technical notes Journal of proteome research

Improved Quantitative Analysis of Mass Spectrometry using Quadratic Equations

Joo Young Yoon,[†] Kyung Young Lim,[†] Sunho Lee,[†] Kunsoo Park,[*,†] Eunok Paek,[*,‡] Un-Beom Kang,[§] Jeonghun Yeom,[§,||] and Cheolju Lee[§,||]

School of Computer Science and Engineering, Seoul National University, Seoul, 151-742, Korea, Department of Mechanical and Information Engineering, University of Seoul, Seoul, 130-743, Korea, Life Sciences Division, Korea Institute of Science and Technology, Seoul, 136-791, Korea, and Department of Biomolecular Science, University of Science and Technology, Daejeon, 305-333, Korea

Received March 1, 2010

Abstract: Protein quantification is one of the principal computational problems in mass spectrometry (MS) based proteomics. For robust and trustworthy protein quantification, accurate peptide quantification must be preceded. In recent years, stable isotope labeling has become the most popular method for relative quantification of peptides. However, some stable isotope labeling methods may carry a critical problem, which is an overlap of isotopic clusters. If the mass difference between the light- and heavy-labeled peptides is very small, the overlap of their isotopic clusters becomes larger as the mass of original peptide increases. Here we propose a new algorithm for peptide quantification that separates overlapping isotopic clusters using quadratic equations. It can be easily applied in Trans-Proteomic Pipeline (TPP) instead of XPRESS. For the mTRAQ-labeled peptides obtained by an Orbitrap mass spectrometer, it showed more accurate ratios and better standard deviations than XPRESS. Especially, for the peptides that do not contain lysine, the ratio difference between XPRESS and our algorithm became larger as the peptide masses increased. We expect that this algorithm can also be applied to other labeling methods such as ^{18}O labeling and acrylamide labeling.

Keywords: mass spectrometry • peptide quantification • mTRAQ labeling • quadratic equation • stable isotope labeling

Introduction

In recent decades, many people have researched structures, functions, and regulations of proteins in biological systems. Mass spectrometry (MS) led to a great progress in proteomics as a powerful and robust tool for protein analysis. In many MS-based proteomic experiments, proteins are digested to peptides by enzyme and then analyzed by liquid chromatography coupled to online mass spectrometry (LC−MS) that generates a large amount (often as big as 1 GB) of raw data. By analyzing mass spectra, monoisotopic masses of peptides included in the mixture can be determined,[1,2] and the introduction of tandem MS (MS/MS)[3] enables the determination of peptide sequences.[4−6]

Protein quantification is one of many interesting computational problems in MS. Because most MS experiments are executed using peptides, protein quantification must be inferred from peptide quantification. Therefore, accurate peptide quantification is essential to correctly quantify the ratios of proteins. There are three major experimental strategies for quantitative proteomics:[7] spectral counting, stable isotope labeling, and label-free quantification. Among these, stable isotope labeling is considered as the most reliable and accurate method. There are various labeling techniques: ICAT,[8] SILAC,[9] ^{18}O labeling,[10,11] acrylamide labeling,[12,13] mTRAQ,[14] and so on. SILAC exploits metabolic labeling of cultured cells in which the medium is supplemented with amino acids containing stable isotopes. ICAT incorporates isotopes on the thiol moiety of cysteines chemically. Similarly, up to two ^{18}O atoms can be incorporated into carboxyl groups of peptides by digestion with proteases in the presence of $H_2{}^{18}O$. Acrylamide labeling uses cysteine alkylation, which occurs during polyacrylamide gel electrophoresis. Regardless of the chemical nature of the incorporated labels, quantitation is performed by analyzing MS scans. On the other hand, multiplexed isobaric tags (e.g., iTRAQ and TMT) represent another group of chemical-tagging reagents, which enables MS/MS-based quantitation.[15,16]

The isotope label mTRAQ is a nonisobaric version of iTRAQ and was originally designed for multiple reaction monitoring.[14] The labels come in chemically identical versions but with different masses. By virtue of this mass difference, unique multiple reaction monitoring (MRM) transitions are generated for any given peptide labeled with mTRAQ. In turn, this implies that when peptides from different sources are tagged separately with the different versions of mTRAQ labels, they can be quantitated based on MS full scans. We recently verified the utility of mTRAQ for MS-based relative quantification (manuscript under review at Journal of Proteome Research) by demonstrating that mTRAQ-labeled peptides show better qual-

* To whom correspondence should be addressed. Kunsoo Park, School of Computer Science and Engineering, Seoul National University, Seoul, 151-742, Korea. Tel: +82-2-880-8381. Fax: +82-2-885-3141. E-mail: kpark@theory.snu.ac.kr. Eunok Paek. Dept. of Mechanical and Information Engineering, University of Seoul, 90 Jeonnong-dong, Dongdaemun-gu, Seoul, Korea, 130-743. Tel: 82-2-2210-2680. Fax: 82-2-2210-5575. E-mail: paek@uos.ac.kr.

† Seoul National University.

‡ University of Seoul.

§ Korea Institute of Science and Technology.

|| University of Science and Technology.

10.1021/pr100183t

Published on Web 03/23/2010

ity MS/MS spectra with increased cross-correlation (XCorr) values from a SEQUEST search than the corresponding unlabeled peptides.

On the other hand, some isotopic labeling can potentially carry a critical problem, which results from an overlap of isotopic clusters. Overlapping isotopic clusters are found when the mass difference of light- and heavy-labeled peptides is very small (i.e., 4 Da in the case of mTRAQ if the original peptide has no lysine). In case of an overlap, a naïve approach can result in inaccurate quantification, which is biased toward the heavy-labeled peptide. Fortunately, isotopic peaks can now be better separated as high-resolution mass spectrometers, such as Fourier transform ion cyclotron resonance (FT-ICR) or Orbitrap, are used. Using the theoretical isotopic distribution that can be modeled based on the existential probabilities of isotope composition, we can identify isotopic clusters of peptides in complex mass spectra and quantify the ratios of the peptides accurately using these isotopic clusters.

Various computational tools for the stable isotope labeling have also been developed.[17−27] Some software tools such as VIPER[24] use only MS scans for quantification, but the most of them require a list of identified peptides as additional input. Q3[23] and IEMM[27] are designed to overcome overlapping isotopic cluster problems occurring in acrylamide labeling and ^{18}O labeling, respectively. Q3 uses an isotopic distribution prediction algorithm based on a Poisson approximation. IEMM uses isotopic envelope mixture modeling. Since most of them are designed for a specific isotopic labeling, it is difficult to compare the performance of various tools. XPRESS[17] is one of quantification tools that can analyze data with various isotopic labels including mTRAQ.

In this paper, we present a new algorithm to improve the accuracy of peptide quantification when mTRAQ-labeling is used. Most of our analysis is performed using Trans-Proteomic Pipeline (TPP) except that we use our new algorithm instead of XPRESS to quantify the ratios of peptides. The overall framework is shown in Figure 1. Because database search is performed considering the mass modification caused by mTRAQ, we can find not only the sequences of original peptides, but also the label types of these peptides. Using this information, we first extract isotopic clusters of labeled peptide pairs and then separate them using quadratic equations if they are overlapped. We tried three different methods to integrate the ratios computed from these isotopic cluster pairs and chose the one with the best performance. We compared the result of our algorithm with that of XPRESS for several experimental data sets in which we know the expected ratios of the proteins included. Our algorithm resulted in more accurate ratios for most peptides, especially in the case of peptides with no lysine. The potential use of this algorithm for other labeling is also discussed.

Experimental Section

Samples. Human plasma was obtained from healthy volunteers. The six most abundant proteins (serum albumin, immunoglobulin G, immunoglobulin A, transferrin, haptoglobin, antitrypsin) were depleted using an antibody-based depletion system (MARS column, Agilent Technologies, Palo Alto, CA). The unbound fraction was concentrated using Microcon (3000 Da cutoff, Millipore), and proteins were precipitated by letting stand in 6.5 volumes of cold acetone for 15 min at −20 °C. The precipitate was dissolved in a buffer containing 50 mM Tris-HCl (pH 8.0) and 6 M urea. Protein concentration was determined by the Bradford method.

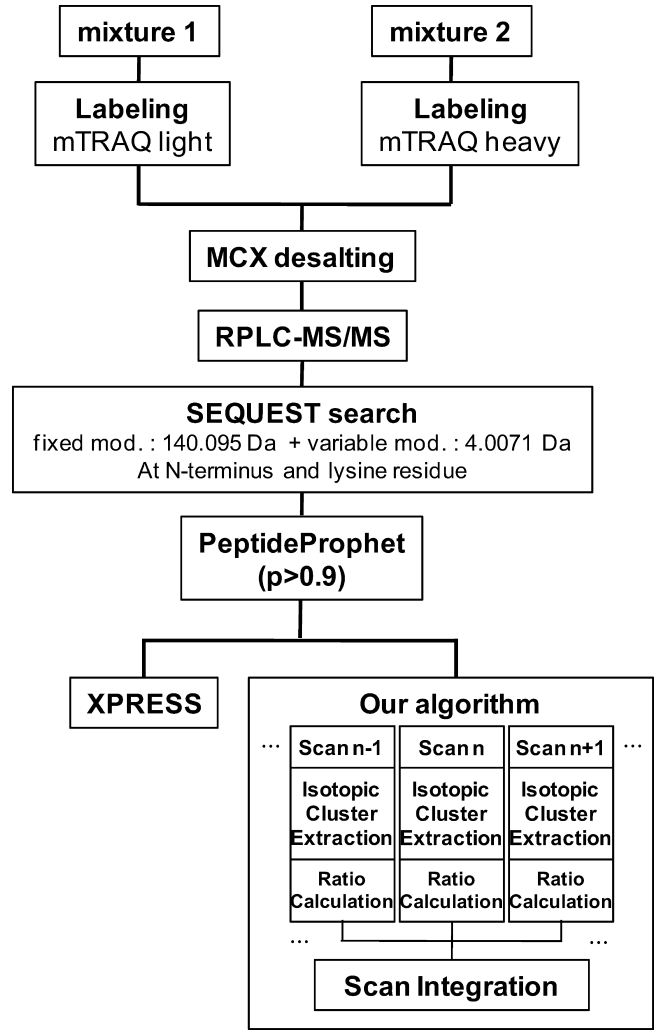

Figure 1. Overall framework of experiments.

Two kinds of standard protein mixtures were prepared for mTRAQ quantification tests. Each mixture consisted of alpha-lactalbumin, beta-casein, serotransferrin, alpha-S1-casein, alpha-S2-casein and pancreatic ribonuclease in 50 mM Tris pH 8.0 at different amounts: 10, 10, 20, 25, 25, and 10 μg for standard mixture 1 (Std1); 10, 20, 10, 5, 5, and 50 μg for standard mixture 2 (Std2).

To experiment on a variety of ratios, we mixed two standard protein mixtures (Std1 and Std2) in various ratios (1:1, 1:5, 5:1). In addition to this, prior to MS analysis, 0.4 mg of the mTRAQ labeled standard protein mixture was added to 1 mg of the trypsin-digested unlabeled plasma proteome in order to test performance under more realistic conditions (Table 1).

mTRAQ Labeling. The standard protein mixtures and the depleted plasma sample were labeled with mTRAQ reagent according to the manufacturer's protocol. Samples were reduced with 50 mM tris (2-carboxyethyl) phosphine (TCEP) for 1 h at 60 °C, treated with 200 mM methyl methanethiosulfonate (MMTS) for 10 min at 25 °C, and then diluted 10-fold with 50 mM Tris (pH 8.0), and digested with sequencing-grade trypsin (Promega, Madison, WI) at 37 °C overnight at the protein: trypsin molar ratio of 10:1. Tryptic digests were desalted with C18 solid-phase extraction (SPE) cartridge and dried in vacuo. The dried samples were reconstituted in 500 mM triethyl ammonium bicarbonate and incubated with appropriate mTRAQ

Table 1. Sample Description[a]

sample	light-labeled mixture	heavy-labeled mixture	plasma	mix ratios
S1L1_S2H1	Std1	Std2	No	1:1
S1H1_S2L1	Std2	Std1	No	1:1
S1L1_S2H5	Std1	Std2	No	1:5
S1L5_S2H1	Std1	Std2	No	5:1
PLASMA_S1L1_S2H1	Std1	Std2	Yes	1:1
PLASMA_S1H1_S2L1	Std2	Std1	Yes	1:1

[a] Each standard protein mixture (Std1 and Std2) was labeled with heavy or light mTRAQ label and mixed with various ratios (1:1, 1:5, 5:1). They were also spiked into unlabeled human plasma sample. More detailed description can be found in Experimental Section.

reagents at 25 °C for 1 h. After the labeling reaction, samples were dried in vacuo, redissolved in 0.1% trifluoroacetic acid (TFA), desalted with a mixed-mode strong cation-exchange (MCX) cartridge and dried again.

Liquid Chromatography and Tandem Mass Spectrometry. Tagged peptide samples were reconstituted in 0.4% acetic acid and an aliquot (~1 μg) was injected to a reversed-phase Magic C18aq column (15 cm × 75 μm) on an Eksigent multidimensional liquid chromatography (MDLC) system at the flow rate of 300 nL/min. The column was equilibrated with 95% buffer A (0.1% formic acid in H_2O) + 5% buffer B (0.1% formic acid in acetonitrile) prior to use. The peptides were eluted with a linear gradient of 10 to 40% Buffer B over 40 min.

The high performance liquid chromatography (HPLC) system was coupled to a linear trap quadrupole (LTQ) XL-Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA). The spray voltage was set to 1.9 kV, and the temperature of the heated capillary was set to 250 °C. Survey full-scan MS spectra (m/z 300−2000) were acquired in the Orbitrap with 1 microscan and a resolution of 100 000, allowing the preview mode for precursor selection and charge-state determination. MS/MS spectra of the five most intense ions from the preview survey scan were acquired in the ion-trap concurrently with full-scan acquisition in the Orbitrap with the following options: isolation width, ±10 ppm; normalized collision energy, 35%; dynamic exclusion duration, 30 s. Precursors with unmatched charge states were discarded during data dependent acquisition. Data were acquired using the Xcalibur software v2.0.7 (Thermo Electron, San Jose, CA).

Database Search. The dta files for tandem mass spectra were generated by Extract-msn program (v3) of Bioworks software (v3.2, Thermo Electron, San Jose, CA) with the following parameters: minimum ion count threshold, 15; minimum intensity, 100. The MS/MS spectra acquired from plasma samples were searched using SEQUEST[5] (TurboSequest version 27, revision 12) against the International Protein Index human database (IPI human, version 3.44, European Bioinformatics Institute, http://www.ebi.ac.uk/IPI) including known contaminants totaling 72 065 protein entries, allowing the options of no enzyme, ±0.5 Da mass tolerance for MS/MS, ±15 ppm mass tolerance for MS, fixed modification of 140.095 Da plus variable modification of +4.0071 Da on peptide N-terminus and Lys residue. Variable modification of methionine oxidation (+15.9949 Da) was allowed and a fixed modification of 45.9877 Da on Cys residue (MMTS) was used. In case of MS/MS spectra for the standard protein mixtures, a compound database consisting of IPI human v3.56 and IPI bovine v3.30 was used with all the other options the same as the plasma data set.

Peptide assignment and quantification were performed with the Trans-Proteomic Pipeline (TPP, version 4.0, http://www.proteomecenter.org). The SEQUEST search output was used as an input for pepXML module allowing trypsin restriction and 'monoisotopic masses' options. Then PeptideProphet[28] was applied with the 'accurate mass binning' option. Peptides with probabilities greater than 0.9 were included in the subsequent quantitation analysis.

Peptide Quantification. Relative quantitative ratios of peptides were determined using both XPRESS[17] and our algorithm. Relative quantification using XPRESS was performed during PeptideProphet analysis with a parent mass tolerance of 0.05 Da and mass difference of 4.0071 Da on N-terminus and Lys. We used two methods to determine elution areas of the precursor ions: (1) by automatic determination and (2) in ±30 scans from the peak apex. The same quantification was also performed by using our algorithm and the results were compared.

Methods

Model for Overlapping Isotopic Clusters. The mass difference between light and heavy mTRAQ-labeled peptides is about 4 Da if the original peptide has no lysine. These pairs have an overlap in their isotopic clusters if the light-labeled isotopic cluster has five or more isotopic peaks (Figure 2). In this work, we assume

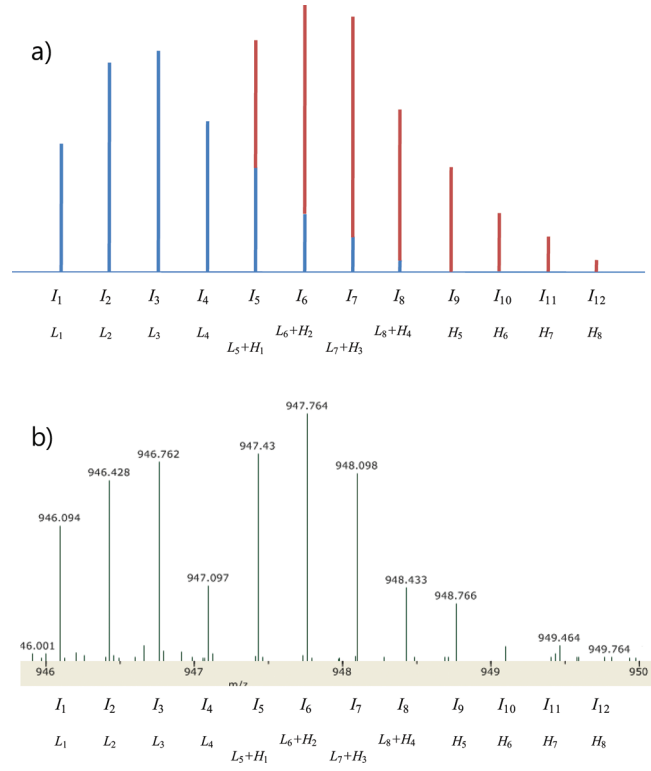

Figure 2. Examples of overlapping isotopic clusters. (a) Schematic illustration of overlapping isotopic clusters. Blue and red lines represent isotopic peaks of light- and heavy-labeled peptides, respectively. Because the mass difference between light- and heavy-labeled peptides is about 4 Da and the number of peaks in an isotopic cluster is 8, the fifth to the eighth peaks of the blue distribution overlap with the first to the fourth from the red distribution. To obtain the accurate ratio between light- and heavy-labeled peptides, the separation of the overlapping peaks is very important. (b) Experimental overlapping isotopic clusters of "HQPQEFPTYVEPTNDEIcEAFR" ("c" represents a modified cysteine by MMTS). By our method, we calculated four partial H/L ratios $\alpha_1 = 1.167463$, $\alpha_2 = 1.304014$, $\alpha_3 = 0.899859$, $\alpha_4 = 0.919235$ and finally obtained the H/L ratio 1.093693 from the four values.

that an isotopic cluster of a peptide has 8 or less peaks. It is reasonable for peptides whose masses are less than 4000 Da because the relative intensity of the ninth peak in the theoretical distribution of an averagine[29] whose mass is 4000 Da is only 0.56%. The intensity, I_k, of the k-th peak of a theoretical distribution of overlapping isotopic clusters is given as follows:

$$I_k = \begin{cases} L_k & if \quad k \le 4 \\ L_k + H_{k-4} & if \quad 4 < k \le n \\ H_{k-4} & if \quad k > n \end{cases}$$

where n is the number of peaks in the isotopic distribution of a peptide, L_k is the intensity of the k-th peak of the isotopic distribution of the light-labeled peptide, and H_k is the intensity of the k-th peak of the isotopic distribution of the heavy-labeled peptide.

Let α be the heavy-to-light (H/L) ratio, that is, $H_k = \alpha L_k$. Then we can calculate α from I_k values. First, we induce a quadratic equation $I_1\alpha^2 - I_5\alpha + I_9 = 0$ from $I_9 = \alpha L_5$ and $I_5 = L_5 + \alpha L_1$. Using the quadratic formula, we obtain two solutions:

$$\alpha = \frac{I_5 \pm \sqrt{I_5^2 - 4I_1I_9}}{2I_1}$$

To find an exact solution, we transform them to the equations of L_k:

$$\begin{aligned} &\frac{I_5 \pm \sqrt{I_5^2 - 4I_1I_9}}{2I_1} \\ &\quad = \frac{\alpha L_1 + L_5 \pm \sqrt{(\alpha L_1)^2 + 2\alpha L_1 L_5 + L_5^2 - 4\alpha L_1 L_5}}{2L_1} \\ &\quad = \frac{\alpha L_1 + L_5 \pm \sqrt{(\alpha L_1 - L_5)^2}}{2L_1} \end{aligned}$$

It is easy to see that the larger solution is equal to α if $\alpha \ge L_5/L_1$, and the smaller solution becomes α, otherwise. I_k values are read from the experimental data and L_k can be calculated using the theoretical distribution of the peptide. Therefore, we can calculate α as follows:

$$\alpha = \begin{cases} \dfrac{I_5 + \sqrt{I_5^2 - 4I_1I_9}}{2I_1} & \text{if} \quad \dfrac{I_5 + \sqrt{I_5^2 - 4I_1I_9}}{2I_1} \ge \dfrac{L_5}{L_1} \\ \dfrac{I_5 - \sqrt{I_5^2 - 4I_1I_9}}{2I_1} & \text{otherwise} \end{cases}$$

Similarly, we can induce three more equations from I_k values:

$$\begin{aligned} I_2\alpha^2 - I_6\alpha + I_{10} = 0 \\ I_3\alpha^2 - I_7\alpha + I_{11} = 0 \\ I_4\alpha^2 - I_8\alpha + I_{12} = 0 \end{aligned}$$

We can calculate multiple α values from these four quadratic equations. Theoretically, all the values should be the same, but the α values calculated using experimental data can be different from each other due to various imperfections in experiments such as low sensitivity, chemical noise, and/or experimental errors. Therefore, it is necessary to integrate the four values. Let α_k be the α value calculated from I_k, I_{k+4} and I_{k+8}, that is,

$$\alpha_k = \frac{H_k + H_{k+4}}{L_k + L_{k+4}}$$

Since $I_k + I_{k+4} + I_{k+8} = (L_k + L_{k+4}) + (H_k + H_{k+4})$, we get

$$\begin{aligned} L_k + L_{k+4} = (I_k + I_{k+4} + I_{k+8})/(1 + \alpha_k), \\ H_k + H_{k+4} = (I_k + I_{k+4} + I_{k+8}) \times \alpha_k/(1 + \alpha_k). \end{aligned}$$

By summing these values up, we can calculate the H/L ratio as $\alpha = \sum H_k / \sum L_k$.

Sometimes the quadratic formula for α_k gives no real number solution. This happens when I_k or I_{k+8} are larger than (equivalently, I_{k+4} is smaller than) the theoretically expected intensities. To have at least one real number solution, I_{k+4} must be large enough to satisfy the constraint $I_{k+4}^2 \ge 4I_kI_{k+8}$. Under this assumption, we substitute $2(I_kI_{k+8})^{1/2}$ for I_{k+4} and obtain $\alpha_k = I_{k+4}/2I_k = (I_{k+8}/I_k)^{1/2}$ if $I_{k+4}^2 - 4I_kI_{k+8} < 0$.

Extraction of Isotopic Clusters. For each peptide, we first extract isotopic clusters of the peptide from the precursor MS scan of a MS/MS scan. Because the mass of the peptide and the charge state are obtained from the pepXML file, we can easily locate the first peak of one of the (light or heavy) isotopic clusters. We also find the first peak of the other isotopic cluster depending on the type of label (which is also obtained from the pepXML file). Subsequently, we extract at most 7 next peaks from each cluster if the two labeled peptides have 8 Da or more mass difference, and we extract at most 11 peaks overall if they have 4 Da mass difference. Each peak is found within 10 ppm mass tolerance.

In experimental data, some peaks from other peptides can be overlapped with the extracted isotopic clusters. To avoid including peaks from these other peptides, which can lead to incorrect quantification, we use the least-squares fit values between the extracted isotopic clusters and the theoretical distribution of the peptide. The least-squares fit values are calculated as follows:

$$LSF_L = \frac{\sum (T_k - N_L L_k)^2}{\sum T_k^2}, \; LSF_H = \frac{\sum (T_k - N_H H_k)^2}{\sum T_k^2}$$

Table 2. Performance Comparison between Different Methods for Determining the Ratios[a]

mode		XPRESS	Sum Ratio	Weighted Avg.	Regression
AUTO	Average $\log_{10}$(H/L)	0.071242	0.034487	0.034569	0.022763
	Standard deviation	0.110822	0.121618	0.151820	0.134611
FIX	Average $\log_{10}$(H/L)	0.070026	0.036780	0.037344	0.031133
	Standard deviation	0.115151	0.102963	0.109182	0.101954

[a] It shows averages and standard deviations of $\log_{10}$(H/L) values obtained from 1:1 human plasma sample. There are three different methods in our algorithm to integrate ratios from a set of scans ("Sum Ratio", "Weighted Average", and "Regression"). For the determination of elution areas of precursor ions, all programs were executed in both automatic mode (AUTO) and user-defined mode with ±30 scans (FIX).

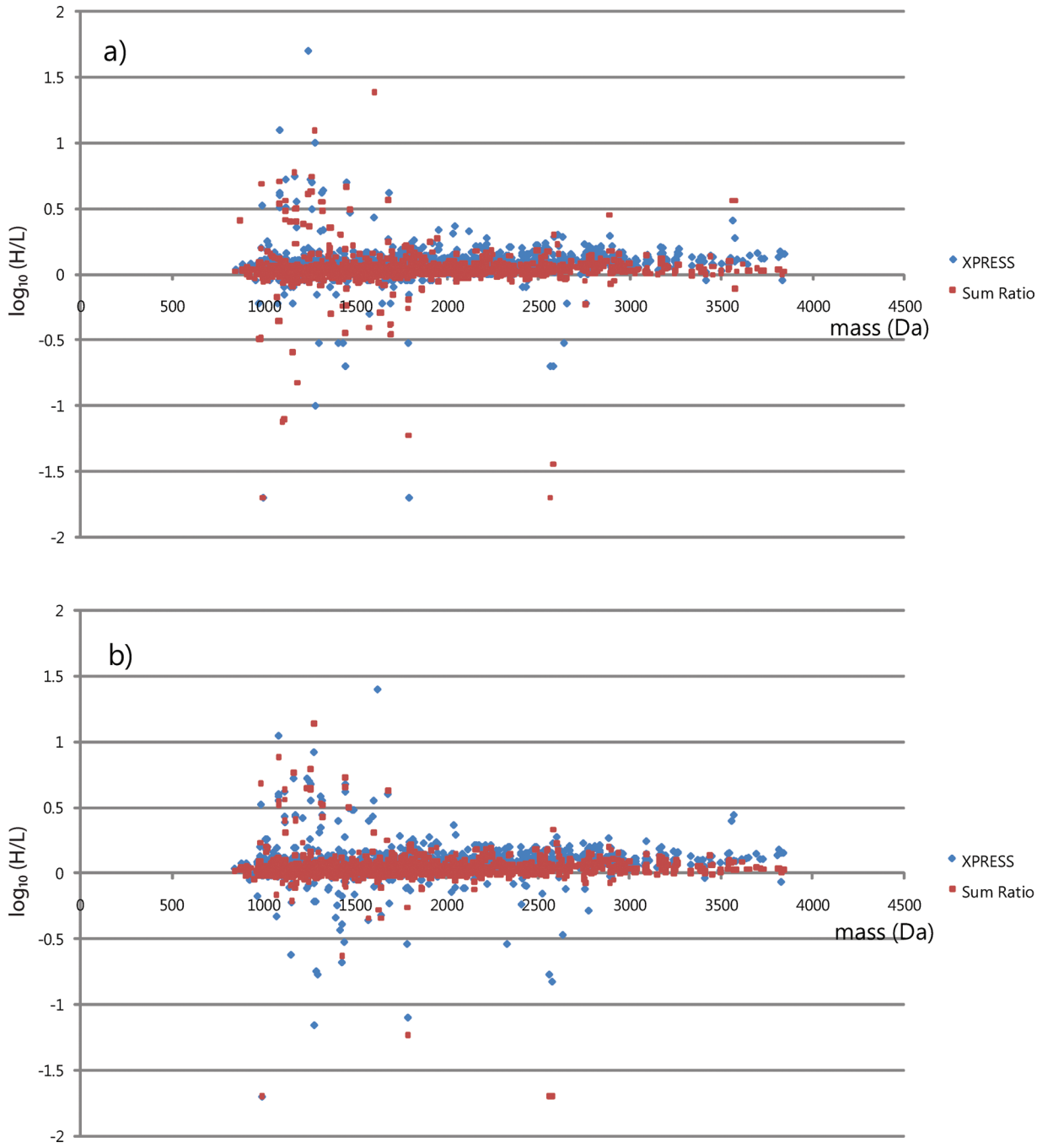

Figure 3. Distribution of $\log_{10}$(H/L) values of peptides from 1:1 human plasma sample under (a) AUTO mode and (b) FIX mode. Though all $\log_{10}$(H/L) values are expected to be 0.0, most of $\log_{10}$(H/L) values calculated by XPRESS were biased toward heavy.

where T_k is the relative intensity of the k-th peak in the theoretical distribution of the peptide, and N_L and N_H are normalization factors. The extracted isotopic clusters are used to quantify the ratio of the peptide if at least one of the two least-squares fit values is less than 20%.

We also extract isotopic clusters from the scans that are adjacent to the precursor scan of the MS/MS scan. We provide an automatic determination of elution areas of the precursor ions. First, we consider all scans during 10 s from the precursor scan of the MS/MS scan. Then, we consider subsequent scans until we find two consecutive holes or five holes in total. (A hole is a scan with no isotopic cluster whose least-squares fit value is less than 20%.) We also consider scans directly prior to the precursor MS scan of the MS/MS scan until we find two consecutive holes or five in total. It is also possible for a user to define a fixed number of scans to be used for quantitation. In this mode, we only consider $\pm$ n MS scans from MS/MS scan where n is a user-defined number.

Integration of Ratios. There may be many scans from which we can calculate a ratio α for a peptide. We need to integrate the ratios obtained from these scans. We consider three integration methods. First, we sum up the intensities for each labeled peptide (for all scans, not only for one scan) and calculate the ratio between summed intensities (called "Sum Ratio"). Second, we calculated the weighted average of all the ratios from each scan (called "Weighted Avg."). In this case, the sum of intensities in a scan is used as the weight. Third, we calculate a linear regression of the sum of intensities for

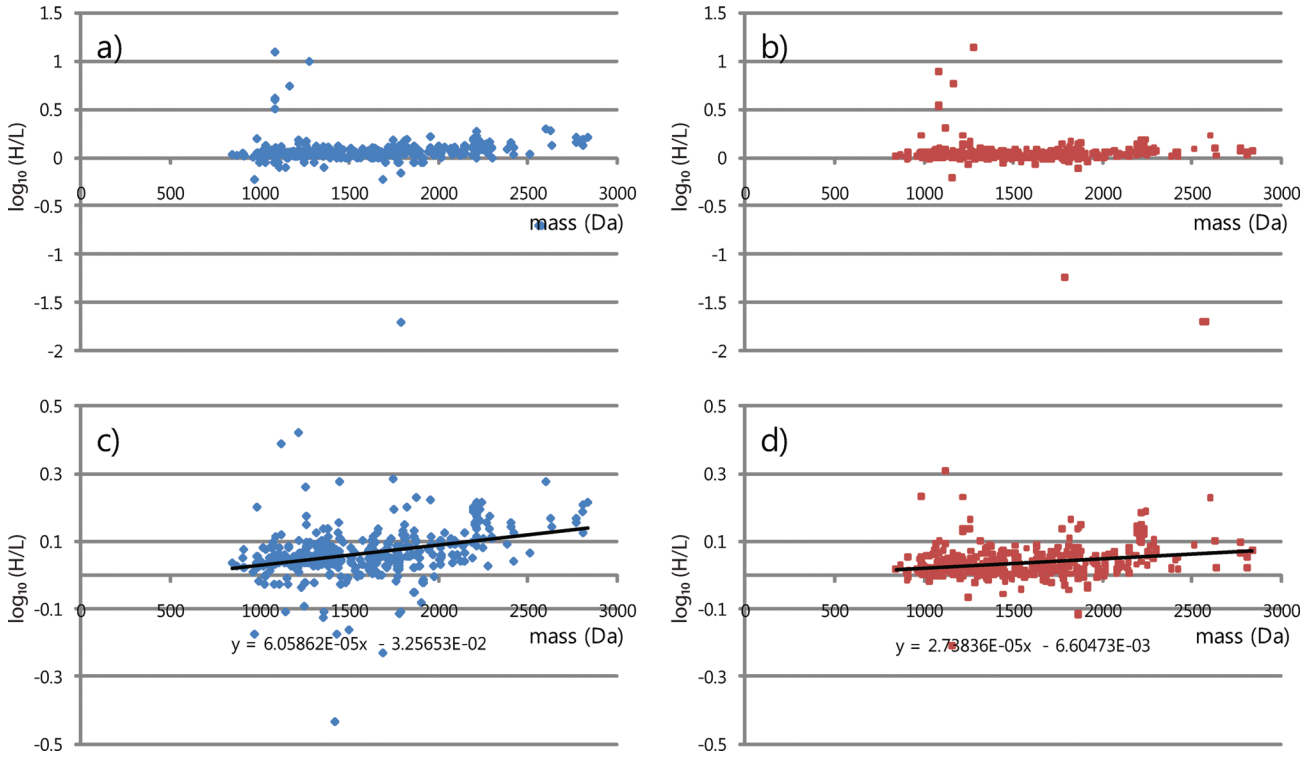

Figure 4. Distribution of $\log_{10}$(H/L) values of peptides with no lysine. (FIX mode) (a) Distribution of $\log_{10}$(H/L) values calculated by XPRESS. (b) Distribution of $\log_{10}$(H/L) values calculated by our algorithm (Sum Ratio). (c) Distribution of $\log_{10}$(H/L) values calculated by XPRESS after removing outliers (>0.5 or <−0.5). H/L ratios increase as peptide masses increase. (d) Distribution of $\log_{10}$(H/L) values calculated by our algorithm (Sum Ratio) after removing outliers. The ratios are more consistent than XPRESS regardless of the masses of peptides.

each labeled peptide using the form "Heavy intensity sum = α × light intensity sum" (called "Regression").

Results

Overview. For mTRAQ quantification tests, we prepared seven experimental data sets. We first mixed the same amount of the human plasma samples labeled differently. Because the same sample was used for light and heavy labeling, all the ratios of the peptides identified from this mixture are expected to be 1.0. We also mixed two standard protein mixtures (Std1 and Std2) in various ratios (Table 1).

We determined peptide ratios by trying three different methods: Sum Ratio, Weighted Avg., and Regression. To evaluate the performance of our algorithm, we compared them with XPRESS developed at the Institute for Systems Biology. All programs were executed on the same PC (Intel E6300 processor 1.86 GHz, 2GB RAM, Windows XP). For the determination of elution areas of precursor ions, all programs were executed in both automatic mode (AUTO) and user-defined mode with ±30 scans (FIX). We averaged all $\log_{10}$(H/L) values of each peptide and calculated standard deviations. Because our tests used sampled data with known ratios, the averages and standard deviations of ratios can be a measure to evaluate the correctness of a method. Especially, low standard deviation can be very important for accurate relative quantification. Some peptides have low labeling efficiency because of their chemical properties. In this case, observed ratios can be very different from expected ratios, but they would be considered more reliable if they are all close to a certain value.

1:1 Human Plasma Sample. The same human plasma samples were labeled with light and heavy mTRAQ reagents, respectively. We mixed light and heavy mTRAQ-labeled peptides in the ratio of 1:1. So the average ratios of all peptides were expected to be 1.0. Overall, 2291 peptides were selected by PeptideProphet from this mixture.

The ratios calculated by eight different methods are shown in Table 2. The averages of $\log_{10}$(H/L) values calculated by XPRESS were larger than 0.07 while the largest average value calculated by our method was 0.037344. In this sample, the standard deviations of our method in AUTO mode were larger than XPRESS, but the overall performance of our method was better considering that our average was a lot closer to 0.0. All three methods in FIX mode gave more accurate ratios and better standard deviations than XPRESS. In this sample, Regression in FIX mode seemed to give better results, but the differences with other methods were negligible. On the other hand, with the standard mixture experiments, Regression showed worse results than Sum Ratio, especially when expected ratios are far from 1.0. Overall, we recommend using Sum Ratio. Distribution of $\log_{10}$(H/L) values of XPRESS and Sum Ratio are shown in Figure 3. It shows that most of $\log_{10}$(H/L) values calculated by XPRESS were biased toward heavy.

Ratio of Peptides with No Lysine. Since mTRAQ is specific to primary amine, the mass difference between heavy- and light-labeled peptides is a multiple of 4 Da depending on the number of Lys residues present, and thus, peptides without any lysine have the smallest mass difference of 4 Da. Our overlap model is especially effective for peptides with no lysine. From 1:1-mixed human plasma sample, we selected the peptides that have no lysine. There were 544 such peptides. Since there are a few outliers whose ratios are far from the expected ratio, it is hard to observe their linearity (Figure 4a and b). Therefore, we removed those outliers whose $\log_{10}$(H/L) values were larger than 0.5 or smaller than −0.5. Then, we fitted

Table 3. Expected Ratios and Computed Ratios for Each of Six Proteins in Standard Mixtures[a]

			XPRESS		our method (Sum Ratio)	
protein	expected ratio		AUTO	FIX	AUTO	FIX
		(a) S1L1_S2H1 sample				
Alpha-lactalbumin	1	Average (H/L)	1.230691	1.219947	1.149738	1.137588
		Average Log(H/L)	0.090149	0.086341	0.060599	0.055985
		Standard deviation	0.117583	0.114267	0.079873	0.078114
Beta-casein	2	Average (H/L)	2.006855	2.095975	1.834193	1.920547
		Average Log(H/L)	0.302516	0.321386	0.263445	0.283425
		Standard deviation	0.107177	0.043626	0.113144	0.076320
Serotransferrin	0.5	Average (H/L)	0.538917	0.536943	0.481471	0.470103
		Average Log(H/L)	−0.268478	−0.270072	−0.317430	−0.327807
		Standard deviation	0.088544	0.084211	0.094122	0.076619
Alpha-S1-casein	0.2	Average (H/L)	0.245641	0.246540	0.204650	0.197205
		Average Log(H/L)	−0.609700	−0.608113	−0.688988	−0.705081
		Standard deviation	0.147435	0.123440	0.075113	0.077004
Alpha-S2-casein	0.2	Average (H/L)	0.276269	0.278549	0.237995	0.239169
		Average Log(H/L)	−0.558668	−0.555098	−0.623433	−0.621295
		Standard deviation	0.167923	0.164208	0.159004	0.160128
Pancreatic ribonuclease	5	Average (H/L)	7.580607	7.552296	7.460948	7.698693
		Average Log(H/L)	0.879704	0.878079	0.872794	0.886417
		Standard deviation	0.283208	0.145355	0.126991	0.130309
		(b) S1H1_S2L1 sample				
Alpha-lactalbumin	1	Average (H/L)	1.038133	1.044059	1.004826	0.956364
		Average Log(H/L)	0.016253	0.018725	0.002091	−0.019377
		Standard deviation	0.082272	0.088696	0.094697	0.058738
Beta-casein	0.5	Average (H/L)	0.579788	0.586198	0.528163	0.514628
		Average Log(H/L)	−0.236731	−0.231956	−0.277232	−0.288507
		Standard deviation	0.078908	0.065993	0.085222	0.060252
Serotransferrin	2	Average (H/L)	2.380324	2.376561	2.271747	2.269645
		Average Log(H/L)	0.376636	0.375949	0.356360	0.355958
		Standard deviation	0.124123	0.125028	0.099330	0.116415
Alpha-S1-casein	5	Average (H/L)	5.277849	5.069837	4.826430	4.951468
		Average Log(H/L)	0.722457	0.704994	0.683626	0.694734
		Standard deviation	0.312771	0.231996	0.088898	0.101004
Alpha-S2-casein	5	Average (H/L)	4.956373	5.189625	4.759058	4.818879
		Average Log(H/L)	0.695164	0.715136	0.677521	0.682946
		Standard deviation	0.163442	0.088622	0.091010	0.102082
Pancreatic ribonuclease	0.2	Average (H/L)	0.168144	0.165319	0.136372	0.133329
		Average Log(H/L)	−0.774318	−0.781678	−0.865275	−0.875077
		Standard deviation	0.182539	0.175129	0.120305	0.123731
		(c) S1L1_S2H5 sample				
Alpha-lactalbumin	5	Average (H/L)	5.755300	5.525073	6.917609	7.053002
		Average Log(H/L)	0.760068	0.742338	0.839956	0.848374
		Standard deviation	0.321660	0.239571	0.078998	0.090011
Beta-casein	10	Average (H/L)	13.714659	12.744682	11.586279	11.883627
		Average Log(H/L)	1.137185	1.105329	1.063944	1.074949
		Standard deviation	0.418943	0.346553	0.129604	0.195382
Serotransferrin	2.5	Average (H/L)	2.994408	2.983878	2.881848	2.869492
		Average Log(H/L)	0.476311	0.474781	0.459671	0.457805
		Standard deviation	0.115859	0.120334	0.068620	0.062601
Alpha-S1-casein	1	Average (H/L)	1.288715	1.288801	1.194904	1.197350
		Average Log(H/L)	0.110157	0.110186	0.077333	0.078221
		Standard deviation	0.054258	0.048935	0.040909	0.050246
Alpha-S2-casein	1	Average (H/L)	1.542403	1.524502	1.387951	1.411569
		Average Log(H/L)	0.188198	0.183128	0.142374	0.149702
		Standard deviation	0.050405	0.060045	0.059325	0.068205
Pancreatic ribonuclease	25	Average (H/L)	51.207079	45.288819	28.540593	34.022091
		Average Log(H/L)	1.709330	1.655991	1.455463	1.531761
		Standard deviation	0.551515	0.506536	0.179232	0.212650
		(d) S1L5_S2H1 sample				
Alpha-lactalbumin	0.2	Average (H/L)	0.295901	0.298147	0.255756	0.258905
		Average Log(H/L)	−0.528854	−0.525569	−0.592175	−0.586859
		Standard deviation	0.153596	0.142507	0.125617	0.139198
Beta-casein	0.4	Average (H/L)	0.648700	0.628520	0.566830	0.566405
		Average Log(H/L)	−0.187956	−0.201681	−0.246547	−0.246873
		Standard deviation	0.349004	0.353914	0.362788	0.364027

Table 3. Continued

protein	expected ratio		XPRESS AUTO	XPRESS FIX	our method (sum ratio) AUTO	our method (sum ratio) FIX
Serotransferrin	0.1	Average (H/L)	0.147768	0.147968	0.123075	0.119674
		Average Log(H/L)	−0.830421	−0.829833	−0.909831	−0.922000
		Standard deviation	0.340103	0.339036	0.341322	0.337484
Alpha-S1-casein	0.04	Average (H/L)	0.071880	0.075151	0.058639	0.053077
		Average Log(H/L)	−1.143394	−1.124066	−1.231815	−1.275091
		Standard deviation	0.240020	0.213598	0.125220	0.142586
Alpha-S2-casein	0.04	Average (H/L)	0.122571	0.125305	0.063392	0.059175
		Average Log(H/L)	−0.911612	−0.902030	−1.197965	−1.227863
		Standard deviation	0.539069	0.547153	0.243583	0.249255
Pancreatic ribonuclease	1	Average (H/L)	2.546091	2.546719	2.251614	2.307985
		Average Log(H/L)	0.405874	0.405981	0.352494	0.363233
		Standard deviation	0.530181	0.555775	0.404159	0.423759
		(e) PLASMA_S1L1_S2H1 sample				
Alpha-lactalbumin	1	Average (H/L)	1.167103	1.173313	1.045890	1.063616
		Average Log(H/L)	0.067109	0.069414	0.019486	0.026785
		Standard deviation	0.071767	0.076176	0.065369	0.069395
Beta-casein	2	Average (H/L)	1.993862	1.988186	1.664604	1.828555
		Average Log(H/L)	0.299695	0.298457	0.221311	0.262108
		Standard deviation	0.040448	0.033685	0.12183	0.051021
Serotransferrin	0.5	Average (H/L)	0.526971	0.529435	0.472549	0.465749
		Average Log(H/L)	−0.278213	−0.276187	−0.325553	−0.331848
		Standard deviation	0.052557	0.065528	0.059061	0.040219
Alpha-S1-casein	0.2	Average (H/L)	0.246812	0.245988	0.194489	0.191135
		Average Log(H/L)	−0.607633	−0.609086	−0.711106	−0.71866
		Standard deviation	0.089503	0.09047	0.038395	0.039765
Alpha-S2-casein	0.2	Average (H/L)	0.278516	0.280265	0.237007	0.239811
		Average Log(H/L)	−0.55515	−0.552431	−0.625239	−0.62013
		Standard deviation	0.058338	0.063346	0.080802	0.081277
Pancreatic ribonuclease	5	Average (H/L)	6.120938	5.973557	3.346162	5.722595
		Average Log(H/L)	0.786818	0.776233	0.524547	0.757593
		Standard deviation	0.159831	0.097629	0.635218	0.099495
		(f) PLASMA_S1H1_S2L1 sample				
Alpha-lactalbumin	1	Average (H/L)	1.048682	1.058293	0.958685	0.959951
		Average Log(H/L)	0.020644	0.024606	−0.018324	−0.017751
		Standard deviation	0.066572	0.063705	0.066239	0.066273
Beta-casein	0.5	Average (H/L)	0.613213	0.609960	0.533394	0.545419
		Average Log(H/L)	−0.212389	−0.214699	−0.272952	−0.26327
		Standard deviation	0.032922	0.035318	0.026962	0.028457
Serotransferrin	2	Average (H/L)	2.327148	2.370915	2.230874	2.267410
		Average Log(H/L)	0.366824	0.374916	0.348475	0.35553
		Standard deviation	0.094981	0.102624	0.070709	0.059478
Alpha-S1-casein	5	Average (H/L)	5.595772	5.491525	5.492954	5.480965
		Average Log(H/L)	0.74786	0.739693	0.739806	0.738857
		Standard deviation	0.081621	0.080087	0.071198	0.070827
Alpha-S2-casein	5	Average (H/L)	4.698346	4.739472	4.640534	4.631237
		Average Log(H/L)	0.671945	0.67573	0.666568	0.665697
		Standard deviation	0.080163	0.080312	0.076332	0.078796
Pancreatic ribonuclease	0.2	Average (H/L)	0.225634	0.220806	0.134457	0.178023
		Average Log(H/L)	−0.646596	−0.65599	−0.871417	−0.749525
		Standard deviation	0.16896	0.161123	0.278224	0.134524

[a] We first calculated the averages and standard deviations of $\log_{10}$(H/L) values of peptides. Then, we transformed the averages into H/L scale to compare them to expected ratios.

$\log_{10}$(H/L) values using linear regression (Figure 4c and d). The slope from the results of XPRESS in FIX mode was 6.05862×10^{-5}, which was more than twice larger than that from our algorithm (Sum Ratio in FIX mode), 2.73836×10^{-5}. The ratios calculated by XPRESS were consistently larger than the ratios calculated by our algorithm, especially in high masses. It strongly indicates that our algorithm shows better performance than XPRESS for the quantification of the stable isotope labeled peptides that have an overlap in their isotopic distributions.

Standard Mixture 1 and Standard Mixture 2. Standard mixture 1 (Std1) and Standard mixture 2 (Std2) consisted of 6 proteins, where proteins are present in different amounts: 10, 10, 20, 25, 25, and 10 μg for standard mixture 1; 10, 20, 10, 5, 5, and 50 μg for standard mixture 2. Both standard mixtures were labeled with heavy and light mTRAQ labels, respectively, and then mixed in various ratios (1:1, 1:5, 5:1), generating six LC−MS data sets (Table 1). By doing so, we could experiment on a variety of ratios. For example, alpha-lactalbumin amounted

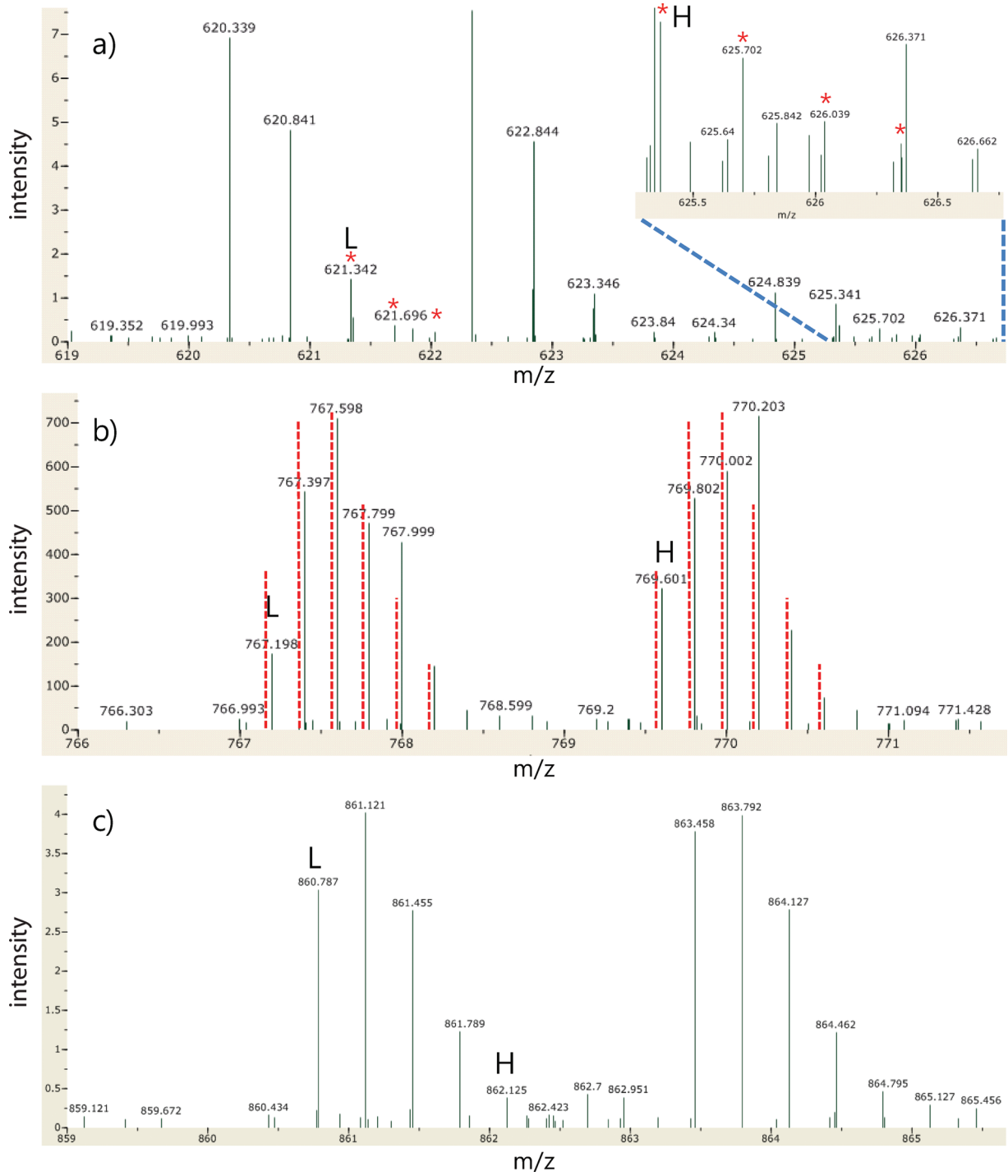

Figure 5. Examples where our method calculated more accurate ratios. "L" and "H" represent the monoisotopic peak of light- and heavy-labeled peptides, respectively. (a) Because the least-squares fit value of light-labeled peptide (LSF_L) is 0.31, our method excluded this mass spectrum. We can verify that the monoisotopic peak of light-labeled peptide is overlapped with the third peak of a peptide whose mass is 1242.68 Da. (b) Red dashed lines represents the theoretical isotopic distribution. XPRESS obtained 1.86 as H/L ratio because it uses only the monoisotopic peak. Our method obtained 0.99 as H/L ratio. The relative intensity of monoisotopic peak becomes smaller as the mass of peptide becomes larger, and therefore all peaks of isotopic cluster should be used for better quantification. (c) XPRESS used the fifth peak of isotopic cluster of light-labeled peptide to quantify heavy-labeled peptide, and obtained 0.15 as H/L ratio. Our methods only found the isotopic cluster of light-labeled peptide, and obtained 0.02 (user-defined minimum ratio) as H/L ratio. In this case, it seems that peptide was misassigned. The correct mass difference between light and heavy labels should be 8 Da.

to 10 μg in Std1 and 10 μg in Std2. The expected ratios of alpha-lactalbumin would be 5.0 in S1L1_S2H5 sample and 0.2 in S1L5_S2H1 sample. Similarly, because alpha-S2-casein amounted to 25 μg in Std1 and 5 μg in Std2, its ratios is expected to be 1.0 in S1L1_S2H5 sample and 0.04 in S1L5_S2H1 sample.

For standard protein mixture samples, we present only the results of Sum Ratio method. (Weighted Avg. and Regression methods showed similar but somewhat worse results.) Expected ratios and computed ratios for each of six proteins are given in Table 3. For all mixtures, our method showed similar or better

average ratios than XPRESS except pancreatic ribonuclease. Especially, the peptides of alpha-S1-casein and alpha-S2-casein showed much better ratios than XPRESS. Furthermore, in spite of inaccurate averages, most of the peptides of pancreatic ribonuclease (except AUTO mode in PLASMA_S1L1_S2H1) showed ratios biased toward Std2. This result implies that there may be unknown error factors that make the ratios of peptides of pancreatic ribonuclease incorrect. Most of standard deviations from our method were also better than XPRESS.

Quantification using Isotopic Cluster. Our method uses the sum of intensities of detected isotopic cluster for quantification while XPRESS uses only the intensity of the monoisotopic peak. By checking the existence of isotopic clusters (not just existence of monoisotopic peaks), we can avoid incorrect quantification in various cases: first, we can exclude peaks of other peptides (Figure 5a); second, we can quantify accurately even when the monoisotopic peak has abnormal intensity (Figure 5b); third, we can avoid using the fifth peak of isotopic cluster of light-labeled peptide as heavy-labeled peptide (Figure 5c).

Discussion

We presented a new algorithm for the peptide quantification that can overcome errors resulting from overlapping isotopic clusters of heavy- and light-labeled peptides. The peptide quantification is the first step in protein quantification. Using the quadratic equations induced from the theoretical distribution model of overlapping isotopic clusters, our algorithm could separate the overlapping isotopic clusters and quantify the ratio of isotope labeled peptides more accurately and reliably than XPRESS, which does not cope with isotopic overlaps at all, especially for the peptides whose mass difference between labels is relatively small. The ratios of our algorithm are more accurate, even when it was given the same elution profiles (FIX mode). It strongly implies that impact of overlapping clusters should not be ignored. We expect that this algorithm can be extended to other labeling methods such as ^{18}O and acrylamide labeling. Because the mass shift in ^{18}O labeling is 2 or 4 Da and the mass shift in acrylamide labeling is 3 Da, it can be analyzed based on the same principle we used for mTRAQ labeled peptides without lysine.

Our algorithm obtained good averages and standard deviations for the peptides whose expected ratios lie between 0.1 and 10.0, but failed to obtain good results for peptides with bigger differences in quantity, similarly with an existing method. It seems that the current methodology is not sensitive enough to handle proteins whose quantitative difference exceeds more than an order of magnitude. Further study is required both in terms of mass spectrometry and data analysis. We exploited high mass accuracy and high resolution of the Orbitrap mass spectrometer, and successfully corrected the biases introduced by overlapped isotope clusters. But we need to explore various other ways we can further improve sensitivity in quantitation. For instance, we can improve chromatographic conditions by introducing multidimensional separation; we can try out different modes of operation for data acquisition by mass spectrometers such as making use of inclusion/exclusion lists; different algorithmic solution for identifying isotopic clusters of labeled peptide pairs can also be sought out.

The determination of elution areas of precursor ions is also important to improve quantitation results. Our method does not use the intensities of isotopic clusters in automatic determination of elution areas. Use of intensity information considering the shape of an elution curve of a peptide would improve the accuracy of our method.

Because our algorithm used the peptide sequence assigned by the database search software such as SEQUEST, it may extract a wrong isotopic cluster as the pair of the isotopic cluster that corresponds to the identified peptide if the search engine assigned an incorrect peptide sequence to the MS/MS of the peptide. To overcome such a problem, different search strategies may be needed, for example, checking all other isotopic clusters that could be paired with the isotopic cluster obtained from peptide mass. However, it leads to a difficult problem (e.g., the peptide identification correction problem) that warrants further research.

Acknowledgment. This work was supported by 21C Frontier Functional Proteomics Project from Korean Ministry of Education, Science & Technology (FPR08-A1-020, FPR08-A1-021, FPR08-A1-090).

References

(1) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. *J. Am. Soc. Mass Spectrom.* **2000**, *11* (4), 320–332.
(2) Park, K.; Yoon, J. Y.; Lee, S.; Paek, E.; Park, H.; Jung, H. J.; Lee, S. W. Isotopic peak intensity ratio based algorithm for determination of isotopic clusters and monoisotopic masses of polypeptides from high-resolution mass spectrometric data. *Anal. Chem.* **2008**, *80* (19), 7294–7303.
(3) Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. *Nature* **2003**, *422* (6928), 198–207.
(4) Mann, M.; Wilm, M. Error Tolerant Identification of Peptides in Sequence Databases by Peptide Sequence Tags. *Anal. Chem.* **1994**, *66* (24), 4390–4399.
(5) Eng, J. K.; Mccormack, A. L.; Yates, J. R. An Approach to Correlate Tandem Mass-Spectral Data of Peptides with Amino-Acid-Sequences in a Protein Database. *J. Am. Soc. Mass Spectrom.* **1994**, *5* (11), 976–989.
(6) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20* (18), 3551–3567.
(7) Mueller, L. N.; Brusniak, M. Y.; Mani, D. R.; Aebersold, R. An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *J. Proteome Res.* **2008**, *7* (1), 51–61.
(8) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **1999**, *17* (10), 994–999.
(9) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M.; Stable isotope labeling by amino acids in cell culture, SILAC. as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002**, *1* (5), 376–386.
(10) Yao, X. D.; Freas, A.; Ramirez, J.; Demirev, P. A.; Fenselau, C. Proteolytic O-18 labeling for comparative proteomics: Model studies with two serotypes of adenovirus. *Anal. Chem.* **2001**, *73* (13), 2836–2842.
(11) Stewart, I. I.; Thomson, T.; Figeys, D. O-18 Labeling: a tool for proteomics. *Rapid Commun. Mass Spectrom.* **2001**, *15* (24), 2456–2465.
(12) Sechi, S.; Chait, B. T. Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal. Chem.* **1998**, *70* (24), 5150–5158.
(13) Sechi, S. A method to identify and simultaneously determine the relative quantities of proteins isolated by gel electrophoresis. *Rapid Commun. Mass Spectrom.* **2002**, *16* (15), 1416–1424.
(14) DeSouza, L. V.; Taylor, A. M.; Li, W.; Minkoff, M. S.; Romaschin, A. D.; Colgan, T. J.; Siu, K. W. M. Multiple reaction monitoring of mTRAQ-labeled peptides enables absolute quantification of endogenous levels of a potential cancer marker in cancerous and normal endometrial tissues. *J. Proteome Res.* **2008**, *7* (8), 3525–3534.
(15) 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. Multiplexed protein quantitation in

Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3* (12), 1154–1169.
(16) Dayon, L.; Hainard, A.; Licker, V.; Turck, N.; Kuhn, K.; Hochstrasser, D. F.; Burkhard, P. R.; Sanchez, J. C. Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Anal. Chem.* **2008**, *80* (8), 2921–2931.
(17) Han, D. K.; Eng, J.; Zhou, H. L.; Aebersold, R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.* **2001**, *19* (10), 946–951.
(18) Li, X. J.; Zhang, H.; Ranish, J. A.; Aebersold, R. Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. *Anal. Chem.* **2003**, *75* (23), 6648–6657.
(19) Schulze, W. X.; Mann, M. A novel proteomic screen for peptide-protein interactions. *J. Biol. Chem.* **2004**, *279* (11), 10756–10764.
(20) Shinkawa, T.; Taoka, M.; Yamauchi, Y.; Ichimura, T.; Kaji, H.; Takahashi, N.; Isobe, T. STEM: A software tool for large-scale proteomic data analyses. *J. Proteome Res.* **2005**, *4* (5), 1826–1831.
(21) Halligan, B. D.; Slyper, R. Y.; Twigger, S. N.; Hicks, W.; Olivier, M.; Greene, A. S. ZoomQuant: An application for the quantitation of stable isotope labeled peptides. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (3), 302–306.
(22) Bellew, M.; Coram, M.; Fitzgibbon, M.; Igra, M.; Randolph, T.; Wang, P.; May, D.; Eng, J.; Fang, R. H.; Lin, C. W.; Chen, J. Z.; Goodlett, D.; Whiteaker, J.; Paulovich, A.; McIntosh, M. A suite of algorithms for the comprehensive analysis of complex protein mixtures using high-resolution LC-MS. *Bioinformatics* **2006**, *22* (15), 1902–1909.
(23) Faca, V.; Coram, M.; Phanstiel, D.; Glukhova, V.; Zhang, Q.; Fitzgibbon, M.; McIntosh, M.; Hanash, S. Quantitative analysis of acrylamide labeled serum proteins by LC−MS/MS. *J. Proteome Res.* **2006**, *5* (8), 2009–2018.
(24) Monroe, M. E.; Tolic, N.; Jaitly, N.; Shaw, J. L.; Adkins, J. N.; Smith, R. D. VIPER: an advanced software package to support high-throughput LC-MS peptide identification. *Bioinformatics* **2007**, *23* (15), 2021–2023.
(25) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367–1372.
(26) Park, S. K.; Venable, J. D.; Xu, T.; Yates, J. R. A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat. Methods* **2008**, *5* (4), 319–322.
(27) Dasari, S.; Wilmarth, P. A.; Reddy, A. P.; Robertson, L. J. G.; Nagalla, S. R.; David, L. L. Quantification of Isotopically Overlapping Deamidated and O-18-Labeled Peptides Using Isotopic Envelope Mixture Modeling. *J. Proteome Res.* **2009**, *8* (3), 1263–1270.
(28) Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **2002**, *74* (20), 5383–5392.
(29) Senko, M. W.; Beu, S. C.; Mclafferty, F. W. Determination of Monoisotopic Masses and Ion Populations for Large Biomolecules from Resolved Isotopic Distributions. *J. Am. Soc. Mass Spectrom.* **1995**, *6* (4), 229–233.

PR100183T