1
Quantitation in Proteomic Experiments Utilizing Mass Spectrometry

ANDREW C. PAOLETTI AND MICHAEL P. WASHBURN*

Stowers Institute for Medical Research, 1000 E. 50th Street, Kansas City, MO 64110, USA

Introduction

Advances in the field of quantitative proteomics are beginning to routinely allow scientists to reliably and accurately determine absolute protein concentrations or changes in protein concentration from complex samples *in vivo*. These changes can be measured in response to differential growth conditions, drug regimens, disease states, and even more specifically, quantitative measurements can determine levels of PTMs, such as phosphorylation in cell signalling pathways or glycosylation levels of haemoglobin in diabetic patients. These measurements are allowing scientists unprecedented insight into the roles individual protein levels and modifications play in the overall view of the system being studied.

This is not meant to be a comprehensive review of the subject, and numerous excellent reviews that emphasize specific topics or methodologies are available

*To whom correspondence may be addressed (MPW@Stowers-Institute.org)

Abbreviations: 1D, one-dimensional; 2D, two-dimensional; AQUA, absolute quantification; CHCA, α-cyano-4-hydroxycinnamic acid; CV, coefficient of variation; EGF, epidermal growth factor; ESI, electrospray ionization; FLAG, recombinant protein with octapeptide tag; GST, glutathione S-transferase; ICAT, isotope coded affinity tag; IEF, isoelectric focusing; MALDI, matrix assisted laser desorption ionization; MS/MS, mass spectrometry/mass spectrometry; NCBI, National Center for Biotechnology Information; PSD, postsynaptic density; PTM, post-translational modification; RA, rheumatoid arthritis; RP, reverse phase; SCX, strong cation exchange; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; SGI, Silicon Graphics, Inc.; SISCAPA, stable isotope standards and capture by anti-peptide antibodies; SRM, selective reaction monitoring; TOF, time of flight; VICAT, visible isotope coded affinity tag.

Abbreviations used for proteins: AAC, α_1-antichymotrypsin; CRP, C-reactive protein; GKAP, disks large-associated protein 1; Glu-C, endoproteinase Glu-C; GLuR1, glutamate receptor 1 precursor; GLuR2, glutamate receptor 2 precursor; HbAc, haemoglobin Ac; HbAlc, haemoglobin Alc; hGV, human recombinant group V; HSA, human serum albumin; Hx, hemopexin; IgG, immunoglobulin G; IL-6, interleukin-6; NmDAR1, glutamate [NMDA] receptor subunit zeta 1 precursor; NPR1, nitrogen permease reactivator protein; PSD-95, presynaptic density protein 95; Shank1, SH3 and multiple ankyrin repeat domains protein 1; TNF-α, tumour necrosis factor-α.

Biotechnology and Genetic Engineering Reviews – Vol. 22, January 2006
0264–8725/06/22/1–19 $20.00 + $0.00

(Aebersold and Mann, 2003; Ranish *et al.*, 2003; Julka and Regnier, 2004). The aim of this review is to familiarize readers with advancements in the various disciplines of proteomics, outline specific procedures and methodologies that are utilized in studies, and to present the most recent, significant experiments that have been performed. Hopefully, this will demonstrate to readers who are less familiar with the field, the manner of hypotheses that may be addressed by these methods, the equipment and skills necessary to carry them out, a basic overview of the data compiled by these studies, and the conclusions that can be reliably established from the data.

Proteomics is the study of the proteome of a species, and individual proteomics experiments analyse large, complex protein samples (Washburn *et al.*, 2001; Durr *et al.*, 2004; Yan *et al.*, 2004). These samples routinely consist of whole cell lysates, large multi-protein complexes, or proteins with a specific chemical make-up (such as phosphorylation), and may consist of hundreds of proteins. In order to carry out these analyses, technologies from many diverse fields are coupled and utilized in order to perform these experiments and identify the proteins present. A brief outline of the major components of a typical proteomics experiment and the technology involved follows.

Chromatography

Although intact proteins may be utilized in experiments, most current studies digest protein samples into peptides (normally with trypsin) and identify the proteins from the resulting peptide sequences. Since thousands of peptides may be analysed in a single experiment, it is necessary for a chromatographic system to resolve this peptide mixture so that individual peptides are introduced into the mass spectrometer and analysed separately.

The chromatography system employed depends on the sample but may consist of single or multiple pumps, switching valves, and columns. Chromatographic methods for analysis may combine several of the following methods; 1D or 2D SDS-PAGE, ion exchange, reverse phase, normal phase, size exclusion, affinity or antibody capture in various orders so that maximum resolution of the sample peptides is achieved (Nagele *et al.*, 2003, 2004).

Spectrometry

Proteomics experiments require that the sample is ionized in the gas phase, masses of the ions can be resolved and differentiated, followed by fragmentation of the parent ion in order to analyse the spectra of the resulting fragments. Ionization usually takes place via ESI, where an electric current ionizes small solvent droplets eluted from a chromatographic system, or MALDI, where the sample is deposited on a matrix containing a molecule (such as CHCA) that, when irradiated by a laser, both sublimates and ionizes the analyte. Mass analysers routinely utilized are quadrupoles (which act as a mass filter allowing specific ranges of mass/charge ratios to be analysed), ion trap (that accumulates ions of specific mass/charge ratios), and TOF (where mass/charge ratios of ions are determined by the flight time of ions from ionization to the detector). Ion trap mass spectrometers are usually coupled to ESI sources and utilized to measure

MS/MS events, while TOF analysers are usually coupled to MALDI sources and generally demonstrate greater mass accuracy. Quadrupole mass analysers may be employed in tandem with TOF or ion traps in order to improve detection and sensitivity (Aebersold and Mann, 2003; Yates, 2004).

After an initial mass spectrum of the injected sample has been taken, routinely a parent ion is chosen (may be based on a specific m/z or highest abundance) for further analyses. The parent ion is then induced to fragment (normally by collision-induced dissociation with an inert gas, such as helium) and a new mass spectrum is taken of one of the fragments from the parent ion (MS/MS). This process may be repeated several times, depending on the type of analysis being performed. These additional experimental mass spectra (normally of digested peptides from the biological sample) are utilized to identify the protein from which the peptide originated. This is accomplished by comparing the experimental mass spectra to the theoretical mass spectra of peptides from protein databases by search algorithms, such as SEQUEST (Eng *et al.*, 1994) or MASCOT (Perkins *et al.*, 1999).

Bioinformatics

The large numbers of peptide spectra generated during proteomics experiments must be searched against protein databases in order to determine which proteins are present in the sample. Computer algorithms, such as SEQUEST (Eng *et al.*, 1994) and MASCOT (Perkins *et al.*, 1999), generate theoretical peptide spectra that are compared to the experimental spectra in order to identify the peptides present. These peptides are then searched against protein databases, such as NCBI (www.ncbi.nlm.nih.gov) or Swiss-Prot (www.isb-sib.ch), in order to identify the proteins represented by the identified peptides. New computer programs, such as Peptide Prophet (von Haller *et al.*, 2003), allow statistical evaluations of the peptides identified to reduce the numbers of incorrectly identified peptides. Identified peptides and proteins can be organized by a computer program like DTASelect (Tabb *et al.*, 2002), which can apply additional, more stringent acceptance criteria to the identified peptides. DTASelect (Tabb *et al.*, 2002) then displays these peptides according to the number of times each was identified and the protein from which it originates, as well as data such as protein sequence coverage and observed and calculated masses. Large data sets from separate analyses of protein samples can be compared by a computer program called Contrast (Tabb *et al.*, 2002), which searches all peptides and proteins identified in each experiment and lists the proteins and peptides identified in all analyses, allowing the researcher to evaluate the reproducibility of the experimental method employed. These analyses require dedicated computer systems, such as a Linux cluster with multiple nodes or SGI SuperComputers, to perform these searches in a time-effective manner.

Quantitation

In order to determine the concentration of a specific protein in a sample, it is necessary to measure and compare the signal from a known standard to the signal from the desired target (can be a protein or a peptide). These standards can be measured either concurrently with the sample (internal) or separately (external), with

the signal intensity measured for different concentrations establishing a calibration curve for the analyte protein. If the standard protein is measured separately, there is no need to isotopically label the standard (with stable isotopes), thereby saving both money and time. Internal standard proteins can be either isotopically labelled (also with stable isotopes) or not, depending on the manner in which the data is to be interpreted and the information that needs to be determined. Isotopically-labelled proteins can determine the absolute concentration of otherwise identical sample proteins if the labelled protein is highly purified and the concentration known. Chemically identical (standards are isotopically labelled) proteins and peptides will elute from the chromatographic system at the same time, and the absolute concentration can be determined by comparing the ratio of the standard/sample signals (ion count measured by the detector for each of the isotopic species present). Isotopically-labelled proteins can also be utilized in a less stringent manner, such as comparing relative protein levels between samples expressed in labelled and unlabelled media. This method is simpler (alleviates the need for purification procedures), but only relative concentrations between the samples can be determined, again by comparison of the signal ratios. Internal standards that are not isotopically labelled need to be chemically different (i.e. not the same protein or peptide) so that they can be distinguished by the mass spectrometer. Relative concentrations of analyte proteins can then be determined by normalization of the signal from the analyte protein to the signal from a known concentration of the internal standard. In this way, the relative protein concentrations from two separate growths can be determined by comparison and normalization with an internal standard protein that is not identical.

STABLE ISOTOPIC LABELLING

There are four main strategies utilized in the labelling of proteins and peptides for use as standards to quantitate individual proteins in mass spectrometric-based proteomic experiments (Hunter and Washburn, 2003), with different approaches having different entry points into the scheme, as shown in *Figure 1.1* (Washburn *et al.*, 2002). This allows the investigator a great deal of flexibility to determine the label most suited for the experiment, control the costs associated with various strategies, and the time and effort required for purification of the labelled standard.

(1) Metabolic labelling: proteins can be expressed in isotopically-enriched media. This can mean that a specific atom is enriched, i.e. ^{15}N, ^{13}C, ^{2}H, or that specific amino acids have stable isotopes already incorporated. Metabolic labelling has been employed with *E. coli* (Marley *et al.*, 2001), *C. elegans* (Krijgsveld *et al.*, 2003), *D. melanogaster* (Krijgsveld *et al.*, 2003), *Rattus norvegicus* (Wu *et al.*, 2004), yeast (Washburn *et al.*, 2002), human cells (Ibarrola *et al.*, 2003), and can be utilized in absolute or relative quantitation, with absolute standards requiring additional time and money spent in purification (*Figure 1.1*).

(2) Amino acid modifications: ICATs (Zhou *et al.*, 2002; Whetstone *et al.*, 2004) are isotopically-labelled, synthetic molecules that allow the labelling of proteins or peptides by covalently binding to cysteine residues via a thiol reactive group incorporated into the isotope tag. ICAT reagents also incorporate biotin that allows purification and enrichment of the desired peptides by avidin capture.

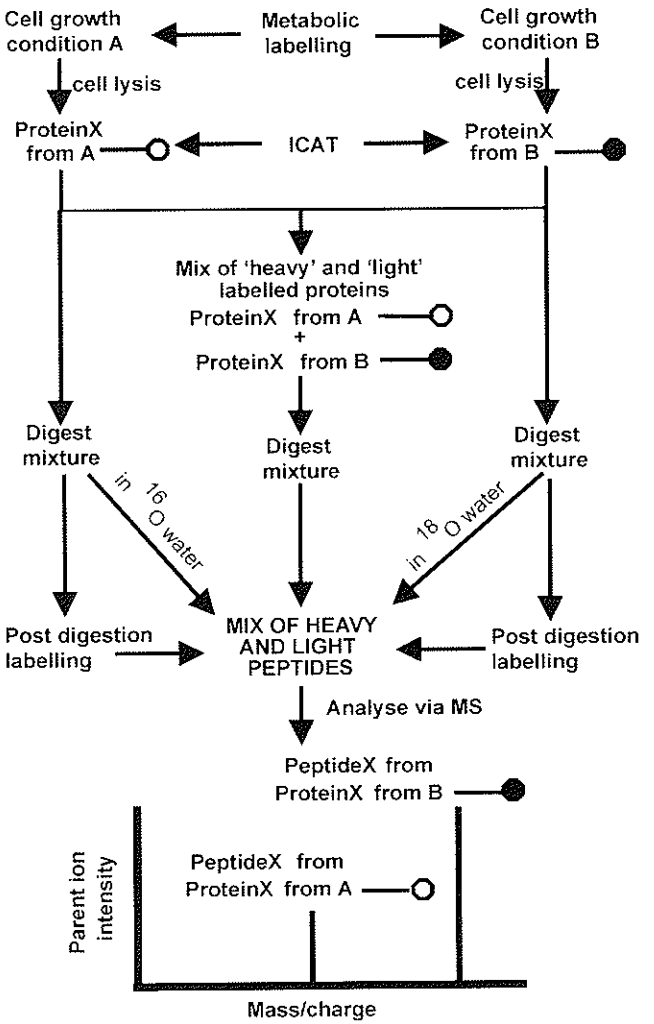

Figure 1.1. Method entry into a quantitative proteomic analysis scheme. When carrying out a quantitative proteomic analysis, the key is for the same peptide from two unique growth conditions to have unique masses when being analysed by a mass spectrometer. 'Heavy' and 'light' peptides may be generated at many points in a sample preparation pathway. Metabolic labelling introduces a label during the growth of the organism, and is therefore the earliest point of introduction of 'heavy' and 'light' labels. Metabolic labelling is followed by ICAT, digestion in ^{16}O and ^{18}O water, and lastly, post-digestion labelling. Only after a label has been introduced can the samples be mixed and further processed. (Reproduced with permission from *Analytical Chemistry* from Washburn *et al.*, 2002.)

Other amino acid functional groups, such as primary amines (Hoang *et al.*, 2003) and carboxylic acids (Goodlett *et al.*, 2001), may also be modified by mass tags. ICAT labelling can be utilized in absolute or relative quantitation, although use in absolute quantitation may become more widely exploited (Lu *et al.*, 2004). This method is expensive and requires extensive time spent in purification (*Figure 1.1*).

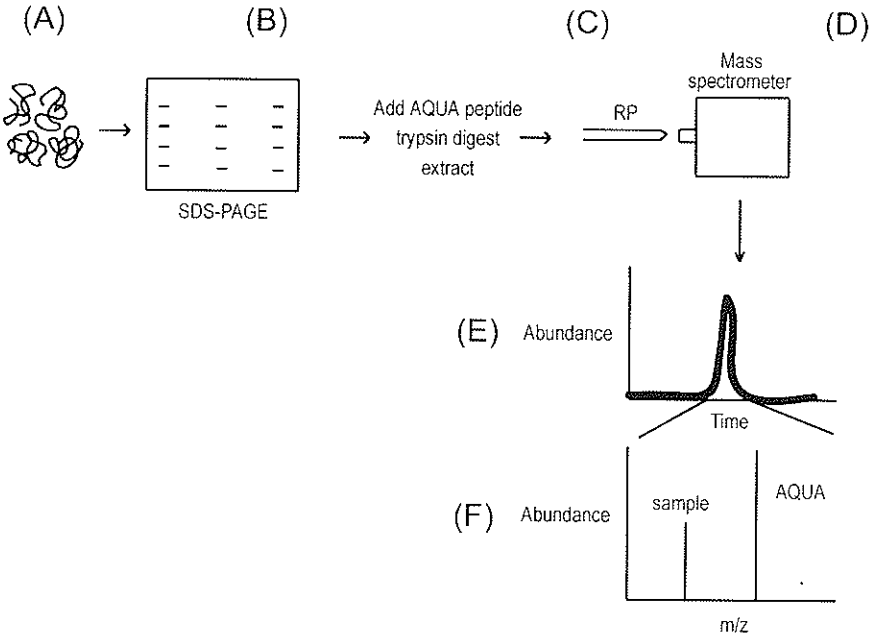

Figure 1.2. General scheme of the AQUA method for absolute protein concentration determination (Gerber *et al.*, 2003). Whole cell lysates are collected (A), and AQUA peptide standards are added to the gel (B), trypsin digested, and the sample is extracted (C). The peptide sample is then resolved by reverse phase chromatography and analysed by mass spectrometer (D). Individual elution peaks that contain both sample and AQUA peptides (E) are quantitated by the mass spectrometer, which measures the ratio of both isotopic species present (F). The concentration of the AQUA peptide standards are known, so the concentration of the sample peptide (protein) is then determined from the ratio of the signals.

(3) Peptides can be synthesized with isotope labels (as in the AQUA method (Gerber *et al.*, 2003)) which correspond to tryptic peptides that would be produced from trypsin digestion of the analyte protein (Nakanishi *et al.*, 2003). This method is employed for absolute quantitation, and purified peptides can be commercially obtained or synthesized with a peptide synthesizer. Prior digestion and mass spectral analysis of the target protein is required to determine that the peptide chosen is quantitatively produced during digestion (*Figure 1.2*).

(4) End labelling: C-terminus end labelling; proteins can be digested (normally by trypsin) in ^{18}O-labelled water, which allows the incorporation of two ^{18}O atoms in the carboxylic terminus of each resulting peptide (Johnson and Muddiman, 2004). N-terminus end labelling; selective labelling of peptide amino termini can be carried out with reagents, such as nicotinyl-N-hydroxysuccinimide (Munchbach *et al.*, 2000). These methods are normally associated with relative quantitation methods and therefore do not require additional purification (*Figure 1.1*).

ABSOLUTE QUANTITATION

The absolute quantitation of individual proteins from complex biological samples has the potential to elucidate the role a specific protein plays within the system being studied. This can be measured both in terms of protein concentration or the level of PTM necessary to translate a signal or stimulate a cellular event. For this reason, absolute quantitation will most likely play a greater role in the future of systems biology.

The laboratories of Turecek, Aebersold, and Gelb have developed a new method of affinity tagging proteins, termed VICAT, which allows for the absolute quantitation of proteins in cell lysates (Bottari *et al.*, 2004; Lu *et al.*, 2004). Three separate VICAT reagents are utilized in these studies; a protein tagging reagent that is bound to the proteins from the analyte sample, an internal standard reagent that is bound to a synthetic peptide corresponding to a tryptic peptide from the analyte protein, and an IEF marker reagent also bound to the synthetic tryptic peptide. All VICAT reagents contain a biotin moiety, photocleavable 2-nitro-benzyl bond, and a thiol reactive site. The internal standard reagent and the IEF marker reagent also incorporate a single ^{14}C atom.

All the reagents are readily distinguishable by mass spectrometric analysis; the protein tagging reagent and the internal standard reagent have an identical chemical skeleton, only the internal standard reagent has four ^{13}C and two ^{15}N atoms in the peptide tag, while the IEF marker reagent peptide tag is two methylene groups shorter than the other tags. The protein sample is reduced, reacted with protein tagging reagent, trypsin digested, and a known quantity of the synthetic tryptic peptide bound to the internal standard reagent is added to the sample. A larger quantity of the tryptic peptide bound to the IEF marker reagent is also added to the sample (this allows the peptides to be visualized on film and minimizes loss of sample during electrophoresis). All of the peptides are then resolved on an IEF gel (Amersham Biotech.), visualized on film, and the peptides with the VICAT reagents are cut and extracted from the gel. These tagged proteins are then bound to streptavidin agarose (Sigma), washed, and released by photocleavage. The tagged peptides are collected, further resolved on a C18 microcolumn before injection into an ESI LCQ Deca XP ion trap mass spectrometer (ThermoFinnigan), utilizing SRM.

This method was used to quantitate human recombinant group V protein (hGV) spiked into crude cell lysates from Sf9 insect cells. This was accomplished by integrating the elution peaks for the various tagged peptides and comparing the ratio of ion counts from the detector to determine the ratio of sample and standard peptide tagged with VICAT reagents. The hGV protein was quantitated from 50–1000 fmole and the signal response was linear in this range, with a line slope of 1.12 (Bottari *et al.*, 2004; Lu *et al.*, 2004). While this is a highly effective methodology, the reagents are not yet widely available, they are expensive, and the gel purification technique employed makes it difficult to automate the process.

An alternative approach relies on synthesized peptides (*Figure 1.2*) that correspond to tryptic peptides from the protein of interest. Gygi's laboratory has developed the AQUA method for the absolute quantitation of proteins and phosphoproteins from cell lysates (Gerber *et al.*, 2003). These synthesized peptides incorporate amino acids that contain stable isotopes (^{13}C, ^{15}N) and are added to the protein sample after

purification by SDS-PAGE. The AQUA peptide internal standard is therefore present during trypsin digestion, extraction from the gel, and reverse phase chromatography, so that any loss of sample happens equally for standard and sample peptide. Samples were analysed by an LCQ Deca XP ion trap or TSQ Quantum triple quadrupole mass spectrometer (ThermoFinnigan) and enhanced by only analysing known retention times for the peptides of interest. MS/MS spectra were collected for all parent ions, and quantitation was accomplished by integration of the peaks from the ion chromatograph and comparison of the ratios of the internal standard to the analyte protein peak. This methodology was originally validated by analysis and quantitation of horse heart myoglobin added to yeast cell lysates. Horse heart myoglobin was analysed from 300 amole to 30 pmole, and the calibration curve for expected versus observed myoglobin demonstrated a slope of one. Finally, the per cent of phosphorylation of Separase at Ser-1126 was determined utilizing two AQUA peptides, one of which was phosphorylated at the serine residue. This methodology was able to determine the dynamic changes in phosphorylation of this protein during premitosis, mitotic entry, and exit from mitosis. This analysis was accomplished from 400 μg of HeLa cell lysate using 0.15 pmole of AQUA peptide standards in the ion trap mass spectrometer, and only 16 μg of cell lysate and 0.1 pmole of AQUA peptide utilizing the triple quadrupole mass spectrometer (Gerber *et al.*, 2003).

A further application of the AQUA method (*Figure 1.2*) by Gygi is the absolute quantitation of six proteins from the rat forebrain PSD (Peng *et al.*, 2004). The proteins GLuR1, GLuR2, NmDAR1, PSD-95, GKAP, and Shank1 were expressed in ^{15}N media, purified by SDS-PAGE and quantified by Coomassie R-250 staining. Four pmoles of these proteins were added to approximately 20 μg of PSD, separated by SDS-PAGE, digested, and combined into one sample. The peptides were resolved on a 75 μm i.d. C18 column, and analysed on an LCQ Deca XP (ThermoFinnigan) ion trap mass spectrometer. Analysis of defined m/z ranges at specific elution times allowed the capture of labelled and unlabelled peptide pairs in the ion trap, which were then subjected to collision-induced dissociation and the MS/MS spectra obtained. Quantification was accomplished by comparison of the peak area ratios for labelled and unlabelled peptides. All six proteins were able to be quantitated by this method, and the per cent variation over two experiments ranged from 2.5% to 14.3% (Peng *et al.*, 2004). While the AQUA method has been demonstrated to be reliable when analysing complex samples, it also relies on gel-based separation techniques and requires testing to verify which peptides are suitable for use as standards.

A methodology also relying on isotopically-labelled peptide standards was utilized by Kuhn and co-workers to determine the amount of CRP in human serum samples (Kuhn *et al.*, 2004). This methodology differs from the AQUA approach in that it relies on the depletion of the most abundant proteins and fractionates protein samples by size exclusion chromatography, not SDS-PAGE. The levels of CRP were measured in subjects with and without RA, in order to determine the differences in concentration of this biomarker protein. Human plasma samples were first depleted of haptoglobin, IgG, and HSA using affinity chromatography (High Trap NHS-activated HP, High Trap Protein G, and HiTrap Blue columns, respectively, Amersham Biosciences). This was necessary since these proteins may constitute 60–90% of sample proteins and may interfere in mass spectrometric analysis during elution,

resulting in ion suppression of the desired peptides. The remainder of the sample was fractionated on a Superdex 200 Hi Load column (Amersham Biosciences) and analysed by SDS-PAGE to determine the fractions rich in CRP. These fractions were pooled, trypsin digested, and four synthetic ^{13}C-labelled tryptic peptides were added at either 250 or 500 fmole/μl. These samples were resolved with MAGIC C18 AQ Media (Michrom Bioresources) in a capillary column (New Objective) and analysed by an API-3000 TQ-MS (Applied Biosystems), MS/MS spectra were collected on doubly charged parent ions, and five subsequent transition ions were analysed. Concentrations of sample CRP were determined from the equation:

$$((\text{peak area ratio})(250 \text{ or } 500 \text{ fmole}/\mu\text{l})(\text{fraction volume, } \mu\text{l})(\text{MW}))/1 \times 10^9. \quad (1.1)$$

The average concentration determinations of CRP from four experiments were found to be 0.75 μg/ml for healthy patients, 2.8 μg/ml for non-erosive RA patients, and 37.65 μg/ml for erosive RA patients. Average CRP concentrations were determined to be approximately twice as high by immunoassay, but a measurable difference was seen for the three sets of patients. Finally, in order to determine the recovery percentage for CRP during the chromatographic process, 126 μg of purified protein was added to samples from non-erosive RA patients (2.8 μg/ml) and the concentration determined as outlined previously. The per cent recovery, based on comparison with the four labelled standard peptides, varied from 0.2% to 66% (Kuhn *et al.*, 2004). This study highlights the need to determine the recovery efficiency of standard and sample proteins, since low abundance proteins that are subjected to extensive purification techniques may themselves be greatly depleted and skew the results.

Isotopically-labelled tryptic peptide standards are an excellent method for quantifying proteins that may be difficult to work with due to solubility or gel separation issues. To overcome solubility problems, Barnidge and colleagues utilized a single peptide cleaved from rhodopsin (tryptic C-terminal peptide) in order to quantify the membrane protein (Barnidge *et al.*, 2003). This study was conducted on purified rhodopsin samples, so resolution of the protein from a complex mixture was not necessary. It was demonstrated that the peptide, TETSQVAPA, could be quantitatively cleaved from rhodopsin under membrane solubilizing and non-solubilizing conditions. The tryptic peptide was quantified with a synthetic peptide that incorporated two deuterated alanine residues. This peptide was utilized as both an internal and external standard in order to compare the quantification results achieved with both methods. The peptides were resolved on a 1 mm diameter column, packed with TARGA media C18 (Higgins Analytical), prior to analysis on a TSQ Quantum triple quadrupole mass spectrometer (ThermoFinnigan). Single reaction monitoring of the peptides was utilized so that only one daughter ion was analysed in the third quadrupole, resulting in the signal measured for quantitation. The linear dynamic range of the peptide standards was evaluated for unlabelled synthetic peptide utilized as an external standard and isotopically-labelled peptides utilized as an internal standard. The external standards (integration of peak area versus concentration) generated a linear standard curve from 100 fmole/μl to 8 pmole/μl with an R^2 value of 0 .9948, while the labelled internal standards (compared by the ratio of the signals for both peptides) was linear over the same concentration range with an R^2 value of 0.9964. Analyses of rhodopsin utilizing this method were carried out on two separate days, with each experiment carried out in triplicate, utilizing solubilized

membranes and with both internal and external standards. The use of external standards resulted in recovery ranging from 89% to 120% and CV values of 3.6 and 6.1. The internal standards yielded recoveries ranging from 84% to 159% and CV values of 4.1 and 17 (Barnidge *et al.*, 2003). In this analysis, the external standards allowed for more precise quantitation of the analyte peptide; this seems reasonable, since samples containing only rhodopsin were analysed, if the sample originated from a more complex mixture of peptides, matrix effects during chromatography and ionization might demonstrate a need for internal standards to overcome these difficulties. This study serves to demonstrate that proteins that are difficult to work with can still be quantitated if a suitable peptide can be utilized as a standard. Since the AQUA method first resolves proteins on a gel, this is an attractive alternative to absolute quantitation, if the protein can be highly purified by other methods (such as affinity purification).

Unlabelled peptides can be utilized as external standards when the sample is highly purified or the elution profile of the target does not overlap with other sample peptides. Jeppsson *et al.* (2002) capitalized on purified haemoglobin samples and employed an unlabelled peptide as an external standard to determine the ratio of glycosylated haemoglobin Alc (HbA1c) peptide to non-glycosylated haemoglobin Ac (HbAc) peptide for the Glu-C digested β-chain of haemoglobin glycosylated at the N-terminal amine group. This is important since there are fifteen widely utilized methodologies (chromatography, electrophoresis, immunology) that may not be entirely specific for the peptide analysed in this method and whose results may vary by 20%. Calibration solutions were prepared that contained 0%, 3%, 6%, 9%, 12%, and 15% of HbA1c peptide, then resolved on a reverse phase cyanopropyl column (ZORBAX), and the first twelve minutes of the elution profile were analysed by an SSQ700 single stage quadropole mass spectrometer (ThermoFinnigan). Utilizing mass spectrometry allows for the unambiguous determination of HbA1c and HbAc peptides. These measurements produced a linear calibration curve for the standards with an R^2 value of 1. Haemoglobin collected from patients was digested with Glu-C and analysed in the same manner as described above. Concentration was determined by the ratio of the signals for the standards versus the patient samples. This method was tested on ten samples in eleven different laboratories in Europe, USA, and Japan, and produced an inter-laboratory CV of 1.76% and an intra-laboratory CV of 1.05%. Comparison with another method (HPLC-CE) developed in this study yielded a linear calibration curve between the methods with an R^2 value of 1. This method has now been approved by the International Federation of Clinical Chemistry and Laboratory Medicine and will serve as the new reference method worldwide (Jeppsson *et al.*, 2002). This study highlights that mass spectral analysis allows the target biomolecule to be positively identified (not the case with chromatographic or gel techniques), while still allowing reliable and accurate quantitation determinations to be made. Mass spectral analysis may prove to be the method of choice for other analyses where positive identification of the target is not possible with the analytical technique being utilized.

While it is more difficult to utilize entire labelled proteins as standards (proteins from complex samples may co-elute and cause signal suppression), they may be useful in specific applications. An example of this method was demonstrated by Ji and colleagues, who utilized an intact, ^{15}N-labelled protein as an internal standard to

quantitate the rK5 protein from plasma samples (Ji *et al.*, 2003). The method was then employed to conduct preclinical trials of rK5 concentration in monkey plasma samples after initiation of a drug regimen. Plasma samples with 200 µl of approximately 5 µg/ml of ^{15}N rK5 were loaded in 96 well plates on a HLB 60 mg solid phase extraction plate (Waters Oasis), washed, and then eluted with 0.8 ml of acetonitrile and 0.2% TFA. These samples were then resolved on a C18 column (Symmetry300) at 40°C to improve peak shape, and analysed on an API 3000 triple quadrupole mass spectrometer (PE Sciex), utilizing selective reaction monitoring for the rK5 and ^{15}N rK5 standard. The peak areas were determined utilizing Sciex MacQuan software version 1.6 in order to determine the ratio of standard and analyte protein. The linear range of this assay was determined to be between 99.23 and 52 920.0 ng/ml, with R^2 values for calibration curves ranging between 0.9972 and 0.9994. Eighteen replicate samples from three analyses were run at the low and high end of the assay range and yielded CV values of 4.8% and 1.7%, and accuracy (per cent theoretical) of 105.3% and 100.3%, respectively. The extraction recovery of two control samples was determined to be 85% and 72% when the ion chromatograph peak area ratios were compared to quality control samples. This methodology was then applied to determine the rK5 level in monkey plasma samples after an intravenous drug was administered to the animal. The toxicokinetic curve was linear over the range studied ($10.5e^4$ to $10e^3$ ng/ml) during a 25-hour period after drug administration and all samples were analysed without dilution (Ji *et al.*, 2003). This study demonstrates that intact proteins may be utilized as internal standards even in complex mammalian samples, but it is important to demonstrate the chromatographic system employed is capable of resolving the protein targets.

The new methodologies for absolute protein concentration determination described here allow for very precise measurements of individual proteins from increasingly complex samples. While these methods will allow very specific hypotheses to be addressed, they do still retain some drawbacks. Some of the reagents are costly and not widely available (VICAT), require extensive purification procedures, the target protein must be extensively analysed to demonstrate it is a viable candidate (digestion and recovery issues), while several methods still rely on gel separation techniques. This means that the process is difficult to automate and precludes the study of certain proteins (high or low molecular weight, hydrophobic, extreme isoelectric point) (Patton, 1999; Gygi *et al.*, 2000; Patton *et al.*, 2002; Shaw and Riederer, 2003). As the methodologies become mature and are further refined, it is likely that many of these issues will be addressed to allow more scientists to employ these methods efficiently.

RELATIVE PROTEIN CONCENTRATION DETERMINATION

While relative protein concentration determinations cannot provide the precision of absolute quantitation studies, important details of the biological system's response to stimuli can be ascertained. In many instances (such as the response to a drug regimen), these studies may provide enough data to allow the investigator to reliably draw the proper conclusions. Relative protein concentration determinations can be measured with and without the use of stable isotopic labelling. Studies that do not incorporate isotopic labelling may use an exogenous protein (or peptide) as

an internal standard that the signal from sample proteins are normalized to, or compare the ratio of the signals from the same protein from two different samples. When stable isotopic labelling is utilized, one sample may be labelled either metabolically (Ibarrola *et al.*, 2003) or during digestion (Bonenfant *et al.*, 2003), with equal quantities of both samples measured concurrently, and the ratio of proteins from each sample determined by comparison of the signals from each isotopic species present. Alternatively, peptides labelled with stable isotopes, corresponding to tryptic peptides from the target protein, may serve as internal standards for each protein sample.

Non-isotopic labelling

Several groups have recently published articles describing the use of unlabelled, exogenous proteins (digested to peptides) as internal standards to allow for relative protein concentration determination between samples (*Figure 1.3*). The following three studies are applications of this methodology.

Bondarenko and colleagues employed horse myoglobin (Sigma) as an internal standard spiked into human plasma samples (Sigma) (both were digested and analysed as peptides) (*Figure 1.3*; Bondarenko *et al.*, 2002). The peptides were resolved on a 75 μm i.d. column packed with BioBasic C18 media (New Objective), and analysed on an LCQ Deca ion trap mass spectrometer (ThermoFinnigan). Two plasma samples with 200 or 400 fmole of horse myoglobin were analysed, and peak areas for the peptides were compared to determine if the difference in concentration could be reliably calculated. Next, the peak area for an internal protein, Apolipoprotein A-1, was calculated by comparison and normalized to horse myoglobin. All chromatographic peak areas were defined as ion intensity times seconds. The total peak areas for horse myoglobin peptides at 200 and 400 fmole were calculated to be 3.48×10^9 and 6.46×10^9, respectively. Next, the horse myoglobin to apolipoprotein A-1 ratio of peak areas was determined and found to be 1.15 and 2.23 for 200 and 400 fmole, respectively and, after normalization, the peak area ratios were determined to be 1 and 1.93, respectively. Repetitive measurements demonstrated that relative peak area measurement error was below 11% (Bondarenko *et al.*, 2002).

Wang and colleagues utilized a similar method in which exogenous proteins could be spiked into a complex protein sample and relative changes in sample protein concentration could be determined by comparison with these internal standards (*Figure 1.3*) (Wang *et al.*, 2003). Bovine carbonic anhydrase and horse myoglobin (Sigma) were spiked into human serum proteome samples (Sigma) at concentrations of 100 fmole to 100 pmole, depleted of serum albumin and IgG proteins by affinity beads (Proteomic Biosciences), digested with trypsin, resolved in a fused silica capillary column packed with BioBasic C18 media (New Objective), and analysed by both LCQ Deca ion trap (ThermoFinnigan) and LCT ESI TOF (Micromass) mass spectrometers. It was demonstrated that peak intensities for these internal standard peptides were linear over this concentration range, with R^2 values ranging from 0.993 to 0.9972 for calibration curves for these peptides. The CV values for 25 runs of the human serum samples gave median and average values of 25.7% and 29%, respectively. Comparison of peak areas from sample peptides with these internal standard peptides will allow for differential protein concentration

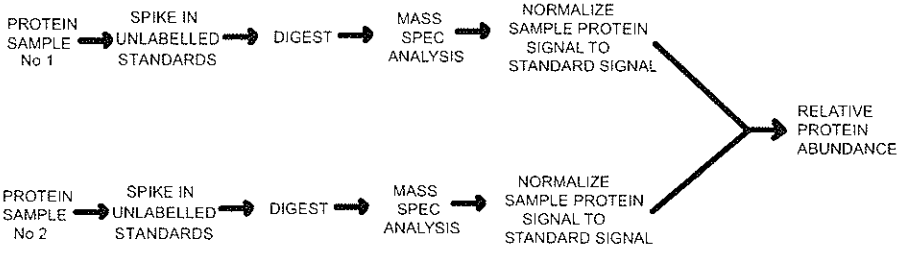

Figure 1.3. General scheme for relative protein abundance determinations between samples utilizing unlabelled standards. Individual protein samples are measured separately with known quantities of unlabelled standard proteins. After digestion and mass spectrum analysis, the signals from the sample peptides are normalized to the signal from a standard peptide. Since both signals are normalized to the same standard whose concentration is known, the ratio of signals from the sample peptides allows the relative abundance of proteins to be determined.

determinations of samples collected under varying conditions (Wang *et al.*, 2003).

Liu and colleagues utilized six unlabelled human proteins spiked into soluble yeast cell lysates as internal standards to be able to quantitate relative protein concentration differences between samples (*Figure 1.3*) (Liu *et al.*, 2004). Six human proteins (BioRad low molecular weight markers), phosphorylase b, serum albumin, ovalbumin, carbonic anhydrase, trypsin inhibitor, and lysozyme were spiked into yeast cell lysates over a range of 0.0417% to 4.17% of the total protein mass and digested with the sample. The peptide mixture was resolved on a biphasic SCX/RP capillary column (Partisphere SCX, Whatman, and XDB C18, Hewlett Packard), and analysed on an LCQ ion trap mass spectrometer (ThermoFinnigan). The number of spectra for each peptide recorded by the mass spectrometer was linear over the two order of magnitude range, and demonstrated R^2 values between 0.9967 and 0.9995 for calibration curves of spectral counts versus amount of internal standard added. These data suggest that, since the spectral counts for these peptides is determined by their concentration, comparison of spectra for analyte proteins compared to these internal standards could allow for the relative concentration determination of proteins from different samples (Liu *et al.*, 2004). The Bondarenko, Wang, and Liu studies demonstrate that low cost, commercially available standards can be utilized to reliably quantitate relative differences in protein concentration (Bondarenko *et al.*, 2002; Wang *et al.*, 2003; Liu *et al.*, 2004).

Software development is improving the quantitative determinations between samples by improved analysis of experimental data (chromatograms, signals) and statistical analysis of associated errors (Li *et al.*, 2003; MacCoss *et al.*, 2003). A new software program that allows analysis of the peak area ratios between individual samples of unlabelled proteins for differential protein quantitation without standards of any type has now been developed. An example of this method is demonstrated by Chelius and Bondarenko (2002), utilizing software developed at ThermoFinnigan. Two cultures of A431 cells were grown and, ten minutes prior to harvest, one sample was treated with EGF. Total protein extracts from both cultures were prepared, digested by trypsin, and analysed separately. Each culture was initially eluted from a strong cation exchange column by a series of salt bumps, each salt bump eluent was

captured by a polystyrene-divinylbenzene peptide trap (Michrom Bioresources), and finally resolved on a 75 μm i.d. column packed with Bio Basic C18 media (New Objective). Mass spectrometric analysis was carried out on an LCQ Deca XP Plus ion trap mass spectrometer (ThermoFinnigan), with two MS/MS spectra obtained for the most intense peaks, utilizing dynamic exclusion. Peak areas were calculated for each precursor ion within a defined elution time based on the signal intensity. Utilizing this methodology, 280 proteins were identified and 12 unique peptides were detected in four analyses (treated and untreated samples were analysed twice). The per cent error in the peak areas calculated for both runs generally ranged from 1.3 to 17.5 (the per cent error for one peptide was 68.2), while the standard deviations for the two runs varied from 0.1 to 0.94. The peak area ratios that were determined to be different by use of an a-table with p-values below 0.05 were considered significant. By these criteria, seven proteins were determined to be differentially expressed in A431 cells after administration of EGF. In order to detect phosphorylation of proteins, the initial SEQUEST (Eng *et al.*, 1994) search was mass modified at seriene, threonine, and tyrosine residues and, utilizing the same process, it was determined that the relative phosphorylation of four proteins increased in the EGF-treated sample (Chelius and Bondarenko, 2002). This study is important since, if this software becomes commercially available, it alleviates the need for more expensive labelling techniques.

Stable isotopic labelling

An example of stable isotope metabolic labelling used to determine differential PTM was demonstrated by Ibarrola and co-workers, who utilized isotopically-labelled Frigg protein to quantitate the differential phosphorylation of Frigg after stimulation with calyculin A at various time points (Ibarrola *et al.*, 2003). Human 293T cells, transfected with FLAG tagged Frigg expression construct, were originally grown in unlabelled media or media containing isotopically-labelled lysine (6 Dalton mass difference). Both growths were incubated with calyculin A for different periods of time (the light isotope media for the shorter period of time) to stimulate phosphorylation. After cell harvest, the cell lysates were mixed, immunoprecipitated with anti-FLAG antibody (Sigma), further resolved by SDS-PAGE, and stained with Coomassie blue. The Frigg protein band was excised, digested with trypsin, and the Frigg peptides resolved on a capillary column packed with C18 media (Vydac) before analysis on a Q-TOF API-US (Micromass) mass spectrometer. The first experiments (utilizing isotopically-labelled lysine amino acid) demonstrated a 1:3 ratio of Frigg phosphorylation in cells treated with calyculin A for 10 or 30 minutes, respectively, and a 1:9 ratio for cells treated for 3 and 30 minutes, respectively. Phosphorylation was confirmed by the loss of 69 mass units (β-elimination of phosphoric acid) during MS/MS scans of the peptide, and the ratio of non-phosphorylated versus phosphorylated peptide was determined by the ratio of ion current for each peptide present in the elution peak. In order to analyse other peptides that may be phosphorylated but do not contain lysine, the previous method was repeated, except that the cells were expressed in media containing isotopically-labelled lysine and arginine. Two new peptides were analysed, one of which was not phosphorylated during the course of the experiment. The other peptide was demon-

strated to be phosphorylated at S7 and S11, at a 1:2 ratio after 10 or 30 minutes, respectively, of calyculin A treatment. Phosphorylation at S7 had not been previously reported (Ibarrola *et al.*, 2003). This study demonstrates that the careful selection of metabolic labelling media can reliably quantitate relative differences in protein concentration, and may aid in the search for novel PTM sites.

Differential PTM can also be determined by C-terminal end labelling, as in the study by Bonenfant and co-workers, who utilized ^{18}O incorporated during trypsin digestion to determine the effect rapamycin treatment has on the phosphorylation of yeast NPR1 protein (Bonenfant *et al.*, 2003). Yeast strains JC19-1A or JC28-1B were transformed with a plasmid expressing GST-tagged NPR1, and either treated with rapamycin for 15 minutes, or left untreated. After harvesting the cells, NPR1 was first purified on a Glutathione Sepharose 4B column (Amersham Pharmacia), followed by capture of phosphoproteins by $FeCl_3$ charged, FAST Flow beads (Amersham Pharmacia). After release, the proteins were dephosphorylated with calf intestinal alkaline phosphatase (CIP), and untreated cells were digested with trypsin in $H_2{}^{16}O$, while rapamycin-treated cells were digested in $H_2{}^{18}O$. Equal quantities of peptides were mixed and spotted on a CHCA matrix for ionization. Mass spectrometric analysis was accomplished on a Bruker Reflex III MALDI-TOF (Bruker Daltonik), and determination of peptide ratios was accomplished by the ratio of the isotope clusters for labelled and unlabelled peptides. It was demonstrated that phosphorylation of NPR1 was unaffected by rapamycin treatment at amino acids 707–721, while phosphorylation of the peptide encompassing amino acids 353–362 decreased 3.6-fold after rapamycin treatment (Bonenfant *et al.*, 2003). This method of isotopic labelling is fairly inexpensive, straightforward, and may be applied in most laboratories, while still allowing accurate relative protein concentrations to be determined.

While not widely employed, effort is being made to utilize antibodies in the chromatographic purification of peptides from biological samples. It is believed that the specificity of antibodies will allow enrichment of target peptides and increase the resolution of the chromatographic system. Anderson and co-workers have developed an approach termed SISCAPA to enrich complex samples for specific peptides to aid in quantitation measurements (Anderson *et al.*, 2004). Single peptides from plasma proteins Hx, AAC, IL-6, and TNF-α were selected to serve as detection and quantification markers for these proteins. A rabbit antibody was raised for each peptide, bound to Self-Pack POROS beads (Applied Biosystems), and each antibody was separately packed into a 100 μm i.d. fused silica capillary to a bed length of approximately 1 cm. After elution from the antibody column, the flow was captured by a Pepmap C18 trap cartridge (LC Packings) before loading onto a Pepmap C18 analytical column (LC Packings) for final resolution before injection into a Q-TRAP triple quadrupole/linear ion trap mass spectrometer (Applied Biosystems). Selective reaction monitoring of known fragment ions increased specificity and was used to quantitate peptides, while the quadrupole mass filters allowed a range of ions within 4 amu to enter the ion trap to increase sensitivity. When a mixture of the four standard isotopically-labelled peptides was applied to each antibody column, all four peptides were detected in approximately a 1:1 ratio, and gave an average enrichment of 120-fold for the target peptide. However, when a complex peptide mixture from human plasma was applied to the columns, only the peptides from Hx and AAC were detected. The proteins IL-6 and TNFα are low abundance proteins,

and no peptides were detected by mass spectrum analysis (Anderson *et al.*, 2004). While the use of antibodies holds promise for greater enrichment and resolution of target peptides, this study demonstrates that the field has to overcome many obstacles inherent in antibody chromatography (slower binding kinetics, selectivity and lifetime issues).

Relative protein concentration determinations have more traditionally employed stable isotopic labelling to measure concentration differences between samples (Washburn *et al.*, 2002; Hardwidge *et al.*, 2004). This method has been demonstrated to be reliable, but relatively costly, due to the stable isotopes utilized. New computer algorithms are beginning to allow differential protein concentrations to be measured without the need for stable isotopes (Bondarenko *et al.*, 2002; Chelius and Bondarenko, 2002; Wang *et al.*, 2003; Liu *et al.*, 2004). These methods allow unlabelled, commercially available standards (cheaper than labelled samples) to be utilized as internal standards, or for the signals from individual samples to be compared without the need for standards at all. As these new methods are improved and demonstrated to be reliable, the costs associated with current studies should decrease and allow increased numbers of researchers to employ these methods.

Conclusion

Absolute protein concentration determinations allow for very precise measurements of individual proteins from complex samples, but are costly and more time consuming due to the expensive stable isotope labelled standards and the extensive purification protocols utilized. Relative protein concentration determinations can be utilized when only changes in concentrations need to be measured (such as in response to a drug regimen). While these measurements have utilized stable isotopic labelling, they do not require extensive purification protocols and new methodologies described here are now alleviating the need for isotopic labelling.

The tools available to scientists today allow for great flexibility in determining protein concentrations from biological samples. This can range from simple differential expression experiments with unlabelled standards to absolute concentration determinations using isotopically-labelled standards and affinity-purified complexes from transfected cells. These methods allow the examination, in ever-greater detail, of the role individual proteins (or PTMs) are playing in the context of the entire system. The technology necessary to carry out these experiments is either commercially available (mass spectrometers, chromatographic systems, search algorithms) or available for public use through government-funded initiatives (NCBI, Swiss-Prot). The contribution made by mass spectrometric-based proteomics to the biological field should increase dramatically as more scientists become aware of these technologies and become familiar with their use.

References

AEBERSOLD, R. AND MANN, M. (2003). Mass spectrometry-based proteomics. *Nature* **422** (6928), 198–207.

ANDERSON, N.L., ANDERSON, N.G., HAINES, L.R., HARDIE, D.B., OLAFSON, R.W. AND PEARSON, T.W. (2004). Mass spectrometric quantitation of peptides and proteins using

Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *Journal of Proteome Research* **3** (2), 235–244.

BARNIDGE, D.R., DRATZ, E.A., MARTIN, T., BONILLA, L.E., MORAN, L.B. AND LINDALL, A. (2003). Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. *Analytical Chemistry* **75** (3), 445–451.

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. *Analytical Chemistry* **74** (18), 4741–4749.

BONENFANT, D., SCHMELZLE, T., JACINTO, E. *ET AL.* (2003). Quantitation of changes in protein phosphorylation: a simple method based on stable isotope labelling and mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **100** (3), 880–885.

BOTTARI, P., AEBERSOLD, R., TURECEK, F. AND GELB, M.H. (2004). Design and synthesis of visible isotope-coded affinity tags for the absolute quantification of specific proteins in complex mixtures. *Bioconjugate Chemistry* **15** (2), 380–388.

CHELIUS, D. AND BONDARENKO, P.V. (2002). Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *Journal of Proteome Research* **1** (4), 317–323.

DURR, E., YU, J., KRASINSKA, K.M. *ET AL.* (2004). Direct proteomic mapping of the lung microvascular endothelial cell surface *in vivo* and in cell culture. *Nature Biotechnology* **22** (8), 985–992.

ENG, J., MCCORMACK, A.L. AND YATES, J.R.I. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of American Mass Spectrometry* **5**, 976–989.

GERBER, S.A., RUSH, J., STEMMAN, O., KIRSCHNER, M.W. AND GYGI, S.P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences of the United States of America* **100** (12), 6940–6945.

GOODLETT, D.R., KELLER, A., WATTS, J.D. *ET AL.* (2001). Differential stable isotope labelling of peptides for quantitation and *de novo* sequence derivation. *Rapid Communication Mass Spectrometry* **15** (14), 1214–1221.

GYGI, S.P., CORTHALS, G.L., ZHANG, Y., ROCHON, Y. AND AEBERSOLD, R. (2000). Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proceedings of the National Academy of Sciences of the United States of America* **97** (17), 9390–9395.

HARDWIDGE, P.R., RODRIGUEZ-ESCUDERO, I., GOODE, D. *ET AL.* (2004). Proteomic analysis of the intestinal epithelial cell response to enteropathogenic *Escherichia coli*. *Journal of Biological Chemistry* **279** (19), 20127–20136.

HOANG, V.M., CONRADS, T.P., VEENSTRA, T.D. *ET AL.* (2003). Quantitative proteomics employing primary amine affinity tags. *Journal of Biomolecular Technology* **14** (3), 216–223.

HUNTER, T. AND WASHBURN, M. (2003). Integration of chromatography and peptide mass modifications for quantitative proteomics. *Journal of Liquid Chromatography and Related Technologies* **26** (14), 2285–2301.

IBARROLA, N., KALUME, D.E., GRONBORG, M., IWAHORI, A. AND PANDEY, A. (2003). A proteomic approach for quantitation of phosphorylation using stable isotope labelling in cell culture. *Analytical Chemistry* **75** (22), 6043–6049.

JEPPSSON, J.O., KOBOLD, U., BARR, J. *ET AL.* (2002). Approved IFCC reference method for the measurement of HbA1c in human blood. *Clinical Chemistry and Laboratory Medicine* **40** (1), 78–89.

JI, Q.C., RODILA, R., GAGE, E.M. AND EL-SHOURBAGY, T.A. (2003). A strategy of plasma protein quantitation by selective reaction monitoring of an intact protein. *Analytical Chemistry* **75** (24), 7008–7014.

JOHNSON, K.L. AND MUDDIMAN, D.C. (2004). A method for calculating 16O/18O peptide ion ratios for the relative quantification of proteomes. *Journal of the American Society of Mass Spectrometry* **15** (4), 437–445.

JULKA, S. AND REGNIER, F. (2004). Quantification in proteomics through stable isotope coding: a review. *Journal of Proteome Research* **3** (3), 350–363.

KRIJGSVELD, J., KETTING, R.F., MAHMOUDI, T. *ET AL.* (2003). Metabolic labelling of *C. elegans* and *D. melanogaster* for quantitative proteomics. *Nature Biotechnology* **21** (8), 927–931.

KUHN, E., WU, J., KARL, J., LIAO, H., ZOLG, W. AND GUILD, B. (2004). Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and 13C-labelled peptide standards. *Proteomics* **4** (4), 1175–1186.

LI, X.J., ZHANG, H., RANISH, J.A. AND AEBERSOLD, R. (2003). Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. *Analytical Chemistry* **75** (23), 6648–6657.

LIU, H., SADYGOV, R.G. AND YATES, J.R., 3RD (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Analytical Chemistry* **76** (14), 4193–4201.

LU, Y., BOTTARI, P., TURECEK, F., AEBERSOLD, R. AND GELB, M.H. (2004). Absolute quantification of specific proteins in complex mixtures using visible isotope-coded affinity tags. *Analytical Chemistry* **76** (14), 4104–4111.

MACCOSS, M.J., WU, C.C., LIU, H., SADYGOV, R. AND YATES, J.R., 3RD (2003). A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. *Analytical Chemistry* **75** (24), 6912–6921.

MARLEY, J., LU, M. AND BRACKEN, C. (2001). A method for efficient isotopic labelling of recombinant proteins. *Journal of Biomolecular NMR* **20** (1), 71–75.

MUNCHBACH, M., QUADRONI, M., MIOTTO, G. AND JAMES, P. (2000). Quantitation and facilitated *de novo* sequencing of proteins by isotopic N-terminal labelling of peptides with a fragmentation-directing moiety. *Analytical Chemistry* **72** (17), 4047–4057.

NAGELE, E., VOLLMER, M. AND HORTH, P. (2003). Two-dimensional nano-liquid chromatography-mass spectrometry system for applications in proteomics. *Journal of Chromatography A* **1009** (1–2), 197–205.

NAGELE, E., VOLLMER, M. AND HORTH, P. (2004). Improved 2D nano-LC/MS for proteomics applications: a comparative analysis using yeast proteome. *Journal of Biomolecular Techniques* **15** (2), 134–143.

NAKANISHI, T., IGUCHI, K. AND SHIMIZU, A. (2003). Method for haemoglobin A(1c) measurement based on peptide analysis by electrospray ionization mass spectrometry with deuterium-labelled synthetic peptides as internal standards. *Clinical Chemistry* **49** (5), 829–831.

PATTON, W.F. (1999). Proteome analysis. II. Protein subcellular redistribution: linking physiology to genomics via the proteome and separation technologies involved. *Journal of Chromatography B Biomedical Science Applications* **722** (1–2), 203–223.

PATTON, W.F., SCHULENBERG, B. AND STEINBERG, T.H. (2002). Two-dimensional gel electrophoresis; better than a poke in the ICAT? *Current Opinion in Biotechnology* **13** (4), 321–328.

PENG, J., KIM, M.J., CHENG, D., DUONG, D.M., GYGI, S.P. AND SHENG, M. (2004). Semiquantitative proteomic analysis of rat forebrain postsynaptic density fractions by mass spectrometry. *Journal of Biological Chemistry* **279** (20), 21003–21011.

PERKINS, D.N., PAPPIN, D.J., CREASY, D.M. AND COTTRELL, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20** (18), 3551–3567.

RANISH, J.A., YI, E.C., LESLIE, D.M. *ET AL.* (2003). The study of macromolecular complexes by quantitative proteomics. *Nature Genetics* **33** (3), 349–355.

SHAW, M.M. AND RIEDERER, B.M. (2003). Sample preparation for two-dimensional gel electrophoresis. *Proteomics* **3** (8), 1408–1417.

TABB, D.L., MCDONALD, W.H. AND YATES, J.R., 3RD (2002). DTASelect and contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *Journal of Proteome Research* **1** (1), 21–26.

VON HALLER, P.D., YI, E., DONOHOE, S. *ET AL.* (2003). The application of new software tools to quantitative protein profiling via isotope coded affinity tag (ICAT) and tandem mass

spectrometry: I. Statistically annotated datasets for peptide sequences and proteins identified via the application of ICAT and tandem mass spectrometry to proteins copurifying with T cell lipid rafts. *Molecular Cell Proteomics* **2** (7), 426–427.

WANG, W., ZHOU, H., LIN, H. *ET AL.* (2003). Quantification of proteins and metabolites by mass spectrometry without isotopic labelling or spiked standards. *Analytical Chemistry* **75** (18), 4818–4826.

WASHBURN, M.P., WOLTERS, D. AND YATES, J.R., 3RD (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology* **19** (3), 242–247.

WASHBURN, M.P., ULASZEK, R., DECIU, C., SCHIELTZ, D.M. AND YATES, J.R., 3RD (2002). Analysis of quantitative proteomic data generated via multidimensional protein identification technology. *Analytical Chemistry* **74** (7), 1650–1657.

WHETSTONE, P.A., BUTLIN, N.G., CORNEILLIE, T.M. AND MEARES, C.F. (2004). Element-coded affinity tags for peptides and proteins. *Bioconjugate Chemistry* **15** (1), 3–6.

WU, C.C., MACCOSS, M.J., HOWELL, K.E., MATTHEWS, D.E. AND YATES, J.R.I. (2004). Metabolic labelling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Analytical Chemistry* **76** (17), 4951–4959.

YAN, W., LEE, H., DEUTSCH, E.W. *ET AL.* (2004). A dataset of human liver proteins identified by protein profiling via isotope coded affinity tag (ICAT) and tandem mass spectrometry. *Molecular and Cellular Proteomics* **3** (10), 1039–1041.

YATES, J.R., 3RD (2004). Mass spectral analysis in proteomics. *Annual Review of Biophysical and Biomolecular Structure* **33**, 297–316.

ZHOU, H., RANISH, J.A., WATTS, J.D. AND AEBERSOLD, R. (2002). Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. *Nature Biotechnology* **20** (5), 512–515.