Mass spectrometry–based proteomics in cell biology

Tobias C. Walther[1,3] and Matthias Mann[2]

[1]Organelle Architecture and Dynamics and [2]Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, D-82152 Martinsried, Germany
[3]Department of Cell Biology, Yale University School of Medicine, New Haven, CT 06520

The global analysis of protein composition, modifications, and dynamics are important goals in cell biology. Mass spectrometry (MS)–based proteomics has matured into an attractive technology for this purpose. Particularly, high resolution MS methods have been extremely successful for quantitative analysis of cellular and organellar proteomes. Rapid advances in all areas of the proteomic workflow, including sample preparation, MS, and computational analysis, should make the technology more easily available to a broad community and turn it into a staple methodology for cell biologists.

The principle challenge of cell biology is to reveal the mechanisms and inner workings of cells. In this quest, cells are more and more perceived as systems in which the dynamic interplay of a large number of components determines the output of many biological processes occurring in parallel. To characterize these processes and to reveal their underlying principles, one needs to evaluate the dynamic composition and localization of the molecular components. All cellular processes involve proteins and their characterization has therefore drawn most interest over the years. However, it has been technically challenging to determine their abundance, modification state, and localization in a systematic way. In the absence of system-wide technologies, targeted approaches are currently used to measure the abundance and localization of specific proteins of interest. These rely on the availability of antibodies or epitope-tagged versions of the proteins to detect them by Western blot or microscopy. These workhorse techniques of cell biologists have allowed for the extensive characterization of many cellular processes. However, they often just open a small window into the complex world governing the organization of the cell and highlight only a small part of a large interconnected network of functionally and physically interacting proteins.

For these reasons, there is a great need for techniques that allow the unbiased analysis of cellular compositions under changing conditions. A great breakthrough in this direction was the advent of microarrays, which enable the global quantification of gene expression. By default, mRNA quantification was also used as a proxy for measuring changes of protein abundance. Although this has resulted in a dramatic increase in our knowledge of many cellular responses, studying protein levels directly would be advantageous because mRNA levels often do not correlate well with protein abundance. This is because protein levels are determined by complex posttranscriptional processes, where every step in the life-cycle of a protein, from its synthesis to its degradation, is subject to regulatory input. Furthermore, the central role of covalent protein modifications such as phosphorylation, acetylation, and glycosylation in cellular physiology as signals in information processing or as marks mediating protein associations is becoming increasingly appreciated. These modifications can also guide assembly of proteins into large macromolecular machines or instruct their localization to different organelles.

Among different possible approaches to study proteins, mass spectrometry (MS)–based proteomics is increasingly used to acquire the data important for understanding these processes. This technology is rapidly advancing and in modern proteomics it has essentially completely replaced previous tools such as two-dimensional gel electrophoresis. Within the field of MS-based proteomics there is still a great variety of approaches and instruments, which can be confusing to the outsider. Here, we will mainly focus on one particular pipeline for high resolution MS-based proteomics that has proven robust and successful in our hands (Fig. 1). It can be used to derive the protein composition of a cell, to determine the members of proteins complexes, their architecture, the protein inventory of organelles, and the dynamics of these processes. It can also be readily combined with the analysis of the posttranslational modification state of proteins and their dynamics. However, particularly for these more advanced applications, routine availability of proteomics in core facilities lags far behind pioneering studies reported in the literature (Bell et al., 2009). We hope that by focusing on one prototypical and robust workflow and exemplary applications, we will help to break down communication barriers between cell biologists and mass spectrometrists. Of course, different setups are also used very productively and alternative approaches exist for each step. We refer the reader to in-depth reviews on these topics here and below (Yates et al., 2005; Domon and Aebersold,

Figure 1. **A prototypical proteomics pipeline.** Different samples are processed and fed into the proteomics pipeline, yielding different results depending on the application. The pipeline consists of several steps, listed in the different panels: proteolytic digest (1), the separation and ionization of peptides by liquid chromatography (LC) coupled to electrospray ionization (2), their analysis by MS (3), fragmentation of selected peptides and analysis of the resulting MS/MS spectra (4) and data analysis, including identification and quantification of proteins from several detected peptides, as well as downstream bioinformatic analysis depending on the specific application (5). The peptide peaks in panel 3 represent isotopically labeled variants of the same peptide, in this case using the SILAC methodology. Their intensity ratio is used for relative quantitation of proteins present in two samples. In the workflow described here, mass spectrometry is performed with a linear ion trap–Orbitrap instrument (LTQ-Orbitrap) and computational proteomics is done within the MaxQuant framework. LC-MS/MS, liquid chromatography coupled online to mass spectrometry; MS, mass spectrometry; MS/MS, tandem mass spectrometry; PTM, post-translational protein modification; Protein ID, protein identification; m/z, mass- to-charge ratio.
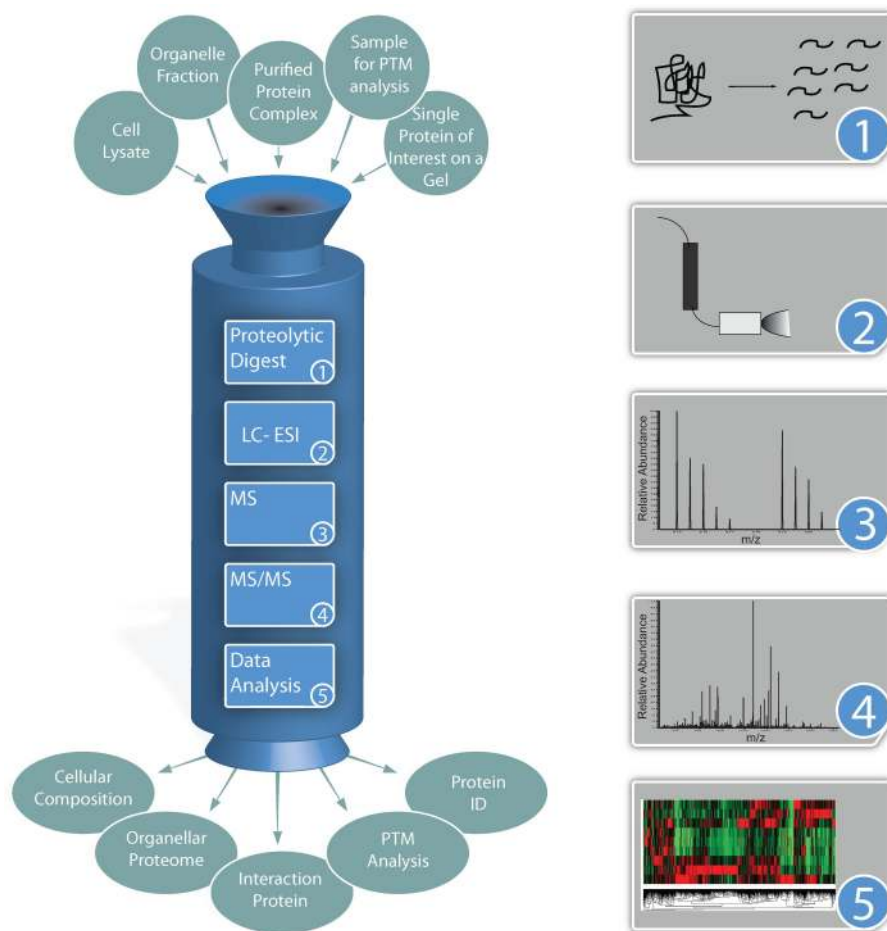
2006; Jensen, 2006; Cravatt et al., 2007; Gingras et al., 2007; Trinkle-Mulcahy et al., 2008).

The language and principles of MS-based proteomics

MS is a way to accurately measure the weight of a molecule—or more accurately its mass-to-charge ratio (m/z). Because mass analysis uses electromagnetic fields in a vacuum, molecules must first be electrically charged and transferred into the gas phase. In the pipeline that we describe here, both tasks are accomplished by electrospray ionization, which was developed by John B. Fenn and for which he shared the Nobel Prize in chemistry in 2002 (Fenn et al., 1989). Once in the gas phase, the m/z ratio of molecules is determined by their trajectories in a static or dynamic electric field. For example, a quadrupole mass filter can be set to only transmit ions of a particular m/z and by scanning through a range of m/z values a mass spectrum can be obtained. Other popular MS instrument types include quadrupole–time of flight (TOF) instruments, in which a quadrupole mass filter is coupled to a TOF analyzer that distinguishes the molecules by their arrival times at a detector (Ens and Standing, 2005). Alternatively, ions are captured by the field of an ion trap where they can be accumulated and manipulated for further analysis. In the described pipeline, we use a combination of a linear ion trap with an Orbitrap, in which ions circulate around a central, spindle-shaped electrode (Makarov, 2000; Hardman

and Makarov, 2003; Scigelova and Makarov, 2006). The axial frequency of oscillations of the ions on this trajectory is proportional to the square root of m/z. Because this frequency can be determined with high precision, the m/z is measured very accurately. The Orbitrap also has very high mass spectrometric resolution, which is defined as the width of the peak at half height divided by the mass of the peak (and is therefore a dimensionless number). Resolution commonly achieved in proteomics has risen within the last decade from just a few hundred in ion traps to 60,000 in current Orbitraps. Resolution is just as important in proteomics as it is in microscopy or in structural biology: with low resolution, peptides are effectively merged into common peaks, whereas high resolution allows the mass spectrometer to distinguish hundreds of thousands of different peptides from each other, a precondition for their accurate identification and quantification.

To determine the mass of an analyte, such as a peptide, from the m/z value, the charge state of the molecule is first derived from the pattern of naturally occurring isotopes of different masses. This pattern is mainly caused by ^{13}C, which occurs with a low natural frequency ($\sim$1% of the main ^{12}C isotope). Natural compounds have many carbon atoms and therefore show a family of peaks representing one, two, or more ^{13}C atoms integrated into the molecule. If the distance between peaks is one unit on the m/z scale, the charge of the peptide was one, if it was 0.5, then the charge was two.

For proteomics, the first idea may be to characterize proteins by their unique weight, which is a function of their composition. For large proteins the mass differences between different proteins with similar composition is small and entire proteins are anyway difficult to measure (this is the topic of a proteomic specialty called "top down proteomics"; McLafferty et al., 2007). Therefore, for most experiments, not the mass of entire proteins, but of peptides derived from them by enzymatic cleavage, is measured ("bottom up proteomics"). For a mixture of peptides, this yields the MS-spectrum of mass-to-charge ratios plotted against their mass spectrometric signal, the ion current. To determine the identity (i.e., sequence), in addition to the exact mass of a peptide, it is fragmented along its backbone, usually by collision with an inert gas such as helium or nitrogen at low pressure (CID, collision induced dissociation). The resulting spectrum, called an MS/MS (or tandem or MS^2) spectrum, is basically a list of m/z ratios for different fragments with some of the differences corresponding to the specific mass of one amino acid. In principle, connecting the fragments with increasing size from the N terminus (b-ion series) or C terminus (y-ion series) allows for the deduction of the peptide sequence from the series of specific mass differences, each corresponding to a successive amino acid (de novo sequencing). However, the gas phase chemistry of decomposing protonated peptides is quite intricate (Paizs and Suhai, 2005),

Glossary of terms:

Liquid chromatography coupled online to mass spectrometry (LC-MS): Liquid chromatography by HPLC and usually reverse-phase chromatography, directly injecting samples via electrospray ionization into a mass spectrometer for analysis during the chromatographic run.

Electrospray ionization: Analyte molecules such as peptides are dissolved in liquid that passes through a needle at high electric potential. The applied voltage causes the liquid to disperse into small, highly charged droplets, which evaporate and transfer the analyte molecules into atmosphere in an ionized form (usually by multiple protons). From here they are transferred into the vacuum of a mass spectrometer.

Time of flight mass analyzer (TOF): Mass analyzer where the m/z is determined by the time ions need to travel through an electric field. Due to the pulsed nature of the measurement, this type of analyzer was most often coupled to matrix-assisted laser desorption (MALDI) of peptides, where a pulse laser leads to the release and ionization of peptides from a dried mixture of sample and matrix. Today, it is also frequently used with electrospray in hybrid mass spectrometers of a quadrupole–time of flight format.

Linear ion trap–Orbitrap: A hybrid mass spectrometer consisting of a combination of a linear ion trap with an Orbitrap analyzer. The Orbitrap is a type of Fourier transform (FT) ion trap mass analyzer developed by A. Makarov where the ions oscillate along and around a central spindle. The square root of the m/z of trapped ions is proportional to the frequency of oscillations along the axis, which can be measured with very high precision. In Fourier transform mass spectrometers, the frequencies of the different components of the spectrum are transformed by a Fourier transform to yield the mass spectrum.

Collision-induced dissociation (CID): Fragmentation of peptide ions by energy acquired during collisions with a chemically inert gas. In ion tap instruments, CID is most often performed by an auxiliary, resonating electric field that specifically excites the ions with the targeted mass-to-charge ratio. In quadrupole–time of flight (TOF) instruments, the kinetic energy is acquired by acceleration of peptides into a collision chamber containing the inert gas at a low pressure.

Stable isotope labeling by amino acids in cell culture (SILAC): SILAC is one of the metabolic labeling techniques in quantitative proteomics. Usually "heavy" ^{13}C- and ^{15}N-labeled forms of lysine, arginine, or both are incorporated in a cell line or organism using metabolic labeling. In the MS analysis, these heavy amino acids are distinguished from their light, unlabeled counterparts by a characteristic mass shift. Note that these are nonradioactive forms of the amino acids.

iTRAQ: Quantitative proteomics technique that uses isobaric tags to chemically modify sample and control proteins. Peptides derived from these proteins can be distinguished via low mass reporter ions in their fragmentation spectra. The ratio of the reporter ions indicates the relative amounts of a peptide in each sample.

Single reaction monitoring (SRM) and multiple reaction monitoring (MRM): Specialized MS technique in which one quadrupole mass analyzer is set to transmit a peptide precursor ion of choice, a second quadrupole functions as a collision chamber (see CID), and a third quadrupole is set to a previously determined fragment mass of the peptide to be monitored. In SRM, a single transition is probed but to increase specificity of this technique it is more common to measure 3 to 8 transitions for each targeted peptide (MRM).

Protein correlating profiles (PCP): The localization of a protein in a specific organelle is determined by quantitatively comparing its abundance in the fractions of a purification with the behavior of marker proteins for that organelle.

and it is much easier and more common to match the measured fragment spectrum and peptide mass against a protein database with a search engine (Sadygov et al., 2004). For each protein, several peptides are measured and each contributes with a database identification score, which should lead to highly confident identification. The most popular commercially available peptide search engines are Mascot (Perkins et al., 1999) and SEQUEST (Eng et al., 1994), whereas X!Tandem is an example of an open source program (Craig and Beavis, 2004). Despite the automated nature of searching database with MS data, the cell biologist should keep a critical attitude toward search results, and if possible, try to verify key identifications using the underlying primary data. This is particularly important in the case of low resolution spectra and when identifying modified peptides.

For most experiments, such as the determination of the proteome composition of an organelle or a protein complex, one

needs to detect proteins in complex mixtures. When proteins are digested to peptides, mixtures of even higher complexity are generated. The most robust method to reduce this complexity is one-dimensional SDS gel electrophoresis of proteins, followed by in-gel digestion by a protease (typically trypsin) and by extraction of the resulting peptides from the gel (Shevchenko et al., 1996). The alternative is to digest proteins in solution, avoiding the tedious gel separation and extraction step (Link et al., 1999; Washburn et al., 2001). The recently introduced filter-aided sample preparation (FASP) method combines the advantages of complete solubilization in SDS with the advantages of in-solution digestion (Wiśniewski et al., 2009).

A key challenge in MS-based proteomics is to resolve and detect peptides in very complex mixtures. To this end, very low flow and narrow-bore liquid chromatography is coupled directly via electrospray to the mass spectrometer (LC-MS/MS). In this

mode, the sample is analyzed continuously as it elutes from an HPLC column. If very deep coverage of the composition of the sample is desired or if specific peptides should be enriched, e.g., modified by phosphorylation, an additional separation step, often using strong cation exchange (SCX), is used before analyzing the samples. Additional fractionation is only seldom applied because the number of samples that need to be measured multiply with each such step. In addition, classical biochemical fractionations, even if very vigorously performed, often yield fractions overlapping in composition. For MS, such additional fractionation steps often do not outweigh the increased analysis time. Instead, the very high mass accuracy and resolution of modern MS analyzers is used to distinguish between closely spaced peaks in spectra. In many proteomics applications the combination of such high resolution MS with good HPLC separation yields peptide spectra sufficiently well resolved for a comprehensive detection of proteins present in a sample.

In the analysis pipeline described here, an MS spectrum is obtained roughly every 2 s. From this, typically up to 10 peptides are selected for fragmentation and the MS/MS spectrum is recorded from each simultaneously with the MS spectrum. Over the usual 2 h chromatography run, an enormous amount of data are therefore collected and its analysis is a severe challenge for proteomics. Many efforts have been made to automate parts or all of the computational tasks associated with proteomics. One of the earliest was the trans-proteomic pipeline (Deutsch et al., 2010), but there are many others, such as msInspect (Bellew et al., 2006) or Census (Park et al., 2008).

In the pipeline described here, all computational proteomics tasks are handled by the freely available MaxQuant environment (Cox and Mann, 2008), which is equipped with its own search engine called *Andromeda* (also freely available) as well as with visualization tools that allow verification of database identifications. Analysis in MaxQuant results in much-improved mass accuracy and percentage of successfully identified MS/MS spectra (typically >50%). In addition to the computational proteomics analysis packages, many of the established tools for the analysis of microarray gene expression data within the Bioconductor/R software are applicable to the downstream interpretation of proteomics datasets (Kumar and Mann, 2009).

Quantitative proteomics

Most often it is more important to determine how protein levels change from one condition to the next than it is to know just whether a protein is present or not. Relative abundance of proteins between two cellular states, for example between control and specific perturbation, is therefore a crucial variable in cell biological experiments. The goal of biochemical purifications of organelles or protein complexes is most often to determine whether a protein is enriched in the purified fraction, as opposed to being present merely as a contaminant. The most interesting question may be how the composition of an organelle or a protein complex changes under different conditions, or in different cell types. Similarly, changes in abundance of posttranslational modifications are important, particularly for studying cellular signaling, as the level rather than just the presence of many modifications transmit crucial information in the cell. For example,

in order to understand the systems' response to receptor stimulation, changes of modifications over time should be quantified.

From an MS perspective, all these challenges necessitate the quantification of the relative abundance of peptides in different samples. Comparing two protein complexes or large assemblies, such as the nucleolus, requires knowing the relative abundance of peptides derived from proteins present in two preparations, whereas elucidation of signal response requires measuring changes in the levels of phosphopeptides. Several approaches have been developed to tackle this problem. Mostly, we will focus here on two technologies that are frequently used and that can be combined seamlessly with the analysis pipeline described: isotope labeling and label-free quantification. In one of the metabolic labeling approaches, proteins fully incorporate "heavy", nonradioactive isotope-containing amino acids (stable isotope labeling with amino acids in cell culture [SILAC]; Ong et al., 2002) and are analyzed by high resolution MS. To this end, arginine and/or lysine that is labeled with ^{13}C atoms and/or ^{15}N atoms is fed to cells, which integrate these amino acids into all proteins in the course of several cell doublings. Digestion of these proteins with trypsin, which cuts after arginine or lysine, leads to peptides with a heavy amino acid at their C terminus. The heavy labeled proteome remains distinguishable from the "light" or normal labeled control proteome and the two can be combined at the level of cells or directly after lysis. This prevents differences in sample preparation from influencing quantification accuracy. The resulting mixture contains SILAC peptide pairs recognizable by having the exact mass difference between the heavy and normal amino acids. The relative intensity of the peaks reflects the relative abundance of the proteins in the mixture. Although SILAC was originally developed for work on cell lines, it has been extended to include microorganisms and entire mice and even to the very accurate quantification of the levels of thousands of proteins in human tumor biopsies (Krüger et al., 2008; Geiger et al., 2010).

The alternative to metabolic labeling is chemical modification of peptides by stable isotope–containing tags. The best-known strategy to this end is called iTRAQ. It uses up to eight isobaric tags that react with primary amine groups of peptides. During MS analysis, the tags are fragmented into reporter groups of different mass for each tag. The intensity of the different reporter groups is then used to derive the relative abundance of the corresponding peptides and proteins in the starting mixture (Ross et al., 2004).

Besides stable isotope labeling, so-called "label-free" quantification is increasingly used (Old et al., 2005). The basic idea here is to align separate LC-MS/MS runs of peptide mixtures and to calculate differences in intensities of the same peptides detected in each run. Although quantification using this methodology is less accurate than methods using isotope labels, it may be simpler than isotope-based methods and makes cell types accessible that are difficult or impossible to label with amino acids (Malmström et al., 2009; Luber et al., 2010). In addition to these two, a number of variations and alternative techniques have been developed that are reviewed in detail elsewhere (Ong and Mann, 2005; Bantscheff et al., 2007; Wilm, 2009). As an example, chemical dimethyl labeling of peptides can be

performed economically with isotopically light or heavy reagents (Hsu et al., 2003; Boersema et al., 2009). As usual with chemical techniques in proteomics, it is important to ensure that reactions go to completion. Last not least, if quantification of only a selected subset of proteins is desired, peptides of these proteins can be targeted by a technique called multiple reaction monitoring, or MRM. This requires specialized "triple quadrupole" mass spectrometers, which consist of a selection quadrupole for the precursor ion, a collision cell quadrupole, and a selection quadrupole for the fragments. They are set to exclusively monitor predetermined precursor-to-fragment transitions in rapid succession. In this way the presence and—if an isotope-labeled synthetic peptide analogue is used, the quantity of selected peptides—can be monitored (Wolf-Yadlin et al., 2007; Kitteringham et al., 2009; Unwin et al., 2009).

Analysis of cellular composition and architecture

MS-based proteomics is now closing the gap to gene expression analysis, which measures the abundance of messenger RNAs. We have recently quantified haploid against diploid yeast using the SILAC technology. Essentially all proteins that were found to be expressed by both TAP and GFP tagging of all yeast open reading frames (Ghaemmaghami et al., 2003; Huh et al., 2003) were identified. This demonstrates that complete proteomes can be obtained by MS-based proteomics (de Godoy et al., 2008). MRM has been used on selected proteins in yeast and has identified proteins down to an estimated 43 copies per cell (Picotti et al., 2009). Together with MRM of isotope-labeled reference peptides, the absolute abundance of some proteins was determined in that study. This approach has been combined with cryo-tomography to measure the copy number of cellular structure of the pathogen *Leptospira* (Malmström et al., 2009). The abundance of other cellular proteins can then roughly be approximated by the added peptide intensity for each protein.

All cells, and particularly eukaryotic ones, are characterized by a high degree of spatial organization of biochemical reactions. Therefore, in addition to the inventory of a cell, the localization of proteins is an important second dimension of knowledge. Classically, two approaches are taken to answer this question: microscopy to detect proteins in situ, and biochemical fractionation of cells in organelles and proteins complexes.

Fractionation approaches are now increasingly combined with MS-based proteomics to detect and measure proteins in purified organelles (Yates et al., 2005). Despite the success of early organellar proteomics (Neubauer et al., 1997; Rout et al., 2000), researchers soon realized that in other studies often too many organellar proteins were being identified. In fact, the principle problem of biochemical purifications of organelles is that not all proteins in a fraction are bona fide constituents of the investigated organelle, but might instead be contaminants, copurifying with the organelle. This is particularly a problem for modern MS techniques with extremely high sensitivity. The simplest solution to this problem is "subtractive proteomics", in which the inventory of both a target and a related fraction missing the structure of interest are analyzed. Each proteomic characterization leads to an inventory list, which are "subtracted" from each other.

The remaining proteins only detected in the target fraction are enriched in components of the organelle of interest. Such a strategy was used to characterize nuclear envelope components (Schirmer et al., 2003). However, limitations of this approach include that results depend not only on proteins in the target organelle, but also on those in the control. For the case of the nuclear envelope, the choice of endoplasmic reticulum (ER) as a control is straightforward. For other organelles, such as mitochondria, it is less so. In addition, in very complex fractions only subsets of the proteins may be detected and this detection may be partly stochastic (because the same peptides are not always picked for sequencing by the mass spectrometer). This could lead to false-positive organellar assignments. Conversely, if one achieves to detect all proteins in each fraction, the subtraction list may miss many true hits as they are likely present in minor amounts as contaminants in the control fraction.

Fortunately, this problem of very many low abundant background proteins can be sidestepped using quantitative proteomics, where not only the identity, but also the abundance of proteins in a sample are determined. This information is then used to separate genuine members of organelles from copurifying ones by their quantitative behavior during fractionation. In a crude form, which is nevertheless a large improvement on purely qualitative data reporting on the presence of proteins in a sample, the number of identified peptides of each protein in an organelle fraction or control are compared. These techniques are termed "spectral counting" (Liu et al., 2004; Lu et al., 2007) or "exponentially modified peptide abundance index" (emPAI; Ishihama et al., 2005), and are often built into proteomics data analysis pipelines. They are based on the principle that the likelihood of identifying peptides of a given protein is correlated with the abundance of the protein. True quantification of the signal in MS peaks is much more accurate than spectral counting and can be the basis of bioinformatic, statistical models that assign a probability for localization in the compartment or complex. For example, precise ratios of protein abundance can be obtained by mixing an organelle preparation with different SILAC-labeled control fractions enriched in different organelles. From these ratios a score that reflects the likelihood of localization in each organelle can be developed. For each such combination, one expects a bimodal distribution with proteins truly localized in the target organelle distributing around one mode of ratios and contaminants forming a second mode. The localization score is developed from a Bayesian model assuming overlapping normal distributions of measured ratios. This methodology is particularly suited for samples where stringent purification is either not possible due to limited sample availability or not desirable because many associated proteins are expected to be lost during purification. We have applied this strategy to crude purifications of mitochondria from brown and white adipose tissue, where it revealed significant differences both in the presence and abundance of mitochondrial proteins (including isoforms) between these tissues (Forner et al., 2009).

A second quantitative approach to organellar proteomics is to determine not just abundance of proteins in the target fraction, but to follow their behavior during the purification. The basic idea is that proteins that are together in one organelle

should also copurify with the same abundance pattern over fractions of the purification. These abundance profiles are compared and grouped for similarity with known markers of the target organelle. The resulting protein correlation profiles (PCPs) resemble the calculation of purification factors in classical biochemistry because they reveal not only which proteins are present in the target fraction, but which proteins co-enrich there (Fig. 2). Particularly useful for this method are data derived from purification by density gradients where the profiles of proteins over the length of the gradients are quantified. Applied to centrosome purifications, this led to the recognition of the centrosomal localization of many proteins previously not known to be located there (Andersen et al., 2003). When used in a SILAC format a batch of cells is labeled and used to isolate a target organelle, e.g., by density gradient centrifugation. This labeled sample is then mixed into each fraction of a second purification from unlabeled cells. For each protein, a profile of ratios identified throughout the fractions is then calculated. Similar to their behavior in profiles resulting from other quantitation methods, proteins strongly enriched in the target fraction will have a peak of the ratio only there. Conversely, if a protein is a contaminant, it will peak elsewhere, depending on its main localization in the cell. Another strategy with a similar aim is localization of organelle proteins by isotope tagging (LOPIT). This technology was first demonstrated using principle component analysis of the correlation patterns of proteins in different fractions to reveal membrane proteins residing in the ER or Golgi apparatus of plant cells (Dunkley et al., 2004; Tan et al., 2009).

The literature contains a wide diversity of organellar proteomics datasets, ranging from stringent quantitative experimental designs with very low error rates to purely qualitative lists of proteins in which clearly the majority of proteins do not belong to the organelle under investigation. So how can the quality of organellar assignment be assessed? Currently there is no standard answer to this question, but comparison to data generated independently using different methods may offer valuable positive controls. For example, an atlas of subcellular localization of GFP-tagged fusion proteins can be used to confirm datasets generated for the yeast *Saccharomyces cerevisiae* (Huh et al., 2003). It may also be useful to compare data obtained for different model systems with each other. The nucleolar proteome derived from HeLa cells, for instance, contained homologues of 87% of the yeast nucleolar proteins, implicitly validating both datasets (Andersen et al., 2005).

Gene ontology (GO) terms are defined sets of concepts in cellular components, molecular functions, and biological processes, as well as the relationships between them. These terms describe different aspects of protein characteristics, such as its localization, biochemical activity, and biochemical context. As an easily accessible summary of information, they are often used to assess the quality of a dataset (Ashburner et al., 2000). In addition, data from other sources, such as genetic or RNAi screens is increasingly used to benchmark results from comprehensive proteomics experiments, assuming that knock-down of proteins is likely to result in a phenotype related to the function of the organelle where it is located. In this way, many proteins found by protein correlation profiling to localize to centrosomes
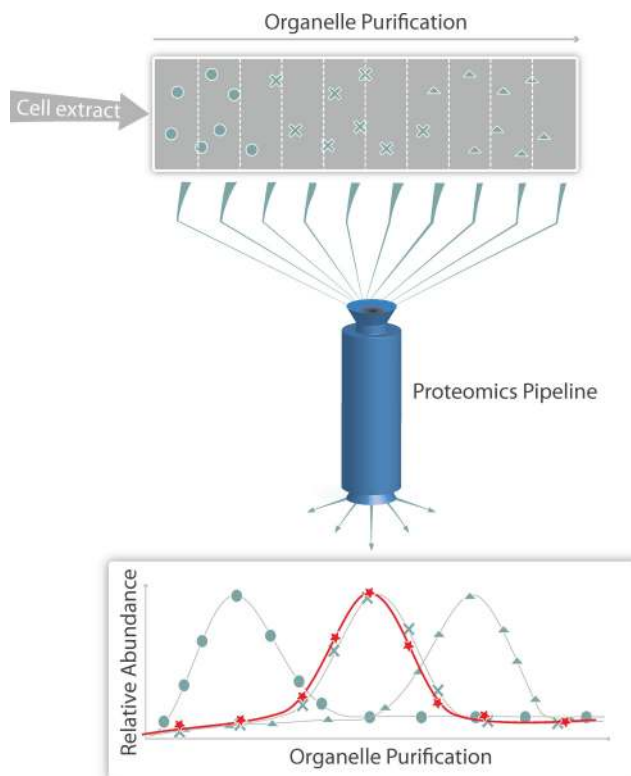

Figure 2. **Organellar proteomics.** Organelles are purified using fractionation of a cell extract, e.g., by gradient centrifugation. The resulting gradient (gray box) is divided into samples and these resulting fractions are analyzed using the proteomics pipeline. Comparison of the abundance profiles of proteins (bottom panel, green lines) throughout all fractions of the purification to the profile of known marker protein (bottom panel, red line) yields classes of proteins that cofractionate with the organelle of interest (protein abundance marked with crosses), or that are present in the peak fraction for this organelle merely as a contaminant (protein abundance marked with circles or triangles).

were later described to cause genetic diseases when mutated. These ciliopathies have the common feature that function of the basal body, the centrosome analogue in post-mitotic cells, is impaired.

In an interesting application of quantitative proteomics to the remodeling of nuclear organelles, Boisvert et al. (2010) measured the quantitative fraction of each of 2,000 proteins in the cytosol, nucleus, and nucleolus, which they term "spatial proteomics." They then studied the redistribution of those proteins as a function of a perturbation, in this case chemically induced DNA damage. The same group also found distinct redistribution of specific nucleolar proteins upon infection of cells with adenovirus (Lam et al., 2010).

In addition to organelles, protein complexes are a fundamental unit of cellular organization. In principle, the same considerations concerning the background of contaminant proteins apply to both organelle and protein complex purifications. Differences in the analysis stem mostly from different purification strategies used for either application. Most often, protein complexes are purified by affinity chromatography of a tagged subunit of the complex. In contrast to organelle purifications that differ strongly for each target, protocols for complex purification

can be standardized to a much higher degree. Similar to the case of organellar purification, the basic principle is to use quantitative proteomics to distinguish background proteins that equally occur in the pulldown with the bait and in the control pulldown from specific binders that preferentially occur in the pulldown with the bait (Fig. 3; Blagoev et al., 2003; Ranish et al., 2003). For a number of purifications, the background can also be estimated from a "bead proteome," representing all proteins bound nonspecifically to the purification bead matrix, or from quantitative proteomics comparing the affinity purification of a target complex with a mock purification (Trinkle-Mulcahy et al., 2008). Because affinity purification yields much higher purification factors and generally fewer contaminants than organelle enrichments, label-free approaches are often easier for determination of specific interactors, but SILAC can equally be applied (Hubner et al., 2010). In either case, quantitation is essential to efficiently separate true binders from hundreds of background proteins. When combined with labeled peptide standards, proteomics can also be used to determinate the stoichiometry of complexes (Wepf et al., 2009).

MS-based proteomics has been applied to determine the interactome of the yeast model organism, using genomically integrated tags for purification (Gavin et al., 2006; Krogan et al., 2006). Initially in proteomics projects, stringent purification with tandem affinity purification (TAP) tags (Rigaut et al., 1999) was widely used, but quantitative methods now allow reverting to milder, single-step purification, increasing the chance to retain weak interactors. GFP is particularly useful as a purification tag because it allows for localization by fluorescence microscopy of the proteins of interest in the same cells where MS analysis is performed (Shou et al., 1999; Cheeseman and Desai, 2005; Cristea et al., 2005). Highly efficient binders to GFP have been developed (Rothbauer et al., 2008).

In such systematic experiments, the availability of larger datasets is an advantage for analysis. Contaminants tend to be shared between many purification experiments. For yeast, where a number of independent genome-wide interactome studies are available, this wealth of information has already been used to integrate the data and develop an interaction map considering the statistical significance of a detected interaction (Collins et al., 2007; Wang et al., 2009). However, similarly to the case of the organellar purifications discussed above, we recommend doing such filtering in a quantitative rather than a qualitative way.

Analysis of cellular dynamics

At the next level of analysis, addition of a time dimension will yield the dynamic properties of cellular systems. This is achieved by quantitative proteomic comparison between different time points. The same quantitative proteomics methods outlined above are applicable here. In such an experiment, the abundance of proteins after a time interval of, for instance, a drug treatment, is compared with the starting abundance (Fig. 4). For example, temporal analysis of protein flux out of an organelle was measured for the nucleolus. A large number of proteins was found to exit the nucleolus after inhibition of transcription with actinomycin D, but interestingly, disassembly of the entire structure was not observed (Andersen et al., 2005).
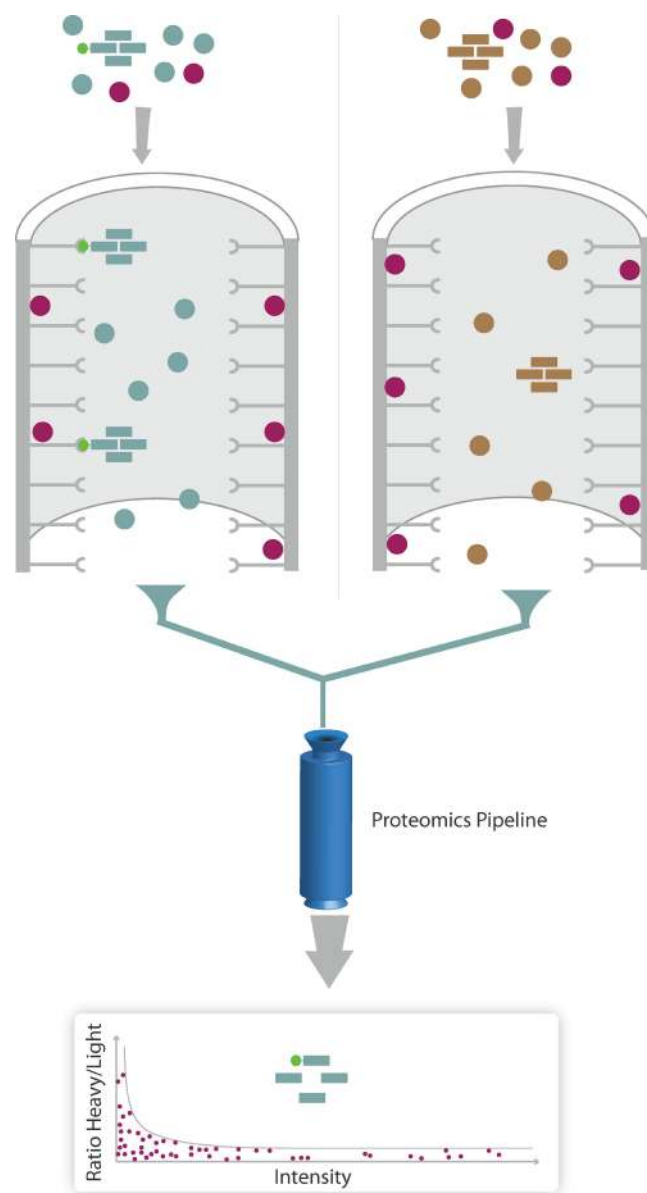

Figure 3. **Interaction proteomics.** To detect specific binders to a tagged protein of interest (e.g., GFP, green sphere on target protein), abundance of potential interaction partners is compared with the abundance of the same protein from a mock purification expressing only untagged versions of the proteins. For example, extracts from "heavy" SILAC-labeled cells expressing tagged proteins (left) or control extracts (right) are subjected to affinity chromatography. The resulting eluates from target and mock purification are mixed and analyzed by the proteomics pipeline. Background binders are present in equal amounts in both purifications (circles in the plot), whereas proteins specifically binding to the bait have high ratios reflecting their specific enrichment from the heavy SILAC-labeled sample (rectangles in the plot).

In a different type of experiment, turnover of proteins in the cell or a specific organelle can be measured. Analogously to pulse-chase experiments, heavy labeled amino acids are used for kinetic measurements. However, in proteomics cells are only pulsed. The proteome can be fully labeled with heavy amino acids followed by a switch to nonlabeled ones at the start of the experiment. The disappearance of the label is then a measure of the protein turnover rate (Beynon and Pratt, 2006).
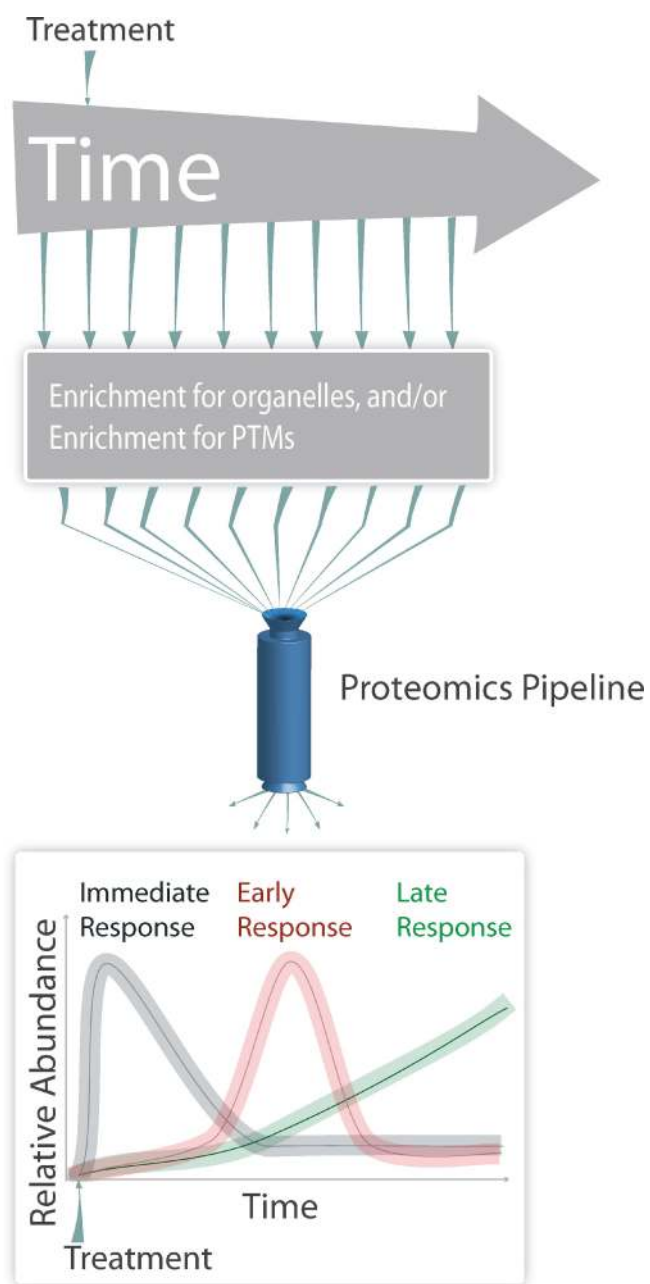
Figure 4. **Time-resolved proteomics.** To detect dynamic proteome changes after a change of conditions, the amount of proteins is followed over time after an initial treatment. For example, cells are treated with a stimulus at the beginning of the experiment and aliquots are analyzed in intervals either directly or after enrichment of organelles, protein complexes, or post-translational modifications (PTMs) by the proteomics pipeline. Plotting the abundance of proteins or modifications over time results in a dynamic view of the abundance of proteins or specific modifications during the time course of the experiment, and may identify different phases of the response (e.g., immediate, early, and late response).

The analysis of proteome dynamics is not restricted to analysis of organelles. Cellular responses to a change in external conditions are often mediated by signal transduction cascades in which the flow of information is represented by dynamic changes in the posttranslational modification of proteins. By far the best studied is phosphorylation of serines, threonines, and tyrosines. Phosphorylated peptides can be enriched up to 100-fold by metal

or antibody affinity chromatography, and this modification is detected by MS as a mass shift of the intact peptide and of the b- and y-fragment ions bearing the phosphorylated amino acid (Wepf et al., 2009).

Combined with quantitative approaches such as SILAC, these types of phosphoproteomics techniques enable the temporal analysis of signaling in a broad, system-wide fashion. Quantitation of the cellular phosphoproteome response to EGF treatment revealed a previously unimagined scope of the cellular response. The temporal dynamics also demonstrated that within one protein, different sites are often regulated with different kinetics (Olsen et al., 2006). Such types of analyses provide an unprecedented density of data for analysis although they are restricted to specialized laboratories for the time being (Nita-Lazar et al., 2008; White, 2008; Boersema et al., 2009).

The analysis of posttranslational modifications is not limited to phosphorylation networks. Systems analysis of protein acetylation on lysines has recently identified more that 3,600 acetylation sites. To reach this depth of coverage, acetylated peptides are enriched using immunoprecipitation by monoclonal antibodies directed against acetylated lysines. Often, this modification occurs on large macromolecular complexes, which suggests that it might regulate these machines (Kim et al., 2006; Choudhary et al., 2009). Consistent with this notion, acetylation of metabolic enzymes in bacterial *Salmonella* cells led to the detection of 191 acetylated sites and together with a study in human liver tissue highlights metabolic regulation by acetylation (Wang et al., 2010; Zhao et al., 2010). Likewise, MS-based proteomics is frequently used for the analysis of ubiquitination in which the Gly-Gly tag is detected as a modification of the substrate site after tryptic digestion, or for SUMOlylation analysis, in which case a larger cross-linked peptide is generated (Andersen et al., 2009). Additionally, a wide variety of other modifications can be detected by MS and large-scale studies of many of them have already been undertaken (Witze et al., 2007).

Outlook

One trend in the future of cell biology is the application of unbiased approaches to questions of cellular behavior. In this regard, MS-based proteomics is particularly attractive because it focuses on proteins, their localization, modifications, and interactions. It is now becoming available to a larger community. The rapid developments of instrumentation and informatics tools for proteomics described here will facilitate this. However, large differences in the quality of data generated in proteomics projects still exist. Particularly, low resolution spectra can lead to misinterpretation of identifications, localization of modifications, and quantitation. Furthermore, in many experiments the trade-off between proteome coverage and the different conditions tested is still a limitation. While in-depth analysis, particularly of very complex samples, is still a difficult problem tackled mainly by expert laboratories, robust and reliable instrumentation increasingly make MS-based proteomics a biochemical technique of choice. So far, it still requires considerable expertise and dedicated personnel, but most cell biochemical approaches driven by proteomics outlined here are now accessible to all. Application of these methods in combination with biochemical techniques used

for characterization of organelles, large protein complexes, and posttranslational modifications will open a new window into the cell. This will tremendously improve our understanding of its behavior, architecture, and dynamics.

References

Andersen, J.S., C.J. Wilkinson, T. Mayor, P. Mortensen, E.A. Nigg, and M. Mann. 2003. Proteomic characterization of the human centrosome by protein correlation profiling. *Nature*. 426:570–574. doi:10.1038/nature02166

Andersen, J.S., Y.W. Lam, A.K. Leung, S.E. Ong, C.E. Lyon, A.I. Lamond, and M. Mann. 2005. Nucleolar proteome dynamics. *Nature*. 433:77–83. doi:10.1038/nature03207

Andersen, J.S., I. Matic, and A.C. Vertegaal. 2009. Identification of SUMO target proteins by quantitative proteomics. *Methods Mol. Biol*. 497:19–31. doi:10.1007/978-1-59745-566-4_2

Ashburner, M., C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, et al; The Gene Ontology Consortium. 2000. Gene ontology: tool for the unification of biology. *Nat. Genet*. 25:25–29. doi:10.1038/75556

Bantscheff, M., D. Eberhard, Y. Abraham, S. Bastuck, M. Boesche, S. Hobson, T. Mathieson, J. Perrin, M. Raida, C. Rau, et al. 2007. Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat. Biotechnol*. 25:1035–1044. doi:10.1038/nbt1328

Bell, A.W., E.W. Deutsch, C.E. Au, R.E. Kearney, R. Beavis, S. Sechi, T. Nilsson, and J.J. Bergeron; HUPO Test Sample Working Group. 2009. A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. *Nat. Methods*. 6:423–430. doi:10.1038/nmeth.1333

Bellew, M., M. Coram, M. Fitzgibbon, M. Igra, T. Randolph, P. Wang, D. May, J. Eng, R. Fang, C. Lin, et al. 2006. A suite of algorithms for the comprehensive analysis of complex protein mixtures using high-resolution LC-MS. *Bioinformatics*. 22:1902–1909. doi:10.1093/bioinformatics/btl276

Beynon, R.J., and J.M. Pratt. 2006. Strategies for measuring dynamics: the temporal component of proteomics. *Methods Biochem. Anal*. 49:15–25.

Blagoev, B., I. Kratchmarova, S.E. Ong, M. Nielsen, L.J. Foster, and M. Mann. 2003. A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol*. 21:315–318. doi:10.1038/nbt790

Boersema, P.J., R. Raijmakers, S. Lemeer, S. Mohammed, and A.J. Heck. 2009. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc*. 4:484–494. doi:10.1038/nprot.2009.21

Boisvert, F.M., Y.W. Lam, D. Lamont, and A.I. Lamond. 2010. A quantitative proteomics analysis of subcellular proteome localization and changes induced by DNA damage. *Mol. Cell. Proteomics*. 9:457–470. doi:10.1074/mcp.M900429-MCP200

Cheeseman, I.M., and A. Desai. 2005. A combined approach for the localization and tandem affinity purification of protein complexes from metazoans. *Sci. STKE*. 2005:pl1. doi:10.1126/stke.2662005pl1

Choudhary, C., C. Kumar, F. Gnad, M.L. Nielsen, M. Rehman, T.C. Walther, J.V. Olsen, and M. Mann. 2009. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science*. 325:834–840. doi:10.1126/science.1175371

Collins, S.R., P. Kemmeren, X.C. Zhao, J.F. Greenblatt, F. Spencer, F.C. Holstege, J.S. Weissman, and N.J. Krogan. 2007. Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics*. 6:439–450.

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

Craig, R., and R.C. Beavis. 2004. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*. 20:1466–1467. doi:10.1093/bioinformatics/bth092

Cravatt, B.F., G.M. Simon, and J.R. Yates III. 2007. The biological impact of mass-spectrometry-based proteomics. *Nature*. 450:991–1000. doi:10.1038/nature06525

Cristea, I.M., R. Williams, B.T. Chait, and M.P. Rout. 2005. Fluorescent proteins as proteomic probes. *Mol. Cell. Proteomics*. 4:1933–1941. doi:10.1074/mcp.M500227-MCP200

de Godoy, L.M., J.V. Olsen, J. Cox, M.L. Nielsen, N.C. Hubner, F. Fröhlich, T.C. Walther, and M. Mann. 2008. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*. 455:1251–1254. doi:10.1038/nature07341

Deutsch, E.W., L. Mendoza, D. Shteynberg, T. Farrah, H. Lam, N. Tasman, Z. Sun, E. Nilsson, B. Pratt, B. Prazen, et al. 2010. A guided tour of the Trans-Proteomic Pipeline. *Proteomics*. 10:1150–1159. doi:10.1002/pmic.200900375

Domon, B., and R. Aebersold. 2006. Mass spectrometry and protein analysis. *Science*. 312:212–217. doi:10.1126/science.1124619

Dunkley, T.P., R. Watson, J.L. Griffin, P. Dupree, and K.S. Lilley. 2004. Localization of organelle proteins by isotope tagging (LOPIT). *Mol. Cell. Proteomics*. 3:1128–1134. doi:10.1074/mcp.T400009-MCP200

Eng, J.K., A.L. McCormack, and J.R. Yates. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom*. 5:976–989. doi:10.1016/1044-0305(94)80016-2

Ens, W., and K.G. Standing. 2005. Hybrid quadrupole/time-of-flight mass spectrometers for analysis of biomolecules. *Methods Enzymol*. 402:49–78. doi:10.1016/S0076-6879(05)02002-1

Fenn, J.B., M. Mann, C.K. Meng, S.F. Wong, and C.M. Whitehouse. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science*. 246:64–71. doi:10.1126/science.2675315

Forner, F., C. Kumar, C.A. Luber, T. Fromme, M. Klingenspor, and M. Mann. 2009. Proteome differences between brown and white fat mitochondria reveal specialized metabolic functions. *Cell Metab*. 10:324–335. doi:10.1016/j.cmet.2009.08.014

Gavin, A.C., P. Aloy, P. Grandi, R. Krause, M. Boesche, M. Marzioch, C. Rau, L.J. Jensen, S. Bastuck, B. Dümpelfeld, et al. 2006. Proteome survey reveals modularity of the yeast cell machinery. *Nature*. 440:631–636. doi:10.1038/nature04532

Geiger, T., J. Cox, P. Ostasiewicz, J.R. Wisniewski, and M. Mann. 2010. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat. Methods*. 7:383–385. doi:10.1038/nmeth.1446

Ghaemmaghami, S., W.K. Huh, K. Bower, R.W. Howson, A. Belle, N. Dephoure, E.K. O'Shea, and J.S. Weissman. 2003. Global analysis of protein expression in yeast. *Nature*. 425:737–741. doi:10.1038/nature02046

Gingras, A.C., M. Gstaiger, B. Raught, and R. Aebersold. 2007. Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol*. 8:645–654. doi:10.1038/nrm2208

Hardman, M., and A.A. Makarov. 2003. Interfacing the orbitrap mass analyzer to an electrospray ion source. *Anal. Chem*. 75:1699–1705. doi:10.1021/ac0258047

Hsu, J.L., S.Y. Huang, N.H. Chow, and S.H. Chen. 2003. Stable-isotope dimethyl labeling for quantitative proteomics. *Anal. Chem*. 75:6843–6852. doi:10.1021/ac0348625

Hubner, N.C., A.W. Bird, J. Cox, B. Splettstoesser, P. Bandilla, I. Poser, A. Hyman, and M. Mann. 2010. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J. Cell Biol*. 189:739–754. doi:10.1083/jcb.200911091

Huh, W.K., J.V. Falvo, L.C. Gerke, A.S. Carroll, R.W. Howson, J.S. Weissman, and E.K. O'Shea. 2003. Global analysis of protein localization in budding yeast. *Nature*. 425:686–691. doi:10.1038/nature02026

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

Jensen, O.N. 2006. Interpreting the protein language using proteomics. *Nat. Rev. Mol. Cell Biol*. 7:391–403. doi:10.1038/nrm1939

Kim, S.C., R. Sprung, Y. Chen, Y. Xu, H. Ball, J. Pei, T. Cheng, Y. Kho, H. Xiao, L. Xiao, et al. 2006. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell*. 23:607–618. doi:10.1016/j.molcel.2006.06.026

Kitteringham, N.R., R.E. Jenkins, C.S. Lane, V.L. Elliott, and B.K. Park. 2009. Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci*. 877:1229–1239. doi:10.1016/j.jchromb.2008.11.013

Krogan, N.J., G. Cagney, H. Yu, G. Zhong, X. Guo, A. Ignatchenko, J. Li, S. Pu, N. Datta, A.P. Tikuisis, et al. 2006. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature*. 440:637–643. doi:10.1038/nature04670

Krüger, M., M. Moser, S. Ussar, I. Thievessen, C.A. Luber, F. Forner, S. Schmidt, S. Zanivan, R. Fässler, and M. Mann. 2008. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell.* 134:353–364. doi:10.1016/j.cell.2008.05.033

Kumar, C., and M. Mann. 2009. Bioinformatics analysis of mass spectrometry-based proteomics data sets. *FEBS Lett.* 583:1703–1712. doi:10.1016/j.febslet.2009.03.035

Lam, Y.W., V.C. Evans, K.J. Heesom, A.I. Lamond, and D.A. Matthews. 2010. Proteomics analysis of the nucleolus in adenovirus-infected cells. *Mol. Cell. Proteomics.* 9:117–130. doi:10.1074/mcp.M900338-MCP200

Link, A.J., J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, and J.R. Yates III. 1999. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* 17:676–682. doi:10.1038/10890

Liu, H., R.G. Sadygov, and J.R. Yates III. 2004. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* 76:4193–4201. doi:10.1021/ac0498563

Lu, P., C. Vogel, R. Wang, X. Yao, and E.M. Marcotte. 2007. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* 25:117–124. doi:10.1038/nbt1270

Luber, C.A., J. Cox, H. Lauterbach, B. Fancke, M. Selbach, J. Tschopp, S. Akira, M. Wiegand, H. Hochrein, M. O'Keeffe, and M. Mann. 2010. Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. *Immunity.* 32:279–289. doi:10.1016/j.immuni.2010.01.013

Makarov, A. 2000. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal. Chem.* 72:1156–1162. doi:10.1021/ac991131p

Malmström, J., M. Beck, A. Schmidt, V. Lange, E.W. Deutsch, and R. Aebersold. 2009. Proteome-wide cellular protein concentrations of the human pathogen Leptospira interrogans. *Nature.* 460:762–765. doi:10.1038/nature08184

McLafferty, F.W., K. Breuker, M. Jin, X. Han, G. Infusini, H. Jiang, X. Kong, and T.P. Begley. 2007. Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS J.* 274:6256–6268.

Neubauer, G., A. Gottschalk, P. Fabrizio, B. Séraphin, R. Lührmann, and M. Mann. 1997. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. *Proc. Natl. Acad. Sci. USA.* 94:385–390. doi:10.1073/pnas.94.2.385

Nita-Lazar, A., H. Saito-Benz, and F.M. White. 2008. Quantitative phosphoproteomics by mass spectrometry: past, present, and future. *Proteomics.* 8:4433–4443. doi:10.1002/pmic.200800231

Old, W.M., K. Meyer-Arendt, L. Aveline-Wolf, K.G. Pierce, A. Mendoza, J.R. Sevinsky, K.A. Resing, and N.G. Ahn. 2005. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol. Cell. Proteomics.* 4:1487–1502. doi:10.1074/mcp.M500084-MCP200

Olsen, J.V., B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, and M. Mann. 2006. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell.* 127:635–648. doi:10.1016/j.cell.2006.09.026

Ong, S.E., and M. Mann. 2005. Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* 1:252–262. doi:10.1038/nchembio736

Ong, S.E., B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, and M. Mann. 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics.* 1:376–386. doi:10.1074/mcp.M200025-MCP200

Paizs, B., and S. Suhai. 2005. Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* 24:508–548. doi:10.1002/mas.20024

Park, S.K., J.D. Venable, T. Xu, and J.R. Yates III. 2008. A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat. Methods.* 5:319–322.

Perkins, D.N., D.J. Pappin, D.M. Creasy, and J.S. Cottrell. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis.* 20:3551–3567. doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2

Picotti, P., B. Bodenmiller, L.N. Mueller, B. Domon, and R. Aebersold. 2009. Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. *Cell.* 138:795–806. doi:10.1016/j.cell.2009.05.051

Ranish, J.A., E.C. Yi, D.M. Leslie, S.O. Purvine, D.R. Goodlett, J. Eng, and R. Aebersold. 2003. The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* 33:349–355. doi:10.1038/ng1101

Rigaut, G., A. Shevchenko, B. Rutz, M. Wilm, M. Mann, and B. Séraphin. 1999. A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17:1030–1032. doi:10.1038/13732

Ross, P.L., Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, et al. 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics.* 3:1154–1169. doi:10.1074/mcp.M400129-MCP200

Rothbauer, U., K. Zolghadr, S. Muyldermans, A. Schepers, M.C. Cardoso, and H. Leonhardt. 2008. A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol. Cell. Proteomics.* 7:282–289.

Rout, M.P., J.D. Aitchison, A. Suprapto, K. Hjertaas, Y. Zhao, and B.T. Chait. 2000. The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J. Cell Biol.* 148:635–651. doi:10.1083/jcb.148.4.635

Sadygov, R.G., D. Cociorva, and J.R. Yates III. 2004. Large-scale database searching using tandem mass spectra: looking up the answer in the back of the book. *Nat. Methods.* 1:195–202. doi:10.1038/nmeth725

Schirmer, E.C., L. Florens, T. Guan, J.R. Yates III, and L. Gerace. 2003. Nuclear membrane proteins with potential disease links found by subtractive proteomics. *Science.* 301:1380–1382. doi:10.1126/science.1088176

Scigelova, M., and A. Makarov. 2006. Orbitrap mass analyzer—overview and applications in proteomics. *Proteomics.* 6(Suppl 2):16–21. doi:10.1002/pmic.200600528

Shevchenko, A., M. Wilm, O. Vorm, and M. Mann. 1996. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* 68:850–858. doi:10.1021/ac950914h

Shou, W., J.H. Seol, A. Shevchenko, C. Baskerville, D. Moazed, Z.W. Chen, J. Jang, A. Shevchenko, H. Charbonneau, and R.J. Deshaies. 1999. Exit from mitosis is triggered by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex. *Cell.* 97:233–244. doi:10.1016/S0092-8674(00)80733-3

Tan, D.J., H. Dvinge, A. Christoforou, P. Bertone, A. Martinez Arias, and K.S. Lilley. 2009. Mapping organelle proteins and protein complexes in *Drosophila melanogaster*. *J. Proteome Res.* 8:2667–2678. doi:10.1021/pr800866n

Trinkle-Mulcahy, L., S. Boulon, Y.W. Lam, R. Urcia, F.M. Boisvert, F. Vandermoere, N.A. Morrice, S. Swift, U. Rothbauer, H. Leonhardt, and A. Lamond. 2008. Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J. Cell Biol.* 183:223–239. doi:10.1083/jcb.200805092

Unwin, R.D., J.R. Griffiths, and A.D. Whetton. 2009. A sensitive mass spectrometric method for hypothesis-driven detection of peptide post-translational modifications: multiple reaction monitoring-initiated detection and sequencing (MIDAS). *Nat. Protoc.* 4:870–877. doi:10.1038/nprot.2009.57

Wang, H., B. Kakaradov, S.R. Collins, L. Karotki, D. Fiedler, M. Shales, K.M. Shokat, T.C. Walther, N.J. Krogan, and D. Koller. 2009. A complex-based reconstruction of the *Saccharomyces cerevisiae* interactome. *Mol. Cell. Proteomics.* 8:1361–1381. doi:10.1074/mcp.M800490-MCP200

Wang, Q., Y. Zhang, C. Yang, H. Xiong, Y. Lin, J. Yao, H. Li, L. Xie, W. Zhao, Y. Yao, et al. 2010. Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science.* 327:1004–1007. doi:10.1126/science.1179687

Washburn, M.P., D. Wolters, and J.R. Yates III. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 19:242–247. doi:10.1038/85686

Wepf, A., T. Glatter, A. Schmidt, R. Aebersold, and M. Gstaiger. 2009. Quantitative interaction proteomics using mass spectrometry. *Nat. Methods.* 6:203–205. doi:10.1038/nmeth.1302

White, F.M. 2008. Quantitative phosphoproteomic analysis of signaling network dynamics. *Curr. Opin. Biotechnol.* 19:404–409. doi:10.1016/j.copbio.2008.06.006

Wilm, M. 2009. Quantitative proteomics in biological research. *Proteomics.* 9:4590–4605. doi:10.1002/pmic.200900299

Wiśniewski, J.R., A. Zougman, N. Nagaraj, and M. Mann. 2009. Universal sample preparation method for proteome analysis. *Nat. Methods.* 6:359–362. doi:10.1038/nmeth.1322

Witze, E.S., W.M. Old, K.A. Resing, and N.G. Ahn. 2007. Mapping protein post-translational modifications with mass spectrometry. *Nat. Methods.* 4:798–806. doi:10.1038/nmeth1100

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

Yates, J.R. III, A. Gilchrist, K.E. Howell, and J.J. Bergeron. 2005. Proteomics of organelles and large cellular structures. *Nat. Rev. Mol. Cell Biol.* 6:702–714. doi:10.1038/nrm1711

Zhao, S., W. Xu, W. Jiang, W. Yu, Y. Lin, T. Zhang, J. Yao, L. Zhou, Y. Zeng, H. Li, et al. 2010. Regulation of cellular metabolism by protein lysine acetylation. *Science.* 327:1000–1004. doi:10.1126/science.1179689