Mass spectrometry-based quantitative proteomic profiling

Wei Yan and Sharon S. Chen

Date received: 2nd December 2004

Wei Yan
is a research scientist at the Institute for Systems Biology and in the Department of Pathology at the University of Washington.

Sharon S. Chen
is a graduate student in the Department of Bioengineering at the University of Washington.

Keywords: *proteomics, stable isotope labelling, mass spectrometry, quantitative protein profiling, two-dimensional electrophoresis, liquid chromatography*

Wei Yan,
Institute for Systems Biology,
1441 N. 34th Street ,
Seattle, WA 98103-8904, USA

E-mail: wyan@systemsbiology.org

Sharon S. Chen,
University of Washington,
Department of Bioengineering,
Box #357962,
Seattle, WA 98115, USA

E-mail: chenss@u.washington.edu

Abstract

Quantitative proteomics involves the identification and quantitation of protein components in various biological systems. Stable isotope labelling technology, by both metabolic and chemical methods, has been the most commonly used approach for global proteome-wide profiling. Recently, its capability has been extended from labelled pairs to multiple labels, allowing for the simultaneous quantification of multiplex samples. The ion intensity-based quantitative approach has progressively gained more popularity as mass spectrometry performance has improved significantly. Although some success has been reported, it remains difficult comprehensively to characterise the global proteome, due to its enormous complexity and dynamic range. The use of sub-proteome fractionation techniques permits a simplification of the proteome, and provides a practical step towards the ultimate dissection of the entire proteome. Further development of the technology for targeting sub-proteomes on a functional basis, such as selecting proteins with differential expression profiles from mass spectrometric analyses, for further mass spectrometric sequencing in an intelligent manner, is expected in the near future.

INTRODUCTION

Aided by the proliferation in sequenced genomes, including the completion of the rough draft sequence of the human genome, proteomics has emerged as a field for studying global gene expression profiles at the protein level. In general, proteomics involves the identification of protein components and the measurement of protein abundance in biological systems such as cultured cells or tissue samples. While most of the initial efforts in proteomics have focused on protein identification, recent mass spectrometry (MS)-based technology developments have provided useful platforms for the study of quantitative changes in protein components within the cell. Quantitative analysis of global protein levels, termed 'quantitative proteomics', is important for the system-based understanding of the molecular function of each protein component and is expected to provide insights into molecular mechanisms of various biological processes and systems.

Currently, there are several widely used methods to generate global quantitative protein profiles, including two-dimensional gel electrophoresis (2DE) followed by MS analysis, stable isotope labelling-based quantitation, MS signal intensity-based quantitation and protein array-based quantitation. The protein array method generally spots a large number of protein-interacting elements (for example, antibodies) on solid supports in a distribution-regulated manner and is designed for high-throughput analysis. As a successor to DNA microarray technology, it has been successfully applied to detect proteins interacting with individual target protein elements on the array, both qualitatively and quantitatively. The utility of this approach, however, is limited by difficulties in handling and immobilising large numbers of proteins in the array and maintaining the proteins in an active state. Since MS is not a major tool for this method, it will not be discussed in depth in this paper, and readers are referred to other recent reviews.[1,2]

Article number = K123

In this paper, we will focus on recent advances in quantitative proteomics using MS technology. Currently, MS has been overwhelmingly applied as the technology base for proteomics analysis.[3,4] Proteins have been identified and quantified by characterisation of their derived peptides (from proteolysis) using either electrospray ionisation (ESI) or matrix-assisted laser desorption/ionisation (MALDI)-based MS analysis. Several major MS-based methodologies (Figure 1) towards quantitations of the proteome, including 2DE, stable isotope labelling and quantitation using the absolute peptide ion intensity, will be discussed. Particular focus will be on those techniques using liquid chromatography followed by tandem mass spectrometry (LC-MS/MS).

QUANTITATION VIA 2DE AND MS

An older and still commonly used approach for proteomic quantitation is the combination of 2DE and MS. In this procedure, proteins are separated by 2DE and quantified based on the intensity of the protein spots of individual gels. Following proteolysis, the derived peptides are extracted from the gel spots and subjected to MS analysis for identification. Since its introduction in the mid-1990s,[1] this procedure has been routinely used for large-scale quantitative protein profiling of protein complex mixtures. A potential problem with 2DE-based quantitation is that the gel spots corresponding between different experiments can be difficult to measure reproducibly. This has been significantly improved by the recent development of a two-colour fluorescence labelling system which allows parallel comparison of two protein samples within the same gel.[5] Nevertheless, global proteomics analysis by 2DE/MS is still limited by the bias of the method against certain protein groups, such as membrane proteins, excessively large or small proteins, and very acidic or basic proteins. Moreover, the system is limited in its ability to resolve proteins in complex mixtures. Some proteins co-migrate in 2D gels. The presence of more than one protein in a single spot can produce ambiguous

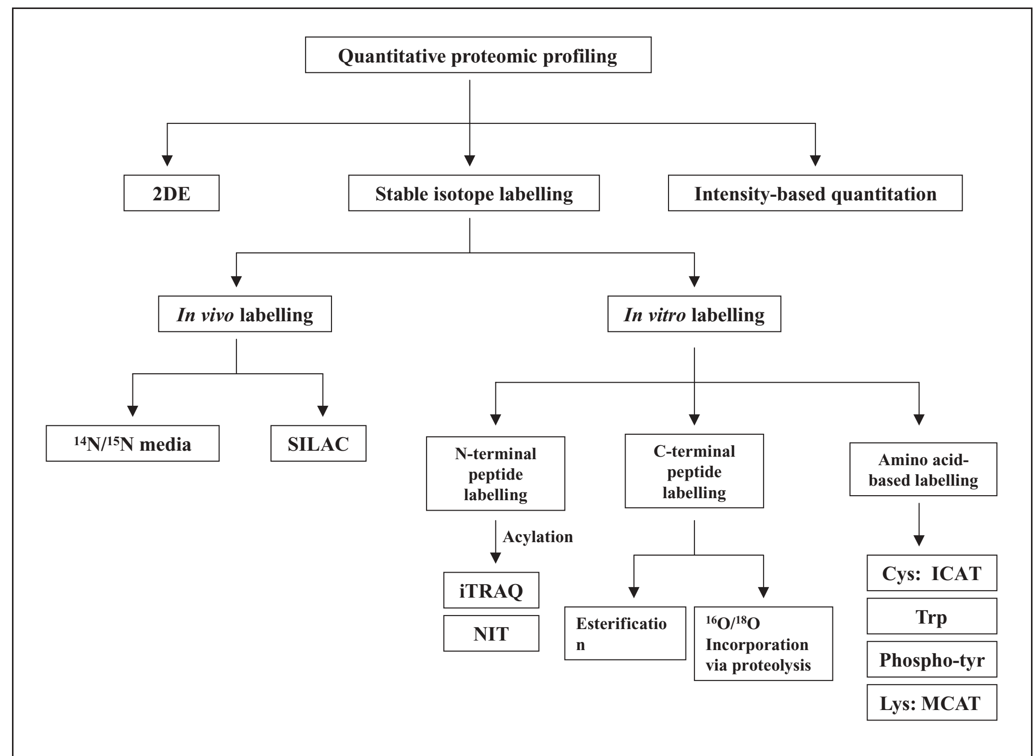

Figure 1: Strategies for quantitative proteomic profiling. 2DE: two-dimensional gel electrophoresis; SILAC: stable isotope labelling with amino acids in cell culture; iTRAQ: isobaric tags for relative and absolute quantitation; ICAT: isotope-coded affinity tags; NIT: N-terminal isotope-encoded tagging; MCAT: mass-coded abundance tagging.

quantitative results.[6] For a more detailed discussion of the 2DE/MS technology, we refer the reader to several recent reviews on this topic.[7–9]

STABLE ISOTOPE LABELLING FOR QUANTITATION

The absolute signal intensity of a peptide ion measured in an MS run does not always reflect the abundance of peptide present in the analysed sample. This is due to variability in ionisation with ESI or MALDI peptide ionisation, and the influence of other ions in the sample on the measured ion intensity of a specific peptide ion. A reliable internal standard is often required to normalise quantitative variations among different MS measurements. An ideal internal standard should be chemically and physically as similar as possible to the analysed peptide. Thus, the best internal standard for a peptide is a peptide of identical sequence but labelled with different stable isotopes. This has led to the development of several quantitative proteomics technologies via incorporation of stable isotope tags *in vivo* or *in vitro*.

In vivo labelling via metabolic incorporation

The *in vivo* labelling approach involves metabolically incorporating stable isotopes into proteins of cells grown in special media containing these isotopes. Early efforts have been successfully reported in cell cultures of yeast and bacteria.[9–11] Cells were grown in two media containing either ^{14}N (light) or ^{15}N (heavy) isotopes, respectively. Following isotopic labelling, the two cell cultures were combined, processed and analysed by MS. Incorporating the ^{15}N isotope into the proteins of the cells grown in the heavy media, and the ^{14}N isotope into the proteins of the cells grown in the normal media, caused a mass shift in the resultant peptides from the two types of media. This mass shift is readily observed during MS analysis. The relative quantity of the proteins in the two samples was determined from the ratio of the measured ion intensities of the isotopically labelled peptide pairs. Since the method labels both the backbone and side-chain nitrogen atoms, however, the mass shift cannot be predicted for peptides whose sequence is unknown. This makes it difficult to quantify and identify peptides in highly complex samples such as cell lysates unless a high accuracy mass spectrometer is used and/or other fractionation procedures, such as cysteine capture, are applied to reduce the sample complexity.[12]

In an alternative approach to *in vivo* stable isotope labelling, termed 'stable isotope labelling with amino acids in cell culture (SILAC)', proteins are labelled by growing cells in media containing isotpically labelled amino acids including ^{2}H-leucine,[13] ^{13}C-lysine,[14] ^{13}C-tyrosine,[15] ^{13}C-arginine[16] and $^{13}C/^{15}N$-arginine.[17] This approach has recently gained popularity because of the predictability of the mass shift which can be analysed by conventional mass spectrometers (see review by Ong *et al.*[16] for more detail).

In general, stable isotope labelling *in vivo* has proven to be an effective method of quantitative proteomics analysis. One advantage is that the stable isotopic tags are incorporated into the early stages of sample preparation and thus reduce variation between samples, yielding highly accurate quantification. A disadvantage, however, is that this method is not practical for analysing biological samples that cannot be grown in culture, such as tissues or body fluids. In addition, it requires a relatively long labelling incubation time in cell culture (ie five population doublings) to satisfactorily incorporate the isotopic labels.[16]

In vitro labelling via chemical reactions

The *in vitro* labelling technology involves incorporation of the stable isotopic tags onto selective sites on peptides via *in vitro* chemical reactions. It includes isotopic

labelling of target peptides at their amino- (N-) or carboxyl- (C-) terminal or on specific amino acid residues, such as cysteine, lysine, tyrosine, etc.

Cysteine-containing peptide labelling

In vitro stable isotope labelling was first introduced by Gygi *et al.*[18] as a method termed 'isotope-coded affinity tags (ICAT)'. The ICAT reagent consisted of a biotin affinity tag for selective purification, a linker that incorporated stable isotopes (ie ^{1}H or ^{2}H), and an iodoacetamide reactive group that specifically reacted with cysteinyl thiols (Figure 2). Proteins from two samples were labelled at their cysteine residues with either isotopically light (^{1}H) or heavy (^{2}H) ICAT reagents, respectively. The light- and heavy- labelled samples were then combined, proteolysed to peptides, fractionated by multi-dimensional chromatography and quantitatively analysed by MS. Ion intensity ratios between the light and heavy forms of a specific peptide (with a mass shift of 8 Da per labelled cysteine residue) indicated their relative abundance.

Since its introduction in 1999, the ICAT method has been applied to the comparison of the relative quantity of proteins between samples (ie control versus perturbation) in a variety of species including bacteria,[19] yeast,[20] mouse[21] and humans.[22,23] In addition to the cultured cells, studies on tissues[24,25] and body fluids[26] have also been reported. Recently, an improved ICAT reagent has been made commercially available (http://www.appliedbiosystems.com). This second-generation ICAT reagent contains an acid-cleavable linker that allows for removal of the large affinity tag prior to MS. This improves peptide identification by removing affinity tag-derived fragment ions from MS/MS spectra. In addition, the second-

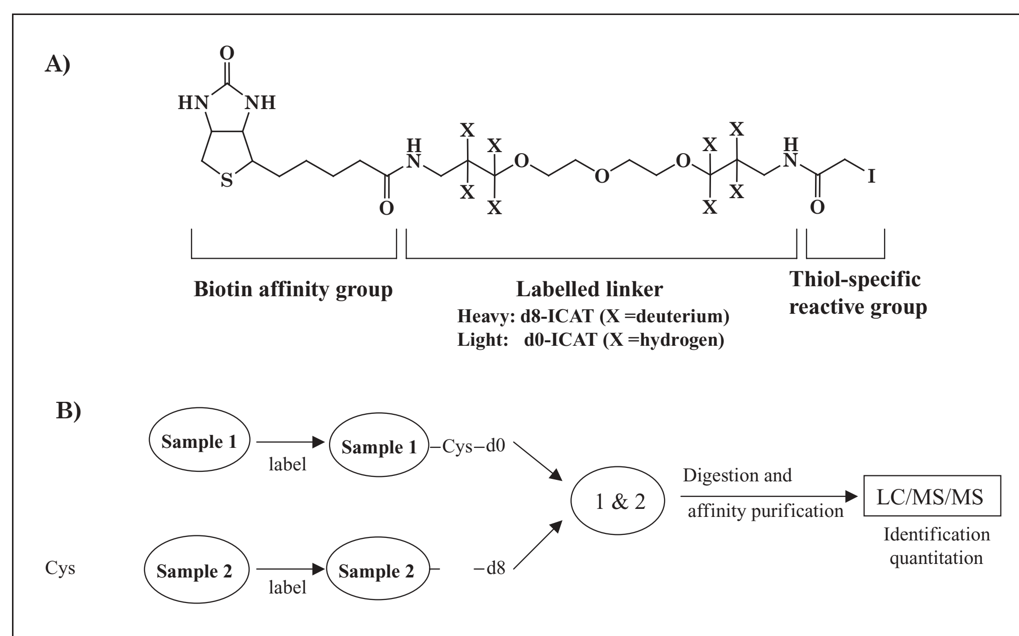

Figure 2: (A) Structure of the isotope-coded affinity tag (ICAT), which consists of a biotin affinity tag for selective purification, a linker that incorporates stable isotopes, and an iodoacetamide reactive group that specifically reacts with cysteinyl thiols; (B) Strategy for quantitation by ICAT. Two protein mixtures from different samples are labelled with heavy (d8) and light (d0) ICAT reagents, respectively. The labelled mixtures are then combined, digested into peptides using trypsin and subjected to avidin affinity chromatography to enrich the labelled peptides that carry biotin tags. The peptides obtained are then fractionated by liquid chromatography prior to identification and quantitation by mass spectrometry.

generation ICAT incorporates $^{12}C/^{13}C$, instead of the $^{1}H/^{2}H$ incorporated in the original ICAT, to prevent possible chromatography shifts.[27] Recently, a solid-phase variant of the ICAT procedure has been reported to allow for simpler enrichment of target peptides and potential automation.[28] In this method, cysteine-containing peptides were covalently captured and labelled with stable isotope tags on glass beads in one step. The isoptically labelled peptides were released by photo-cleavage of the linker before MS analysis.

N-terminal peptide labelling

The ICAT method selectively purifies the cysteine-containing peptides and thus dramatically reduces sample complexity, allowing for the detection and quantitation of non-abundant proteins. In the human proteome, about 26.6 per cent of the total tryptic peptides contain at least one cysteine residue and cover 96.1 per cent of the human proteome (Human IPI database v2.23).[29] Application of this cysteine-based enrichment theoretically reduces sample complexity by at least fourfold while missing fewer than 4 per cent of the proteins. To obtain complete protein coverage, an ideal target for introducing an isotopic tag is the N- or C- termini of peptides. Incorporation of stable isotopes during acylation of the primary amine group (including N-terminal and ε-amino group of lysine residues) has been successfully applied to quantify proteins in relatively simple complex mixtures.[30–32] An improvement of this N-terminal acylation-based approach, termed N-terminal isotope-encoded tagging (NIT), specifically incorporates stable isotopes at the N-termini of peptides by firstly converting the lysine residue to homoarginine using *O*-methylisourea.[33] Unwanted acylation of the ε-amino group of lysine is thereby eliminated and data analysis simplified. This process also prevents the acylation-mediated loss of the positive charge on the lysine side chain, which could reduce MS-based peptide detection sensitivity.

Recently, an amine group-based isotope labelling methodology, termed isobaric tags for relative and absolute quantitation (iTRAQ), was developed.[34] The iTRAQ reagent consists of a reporter group, a balance group and a peptide reactive group (Figure 3). The peptide-reactive group specifically reacts with primary amine groups of peptides similar to the method described above. The reporter group is a tag with a mass of 114, 115, 116 or 117 Da, depending on differential isotopic combinations of $^{12}C/^{13}C$ and $^{16}O/^{18}O$ in each individual reagent. The balance group ranges in mass from 28 to 31 Da to ensure the combined mass of the reporter and balance groups remains constant (145 Da) for all four reagents. Therefore, peptides labelled with different isotopes are isobaric and are chromatographically indistinguishable, a factor that is important for accurate quantitation. During collision-induced dissociation (CID), the reporter group ions fragment from the backbone peptides, displaying distinct masses of 114 to117 Da. The intensity of these fragments is used for quantitation of the individual representative peptides. Unlike other stable isotope labelling approaches that use MS spectra for quantitation, iTRAQ quantifies the relative peptide abundance from MS/MS spectra. The most significant advantage of this technology is that it allows labelling of up to four different samples within a single experiment. This four-multiplex labelling strategy is useful for quantifying proteins from multiplex samples, such as those in a time course study, replicate measurements of the same sample, or simultaneous comparison of normal, diseased and drug-treated samples. In addition, the isotopically labelled peptides are isobaric and all contribute to one ion species that is observed in the MS and used for CID. This results in increased signal intensity and an increased probability of correct peptide identification, particularly for non-abundant proteins, which are often biologically meaningful. Shortly after its initial introduction at the 52nd The

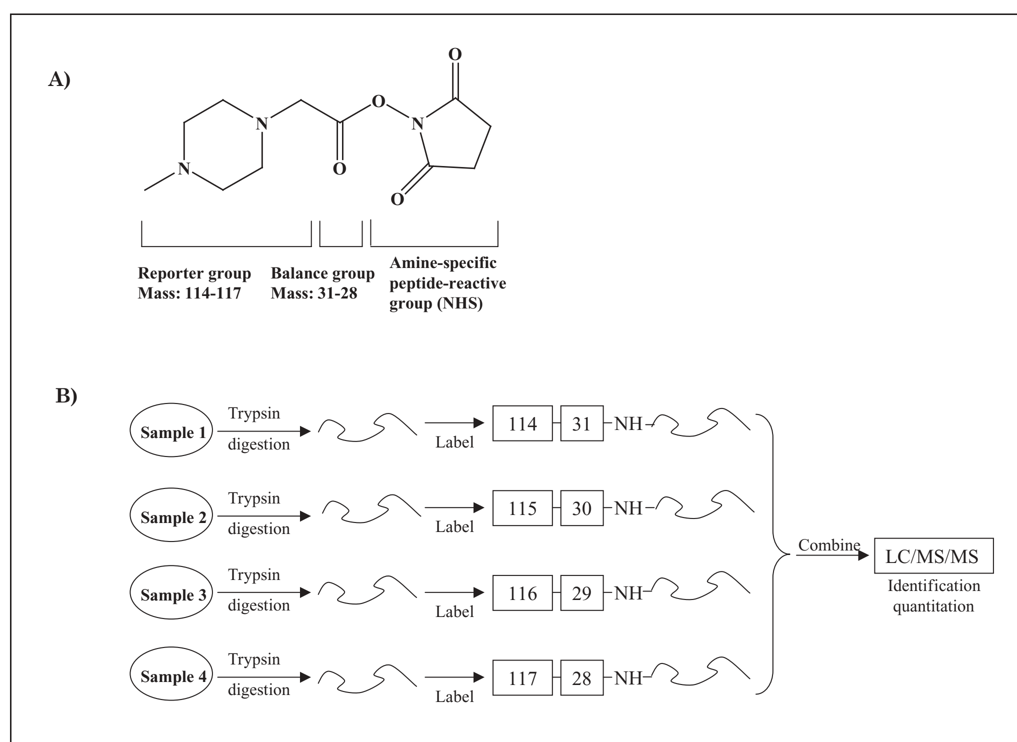

Figure 3: (A) Structure for reagents of isobaric tags for relative and absolute quantitation (iTRAQ), which consists of a reporter group with a mass ranging from 114 to 117 Da, a balance group with a mass ranging from 31 to 28 Da and an amine-specific peptide-reactive group. (B) Strategy for quantitation by iTRAQ. Protein mixtures from up to four different samples are proteolysed by trypsin. The resultant peptides are labelled with individual iTRAQ reagents (114, 115, 116 or 117 Da), respectively. The labelled peptides are combined and analysed by liquid chromatography and mass spectrometry (LC/MS/MS) for identification and quantitation.

American Society for Mass Spectrometry conference in May 2004, iTRAQ has gained much attention and has been successfully applied to the quantification of proteins from yeast[34] and mouse macrophages (Yan and Aebersold, unpublished data). The iTRAQ reagent is commercially available (http://www.appliedbiosystems.com).

Carboxyl terminal peptide labelling

Analogous to labelling the N-terminal amine groups, tryptic peptides have been labelled with stable isotopes at the carboxyl termini. One approach is to introduce stable isotopes via esterification of their carboxyl groups. Goodlett and colleagues reported such a method using $^{1}H/^{2}H$ labelled methanol to convert carboxylic acids into isotopic methyl esters to identify and quantify the parent protein of the corresponding methylated peptides.[35] A disadvantage of this method is that esterification is not specific to the carboxyl terminal of a peptide. Aspartate and glutamate residues can also be labelled by this method, complicating data analysis. Furthermore, the harsh reaction conditions used in this experiment can potentially result in partial deamidation of asparagine and glutamine residues and thus increase sample heterogeneity.

An alternative to chemically introducing stable isotopes specifically to the C-termini of peptides is via biochemical reactions using enzymes such as trypsin. During trypsin proteolysis, ^{16}O or ^{18}O isotopes can be incorporated into the C-termini of peptides in the presence of ^{16}O or ^{18}O water. The relative

quantity of proteins is determined by the ratio of ion intensities of ^{16}O- to ^{18}O-labelled peptides measured by MS.[36–39] A pitfall of this method is that the possible loss or incomplete incorporation of the isotopic labels may complicate the quantitation.

While this type of N-terminal or C-terminal peptide labelling approach is expected to have complete protein coverage, it remains a challenge to apply it to global proteome-wide quantitative profiling, due to the high sample complexity. Currently, this approach is particularly useful in quantifying specific subsets of the proteome, such as protein mixtures separated by biochemical (ie immunoprecipitation, organelle fractionation, etc) or chemical (ie.amino acid-based purification, as discussed below) procedures.

Peptide labelling with amino acid constraint

One approach to reducing sample complexity is to enrich for peptides containing a specific amino acid by chemically modifying the amino acid and affinity purifying the modified peptides. The ICAT labelling technology described above is a prime example of this approach, in which the cysteine-containing peptides are purified by avidin affinity chromatography via the incorporated biotin tags. In addition to cysteine residues labelled by the ICAT reagent, other amino acids have been reported for enrichment. Kuyama *et al.* reported an approach to enrich tryptophan-containing peptides by modifying the tryptophan residues with isotopic ($^{12}C/^{13}C$) 2-nitrobenzenesulfenyl chloride (NBSCl).[40] The labelled tryptophan-containing peptides were then enriched by Sephadex chromatography, taking advantage of increased hydrophobicity of the tryptophan-containing peptides after the NBSCl modification. The enriched peptides were then identified by MS and quantified by the ratio of the ion intensities of the isotopic peaks. Similarly, isotopically labelled phosphotyrosine-containing peptides were reported to be quantified after enrichment of the phosphotyrosine-containing peptides through an anti-phosphotyrosine antibody-based purification procedure.[15]

Another effort to target specific amino acid-containing peptides, reported by Cagney *et al.*, analysed lysine-containing peptides by a method termed mass-coded abundance tagging (MCAT).[41] Contrary to the NIT method discussed above, the MCAT procedure specifically labels the ε-amino group of lysine by guanidinylation using *O*-methylisourea. The labelled sample is then compared with the non-labelled sample to determine relative quantities. This is not a strict stable isotope labelling approach, as the internal standard (unlabelled peptide) is chemically different from the labelled sample by more than just isotopic atoms, and the physicochemical difference between the labelled and unlabelled peptides may reduce the accuracy of the quantitation.

QUANTITATION WITHOUT STABLE ISOTOPE TAGS

Isotopic labelling of proteins is not always practical and has several disadvantages. For example, labelling with stable isotopes is expensive, and sometimes the isotopic labels exhibit chromatography shifts that can make quantitation of differentially labelled peptides computationally difficult. Moreover, there may not be enough different isotopes to allow for simultaneous quantitation of proteins from multiple samples. As an alternative, several groups have presented methods of peptide and protein quantitation without isotopic tags by comparing peptide signal intensities measured in sequential MS analyses.

For example, signal intensity-based quantitation has been applied to quantify differentially expressed proteins from samples analysed by LC-MS. Pasa-Tolic *et al.* have presented a system that uses the mass of a peptide coupled with its corresponding chromatographic elution

time as peptide properties that uniquely define a peptide sequence, a method termed 'the accurate mass and time (AMT) tag approach'.[42] Using LC coupled with Fourier transform ion cyclotron resonance (LC-FTICR) MS to obtain the chromatographic and high mass accuracy information, they identified peptide sequences by matching the AMT tags to previously acquired LC-MS/MS sequence information stored in a database. By taking advantage of the observed linear correlation between peak area of measured peptides and their abundance,[43,44] these peptides were relatively quantified by the signal intensity ratio of their corresponding peaks compared between MS runs. Using this method to analyse proteins from *Shewanella oneidensis* grown under aerobic or low oxygen conditions, they were able to observe changes in the abundance ratio of several proteins previously known to change expression levels under differential oxygen levels by as much as tenfold (Ruihua Fang and Richard Smith, personal communication). The primary advantage of this method is that by not selecting masses for CID, as is carried out with traditional LC-MS/MS, one can obtain higher run to run identification reproducibility, consequently allowing for accurate quantitation of more peptides.

A major disadvantage of peptide quantitation by the signal intensity is that it often includes experimental variation and signal noise which can affect the quantitative value and accuracy. To circumvent this problem, Wiener and colleagues at Merck Research Laboratories applied statistics tools such as the Student's t-test to analyse data from multiple LC-MS runs (ie ten times in their experiment) for each sample.[45] At each point of acquisition time and m/z, the amplitudes of signal intensities from multiple LC-MS runs were compared between two samples to detect peptides with statistically significant differences in abundance between samples. This statistically validated approach of signal intensity comparison, focusing on signal variability between samples, appears to be more sensitive and robust than the traditional approach focusing on absolute signal intensity acquired from background/noise subtraction. The major benefit of this approach is its ability to detect low-abundance peptides that are significantly different between samples but are of such low abundance that they would be ignored by traditional intensity-dependent acquisition. As a consequence, more peptides can be identified and quantified by this approach.

The use of signal intensities for quantitation is of particular interest for cases in which isotopic labelling is impractical or not feasible. For example, it has been used for the quantitation of proteins present in multiple cellular fractions. Andersen *et al.*[46] demonstrated the effectiveness of signal intensity-based quantitation by the analysis of distribution profiles for centrosomal proteins present in multiple sucrose gradient fractions. Using LC-MS/MS, they analysed consecutive fractions from a sucrose gradient and quantified LC-time matched peptides using the area under the extracted ion peaks. By calculating the relative abundance of peptides present in samples from each sucrose gradient fraction, they identified signature distribution profiles for protein groups present in different cellular organelles, including the centrosome. From the obtained protein profiles, they identified most of the known centrosomal proteins, as well as 23 novel components and 41 potential new components that would have been difficult to identify by other isotopic labelling methods.

EFFECTIVE QUANTITATIVE PROTEIN PROFILING TOWARDS A SUB-PROTEOME

Although quantitative protein profiling technologies have developed rapidly, global proteome analysis of highly complex samples remains a challenge using currently available methods. It appears more practical to take a 'divide

and conquer' strategy to study specific subsets of the proteome comprehensively. Several approaches have been adapted to acquire sub-proteomes for comprehensive identification and quantitation, including those described above, to isolate peptides containing specific amino acids via chemical reactions. Other approaches towards acquiring sub-proteomes, not discussed indepth in this paper, involve isolation and quantitation of sub-proteomes based on post-translational modifications (PTM) such as phosphorylation,[47,48] glycosylation,[49] ubiquitination[50] and sumoylation.[51,52] The combination of quantitative proteomics and the sub-proteome fractionation approach will lead to quantitative characterisation of many biologically important sub-proteomes, and ultimately towards accomplishing global proteome characterisation. For a detailed discussion of PTM-based proteomics, readers are referred to recent reviews.[29,53,54]

An ideal approach is to target intelligently only the differentially expressed proteins (believed to be biologically interesting) for subsequent MS analysis. Griffin and coworkers reported a novel approach towards this goal by performing an offline MS analysis to select potential target peptides/proteins for further identification.[55,56] In this procedure, ICAT-labelled peptides were first analysed in MS mode to quantify the ratio between paired isotopically labelled peptides. Only those peptides with different abundance ratios between experimental and control samples, which are often biologically interesting from within a background of constitutively represented peptides, were selected for further identification by MS/MS. This abundance ratio-dependent proteomics approach selectively detects proteins on a functional basis. A similar strategy can also be applied in the signal intensity-based approach. Using similar approaches to those described above using comparison of signal intensities,[45] peptides/proteins with a statistically significant difference in abundance between two samples can be selected for further analysis and identification. .Another example is the approach using surface-enhanced laser desorption/ionisation (SELDI) protein chip technology;[57] however, only protein expression profiles without identification of responsible proteins were obtained in the SELDI technology. With recent improvement in the reproducibility of MS measurements, a combinatorial approach using MS-based peptide profiling technology to select targets for subsequent protein identification by MS/MS is expected to be developed in the next few years for intelligent quantitative proteomics profiling.

CONCLUSIONS

With steady advances in MS technologies, quantitative proteomics has progressed dramatically in the past few years. Efforts to analyse proteomes of many species, both qualitatively and quantitatively, have generated an abundance of data in a variety of biological systems, from bacteria to humans. While 2DE-based quantitation continues to be practised, and signal intensity-based quantitation has shown a promising future for further development, current approaches of quantitative proteomics are mainly based on incorporation of stable isotope tags into proteins/peptides.

The *in vivo* stable isotope labelling technology provides a consistent and accurate quantity for measuring protein abundance. The limitation of this approach is that it can only be applied to cultured cells and is thus restrictive for applications to tissues or body fluids, which are of particular interest for medical research. The *in vitro* labelling technology, including the commercially available ICAT and iTRAQ methods, can be used on all kinds of biological samples. The ICAT method, which focuses on cysteine-containing peptides only, has been successfully applied to the global quantitation of many proteomes. The recently introduced iTRAQ method, which can be used to label all peptides at

their N-termini, is particularly useful for quantifying proteins from multiplex samples with less complexity. Because of the enormous sample complexity of the whole proteome, a current practical and efficient method of quantitative proteomic profiling is to simplify biological samples by separating them into several subsets (sub-proteomes) by various fractionation methods. Comprehensive analyses of these biologically interesting sub-proteomes, and integration of these datasets by computational approaches, will ultimately lead to a more thorough molecular understanding of complex biological systems.

Acknowledgments
The authors thank Dr Jeffery A. Ranish for critical reading of the manuscript. This work was supported in part by Grants from the National Heart, Lung, and Blood Institute Proteomics Center at the Institute for Systems Biology, USA (N01-HV-28179) and from the National Cancer Institute, USA (1R33CA93302).

References

1. Glokler, J. and Angenendt, P. (2003), 'Protein and antibody microarray technology', *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, Vol. 797, pp. 229–240.
2. Espina, V., Mehta, A. I., Winters, M. E. *et al.* (2003), 'Protein microarrays: Molecular profiling technologies for clinical specimens', *Proteomics* Vol. 3,pp. 2091–2100.
3. Aebersold, R. and Mann, M. (2003), 'Mass spectrometry-based proteomics', *Nature*, Vol. 422, pp. 198–207.
4. Aebersold, R. and Goodlett, D. R. (2001), 'Mass spectrometry in proteomics', *Chem. Rev.*, Vol. 101, pp. 269–295.
5. Patton, W. F. and Beechem, J. M. (2002), 'Rainbow's end: The quest for multiplexed fluorescence quantitative analysis in proteomics', *Curr. Opin. Chem. Biol.*, Vol. 6, pp. 63–69.
6. Peng, J. and Gygi, S. P. (2001), 'Proteomics: The move to mixtures', *J. Mass Spectrom.*, Vol. 36, pp. 1083–1091.
7. Gorg, A., Weiss, W. and Dunn, M. J. (2004), 'Current two-dimensional electrophoresis technology for proteomics', *Proteomics*, Vol. 4, pp. 3665–3685.
8. Rabilloud, T. (2002), 'Two-dimensional gel electrophoresis in proteomics: Old, old fashioned, but it still climbs up the mountains', *Proteomics*, Vol. 2, pp. 3–10.
9. Lahm, H. W. and Langen, H. (2000), 'Mass spectrometry: A tool for the identification of proteins separated by gels', *Electrophoresis*, Vol. 21, pp. 2105–2114.
10. Oda, Y., Huang, K., Cross, F. R. *et al.* (1999), 'Accurate quantitation of protein expression and site-specific phosphorylation', *Proc. Natl. Acad. Sci. USA*, Vol. 96, pp.6591–6596.
11. Washburn, M. P., Ulaszek, R., Deciu, C. *et al.* (2002), 'Analysis of quantitative proteomic data generated via multidimensional protein identification technology', *Anal. Chem.*, Vol. 74, pp. 1650–1657.
12. Conrads, T. P., Alving, K., Veenstra, T. D. *et al.* (2001), 'Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and 15N-metabolic labeling', *Anal. Chem.*, Vol. 73, pp. 2132–2139.
13. Ong, S. E., Blagoev, B., Kratchmarova, I. *et al.* (2002), 'Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics', *Mol. Cell. Proteomics* Vol. 1, pp. 376–386.
14. Everley, P. A., Krijgsveld, J., Zetter, B. R. and Gygi, S. P. (2004), 'Quantitative cancer proteomics: Stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research', *Mol. Cell. Proteomics*, Vol. 3, pp. 729–735.
15. Ibarrola, N., Molina, H., Iwahori, A. and Pandey, A. (2004), 'A novel proteomic approach for specific identification of tyrosine kinase substrates using [13C]tyrosine', *J. Biol. Chem.*, Vol. 279, pp. 15805–15813.
16. Ong, S. E., Foster, L. J. and Mann, M. (2003), 'Mass spectrometric-based approaches in quantitative proteomics', *Methods'*, Vol. 29, pp. 124-130.
17. Blagoev, B., Ong, S.-E., Kratchmarova, I. and Mann, M. (2004), 'Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics', Vol. 22, pp. 1139–1145.
18. Gygi, S. P., Rist, B., Gerber, S. A. *et al.*(1999), 'Quantitative analysis of complex protein mixtures using isotope-coded affinity tags', *Nat. Biotechnol.*, Vol. 17, pp. 994–999.
19. Hardwidge, P. R., Rodriguez-Escudero, I., Goode, D. *et al.* (2004), 'Proteomic analysis of the intestinal epithelial cell response to enteropathogenic *Escherichia coli'*, *J. Biol. Chem.*, Vol. 279, pp. 20127–20136.
20. Ranish, J. A., Yi, E. C., Leslie, D. M. *et al.* (2003), 'The study of macromolecular complexes by quantitative proteomics', *Nat. Genet.*, Vol. 33, pp. 349–355.
21. Brand, M., Ranish, J. A., Kummer, N. T. *et al.* (2004), 'Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative

proteomics', *Nat. Struct. Mol. Biol.* Vol. 11, pp. 73–80.

22. Han, D. K., Eng, J., Zhou, H. and Aebersold, R. (2001), 'Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry', *Nat. Biotechnol.*, Vol. 19, pp. 946–951.
23. Yan, W., Lee, H., Yi ,E. C. *et al.* (2004), 'System-based proteomic analysis of the interferon response in human liver cells', *Genome Biol.*, Vol. 5, p. R54.
24. Jiang, X.-S., Dai, J., Sheng, Q.-H. *et al.* (2004), 'A comparative proteomic strategy for subcellular proteome research: icat approach coupled with bioinformatics prediction to ascertain rat liver mitochondrial proteins and indication of mitochondrial localization for catalase', *Mol. Cell. Proteomics*, Vol. 4, pp. 4–34.
25. Harris, R. D., Nindl, G., Balcavage, W. X. *et al.* (2003), 'Use of proteomics methodology to evaluate inflammatory protein expression in tendinitis', *Biomed. Sci. Instrum.*, Vol. 39, pp. 493–499.
26. Wu, S. L., Choudhary, G., Ramstrom, M. *et al.* (2003), 'Evaluation of shotgun sequencing for proteomic analysis of human plasma using HPLC coupled with either ion trap or Fourier transform mass spectrometry', *J. Proteome Res.*, Vol. 2, p.p. 383–393.
27. Zhang, R. and Regnier, F. E. (2002), 'Minimizing resolution of isotopically coded peptides in comparative proteomics', *J. Proteome Res.*, Vol. 1, pp. 139–147.
28. Zhou, H., Ranish, J. A., Watts, J. D. and Aebersold, R. (2002), 'Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry', *Nat. Biotechnol.*, Vol. 20, pp. 512–515.
29. Zhang, H., Yan, W. and Aebersold, R. (2004), 'Chemical probes and tandem mass spectrometry: A strategy for the quantitative analysis of proteomes and subproteomes', *Curr. Opin. Chem. Biol.*, Vol. 8, pp. 66–75.
30. Ji, J., Chakraborty, A., Geng, M. *et al.* (2000), 'Strategy for qualitative and quantitative analysis in proteomics based on signature peptides', *J. Chromatogr. B Biomed. Sci. Appl.*, Vol. 745, pp. 197–210.
31. Munchbach, M., Quadroni, M., Miotto, G. and James, P. (2000), 'Quantitation and facilitated *de novo* sequencing of proteins by isotopic N-terminal labeling of peptides with a fragmentation-directing moiety', *Anal. Chem.*, Vol. 72, pp. 4047–4057.
32. Mason, D. E. and Liebler, D. C. (2003), 'Quantitative analysis of modified proteins by LC-MS/MS of peptides labeled with phenyl isocyanate', *J. Proteome Res.*, Vol. 2, pp. 265–272.
33. Zhang, X., Jin, Q. K., Carr, S. A. and Annan, R. S. (2002), '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.*, Vol. 16, pp. 2325–2332.
34. Ross, P. L., Huang, Y. N., Marchese, J. *et al.* (2004), 'Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents', *Mol. Cell. Proteomics* , Vol. 3, pp. 1154–1169.
35. Goodlett, D. R., Keller, A., Watts, J. D. *et al.* (2001), 'Differential stable isotope labeling of peptides for quantitation and *de novo* sequence derivation', *Rapid Commun. Mass Spectrom.*, Vol. 15, pp. 1214–1221.
36. Mirgorodskaya, O. A., Kozmin, Y. P., Titov, M. I. *et al.* (2000), 'Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)O-labeled internal standards', *Rapid Commun. Mass Spectrom.*, Vol. 14, pp. 1226–1232.
37. Wang, Y. K., Ma, Z., Quinn, D. F. and Fu, E. W. (2001), 'Inverse ^{18}O labeling mass spectrometry for the rapid identification of marker/target proteins', *Anal. Chem.*, Vol. 73, pp. 3742–3750.
38. Stewart, I. I., Thomson, T. and Figeys, D. (2001), '^{18}O labeling: A tool for proteomics', *Rapid Commun. Mass Spectrom.*, Vol. 15, pp. 2456–2465.
39. Yao, X., Freas, A., Ramirez, J. (2001), 'Proteolytic ^{18}O labeling for comparative proteomics: Model studies with two serotypes of adenovirus', *Anal. Chem.*, Vol. 73, pp. 2836–2842.
40. Kuyama, H., Watanabe, M., Toda, C. *et al.* (2003), 'An approach to quantitative proteome analysis by labeling tryptophan residues', *Rapid Commun. Mass Spectrom.*, Vol. 17, pp. 1642–1650.
41. Cagney, G. and Emili, A. (2002), '*De novo* peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging', *Nat. Biotechnol.*, Vol. 20, pp. 163–170.
42. Pasa-Tolic, L., Masselon, C., Barry, R. C. *et al.* (2004), 'Proteomic analyses using an accurate mass and time tag strategy', *Biotechniques*, Vol. 37, pp. 621–624, 626–33, 636 passim.
43. Tang, K., Page, J. S. and Smith, R. D. (2004), 'Charge competition and the linear dynamic range of detection in electrospray ionization mass spectrometry', *J. Am. Soc. Mass Spectrom.*, Vol. 15, pp. 1416–1423.
44. Chelius, D. and Bondarenko, P. V. (2002), 'Quantitative profiling of proteins in complex mixtures using liquid chromatography and

mass spectrometry', *J. Proteome Res.*, Vol. 1, pp. 317–323.

45. 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.*, Vol. 76, pp. 6085–6096.
46. Andersen, J. S., Wilkinson, C. J., Mayor, T. *et al.* (2003), 'Proteomic characterization of the human centrosome by protein correlation profiling', *Nature,* Vol. 426, pp. 570–574.
47. Salih, E. (2004), 'Phosphoproteomics by mass spectrometry and classical protein chemistry approaches', *Mass Spectrom. Rev.*, Vol ??, pp.??.
48. Mann, M., Ong, S. E., Gronborg, M. *et al.* (2002), 'Analysis of protein phosphorylation using mass spectrometry: Deciphering the phosphoproteome. *Trends Biotechnol.*, Vol. 20, pp. 261–268.
49. Zaia, J. (2004), 'Mass spectrometry of oligosaccharides', *Mass Spectrom. Rev.*, Vol. 23, pp. 161–227.
50. Peng, J., Schwartz, D., Elias, J. E. *et al.* (2003), 'A proteomics approach to understanding protein ubiquitination', *Nat. Biotechnol.*, Vol. 21, pp. 921–926.
51. Denison, C., Rudner, A. D., Gerber, S. A. *et al.* (2004), 'A proteomic strategy for gaining insights into protein sumoylation in yeast', *Mol. Cell. Proteomics,* Vol. ??, pp. M400154–MCP200.
52. Panse, V. G., Hardeland, U., Werner, T. *et al.* (2004), 'A proteome-wide approach identifies sumoylated substrate proteins in yeast', *J. Biol. Chem.*, Vol. 279, pp. 41346–41351.
53. Mann, M. and Jensen, O. N. (2003), 'Proteomic analysis of post-translational modifications', *Nat. Biotechnol.*, Vol. 21, pp. 255–261.
54. Reinders, J., Lewandrowski, U., Moebius, J. *et al.* (2004), 'Challenges in mass spectrometry-based proteomics', *Proteomics* Vol. ?, pp. ??.
55. Griffin, T. J., Lock, C. M., Li, X. J. *et al.*, (2003) 'Abundance ratio-dependent proteomic analysis by mass spectrometry', *Anal. Chem.*, Vol. 75, pp. 867–874.
56. Griffin, T. J., Gygi, S. P., Rist, B. *et al.* (2001), 'Quantitative proteomic analysis using a MALDI quadrupole time-of-flight mass spectrometer', *Anal. Chem.*, Vol. 73, pp. 978–986.
57. Weinberger, S. R., Dalmasso, E. A. and Fung, E. T. (2002), 'Current achievements using ProteinChip Array technology', *Curr. Opin. Chem. Biol.*, Vol. 6, pp. 86–91.