Pseudo Internal Standard Approach for Label-Free Quantitative Proteomics

ARTICLE *in* ANALYTICAL CHEMISTRY · DECEMBER 2007

Impact Factor: 5.64 · DOI: 10.1021/ac701628m · Source: PubMed

4 AUTHORS, INCLUDING:

Junro Kuromitsu
Eisai Japan
39 PUBLICATIONS **1,694** CITATIONS
SEE PROFILE

Yoshiya Oda
Eisai Japan
107 PUBLICATIONS **6,192** CITATIONS
SEE PROFILE

Anal. Chem. **2007,** *79,* 8440−8445

Pseudo Internal Standard Approach for Label-Free Quantitative Proteomics

Tsuyoshi Tabata,[†,‡] Toshitaka Sato,[†] Junro Kuromitsu,[†] and Yoshiya Oda*[,†,‡]

Eisai Co., Ltd. Laboratory of Core Technology, Tokodai 5-1-3, Tsukuba, Ibaraki 300-2635, Japan, and Core Research for Evolutional Science and Technology, Japan Science and Technology Corporation, Saitama 332-0012, Japan

Quantitative proteome analysis has become a versatile tool to understand biological functions. Although stable isotope labeling is the most reliable method for quantitative mass spectrometry, preparation of isotope-labeled compounds is time-consuming and expensive. Simple label-free approaches have been introduced, but intensity-based quantitation without standards is not generally accepted as reliable, especially for small molecules. We have developed a novel label-free quantitative proteome analysis using pseudo internal standards (PISs). This idea was derived from northern blotting analysis, in which housekeeping genes are used as standards to normalize and compare target gene expression levels in different samples. In many proteomics studies, most proteins do not change their expression levels under different conditions, and therefore, these proteins can be employed as pseudo internal standards. This new approach is simple and does not require additional standards or labeling reagents. The PIS method represents a novel approach for mass spectrometry-based comprehensive quantitatitation and may also be applicable to quantitative metabolome analysis.

Nanoscale liquid chromatography/mass spectrometry (LC/MS) with electrospray ionization has become an indispensable tool for high-throughput peptide identification with high sensitivity in proteomics,[5−7] as well as for analysis of small molecules, such as pharmaceuticals.[8,9] However, in proteomics, unreliable analytical data can lead to an incorrect interpretation of protein functions. Label-free or standard-free LC/MS based on ion intensities is therefore problematic, because the results are influenced by many variables. For example, matrix effects can result in ion suppression due to the presence of unknown/unexpected components in the sample that coelute with target analytes,[10] and consequently, the response of an analyte standard in pure solvent can differ significantly from that in matrix samples. The matrix may suppress or enhance ionization of analytes, which would lead to a decreased or an increased MS response, respectively, and the level of the effect in the same matrix may vary depending on the amount/content of the matrix and the nature of the target. One possible approach to address this problem is to clean up samples from biological sources prior to LC/MS analysis in order to reduce the amount of matrix components entering the MS at the same time as the analyte.[11,12] However, extensive sample cleanup or analyte extraction reduces reproducibility among samples and entails the risk of loss of analytes. Chromatographic separation is another important process for reducing matrix interference, by removing components that would coelute in LC/MS from the original biological sample. However, in proteomics, each sample contains a huge number of analytes, and many may coelute or overlap with a target peptide in chromatography, so that complete separation is not practical for highly complex samples. Further, the retention order is not necessarily consistent in LC/MS analyses. Although matrix effects may be compensated by addition of internal standards to each sample, several problems remain: the processing time is significantly increased, matrix effects can strongly depend upon the chromatographic retention time, and plural internal standards may be needed for multiple target analytes.[13] Thus, the internal standard should have a retention time very close to that of the analyte.[12] Stable isotope-labeled internal standards are well-suited for this purpose. An internal standard that is structurally analogous to the analytical target is expected to experience a matrix effect very similar to that of the target in any batch of biological samples, so that the accuracy/precision of the method is largely unaffected by the matrix effect, or by variations in sample preparation. Metabolic labeling of the proteome is probably the gold standard for reliable quantitation in cultured cells,[14,15] and this approach has been applied to animal tissue proteome analysis.[16,17] Chemical tags with different molecular

* To whom correspondence should be addressed. E-mail: y-oda@hhc.eisai.co.jp.
† Eisai Co., Ltd. Laboratory of Core Technology.
‡ Core Research for Evolutional Science and Technology.

(1) Sechi, S.; Oda, Y. *Curr. Opin. Chem. Biol.* **2003**, *7*, 70−77.
(2) Ong, S. E.; Mann, M. *Nat. Chem. Biol.* **2005**, *1*, 252−262.
(3) Rinner, O.; Mueller, L. N.; Hubalek, M.; Muller, M.; Gstaiger, M.; Aebersold, R. *Nat. Biotechnol.* **2007**, *25*, 345−352.
(4) Fang, R.; Elias, D. A.; Monroe, M. E.; Shen, Y.; McIntosh, M.; Wang, P.; Goddard, C. D.; Callister, S. J.; Moore, R. J.; Gorby, Y. A.; Adkins, J. N.; Fredrickson, J. K.; Lipton, M. S.; Smith, R. D. *Mol. Cell. Proteomics* **2006**, *5*, 714−725.
(5) Wu, C. C.; Yates, J. R., III. *Nat. Biotechnol.* **2003**, *21*, 262−267.
(6) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198−207.
(7) Domon, B.; Aebersold, R. *Science* **2006**, *312*, 212−217.
(8) Kostiainen, R.; Kotiaho, T.; Kuuranne, T.; Auriola, S. *J. Mass Spectrom.* **2003**, *38*, 357−372.
(9) Maurer, H. H. *J. Mass Spectrom.* **2006**, *41*, 1399−1413.
(10) Jemal, M. *Biomed. Chromatogr.* **2000**, *14*, 422−429.
(11) Farre, M.; Petrovic, M.; Barcelo, D. *Anal. Bioanal. Chem.* **2007**, *387*, 1203−1214.
(12) Picó, Y.; Blasco, C.; Font, G. *Mass Spectrom. Rev.* **2004**, *23*, 45−85.
(13) King, R.; Bonfiglio, R.; Fernandez-Metzler, C.; Miller-Stein, C.; Olah, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 942−950.
(14) Oda, Y.; Huang, K.; Cross, F. R.; Cowburn, D.; Chait, B. T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6591−6596.

10.1021/ac701628m CCC: $37.00
Published on Web 10/12/2007

weights, with/without stable isotopes, also provide reliable quantitative results, especially in human clinical samples.[18,19] However, the drawbacks of these isotopically labeled analogues as internal standards are that extra steps are needed to prepare the labeled standards, and the cost is very high. Therefore, an unlabeled structural analogue, which may even be a type of compound quite different from the target, is often used as an internal standard to adjust recoveries during sample processing, as well as ion suppression effects in MS; it is spiked into the sample at the beginning of sample processing.[9−12,17] This type of internal standard is sometimes called a surrogate standard.[20−23] This strategy has been successfully used in a number of LC/MS assays for various classes of small molecules, including drugs. Signal suppression by matrix impurities still occurs, because the surrogate standard does not coelute with the target molecules, but this limitation may be overcome by employing multiple surrogate standards, because some of them may not be subject to the matrix effect in MS, or the matrix effects can be averaged.[11] Further, since multiple surrogate standards may normalize variations during sample processing, the relative recoveries of target and surrogate standards should be comparable, when the surrogate standards have physicochemical properties similar to the target molecule.[20−23] However, many types of peptides are analyzed in proteomics, and it is not easy to prepare multiple unlabeled surrogate standards for the highly complex proteome, because molecules with the same sequences as some standards may already be contained in target samples, and preparation/addition of large numbers of standards is time-consuming and sometimes not practical. An important point is that a constant amount of surrogate standards should be added to a constant amount/volume of biological samples at the beginning of sample processing. We had noticed that many proteome analyses show that most proteins do not change their expression levels under different conditions, such as disease state versus normal. Therefore, the relative amounts of these unchanged proteins to the original sample is expected to be constant among samples to be tested. We considered that these proteins with unchanged expression might be useful as surrogate standards, and we designated them as pseudo internal standards (PISs).

MATERIALS AND METHODS

Materials and Reagents. TPCK-treated, sequencing-grade modified trypsin was obtained from Promega (Madison, WI). C18 Empore disk was from 3M (St. Paul, MN). Gels with a thickness of 1.0 mm (Tris-HCl, 5−20% T) were obtained from DRC (Tokyo, Japan). All other reagents were of analytical grade, and water was obtained from a Milli-Q system (Millipore, Bedford, MA).

Neuro2a cells and HCT116 cells were grown in RPMI-1640 medium, supplemented with 10% fetal bovine serum plus antibiotics. For each experiment, one dish (5×10^7 cells/15-cm-diameter dish) of confluent cells was used. Cells were homogenized in ice-cold Tris buffer (including 100 mM Tris pH 8, protease inhibitors (Complete, Roche, Indianapolis, IN)), with a Teflon Potter-type homogenizer. Homogenates were cleared by centrifugation at 1500*g* for 10 min. Urea was dissolved in the supernatant to make an 8 M solution. Proteins were reduced with 20 mM dithiothreitol and then alkylated with 50 mM iodoacetamide. Sample solutions were treated with 5% (w/w) Lys-C (Wako Pure Chemicals) to digest proteins, diluted to 2 M urea in 0.1 M tris buffer, pH 8 (final concentration), and then further treated with trypsin.

Mouse Brain Sample Preparation. Whole brains of female APP/tau double-mutated (hemizygous) mice at eight weeks old were purchased from Taconic Co. (Gernabtown, NY), and forebrains of male Hip1 (huntingtin interacting protein 1) transgenic mice produced by the gene trap method at two weeks old were obtained from TransGenic Inc. (Kumamoto, Japan). The procedures for whole protein extraction from brain samples were the same as those described above for neuro2a/HCT116 cells. For iTRAQ analysis, iTRAQ kits were used according to the protocol recommended by Applied Biosystems (Foster City, CA) after obtaining tryptic peptides desalted with the C18 Empore disk.

Identification of Proteins Enriched by Affinity Columns. Proteins binding to E7070 analogue or amide analogue columns were enriched as described previously.[29]

In-Gel Digestion. After SDS-PAGE, target bands were excised, and in-gel digestion was carried out as described previously.[30]

Mass Spectrometric Analysis. Samples were redissolved in 10 μL of acetonitrile (ACN)/water/TFA = 5/95/0.1, and then 3-μL aliquots of samples were loaded onto the LC/MS system. The peptide fractions were separated on a 100 μm × 15 cm in-house-packed C-18 column with a 4−32% ACN gradient for 110 min in 0.2% acetic acid at a flow rate of 500 nL/min, using an Ultimate 3000 nanopump from Dionex LC Packings (San Francisco, CA) and an HTC-PAL autosampler (CTC Analytics AG, Zwingen, Switzerland) equipped with Valco C2 valves with 150-μm ports. The liquid chromatography (LC) eluate was coupled to an in-house-built nanospray source attached to a quadrupole ion trap mass spectrometer (model LTQ) from Thermo Electron Corp.

(15) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. *Mol. Cell. Proteomics* **2002**, *1*, 376−386.
(16) Wu, C. C.; MacCoss, M. J.; Howell, K. E.; Matthews, D. E.; Yates, J. R., III. *Anal. Chem.* **2004**, *76*, 4951−4959.
(17) Ishihama, Y.; Sato. T.; Tabata, T.; Miyamoto, N.; Sagane, K.; Nagasu, T.; Oda, Y. *Nat. Biotechnol.* **2005**, *23*, 617−621.
(18) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17*, 994−999.
(19) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. *Mol. Cell. Proteomics* **2004**, *3*, 1154−1169.
(20) Risha, K.; Flaherty, J.; Wille, R.; Buck, W.; Morandi, F.; Isemura, T. *Anal. Chem.* **2005**, *77*, 1503−1508.
(21) Ishii, Y.; Ogara, A.; Okamura, T.; Umemura, T.; Nishikawa, A.; Iwasaki, Y.; Ito, R.; Saito, K.; Hirose, M.; Nakazawa, H. *J. Pharm. Biomed. Anal.* **2007**, *43*, 1737−1743.
(22) Ternes, T.; Bonerz, M.; Schmidt, T. *J. Chromatogr., A* **2001**, *938*, 175−185.
(23) Löffler, D.; Ternes, T. A. *J. Chromatogr., A* **2003**, *1000*, 583−588.
(24) Lewis, J.; Dickson, D. W.; Lin, W.-L.; Chisholm, L.; Corral, A.; Jones, G.; Yen, S.-H.; Sahara, N.; Skipper, L.; Yager, D.; Eckman, C.; Hardy, J.; Hutton, M.; McGowan, E. *Science* **2001**, *293*, 1487−1491.
(25) Rappsilber, J.; Ryder, U.; Lamond, A. I.; Mann, M. *Genome Res.* **2002**, *12*, 1231−1245.
(26) Gay, S.; Binz, P. A.; Hochstrasser, D. F.; Appel, R. D. *Proteomics* **2002**, *2*, 1374−1391.
(27) Zybailov, B.; Coleman, M. K.; Florens, L.; Washburn, M. P. *Anal. Chem.* **2005**, *77*, 6218−6224.
(28) Ishihama, Y.; Oda, Y.; Tabata, T.; Sato, T.; Nagasu, T.; Rappsilber, J.; Mann, M. *Mol. Cell. Proteomics* **2005**, *4*, 1265−1272.
(29) Oda, Y.; Owa, T.; Sato, T.; Boucher, B.; Daniels, S.; Yamanaka, H.; Shinohara, Y.; Yokoi, A.; Kuromitsu, J.; Nagasu, T. *Anal. Chem.* **2003**, *75*, 2159−2165.
(30) Katayama, H.; Tabata, T.; Ishihama, Y.; Sato, T.; Oda, Y.; Nagasu, T. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2388−2394.

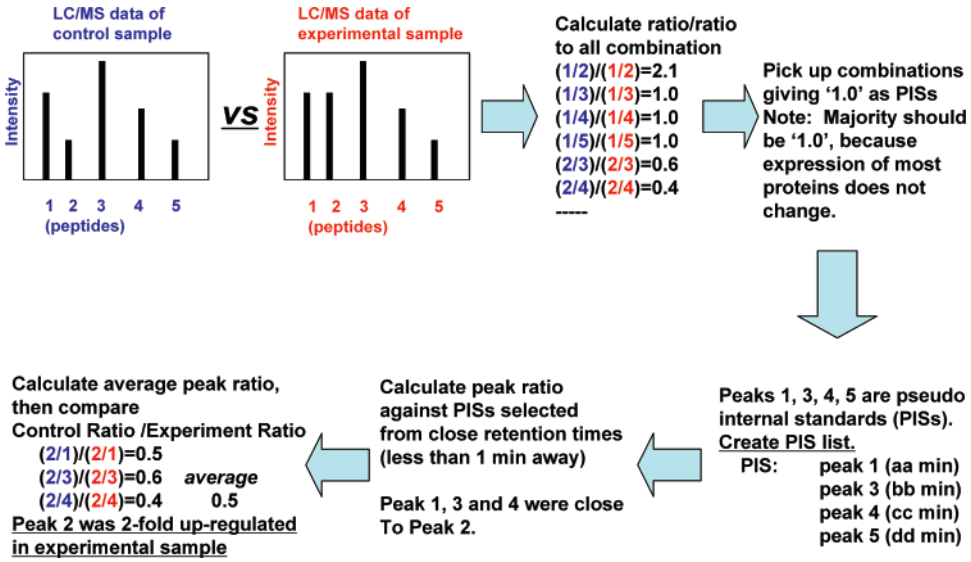

Figure 1. Scheme of PIS-based quantitation.

(San Jose, CA). The spray voltage was 2.1 kV. Peptides were analyzed in positive ion mode, and the MS scan range was 350−1500 m/z (profile mode). LC/MS spectra were collected in a data-dependent mode in which the highest-intensity peaks in each MS scan were chosen for collision-induced dissociation, and the isolation window was 3 Da (precursor m/z 1.5 Da) to select a precursor ion. The dynamic exclusion option, which prevented the same m/z from being selected for 1 min after its acquisition, was applied with a repeat count of 1, a repeat duration of 0.5 min, and an exclusion duration of 1 min.

Data Processing. MS/MS data were analyzed by X! Tandem (GPM, Manitoba, Canada) to identify peptides using the NCBI nonredundant (NCBI nr) *Mus musculus* database as of March 2007. Carbamidomethylation was treated as a fixed modification of Cys residues, whereas oxidation of Met residues was allowed as a variable modification. Peptide tolerances in MS and MS/MS modes for LTQ were 0.5 Da (before)/1.5 Da (after) and 0.4 Da, respectively. One missed cleavage of trypsin was allowed. Only peptides with clear mass spectra with a score of >95% reliability were considered in this study. SIEVE software (Thermo Electron) was used for retention time alignment, and peak intensities on mass chromatograms were obtained with Mass Navigator software (MKI, Tokyo, Japan). Peak ratio and the ratio of ratios for all combinations were calculated with in-house software, and manual confirmation was sometimes done to correct peak height. Peaks giving a ratio of 1.0 in any combination of pairs were assigned as PISs. Each peak was quantified relative to corresponding PIS peaks with less than 1-min retention time difference (this program is provides as Supporting Information (SI) 1).

RESULTS AND DISCUSSION

We have developed a label-free and standard-free quantitation approach, which we have named the PIS method, for mass spectrometry-based comprehensive proteomics analysis. The principle is shown in Figure 1. In the PIS method, relative peak intensities of analytes are calculated with respect to the PISs. Although candidate PISs are unknown before quantitative analysis, relative mass spectral peak ratios between PISs in two samples should be 1.0 (practically, 1.0 ± 0.2). For instance, if peptide peaks X, Y, and Z in the mass spectra are PIS candidates, the ratio of X/Y (or X/Z, Y/Z) in a control sample versus the ratio of X/Y (or X/Z, Y/Z) in an experimental sample should be 1.0. Therefore, calculation of all combinations of control peak ratio to experimental peak ratio identifies candidate PISs. This group will also include some peptides derived from different proteins that coincidentally show the same change in the experimental sample (e.g., 2-fold increase for both proteins). Therefore, after picking up PIS candidates with less than 1-min retention time difference from the target peak, we calculated the relative quantitative values between control and experimental samples, and it was confirmed that most proteins did not show changed expression levels in the experimental sample versus the control. Since the PIS method calculates a ratio of ratios to obtain relative quantitative values, it offers the advantage that normalization of sample amount (volume, number of cells, etc.) is not required. A further advantage is that no extra steps are required for the preparation/addition of standards, because PISs are inherently contained in every sample.

To confirm our new methodology using PISs, we prepared transgenic mice brains as well as their wild types, digested them with trypsin, repeated LC/MS analysis, and compared precision data, which is the variation among a series of measurements within an assay, obtained by the PIS method and by ion intensity-based quantitation. As a first step, we picked up 2129 peptides commonly identified from both a wild mouse brain and a transgenic one (SI 2). Among them, we selected 692 peptides as PISs (SI 3), and then precision data in experiment 1 as given in Table 1a were calculated, the parameters of which are mostly expressed as coefficient of variation (CV) (the detailed information is in SI 4−7). As a result, precision of the PIS method was slightly better than that of ion intensity-based quantitation when real biological samples were analyzed by LC/MS on the same day (Table 1a) and the PIS method was also better than the ion intensity-based quantitation when LC/MS analyses were run on the separate days (Table 1b and SI 9), though the CV values in this case (different day LC/MS analysis) were not as good as that of stable isotope-based quantitation (usually better than CV 20%). Next, a known amount of human serum albumin (HSA) was added to the

Table 1. Examples of Precision Data of PIS-Based and Ion Intensity-Based Quantitation

(a) Using Hip1 Transgenic Mouse Whole Brains[a,b]

	expt 1 (intraday 1)	expt 2 (intraday 2)	expt 3 (intraday 3)	expt 4 (intraday 4)	expt 1−4 (interday)
PIS based average CV (%)	7.996	7.612	7.400	6.901	19.50
ion intensity-based average CV (%)	8.065	8.974	8.516	14.25	28.71

(b) Using APP/tau Transgenic Mouse Whole Brains Obtained by Different Day LC/MS Analysis[c,d]

	expt 1	expt 2	expt 3
PIS-based average CV (%)	69.52	57.72	62.80
ion intensity-basedaverage CV (%)	89.14	96.73	87.97

[a] Each experiment was compared a wild mouse whole brain with a Hip1 transgenic mouse whole brain. Each experiment was $n = 3$ (wild) and $n = 3$ (transgenic), and LC/MS measurements were done on the same day. Experiments 1−4 were used different mice (total four wild and four transgenic mice). Total 2264 different peptides were identified, but only 213 peptides were always identified from all 24 LC/MS analyses ($n = 3 \times 2 \times 4 = 24$). [b] A total of 213 different sequences were quantified and evaluated. [c] Each experiment was compared a wild mouse whole brain with an APP/τ transgenic mouse whole brain. Each experiment was $n = 3$ (wild) and $n = 3$ (transgenic), and LC/MS measurements were done on separate days. Experiments 1−3 used different mice (total three wild and three transgenic mice). A total of 1073 different peptides were identified, but only 51 peptides were always identified from all 18 LC/MS analysis ($n = 3 \times 2 \times 3 = 18$). [d] A total of 51 different sequences were quantified and evaluated.

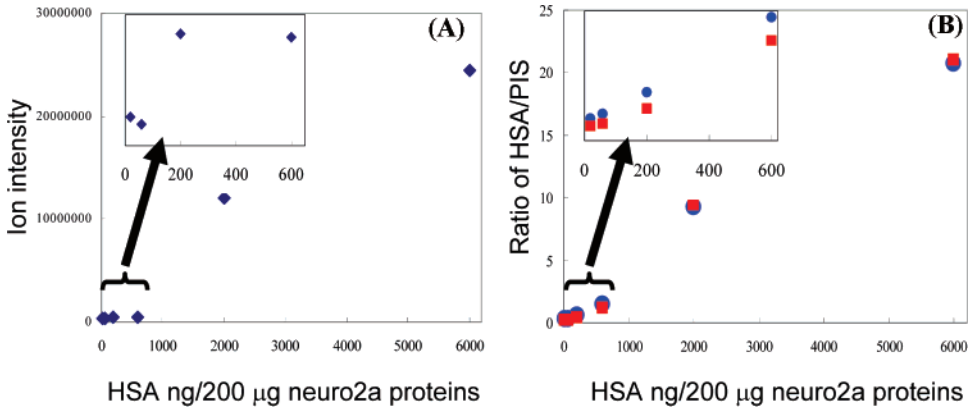

Figure 2. Linearity of (A) ion intensity-based and (B) PIS-based quantitation. Known amounts of HSA were added to cell lysates of neuro2a. After reduction and alkylation of cysteine residues, protein mixtures were digested with trypsin. Desalted samples were analyzed by LC/MS. Quantitation values were obtained from residues 473−481 (CCTESLVNR). (B) Red square represents single PIS (*m*/*z* 404)-based, and blue circle represents multiple PIS (six PISs: *m*/*z* 404, 485, 472, 395, 366, 642)-based quantitation.

neuro2A cells lysate. After tryptic digestion, samples were analyzed by LC/MS. Although the range of linearity was nearly the same for the PIS method (RSQ = 0.985) and ion intensity-based quantitation (RSQ = 0.964), the PIS method showed clearly better linearity in the lower concentration range (Figure 2). This is probably because low-intensity peaks are more subject to matrix effects. Also, multiple PISs was better than single PIS, because the reliability was increased by averaging of PIS intensities.

Accuracy as well as precision is one of the most important criteria in the assessment of an analytical method. Accuracy is the closeness of agreement to the true value of the analyte. In proteomics, target analytes are unknown until the data are analyzed by a search engine; therefore, it is not easy to evaluate accuracy of each analyte using real unknown biological samples. Since quantitation based on stable isotope-labeled internal standards is still the gold standard and probably shows the most accurate results in proteomics, we compared the quantitation data of iTRAQ[19] with that of the PIS method and the ion intensity-based approach by using whole brain samples from transgenic mice doubly mutated in the amyloid precursor protein (APP) and microtubule-associated protein tau genes. APP/tau double mutant transgenic mice carry the transgene coding for the 695-amino acid isoform of human Alzheimer β-amyloid (Aβ) protein, in addition to the transgene for the human P301L mutant of the tau gene, and develop β-amyloid plaques and neurofibrillary tangles in the brain and overt neuronal loss in selectively vulnerable brain limbic areas.[24] Coexpression of human mutant APP and tau in the mouse model enhanced both Aβ and tau pathology and nearly mimics Alzheimer's disease. After extraction of whole proteins, samples were digested with trypsin. The samples were divided into two parts, of which one was labeled with iTRAQ reagents before LC/MS analysis and the other was directly analyzed by LC/MS. The major innovation of the iTRAQ approach was to allow four- or eight-plex sample analysis at the same time, and therefore, we prepared duplicate samples (total of 4 different samples) and analyzed them at once. On the other hand, an important problem that may affect quantitation data is the use of low peaks in mass spectra (not MS/MS spectra), especially ion trap-type MS, which can obtain high-quality MS/MS spectra from very weak precursor ions. In this case, the signal of the peak will be overestimated,

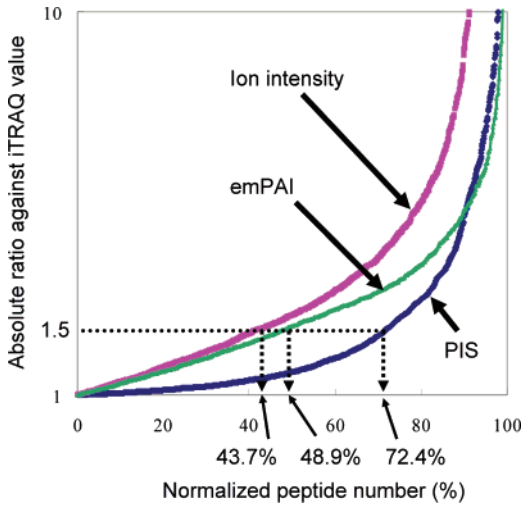

Figure 3. Accuracy data of label-free quantitation based on iTRAQ results: PIS/iTRAQ (blue line), Ion intensity/iTRAQ (pink line) and emPAI/iTRAQ (green line). 1810 different peptides were quantitated by PIS-based, ion intensity-based, and iTRAQ-based approaches. The emPAI result (total of 877 different proteins) was compared with iTRAQ at the protein level.

inevitably leading to underestimation of the analyte concentration in the sample. In order to minimize such problems, the PISs should be chosen from among the intense peaks. The iTRAQ method also provides a higher signal-to-noise ratio of target peaks (reporter ions in this case), because MS/MS spectra contain less noise peaks than MS spectra. We evaluated the relative fold-change values of 1810 different peptides using iTRAQ data and PIS (or ion intensity) data as the accuracy of PIS data. We found that 72.4% of PIS data showed less than 1.5-fold difference from the iTRAQ data, while only 43.7% of ion intensity-derived data met the same criterion, as shown in Figure 3 (also SI 9). Our results indicated that the PIS approach was reasonably consistent with iTRAQ (83.6% of the data showed less than 2-fold difference), suggesting that the method might be practically applicable. It is important to note that although iTRAQ would be more reliable in terms of quantitation, the number of identified peptides was one-fifth less than that with the PIS approach. iTRAQ cannot achieve 100% labeling efficiency/recovery, and PQD in LTQ was not suitable for identification, so we selected the triple play mode (precursor scan-CID-PQD), resulting in a smaller number of CID spectra and a smaller number of identifications. Indeed, we only identified 3441 different peptides using the iTRAQ approach, compared with 14 949 peptides using the PIS/ion intensity-based method. Protein abundance index (PAI)[25] or spectral counting[26,27] is an alternative approach to label-free quantitation, and emPAI is an improvement of PAI designed to obtain a linear relationship between PAI and protein concentration.[28] These label-free approaches do not require MS peak intensities, so quantitative

Table 2. Reproducibility of (A) PIS-Based Quantitation and (B) Ion Intensity-Based Quantitation after Affinity Chromatography, SDS-PAGE Separation, and LC/MS Analysis[a]

(A) specificity (peak intensity on E7070 column/ peak intensity on amide column)

gene name	expt 1	expt 2	expt 3	expt 4	expt 5	average	CV (%)
MDH1	64.33	69.80	51.68	58.47	36.82	56.22	22.72
RANBP5	14.49	46.59	10.31	nd[b]	nd	23.80	83.42
IPO9	10.40	28.64	nd	nd	nd	19.52	66.11
MTAP	21.69	21.70	nd	9.50	14.63	16.88	35.19
CSE1L	19.24	25.44	8.56	nd	12.79	16.51	44.84
CLTA	17.81	12.99	nd	11.77	21.94	16.13	28.95
VCL	12.63	nd	23.95	11.11	nd	15.90	44.12
TNPO1	nd	19.33	nd	nd	11.07	15.20	38.45
TUBB2A	15.17	12.61	nd	nd	nd	13.89	13.03
HSP90AA1	nd	19.28	nd	7.47	nd	13.37	62.45
PKM2	14.06	15.80	9.39	8.97	10.52	11.75	25.71
FASN	11.83	15.99	10.26	7.12	nd	11.30	32.66
average CV (%)							41.47%

(B) specificity (peak intensity on E7070 column/ peak intensity on amide column)

gene name	expt 1	expt 2	expt 3	expt 4	expt 5	average	CV (%)
TUBB2A	80.78	149.72	nd	nd	ND	115.25	42.30
MDH1	113.18	270.82	32.59	31.85	24.43	94.58	111.03
MTAP	112.00	217.15	nd	9.43	11.18	87.44	113.08
PKM2	58.59	166.77	15.34	10.27	7.26	51.65	130.98
RANBP5	78.06	45.04	12.30	nd	nd	45.13	72.86
CSE1L	104.14	42.84	14.99	nd	9.87	42.96	100.75
CLTA	61.26	24.54	nd	22.91	23.30	33.00	57.12
IPO9	29.43	31.12	nd	nd	nd	30.27	3.96
VCL	36.00	nd	32.13	20.22	nd	29.45	27.92
FASN	46.97	37.70	10.80	9.65	nd	26.28	72.02
TNPO1	nd	40.03	nd	nd	8.56	24.30	91.59
HSP90AA1	nd	28.94	nd	8.26	nd	18.60	78.65
average CV (%)							75.19

[a] Procedures were the same as those of the previous method. Affinity procedures were repeated 5 times repeated. The top 12 specific binders to E7070 are listed. [b] nd, not detect.

information can basically be obtained from the literature without MS raw data. We also compared the iTRAQ results with emPAI calculation and found that emPAI was better than ion intensity-based quantitation, but worse than PIS (Figure 3 and SI 10). A disadvantage of emPAI is its insensitivity to low-abundance proteins, which were identified from less than three peptides per protein. Since the MS/MS duty cycle is quite limited, identification of low-abundance proteins depends on chance; emPAI thus provides only a rough estimation, and changes of less than 5-fold should not be considered reliable.

Surrogate standards for small-molecule quantitation are also used to normalize fluctuations during sample processing. Thus, we examined whether the PIS method could pick up appropriate targets after complex sample processing. Cytosolic malate dehydrogenase (cMDH) selectively binds to an E7070 analogue (positive) affinity column, but not to an amide analogue (negative) column.[29] Whole cell lysates from HCT116 cells were applied to an affinity column with immobilized E7070 analogue or amide analogue. The column was washed, and then bound proteins were eluted and separated by SDS-PAGE. In-gel-digested peptides were analyzed by LC/MS. After having repeated these processes five times, relative binding amounts between the two columns were calculated by the PIS method, and it was found that cMDH was the most specific binder to the E7070 analogue column, as shown in Table 2 A. Based on the peak intensities in the mass spectra, cMDH was not clearly identified as a specific binder to the E7070 analogue column compared with the amide analogue column (Table 2B). In addition to the instability of absolute peak intensities in mass spectra, affinity purification steps are usually quite variable, as is SDS-PAGE followed by in-gel digestion; nevertheless, PIS methodology could still pick up the target.

Although it may be questioned whether peptides having different sequences and different retention times can play a role as internal standards, we have already shown that this is feasible in culture-derived isotope tag methodology. The quantitative precision did indeed become worse when different-sequence peptides were selected as internal standards instead of same-sequence isotope-labeled peptides, but the difference was only 1.29 $\pm$ 19.3% ($n = 155$), and normalization for "semiquantification" was confirmed to be possible by using different-sequence peptides.[17] The PIS method provides less reliable quantitative data than quantitation based on real internal standards, but it is simple and convenient for identifying specific binding proteins from protein mixtures, as well as for quantitative proteomics. The PIS method should also be applicable to other mass spectrometry-based comprehensive analyses, such as metabolomics.

ACKNOWLEDGMENT

This work was supported by funds from the Japan Science and Technology Agency.The authors declare that they have no competing financial interests.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review July 31, 2007. Accepted September 25, 2007.

AC701628M