Automated Comparative Proteomics Based on Multiplex Tandem Mass Spectrometry and Stable Isotope Labeling*

Guoan Zhang and Thomas A. Neubert‡

Comparative proteomic approaches using isotopic labeling and MS have become increasingly popular. Conventionally quantification is based on MS or extracted ion chromatogram (XIC) signals of differentially labeled peptides. However, in these MS-based experiments, the accuracy and dynamic range of quantification are limited by the high noise levels of MS/XIC data. Here we report a quantitative strategy based on multiplex (derived from multiple precursor ions) MS/MS data. One set of proteins was metabolically labeled with [13C6]lysine and [15N4]arginine; the other set was unlabeled. For peptide analysis after tryptic digestion of the labeled proteins, a wide precursor window was used to include both the light and heavy versions of each peptide for fragmentation. The multiplex MS/MS data were used for both protein identification and quantification. The use of the wide precursor window increased sensitivity, and the y ion pairs in the multiplex MS/MS spectra from peptides containing labeled and unlabeled lysine or arginine offered more information for, and thus the potential for improving, protein identification. Protein ratios were obtained by comparing intensities of y ions derived from the light and heavy peptides. Our results indicated that this method offers several advantages over the conventional XIC-based approach, including increased sensitivity for protein identification and more accurate quantification with more than a 10-fold increase in dynamic range. In addition, the quantification calculation process was fast, fully automated, and independent of instrument and data type. This method was further validated by quantitative analysis of signaling proteins in the EphB2 pathway in NG108 cells. Molecular & Cellular Proteomics 5:401–411, 2006.

In recent years, various strategies based on stable isotope coding and MS have been established and have become increasingly popular as alternatives to two-dimensional PAGE-based methods (1, 2) for quantitative proteomics (3–8). In these methods, proteins or peptides from different samples are differentially labeled using stable isotopes, and protein quantification is achieved by comparing their relative abundances in MS. Stable isotopes can be introduced through chemical reactions or metabolic incorporation during cell culture. Although both strategies have characteristic strengths and limitations (6), metabolic labeling seems to have gained more attention recently (4, 5, 8, 9). The labeling process is straightforward and highly efficient. Unlabeled and labeled samples can be combined directly after, or sometimes even prior to, cell lysis and treated as a single sample in all subsequent steps, thus minimizing errors introduced during sample preparation.

In most comparative proteomic studies, differentially labeled peptides are analyzed by LC-MS/MS in a data-dependent manner in which MS/MS scans are automatically triggered after MS survey scans. In these experiments, MS/MS data are used for protein identification, and MS data are used for quantification. Protein ratios can then be determined by comparing the relative intensities of MS signals from differentially labeled peptides. Alternatively quantification can be based on extracted ion chromatograms (XICs)† of the MS signal. However, both MS and XIC signals are often strongly affected by background noise, which can limit quantification accuracy and dynamic range (10, 11) especially for certain types of mass spectrometers such as ion traps (11). Background noise is especially problematic in quantitative experiments because the relative contribution of even a constant background signal is greater for low abundance peptides than for higher abundance peptides. In addition, when samples are of high complexity, co-eluting isobaric peptides or contaminants can interfere with target signals.

An alternative strategy is to use MS/MS data for quantification. Although much less frequently used than MS-based strategies, MS/MS-based quantification offers great potential for enhancing accuracy and dynamic range because MS/MS data generally have much better s/n than XIC and MS data (9, 11–13). Especially in ion traps, small signals with low s/n can be accumulated in the trap before fragmentation to achieve MS/MS spectra with much higher s/n than MS spectra (11). Several previous studies have used this strategy to measure protein ratios. In two recent reports, peptides were labeled with isobaric tags that contain labile bonds easily cleaved

---

† The abbreviations used are: XIC, extracted ion chromatogram; IT, ion trap; s/n, signal to noise ratio.
Comparative Proteomics Based on Multiplex MS/MS

during MS/MS, and the resulting fragment ions are of different masses and can be used for quantification (14, 15). However, both methods rely on fragment ions from the tags and thus cannot be used for metabolic isotope labeling experiments, which have the advantages discussed above. Backbone fragment ions have also been used for quantification in small scale analyses. In one study, a target protein of known sequence was quantified by using fragment ions from peptides labeled with d<sub>0</sub>/d<sub>3</sub> on cysteine residues (12). In another study, Heller et al. (16) used MS/MS spectra to quantify 16O/18O-labeled peptides, but their result suggested MS/MS quantification did not provide an advantage over MS and MS-XIC approaches. A more recent study by Venable et al. (11) established a novel quantitative approach using "data-independent" acquisition, which is based on the sequential isolation and fragmentation of 10 m/z precursor window increments within the ion trap until a desired mass range has been covered. Protein quantification is achieved by reconstructing XICs from fragment ion intensities of MS/MS spectra. However, this approach requires very fast scan speeds by the mass spectrometer used, and use of predetermined precursor ion selection windows can differentially influence signal intensities for each member of a labeled and unlabeled peptide pair when one member of the pair is near a selection window boundary.

Another difficulty for comparative proteomics has been the lack of automated software tools for calculations of protein ratios. Although a number of computer programs have been developed, to our knowledge, only a few of them have achieved relative popularity (11, 17, 18). To date, a considerable proportion of quantitative proteomic analysis has been carried out manually. This is most likely due to the fact that most software programs are not highly automated, and more importantly, instruments and raw data are format-specific, which has greatly limited their use.

To address these difficulties in quantitative proteomic analysis and to further exploit the potential of MS/MS-based quantification, we developed an MS/MS-based strategy for comparative proteomics. In our method, a wide precursor window was applied to include both the light and heavy peptides for simultaneous fragmentation. The resulting multiplex MS/MS data were used for protein identification by Mascot and quantification by a simple Perl program (MS2Ratio) that compares fragment ion intensities derived from the light and heavy peptides. Our results indicated that this method offers several advantages over conventional MS-based ones, including increased sensitivity for protein identification and more accurate quantification with an extended dynamic range. In addition, the quantification process was fast, fully automated, and independent of instrument data type. The method was further validated by analyzing signaling proteins in the EphB2 signaling pathway.

EXPERIMENTAL PROCEDURES

Cell Culture and Metabolic Labeling—Two populations of NG108 cells (mouse neuroblastoma and rat glioma hybrid) stably overexpressing EphB2 were maintained in Lys- and Arg-depleted Dulbeccco’s modified Eagle’s medium (Special Media, Philisburg, NJ) supplemented with 10% dialyzed fetal bovine serum (Invitrogen), hypoxanthine-aminopterin-thymidine (Sigma), 100 units/ml penicillin/streptomycin (Invitrogen), and either normal or [13C6]Lys and [15N4]Arg (Cambridge Isotope Labs, Andover, MA), respectively. Cells were grown for at least six divisions to allow full incorporation of labeling amino acids.

Preparation of Labeled/Unlabeled Samples of Known Ratios—After metabolic labeling, the cells were lysed in lysis buffer containing 1% Triton X-100, 150 mM NaCl, 20 mM Tris, pH 8, 0,2 mM EDTA, 2 mM Na<sub>2</sub>VO<sub>4</sub>, 2 mM NaF, and protease inhibitors (Complete tablet; Roche Applied Science). Lysates were clarified by centrifugation at 14,000 × g for 10 min. Labeled and unlabeled lysates were mixed at 1:1, 1:3, 1:10, 1:30, and 1:100 ratios. The labeled sample and the 1:1 mixed sample were analyzed by MALDI MS (Micromass Q-TOF Ultima, Waters) after digestion to ensure complete labeling and equal mixing. Mixed lysates were separated on a 7.5% Tris-HCl gel (Bio-Rad). The gel was stained with Coomassie Blue for protein visualization. For each sample, the protein band corresponding to ~75 kDa was excised and digested in-gel with trypsin.

Preparation of Samples for Study of EphB2 Signaling—After metabolic labeling, the cells were serum-starved for 24 h. Cells cultured in heavy Lys/Arg medium (about 1 × 10<sup>7</sup>) were treated for 45 min with 2 µg/ml ephrinB1-Fc (Sigma) that was preclustered using goat anti-human Fc (Jackson Immunoresearch Laboratories, West Grove, PA) by incubating ephrinB1-Fc (250 µg/ml) and anti-human Fc (65 µg/ml) at 4 °C for 1.5 h. The cells were then lysed in 1 ml of lysis buffer containing 1% Triton X-100, 150 mM NaCl, 20 mM Tris, pH 8, 0.2 mM EDTA, 2 mM Na<sub>2</sub>VO<sub>4</sub>, 2 mM NaF, and protease inhibitors (Complete tablet; Roche Applied Science). The lysate was clarified by centrifugation at 14,000 × g for 10 min and precleared by incubating with 20 µl of protein A-agarose beads (Pierce) at 4 °C for 1 h. The lysate was incubated with 20 µl (10 µg) of agarose-conjugated anti-Tyr(P) antibody PY-99 (Santa Cruz Biotechnology, Santa Cruz, CA) for 4 h. After incubation, the beads were washed four times, each with 1 ml of lysis buffer for 1 min. Precipitated proteins were then eluted by boiling the beads in SDS-PAGE sample buffer for 5 min and separated on a 10% Tris-HCl gel (Bio-Rad). The gel was stained with Coomassie Blue for protein visualization.

In-solution and In-gel Digestion—For in-solution digestion, BSA was incubated in 25 mM NH<sub>4</sub>HCO<sub>3</sub> with trypsin (Promega, Madison, WI) at a ratio of 1:50 (enzyme/protein) for 4 h at 37 °C after heat denaturation of target proteins at 95 °C for 5 min. In-gel digestion was performed using a modified version of the protocol developed by Shevchenko et al. (19). Briefly, excised gel bands were cut into small pieces and destained in 25 mM NH<sub>4</sub>HCO<sub>3</sub> 50% acetonitrile; dehydrated with acetonitrile; and dried. Then the gel pieces were rehydrated with 12.5 ng/µl trypsin solution (in 25 mM NH<sub>4</sub>HCO<sub>3</sub>) and incubated overnight at 37 °C. Peptides were extracted twice with 5% formic acid, 50% acetonitrile followed by a final extraction with acetonitrile. Samples were dried with vacuum centrifugation before further preparation or analysis.

Mass Spectrometry—An LCO-DecaXP (ThermoFinnigan, San Jose, CA) equipped with a nano-ESI source (Jake Hill Instrument Services, Arlington, MA) was used for all ion trap experiments. For direct infusion analysis, the samples were delivered with an infusion pump at a flow rate of 1 µl/min. For LC-IT-MS/MS analysis, an Ultimate HPLC system (LC Packings, Sunnyvale, CA) equipped with a 180-µm × 15-cm C<sub>18</sub> column (LC Packings PepMap) was coupled to the ion trap instrument via a 10-µm-inner diameter PicoTip™ emitter.
Quantification is based directly or indirectly (through XICs) on MS data. A small precursor window is used to isolate a single peptide to obtain MS/MS data for protein identification. Quantification is based directly or indirectly (through XICs) on MS data.

**RESULTS**

The MS/MS-based Strategy—In conventional methods for quantitative proteomic analysis by LC-MS/MS, protein identification is based on MS/MS spectra, whereas quantification is based directly or indirectly (through XICs) on MS spectra (Fig. 1A). In our multiplex MS/MS-based method, a wide precursor window (10 m/z) was applied so that both the labeled and unlabeled peptide peaks were included for simultaneous fragmentation. Both Lys and Arg labeling was used in combination with tryptic digestion so that, in multiplex MS/MS spectra, the C-terminal ions (typically y ions) appeared as doublets with predictable mass differences, whereas N-terminal ions (typically b ions) appeared as singlet peaks. Because the multiplex MS/MS spectra contain all the sequence information of standard uniplex MS/MS spectra, these multiplex spectra can be used for protein identification. Because the relative intensities of the C-terminal ions reflect the relative...
Comparative Proteomics Based on Multiplex MS/MS

peptide abundances, the same MS/MS data also can be used for quantification (shown in Fig. 1B). To ensure accurate quantification, the labeled and unlabeled peptides must have the same retention time in LC separations. Here we used 15N- and 13C-amino acids for labeling; they have been shown to have no detectable effect on peptide retention time in reversed phase LC (10, 20). Lys and Arg labelings produce different mass shifts, so that, with a multiplex MS/MS spectrum, the C terminus of the peptide can be easily determined from the mass difference of the y ion pairs, which provides additional information for protein identification (13).

Software for Quantification—In our method, Mascot (www.matrixscience.com/) was used for protein identification. Mascot is a powerful commercially available search engine that has achieved considerable popularity for proteomic studies (21). It supports data from most commercial MS instruments and provides a friendly user interface to allow convenient presentation of results and data visualization. We took advantage of these features by incorporating the quantification program MS2Ratio into the reporting system of Mascot. The procedure for calculating relative protein ratios using MS2Ratio is outlined in Fig. 2, and a more detailed description can be found under “Experimental Procedures.”

During development of the software, we found that two factors need to be taken into consideration to ensure accurate quantification. 1) It is not unusual for an MS/MS spectrum to show isobaric y and b or other fragment ions. In that case, the ratio between the y ion pair no longer reflects the actual peptide ratio. Therefore a stringent filter was applied by MS2Ratio to remove potential interference from isobaric fragment ions. 2) Ratios calculated from low intensity y ion pairs are less accurate than those from high intensity ions, as also mentioned in a previous report (16), usually due to insufficient ion counts. This problem was alleviated in two ways. First, a filter was applied to exclude low intensity y ions. In an MS/MS spectrum, all y ions with intensities below 10% of the most intense y ion were discarded. Second, instead of simply averaging ratios of all quantifiable y ion pairs, the peptide ratio was calculated as the ratio between the sum of all y ion intensities from the labeled and unlabeled peptides such that y ions of higher intensity have more weight than lower intensity ions in the final result.

Protein Identification Using Multiplex MS/MS Data—We used a precursor window of 10 m/z to include peptide doublets. The effect of a wide precursor selection window on MS/MS quality has been adequately addressed in a previous study (11), which showed that a wide window had no significant effect on spectrum quality. For an isolation window as wide as 25 m/z, the precursor ion position relative to the center of the selection window had no effect on fragmentation and Sequest scores (11).

To further test the effect of the 10 m/z window on protein identification of labeled/unlabeled samples, one mixed labeled/unlabeled sample with known ratio of 1:1 and one sample with ratio 1:3 were separated by SDS-PAGE in separate lanes as described under “Experimental Procedures.” One band (apparent molecular weight about 75,000) from the 1:3 lane and another band (molecular weight about 25,000) from the 1:1 lane were excised, digested with trypsin, and combined for LC-IT-MS/MS analysis. Based on Coomassie staining intensities (data not shown), the 25-kDa band from the 1:1 lane contained much more protein than the 75-kDa band from the 1:3 lane; this allowed the mixed sample to mimic a real sample because usually only a small proportion of proteins in metabolic labeling experiments shows significant changes in abundance. Two different precursor windows, 2.5 and 10 m/z, were used for data acquisition, and for each window width, analysis was conducted three times. The Mascot search result (Fig. 3A) shows that more proteins were identified using the 10 m/z window. Based on their quantification results and distinctly different molecular masses (i.e. the ~75-kDa proteins were in a known ratio of 1:3, and 25-kDa proteins were in a known ratio of 1:1), the proteins were assigned into two groups with 1:1 and 1:3 ratios, respectively. In both groups, more proteins were identified using a 10 m/z precursor window, more peptides were identified, and the average number of unique peptides per protein increased from 3.2 to 4.1. However, the average peptide score decreased from 47 to 36 as the selection window increased from 2.5 to 10 m/z.

The decrease in peptide score was not unexpected. One problem we encountered when using multiplex MS/MS data for database searching was that the additional C-terminal ions (from the labeled/unlabeled peptide pairs) could not be matched to theoretical MS/MS spectra derived from database sequences, decreasing Mascot scores. To address this problem, we defined neutral loss modifications for Arg and Lys, respectively, and used a large peptide mass tolerance (7.2 Da)
in database searching. This allowed the extra C-terminal ions to be matched as neutral loss ions. According to the scoring algorithm of Mascot, the extra (neutral loss) matches do not contribute to increased peptide scores. At the same time, the wider selection window allows the introduction of more ions that cannot be matched resulting in the decrease in scores that we observed. However, in general the score decrease seemed to have little effect on success in protein identification. The increased mass tolerance can potentially increase false positive rates for protein identification as suggested by some studies (22, 23), although at least one report indicated that varying mass tolerance had little effect on the accuracy of protein identification (24). This concern is addressed by Mascot by increasing the identification threshold for statistical significance when the mass tolerance increases.

In "Experimental Procedures," the threshold for statistically significant identifications using a 7.2-Da mass tolerance was 46 compared with 39 for a mass error search tolerance of 1.2 Da. It should be noted that Mascot is not tailored for multiplex MS/MS data as our results showed that the more informative information in multiplex MS/MS spectra did not translate into higher Mascot scores. Development of novel correlation algorithms to make full use of the extra C-terminal ion information would further improve the accuracy of protein identification.

Use of a wide precursor ion selection window most likely leads to more peptide and protein identifications because of improved MS/MS sensitivity due to fragmentation of both labeled and unlabeled peptides. Peptides with label abundance ratios near 1:1 benefit the most from this effect. Fig. 3 shows that proteins in a 1:3 ratio also benefited from increased sensitivity based on the increased number of identified proteins. To measure the impact of precursor window width on peptides with more dramatic abundance differences, BSA tryptic peptides (1 pmol/µl) were directly infused into the ion trap. The data were acquired in zoom scan mode with automatic gain control turned off and an injection time of 5 ms/scan. A BSA peptide, LGEYGFQNALIVR (740.4 m/z), was isolated using different precursor selection window widths with the fragmentation energy set to 0 so that no fragmentation of the peptides occurred. For each window width, data were acquired for 5 min, and the average peak areas (including all isotopic peaks) from each experiment were compared to indicate the isolation efficiency. Fig. 3 clearly shows that the 10 m/z window allowed accumulation of considerably more ions than the 2.5 m/z window, suggesting the large window can improve sensitivity even when peptides are of an

**Fig. 3.** Effect of a wide precursor ion selection window on protein identification. A, effect of precursor window width on number of identified proteins and matched peptides. A sample containing proteins of 1:1 labeled/unlabeled ratio and another sample containing different proteins at a 1:3 ratio were combined for tryptic digestion, and the resulting peptide mixture was analyzed by LC-IT-MS/MS using 2.5 and 10 m/z precursor selection windows. Panels from left to right, total number of proteins identified, number of identified proteins of 1:3 ratio, and total number of identified peptides. Mascot was used for database searching. B, effect of precursor window width on intensity of selected ions. Tryptic peptides from unlabeled BSA were analyzed by ESI-ion trap MS by direct infusion. An ion of m/z 740.4 corresponding to the doubly charged peptide LGEYGFQNALIVR was used for this experiment. No CID energy was applied, and therefore the peptide was not fragmented. Normalized peak areas representing all isotopic peaks were plotted against precursor window width.

**Fig. 4.** The effect of a precursor window of 10 m/z on selection of precursor ions for quantification using MS/MS spectra. Tryptic peptides from BSA were detected by ESI-ion trap MS with direct infusion. Ions of m/z 461.7 and 464.3 are from the doubly charged peptides AEFVEVTK and YLYEIAF, respectively. A, MS scan. B, ion of m/z 461.7 was selected. C, ion of m/z 464.3 was selected. No CID energy was applied in these experiments, therefore the peptides were maintained in their intact forms. Int, relative intensity.
Comparative Proteomics Based on Multiplex MS/MS

Fig. 5. Comparison of protein quantification using the multiplex MS/MS-based method and the XIC-based method. The expected protein labeled/unlabeled ratios are 1:1 (A), 1:3 (B), 1:10 (C), 1:30 (D), and 1:100 (E). Each bar in the figures represents the calculated ratios for each protein with solid bars representing ratios obtained by the multiplex MS/MS method and open bars representing XIC-based results. The numbers of total identified unique peptides/numbers of quantifiable unique peptides obtained by the two methods are shown below the bars. The peptide numbers for the multiplex MS/MS method are given first followed by the numbers for the XIC-based method in parentheses.
Quantification with Multiplex MS/MS Data—One major concern we had about quantification with multiplex MS/MS data is that the labeled and unlabeled precursor ions might be selected with unequal efficiency, which would introduce an error in relative quantification by MS/MS. To evaluate the impact of the relative position of the precursor ion within the selection window on selection efficiency, pseudo-MS/MS spectra of two BSA peptides with similar masses, AEFVEVTK and YLYEIAR, were collected with the 10 m/z precursor window centered at their doubly charged ion masses 461.7 and 464.3 m/z, respectively. To compare isolation efficiencies, spectra were acquired in MS mode with fragmentation energy set to 0. Fig. 4 shows that the 10 m/z window selected the peptides of the doublet almost equally well with the peak in the window center having slightly higher selection efficiency than the other peak (1:0.97). This value was used by MS2Ratio to correct the final calculations of labeled/unlabeled peptide ratios.

To evaluate the performance of the new method, we carried out LC-IT-MS/MS analyses on mixtures of [13C6]Lys/[15N4]Arg-labeled and unlabeled protein at ratios of 1:1, 1:3, 1:10, 1:30, and 1:100. Labeled and unlabeled NG108-EphB2 cell lysates at those ratios were subjected to SDS-PAGE. A gel band of approximate molecular mass 75 kDa was digested with trypsin to produce peptides for mass spectrometry analysis. DTA files were generated from the data and submitted to Mascot for protein identification before MS2Ratio quantification. To present a comparison between the MS/MS-based strategy and the conventional XIC-based strategy, the same DTA files were used for protein identification before MS2Ratio quantification. To present a comparison between the MS/MS-based strategy and the conventional XIC-based strategy, the same DTA files were used for protein identification by Sequest, and the identified proteins were quantified with ASAPRatio, an XIC-based quantification tool that has been used in many previous studies (18). Fig. 5 shows a comparison of quantification results for proteins that were confidently identified in both analyses.

For the sample of 1:1 ratio, the MS/MS-based method calculated ratios within 30% of the true values for each of the identified proteins, indicating the method is capable of detecting relatively small changes in protein abundance. Another important observation about the MS/MS quantification result is the wide dynamic range as ratios calculated from the 1:100 samples were still reasonably accurate. The plot of the calculated ratios against the true ratios indicates a dynamic range of 2 orders of magnitude (Fig. 6A). In comparison to the MS/MS-based approach, the XIC-based method exhibited more significant errors and a considerably narrower dynamic range (Figs. 5 and 6B). This result appears inconsistent with a previous report (18) on ASAPRatio. However, in that report, the method validation was solely based on the quantification result of one standard protein of high abundance. Our data suggest that, for low abundance proteins (generally proteins with fewer matched peptides), the quantification using this method is much less reliable (Fig. 5).

In our experiments, the number of quantifiable proteins was not affected by the filters applied by MS2Ratio to remove non-quantifiable MS/MS spectra and y ions. In most cases non-quantifiable MS/MS spectra were from singly charged peptides and peptides with missed cleavage sites and accounted for only 15% of the 891 confidently matched unique peptides in this study. In contrast, with the XIC-based approach, 38% of peptides were not quantifiable because of low s/n XIC peaks. This is especially a problem with low abundance proteins of extreme ratios as shown in Fig. 5. In fact, with the MS/MS-based strategy, all proteins identified in this study were quantifiable by MS2Ratio, including proteins identified by only one or two unique peptides. This most likely was not a coincidence because singly charged precursor ions generally do not generate MS/MS spectra with sufficient quality to allow confident protein identification based on one or two peptides. High quality MS/MS spectra can, in turn, provide accurate quantification, which might explain the fact that accuracy and precision of quantification by MS2Ratio did not decrease with low numbers of unique peptides matched per protein.

In each MS/MS spectrum from the 891 uniquely identified peptides, an average of 7.9 y ion pairs was matched by Mascot. This large number offered enough flexibility to avoid interfering peaks for quantification. Even after exclusion of low intensity ion pairs and non-quantifiable ion pairs by our stringent filters, an average of 3.9 ion pairs (calculated from 756
quantifiable unique peptides) could be used. We expect that the use of a high resolution instrument such as an FT-ICR system (with the possible caveat of relatively low sensitivity in MS/MS mode) or Orbitrap in MS/MS mode would further increase the number of quantifiable y ions because the high resolution would allow better discrimination of y ions from isobaric interfering ions. The fact that quantification based on multiplex MS/MS produced much better results than the strategy based on XIC of MS spectra is likely due to the much higher signal to noise ratio of MS/MS data as mentioned in previous reports (11, 12) and as was frequently observed in our experiments. Fig. 7 shows a comparison between the XIC, MS, and MS/MS data for a peptide pair from the 1:3 labeled/unlabeled sample. The low frequency of MS survey scans (approximately one scan every 5–6 s) and limited resolution of the ion trap led to poorly defined XIC peaks, which in turn made quantification very difficult. In the MS spectrum (Fig. 7B), the peptide s/n was only about 2, which will inevitably lead to a large quantitative error. In contrast, the MS/MS spectrum was much cleaner. With dominant y ions of good s/n and free of isobaric interference, the ratio was able to be measured accurately.

**Demonstration of Multiplex MS/MS Quantification Using Signaling Proteins in the EphB2 Pathway**—For this experiment, NG108-EphB2 cells were cultured in [13C6]Lys/[13C6]Arg medium or control medium with amino acids of natural isotopic abundance. The [13C]-labeled cells were stimulated with ephrinB1 to activate EphB2 receptors, the cells were lysed, and stimulated and unstimulated cell lysates were combined at a 1:1 ratio. Tyr(P) proteins were isolated by Tyr(P) immunoprecipitation and then separated by SDS-PAGE. Western blotting of the same sample using a Tyr(P) antibody (PY99) indicated that proteins with apparent molecular weight of 100,000–150,000 were heavily tyrosine phosphorylated (data not shown). The gel section corresponding to this molecular weight range was excised accordingly and digested with trypsin before analysis by LC-IT-MS/MS. The quantification results are shown in Table I.

Three proteins in this molecular weight range, including the EphB2 receptor, were identified as significantly regulated proteins upon ligand stimulation. RAS p21 protein activator 1 is a known effector in EphB2 signaling (25), and the up-regulation of Hrs was not surprising as EphB2 activation has been shown to lead to rapid endocytosis (26–28). Two proteins with ratios slightly less than 1, CAS and PTK2, have been shown to be dephosphorylated in response to ephrin signaling (29). Each of these results agrees closely with a biologically similar MS-based SILAC (stable isotope labeling by amino acids in cell culture) experiment performed in our laboratory using 10 times as many cultured cells (ratios were 14 for EphB2, 4.0 for RAS p21 protein activator 1, 2.9 for Hrs, 0.75 for CAS, and 0.67 for PTK2). Interestingly the increased ratios for EphB2 obtained in this experiment compared with the MS-based experiment may reflect the increased dynamic range of the
TABLE I
Quantification of proteins identified in the EphB2 experiment

<table>
<thead>
<tr>
<th>Gi no.</th>
<th>Protein name</th>
<th>Peptide no.</th>
<th>Ratio (heavy/light)</th>
</tr>
</thead>
<tbody>
<tr>
<td>47777351</td>
<td>EphB2</td>
<td>12 (10)</td>
<td>26.90 ± 0.10</td>
</tr>
<tr>
<td>6981460</td>
<td>RAS p21 protein activator 1</td>
<td>2 (2)</td>
<td>6.20 ± 2.22</td>
</tr>
<tr>
<td>9506795</td>
<td>Hrs</td>
<td>2 (2)</td>
<td>4.75 ± 0.56</td>
</tr>
<tr>
<td>31543315</td>
<td>Nucleolin</td>
<td>4 (4)</td>
<td>1.20 ± 0.27</td>
</tr>
<tr>
<td>7949051</td>
<td>hnRNP U</td>
<td>2 (2)</td>
<td>1.13 ± 0.17</td>
</tr>
<tr>
<td>42558248</td>
<td>GPI-anchored membrane protein 1</td>
<td>2 (2)</td>
<td>1.12 ± 0.14</td>
</tr>
<tr>
<td>14389299</td>
<td>Vimentin</td>
<td>5 (5)</td>
<td>0.89 ± 0.17</td>
</tr>
<tr>
<td>6755863</td>
<td>Tumor rejection antigen gp96</td>
<td>3 (3)</td>
<td>0.86 ± 0.15</td>
</tr>
<tr>
<td>6881236</td>
<td>Myosin</td>
<td>2 (2)</td>
<td>0.84 ± 0.06</td>
</tr>
<tr>
<td>40254593</td>
<td>CAS</td>
<td>5 (5)</td>
<td>0.80 ± 0.17</td>
</tr>
<tr>
<td>6679741</td>
<td>PTK2</td>
<td>12 (10)</td>
<td>0.74 ± 0.18</td>
</tr>
<tr>
<td>6981460</td>
<td>Pancreatic trypsin 1</td>
<td>2 (2)</td>
<td>0.00 ± 0.00</td>
</tr>
<tr>
<td>47059057</td>
<td>Immunoglobulin heavy chain</td>
<td>4 (3)</td>
<td>0.01 ± 0.01</td>
</tr>
<tr>
<td>34868298</td>
<td>Keratin</td>
<td>2 (2)</td>
<td>0.01 ± 0.00</td>
</tr>
</tbody>
</table>

MS/MS-based method. Immunoglobulin heavy chain, pancreatic trypsin, and keratin had labeled/unlabeled ratios of nearly 0, which would be expected of completely unlabeled contaminating proteins and which demonstrated the high dynamic range of the new method.

DISCUSSION

We developed a method that uses multiplex MS/MS data for relative quantification of protein abundance because of limitations in conventional quantification methods based on MS spectra that we experienced in our laboratory. Our results indicate that this method shows significant improvement in accuracy, precision, and dynamic range over a conventional quantification method based on XIC of MS data. We feel that the increased dynamic range of our MS/MS-based method is especially important. In previous reports of quantitative proteomic experiments, differences of more than 10-fold in protein levels between samples have rarely been reported; we feel this can be at least partially attributed to insufficient dynamic range of conventional quantitative approaches. For some experiments such as protein turnover and time course studies, dramatic protein ratio changes can be expected, and a wide dynamic range for the analysis becomes very important (30).

The benefits offered by the new method are mainly due to the fact that MS/MS spectra inherently have lower background and less interference from other ions than MS spectra. It has been shown that spectral background is the major factor that greatly limits the accuracy and dynamic range for quantitative methods based directly or indirectly on MS signals (10, 11). This problem is especially serious for low s/n data (11) largely due to unequal relative contribution of background signals to high and low intensity signals generated by peptides of interest. Results from other experiments in our laboratory (data not shown) indicated that the same is true even when the quantification was based directly on MS spectra (not XICs) from a Q-TOF instrument, which is known for its high resolution and low background level compared with ion traps. To overcome this obstacle, some methods (11, 18) have used stringent filters to exclude low s/n signals from quantification. However, a problem with this strategy is that it decreases the number of quantifiable peptides, making it difficult to quantify low abundance proteins. Another strategy to remove noise is to subtract an average level of noise from the spectra. This can be problematic because it is well known that s/n is dependent on spectrum acquisition time, and the s/n level can be increased simply by using longer acquisition time per spectrum or summing more spectra, leading to variability in background when the total acquisition time for compared spectra is not equal. The best solution to this problem, we believe, is to obtain cleaner spectra in the first place rather than to try to differentiate signals from noise. By using MS/MS for quantification, this background problem was readily circumvented.

In this new method, we used multiplex MS/MS data for protein identification that are more informative than conventional MS/MS data and have been shown to be able to greatly assist in peptide sequencing (13, 31–35). In this study we used Lys and Arg labeling; these residues are at the C termini of tryptic peptides. Therefore, in the multiplex MS/MS spectra, C-terminal fragment ions can be easily recognized as doublet peaks. In addition, Lys and Arg labeling introduced additional mass shifts, which helped to determine the C-terminal residue of the peptides. This additional information is greatly helpful for manual inspection of spectra or de novo sequencing. With the development of new database searching and de novo peptide sequencing algorithms that can make full use of the rich information in multiplex MS/MS data, we propose that the protein identifications could be further improved by using multiplex MS/MS data. Another benefit is that the wide precursor window we used to acquire multiplex MS/MS data allowed precursor ions to be isolated at higher efficiency and resulted in improved sensitivity for protein identification.

For quantitative analysis based on data-dependent acquisition when using MS/XIC-based quantification approaches, more frequent MS scans are desirable for better quantification. However, this is at the expense of protein identification because less time is spent on MS/MS scans. Therefore a careful compromise has to be made between identification and quantification especially when LC peaks are narrow. In contrast, this is not a problem with our MS/MS-based strategy because MS/MS data are used for both identification and quantification. This provides more flexibility in experiment
Comparative Proteomics Based on Multiplex MS/MS

design such that conditions can be optimized for protein identification.

In our strategy, MS/MS data in the format of simple text files were used for quantification, offering several advantages. 1) No raw data are needed for quantification, and therefore the program can work with data from any type of instruments. 2) Quantification can be done with simple and straightforward algorithms, and programming is greatly simplified. 3) Quantification is fully automated and much faster than methods in which raw data are involved. For data from a typical LC-MS/MS run containing hundreds of identified peptides, protein ratios can be calculated within seconds.

In this study, we used the multiplex MS/MS method for relative quantification of a metabolic labeling experiment. However, the method should also be applicable to other C/N-terminal labeling techniques provided that the labeling does not cause chromatographic separation and the mass shift is small enough to allow the use of a precursor window of reasonable width. Many previously reported labeling methods such as 18O labeling through tryptic digestion (36, 37) can be used without modifying the established protocols.

In summary, we developed an automated quantification method based on multiplex MS/MS data. The applicability of the new method was demonstrated with the analysis of a subset of signaling proteins from the EphB2 pathway. The new method offers accurate quantification with a dynamic range of 2 orders of magnitude. The wide precursor window is small enough to allow the use of a precursor window of reasonable width. Many previously reported labeling methods such as 18O labeling through tryptic digestion (36, 37) can be used without modifying the established protocols.

Acknowledgments—Brian Fernholz is greatly appreciated for help with the software development and valuable discussions about Mascot software. We thank Dr. Vivek Shetty for expert help in LC-MS/MS experiments and Dr. John Cottrell from Matrix Science for advice on database searching.

* This work was supported by National Institutes of Health Grants R21 NS 44184, P30 NS050276, and S10 RR 017990-01 and NCI, National Institutes of Health Core Grant 2P30 CA 016087 (to T. A. N.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 212-263-7265; Fax: 212-263-8214; E-mail: neubert@saturn.med.nyu.edu.

REFERENCES

Juxtamembrane tyrosine residues couple the Eph family receptor EphB2/Nuk to specific SH2 domain proteins in neuronal cells. EMBO J. 16, 3877–3888