P. Gygi, *Nature* **2016**, *534*, 500 – 505; b) J. D. Lapek, Jr., P. Greninger, R. Morris, A. Amzallag, I. Pruteanu-Malinici, C. H. Benes, W. Haas, *Nat. Biotechnol.* **2017**, *35*, 983 – 989; c) P. A. Rudnick, S. P. Markey, J. Roth, Y. Mirokhin, X. Yan, D. V. Tchekhovskoi, N. J. Edwards, R. R. Thangudu, K. A. Ketchum, C. R. Kinsinger, M. Mesri, H. Rodriguez, S. E. Stein, *J. Proteome Res.* **2016**, *15*, 1023 – 1032.

Manuscript received: October 23, 2018
Revised manuscript received: December 27, 2018
Accepted manuscript online: January 4, 2019
Version of record online: ■■ ▮▮, 0000

*N. Pappireddi, L. Martin, M. Wühr**

■■ – ■■

A Review on Quantitative Multiplexed Proteomics



Finding a needle in a haystack: Multiplexed proteomics allows the quantification of thousands of proteins among multiple samples. Peptides from multiple samples, labeled by isobaric tags, are indistinguishable in the MS1 spectrum, but separately quantifiable in the MS2 spectrum. The current state of various multiplexed proteomics techniques is reviewed and compared with alternative techniques.