Biotechnology and Genetic Engineering Reviews - Vol. 24, 147-164 (2007)

Mass spectrometry-based quantitative proteomics

TATSUJI NAKAMURA AND YOSHIYA ODA*

Laboratory of Seeds Finding Technology, Eisai Co., Ltd., Tokodai 5-1-3, Tsukuba, Ibaraki 300-2635, Japan

Introduction

The genomic era in the 1990s provided massive amounts of information about DNA sequences from many species. This wealth of data has been annotated and is continuing to be analyzed by new bioinformatics algorithms. The aim of proteomics is to determine the structure, function, and expression of all proteins and their isoforms of a genome. Proteomics may be defined broadly as the study of all proteins expressed by a genome. The recent revolution in the development of diverse comprehensive analytical methodologies for providing extensive knowledge about the proteome has led to unprecedented interest. In particular, advances in mass spectrometry (MS) have enabled the identification of thousands of both known and novel proteins (Domon and Aebersold 2006; Aebersold and Mann 2003; Elias et al, 2005). However, these methods have typically provided qualitative and static information, therefore, many of the functional proteomic studies described to date are qualitative and discovery based. There is also a need for more quantitative analysis of functional proteomics. For example, the distinction between normal states and stimulated/disease states are very important for elucidating indicators of drug safety, mechanism of action, efficacy, and disease state progression. These indicators, referred as *biomarkers* may dramatically improve the efficiency of drug discovery and development. Today, several high-throughput methods are available that provide quantitative information. The most commonly used technology for monitoring changes in the expression of complex protein mixtures is still two-dimensional gel electrophoresis (2-DE) (Gorg et al, 2004).

*To whom correspondence may be addressed (y-oda@hhc.eisai.co.jp)

Abbreviations: 2-DE, 2-dimensional gel electrophoresis; AQUA, absolute quantitation; CDIT, culture derived isotope tag; ICAT, isotope coated affinity tag; LC, liquid chromatography; MS, mass spectrometry; PAI, protein abundance index; emPAI, exponential form of PAI minus 1.

However, 2-DE does not allow the identification of proteins. The proteins of interest must be picked out from the gel one by one, treated with tryptic digestion, and analyzed by MS to identify proteins combination with search engines. Although 2-DE is still the method of choice for proteomics, there are a lot of unsettled issues such as reproducibility, difficulty in detection of scarce proteins, and incompatibility for a hydrophobic, high molecular weight, or high pI protein analysis. Considering the reasons mentioned above, we have decided to focus on reviewing the methods of MS-based quantitative proteomics since MS has been an essential tool for protein identification and analysis of post-translational modification, and MS offers a sensitive and selective detection system.

Absolute quantitation by mass spectrometry

MS based quantitation has been mainly classified into two categories, with and without internal standard approaches. Also the term of *quantitation* has two meanings, namely absolute and relative values. Traditionally, MS has been widely used for the quantitative measurement of specific small molecules, drug metabolites, and hormones, with excellent precision and high specificity and very high throughput. However, the ionization process in MS is usually low in reproducibility and this is in addition to variable sample preparation procedures. In order to overcome these disadvantages, normalization to reduce error is very important for MS based quantitation. These errors can be significantly reduced using internal standards. Most importantly, the internal standards should be very well characterized both physically and chemically, and be appropriate for the particular analyte being analyzed. For this reason, stable isotopes have been used as internal standards and small amounts of peptides have been quantified by MS against stable isotope-labeled standards in the 1990s (Desidero, 1992; Kippen et al, 272; Wang et al, 1996; Kusmierz et al, 1990); commonly used isotopes include ^{2}H, ^{13}C, ^{15}N and ^{18}O. In this approach, calibration curves are usually prepared before real sample analysis by using standard samples, therefore this is absolute quantitation. In order to quantitate proteins in a sample, each protein must be quantitatively digested into its constituent peptides by complete chemical or enzymatic cleavage. Within this digest one can select a monitor peptide to serve as a quantitative surrogate for the protein then achieve accurate quantitation by spiking with a stable isotope-labeled version of the same peptide as internal standard (*Figure 1*). This approach is called isotope dilution or absolute quantitation (AQUA), and is convenient to compare the data with others. Unfortunately, these studies were limited to the quantitation of analytes, because internal standards have to be prepared.

Relative quantitation by mass spectrometry

Relative expression levels of cellular proteins under different conditions, e.g., normal and disease states; cells subjected to different stimuli and a similar fashion to genomic analyses, have been measured using stable isotope labeling strategies. Seminal work on differential isotopic labeling proteins by Oda et al (1999) has led to the development of tools such as isotope coded affinity tags or ICAT (Gygi et al, 1999). A common

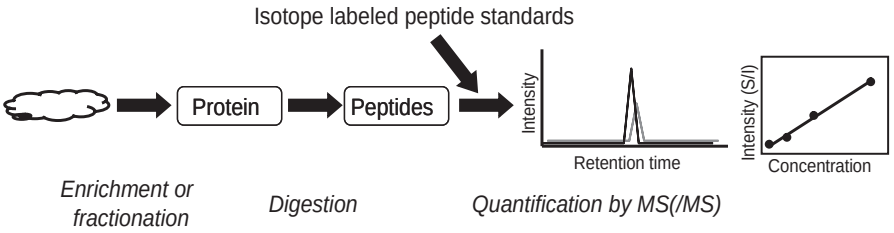

Figure 1. Strategy for absolute quantitation by mass spectrometry. Protein are enriched or fractionated from a sample, then digested to peptides. Isotope labeled peptide standards are spiked into samples as internal standards (IS). Identification of protein and absolute concentration of peptides in the sample are then analyzed by MS.

feature of these techniques (*Figure 2*) is that protein profiling can be performed by comparing the amount of proteins present in two different cell states in which a reference is labeled with a light isotope labeled with another labeled with a heavy isotope. The reference and heavy-isotope labeled samples are then combined where individual peptides have the same amino acid sequence, therefore they exhibit identical chemical properties (i.e. they behave identically throughout any separation steps and ionization process in MS), but differ in mass by the weight of the isotopic label which is easily observed in MS. The ratios of the peaks from the MS spectra representing labeled and unlabelled species then accurately reflects the relative quantities of the two within a mixture. Thus, isotope labeling can be used to obtain precise quantitative proteomics information. This includes *in vivo* (metabolic) labeling, enzyme labeling combination of $^{18}O/^{16}O$ and trypsin (Xao et al 2001; Mirgorodskaya et al, 2000; Wang et al, 2001), and chemical labeling such as ICAT (isotope-coded affinity technology) (Gygi et al, 1999) or iTRAQ (Ross et al, 2004).

In vivo (metabolic) labeling

In an original publication about in vivo labeling of proteins using heavy isotope incorporation (Oda et al, 1999; Ong et al, 2002) protein expression profiles were quantitated through the use of whole-cell stable isotope labeling. The basic premise of this approach (*Figure 3*) involves growing one cell on a medium lacking an essential nutrition and supplementing the medium with a stable isotopic amino acid. After several doublings, almost all proteins are saturated with stable isotope labeled amino acids. The *precursor* ion scan is used to quantify the relative abundances of the heavy versus light versions of each peptide, and the *product* ion scan is for peptide identification. The advantages of this technique include amongst other things minimal sample manipulation. After the labeling, reference and sample cultures can be mixed thus subsequent cell lysis, protein extraction, sample work-up and separation on two-dimensional gels or liquid chromatography are identical for each, eliminating the possibility that variability in these steps may affect the calculated ratio of proteins of interest. For instance, a practical proteomic approach may focus on a more limited proteome from cell fractionation to a cellular organelle (e.g. mitochondria), component,

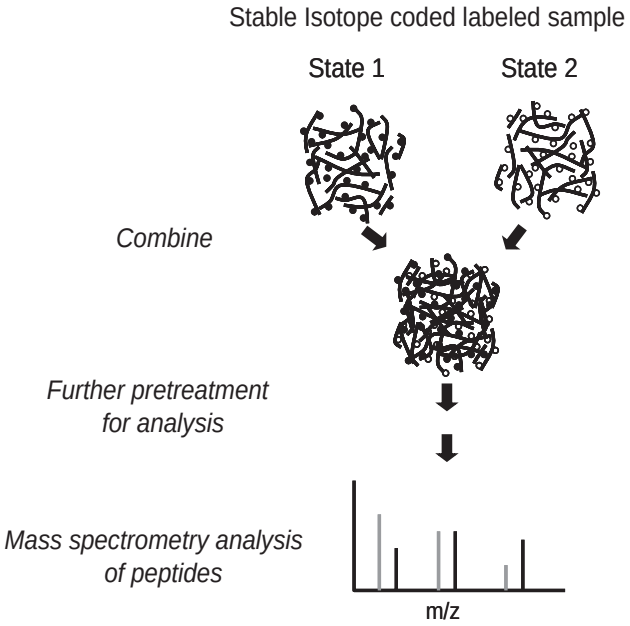

Figure 2. Overall strategies for relative quantitation by mass spectrometry. Samples of two states are labeled with chemically identical reagents differing in their molecular weight. Then the samples are combined and processed for MS analysis. Relative quantification is then performed by comparison of the obtained peptide signal intensities.

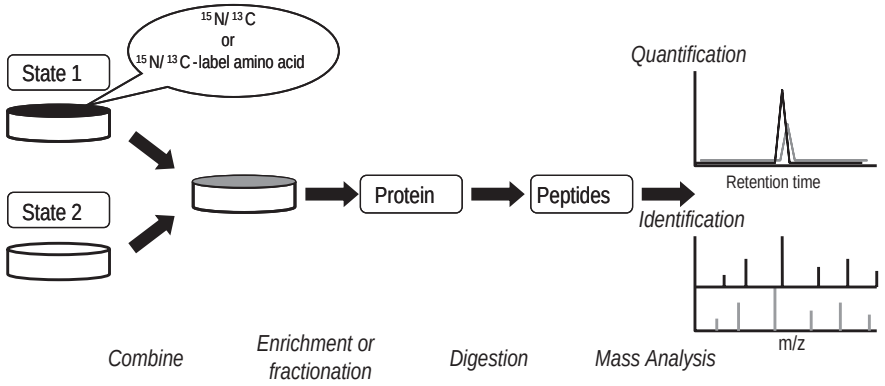

Figure 3. Strategy for quantitation by *in vivo* labeling. Cells are grown in two media, each containing either light or heavy isotope. Following isotopic labeling, the two cell cultures are combined, processed then, analyzed by MS

such as the nucleosome (Anderesen and Mann, 2006; Brunet et al, 2003), or specific modifications, such as phosphorylation (Kokubu et al, 2005; Morandell et al, 2006). Many investigators do emphasize the importance of a quantitative, dynamic description of a focused proteome. Subcellular fractionation (e.g. organelle proteome) or

enrichment of modified peptides (e.g. phosphoproteome) allows identification of less abundant proteins. There is no amplification method for proteins, analogous to the polymerase chain reaction method for DNA or RNA amplification. The absence of an amplification method means that fractionation or enrichment of biological samples to remove highly abundant proteins is a quite important step. For instance, subcellular fractionation consists of two major laborious steps, homogenization and fractionation of the homogenate to separate the different populations of organelles. Although many laboratories have begun to apply these traditional subcellular fractionation procedures within proteome studies, the development of standardized and ready-to-go procedures for subcellular fractionation, or highly reproducible phosphoproteomics is still difficult. To normalize these variations during sample preparation, *in vivo* labeling is the most robust and practical method for quantitative proteomics. However, in vivo labeling of proteins in living cells with stable isotopes requires the metabolic pathway to be accessible to the label thus several samples, particularly animal tissues, are very difficult to be labeled with stable isotopes. Recently, Yates and coworkers (Wu et al, 2004) have reported a method to label mammalian organisms with a diet enriched by long-term in vivo labeling. In their paper, a metabolically labeled rat with stable isotopes was used for tissue-specific internal standards. Subsequent quantitative shotgun comparison of rat tissue from drug-treated and untreated rats was done by adding internal standard tissue to each sample tissue. They quantified 310 different proteins and 127 proteins were altered the protein level. Although their approach was valuable, the drawbacks of their method are that it takes long time (44 days) to get a labeled rat with expensive diet, and some tissues such as brain are not completely labeled with stable isotopes.

Another quantitative approach has been developed for mammalian organism proteome using in vivo labeling. Mouse neuroblastoma Neuro2A cells were cultured in ^{13}C-labeled leucine rich media, which was named as Culture-Derived Isotope Tag (CDIT) cells. CDIT were used as comprehensive internal standards to normalize the variations of sample preparation and the analysis (Ishihama et al, 2005b). Absolute quantitation of proteins in samples requires the preparation of known amounts of isotopically labeled standard peptides to spike into a sample, which means that one of the proteolytic peptides of a particular target protein needs to be synthesized using isotope-labeled reagents and then the absolute amount is measured. However, this method is difficult to apply to large number of quantitative protein analyses because peptide synthesis usually requires 10-fold excess amount of reagents (expensive isotopically labeled reagents in this case), and also because the conventional peptide synthesis is performed in microgram to milligram amounts, while MS requires only femtogram to picogram amounts of peptide. In the CDIT approach, absolute amounts of target proteins in a sample can be measured in a cost effective way. Since all the proteins in the CDIT cells have already been labeled with stable isotopes, conventional unlabeled synthetic peptides can be used to identify the absolute amounts of target proteins. We used all labeled peptides from the CDIT cells for quantification of the unlabeled tissue sample spiked with CDIT cells at the second step.

In vitro (chemical) labeling

An alternative strategy for ICAT analysis (Gygi et al, 1999) (*Figure 4*) involves a

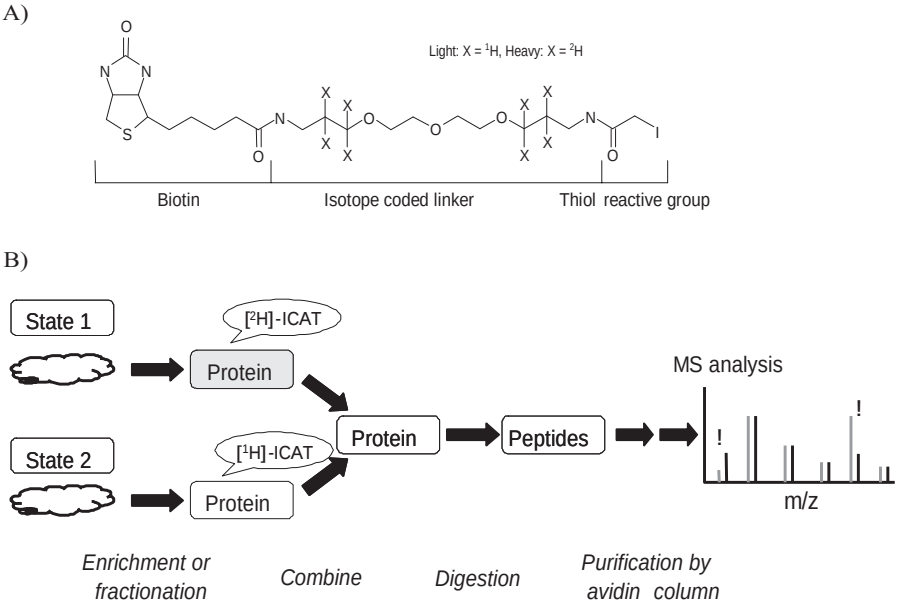

Figure 4. (A) ICAT reagent and (B) strategy for quantitation by ICAT. Two protein mixtures from different states (cells or tissue) are labeled with heavy and light ICAT reagent. The labeled mixtures are then combined, digested into peptides by trypsin and subjected to avidin affinity chromatography to enrich the labeled peptides. The obtained peptides are subjected to MS analysis

reagent with three different parts, an iodoacetamide group that reacts with peptide sulfhydryl groups, an intermediate bridge that contains the isotopic carbon or hydrogen atoms allowing for differential labeling, and a biotin group for affinity capture of the derivatized peptides onto a streptavidin column. The ICAT labeling enables identification of less abundant proteins from the large number of protein mixtures. Also the ICAT-labeled peptides are separated from unlabelled peptides via a biotin affinity handle, resulting in a >10-fold reduction in complexity of the overall peptide mixture. However, looking at ICAT from a different point of view, it can only be applied to proteins that have a cysteine group (approximately 80% of the proteome). In the original version of the ICAT reagent, deuterium labeling of the bridge region was used. However, the d0 and d8-labeled peptides are chromatographically partially resolvable, resulting in an inconsistent d0/d8 ratio across the peptide peak. This problem has been overcome by using 13C labeling of the ICAT reagent [22-24] - this isotopic label has no effect on the chromatographic retention time of the ICAT peptide. *In vitro* chemical labeling such as ICAT has also tremendous advances over the analysis of proteins in tissues, or body fluids, for which the *in vivo* labeling method is difficult to apply. In the ICAT strategy, some well-documented problems are found with the original ICAT reagent including its relatively high mass fragmentation of the tag itself, which might make identification of protein ambiguous. Furthermore, the commercially available reagents are expensive, and time-consuming steps are required to attach ICAT tags and remove excess reagents, which can lead to sample losses,

reducing the sensitivity. There is also the possibility of undesirable side-reactions in the chemical modification such as ICAT. Other drawbacks of the ICAT strategy are that ICAT may only provide information about one peptide from that protein, and which may lose post-translation modificational information. Finally a practical limitation of this method is that the elution of the biotinylated peptides from the streptavidin column is not quantitative, particularly for the less abundance proteins.

In vivo labeling, ICAT, enzyme labeling of $^{16}O/^{18}O$ with trypsin, and most other chemical labelings are based on a mass difference on MS spectra for relative quantitation. There are, however, some limitations imposed by mass-difference labeling. The mass-difference concept for many practical purposes is limited to a binary (2-plex) of ternary (3-plex) set of reagents. This limitation makes comparison of multiple states difficult to undertake and increases the complexity of MS using mass difference labels. Therefore a multiplexed set of reagents for quantitative protein analysis has been developed. In a digest mixture N termini and lysine side chains of peptides are labeled with isobaric mass reagents, which are differentially isotopically labeled so that all derivatized peptides are isobaric and chromatographically indistinguishable. Later the signature or reporter ions following MS/MS can be used to identify and quantify individual members of the multiplex set. This reagent is now commercially available as iTRAQ reagents (Ross et al, 2004), which allow 4-plex analysis to be done at once (*Figure 5*). Peptide coverage of iTRAQ approach was found to be significantly increased relative to ICAT.

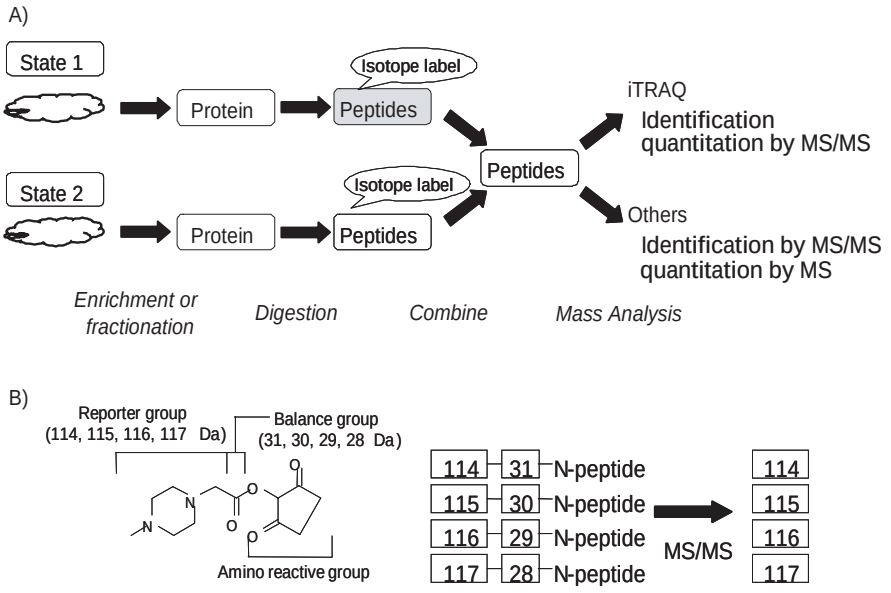

Figure 5. (A)Strategy for quantitation by peptide labeling. (B) Structure of iTRAQ reagent and principles of analysis by MS/MS. Protein mixtures (up to four different samples) are digested by trypsin. Next, the resultant peptides are labeled with individual iTRAQ reagents. Then, the labeled peptides are combined and analyzed for identification and quantitation by LC-MS/MS. The labeled peptides from different state are isobaric, and there are four tags that produce fragment ions of m/z 114, 115, 116, and 117 which show different behavior only in the MS/MS spectra.

Quantitation of post-translational modified proteins

Identification of proteins is not always sufficient for the interpretation of its biological function, because many of the naturally occurring proteins are post-translationally modified. One such ubiquitous modification is phosphorylation that affects a significant subset of the proteome and plays an especially important role in numerous regulatory pathways and cell cycle control in all living cells. Studies in protein phosphorylation may hold particular promise for dissecting signaling pathways, molecular classification of diseases, and profiling of novel kinase-inhibitors (Kokubu et al, 2004). Defining the sites of phosphorylation in specified proteins is frequently a mandatory step for gaining mechanistic insight into a given biological event. MS approaches for identifying phosphorylated peptides are attractive, and can achieve extremely rapid, accurate phosphorylated peptide identification. Modified peptides such as those containing a phosphorylated amino acid can be detected by using MS and sequence databases. If the peptide under investigation can be matched correctly to the mass value predicted on the basis of genomic sequencing, then it can be considered unmodified. However, if the peptide does not match the database, the type and site of the specific modification can be accessed directly from the difference between the observed and expected fragment masses. Therefore MS has been successfully used for their identification. On the other hand, in many cases, phosphorylation is a dynamic process with complex kinetics involving several amino acids within a single protein. Quantitation of changes in site-specific modifications in proteins (e.g., phosphorylation) has been an even greater analytical challenge than the determination of changes in gross protein level. A stable isotope labeling approach combined with a modification specific enrichment technique, however, has enabled to address this issue. Some of the *in vitro* chemical labeling procedures are not suitable for the quantitation of phosphorylation because they isolate specific peptides (e.g. the ICAT procedure isolates cysteine-containing peptides). On the contrary, a procedure such as *in vivo* labeling is suitable for this kind of study and it has been used for the identification and quantification of several sites of the protein (Oda et al, 1999). Labeled synthetic standards were also used for the quantification of several phosphorylation sites (Ruse et al, 2002; Stemmann et al, 2001). Unfortunately, the requirement of having to synthesize a labeled peptide for each site investigated has limited this approach to a few focused studies.

To investigate the phosphoproteome, several enrichment procedures have been developed. In some cases, these procedures involved specific chemical reaction methods for the phosphate group Oda et al, 2001; Adamczyk et al, 2001; Zhou et al, 2001; Goshe et al, 2001; Qian et al, 2003), metal affinity column (Kokubu et al, 2005; Ndassa et al, 2006), or antibodies (Lim et al, 2003). *Figure 6* describes the general strategies for the quantitation of phosphorylation sites. A major difference between the various strategies is that when using an enrichment method, only changes in phosphorylation can be quantified and the information about changes in gene expression are lost. These changes could be quantified only when the total protein (unmodified + modified (e.g. phosphorylated)) is kept in consideration. On the other hand, an enrichment procedure is necessary for obtaining a better coverage of the phosphoproteome.

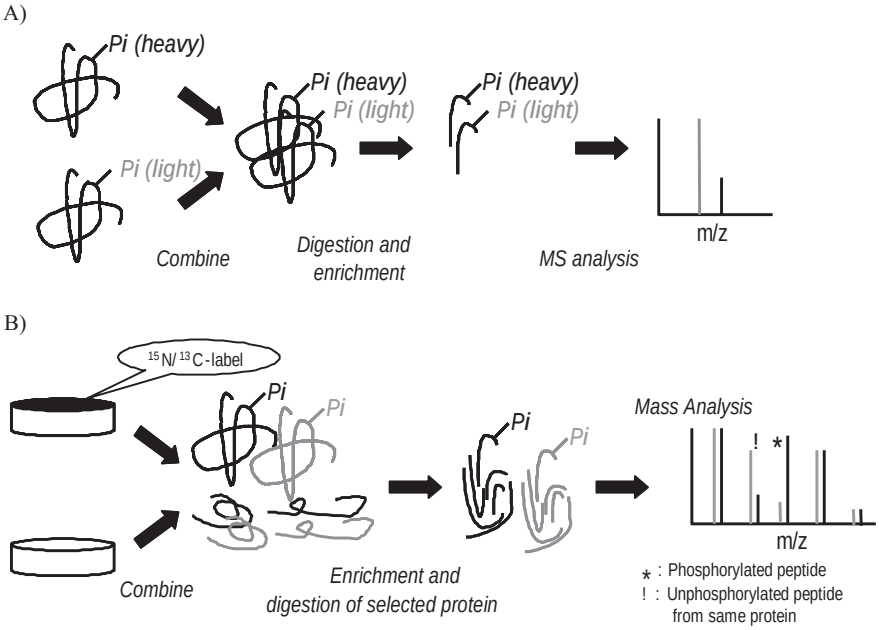

Figure 6. Schematic representations of the quantitative phosphoptroteomics. (A) In some strategies the phosphopeptides are labeled and these labels are used for affinity purification. (B) In other cases, the in vivo labeling procedure is used and all peptides in digest are labeled. With the enrichment procedure, only phosphopeptides are detected and quantified (upper chart), while the in vivo labeling approach has the potential of quantifying all peptides in the digest (lower chart).

Label-free quantitative proteomics

Although stable isotope labeled approaches are reliable quantitation methods and have a long history which includes small molecule analysis, the stable isotopes are expensive, require an extra step to introduce isotopes into samples, and labeling approaches need pairwise- comparisons between samples which prevents retrospective comparisons and complicate the process in larger studies.

Current LC-MS/MS analyses can generate a lot of information about identified proteins, such as the probability score (Allet et al, 2004), the number of identified peptides per protein, ion counts of identified peptides, and LC retention time. Some of these qualitative parameters can be considered as indicators for protein abundance in the analyzed sample. Amongst them, the integrated ion counts of the peptides identifying each protein would be the most direct parameter to describe the abundance. The ion counts have been used to compare protein expression in different states (*Figure 7A*) (Lasonder et al, 2002; Wiener at al, 2004) because mass spectral peak intensities of peptide ions correlate well with protein abundances in complex samples. Linear responses and reproducible results have obtained by several groups (Bondarenko et al, 2002; Chelius and Bondarenko, 2002; Wang et al, 2003; Higgs et al, 2005; Li et al, 2005), basically using known amounts of standard proteins spiked into plasma/

serum. In these test studies, the sample processes before LC/MS analysis were relatively simple and straightforward, and therefore any variations due to sample preparation were minimized. Also, MS is not as versatile as an absorbance detector due to background and ionization suppression effects. Practically, many of peptides peaks on LC/MS spectra are overlapped in a highly complex sample, and low signal intensities for low abundant proteins make it difficult to extract true ion intensities, especially ion-trap type MS, which allows high quality MS/MS data from low intensities of precursor ions.

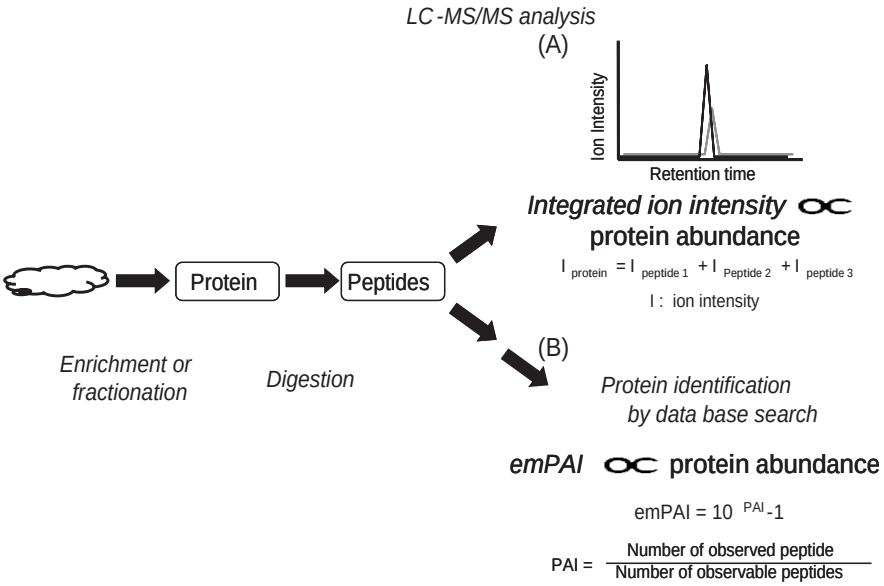

Figure 7. Strategy for quantitation by label-free quantitative method, (A) integrated ion intensity approach and (B) spectral count approach.

Therefore, it is necessary to normalize these parameters to obtain at least approximate quantitative information. The first approach to achieve normalization, to our knowledge, was to use the number of peptides per protein normalized by the theoretical number of peptides - the so-called protein abundance index (PAI) (Lamond and Mann, 2002). PAI is superior to the un-normalized approach since it takes account of the fact that for the same number of molecules, larger proteins and proteins with many peptides in the preferred mass range for MS will generate more observed peptides. Then the exponential form of PAI minus one (emPAI) is roughly proportional to protein absolute abundance (*Figure 7B*) (Ishihama et al, 2005a). Because emPAI is easily calculated from the output information of database search engines such as Mascot, it is possible to apply this approach to previously measured or published datasets to add quantitative information without any additional steps. emPAI can also be used for relative quantitation. This is particularly relevant for cases where isotope-based approaches cannot be applied - where quantitative changes are too large for accurate measurements of ratios, or probably because isotope labeling is not possible, or because sensitivity constraints do not allow chemical labeling techniques. In such

cases, emPAI values of proteins in one sample can be compared with those in another sample, and the outliers from the emPAI correlation between two samples can be determined as increasing or decreasing proteins. This emPAI approach has been applied to multidimensional separation-MS/MS to extend the coverage of proteins. Independently, a similar index for label-free method termed spectral counting (comparison of the number of MS/MS spectra assigned to each protein) has been also introduced for quantitative proteomics (Old et al, 2005). An advantage of spectral counting is that, in principle, relative abundances of different proteins can be measured. Thus, it has been possible to show significant correlations between spectral counts and independent estimates of protein copy number in a sample.

However, since emPAI and spectral counting are highly dependent on peptide identification, they have a major disadvantage. Due to limitations in peptide sampling in MS (data-dependent MS/MS scan), the absence of a protein in the negative control sample hit list is not sufficient proof to state that a protein identified in the experimental sample has beeen isolated in a specific fashion. Even if the same complex sample was analyzed twice by MS, the overlap between the peptides sequenced is never 100%, as long as the number of peptides present in the sample exceeds the number of sequencing cycles available during MS/MS analysis. The apparent absence of a peptide/protein in the negative control may be solely due to the peptide under sampling. Therefore lists of low number of emPAI/spectral counting are less reliable as a quantitative data. Indeed 60-70% of proteins in a highly complex sample are identified based on one to three peptides from a corresponding protein, therefore counting based quantitation seem to be unsuitable for large scale comprehensive proteome analysis.

Quantitative approaches to protein interactions

Proteins often interact with each other or with non-protein molecules such as DNA to form transient or stable complexes that mediate biological activities. Identification of direct specific components in a protein complex, such as protein-protein, DNA/RNA–protein, or chemical–protein, provides an important clue to the functional units of cellular molecular machinery. For this reason, the study of protein interactions has emerged as a valuable method for finding novel components of for example signaling pathways. The yeast two-hybrid screening (Fields and Song, 1989) is a genetic selection procedure designed to detect binary interactions. Some negative points of this procedure are that the approach only records interactions between pairs of genes, it misses interactions stabilized by more than two partners, and does not necessarily reflect the physiological environment.

Another approach involves affinity purification (Oda et al, 2003, Guerrero et al, 2006; Ranish et al, 2003; Blagoev et al, 2003; Kerner et al, 2005), where target proteins are enriched or isolated by affinity techniques using a bait-conjugated matrix. A tag is introduced onto a target protein then the tagged protein is purified from a whole cell lysate with associated proteins, which are subsequently characterized by MS. Affinity purification appears to be the most efficient and gentle discriminatory separation technique for the retrieval of protein complexes. In particular, well-characterized small affinity tags such as FLAG, His6, or glutathione S-transferase have been widely used in affinity purification of recombinant proteins or protein complexes.

MS can rapidly and reliably identify the binding partners in a protein complex, and therefore the current limiting step appears to be the finding of specific proteins, rather than protein identification. An affinity approach may fail to purify proteins of less abundance due to non-specific contamination during purification thus it remains difficult to distinguish specific from nonspecific interactions. The use of more stringent washes to reduce contaminating proteins may also affect the binding of weaker interactors. For example, regulatory subunits, which are often not tightly associated with the enzymes they regulate, may be lost during the purification procedure. Each protein has unique properties, which can be exploited during its purification. However it is difficult to design a procedure valid for all cases. Thus, comprehensive analysis of macromolecular complexes has been hindered by the lack of a general and efficient purification strategy. In biochemical studies to identify drug-binding proteins (drug target proteins), compound-conjugated affinity matrix reagents have played an important role; however, the affinity and specificity of synthetic small molecules for their protein targets are rather low in many cases. Nonspecific interactions often lead to difficulty in specifying the primary binding partners of a synthetic compound. Stable isotope labeling strategies (Oda et al, 2003, Guerrero et al, 2006; Ranish et al, 2003; Blagoev et al, 2003; Kerner et al, 2005) have proven particularly advantageous for the discrimination of proteins specifically associated with the target population from non-specifically co-purified contaminants. Functionally important specific interactions can be picked out of the background binding through the detection of isotope ratios on MS, avoiding a trade-off between false-positive binding and the ability to detect weak components.

The principal strategy is as follows: To distinguish specific complex components from co-purifying proteins, control purification (negative complex, i.e., point-mutated bait) is carried out in which the complex of interest is not enriched. To discriminate non-specific binding partners of a complex isolated from cells, the samples are prepared identically. Stable isotope labeling is performed metabolically at the cell culture level, or by chemically attaching isotopically heavy or light tags after affinity purification. After combining two binding samples, followed by proteolysis, the peak ratios are calculated to determine the specificity of each component of a complex. In addition to defining static interaction networks, both types of approaches are compatible with the analysis of the dynamics of protein complex formation.

Concluding remarks

Although numerous reports regarding MS-based quantitative proteomics have been published, there are still many areas that require further development (*Table 1*). The *in vivo* labeling procedures are potentially the most accurate because the label is introduced early in the process and making it applicable to a post-translational modification analysis combined with specific enrichment techniques, and facilitates protein interaction studies. However, there are some obvious limitations, such as cost, the existence of inaccessible samples for labeling, large number of sample sets and the dynamic range. Therefore choosing the right method for quantitation is the real key to success of a particular experiment. As protocols and methods for new quantitative high-throughput proteomics are constantly being improved, proteomics is likely to continue to reveal a deep insight into human disease and the mechanistic regulation of molecular interaction-networks.

Table 1. Summary of MS-based quantitation proteomics strategies.

Methods	Pros	Cons
Absolute	**Absolute abundance can be obtained**	
Spiked peptides	Common approach for drug metabolism study Applicable to protein analysis	Laborious and costly step in preparation of internal standards
In vivo labeling	**Simple, Accurate**	
Cell culture	Applicable for post-translational modification and protein interaction	Limited to culture cell sample
Diet enrich	Applicable for tissue sample	Time consuming, Costly Sometime sample is not labeled completely
CDIT	Applicable for tissue sample Absolute amount can be obtained	Difficult to apply biological fluid Decrease accuracy when corresponding protein is absent in CDIT cells
In vitro labeling	**Applicable for tissue sample**	**Increase the possibility for false quantitation since all the steps including digestion is done separately**
ICAT	Reduce complexity by enrichment process	Costly. Limited analysis of proteins containing cysteine
$^{16}O/^{18}O$ with Trypsin	Easy derivatization	Complex analysis, Need high resolution MS Possible incomplete incorporation of the label
iTRAQ	Multiplexing four samples High peptide coverage Easy analysis	Costly
Label-free	**Simple, low cost** **Allow large number of sample**	**Cannot normalize variation during sample preparation**
Integrated ion count		May increase false positive. Necessary to normalize
emPAI / spectral count	Easy to calculate	Not suitable for highly complex samples

Acknowledgement

This work was supported by funds from Japan Science and Technology Agency (JST).

References

Adamczyk, M., Gebler, J.C. and Wu, J. (2001) Selective analysis of phosphopeptides within a protein mixture by chemical modification, reversible biotinylation and mass spectrometry. *Rapid Commun. Mass Spectrom*. **15**, 1481–1488.

AEBERSOLD R, AND MANN M. (2003) Mass spectrometry-based proteomics. *Nature*, **422**, 198-207.

ALLET, N., BARRILLAT, N., BAUSSANT, T,. BOITEAU, C., BOTTI, P., BOUGUELERET, L., BUDIN, N., CANET, D., CARRAUD, S., CHIAPPE, D., CHRISTMANN, N., COLINGE, J., CUSIN, I., DAFFLON, N., DEPRESLE, B., FASSO, I., FRAUCHIGER, P., GAERTNER, H., GLEIZES, A., GONZALEZ-COUTO, E., JEANDENANS, C., KARMIME, A., KOWALL, T., LAGACHE, S., MAHE, E., MASSELOT, A., MATTOU, H., MONIATTE, M., NIKNEJAD, A., PAOLINI, M., PERRET, F., PINAUD, N., RANNO, F., RAIMONDI, S., REFFAS, S., REGAMEY, P.O., REY, P.A., RODRIGUEZ-TOME, P., ROSE, K., ROSSELLAT, G., SAUDRAIS, C., SCHMIDT C., VILLAIN, M. AND ZWAHLEN, C. (2004) In vitro and in silico processes to identify differentially expressed proteins. *Proteomics* **4**, 2333-2351

ANDERSEN, J.S. AND MANN, M. (2006) Organellar proteomics: turning inventories into insights. *EMBO Rep*. **7**, 874-879.

BLAGOEV, B., KRATCHMAROVA, I., ONG, S.E., NIELSEN, M., FOSTER, L.J. AND MANN, M. (2003) A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nature Biotechnol.* **21**, 315-318.

BONDARENKO, P.V., CHELIUS, D. AND SHALER, T.A. (2002) Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. *Anal. Chem*. **74**, 4741-4749.

BRUNET, S., THIBAULT, P., GAGNON, E., KEARNEY, P., BERGERON, J.J. AND DESJARDINS, M. (2003) Organelle proteomics: looking at less to see more, *Trends Cell Biol.* **13**, 629-638.

CHELIUS, D. AND BONDARENKO, P.V. (2002) Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J. Proteome Res*. **1**, 317-323.

DESIDERIO, D.M. (1992) Quantitative analytical mass spectrometry of endogenous neuropeptides in human pituitaries. *Life Science* **51**, 169–176.

DOMON, B. AND AEBERSOLD, R. (2006) Mass spectrometry and protein analysis. *Science* **312**, 212-217.

ELIAS, J.E., HAAS, W., FAHERTY, B.K., GYGI, S.P. (2005) Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nature Methods* **2**, 667-675

FIELDS, S. AND SONG, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature*, **340**, 245-246

GORG, A., WEISS, W. AND DUNN, M.J. (2004) Current two-dimensional electrophoresis technology for proteomics. Proteomics, 4, 3665-3685.

GOSHE, M.B., CONRADS, T.P., PANISKO, E.A., ANGELL, N.H., VEENSTRA, T.D. AND SMITH, R.D. (2001) Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. *Anal Chem.* **73**, 2578-2586.

GUERRERO, C., TAGWERKER, C., KAISER, P. AND HUANG, L. (2006) An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26S proteasome-interacting network, *Mol. Cell Proteomics* **5**, 366–378.

GYGI, S.P., RIST, B., GERBER, S.A., TURECEK, F., GELB, M.H. AND AEBERSOLD, R.(1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnol.* **17**, 994-999

HANSEN, K.C., SCHMITT-ULMS, G., CHALKLEY, R.J., HIRSCH, J., BALDWIN, M.A. AND BURLINGAME, A.L. (2003) Mass spectrometric analysis of protein mixtures at low levels using cleavable 13C-isotope-coded affinity tag and multidimensional chromatography. *Mol Cell Proteomic* **2**, 299-314

HIGGS, R.E., KNIERMAN, M.D., GELFANOVA, V., BUTLER. J.P. AND HALE, J.E. (2005) Comprehensive label-free method for the relative quantification of proteins from biological samples. *J. Proteome Res.* **4**, 1442-1450.

ISHIHAMA, Y., ODA, Y., TABATA, T., SATO, T., NAGASU, T., RAPPSILBER, J. AND MANN, M. (2005a) 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.

ISHIHAMA, Y., SATO, T., TABATA, T., MIYAMOTO, N., SAGANE, K., NAGASU, T. AND ODA, Y. (2005b) Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. *Nature Biotechnol.* **23,** 617-621.

KERNER, M.J., NAYLOR, D.J., ISHIHAMA, Y., MAIER, T., CHANG, H.C., STINES, A.P., GEORGOPOULOS, C., FRISHMAN, D., HAYER-HARTL, M., MANN, M. AND HARTL, F.U. (2005) Proteome-wide Analysis of Chaperonin-Dependent Protein Folding in Escherichia coli, *Cell* **122**, 209-220.

KIPPEN, A.D., CERINI, F., VADAS, L., STOCKLIN, R., VU, L., OFFORD, R.E. AND ROSE, K. (1997) Development of an isotope dilution assay for precise determination of insulin, C-peptide, and proinsulin levels in non-diabetic and type II diabetic individuals with comparison to immunoassay. *J. Biol. Chem.* **272**, 12513–12522.

KOKUBU, M., ISHIHAMA, Y., SATO, T., NAGASU, T., AND ODA, Y. (2005) Specificity of immobilized metal affinity-based IMAC/C18 tip enrichment of phosphopeptides for protein phosphorylation analysis. *Anal Chem.***77**, 5144-5154.

KUSMIERZ, J.J., SUMRADA, R. AND DESIDERIO, D.M. (1990) Fast atom bombardment mass spectrometric quantitative analysis of methionine-enkephalin in human pituitary tissues. *Anal. Chem.* **62**, 2395–2400

LAMOND, A.I. AND MANN, M. (2002) Large-scale proteomic analysis of the human spliceosome. *Genome Res*. **12**, 1231-1245.

LASONDER, E., ISHIHAMA, Y., ANDERSEN, J.S., VERMUNT, A.M., PAIN, A., SAUERWEIN, R.W., ELING, W.M., HALL, N., WATERS, A.P., STUNNENBERG, H.G. AND MANN, M. (2002) Analysis of the Plasmodium falciparum proteome by high-accuracy mass spectrometry. *Nature* **419**, 537-542.

LI, X.J., YI, E.C., KEMP, C.J., ZHANG, H. AND AEBERSOLD. R. (2005) A software suite for the generation and comparison of peptide arrays from sets of data collected by liquid chromatography-mass spectrometry *Mol Cell Proteomics* **4**, 1328-1340..

LIM, Y.P., DIONG, L.S., QI, R., DRUKER, B.J. AND EPSTEIN, R.J. (2003) Phosphoproteomic fingerprinting of epidermal growth factor signaling and anticancer drug action in human tumor cells. *Mol. Cancer Ther*. **2**, 1369-1377.

MIRGORODSKAYA, O.A., KOZMIN, Y.P., TITOV, M.I., KORNER, R., SONKSEN, C.P. AND ROEPSTORFF, P. (2000) 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.

MORANDELL, S., STASYK, T., GROSSTESSNER-HAIN, K., ROITINGER, E., MECHTLER, K., BONN, G.K. AND HUBER, L.A. (2006) Phosphoproteomics strategies for the functional analysis of signal transduction. *Proteomics* **6**, 4047-4056.

NDASSA, Y.M., ORSI, C., MARTO, J.A., CHEN, S. AND ROSS, M.M. (2006) Improved immobilized metal affinity chromatography for large-scale phosphoproteomics applications. *J. Proteome Res*. **5**, 2789-2799.

ODA, Y., HUANG, K., CROSS, F.R., COWBURN, D. AND CHAIT, B.T. (1999) Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. USA*. **96**, 6591-6596.

ODA, Y., NAGASU T. AND CHAIT, B.T. (2001) Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nature. Biotechnol.* **19**, 379–382.

ODA, Y., OWA, T., SATO, T., BOUCHER, B., DANIELS, S., YAMANAKA, H., SHINOHARA, Y., YOKOI, A., KUROMITSU, J. AND NAGASU, T. (2003) Quantitative chemical proteomics for identifying candidate drug targets 75, 2159-2165.

OLD, W.M., MEYER-ARENDT, K., AVELINE-WOLF, L., PIERCE, K.G., MENDOZA, A., SEVINSKY, J.R., RESING, K.A. AND AHN, N.G. (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics.*Mol. Cell Proteomics* **4**, 487-502.

ONG, S.E., BLAGOEV, B., KRATCHMAROVA, I., KRISTENSEN, D.B., STEEN, H., PANDEY, A. AND MANN, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell Proteomics* **1**, 376-86

QIAN, W.J., GOSHE, M.B., CAMP, D.G. 2ND, YU, L.R., TANG, K. AND SMITH, R.D. (2003) Phosphoprotein isotope-coded solid-phase tag approach for enrichment and quantitative analysis of phosphopeptides from complex mixtures. *Anal Chem.* **75**, 5441-5450.

RANISH, J.A., YI, E.C., LESLIE, D.M., PURVINE, S.O., GOODLETT, D.R., ENG, J. AND AEBERSOLD, R. (2003) The study of macromolecular complexes by quantitative proteomics. *Nature Genetics* **33,** 349-355.

ROSS, P.L., HUANG, Y.N., MARCHESE, J.N., WILLIAMSON, B., PARKER, K., HATTAN, S., KHAINOVSKI, N., PILLAI, S., DEY, S., DANIELS, S., PURKAYASTHA, S., JUHASZ, P., MARTIN, S., BARTLET-JONES, M., HE, F., JACOBSON, A. AND PAPPIN, D.J. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* **3**, 1154-1169.

RUSE, C.I., WILLARD, B., JIN, J.P., HAAS, T., KINTER, M. AND BOND, M. (2002) Quantitative dynamics of site-specific protein phosphorylation determined using liquid chromatography electrospray ionization mass spectrometry. *Anal. Chem*. **74**, 1658–1664.

RUSH, J., MORITZ, A., LEE, K.A., GUO, A., GOSS, V.L., SPEK, E.J., ZHANG, H., ZHA, X.M., POLAKIEWICZ, R.D. AND COMB, M.J. (2005) Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nature Biotechnol*. **23**, 94-101.

STEMMANN, O., ZOU, H., GERBER, S.A., GYGI, S.P. AND KIRSCHNER, M.W. (2001) Dual inhibition of sister chromatid separation at metaphase. *Cell* **107**, 715–726.

WANG, R., SWEENEY, D., GANDY, S.E. AND SISODIA, S.S. (1996) Detection and quantification of amyloid beta protein and variants by immunoprecipitation-mass spectrometry. *J. Biol. Chem*. **271,** 31894–31902.

WANG, W., ZHOU, H., LIN, H., ROY, S., SHALER, T.A., HILL, L.R., NORTON, S., KUMAR, P., ANDERLE, M. AND BECKER, C.H. (2003) Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal. Chem.* **75**, 4818-4826.

WANG, Y.K., MA, Z., QUINN, D.F. AND FU, E.W. (2001) Inverse 18O labeling mass spectrometry for the rapid identification of marker/target proteins. *Anal Chem*. **73**, 3742-50.

WIENER, M.C., SACHS, J.R., DEYANOVA, E.G. AND YATES, N.A. (2004) Differential mass spectrometry: a label-free LC-MS method for finding significant differences in complex peptide and protein mixtures. *Anal. Chem*. **76**, 6085-6096.

WU, C.C., MACCOSS, M.J., HOWELL, K.E., MATTHEWS, D.E. AND YATES, J.R. 3RD. (2004) Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal Chem.* **76**, 4951-4959.

YAO, X., FREAS, A., RAMIREZ, J., DEMIREV, P.A. AND FENSELAU, C. (2001) Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal Chem*. **73**, 2836-2842. Erratum in: *Anal Chem*. (2004) **76**, 2675.

YI, E.C., LI, X.J., COOKE, K., LEE, H., RAUGHT, B., PAGE, A., ANELIUNAS, V., HIETER, P., GOODLETT, D.R. AND AEBERSOLD, R.(2003) Increased quantitative proteome coverage with (13)C/(12)C-based, acid-cleavable isotope-coded affinity tag reagent and modified data acquisition scheme. *Proteomics* **5**, 380-387

ZHOU, H., WATTS, J.D. AND AEBERSOLD, R. (2001) A systematic approach to the analysis of protein phosphorylation. *Nature Biotechnol.* **19** , 375–378.