REVIEW

nature chemical biology

Mass spectrometry–based proteomics turns quantitative

Shao-En Ong[1] & Matthias Mann[2]

The field of proteomics is built on technologies to analyze large numbers of proteins—ideally the entire proteome—in the same experiment. Mass spectrometry (MS) has been successfully used to characterize proteins in complex mixtures, but results so far have largely been qualitative. Two recently developed methodologies offer the opportunity to obtain quantitative proteomic information. Comparing the signals from the same peptide under different conditions yields a rough estimate of relative protein abundance between two proteomes. Alternatively, and more accurately, peptides are labeled with stable isotopes, introducing a predictable mass difference between peptides from two experimental conditions. Stable isotope labels can be incorporated 'post-harvest', by chemical approaches or in live cells through metabolic incorporation. This isotopic handle facilitates direct quantification from the mass spectra. Using these quantitative approaches, precise functional information as well as temporal changes in the proteome can be captured by MS.

Proteomics is a relatively new 'post-genomic' science with tremendous potential. In contrast to gene expression studies employing oligonucleotide chips ('transcriptomics'), proteomics directly addresses the level of gene products present in a given cell state and can further characterize protein activities, interactions and subcellular distributions. Proteomics has been successfully applied to areas as diverse as determining the protein composition of organelles, systematic elucidation of protein-protein interactions and the large-scale mapping of protein phosphorylation in response to a stimulus.

The term 'proteomics' was coined in the context of two-dimensional gel electrophoresis (2-DE)[1]. In 2-DE experiments, the staining pattern of proteins from two samples are compared and 'up-' and 'downregulated' proteins identified. Protein staining also provides a rough idea of the relative amounts of each protein within the sample. However, 2-DE has limitations, such as low resolution and bias against membrane proteins. Furthermore, the dynamic range of protein expression can vary by as much as 7 to 12 orders of magnitude within a biological sample like serum[2,3], and 2-DE can only visualize the most abundant of these proteins. As a result, 2-DE gels have largely been superseded by MS-based proteomics. (For an overview of different aspects of MS-based proteomics, see refs. 4–11, and for more specialized reviews on quantitative proteomics, see refs. 12,13.)

The dominant sample analysis workflow in proteomics utilizes site-specific enzymatic proteases such as trypsin to digest proteins to peptides[14]. Peptides are fragmented in the mass spectrometer and the resulting 'MS/MS' spectra are used to retrieve the corresponding peptide sequence from a database. Protein identification is straightforward, as only two unique peptides are usually sufficient to recognize a protein. Successful detection and identification of protein-specific peptides confirm their presence within the sample. However, the failure to identify or detect a peptide does not necessarily mean that the protein is absent, as the peptides may simply be below the threshold of detection. Therefore, the Boolean nature of MS protein identification schemes provides a very limited picture of protein abundance in a sample. Thus, although sensitive MS-based proteomic approaches readily identify a large number of proteins, bypassing the gel-visualization step deprives us of any measure of protein abundance in the sample. Moreover, most changes resulting from a targeted perturbation of a biological system are only detectable if some quantitative information is obtained. Furthermore, biology in general and systems biology in particular increasingly require quantitative data as an input for modeling.

Quantitative data comes in two forms, the absolute amount of the protein in the sample or the relative change in protein amount between two states. Absolute quantification is the determination of the amount of the substance in question, for example, ng ml^{-1} of a biomarker or the copy number of a protein per cell. In relative quantification, the amount of a substance is defined in relation to another measure of the same substance, such as a fold change of protein abundance resulting from drug treatment. In principle, absolute quantification encompasses relative comparisons; if the absolute amounts of the proteins are known in two samples, their relative ratios can be calculated easily.

Proteomics researchers commonly try to extract at least some quantitative information from lists of protein identifications in an effort to rank the abundances of proteins in relation to each other within a single sample. For example, a protein's score is a sum of identification scores of its peptides, and one might surmise that a high protein-identification score would correlate with a higher abundance of that protein[15]. Perhaps surprisingly then, identification scores were found to provide a poor estimate of protein abundance even when tested

[1]The Broad Institute of MIT and Harvard, 320 Bent Street, Cambridge, Massachusetts 02141, USA. [2]Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany. Correspondence should be addressed to S.-E.O. (song@broad.mit.edu) or M.M. (mmann@biochem.mpg.de).

Published online 20 September 2005; doi:10.1038/nchembio736

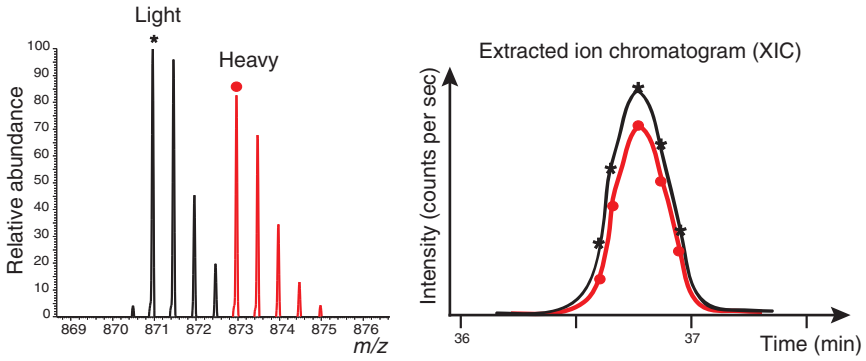


Figure 1 Extracting quantitative data from mass spectra. Mass spectra are acquired, resulting in isotope clusters for each peptide (doubly charged peptide, black in top graph). As the peptide elutes from the column, the signal is sampled several times, forming the black curve in the right panel. The area under the black curve is the XIC, a measure that is proportional to the peptide's abundance. The red isotope cluster is a heavy isotope–labeled analog of the black peptide, 4 Da higher in mass and present at 85% of the unlabeled peptide. The 85% ratio can also be determined by comparing the areas under the red and black curves. Several thousand such peptide XICs can be extracted from a data file obtained in typical complex mixture analysis.

with simple mixtures of purified proteins[16]. This is because database search algorithms score peptides based on the numbers of matching sequence fragments rather than by their absolute intensities. Indeed, it is not unusual for a weak peptide signal to give a clear sequence ladder leading to a high identification score[10,17]. We therefore do not recommend using database identification scores as a quantitative measure of protein abundance. Instead, more accurate 'protein abundance indices' (PAIs) based on various observable parameters have been developed. For instance, the number of peptides identifying a protein increases with increasing protein amount, and a larger protein will generate more measurable peptides than a smaller one. This is the basis of a simple PAI in which the number of observed peptides is normalized to the number of observable peptides for the protein under consideration[18,19]. Ishihama *et al.* have noticed that the relationship between the number of peptides observed and the protein amount within a given sample is logarithmic, leading to the concept of an exponentially modified PAI (emPAI)[20].

In an MS experiment, the intensity of the signal as the peptide elutes from the chromatographic column can be plotted over time (**Fig. 1**). The area under this curve is the extracted ion current (XIC) and, for the same peptide and experimental conditions, is linearly related to its amount. It is not possible to predict the MS detector response to any particular peptide because of unknown extraction and peptide ionization properties and therefore XICs of different peptides of the same protein are also very different. However, the average of the three most intense peptide XICs of a protein is a quantitative measure, and this 'xPAI'[16] has been used to distinguish between matrix and regulatory components of the centrosome[21]. These abundances can be extracted from the raw data with simple scripts and without additional data acquisition. PAIs provide rough estimates of protein abundance that are, in our experience, typically within a factor of 3 to 5 of the true value. Even this very modest accuracy can be very helpful in interpreting proteomic experiments, and indeed is comparable to that of staining approaches, as there can be a several-fold difference in the staining characteristics of different proteins.

There are major caveats of any quantitative method based on counting peptides. Not all peptides are analyzable by MS; some are too small or large and fall beyond the mass range analyzed, while others may not be favorably retained on the chromatographic column during analysis. Furthermore, a direct measurement of abundance suffers from propagation of quantitative errors resulting from variability in the processing steps required for MS analyses. Unknown losses occur at each step—beginning with preparation of the protein sample in a microcentrifuge tube, during sample introduction and the ionization process, in transit from the source region to the detector, and ending with saturation of the detector. Quantitative measures that do not take such factors into account are affected by these nonsystematic errors and are therefore less accurate.

Fortunately, recent developments now overcome these issues and provide robust and practical tools for quantitative proteomics. We will first discuss methods based on comparing the ion currents of the same peptides in different experiments (ion current-based quantification). We then turn to by far the most promising approaches, namely those involving stable isotope labeling, which permits direct comparison of two proteome states in the same analysis. We will also highlight exciting perspectives for functional proteomics based on these approaches.

XIC-based quantification

In order to improve on the simple presence or absence information provided by a purely qualitative experiment, we can integrate and compare the intensities of the same peptides between two states. Although comparing intensities between different peptides is not possible because of different extraction yields during processing and different ionization efficiencies, these sources of error do not apply when comparing the same peptide in different chromatographic runs. The two proteomes to be compared are processed and analyzed one after another and in exactly the same way. This can be achieved with the aid of autosamplers and standard operating procedures. Intensities of the same peptide observed in two separate runs are compared to determine their relative abundance. In complex mixture analysis, not all peptides are selected for fragmentation in every run[8]. Therefore, a critical requirement is the ability to find and quantify the peptide in different runs, even if it has only been sequenced in one. Modern high mass accuracy and high-resolution mass spectrometers, as well as the development of the required software, now make this task much more feasible. Under such conditions, and if signal-to-noise ratios are good, greater than two-fold changes can easily be measured. For peptide signals close to the background, the quantitative accuracy will mainly be determined by the variation in that background.

An obvious advantage of XIC-based quantification is that no labeling strategy is used and that it can be performed with any type of sample. Frequently, a rough and ready quantification as provided by this method will be sufficient to guide subsequent follow-up biological studies or more accurate quantitative experiments. Clear disadvantages are the multiple occasions for quantification error to occur during sample processing and analysis, as well as the presence of interfering substances (detergents, abundant background proteins) in one of the states to be compared. Some of these limitations can be minimized by normalizing between runs with spiked-in calibrants or by using abundant nonchanging peptides as landmarks between runs (for example, see refs. 22,23). Because of the requirement for extremely reproducible

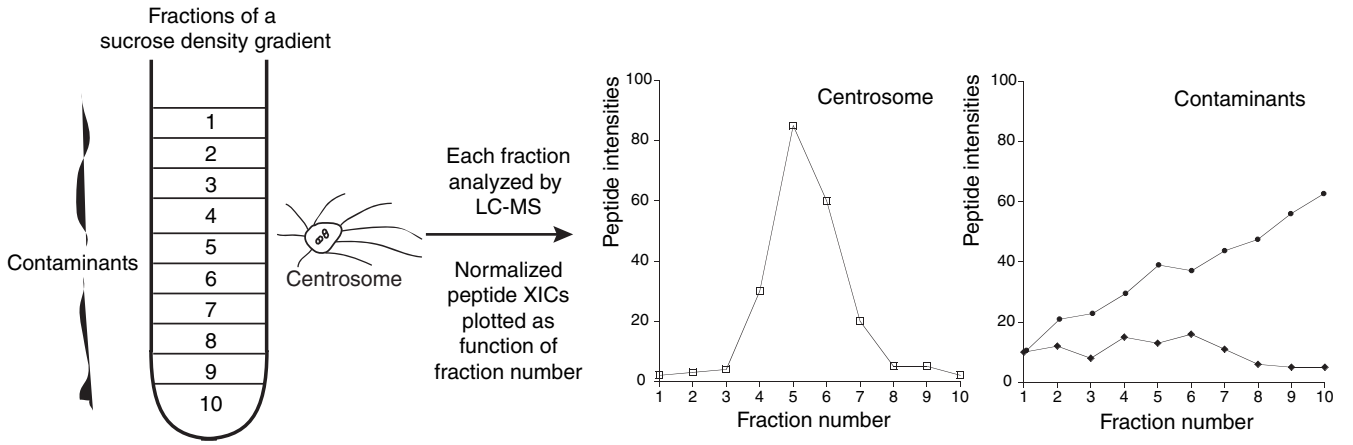


Figure 2 PCP to determine the centrosomal proteome. Centrosomes are enriched by density centrifugation. The XICs of peptides identified in any one of the fractions are normalized and averaged for each protein. New centrosomal proteins coelute with known centrosomal proteins, whereas background proteins have a different distribution across the centrifugation fractions. The degree of cofractionation can objectively be expressed as a χ^2 value—the square deviation from the centrosomal consensus profile[21].

sample preparation, multistage sample purification protocols are not amenable to this quantification method.

Ion current–based quantification can be extended to provide a protein abundance profile to co-localize proteins in a fractionation procedure analogous to western blotting for marker proteins and proteins of interest. For instance, Andersen *et al.* analyzed separate fractions of a sucrose density gradient fractionation of a centrosomal preparation[21]. The abundance of thousands of peptides was followed through several sucrose density fractions. Peptides derived from centrosomal proteins had maximum abundances in the same fraction and also showed the same protein abundance profile as a result of copurifying in the same sucrose density fractions. Background proteins, which vastly outnumbered the centrosomal proteins in this preparation, showed distinct protein profiles and were easily distinguished. This method, termed protein correlation profiling (PCP)[21], can be applied to determine the constituents of any protein assembly that can be enriched, but not purified to homogeneity (**Fig. 2**).

The important problem of phosphorylation stoichiometry has also been addressed by XIC-based quantification. Signal intensities of nonphosphorylated and phosphorylated peptide forms were measured. Then these levels were remeasured after changing the stoichiometry of phosphorylation (by dephosphorylation or in another experimental condition). As the amount of phosphorylated and nonphosphorylated peptide must add up to the same value in both experiments, the relative ionization efficiency and hence the stochiometry of the phosphopeptide can be determined[24–26].

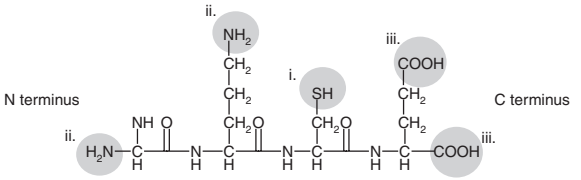
Stable isotope labeling in MS-based quantification

For several decades, metabolic studies have used MS combined with stable isotope labeling for quantification[27]. Analogs of the drug to be tested are synthesized with nonradioactive isotope labels such as ^{13}C, ^{15}N and ^{2}H, and a known quantity of the analog is introduced into the sample. Since their ionization efficiencies are the same, both forms produce the same MS response signal in the mass spectrometer, only offset by the mass difference introduced by the stable isotopes (**Fig. 1**). This highly accurate method of quantification has recently also become the gold standard of MS-based proteomic quantification. Importantly, stable isotope labeling allows the use of matrix-assisted laser desorption-ionization (MALDI) instrumentation in quantitative studies. MALDI is not normally an ionization technique suited for quantification because of the large dependence of signal intensity on analyte-matrix co-crystallization conditions. However, since the isotopic pair behaves identically, quantification is unaffected. On the other hand, surface-enhanced laser desorption-ionization (SELDI, for a review, see ref. 28), a derivative of MALDI, is not amenable to quantitative studies because of the low mass resolution of the instruments employed and the fact that whole proteins rather than peptides would have to be chemically modified (see below).

Several methods have been developed for stable isotope-based quantification in proteomics which are mainly distinguished by the way stable isotope labels are introduced into the peptide or protein. They fall into the categories of (i) spiking in an isotopically labeled analog, (ii) incorporation through an enzyme during protein digestion, (iii) introducing a chemical, isotopically labeled tag onto peptides or proteins and (iv) having cells incorporate the label metabolically. These four methods will be considered in turn below. Regardless of the method chosen for labeling, the mass difference imparted by the stable isotope atoms should at least result in a 3- or 4-Da mass shift of the 'heavy' from the 'light' peptide, to minimize quantitative errors from isotopic overlap (**Fig. 1**). Deuterated peptides can separate in reversed-phase chromatography[29], and therefore the more expensive ^{13}C- and ^{15}N-based reagents may be preferable (these exhibit significantly less separation, most often coeluting). The quantitative ratios can be determined from a relative comparison of signal intensities within a single MS scan (in the case of coeluting peaks) or with peak areas determined from XICs of the heavy and the light peptide. In either case, the exact same experimental conditions prevail for both forms, leading to accurate determination of the ratios.

The simplest approach for the introduction of stable isotope-labeled peptides is to chemically synthesize them and 'spike' known quantities into the sample as internal standards. This was demonstrated by Desiderio and co-workers in the early 1990s for the quantification of neuropeptides[30], in extension of the well-established stable isotope–dilution approaches used in the pharmaceutical industry for the quantification of small molecules. Gygi and co-workers use this absolute quantitation (AQUA) strategy to quantify the phosphorylation status of separase during the *Xenopus* cell cycle[31]. This approach could be applied to the quantification of candidate biomarkers in body fluids

Table 1 Chemical-labeling strategies for quantitative proteomics



i. Sulfhydryl group–directed (cysteine)

ii. Amine-directed (amino terminus and ε-amino group of lysine)

iii. Carboxyl-directed (carboxyl terminus, aspartic and glutamic acid)

Target	Name of method or reagent	Isotopes	Refs.	Notes
Sulfhydryl	Isotope-coded affinity tag (ICAT)	D	45	Iodoacetamide-based, enrichable, not cleavable
	Cleavable ICAT	^{13}C	46–48	Acid cleavable
	Acrylamide	D	101	
	Isotope-coded reduction off of a chromatographic support (ICROC)	D	102	*N*-ethyl-iodoacetamide
	2-vinyl-pyridine	D	103	
	N-*t*-butyliodoacetamide	D	104	
	Iodoacetanilide	D	104	
	HysTag	D	49,50	2-thiopyridyl disulfide, a peptide reagent, 6xHis, enzyme cleavable
	Solid-phase ICAT	D	105	UV photocleavable
	Acid-labile isotope-coded extractants (ALICE)	D	106	Maleimide-based, enrichable, acid cleavable
	Solid phase mass tagging	^{13}C	107,108	Iodoacetamide-based, enrichable, acid cleavable
Amines	Tandem mass tags (TMT)	D	109	Isobaric tag, quantification in MS/MS
	Succinic anhydride	D	110	
	N-acetoxysuccinamide	D	56	
	Acetic anhydride	D	111	
	Propionic anhydride	D	58	
	Nicotinoyloxy succinimide (Nic-NHS)	D	55	Labeling of peptide
	Isotope-coded protein label (ICPL, Nic-NHS)	D	57	Applied to labeling of proteins
	Phenyl isocyanate	D	112	
	Isotope-coded *n*-terminal sulfonation (ICenS) 4-sulfophenyl isothiocyanate	^{13}C	113	
	Sulfo-NHS-SS-biotin and $^{13}C,D_3$-methyl iodide	^{13}C, D	114	
	Formaldehyde	D	115	
	Isobaric tag for relative and absolute quantitation (iTRAQ)	^{13}C, 15N & ^{18}O	59	Isobaric tag, quantitation in MS/MS, four-plex
Lysines	Guanidination (*O*-methyl-isourea) mass-coded abundance tagging (MCAT)	No isotope	116	Derivatization of one state, different ionization between states, LC separation
	Guanidination (*O*-methyl-isourea)	^{13}C, ^{15}N	117,118	Isotopic analogs of *O*-methyl-isourea
	Quantitation using enhanced sequence tags (QUEST)	No isotope	119	Difference of a methylene group, LC separation
	2-Methoxy-4,5-1*H*-imidazole	D	120	More specific than guanidination
Carboxyl	Methyl esterification	D	51	
	Ethyl esterification	D	121	
Others				
Tryptophan	2-nitrobenzenesulfenyl chloride (NBSCl)	^{13}C	122	Targets the indole ring of tryptophan
pSer/pThr	Phosphoprotein isotope-coded affinity tag (PhIAT)	D	123	β-elimination, Michael addition, ethanedithiol-based, biotinylated
	Phosphoprotein isotope-coded solid-phase tag (PhIST)	^{13}C, ^{15}N	124	β-elimination, Michael addition, ethanedithiol-based, UV photocleavable
O-GlcNac	Beta elimination and Michael addition with dithiothreitol (BEMAD)	D	125–127	Cross-reacts with phosphoserine and phosphothreonine as well

*Although we have striven to be comprehensive, we may have overlooked some strategies for quantitative proteomics. Note also that many of the approaches mentioned in the table differ only slightly from each other, despite their different acronyms.

(for example, see ref. 32). To reduce interference from background ions, quantification can be performed on specific fragments of the peptide generated in the mass spectrometer using selected- or multiple-reaction monitoring (SRM or MRM, respectively; in these methods, the MS is set to detect a preprogrammed precursor-fragment combination with very high sensitivity and specificity)[33,34]. The internal peptide standard is introduced at a late stage in the sample processing workflow during or after protein digestion. Any prior quantitative variations are not corrected by the internal standard[35]. Because suitable internal standards need to be identified and synthesized, this approach is usually limited to a small number of preselected proteins. However, Aebersold has suggested the generation of reagents for every protein which would then allow quantification of whole proteomes[36]. The daunting task of synthesizing tens of thousands of peptides might be alleviated by recent proposals to generate them by metabolic incorporation[37].

The stable isotopes necessary for quantification can also be introduced into the peptides by the protease degrading the protein to a peptide mixture[38–40]. The water molecule introduced during trypsin digestion will contain ^{18}O if the digestion is performed in $H_2{}^{18}O$ water. However, one or both carboxyl oxygens may be exchanged, leading to variability in the quantification. If only one ^{18}O atom is incorporated (such as by the use of endoproteinase Lys-N[41]), the mass offset of 2 Da is not sufficient to separate the isotopic envelopes, complicating quantification. As a result, the ^{18}O method has not been widely applied in quantitative proteomics. Postdigestion incubation of peptides with protease in a small volume of $H_2{}^{18}O$ water[42,43] or deactivating the protease through a reduction and alkylation step[44] may lead to more widespread use.

Chemical tagging approaches in MS-based quantification

Another approach to introduce stable isotope labels is to chemically modify the two proteomes under study—one with a light and the other with a heavy chemical reagent. In chemical modification–based approaches, stable isotope–bearing chemical reagents are targeted toward reactive sites on a protein or peptide (**Table 1**). The prototypical example is the isotope-coded affinity tag (ICAT) described in 1999 by Aebersold and co-workers[45]. The ICAT reagent consists of a reactive group that is cysteine-directed, a polyether linker region with eight deuteriums, and a biotin group that allows recovery of labeled peptides. In the classical ICAT experiment, isolated proteins from the experiment and control are denatured, reduced and modified with the heavy or light ICAT reagent, respectively. The combined proteomes are digested and labeled peptides are isolated with an avidin column. ICAT-labeled peptides are eluted from the column and quantified with MS. Because only cysteine-containing peptides are isolated, peptide mixtures are

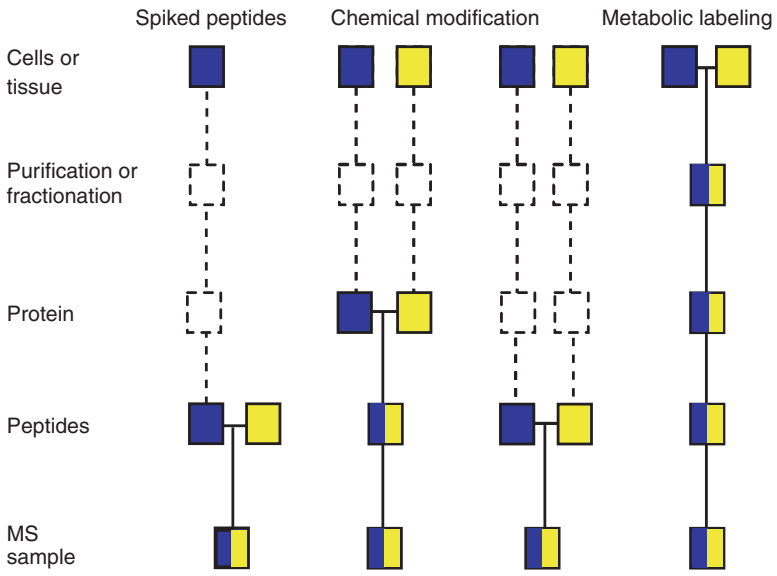


Figure 3 Stages of incorporation of stable isotope labels and their impact on quantitative accuracy. The scheme outlines typical labeling workflows in quantitative proteomics from the cell or tissue stage through purification and protein digestion to MS analysis. The blue and the yellow boxes represent the two cell states to be differentially labeled and compared. The horizontal line denotes the stage where samples are combined. When samples have to be processed in parallel, uncompensated quantification errors can occur (dashed lines).

less complex, which can be an advantage for quantification of complex mixtures. On the other hand, some proteins contain no cysteines and others have to be quantified on the basis of a single peptide. The large ICAT tag significantly influences fragmentation spectra, complicating peptide identification, and the deuterium tag results in separation of light and heavy peptides in reversed-phase chromatography. Subsequent iterations of the ICAT approach made the method more practical by substituting a cleavable and coeluting tag[46–48]. A related method, HysTag, uses a 6× histidine tag to allow enrichment, a 2-thiopyridyl disulfide group to react with thiols, a deuterium-labeled alanine and a tryptic cleavage site to limit the size of the tag[49]. This peptide reagent has been used to quantify cell-surface proteins in mouse brain[50].

With multiple reactive functional groups available in a given polypeptide chain and a choice of different labels, the actual design of the quantitative tag can be varied to suit almost any whim or fancy. This had led to a tremendous number of chemical-labeling methods described in the literature (**Table 1**). Notably, relatively few of these chemical modification methods have been applied in real experimental biology. This is largely due to the nonspecificity and incomplete labeling resulting from certain modification chemistries, which complicate MS identification as well as quantification. Even small amounts of side reactions lead to a tremendous increase in the complexity of the peptide mixture to be analyzed: a 0.1% side reaction of a major component could still be much more abundant than the proteins of interest. Chemical reactions that require several purification steps are generally incompatible with high-sensitivity proteomics experiments. Chemistries instead need to be extremely specific, proceed to completion and involve minimal sample handling.

Apart from cysteine's sulfhydryl group, the primary amines or the carboxylic groups have been successfully labeled (**Table 1**). Amino and carboxyl group–directed approaches can, in principle, quantify every observed peptide. In particular, relative changes in the level of post-translational modifications can also be quantified as modified peptides are also quantifiable.

Carboxylic groups (side chains of aspartic and glutamic acids as well as carboxyl termini) can be modified in a simple permethylation esterification reaction. For quantification, the reaction is carried out with either d0 or d3 methanol[51]. Ficarro *et al.* blocked acidic groups with the same reaction to prevent unspecific binding while purifying phosphopeptides on immobilized metal affinity columns[52]. In this scheme, quantification of phosphopeptides by d3-esterification is accomplished without any additional steps[53]. It is important to drive esterification to completion, especially in peptides with several acidic residues[12,54].

Reactions with primary amine groups (lysine and amino termini) can be specific and largely complete. The sample can be derivatized either at the protein or peptide level with amine-directed reagents. Peptides will generally have one label for the amino terminus as well as one for each internal lysine (trypsin will not cleave at derivatized lysines). James and co-workers use succinylation to protect the amine group on lysines and label amino termini of peptides with d0 or d4-nicotinoyloxy succinimide (d0/d4-Nic-NHS) esters[55]. Several other chemical-modification strategies targeting the amino groups have been developed[56–58]. A method that has recently gained popularity employs the isobaric Tag for Relative and Absolute Quantitation (iTRAQ). iTRAQ uses the same NHS chemistry mentioned previously, but adds an innovative concept, namely a tag that generates a specific reporter ion in fragmentation spectra[59]. There are four tags that produce fragment ions of mass 114, 115, 116 or 117 Da but, through a carbonyl balance group, add up to the same mass. Because labeled peptides from different states are isobaric, the mass spectra are relatively simple, and differential behavior is only revealed in fragmentation spectra. Moreover, this multiplexing strategy allows analysis of four separately labeled pools of protein in a single analysis, increasing analytical throughput. The method has been combined with offline liquid chromatography onto sample plates, which are interrogated by matrix assisted laser desorption-ionization time-of-flight instruments capable of peptide fragmentation (MALDI-TOF-TOF). An advantage of offline sampling by MALDI is the ability to reexamine fractions of special interest and acquire additional quantitative information. With iTRAQ, good peptide separation is important, as coeluting peptides of similar mass would contribute to the same reporter ions, complicating quantification.

In some approaches, the mass offset necessary to distinguish peptides from the two states is achieved by derivatization with two chemically different tags rather than an isotopic variant of the same tag (**Table 1**). For example, lysines may be derivatized with an amidine label that only differs by a methylene group. Although economical, these strategies can severely compromise accuracy of quantification because the two different tags generally have different reaction rates toward substrates, have distinct retention characteristics in chromatography and confer different ionization and extraction efficiencies upon derivatized peptides. If such strategies are used, the experiment should, at a minimum, be repeated using reverse labeling.

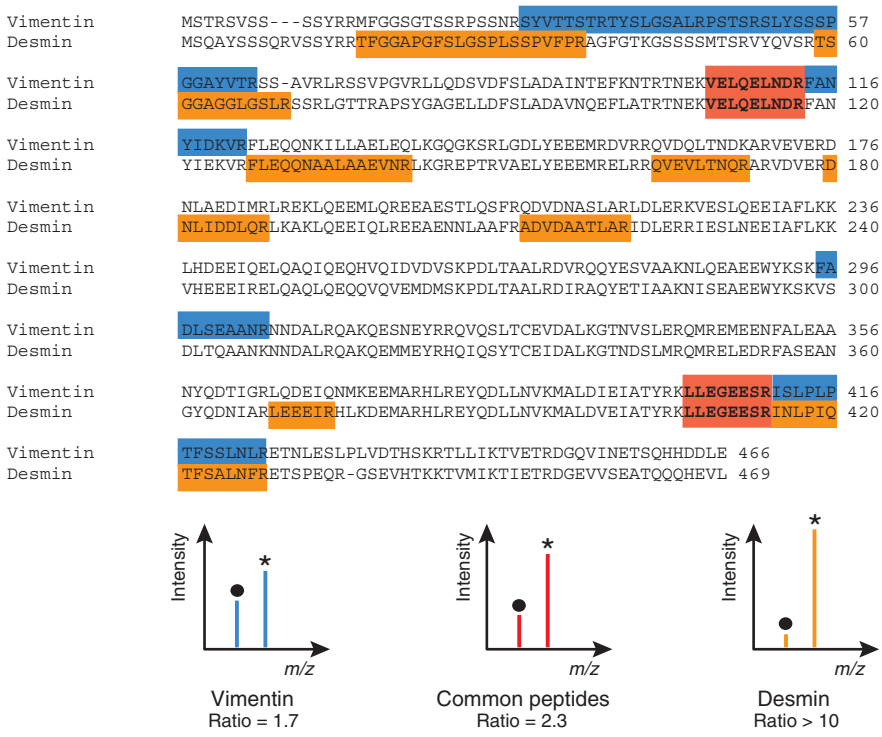


Figure 4 Quantification of proteins with extensive sequence identity. Mouse myoblast C2C12 cells were differentiated into myotubes and measured in a SILAC experiment. Peptides unique to vimentin (blue) show a ratio of 1.7, peptides unique to desmin (orange) show a ratio larger than 10 and peptides common (red) to the two proteins show an intermediate ratio.

Metabolic incorporation of stable isotopes

The other major class of labeling depends on metabolic incorporation of the stable isotope. This approach has been used for decades for quantification and trace monitoring of metabolites with radioactive isotopes. The critical distinction here is that unlike a pulse-chase experiment with a small amount of a radioactive tracer, metabolic labeling for quantitative proteomics replaces whole, unlabeled protein populations with stable isotope–labeled versions before the start of the experiment. To achieve this, metabolic precursors bearing stable isotopes are introduced in growth media of living cells. Through cell growth and protein turnover, these stable isotope–labeled precursors are incorporated into all cellular proteins. The number of stable isotope atoms introduced is constrained by the chemical structure of these metabolic precursors. Cells are grown in light and heavy media, giving rise to two cell populations, distinguishable only by MS.

A principal advantage of metabolic incorporation over chemical labeling is that the label is present in live cells. Cells from state A and state B can be mixed before lysis, fractionation and purification, meaning that accuracy of quantification will not be affected by these steps (**Fig. 3**). The absence of side reactions, due to the very high specificity of enzymatic reactions, is another advantage of metabolic labeling.

In 1998, Langen made use of ^{15}N- and ^{13}C-labeled precursors for the quantitative comparison of proteins with 2-DE gels[60]. This was closely followed in 1999 by the quantification of yeast phosphopeptides by Oda *et al.* with ^{15}N-labeling[61] and the quantification of a mammalian cell line by Smith and co-workers[62]. ^{15}N-labeling of simple organisms like bacteria and yeast through the use of ^{15}N-ammonium salt results in the complete labeling of amino acids within these cells—every peptide observable in MS is therefore quantifiable. These ^{15}N-labeled microorganisms can in turn be fed to small organisms such as *Caenorhabditis elegans*, *Drosophila melanogaster*[63] and even a rat has been labeled by feeding it ^{15}N-labeled algae[64]. Wu and co-workers observed that labeled protein pools in their labeled rat did not reach complete incorporation even after an extended feeding regime[64]. Though a tantalizing prospect, the cost of labeling small mammals with stable isotopes is very high. Therefore, to avoid labeling multiple animals, sample from a single labeled animal was mixed in with separate experimental animals, in effect serving as an internal standard.

There are however, some caveats in using ^{15}N labeling. The mass difference introduced between the unlabeled and the labeled form of the peptide depends on the amino acid sequence of the peptide in question. Although this additional information determines the number of nitrogen atoms and thus adds specificity in peptide identification, presently available peptide identification programs do not handle ^{15}N-labeled samples appropriately. It is also important to use extremely enriched nitrogen to avoid complicated isotopic distributions resulting from partially labeled peptides because of the substantial number of nitrogens in a typical peptide.

Stable isotope labeling by amino acids in cell culture (SILAC)[65] has proven to be a simple yet very powerful approach to quantitative proteomics. Amino acids containing stable isotopes, such as arginine bearing six ^{13}C atoms, are supplied in growth media, thereby introducing this heavy amino acid to newly synthesized polypeptides in a sequence-specific fashion. Metabolic labeling with radioactive amino acids for a short period of time (pulse-chase labeling) has long been employed in biological research. SILAC differs from this approach in that it does not involve radioactive material, labels two proteomes to completion and employs MS as the readout. In proteomics, labeled amino acids have previously been used to increase specificity in identification schema[66,67] but have only recently been applied to quantification[65,68,69]. The SILAC procedure simply consists of growing two cell populations in media, containing either a light or heavy form of an essential amino acid. After several cell doublings, the complete cellular proteome will incorporate the supplied amino acid in each protein of the proteome (even proteins with no turnover will be diluted to a few percent after five cell doublings). Every peptide pair is separated by the mass difference introduced by the labeled amino acid. Accuracy of quantification is only limited by the peptide signal observed (apart from biological variation) and, for abundant proteins, can be as low as a few percent[70]. Several amino acids have been used in the SILAC approach. Leucine labels approximately 70% of tryptic peptides[65,71] and is available in an economical deuterated form, although the deuterium-labeled pairs may separate considerably (as much as a minute) in reversed phase chromatography. SILAC labeling with arginine and lysine while digesting with trypsin results in labeling of every peptide but the carboxyl-terminal peptide of the protein[72], as does labeling with lysine when

digested with Lys-C[73]. Amino acids of special interest such as tyrosine[74] or methionine[75] can also be labeled. *S*-adenosylmethionine is the primary methylation donor in biological systems; therefore, labeling the methionine results in labeling sites of protein methylation. This can be used to directly quantify relative levels of protein methylation[75]. Note that arginine is an essential amino acid in cell culture[76] and that some nonessential amino acids can be incorporated by offering a large excess to the cell or microorganism. If a microorganism is amenable to genetic manipulation, however, a general method for SILAC labeling is the generation of auxotrophic strains, such as Arg^{-}/Lys^{-} double auxotrophic yeast[77].

SILAC labeling for quantitative proteomics is a simple process. The only necessary step is the preparation of the growth media. Any medium formulation can be adapted for use in SILAC as long as the supplemented labeled amino acids are the only available source to growing cells. For that reason, dialyzed serum is used in SILAC. Some cell lines do not grow well in dialyzed serum, most likely because of the loss of essential growth factors during the dialysis process. Combining dialyzed serum with a small amount of normal, undialyzed serum can alleviate this problem[78] and only introduces small quantification errors. We have observed that arginine can metabolically interconvert to proline in some cell types[70]. This interconversion can be minimized by titrating the amount of arginine used in media. Alternatively, peptides containing proline can be excluded from quantification. We generally recommend titration of the SILAC amino acids for economic reasons (even though it is generally not the most expensive component in SILAC media). We maintain a website with practical information about SILAC (www.biochem.mpg.de/mann/), and SILAC protocols have recently been published[79,80].

Although chemical or enzymatic approaches are usually used to quantify between the tissues of two animals, it has recently been shown that it may also be possible to use SILAC for that purpose. Oda and co-workers' have developed the culture-derived isotope tags (CDIT) approach, which uses SILAC labeled cells as the bridging internal standard between two tissue samples[81]. Impressively, they quantify more than 90% of approximately 1,000 proteins identified from mouse brain with SILAC labeled Neuro2A cell line as internal standards. The advantages of this approach over a chemical-modification method include the lack of any special sample processing and relatively low cost compared with chemically labeling milligram quantities of protein, which requires large amounts of chemical reagent.

Why and when to use quantitative proteomics

Profiling the transcriptome with DNA chips is a widespread and powerful method of functional genomics. Current technology, though still imperfect, permits monitoring of whole transcriptomes and processing of large numbers of samples in standard experimental formats. In contrast, quantitative proteomics is still far from characterizing whole proteomes comprehensively and is comparatively low throughput. However, biological processes are mainly controlled by proteins, and mRNA is merely a proxy for determining protein abundance. Many factors apart from mRNA abundance determine the levels and activities of proteins, including regulated destruction of proteins and post-translational modification. Therefore, it is desirable and often necessary to study proteins directly or to compare protein and mRNA levels. Quantitative proteomics has the potential to provide a more accurate picture of protein-directed biological processes than measurements at the mRNA level.

In principle, global levels of protein abundance could also be measured by protein microarrays (for example, see refs. 82,83). These would use capture antibodies in a chip layout similar to DNA arrays. However, it has been difficult to obtain a global set of capture reagents against the whole proteome and to ensure the requisite specificity. These and other technical difficulties have so far prevented widespread adoption of this approach, and at this point, MS appears to be the most powerful approach to quantitative proteomics.

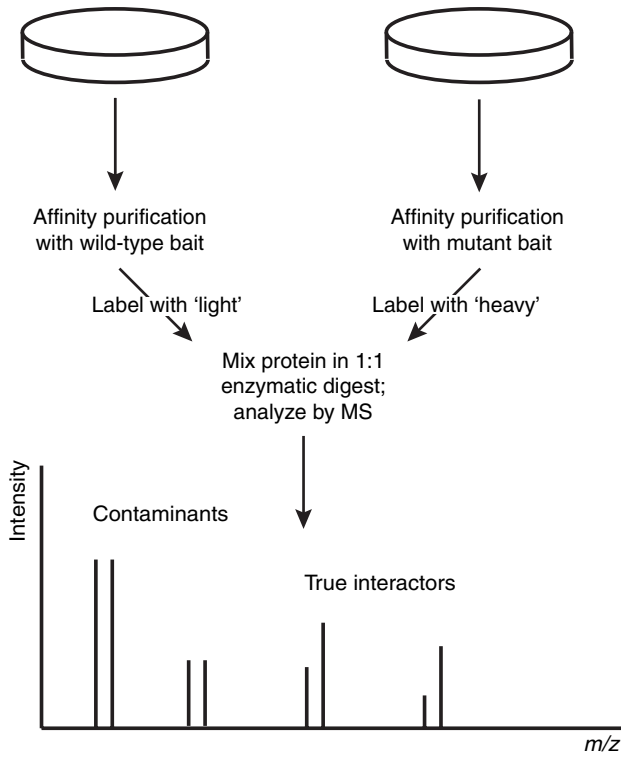


Figure 5 Identification of specific bait-prey interactions in affinity-precipitation experiments. In a coprecipitation experiment, the goal is to identify proteins that bind differentially to wild-type and mutant baits. When performed in combination with quantitative proteomics, proteins that bind unspecifically to the affinity support or beads will show ratios similar to the mixing ratio. Proteins that bind specifically to the bait or secondary interactors will give a ratio indicative of increased binding and enrichment. Repeating the experiment with switched labels should result in inverse ratios, further increasing the specificity of the assay.

What does it mean to quantify a protein in proteomics?

It is now widely recognized that the proteome is much more complex than the genome and transcriptome. Often, a single gene gives rise to several protein isoforms with different tissue specificity and/or expression, and gene families give rise to proteins with closely related sequences. Importantly, these different protein forms can mediate different functions and can be regulated differentially. DNA arrays do not typically capture such subtleties, and the standard quantitative proteomics approaches also fail to do so. This is a recognized problem in protein identification, as peptides unique to an isoform may or may not have been identified[84,85]. There is the analogous pitfall when quantifying proteins with extensive sequence identity (**Fig. 4**). As can be seen in the figure, the two proteins show very different fold changes, as revealed by peptides unique to them, whereas shared peptides exhibit an intermediate ratio.

The statistical reliability of quantitative proteomics is a central question when reporting results in the literature. Generally, high-resolution mass spectrometers help in separating signal from back-

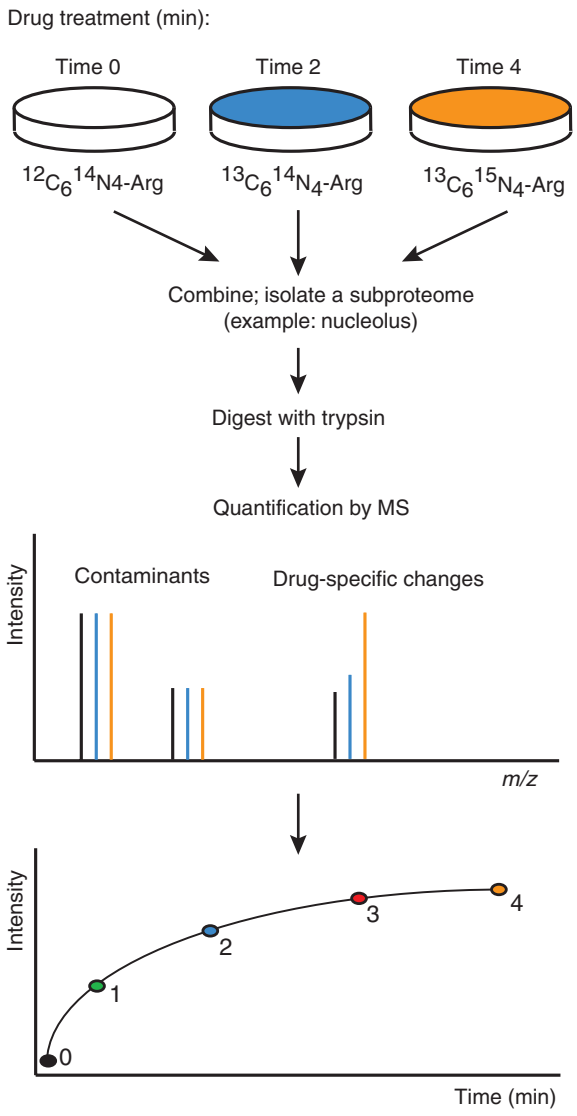


Figure 6 Temporal changes in the nucleolar proteome upon transcriptional inhibition. The SILAC two-state metabolic labeling experiment can be extended to triple encoding with three forms of arginine ($^{12}C_6{}^{14}N_4$-Arg, $^{13}C_6{}^{14}N_4$-Arg, $^{13}C_6{}^{15}N_4$-Arg) to label three cell states, allowing their quantification in a single MS analysis. Each triplet represents the abundance of the corresponding protein in the three proteomes (transcription-inhibited nucleolus). A similar experiment with time points 0, 1 and 3 is performed (data not shown). The ratios from both experiments are combined, generating the curve in the lower panel. In this example, a protein is recruited into the nucleolus as a result of transcriptional inhibition.

ground as well as from coeluting peptides with similar mass; high-sensitivity instruments help in increasing signal of weak peaks, and instruments with a fast sequencing cycle allow quantification of the protein on the basis of several peptides. Because accuracy of quantification is mainly influenced by signal-to-noise, larger peaks can be quantified more accurately than smaller peaks, and one cannot necessarily state a common significance threshold for the entire experiment.

Statistical validity of reported quantification results can be based on the standard deviations obtained in separate runs, from different peptides quantifying the same protein and, for a single peptide, from consecutive scans of coeluting peptide pairs. Ideally, all sources of variability in quantification should be tracked, and a compound standard deviation should be determined. Because quantitative proteomics is a relatively new field, no consensus on how to report statistical confidence has been agreed upon. Further research and standards in this area are urgently needed to ensure quality and transparency. In any case, reporting protein fold changes either without any error intervals or with error intervals only for the whole experiment is not acceptable. Several open-source software packages help in the determination of these ratios, in either a completely automated process[86,87] or a process allowing user intervention (MSQuant[88]). Because the human brain is an excellent judge of visual patterns, human intervention can help in the quantification of important but low-level, noisy or overlapping signals.

Functional studies with quantitative proteomics

The concept of differential display proteomics had its start with classical 2-DE strategies to identify differences in gel patterns. MS-based quantitative proteomics' first applications had also developed along these lines[86,89], and it is still widely used for protein profiling (for example, see refs. 90,91). Although differential display can provide important insights into molecular differences from diseased and normal cells, it is often difficult to tell primary causes apart from secondary effects. Quantitative proteomics uniquely allows the development of targeted assays that can more directly lead to understanding of biological mechanism than these large-scale protein profiling experiments. This will be illustrated with two areas in which quantitative proteomics has been applied very successfully, namely interaction proteomics and measuring the temporal aspects of proteomes.

Interaction proteomics

Mapping protein interaction partners is an excellent starting point in trying to understand a protein's function. These interaction partners may directly implicate the protein in a cellular process. Interaction maps of the yeast proteome have already been obtained by MS-based proteomics[92,93]. However, such assays can be challenging, as interacting partners could be expressed endogenously at low levels in cells, interactions may be transient and a large background of coprecipitating proteins may be present. With biochemical purification, one always has to contend with the tradeoff between sensitivity and specificity. While trying to increase specificity in such pull-downs, one could increase the washing steps to remove contaminating proteins. This rapidly leads to a diminishing number of purified proteins and even potentially losing true interacting partners along the way. Quantitative proteomics can break this *yin* and *yang* of protein biochemistry[94,95] (**Fig. 5**). Briefly, bait and a closely related control are separately incubated with cell lysate. Proteins interacting with bait and control are differentially labeled (by chemical or metabolic labeling) and combined. Background proteins bind equally to bait and to control and therefore lead to one-to-one ratios, whereas specific interactions with the bait result in differential ratios. This principle can be applied to any class of affinity purification, including protein-protein, protein-oligonucleotide and protein-drug[48] interactions.

Temporal dimension in proteomics

So far, proteomics has provided static pictures of cellular proteomes. The study of temporal changes is especially interesting as it effectively converts our observation of biological processes from a 'snapshot' to that of a 'molecular movie'. In order to capture the temporal dimension of events occurring within the cell, experiments have to be compared over a set of varying time points. This is both time consuming and

laborious. Multiplexing these analytical techniques has obvious benefits, as experiments can be designed to compare two experimental states to a control at the same time or to study a time course of perturbation. Blagoev *et al.* used three different forms of arginine in SILAC to compare the phosphotyrosine proteome at five different time points of EGF stimulation[96]. This resulted in temporal profiles for virtually the entire EGF-induced phosphotyrosine-proteome. White and co-workers employed iTRAQ to determine phosphorylation changes at four time points of EGF treatment and clustered these kinetic curves with self-organizing maps[97]. Andersen *et. al.*[73] studied the dynamic composition of a large organelle—the human nucleolus, in response to transcriptional inhibitors and various other drugs (**Fig. 6**). Intriguingly, small and large subunits of the ribosome leave the nucleolus at quite different rates, and proteins in functional or physical subcomplexes show correlated temporal profiles. Time courses of protein synthesis can also be obtained by metabolic labeling methods in quantitative proteomics by measuring the incorporation of the metabolic label[98].

Perspectives

As we have shown, powerful methods for quantitative proteomics have now been developed, and we predict that some form of quantification will soon be a necessity in virtually all proteomics experiments. With the increasing speed and sensitivity of MS instruments, purely qualitative analyses are increasingly difficult to interpret. Quantitative proteomics allows us to distinguish pertinent changes from background proteins and thereby simplifies biological understanding and follow-up of functional experiments.

Specific challenges for quantitative proteomics remain. To enable truly proteome-wide quantification, already impressive MS instrument performance has to be further improved. For example, at least 10,000 peptides in the yeast proteome will have to be quantified in a single analysis. To quantify low-level putative biomarkers in human body fluids, the dynamic range has to be increased by at least 100-fold. Many of these advances can be achieved through improved MS instrumentation, more intelligent algorithms and software, whereas more extensive sample fractionation is incompatible with high-throughput analyses. However, it should be kept in mind that virtually all successes of the field have come not from brute-force analyses, but from clever application of quantitative principles to focus on a functionally relevant subset of the proteome. Where do chemical biologists fit into this picture of quantitative proteomics? On the one hand, they are vital for developing the necessary chemical-labeling reagents as described above. On the other hand, many of their 'chemical proteomics' strategies for targeting specific protein classes[99,100] would also be enhanced tremendously by the quantitative proteomics technologies described in this review.

ACKNOWLEDGMENTS

We gratefully acknowledge our colleagues at the Broad Institute of MIT and Harvard, the Center for Experimental BioInformatics at University of Southern Denmark and the Max Planck Institute for Biochemistry for useful discussions and support. We thank S.A. Carr, B. Küster, L.J. Foster and F. White for critical comments.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturechemicalbiology/

1. Wilkins, M.R. *et al.* Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol. Genet. Eng. Rev.* **13**, 19–50 (1996).
2. Anderson, N.L. & Anderson, N.G. Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* **19**, 1853–1861 (1998).
3. Corthals, G.L., Wasinger, V.C., Hochstrasser, D.F. & Sanchez, J.C. The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* **21**, 1104–1115 (2000).
4. Aebersold, R. & Mann, M. Mass spectrometry–based proteomics. *Nature* **422**, 198–207 (2003).
5. Peng, J. & Gygi, S.P. Proteomics: the move to mixtures. *J. Mass Spectrom.* **36**, 1083–1091 (2001).
6. Brunet, S. *et al.* Organelle proteomics: looking at less to see more. *Trends Cell Biol.* **13**, 629–638 (2003).
7. Garcia, B.A., Shabanowitz, J. & Hunt, D.F. Analysis of protein phosphorylation by mass spectrometry. *Methods* **35**, 256–264 (2005).
8. Kuster, B., Schirle, M., Mallick, P. & Aebersold, R. Scoring proteomes with proteotypic peptide probes. *Nat. Rev. Mol. Cell Biol.* **6**, 577–583 (2005).
9. Lill, J. Proteomic tools for quantitation by mass spectrometry. *Mass Spectrom. Rev.* **22**, 182–194 (2003).
10. Sadygov, R.G., Cociorva, D. & Yates, J.R. Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book. *Nat Methods* **1**, 195–202 (2004).
11. Wysocki, V.H., Resing, K.A., Zhang, Q. & Cheng, G. Mass spectrometry of peptides and proteins. *Methods* **35**, 211–222 (2005).
12. Julka, S. & Regnier, F. Quantification in proteomics through stable isotope coding: a review. *J. Proteome Res.* **3**, 350–363 (2004).
13. Leitner, A. & Lindner, W. Current chemical tagging strategies for proteome analysis by mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **813**, 1–26 (2004).
14. Olsen, J.V., Ong, S.E. & Mann, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* **3**, 608–614 (2004).
15. Allet, N. *et al. In vitro* and *in silico* processes to identify differentially expressed proteins. *Proteomics* **4**, 2333–2351 (2004).
16. Rappsilber, J., Ishihama, I., Foster, L.J., Mittler, G. & Mann, M. in *Proceedings of the 51st American Society for Mass Spectrometry Conference on Mass Spectrometry* (American Society for Mass Spectrometry, Montreal 2003).
17. Steen, H. & Mann, M. The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* **5**, 699–711 (2004).
18. Rappsilber, J., Ryder, U., Lamond, A.I. & Mann, M. Large-scale proteomic analysis of the human spliceosome. *Genome Res.* **12**, 1231–1245 (2002).
19. Sanders, S.L., Jennings, J., Canutescu, A., Link, A.J. & Weil, P.A. Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. *Mol. Cell. Biol.* **22**, 4723–4738 (2002).
20. Ishihama, Y. *et al.* 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*, advance online publication 14 June 2005 (10.174/mcp.M500061-MCP200).
21. Andersen, J.S. *et al.* Proteomic characterization of the human centrosome by protein correlation profiling. *Nature* **426**, 570–574 (2003).
22. Chelius, D. & Bondarenko, P.V. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J. Proteome Res.* **1**, 317–323 (2002).
23. Gillette, M.A., Mani, D.R. & Carr, S.A. Place of pattern in proteomic biomarker discovery. *J. Proteome Res.* **4**, 1143–1154 (2005).
24. Carr, S.A., Huddleston, M.J. & Annan, R.S. Selective detection and sequencing of phosphopeptides at the femtomole level by mass spectrometry. *Anal. Biochem.* **239**, 180–192 (1996).
25. Verma, R. *et al.* Phosphorylation of Sic1p by G1 Cdk required for its degradation and entry into S phase. *Science* **278**, 455–460 (1997).
26. Steen, H., Jebanathirajah, J.A., Springer, M. & Kirschner, M.W. Stable isotope-free relative and absolute quantitation of protein phosphorylation stoichiometry by MS. *Proc. Natl. Acad. Sci. USA* **102**, 3948–3953 (2005).
27. De Leenheer, A.P. & Thienpont, L.M. Applications of isotope dilution-mass spectrometry in clinical chemistry, pharmcokinetics, and toxicology. *Mass Spectrom. Rev.* **11**, 249–307 (1992).
28. Wulfkuhle, J.D., Liotta, L.A. & Petricoin, E.F. Proteomic applications for the early detection of cancer. *Nat. Rev. Cancer* **3**, 267–275 (2003).
29. Zhang, R. & Regnier, F.E. Minimizing resolution of isotopically coded peptides in comparative proteomics. *J. Proteome Res.* **1**, 139–147 (2002).
30. Kusmierz, J.J., Sumrada, R. & Desiderio, D.M. Fast atom bombardment mass spectrometric quantitative analysis of methionine-enkephalin in human pituitary tissues. *Anal. Chem.* **62**, 2395–2400 (1990).
31. Stemmann, O., Zou, H., Gerber, S.A., Gygi, S.P. & Kirschner, M.W. Dual inhibition of sister chromatid separation at metaphase. *Cell* **107**, 715–726 (2001).
32. Kuhn, E. *et al.* Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and 13C-labeled peptide standards. *Proteomics* **4**, 1175–1186 (2004).
33. Hopfgartner, G. *et al.* Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules. *J. Mass Spectrom.* **39**, 845–855 (2004).
34. Kirkpatrick, D.S., Gerber, S.A. & Gygi, S.P. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* **35**, 265–273 (2005).
35. Havlis, J. & Shevchenko, A. Absolute quantification of proteins in solutions and in polyacrylamide gels by mass spectrometry. *Anal. Chem.* **76**, 3029–3036 (2004).
36. Aebersold, R. Constellations in a cellular universe. *Nature* **422**, 115–116 (2003).
37. Beynon, R.J., Doherty, M.K., Pratt, J.M. & Gaskell, S.J. Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signatures. *Nature Methods* **2**, 587–589 (2005).

38. Rose, K. *et al.* A new mass-spectrometric C-terminal sequencing technique finds a similarity between γ-interferon and α 2-interferon and identifies a proteolytically clipped γ-interferon that retains full antiviral activity. *Biochem. J.* **215**, 273–277 (1983).
39. Mirgorodskaya, O.A. *et al.* Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)O-labeled internal standards. *Rapid Commun. Mass Spectrom.* **14**, 1226–1232 (2000).
40. Yao, X., Freas, A., Ramirez, J., Demirev, P.A. & Fenselau, C. Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.* **73**, 2836–2842 (2001).
41. Rao, K.C., Palamalai, V., Dunlevy, J.R. & Miyagi, M. Lys-N catalyzed 18O labeling for comparative proteomics: Application to cytokines/LPS treated human retinal pigment epithelium cell line. *Mol. Cell Proteomics*, advance online publication 5 July 2005 (10.1074/mcp.M500150-MCP200).
42. Yao, X., Afonso, C. & Fenselau, C. Dissection of proteolytic 18O labeling: endoprotease-catalyzed 16O-to-18O exchange of truncated peptide substrates. *J. Proteome Res.* **2**, 147–152 (2003).
43. Bantscheff, M., Dumpelfeld, B. & Kuster, B. Femtomol sensitivity post-digest (18)O labeling for relative quantification of differential protein complex composition. *Rapid Commun. Mass Spectrom.* **18**, 869–876 (2004).
44. Staes, A. *et al.* Global differential non-gel proteomics by quantitative and stable labeling of tryptic peptides with oxygen-18. *J. Proteome Res.* **3**, 786–791 (2004).
45. Gygi, S.P. *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999 (1999).
46. Hansen, K.C. *et al.* Mass spectrometric analysis of protein mixtures at low levels using cleavable 13C-isotope-coded affinity tag and multidimensional chromatography. *Mol. Cell. Proteomics* **2**, 299–314 (2003).
47. Li, J., Steen, H. & Gygi, S.P. Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents: the yeast salinity stress response. *Mol. Cell. Proteomics* **2**, 1198–1204 (2003).
48. Oda, Y. *et al.* Quantitative chemical proteomics for identifying candidate drug targets. *Anal. Chem.* **75**, 2159–2165 (2003).
49. Olsen, J.V. *et al.* HysTag–a novel proteomic quantification tool applied to differential display analysis of membrane proteins from distinct areas of mouse brain. *Mol. Cell. Proteomics* **3**, 82–92 (2004).
50. Nielsen, P.A. *et al.* Proteomic mapping of brain plasma membrane proteins. *Mol. Cell. Proteomics* **4**, 402–408 (2005).
51. Goodlett, D.R. *et al.* Differential stable isotope labeling of peptides for quantitation and *de novo* sequence derivation. *Rapid Commun. Mass Spectrom.* **15**, 1214–1221 (2001).
52. Ficarro, S.B. *et al.* Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **20**, 301–305 (2002).
53. Salomon, A.R. *et al.* Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry. *Proc. Natl. Acad. Sci. USA* **100**, 443–448 (2003).
54. Brill, L.M. *et al.* Robust phosphoproteomic profiling of tyrosine phosphorylation sites from human T cells using immobilized metal affinity chromatography and tandem mass spectrometry. *Anal. Chem.* **76**, 2763–2772 (2004).
55. Munchbach, M., Quadroni, M., Miotto, G. & James, P. Quantitation and facilitated *de novo* sequencing of proteins by isotopic N-terminal labeling of peptides with a fragmentation-directing moiety. *Anal. Chem.* **72**, 4047–4057 (2000).
56. Ji, J. *et al.* Strategy for qualitative and quantitative analysis in proteomics based on signature peptides. *J. Chromatogr. B Biomed. Sci. Appl.* **745**, 197–210 (2000).
57. Schmidt, A., Kellermann, J. & Lottspeich, F. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **5**, 4–15 (2005).
58. Zhang, X., Jin, Q.K., Carr, S.A. & Annan, R.S. N-Terminal peptide labeling strategy for incorporation of isotopic tags: a method for the determination of site-specific absolute phosphorylation stoichiometry. *Rapid Commun. Mass Spectrom.* **16**, 2325–2332 (2002).
59. Ross, P.L. *et al.* Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3**, 1154–1169 (2004).
60. Langen, H., Fountoulakis, M., Evers, S., Wipf, B. & Berndt, P. in *From Genome to Proteome, 3rd Siena 2D Electrophoresis Meeting* (Wiley-VCH, Weinheim, Germany, Siena, 1998).
61. Oda, Y., Huang, K., Cross, F.R., Cowburn, D. & Chait, B.T. Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. USA* **96**, 6591–6596 (1999).
62. Conrads, T.P. *et al.* Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and 15N-metabolic labeling. *Anal. Chem.* **73**, 2132–2139 (2001).
63. Krijgsveld, J. *et al.* Metabolic labeling of *C. elegans* and *D. melanogaster* for quantitative proteomics. *Nat. Biotechnol.* **21**, 927–931 (2003).
64. Wu, C.C., MacCoss, M.J., Howell, K.E., Matthews, D.E. & Yates, J.R.I. Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal. Chem.* **76**, 4951–4959 (2004).
65. Ong, S.E. *et al.* 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 (2002).
66. Chen, X., Smith, L.M. & Bradbury, E.M. Site-specific mass tagging with stable isotopes in proteins for accurate and efficient protein identification. *Anal. Chem.* **72**, 1134–1143 (2000).
67. Veenstra, T.D., Martinovic, S., Anderson, G.A., Pasa-Tolic, L. & Smith, R.D. Proteome analysis using selective incorporation of isotopically labeled amino acids. *J. Am. Soc. Mass Spectrom.* **11**, 78–82 (2000).
68. Zhu, H., Pan, S., Gu, S., Bradbury, E.M. & Chen, X. Amino acid residue specific stable isotope labeling for quantitative proteomics. *Rapid Commun. Mass Spectrom.* **16**, 2115–2123 (2002).
69. Jiang, H. & English, A.M. Quantitative analysis of the yeast proteome by incorporation of isotopically labeled leucine. *J. Proteome Res.* **1**, 345–350 (2002).
70. Ong, S.E., Kratchmarova, I. & Mann, M. Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J. Proteome Res.* **2**, 173–181 (2003).
71. Foster, L.J., De Hoog, C.L. & Mann, M. Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proc. Natl. Acad. Sci. USA* **100**, 5813–5818 (2003).
72. Ibarrola, N., Kalume, D.E., Gronborg, M., Iwahori, A. & Pandey, A. A proteomic approach for quantitation of phosphorylation using stable isotope labeling in cell culture. *Anal. Chem.* **75**, 6043–6049 (2003).
73. Andersen, J.S. *et al.* Nucleolar proteome dynamics. *Nature* **433**, 77–83 (2005).
74. Ibarrola, N., Molina, H., Iwahori, A. & Pandey, A. A novel proteomic approach for specific identification of tyrosine kinase substrates using [13C]tyrosine. *J. Biol. Chem.* **279**, 15805–15813 (2004).
75. Ong, S.E., Mittler, G. & Mann, M. Identifying and quantifying *in vivo* methylation sites by heavy methyl SILAC. *Nat. Methods* **1**, 119–126 (2004).
76. Scott, L., Lamb, J., Smith, S. & Wheatley, D.N. Single amino acid (arginine) deprivation: rapid and selective death of cultured transformed and malignant cells. *Br. J. Cancer* **83**, 800–810 (2000).
77. Gruhler, A. *et al.* Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* **4**, 310–327 (2005).
78. Gehrmann, M.L., Hathout, Y. & Fenselau, C. Evaluation of metabolic labeling for comparative proteomics in breast cancer cells. *J. Proteome Res.* **3**, 1063–1068 (2004).
79. Amanchy, R., Kalume, D.E. & Pandey, A. Stable isotope labeling with amino acids in cell culture (SILAC) for studying dynamics of protein abundance and posttranslational modifications. *Sci. STKE* **267**, l2 (2005).
80. Ong, S.E. *et al.* Stable isotope labeling by amino acids in cell culture (SILAC) for quantitative proteomics. in *Handbook of Cell Biology* 3rd edn. Vol. IV (ed. Celis, J.) (Academic Press, San Diego, 2005).
81. Ishihama, Y. *et al.* Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. *Nat. Biotechnol.* **23**, 617–621 (2005).
82. Schweitzer, B. & Kingsmore, S.F. Measuring proteins on microarrays. *Curr. Opin. Biotechnol.* **13**, 14–19 (2002).
83. Zhu, H., Bilgin, M. & Snyder, M. Proteomics. *Annu. Rev. Biochem.* **72**, 783–812 (2003).
84. Rappsilber, J. & Mann, M. What does it mean to identify a protein in proteomics? *Trends Biochem. Sci.* **27**, 74–78 (2002).
85. Nesvizhskii, A.I. & Aebersold, R. Interpretation of shotgun proteomics data: The protein inference problem. *Mol. Cell Proteomics*, advance online publication 11 July 2005 (10.1074/mcp.R500012-MCP200).
86. Han, D.K., Eng, J., Zhou, H. & Aebersold, R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.* **19**, 946–951 (2001).
87. MacCoss, M.J., Wu, C.C., Liu, H., Sadygov, R. & Yates, J.R., III. A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. *Anal. Chem.* **75**, 6912–6921 (2003).
88. Schulze, W.X. & Mann, M. A novel proteomic screen for peptide-protein interactions. *J. Biol. Chem.* **279**, 10756–10764 (2004).
89. Shiio, Y. *et al.* Quantitative proteomic analysis of Myc oncoprotein function. *EMBO J.* **21**, 5088–5096 (2002).
90. Everley, P.A., Krijgsveld, J., Zetter, B.R. & Gygi, S.P. Quantitative cancer proteomics: stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research. *Mol. Cell. Proteomics* **3**, 729–735 (2004).
91. Gu, S. *et al.* Large-scale quantitative proteomic study of PUMA-induced apoptosis using two-dimensional liquid chromatography-mass spectrometry coupled with amino acid-coded mass tagging. *J. Proteome Res.* **3**, 1191–1200 (2004).
92. Ho, Y. *et al.* Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415**, 180–183 (2002).
93. Gavin, A.C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
94. Blagoev, B. *et al.* A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* **21**, 315–318 (2003).
95. Ranish, J.A. *et al.* The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* **33**, 349–355 (2003).
96. Blagoev, B., Ong, S.E., Kratchmarova, I. & Mann, M. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* **22**, 1139–1145 (2004).
97. Zhang, Y. *et al.* Time-resolved mass spectrometry of tyrosine phosphorylatiopn sites in the EGF receptor signaling network reveals dynamic modules. *Mol. Cell Proteomics*, advance online publication 11 June 2005 (10.1074/mcp.M500089-MCP200).
98. Pratt, J.M. *et al.* Dynamics of protein turnover, a missing dimension in proteomics. *Mol. Cell. Proteomics* **1**, 579–591 (2002).
99. Saghatelian, A. & Cravatt, B.F. Assignment of protein function in the postgenomic era. *Nat Chem. Biol.* **1**, 130–142 (2005).
100. Verhelst, S.H. & Bogyo, M. Chemical proteomics applied to target identification and drug discovery. *Biotechniques* **38**, 175–177 (2005).
101. Sechi, S. & Chait, B.T. Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal. Chem.* **70**, 5150–5158 (1998).

102. Shen, M. *et al.* Isolation and isotope labeling of cysteine- and methionine-containing tryptic peptides: application to the study of cell surface proteolysis. *Mol. Cell. Proteomics* **2**, 315–324 (2003).
103. Sebastiano, R., Citterio, A., Lapadula, M. & Righetti, P.G. A new deuterated alkylating agent for quantitative proteomics. *Rapid Commun. Mass Spectrom.* **17**, 2380–2386 (2003).
104. Pasquarello, C., Sanchez, J.C., Hochstrasser, D.F. & Corthals, G.L. N-t-butyliodoacetamide and iodoacetanilide: two new cysteine alkylating reagents for relative quantitation of proteins. *Rapid Commun. Mass Spectrom.* **18**, 117–127 (2004).
105. Zhou, H., Ranish, J.A., Watts, J.D. & Aebersold, R. Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. *Nat. Biotechnol.* **20**, 512–515 (2002).
106. Qiu, Y., Sousa, E.A., Hewick, R.M. & Wang, J.H. Acid-labile isotope-coded extractants: a class of reagents for quantitative mass spectrometric analysis of complex protein mixtures. *Anal. Chem.* **74**, 4969–4979 (2002).
107. Shi, Y. *et al.* A simple solid phase mass tagging approach for quantitative proteomics. *J. Proteome Res.* **3**, 104–111 (2004).
108. Shi, Y., Xiang, R., Horvath, C. & Wilkins, J.A. Quantitative analysis of membrane proteins from breast cancer cell lines BT474 and MCF7 using multistep solid phase mass tagging and 2D LC/MS. *J. Proteome Res.* **4**, 1427–1433 (2005).
109. Thompson, A. *et al.* Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **75**, 1895–1904 (2003).
110. Wang, S. & Regnier, F.E. Proteomics based on selecting and quantifying cysteine containing peptides by covalent chromatography. *J. Chromatogr. A.* **924**, 345–357 (2001).
111. Che, F.Y. & Fricker, L.D. Quantitation of neuropeptides in Cpe(fat)/Cpe(fat) mice using differential isotopic tags and mass spectrometry. *Anal. Chem.* **74**, 3190–3198 (2002).
112. Mason, D.E. & Liebler, D.C. Quantitative analysis of modified proteins by LC-MS/MS of peptides labeled with phenyl isocyanate. *J. Proteome Res.* **2**, 265–272 (2003).
113. Lee, Y.H., Han, H., Chang, S.B. & Lee, S.W. Isotope-coded N-terminal sulfonation of peptides allows quantitative proteomic analysis with increased *de novo* peptide sequencing capability. *Rapid Commun. Mass Spectrom.* **18**, 3019–3027 (2004).
114. Hoang, V.M. *et al.* Quantitative proteomics employing primary amine affinity tags. *J. Biomol. Tech.* **14**, 216–223 (2003).
115. Hsu, J.L., Huang, S.Y., Chow, N.H. & Chen, S.H. Stable-isotope dimethyl labeling for quantitative proteomics. *Anal. Chem.* **75**, 6843–6852 (2003).
116. Cagney, G. & Emili, A. *De novo* peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. *Nat. Biotechnol.* **20**, 163–170 (2002).
117. Brancia, F.L., Openshaw, M.E. & Kumashiro, S. Investigation of the electrospray response of lysine-, arginine-, and homoarginine-terminal peptide mixtures by liquid chromatography/mass spectrometry. *Rapid Commun. Mass Spectrom.* **16**, 2255–2259 (2002).
118. Brancia, F.L., Montgomery, H., Tanaka, K. & Kumashiro, S. Guanidino labeling derivatization strategy for global characterization of peptide mixtures by liquid chromatography matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* **76**, 2748–2755 (2004).
119. Beardsley, R.L. & Reilly, J.P. Quantitation using enhanced signal tags: a technique for comparative proteomics. *J. Proteome Res.* **2**, 15–21 (2003).
120. Peters, E.C., Horn, D.M., Tully, D.C. & Brock, A. A novel multifunctional labeling reagent for enhanced protein characterization with mass spectrometry. *Rapid Commun. Mass Spectrom.* **15**, 2387–2392 (2001).
121. Syka, J.E. *et al.* Novel linear quadrupole ion trap/FT mass spectrometer: performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J. Proteome Res.* **3**, 621–626 (2004).
122. Kuyama, H. *et al.* An approach to quantitative proteome analysis by labeling tryptophan residues. *Rapid Commun. Mass Spectrom.* **17**, 1642–1650 (2003).
123. Goshe, M.B. *et al.* Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. *Anal. Chem.* **73**, 2578–2586 (2001).
124. Qian, W.J. *et al.* Phosphoprotein isotope-coded solid-phase tag approach for enrichment and quantitative analysis of phosphopeptides from complex mixtures. *Anal. Chem.* **75**, 5441–5450 (2003).
125. Wells, L. *et al.* Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. *Mol. Cell. Proteomics* **1**, 791–804 (2002).
126. Amoresano, A., Marino, G., Cirulli, C. & Quemeneur, E. Mapping phosphorylation sites: a new strategy based on the use of isotopically labelled DTT and mass spectrometry. *Eur. J. Mass Spectrom. (Chichester, Eng)* **10**, 401–412 (2004).
127. Vosseller, K. *et al.* Quantitative analysis of both protein expression and serine / threonine post-translational modifications through stable isotope labeling with dithiothreitol. *Proteomics* **5**, 388–398 (2005).