ELSEVIER

Review

A review of quantitative methods for proteomic studies[☆]

Catherine Fenselau[*]

Department of Chemistry & Biochemistry, and Marlene & Stewart Greenebaum Cancer Center,
University of Maryland, College Park, MD 20742, USA

Abstract

An overview is provided of six strategies for relative or absolute quantitation of protein abundances that are widely used in proteomic studies. Strengths and limitations are discussed. Four of these involve stable isotope labeling and isotope ratio measurements by mass spectrometry. In another, mass spectra are used to deconvolute overlapping peptide HPLC peaks to provide relative quantitation based on peak areas. The sixth provides relative abundances of proteins based on 2-D gel arrays. It should be noted that these strategies measure peptide and protein abundances, and cannot directly assess changes in regulation or expression.
© 2006 Elsevier B.V. All rights reserved.

Contents

1. Introduction

This paper is submitted in honor of the 31st annual meeting of the Japanese Society for Medical Mass Spectrometry. The Japanese Society for Medical Mass spectrometry has a long history of interest in biomarker discovery and in the analysis of biomarkers in clinical diagnosis. One of the important criteria for biomarker discovery is the quantitative comparison of relative and absolute amounts of chemical markers in samples of urine, blood, or other samples, from well and sick subjects. In recent years, interest has been renewed in the use of proteins as biomarkers for diseases and for treatment efficacy. This review discusses the comparative quantitation of proteins in biomarker searches that use proteomic strategies. Six methods will be described, and three will be illustrated from our own studies of changes in protein abundances in cancer cells that have become resistant to chemotherapeutic agents. It should be emphasized that these methods do not measure changes in expression or regulation, but simply changes in protein abundances. Such abundances reflect synthesis, degradation and modification. These methods include:
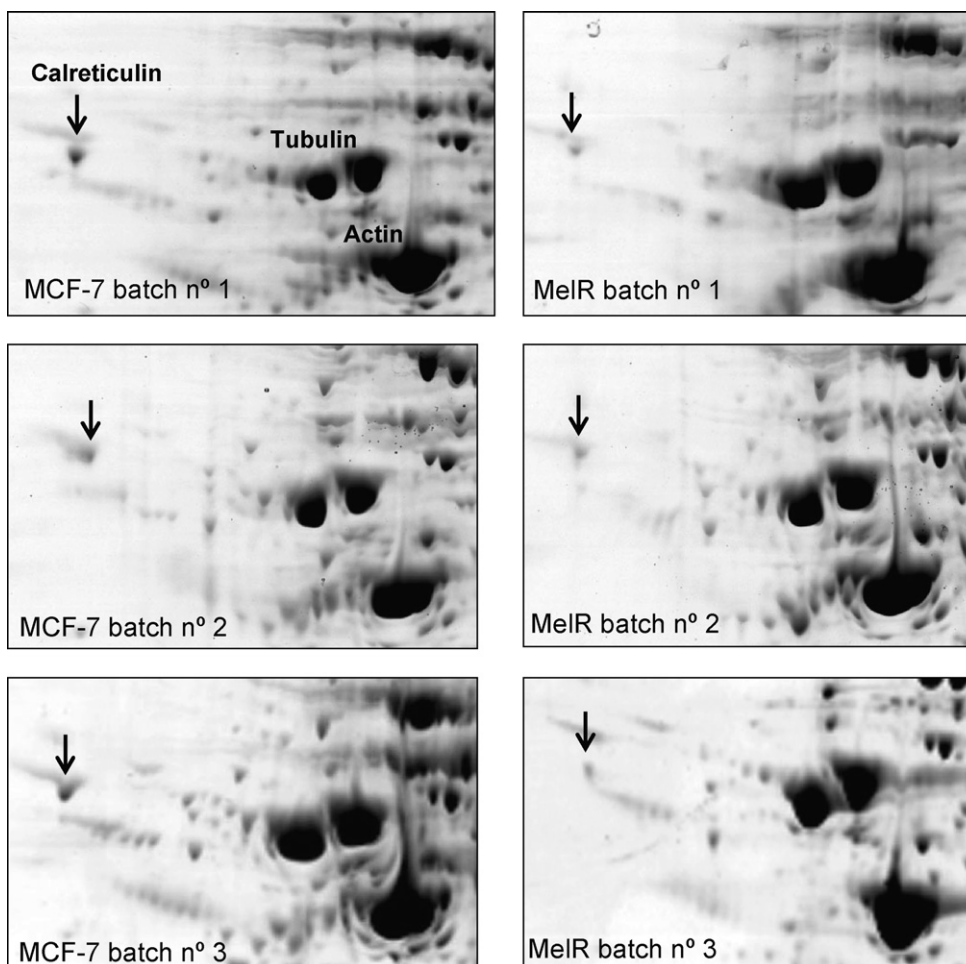
Fig. 1. Replicate arrays obtained by two dimensional gel electrophoresis, Coomassie blue staining and computer-supported comparative densitometry showing the reduced abundance of calreticulin in MCF-7 cells selected for resistance to melphalan [2]. Reproduced with permission from the American Chemical society.

- Gel-based densitometry/fluorescence.
- Metabolic isotope incorporation.
- Enzyme-catalyzed isotope incorporation.
- Isotope labeling by chemical derivatization.
- Addition of synthetic internal standards.
- Areas of corresponding deconvoluted HPLC peaks.

1.1. Densitometric comparison of protein abundances in two dimensional gel arrays

Two-dimensional gel electrophoresis continues to be an important and widely used technique for protein separation [1]. Protein mixtures are separated based on isoelectric point and molecular mass. Proteins are usually detected with silver stain, with Coomassie blue, or with fluorescent tags. These arrays or images may be digitized and aligned digitally, and pairs of gels may be compared, spot by spot to reveal changes in abundances. There are many software programs available commercially to carry out the alignment and automated quantitative comparison. Fig. 1 shows partial images from three pairs of 2-D gel arrays of cytosolic proteins recovered from human breast cancer MCF-7 cells susceptible and resistant to the anti-cancer alkylating agent, melphalan. The entire experiment was carried out in triplicate and the average ratio of calreticulin abundance in MCF-7 cancer cells resistant to melphalan and susceptible to melphalan was found to be 0.4 ± 0.1 [2]. With the advent of modern proteomics, the quality and variety of commercially available immobilized pH gradient (IPG) strips and gels has improved greatly. These separations still require considerable manual manipulation, however they provide the separation of several thousand soluble proteins with resolution as yet unequaled by other methods of protein separation. Two dimensional gels also provide visualization of samples with attendant advantages for comparison and record keeping. The outstanding advantage offered by 2-D gel electrophoresis and mass spectrometry for protein analysis is the ready recognition of isoforms, currently unparalleled by any other technique.

1.2. Metabolic labeling

Metabolic labeling is another method adapted for proteomics from classical protein chemistry. Nutrients highly enriched with stable isotopes are provided to cells in culture, and isotope labels are incorporated into all cellular proteins simultaneously. This method provides differential labeling of multiple pools of proteins. Cultured cells are combined and relative abundances of

individual proteins are determined by measuring isotope ratios of peptide pairs of sets with mass spectrometry. The method was introduced for bacterial proteomics by Oda et al. [3]. Subsequent publications extended the method to eukaryotic cells [4,5]. Ong et al. have proposed the name SILAC, stable isotope labeling by amino acids in cell culture [6]. The great strength of this method is that pools of labeled proteins can be mixed even before the cells are lysed, eliminating differential protein loss during lysis, extraction, and fractionation, before the variously labeled proteins are combined. Once the samples are combined, protein losses, e.g., sticking on the sides of containers, are expected to be proportional. The major limitation of metabolic labeling is that it is not readily extended to clinical or animal studies.

When this laboratory used metabolic labeling to study changes in proteins in the plasma membrane [7] and the nucleus [8] of human breast cancer MCF-7 cells, we found that they did not grow in the dialyzed serum required for the introduction of labeled amino acids. These estrogen positive cells require small

molecule hormones, such as estrogen, which are removed by dialysis. We were able to grow the cells when a small amount of whole serum was retained in the culture medium. This compromise limits the extent to which isotopically labeled amino acids can be incorporated into proteins, however the baseline incorporation can be determined and a correction made to isotope ratios measured between pools. The combined mixture of nuclear soluble proteins was fractionated by 2-D gel electrophoresis, and isotope ratios of peptides from the spots of interest were measured by mass spectrometry. Fig. 2 shows examples of measurements of the differential abundances of three proteins isolated from the nucleus of MCF-7 cells susceptible and resistant to mitoxantrone [8]. These illustrate a protein whose abundance is not changed, and examples where abundance is increased and decreased in the drug resistant line. These ratios are derived from both comparative densitometry and isotope ratios. The abundance ratios of these and other nuclear proteins are shown in Table 1, along with standard deviations.
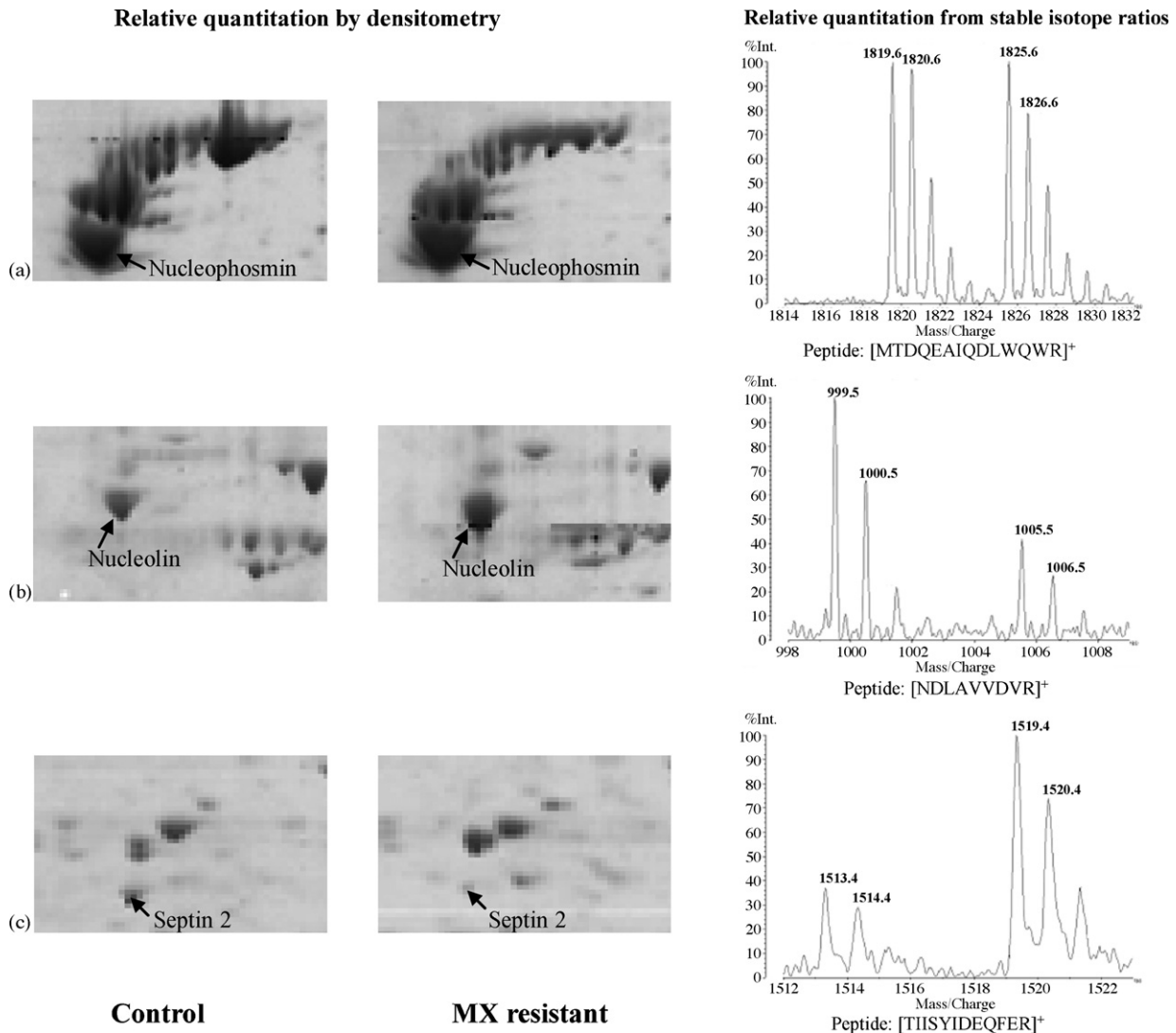

Fig. 2. Changes in abundances of three soluble nuclear proteins determined by (left) two dimensional gel electrophoresis and (right) mass spectrometric determination of isotope ratios of peptides from proteins labeled by metabolic labeling. Adapted from reference [8].

Table 1
Abundance changes in soluble nuclear proteins from mitoxantrone-resistant MCF-7 cells determined using isotope ratios and gel densitometry

Protein name	Isotope ratio MXres/control	Densitometry MXres/control
78 K GRP	4.91 ± 0.46	5.53 ± 0.69
Prohibitin	3.69 ± 0.33	2.81 ± 0.31
HMG-1	2.56 ± 0.18	1.52 ± 0.16
40S ribosomal protein SA	2.35 ± 0.18	1.74 ± 0.19
Cyclophilin B	2.22 ± 0.16	1.66 ± 0.45
Nucleolin (76 K)	2.18 ± 0.16	1.98 ± 0.24
Nucleolin (100 K)	1.58 ± 0.18	1.47 ± 0.26
Cytokeratin 19	0.53 ± 0.09	0.58 ± 0.14
Cytokeratin 8 (truncated)	0.49 ± 0.05	0.43 ± 0.06
PARP-1	0.46 ± 0.06	0.30 ± 0.13
EF-1-beta	0.36 ± 0.06	0.28 ± 0.09
Septin 2	0.11 ± 0.03	0.22 ± 0.03

Ratios are averaged from three harvests and from three to five gels for each of three harvests [8].

1.3. Proteolytic labeling with O-18

We have introduced a method for global quantitative comparisons of abundances of pairs of proteins, which is applicable to samples from animals and clinical specimens [9,10]. Two atoms of O-18 are introduced into the carboxylic acid group of every proteolytic peptide in a protein pool. This set of peptides can be mixed with peptides labeled with O-16, and isotope ratios can be determined by mass spectrometry. The introduction of the label is catalyzed by members of the serine protease family, which include trypsin [11], Glu-C protease [12], Lys-C protease, and (slow) chymotrypsin [10]. In the binding site of these proteases, the residue of choice is covalently bound in a tetrahedral intermediate, which is then disrupted by nucleophilic attack by a water molecule. Fig. 3 illustrates that disruption by a molecule of $H_2^{18}O$ as it incorporates an atom of O-18 into the carboxyl terminus of the peptide. The initial proteolysis can be carried out in $H_2^{18}O$, thus introducing the first atom of O-18 during protein cleavage. The C-terminal residue in each peptide product is still recognized by the protease, e.g., glutamate in the case of Glu-C, and the peptides are bound repeatedly via the same covalent mechanism. If the peptide products are incubated with the catalytic enzyme in $H_2^{18}O$, microreversibility will eventually equilibrate the level of O-18 in the peptides with the level of O-18 in the solvent, preferably >95% [10].

In fact, we argue that there are significant advantages to be gained by separating the proteolysis and labeling steps. These include:
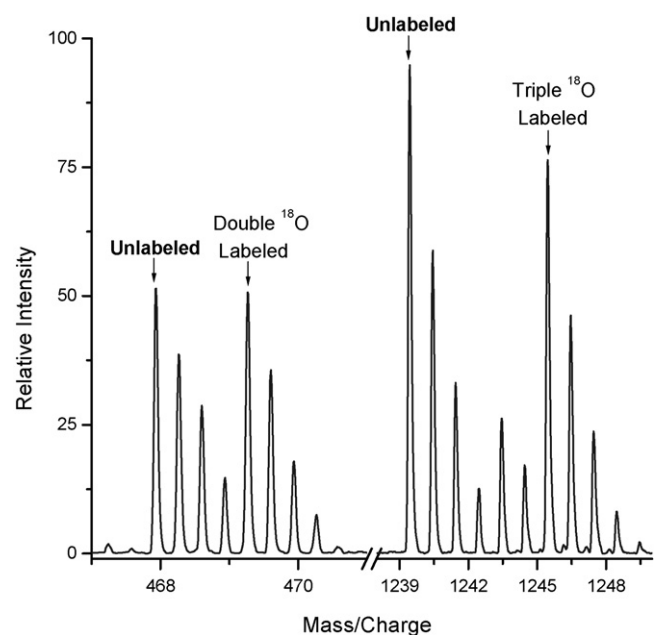

Fig. 4. Partial electrospray mass spectrum of O-18 and O-16 labeled peptides from riboflavin binding protein digested by endoproteinase Glu-C and *N*-glycopeptidase F. A triply charged peptide is detected with a 4 Dalton increment, and a singly charged peptide is detected with a 6 Dalton increment. The latter peptide had been *N*-glycosylated [12]. Reprinted with permission from the American Chemical Society.

- Each operation, proteolysis and exchange, can be carried out under optimal conditions, time, detergents, pH, etc.
- Peptides, not proteins, are dried and redissolved in $H_2^{18}O$. Proteins are often hard to dry and harder to redissolve.
- Less of the expensive $H_2^{18}O$ is required.

There are several advantages to be realized in the use of O-18 labels.

- Every peptide (except the original C-terminus of the protein) is labeled.
- The labeling strategy may be combined with any affinity method for peptide fractionation.
- Incubation of labeled peptides with *N*-glycopeptidase F in $H_2^{18}O$ introduces a third O-18 label at positions where *N*-linked carbohydrates have been released [12]. This is illustrated in Fig. 4.
- The presence of isotope labels in the carboxyl termini of peptides facilitates the assignment of 'y' ions in spectra obtained by tandem mass spectrometry (MSMS) experiments.
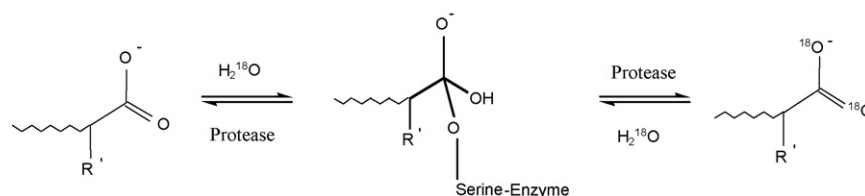

Fig. 3. Scheme illustrating the incorporation of two atoms of O-18 by reversible binding of peptides by members of the serine protease family.
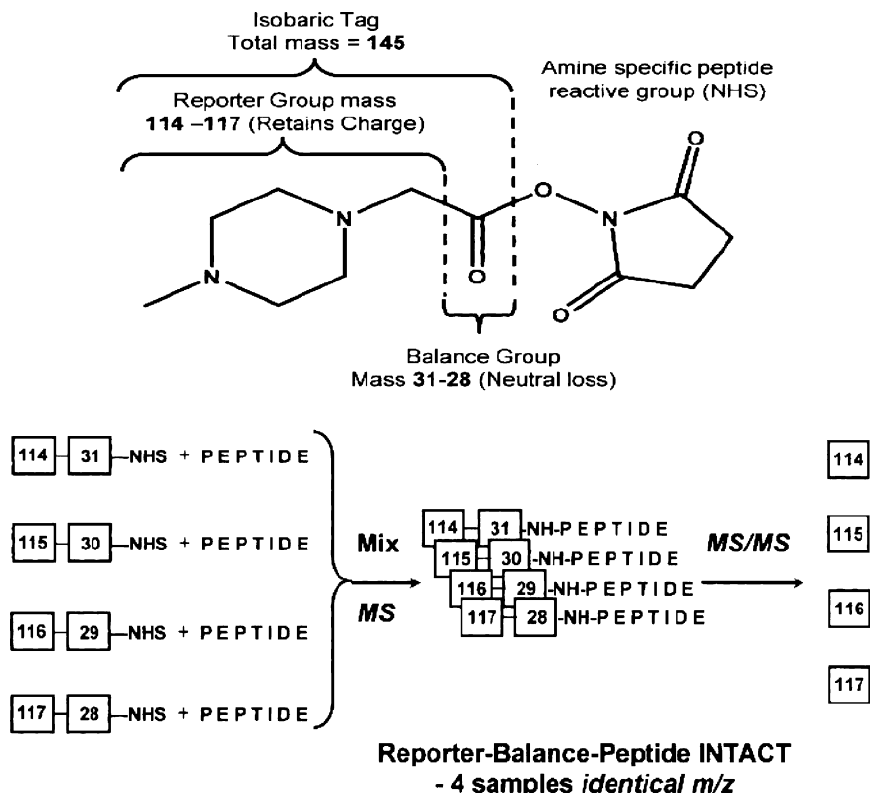
Fig. 5. Scheme for the multiplexed isobaric tagging strategy. The structure of methylpiperazine acetic acid *N*-hydroxy succinimide ester (top) is segmented to indicate the regions that carry isotope labels. In the reaction scheme at the bottom, the isobaric labeled reagents are reacted with proteins or peptides from four different samples, mixed, and analyzed by tandem mass spectrometry. Adapted from reference [19].

- The use of immobilized trypsin or other protease allows the use of a high concentration of enzyme and thus provides more efficient exchange.
- The use of immobilized protease allows the enzyme to be removed readily to prevent back exchange [13,14].
- The secondary product is water, rather than a contaminating chemical. This increases sensitivity [15].

This exchange strategy has the disadvantage that the label is not introduced until after the proteins have been cleaved to peptides. Ideally, manipulation of the sample should be minimized until the label is introduced. Thus, the bulk of the fractionation should be carried out at the peptide level [16]. Software packages have been developed to provide computer-supported recognition of isotope pairs and automated determination of their ratios.

1.4. Isotope incorporation by chemical derivatization

Many strategies have been proposed to introduce isotopes by chemical reactions at either the protein or the peptide level. Most have been given acronyms, such as isotope coded affinity tags (ICAT) [17], global internal standard technology (GIST) [18], nitrobenzene succinyl reagent (NBS) [19] and isobaric tags for relative and absolute quantitation (iTRAQ) [20]. Reagents for several of these strategies are patented, only available commercially, and expensive. The design of the iTRAQ reagents allows quantitative comparison of four (or more) protein pools simul-

taneously [20]. In the scheme in Fig. 5 the reagent derivatizes primary amine groups in four proteins or peptides with four different derivatives, which all have the same total mass. The four derivatized peptide ions are designed to fragment under collisional activation in an MSMS experiment to produce four product ions with differing weights. These product ions all weigh less than 120 Daltons, as shown in the spectra in Fig. 6, and are intended to fall into a clear spot in the MSMS spectra of peptides.

There are a number of criteria for selecting a chemical isotope labeling strategy, whose relative importance will vary, according to the analytical objective. In addition to cost, ease and speed of introduction, the effect on chromatographic separation, the number of peptides available from each protein, and the mass difference between the labeled and unlabeled peptides may be important. Thought should be given to possible changes in hydrophilicity and desorption sensitivity. Procedures involving chemical reactions should go to completion without introducing chemical contamination. It is desirable that the reaction be applicable to samples of small size, and applicable in a high throughput mode. In all cases, the investigator should strive for 100% incorporation of labels, and strategies that facilitate this are desirable. Similarly, strategies are preferable, which permit the parallel samples to be combined early in the process of isolation and fractionation.

It should be pointed out that the precision of the measurement of the isotope ratio and the dynamic range available for the ratio will largely be determined by the mass spectrometer, and
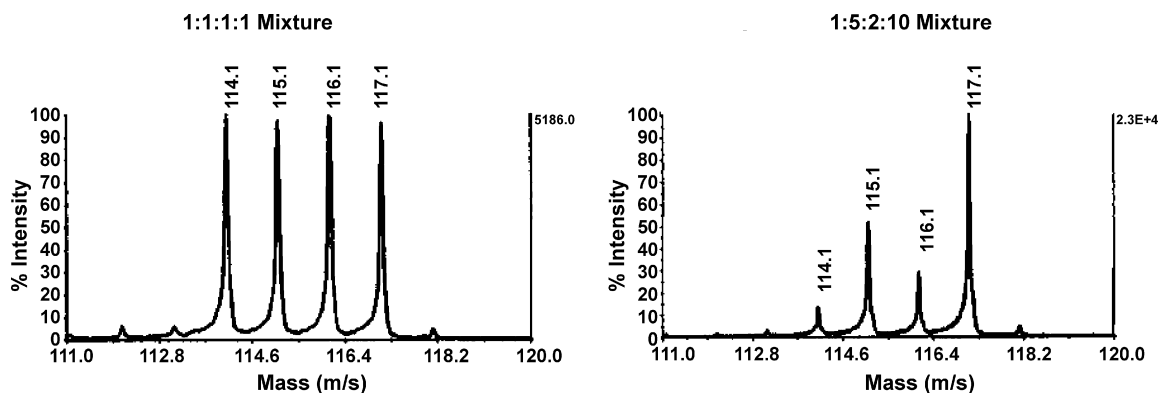
Fig. 6. Partial MSMS spectra of peptides from (left) an equimolar mixture of four protein samples, and (right) a 1:5:2:10 mixture of four protein samples, each labeled with a different isotopically enriched derivative of methylpiperazine acetic acid *N*-hydroxy succinimide ester [20]. Reproduced with permission from the American Society for Biochemistry and Molecular biology.

not by the isotopes or labeling strategy used [21]. The slower scanning speeds available with laser desorption, compared to HPLC-electrospray and nanospray, provide better ion statistics and thus more accurate measurements of isotope ratios. Isotope profiles may be better envisioned when higher resolution is used, e.g., with Fourier transform mass spectrometry. However, relative quantitation is often as accurate when lower resolution analyzers are used.

1.5. Synthetic internal standards

The addition of synthetic internal standards has been used for nearly 40 years for quantitation with mass spectrometry, and is compatible with both LC-ES-MS and MALDI-MS for proteomic analyses. In the most common application to proteomics, peptides are selected, based on uniqueness, stability, sensitivity to ionization techniques, etc., as surrogate markers for proteins of interest, and internal standards are added to allow these proteotypic peptides to be quantified. The best internal standards are peptides that have identical sequences as the biomarker peptides, and carry stable isotopic labels. The isotopes change the mass, but not the chemical behavior. Isotope ratios of spiked peptides and peptides obtained from protein biomarkers are measured by mass spectrometry to provide a good estimate of the absolute quantitation of the endogenous protein. This strategy to provide absolute quantitation has engendered the commercial acronym AQUA [22]. The use of internal standards allows multiple peptides (and thus proteins) to be quantified in a single sample, and facile quantitation of a particular peptide in multiple samples, e.g., at multiple time points. The approach is applicable to quantitation of proteins in tissue samples as well as proteins in solution [23]. Peptides with homologous sequences (and no isotope labels) can also be used as internal standards, supported by standard curves.

Isotope-labeled proteins provide even better internal standards, because they undergo fractionation and digestion along with the biomarker proteins, and provide correction for losses that occur before peptides are produced. Such proteins may be provided from normal or recombinant cells grown in labeled media.

1.6. Comparison of areas of deconvoluted HPLC peaks

Computer algorithms are under development to deconvolute overlapping HPLC peaks in peptide mixtures, in order that peaks can be matched in analyses of multiple peptide pools and peak areas can be compared to provide relative quantitation without the use of isotope labels [24–26]. This strategy has worked very well for small molecules [27], and its capabilities and limits for complex peptide mixtures are currently being evaluated. The use of higher mass accuracy in mass spectrometry, and high resolution and reproducibility in the chromatography are expected to facilitate deconvolution and inter-run correspondence in complex peptide mixtures.

Acknowledgements

References

[1] A. Gorg, W. Weiss, M.J. Dunn, Proteomics 4 (2004) 3665, erratum in Proteomics 5 (2005) 826.
[2] Y. Hathout, K. Riordan, M. Gehrmann, C. Fenselau, J. Proteome Res. 1 (2002) 435.
[3] Y. Oda, K. Huang, F.R. Cross, D. Cowburn, B.T. Chait, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 6591.
[4] H. Jiang, A.M. English, J. Proteome Res. 1 (2002) 345.
[5] S. Gu, S. Pan, E.M. Bradbury, X. Chen, Anal. Chem. 74 (2002) 5774.
[6] S.E. Ong, B. Blagoev, I. Kratchmarove, D.G. Kristensen, H. Steen, A. Pandey, M. Mann, Mol. Prot. 1 (2002) 376.
[7] A. Rahbar, C. Fenselau, J. Proteome Res. 4 (2005) 2148.
[8] Z. Fu, C. Fenselau, J. Proteome Res. 4 (2005) 1583.
[9] X. Yao, A. Freas, J. Ramirez, P. Demirev, C. Fenselau, Anal. Chem. 74 (2001) 2529, erratum in Anal. Chem. 76 (2004) 2675.
[10] X. Yao, C. Afonso, C. Fenselau, J. Proteome Res. 2 (2003) 147.
[11] I.I. Stewart, T. Thomson, D. Figeys, Rapid Commun. Mass Spectrom. 15 (2001) 2456.
[12] K.J. Reynolds, X. Yao, C. Fenselau, J. Proteome Res. 1 (2002) 27.
[13] K.J. Reynolds, C. Fenselau, in: D.W. Speicher (Ed.), Current Protocols in Protein Science, John Wiley & Sons, New York, NY, 2003, p. 2341.

[14] C. Fenselau, X. Yao, in: S. Sechi (Ed.), Quantitative Proteomics, Humana Press Inc., 2006.

[15] L. Zang, D.P. Toy, W.S. Hancock, D.C. Sgroi, B.L. Karger, J. Proteome Res. 3 (2004) 604.

[16] Y. An, Z. Fu, P. Gutierrez, C. Fenselau, J. Proteome Res. 4 (2005) 2126.

[17] S.P. Gigy, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nat. Biotechnol. 17 (1999) 994.

[18] A. Chakraborty, F.E. Regnier, J. Chromatogr. A 949 (2002) 173.

[19] E. Matsuo, C. Toda, M. Watanobe, N. Ojima, S. Izumi, K. Tanaka, S. Tsunasawa, O. Nishimura, Proteomics 6 (2006) 2042.

[20] P.L. Ross, Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, D.J. Pappin, Mol. Cell. Prot. 3 (2004) 1154.

[21] M.F. Lopez, Electrophoresis 21 (2000) 1082.

[22] O. Stemmann, H. Zou, S.A. Gerber, S.P. Gygi, M.W. Kirschner, Cell 107 (2001) 715.

[23] S.I. Hwang, J. Thumar, D.H. Lundgren, K. Rezaul, V. Mayya, I. Wu, J. Eng, M.E. Wright, D.K. Han, Oncogene 25 (2006), published on line June 26.

[24] W. Wang, H. Shou, H. Lin, S. Roy, T. Shaler, L. Hill, S. Norton, P. Kumar, M. Anderle, C.H. Becker, Anal. Chem. 75 (2003) 4818.

[25] T. Nakamura, N. Dohmae, K. Takio, Proteomics 4 (2004) 2558.

[26] D. Radulovic, S. Jelveh, S. Ryu, T.G. Hamilton, D. Foss, Y. Mao, A. Emili, Mol. Cell Prot. 3 (2004) 984.

[27] R.G. Dromey, M.J. Stefik, T.C. Rindfleisch, A.M. Duffield, Anal. Chem. 48 (1976) 1368.