Optimized Method for Computing (18)O/(16)O Ratios of Differentially Stable-Isotope Labeled Peptides in the Context of Postdigestion (18)O Exchange/Labeling

ARTICLE *in* ANALYTICAL CHEMISTRY · JULY 2010

Impact Factor: 5.64 · DOI: 10.1021/ac101284c · Source: PubMed

Anal. Chem. **2010,** *82,* 5878–5886

Optimized Method for Computing $^{18}O/^{16}O$ Ratios of Differentially Stable-Isotope Labeled Peptides in the Context of Postdigestion ^{18}O Exchange/Labeling

Xiaoying Ye,[†] Brian T. Luke,[‡] Donald J. Johann, Jr.,[§] Akira Ono,[||] DaRue A. Prieto,[†] King C. Chan,[†] Haleem J. Issaq,[†] Timothy D. Veenstra,[†] and Josip Blonder*[,†]

Laboratory of Proteomics and Analytical Technologies, Advanced Technology Program, SAIC-Frederick, Inc., National Cancer Institute at Frederick, Maryland 21702, Advanced Biomedical Computing Center, Advanced Technology Program, SAIC-Frederick, Inc., National Cancer Institute at Frederick, Maryland 21702, National Cancer Institute, Center for Cancer Research, Medical Oncology Branch, Bethesda, Maryland 20892, and Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan 48109

Differential $^{18}O/^{16}O$ stable isotope labeling of peptides that relies on enzyme-catalyzed oxygen exchange at their carboxyl termini in the presence of $H_2{}^{18}O$ has been widely used for relative quantitation of peptides/proteins. The role of tryptic proteolysis in bottom-up shotgun proteomics and low reagent costs have made trypsin-catalyzed ^{18}O postdigestion exchange a convenient and affordable stable isotope labeling approach. However, it is known that trypsin-catalyzed ^{18}O exchange at the carboxyl terminus is in many instances inhomogeneous/incomplete. The extent of the ^{18}O exchange/incorporation fluctuates from peptide to peptide mostly due to variable enzyme−substrate affinity. Thus, accurate calculation and interpretation of peptide ratios are analytically complicated and in some regard deficient. Therefore, a computational approach capable of improved measurement of actual ^{18}O incorporation for each differentially labeled peptide pair is needed. In this regard, we have developed an algorithmic method that relies on the trapezoidal rule to integrate peak intensities of all detected isotopic species across a particular peptide ion over the retention time, which fits the isotopic manifold to Poisson distributions. Optimal values for manifold fitting were calculated and then $^{18}O/^{16}O$ ratios derived via evolutionary programming. The algorithm is tested using trypsin-catalyzed ^{18}O postdigestion exchange to differentially label bovine serum albumin (BSA) at *a priori* determined ratios. Both accuracy and precision are improved utilizing this rigorous mathematical approach. We further demonstrate the effectiveness of this method to accurately calculate $^{18}O/^{16}O$ ratios in a large scale proteomic quantitation of detergent resistant membrane microdomains (DRMMs) isolated from cells expressing wild-type HIV-1 Gag and its nonmyristylated mutant.

Protease catalyzed $^{18}O/^{16}O$ stable isotope labeling is a simple method used to differentially label two biologically distinct samples that allows relative quantitation of proteins on a large-scale basis using mass spectrometry (MS)-based proteomics.[1,2] This approach relies on trypsin-catalyzed exchange of two ^{16}O atoms for two ^{18}O atoms at the C-terminal carboxyl group of tryptic peptides, resulting in a mass shift of 4 Da between singly charged, differentially labeled peptides observed in MS^1 spectra.[3] It has been demonstrated that ^{18}O exchange can be decoupled from the protein digestion step, allowing the labeling conditions to be optimized separately.[4] This strategy is currently adopted as a standard $^{18}O/^{16}O$ labeling procedure, allowing minimal consumption of ^{18}O water and simplifying sample handling.[1]

In comparison to other stable isotope labeling techniques, $^{18}O/^{16}O$ stable isotope labeling offers several advantages. First, this labeling technique does not target peptides containing particular amino acids and does not require an additional affinity-based step for labeled peptide enrichment (e.g., ICAT).[5] Second, in contrast to metabolic labeling techniques (e.g., SILAC),[6] $^{18}O/^{16}O$ labeling is amenable to clinically relevant samples, including human plasma/serum or fresh-frozen tissue specimens.[7] Third, during the $^{18}O/^{16}O$ labeling process, every proteolytic fragment is labeled according to an enzyme-specific

* To whom correspondence should be addressed. Dr. Josip Blonder, Laboratory of Proteomics and Analytical Technologies, Advanced Technology Program, SAIC-Frederick, Inc., National Cancer Institute at Frederick, P.O. Box B, Frederick, Maryland 21702. Phone: 301-846-7211. Fax: 301-846-6037. E-mail: blonder@ncifcrf.gov.

[†] Laboratory of Proteomics and Analytical Technologies, Advanced Technology Program, SAIC-Frederick, Inc.

[‡] Advanced Biomedical Computing Center, Advanced Technology Program, SAIC-Frederick, Inc.

[§] National Cancer Institute.

[||] University of Michigan Medical School.

(1) Reynolds, K. J.; Fenselau, C. *Curr. Protoc. Protein Sci.* **2004**, DOI: 10.1002/0471140864.ps2304s34, Unit 23.4.

(2) Ye, X.; Luke, B.; Andresson, T.; Blonder, J. *Brief. Funct. Genomics Proteomics* **2009**, *8*, 136–144.

(3) Schnolzer, M.; Jedrzejewski, P.; Lehmann, W. D. *Electrophoresis* **1996**, *17*, 945–953.

(4) Reynolds, K. J.; Yao, X.; Fenselau, C. *J. Proteome Res.* **2002**, *1*, 27–33.

(5) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17*, 994–999.

(6) Back, J. W.; Notenboom, V.; de Koning, L. J.; Muijsers, A. O.; Sixma, T. K.; de Koster, C. G.; de Jong, L. *Anal. Chem.* **2002**, *74*, 4417–4422.

(7) Qian, W. J.; Monroe, M. E.; Liu, T.; Jacobs, J. M.; Anderson, G. A.; Shen, Y.; Moore, R. J.; Anderson, D. J.; Zhang, R.; Calvano, S. E.; Lowry, S. F.; Xiao, W.; Moldawer, L. L.; Davis, R. W.; Tompkins, R. G.; Camp, D. G., 2nd; Smith, R. D. *Mol. Cell. Proteomics* **2005**, *4*, 700–709.

10.1021/ac101284c
Published on Web 06/11/2010

cleavage pattern (e.g., trypsin, Glu-C). For this reason, $^{18}O/^{16}O$ labeling is well suited for amount-limited samples including laser capture microdissected (LCM) specimens[8] obtained from fresh-frozen and formalin-fixed tissue slices or plasma membrane specimens.[9] Finally, the low cost of reagents along with the ability to be combined with other stable isotope labeling methods represents additional comparative advantages of $^{18}O/^{16}O$ labeling.[10]

However, it is known that the two C-terminal ^{16}O atoms are not always completely exchanged for two ^{18}O atoms during the enzyme catalyzed ^{18}O exchange. This is primarily due to variable peptide-specific enzyme reaction rates K_m[11] and the fact that the majority of labeling experiments are carried out in the presence of 95% enriched $H_2{}^{18}O$. Thus, the isotopic envelope depicting peptide doublets characterized by a variable ^{18}O exchange rate shows a complex pattern.[12] Additionally, the accurate computation of $^{18}O/^{16}O$ ratios for peptides featuring variable/incomplete ^{18}O exchange is not straightforward. It is for these reasons, along with experimental efforts focused on optimization of labeling conditions,[12−14] that several computational methods have been developed for analysis of MS data, in attempt to rectify incomplete ^{18}O incorporation.[12,15−23]

Therefore, we now report and demonstrate results of an algorithm for improved accuracy and reproducibility of $^{18}O/^{16}O$ ratios computation capable of recognizing and accounting for peptide clusters exhibiting variable/incomplete ^{18}O exchange. Our method applies the trapezoidal rule to integrate peak intensities of all detected isotopic species across a particular peptide ion over the retention time. Then it fits the isotopic manifold to Poisson distributions and derives fitting parameters via evolutionary programming. This allows calculation of more accurate $^{18}O/^{16}O$ peptide ratios especially in the presence of variable ^{18}O incorporation. The utility of this algorithm for quantitative proteomics is demonstrated using samples with varied complexity, including ^{18}O labeled bovine serum albumin (BSA) at *a priori* determined ratios, and $^{18}O/^{16}O$ labeled lipid raft membrane proteins from HeLa cells expressing HIV-1 Gag protein in a study of Gag trafficking.

MATERIALS AND METHODS

Reagents. Bovine serum albumin (BSA) and ammonium bicarbonate (NH_4HCO_3) were purchased from Sigma (St. Louis, MO). Sequencing grade trypsin was obtained from Promega (Madison, WI). $H_2{}^{18}O$ (95% pure) was purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA). Bicinchonic acid (BCA) protein assay reagent kit was purchased from Pierce (Rockford, IL). Trifluoroacetic acid (CF_3CO_2H) and formic acid (CHOOH) were purchased from Fluka (Milwaukee, WI). High-performance liquid chromatography (HPLC)-grade methanol (CH_3OH) and acetonitrile (CH_3CN) were obtained from EM Science (Darmstadt, Germany). All solutions were prepared using water purified by a Nanopure II system (Dubuque, IA).

Transfection of HeLa Cells and Isolation of Detergent-Resistant Membrane Microdomains (DRMMs). To express full-length Gag proteins in HeLa cells in the nonbiohazardous context, we transfected approximately 2.5×10^6 HeLa cells with either pCMVNLGagPolRRE/PR$^-$ or pCMVNLGagPolRRE/PR$^-$/1GA, along with pCMV-Rev (a kind gift from Dr. S. Venkatesan) and pCMV-Vphu (a kind gift from Dr. K. Strebel), using Lipofectamine 2000 (Invitrogen, Carlsbad, CA) as described in the manufacturer's instructions. When cotransfected with Rev, pCMVNLGagPolRRE/PR$^-$ expresses wild type Gag and GagPol precursor polyproteins, whereas pCMVNLGagPolRRE/PR$^-$/1GA expresses nonmyristylated derivatives of these polyproteins. pCMV-Vphu encodes the codon-optimized Vpu protein that facilitates release of viruslike particles.[24] After 24 h, homogenates of transfected cells were prepared, postnuclear supernatants of the homogenates were treated with 0.25% Triton X-100, and the detergent-resistant membrane fractions were isolated by equilibrium flotation centrifugation, as described previously.[25] These fractions were diluted 18-fold with 25 mM $(NH_4)_2CO_3$ and subjected to ultracentrifugation at 100 000*g* for 2 h. Pellets were resuspended in 1.1 mL of 25 mM $(NH_4)_2CO_3$ and stored at −80 °C for subsequent proteomic analysis.

Tryptic Digestion and $^{18}O/^{16}O$ Labeling. The differential $^{18}O/^{16}O$ labeling of BSA and DRMMs were performed using a slightly modified method, previously described.[26] Briefly, proteins were first digested by trypsin at an enzyme/protein ratio of 1:50 overnight at 37 °C and lyophilized to dryness by SpeedVac (Thermo Savant, Holbrook, NY). Then the digest was aliquoted and resuspended in 20% CH_3OH/50 mM NH_4HCO_3 prepared in 95% $H_2{}^{18}O$ or $H_2{}^{16}O$, respectively, and incubated with trypsin overnight at 37 °C using an enzyme/protein ratio of 1:20. After labeling, the enzyme activity was terminated by boiling in a water bath for 10 min. Digests were

(8) Zang, L.; Palmer Toy, D.; Hancock, W. S.; Sgroi, D. C.; Karger, B. L. *J. Proteome Res.* **2004**, *3*, 604–612.
(9) Stockwin, L. H.; Blonder, J.; Bumke, M. A.; Lucas, D. A.; Chan, K. C.; Conrads, T. P.; Issaq, H. J.; Veenstra, T. D.; Newton, D. L.; Rybak, S. M. *J. Proteome Res.* **2006**, *5*, 2996–3007.
(10) Blonder, J.; Yu, L. R.; Radeva, G.; Chan, K. C.; Lucas, D. A.; Waybright, T. J.; Issaq, H. J.; Sharom, F. J.; Veenstra, T. D. *J. Proteome Res.* **2006**, *5*, 349–360.
(11) Hedstrom, L. *Chem. Rev.* **2002**, *102*, 4501–4524.
(12) Heller, M.; Mattou, H.; Menzel, C.; Yao, X. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 704–718.
(13) Wang, Y. K.; Ma, Z.; Quinn, D. F.; Fu, E. W. *Anal. Chem.* **2001**, *73*, 3742–3750.
(14) Hajkova, D.; Rao, K. C.; Miyagi, M. *J. Proteome Res.* **2006**, *5*, 1667–1673.
(15) Eckel-Passow, J. E.; Oberg, A. L.; Therneau, T. M.; Mason, C. J.; Mahoney, D. W.; Johnson, K. L.; Olson, J. E.; Bergen, H. R., 3rd *Bioinformatics* **2006**, *22*, 2739–2745.
(16) Halligan, B. D.; Slyper, R. Y.; Twigger, S. N.; Hicks, W.; Olivier, M.; Greene, A. S. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 302–306.
(17) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 320–332.
(18) Mason, C. J.; Therneau, T. M.; Eckel-Passow, J. E.; Johnson, K. L.; Oberg, A. L.; Olson, J. E.; Nair, K. S.; Muddiman, D. C.; Bergen, H. R., 3rd. *Mol. Cell. Proteomics* **2007**, *6*, 305–318.
(19) Ramos-Fernandez, A.; Lopez-Ferrer, D.; Vazquez, J. *Mol. Cell. Proteomics* **2007**, *6*, 1274–1286.
(20) Shinkawa, T.; Taoka, M.; Yamauchi, Y.; Ichimura, T.; Kaji, H.; Takahashi, N.; Isobe, T. *J. Proteome Res.* **2005**, *4*, 1826–1831.
(21) Stewart, I. I.; Thomson, T.; Figeys, D. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 2456–2465.
(22) Park, S. M.; Hwang, I. K.; Kim, S. Y.; Lee, S. J.; Park, K. S.; Lee, S. T. *Proteomics* **2006**, *6*, 1192–1199.
(23) Mueller, L. N.; Brusniak, M. Y.; Mani, D. R.; Aebersold, R. *J. Proteome Res.* **2008**, *7*, 51–61.
(24) Nguyen, K. L.; llano, M.; Akari, H.; Miyagi, E.; Poeschla, E. M.; Strebel, K.; Bour, S. *Virology* **2004**, *319*, 163–175.
(25) Ono, A.; Waheed, A. A.; Joshi, A.; Freed, E. O. *J. Virol.* **2005**, *79*, 14131–14140.

cooled on ice followed by acidification to pH 3 using TFA. The individually digested peptide pools in $H_2{}^{18}O$ and $H_2{}^{16}O$ were mixed lyophilized. The solely ^{18}O labeled BSA digest was also lyophilized in the same fashion. The lyophilized digests were stored at −80 °C and dissolved in 0.1% TFA before LC−MS analysis.

Strong Cation Exchange (SCX) Fractionation of Differentially $^{16}O/^{18}O$-Labeled DRMMs. The SCX of differentially labeled DRMMs digest was performed as previously described.[27] Briefly, the $^{18}O/^{16}O$-labeled digest from HeLa cell DRMM was reconstituted in 100 μL of 45% (v/v) CH_3CN containing 0.1% (v/v) FA immediately prior to SCX chromatography. The digest was resolved into 14 fractions using a microcapillary HPLC system (model 1100, Agilent Technologies Inc., Palo Alto, CA) on a 1 mm × 150 mm, 5 μm, 300 Å, polysulfethyl A SCX column (PolyLC, Inc., Columbia, MD). Mobile phase A was 45% (v/v) CH_3CN and mobile phase B was 45% (v/v) CH_3CN containing 0.5 M ammonium formate (pH 3). Peptide fractions were eluted with an ammonium formate/multistep gradient at a flow rate of 200 μL/min as follows: 1% B/0−2 min, 10% B/62 min, 62% B/82 min, 100% B/85 min.

Nanoflow Reversed-Phase Liquid Chromatography (nanoRPLC)−Tandem Mass Spectrometry (MS^2). NanoflowRPLC−MS^2 analyses were performed in triplicate using an Agilent 1100 nanoflow LC system coupled with hybrid linear ion trap-Fourier transform ion cyclotron resonance (LIT-FTICR) mass spectrometer (LTQ FT Ultra from ThermoElectron, San Jose, CA). Microcapillary RPLC column (75 μm i.d. × 10 cm fused silica capillary with a flame pulled tip) was in-house slurry-packed with 5 μm, 300 Å pore size, Jupiter C-18 stationary phase (Phenomenex, Torrance, CA). After sample injection, the column was washed for 30 min with 98% mobile phase A (0.1% formic acid in water) at 0.5 μL/min. BSA peptides were eluted using a linear gradient of 2% mobile phase B (0.1% formic acid in ACN) to 42% B in 40 min at 0.25 μL/min and then to 98% B for an additional 10 min. DRMM peptides from SCX fractions were eluted using a linear gradient of 2−60% B in 100 min at 0.25 μL/min and then to 98% B for an additional 10 min. The LIT-FTICR-MS was operated in a data dependent mode in which each full MS^1 scan was followed by seven MS^2 scans, wherein the seven most abundant molecular ions were dynamically selected for collision-induced dissociation using a normalized collision energy of 35%.

MS Data Analysis and Interpretation. Acquired mass spectra were searched against the BSA protein sequence for BSA and human database with appended Gag protein sequence for HeLa cell DRMM using the SEQUEST algorithm implemented in the BioWorks 3.2 application (ThermoElectron, San Jose, CA), using 10 ppm precursor ion tolerance and 0.5 Da for fragment ions. Only fully tryptic peptides with up to two miscleavages were considered as positive identification. Charge state dependent cross correlation (Xcorr) ≥ 2.1 for $[M + H]^{1+}$, ≥ 2.5 for $[M + 2H]^{2+}$, and ≥ 3.2 for $[M + 3H]^{3+}$ were used for BSA peptides, and Xcorr ≥ 1.9 for $[M + H]^{1+}$, ≥ 2.2 for $[M + 2H]^{2+}$, and ≥ 3.1 for $[M + 3H]^{3+}$ were used in DRMM samples.

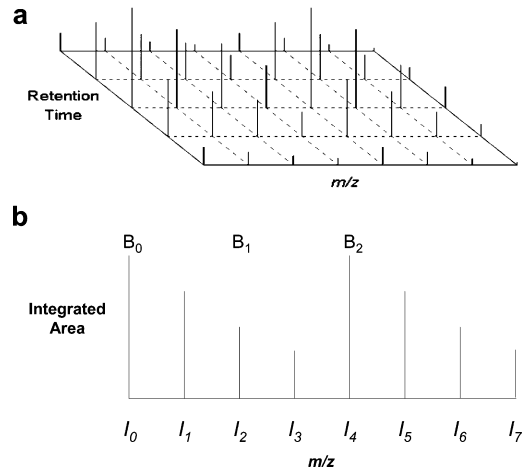

Figure 1. Example of intensities for a given isotopic manifold across multiple MS^1 spectra (a). Example of the integrated areas for peaks within a given isotopic manifold (b).

Calculation of $^{18}O/^{16}O$ Ratios Using the XPRESS Software. The ratios of heavy ^{18}O labeled and light ^{16}O labeled peptides were calculated using the XPRESS algorithm implemented in the BioWorks package (version 3.2). XPRESS reports $^{18}O/^{16}O$ ratios for differentially labeled proteins as an average of ratios of all identified peptides unique to the protein under evaluation, except for C-terminal peptides.

RESULTS

Computation of $^{18}O/^{16}O$ Ratios by Fitting an Isotope Manifold Using a Poisson Distribution along with Evolutionary Programming. The LC−MS profile for each MS experiment was extracted from the respective raw spectral data using a conversion function in BioWorks. Principal extracted variables included identities of differentially labeled peptides, their respective precursor ion masses, charge states, and retention times throughout the LC−MS run/analysis. These parameters provide the location of respective MS^1 spectra for identified peptide ions. Once the MS^1 spectra of these peptides were located, isotopic species for each respective peptide cluster were defined by sequential examination of MS^1 spectra in both directions, obtaining the envelopes for each isotopic species, as shown in Figure 1a. For each direction, the MS^1 spectra were examined and the envelope was stopped for all isotopic species when less than N_i isotopic intensities were observed. In this investigation, the N_i, minimal vector length, was set to 2, meaning that 2 of the 4 isotopic peaks for a given precursor ion had to be present to incorporate these MS^1 intensities in the envelope. The retention time at which the envelope was stopped was then assigned as the start and end point. The intensities at the start and end point were given an arbitrary value of zero to correct background noise.

Following the determination of the start and end points, the peak area was calculated using the trapezoidal rule. For a given peptide, each isotopic species in the isotopic cluster was calculated, resulting in up to eight integrated areas (I_0−I_7), one for each

(26) Blonder, J.; Hale, M. L.; Chan, K. C.; Yu, L. R.; Lucas, D. A.; Conrads, T. P.; Zhou, M.; Popoff, M. R.; Issaq, H. J.; Stiles, B. G.; Veenstra, T. D. *J. Proteome Res.* **2005**, *4*, 523–531.
(27) Blonder, J.; Chan, K. C.; Issaq, H. J.; Veenstra, T. D. *Nat. Protoc.* **2006**, *1*, 2784–2790.

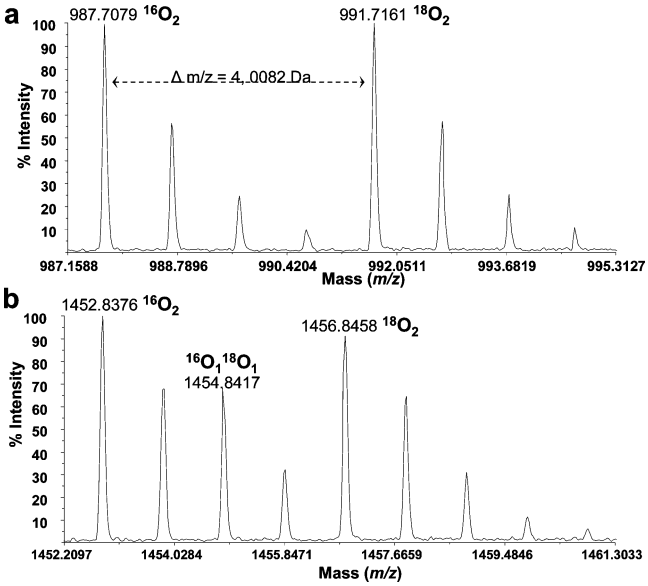

Figure 2. MS[1] of differentially labeled tryptic isotopomeric peptide pair indicating complete ^{18}O exchange/incorporation (a). MS[1] of differentially labeled convoluted peptide clusters indicating variable/incomplete ^{18}O exchange/incorporation exemplified by the presence of the $^{18}O_1$ isotopomer (b).

isotopic species, as shown in Figure 1b. Evidently, the last four integrated areas (I_4–I_7), depicting isotopic envelope of heavy labeled isotopomer, can only originate from the $^{18}O_2$ tagged sample, which is referred to as the B-sample. Figure 2a shows a differentially labeled isotopic cluster exhibiting complete exchange of two ^{16}O atoms for two ^{18}O atoms in the presence of ^{18}O water. Correspondingly, intensities of the first four peaks (I_0–I_3) can originate either from the ^{16}O isotopomer (i.e., incubated with trypsin in the presence of $H_2{}^{16}O$, which is referred to as the A-sample) or in part from the singly $^{18}O_1$ tagged isotopomer present in the B-sample due to incomplete C-terminal labeling (Figure 2b). While earlier investigations fit the series of integrated peak areas using the set of Gaussian functions, we propose that the manifold of isotopic intensities should follow a Poisson distribution as described below:

$$P(k,\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{1}$$

In this expression, λ is the expected number of heavy isotopes present in the peptide and k is the observed number. A given isotopic envelope is possibly perturbed to a small extent by oxygen or sulfur. This situation exists since 0.20% of all oxygen atoms are ^{18}O and 4.29% of all sulfur atoms are ^{34}S, then a corresponding Poisson distribution will cause the intensities to change for $(m + 2)/z$.

In a typical $^{16}O/^{18}O$ labeling experiment, a fraction f of all peptides is ^{18}O tagged. Defining our nomenclature, if the total amount of a given peptide species is B, the total unlabeled amount of this peptide is B_0, the total amount of singly $^{18}O_1$ tagged isotopomer is B_1, and the total amount of doubly $^{18}O_2$ tagged isotopomer is B_2 (Figure 1b). In the same way, on the basis of the reaction kinetics and the assumption that the exchange of the two oxygen atoms is independent of each other, their relationship can be expressed as following:

$$B = B_0 + B_1 + B_2 \tag{2}$$

$$B_0 = B(1 - f)^2 \tag{3}$$

$$B_1 = 2Bf(1 - f) \tag{4}$$

$$B_2 = Bf^2 \tag{5}$$

Assuming that there can only be at most three substitutions that cause a mass shift of one and one substitution that causes a mass shift of two from the naturally occurring isotopes, the calculated intensity of the eight peaks (I_0–I_7) within the isotopic envelope can be given by the following expressions:

$$I_0 = (A + B_0)P(0,\lambda_1)P(0,\lambda_2) \tag{6}$$

$$I_1 = (A + B_0)P(1,\lambda_1)P(0,\lambda_2) \tag{7}$$

$$I_2 = (A + B_0)[P(2,\lambda_1)P(0,\lambda_2) + P(0,\lambda_1)P(1,\lambda_2)] + B_1P(0,\lambda_1)P(0,\lambda_2) \tag{8}$$

$$I_3 = (A + B_0)[P(1,\lambda_1)P(1,\lambda_2) + P(3,\lambda_1)P(0,\lambda_2)] + B_1P(1,\lambda_1)P(0,\lambda_2) \tag{9}$$

$$I_4 = B_1[P(2,\lambda_1)P(0,\lambda_2) + P(0,\lambda_1)P(1,\lambda_2)] + B_2P(0,\lambda_1)P(0,\lambda_2) \tag{10}$$

$$I_5 = B_1[P(1,\lambda_1)P(1,\lambda_2) + P(3,\lambda_1)P(0,\lambda_2)] + B_2P(1,\lambda_1)P(0,\lambda_2) \tag{11}$$

$$I_6 = B_2[P(2,\lambda_1)P(0,\lambda_2) + P(0,\lambda_1)P(1,\lambda_2)] \tag{12}$$

$$I_7 = B_2[P(1,\lambda_1)P(1,\lambda_2) + P(3,\lambda_1)P(0,\lambda_2)] \tag{13}$$

Substituting eqs 3–5 into these equations reduces the parameters down to five: A, B, f, λ_1, and λ_2. Then all calculated areas are divided by I_0 and the first ratio becomes unity, thereby reducing the number of parameters to four: B/A, f, λ_1, and λ_2.

The optimal values of these adjustable parameters used to fit the manifold of integrated area ratios were determined using evolutionary programming. Each parameter was initially set to a random value, while its fitness was determined within a reasonable range. This was done N_{pop} times to create an initial population of putative solutions. In each generation, each putative solution was used to generate a new putative solution by copying the parameter values and then randomly changing them by a small amount. This new putative solution is evaluated and its parameter values and fitness were stored in an offspring population. The parent and offspring populations are then merged and all putative solutions were ordered from lowest to highest fitness. After N_{gen} generations,the process stopped and the final N_{pop} solutions with the highest fitness were examined. All putative solutions with an average error within 10% of the best solution were used to find average values for each of the parameters, which then represented the optimal values for the parameters. In this investigation, the population size, N_{pop}, was set to 2000, the process was run for 4000 generations (N_{gen}

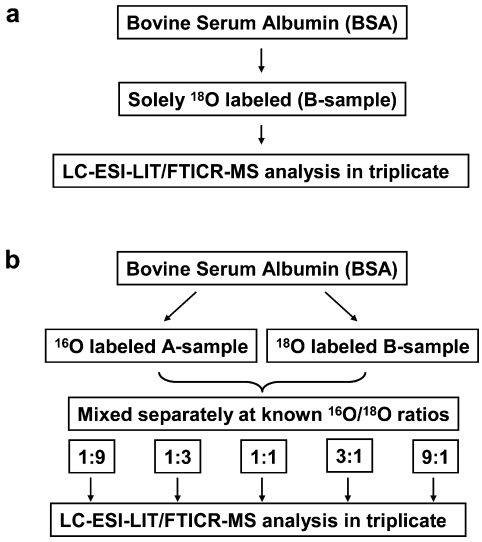

Figure 3. Scheme of experimental design: LC−MS analysis of solely ^{18}O labeled BSA (a). The LC−MS analysis of differentially $^{16}O/^{18}O$ labeled BSA digest mixed at known $^{18}O/^{16}O$ ratios.

= 4000), and the result that yielded the smallest average error was reported.

Assessment of Completeness of ^{18}O Incorporation. To study the completeness of ^{18}O incorporation we used the BSA as a model protein and labeled it in the presence of $H_2{}^{18}O$. To determine how thorough the ^{18}O exchange is and to assess the extent of variable ^{18}O incorporation, we analyzed solely the ^{18}O labeled BSA digest (B-sample) to obtain empirical values of *f* for each identified BSA peptide (Figure 3a). Table 1 summarizes the results obtained by LC−LIT/FTICR analysis of ^{18}O labeled BSA digest.

The peptide sequence, mass, and charge are listed in the first three columns. The fourth column contains the number of calculated *f* values for that peptide. If a given fragment was observed in all three LC−LIT/FTICR runs, then there are a total of three calculated *f* values indicated in the fourth column. The fifth column contains the arithmetic average of the *f* values. The last two rows of Table 1 summarize the average and standard deviation of *f* values for all peptides that end in either lysine or arginine.

Although 50% of the identified peptides did not show a detectable unlabeled B_0 peak, the rest showed the B_0 peak to a variable extent. The average percentage of detectable B_0 peaks in ^{18}O labeled BSA was around 3%, which is expected because 95% pure ^{18}O water was used. In contrast, the average percentage of $^{18}O_1$ monolabeled peptide isotopomers (B_1) was as high as 21%. In one extreme example, the monolabeled $^{18}O_1$ isotopomer of DAFLGSFLYEYSR peptide reached 45%, which is almost as intense as the $^{18}O_2$ dual-labeled isotopomer. We also observed that the labeling efficiency for lysine-ending peptides is different from arginine-ending peptides. On average, the *f* value is significantly higher (*z*-score = 5.778, $P < 4.0 \times 10^{-9}$) if the isotopic substitution occurs at an arginine residue.

Algorithm Validation Using LC−MS Analysis of Differentially $^{18}O/^{16}O$ Labeled BSA Digest. We used BSA as a model protein to assess the utility of the developed algorithm for $^{18}O/^{16}O$-based quantitative proteomics. Figure 3b depicts the experimental approach designed to test the algorithm by analyzing

Table 1. Average *f* Value of BSA Peptides in a Triplicate LC−MS² Analysis of Solely ^{18}O Labeled BSA Digest[a]

peptide	MH^+	*z*	no. of *f*	avg
K.DLGEEHFK[.G	978.4658	2	3	0.76
K.DLGEEHFK[.G	978.4658	1	3	0.78
R.FKDLGEEHFK[.G	1253.629	3	3	0.71
R.FKDLGEEHFK[.G	1253.629	2	3	0.75
K.AEFVEVTK[.L	926.496	2	3	0.71
K.AEFVEVTK[.L	926.496	1	3	0.74
K.HLVDEPQNLIK[.Q	1309.724	2	3	0.66
K.HLVDEPQNLIK[.Q	1309.724	1	3	0.95
K.LVNELTEFAK[.T	1167.639	1	3	0.97
K.KQTALVELLK[.H	1146.722	3	3	0.96
K.YLYEIAR[.R	931.5015	2	3	0.93
K.YLYEIAR[.R	931.5015	1	3	0.96
R.KVPQVSTPTLVEVSR[.S	1643.946	2	3	0.98
R.RHPEYAVSVLLR[.L	1443.82	3	2	0.9
K.VPQVSTPTLVEVSR[.S	1515.851	2	3	0.92
R.RHPEYAVSVLLR[.L	1443.82	2	3	0.91
R.HPEYAVSVLLR[.L	1287.719	3	3	0.97
K.LGEYGFQNALIVR[.Y	1483.803	1	2	0.99
K.DAFLGSFLYEYSRR[.H	1727.852	2	1	0.9
K.DAFLGSFLYEYSR[.R	1571.751	2	3	0.71
R.MPCTEDYLSLILNR[.L	1671.821	2	2	0.9
K.LVTDLTK[.V	793.4797	1	2	0.85
R.KVPQVSTPTLVEVSR[.S	1643.946	3	2	0.94
K.LVNELTEFAK[.T	1167.639	2	1	0.99
K.AEFVEVTKLVTDLTK[.V	1696.95	2	1	0.97
K.AEFVEVTKLVTDLTK[.V	1696.95	3	1	0.93
K.VPQVSTPTLVEVSR[.S	1515.851	3	1	0.93
all K-ending peptides			35	0.82 ± 0.12[b]
all R-ending peptides			31	0.92 ± 0.08[b]

[a] The tagged C-termini are marked by [. [b] Average ± standard deviation was calculated for all K or R ending peptides.

the five differentially $^{18}O/^{16}O$ labeled BSA samples, mixed using *a priori* $^{18}O/^{16}O$ ratios: 9:1, 3:1, 1:1, 1:3, and 1:9. The ^{16}O-labeled BSA sample was conveniently labeled as the A-sample while the ^{18}O-labeled BSA was labeled as the B-sample. After combination of the A-sample and B-sample using predetermined ratios, these samples were analyzed in triplicate by LC−LIT/FTICR. The $^{18}O/^{16}O$ ratios at mixing rates of 3:1, 1:1, and 1:3 were determined using three different analytical approaches for each of the three runs. Table 2 contains summary information and statistics on the computed $^{18}O/^{16}O$ ratios for differentially labeled BSA digests.

The first method denoted as "variable" allows all four parameters (B/A, *f*, λ_1, and λ_2) to vary in value in regard to their roles in determining the ratios of integrated areas in the manifold. The B/A ratio was calculated for each identified BSA peptide. Results in Table 2 list the mean B/A ratio and its standard deviation across these peptides. For the 1:1 mixture of BSA, the calculated B/A ratios were 1.50 ± 1.33, 1.41 ± 0.57, and 1.24 ± 0.58 for runs 1, 2 and 3, respectively, and 1.38 ± 0.88 when averaged across all peptides in all runs. Though these values are reasonably good, the standard deviation is quite large. As seen in Table 2, for the 3:1 and 1:3 mixtures of BSA, this approach yields reasonable B/A ratios, although with relatively large standard deviations.

The second method denoted as "filtered $f > 0.6$" is based on the observation that the experimentally determined *f* value for most peptides was ≥0.60. These results are also reported in Table

Table 2. Computed $^{18}O/^{16}O$ Ratios for BSA Mixed at Predetermined Ratios of 1:1, 3:1, and 1:3 Using Different Adjustments

		average ± SD			
^{18}O:^{16}O	method	run 1	run 2	run 3	all runs
1:1	allow *f* to vary	1.50 ± 1.33	1.41 ± 0.57	1.24 ± 0.58	1.38 ± 0.88
	filtered by $f > 0.6$	1.27 ± 0.48	1.41 ± 0.57	1.18 ± 0.51	1.29 ± 0.53
	predetermined *f*	1.00 ± 0.52	1.18 ± 0.56	1.01 ± 0.47	1.06 ± 0.52
3:1	allow *f* to vary	4.15 ± 2.22	3.58 ± 1.56	3.70 ± 1.67	3.80 ± 1.82
	filtered by $f > 0.6$	4.26 ± 2.31	3.53 ± 1.39	3.72 ± 1.69	3.83 ± 1.83
	predetermined *f*	3.56 ± 5.84	2.68 ± 1.40	3.40 ± 1.96	3.23 ± 3.37
1:3	allow *f* to vary	0.52 ± 0.38	0.45 ± 0.25	0.42 ± 0.13	0.47 ± 0.29
	filtered by $f > 0.6$	0.42 ± 0.12	0.44 ± 0.26	0.43 ± 0.13	0.43 ± 0.17
	predetermined *f*	0.33 ± 0.12	0.34 ± 0.20	0.30 ± 0.15	0.33 ± 0.16

Table 3. Computed $^{18}O/^{16}O$ Ratios in Logarithmic Scale for BSA Mixed at Predetermined Ratios of 9:1, 3:1, 1:1, 1:3, and 1:9 Using Our Algorithm by Allowing *f* to Vary and by XPRESS[a]

	ratio by allowing *f* to vary		ratio by XPRESS	
$\log_2(^{18}O:^{16}O)$	median	range	median	range
2.19	2.45	2.48 (1.40∼3.88)	1.74	3.30 (−0.15∼3.15)
1.09	1.19	2.66 (−0.11∼2.55)	0.64	4.11 (−2.40∼1.71)
0	0.09	1.73 (−0.94∼0.79)	−0.27	3.01 (−2.52∼0.49)
−1.1	−0.92	1.74 (−1.94∼−0.20)	−1.26	3.17 (−3.17∼0.00)
−2.2	−2.26	1.76 (−3.21∼−1.45)	−2.27	3.00 (−3.35∼−0.35)

[a] The smallest and largest ratios are included in the parentheses.

2. Therefore, this method uses the same results obtained by the first method, except it only accepts *f* values greater than 0.6. In general, the average *B*/*A* values showed improvement for the 1:1 and 3:1 mixtures exemplified in an overall decrease in the standard deviation. The results for run 2 in the 1:1 mixture are the same as for the first method, which simply means that all of the calculated *f* values were above 0.60. Notably, the second method did not result in a significant improvement of average *f* values but did improve dispersion measurements as reflected in most standard deviation scores.

The third method denoted as "predetermined *f*" relies on the LC−MS/MS prerun to analyze a portion of ^{18}O labeled sample-B before combining the ^{16}O labeled and ^{18}O labeled sample. This analysis permits the experimental determination of *f* value for each identified peptide, along with an assessment of the extent of variable ^{16}O incorporation. This approach fixes the values of *f* for every peptide identified in the prerun as shown in Table 1. If a peptide is not contained in the prerun list, the corresponding average *f* value from either lysine-ending or arginine-ending cohort shown in Table 1 is used. This choice depends on terminal amino acid (i.e., lysine or arginine). Finally the manifold of the isotopic envelope are then reexamined by using experimentally obtained *f* values and varying only the parameters *B*/*A*, λ_1, and λ_2. As shown in Table 2, the resulting *B*/*A* ratios were more accurate and precise than the ratios when *f* was allowed to vary as exemplified in SD values.

Table 3 shows the ratios obtained using our program, which allowed *f* values to vary versus those obtained by XPRESS for the following mix conditions: 9:1, 3:1, 1:1, 1:3, and 1:9. The ratios calculated using our algorithm is closer to true values as the student *t* test ($P = 0.10$) indicates that there is no significant difference. In contrast, the ratios calculated using XPRESS is significantly lower than true values ($P = 0.01$). In addition, the range of ratios from different peptides obtained using our algorithm is significantly decreased compared to the algorithm of XPRESS ($P < 0.001$).

Algorithm Application to Quantitative Profiling of $^{16}O/^{18}O$ Labeled HeLa DRMMs. To evaluate the algorithm's utility for quantitative profiling of complex protein mixture and investigate HIV Gag trafficking to the plasma membrane, we analyzed differentially $^{16}O/^{18}O$ labeled DRMMs isolated from HeLa cells expressing wild-type Gag or a myristylation-defective Gag mutant, which does not bind DRMMs.[25,28] DRMMs proteins isolated from HeLa cells transfected with a plasmid encoding mutant Gag were digested and labeled in the presence of ^{16}O water, while DRMMs proteins isolated from HeLa cells transfected with a wild-type Gag-expressing plasmid were labeled in the presence of ^{18}O water (Figure 4). Following differential labeling, the digests were mixed at a 1:1 ratio. Differentially labeled digest was then resolved via strong cation exchange (SCX) into 14 fractions. Each fraction was then analyzed by LC−MS. Quantitative analysis using XPRESS algorithm resulted in identification/quantitation of a total of 1740 protein specific peptides (Supplementary Table S-1a in the Supporting Information) that allowed relative quantitation for a total of 558 unique proteins (Supplementary Table S-1b in the Supporting Information) with their ratio distributions shown in Figure 5a. In contrast, our algorithm reported identification/quantitation of a total of 726 protein specific peptides (Supplementary Table S-1b in the Supporting Information) that allowed relative quantitation for a total of 285 unique proteins (Supplementary Table S-1b in the Supporting Information). Higher number of quantified peptides/proteins was reported by XPRESS since this algorithm uses a monoisotopic peak only to calculate the peak area for a given isotopomeric peptide, while our algorithm requires an isotopic manifold for the same calculation. Consequently, a considerable number of identified peptides was not quantified by our algorithm due to missing isotopic peaks in the raw MS data. Figure 5b exhibits a Gaussian distribution around ratio 1, which is expected in large scale quantitative proteome studies. For proteins identified by two or more peptides, 29.1% were quantified by XPRESS and 28.8% were quantified by our algorithm. There was no significant

(28) Ono, A.; Freed, E. O. *J. Virol.* **1999**, *73*, 4136–4144.

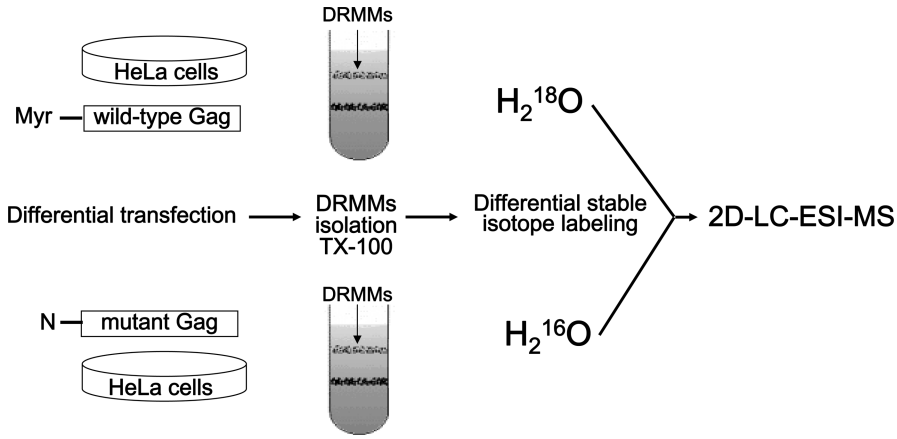

Figure 4. Schematic outline depicting the experimental design for quantitative $^{16}O/^{18}O$-based proteomic analysis of DRMMs transfected with wild-type Gag and a G1A mutant Gag.

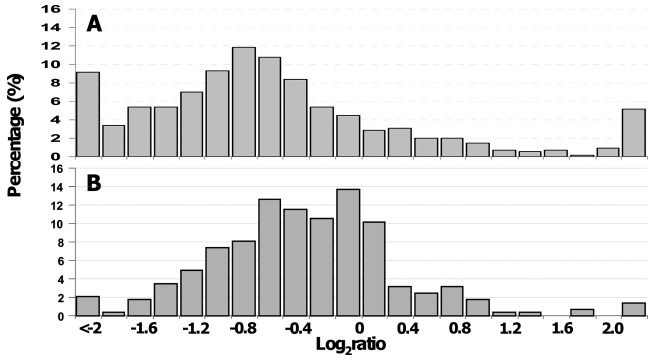

Figure 5. The distribution of the percentage of unique proteins quantified by ^{18}O labeling in the DRMMS of HeLa cells transfected with expression plasmids encoding the wild type Gag and 1GA mutant Gag within binned logarithms of protein abundance ratios obtained using XPRESS (A) and the developed algorithm (B).

difference in the coefficient of variation (CV) distribution between the two algorithms since the Pearson product-moment correlation coefficient was 0.999. Additionally, the median CV from XPRESS and our method were essentially the same (50% and 49%, respectively). Importantly, more than 84% of peptides showed an increased ratio using our algorithm indicating the effectiveness of the present algorithm to detect and account for incomplete labeling. Additionally, our approach is not influenced by cysteine or methionine residues, since there is no significant difference between peptides containing cysteine, methionine, or neither ($P > 0.99$).

In this study, 11 Gag peptides were quantified using our algorithm while 22 were quantified using XPRESS. Table 4 lists the peptides quantified by both algorithms. Of note, all of these peptides have higher ratios quantified by our algorithm versus XPRESS. For f values greater than 0.7, the increase is approximately 150%. In contrast, for f value less than 0.5, the average increase is around 300%, suggesting the necessity of detecting and accounting for incomplete ^{18}O incorporation. Indeed, manual inspection of the MS scans of these peptides showed incomplete labeling at various levels. Among 11 Gag peptides only quantified by XPRESS, 5 are repetitive identification and quantitation of the same peaks due to high peak intensity. The remaining peptides were not quantified using our algorithm due to missing isotopic peaks. The average $^{18}O/^{16}O$ ratio of

Table 4. Computed $^{18}O/^{16}O$ Ratios for Gag Protein in Isolated DRMMs

		ratio			
peptide	SCX fraction	our program	XPRESS	increase (%)	f
R.WIILGLNK[.I	3	8.76	4.19	209	0.59
R.FGEETTTPSQK[.Q	3	18.58	12.48	148	0.76
R.FGEETTTPSQK.Q	3	7.50	5.77	130	0.73
K.ELYPLASLR[.S	3	5.43	1.20	452	0.42
R.FGEETTTPSQK[.Q	4	20.55	12.7	161	0.77
R.FGEETTTPSQK[.Q	4	8.78	5.59	157	0.74
K.ELYPLASLR[.S	5	6.96	2.00	347	0.47
R.QILGQLQPSLQTGSEELR.S	5	4.06	2.18	186	0.31
K.ETINEEAAEWDR[.L	5	13.46	3.48	386	0.34
R.FGEETTTPSQK[.Q	9	27.48	19.06	144	0.75
R.FGEETTTPSQK[.Q	9	8.38	4.92	170	0.73

Gag is 11.81 ± 7.37 using our algorithm while 8.30 ± 8.62 was reported by XPRESS. The quantitative results for Gag obtained by LC−MS analysis in this study are highly consistent with a well-accepted notion that N-terminal myristylation is essential for Gag-DRMMs association.[29]

DISCUSSION

During a typical $^{18}O/^{16}O$ labeling experiment, trypsin binds covalently to C-terminal lysyl or arginyl-residues catalyzing a stepwise substitution of two C-terminal ^{16}O atoms for two ^{18}O atoms. The first hydrolytic reaction $RC^{16}ONHR' + H_2{}^{18}O \rightarrow RC^{16}O^{18}O^- + {}^+H_3NR'$ is followed by a second hydrolytic reaction $RC^{16}O^{18}O^- + H_2{}^{18}O \rightarrow RC^{18}O^{18}O^- + H_2{}^{16}O$. Two complete enzyme-catalyzed turnovers in the presence of the heavy isotopic form of water [$H_2{}^{18}O$] result in a 4 Da mass increase in labeled peptides as shown in Figure 6a. This same difference is present in MS^2 spectra of heavy ^{18}O-labeled peptides. Figure 6a,b shows a mass shift of 4 Da between singly charged y-fragments of light $^{16}O_2$ isotopomer versus heavy $^{18}O_2$ labeled isotopomer, thus confirming the presence of a C-terminal $^{18}O_2$ tag.

However, two complete enzyme turnovers are not always achieved. This results in an incomplete ^{18}O exchange exemplified in co-occurrence of variable amount $^{18}O_1$ isotopomeric species, as illustrated in Figure 2b. This was also confirmed by analyzing the solely ^{18}O labeled BSA specimen sample that was used to

(29) Lindwasser, O. W.; Resh, M. D. *J. Virol.* **2001**, *75*, 7913–7924.

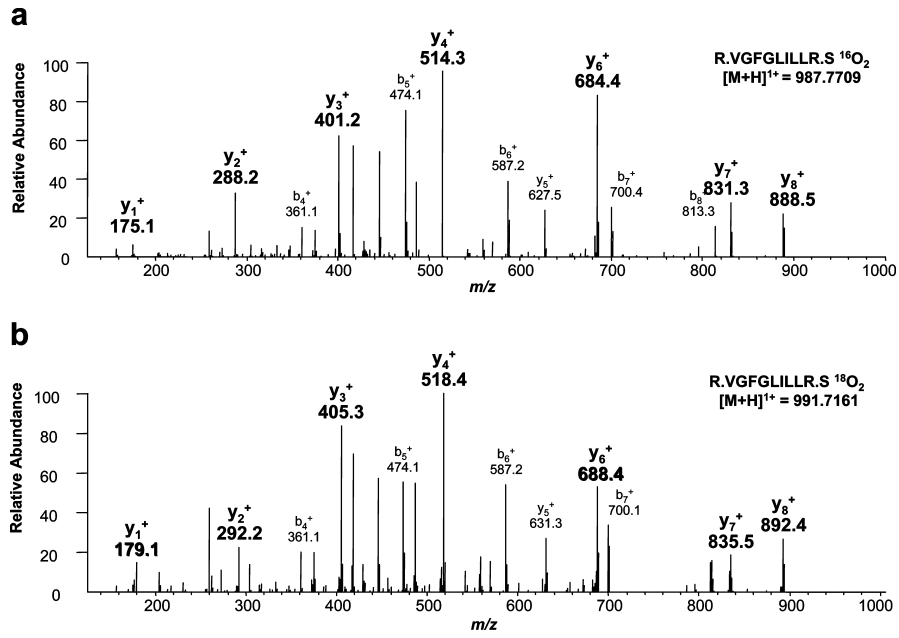

Figure 6. MS^2 spectra of differentially $^{16}O/^{18}O$ labeled peptide pair shown in Figure 2a. MS^2 spectrum showing CID fragments of ^{16}O labeled isotopomer (a). MS^2 spectrum of ^{18}O labeled isotopomer showing that all detected y ion fragments of this heavy labeled isotopomeric peptide exhibit a 4 Da difference when compared to corresponding y ion fragments of ^{16}O labeled peptide (b). This finding confirms the presence of the $^{18}O_2$ tag at the C-terminal.

assess the completeness of ^{18}O exchange/labeling. The average percentage of $^{18}O_1$ monolabeled peptide isotopomers was found to be as high as 21%. In one extreme example, the monolabeled $^{18}O_1$ isotopomer of DAFLGSFLYEYSR peptide reached 45%, which is almost as intense as the $^{18}O_2$ dual-labeled isotopomer. Thus, labeling efficiency can vary and be dependent on the peptide sequence. As shown, the arginine-ending peptides appear to have a higher level of incomplete labeling than lysine-ending peptides. This is consistent with the previous observation and is a consequence of the different substrate affinity of trypsin for arginine and lysine ending peptides.[30]

To account and correct for the effect of inhomogeneous ^{18}O incorporation, we developed an algorithmic method capable of determining the efficiency of ^{18}O exchange. Specifically, we detect and account for partial ^{18}O incorporation. Our approach involves curve fitting the integrated areas of the multifold peaks enclosed by the isotopic envelope and correcting for the incomplete labeling. The $^{18}O/^{16}O$ ratios are formed using a computational approach that employs a Poisson distribution along with a near optimization method that utilizes an evolutionary programming technique.

Compared to current computational methods for $^{18}O/^{16}O$ quantitative labeling, our approach offers advantages in several respects. First, the time window for peak integration is tailored for each identified peak instead of integrating it over a fixed region of the extracted ion chromatogram. This significantly improves the accuracy of isotopic peak intensities and in particular for candidate peptide ions of low intensities. Second, we utilize an overlapping Poisson distribution model to fit isotopic manifolds, which better suits the natural physical properties of the isotope distribution. Several earlier investigations have employed a set of Gaussian functions to fit the isotopic envelope of the identified peptides.[16,31] However, since the numbers of distinct isotope species for the manifold peaks are integers, we propose that the manifold of isotopic intensities should follow a Poisson distribution. This is also supported by our observation that the Gaussian distribution did not completely fit the integrated areas. Third, evolutionary programming is employed to obtain optimized parameters for the fitting with several quality checkpoints to improve accuracy and reproducibility. The program requires that at least two of the four isotopic peaks for a given precursor ion must be present in the manifold envelope to be considered for quantitation. In addition, since λ_1 is the expected number of ^{13}C and λ_2 is the expected number of $^{18}O/^{34}S$ present in a given peptide, these values are calculated for every identified peptide and serve as checkpoints for the quality of fitting.

It should be noted that when the BioWorks XPRESS algorithm is applied for the $^{18}O/^{16}O$ ratio calculation, it assumes the complete ^{18}O exchange for each differentially labeled peptide−ion pair. XPRESS calculates the $^{18}O/^{16}O$ ratio directly from the monoisotopic peaks located at a 4 Da difference from the monoisotopic peak of the identified peptide ion.[32] Therefore, these types of algorithms are not capable of accounting for variable/incomplete ^{18}O incorporation.

As depicted in Table 1, a significant degree of variable/incomplete ^{18}O incorporation is observed for certain tryptic peptides. If the presence of these $^{18}O_1$ monolabeled isotopomers is not taken into consideration for the $^{18}O/^{16}O$ ratios calculation, a variable degree of underestimation of the peptide/protein ratio is to be expected. Evidently, the average ($^{18}O/^{16}O$) ratio

(30) Yao, X.; Afonso, C.; Fenselau, C. *J. Proteome Res.* **2003**, *2*, 147–152.

(31) Jorge, I.; Navarro, P.; Martinez-Acedo, P.; Nunez, E.; Serrano, H.; Alfranca, A.; Redondo, J. M.; Vazquez, J. *Mol. Cell. Proteomics* **2009**, *8*, 1130–1149.

(32) Moulder, R.; Filen, J. J.; Salmi, J.; Katajamaa, M.; Nevalainen, O. S.; Oresic, M.; Aittokallio, T.; Lahesmaa, R.; Nyman, T. A. *Proteomics* **2005**, *5*, 2748–2760.

obtained by XPRESS is smaller than the real values (Table 3). These results are expected since XPRESS software simply divides the peak area exclusively from $^{18}O_2$ tagged peptide isotopomers over that of $^{16}O_2$ naturally tagged peptide species, overlooking the contribution of the monolabeled peptide species. The relative standard deviation for calculated ratios using our developed algorithm was 13.0 ± 9.6% compared to −23.5 ± 13.5% using XPRESS. A significant increase in accuracy and reproducibility is obtained with our method. Of note, regarding proteins with heavy to light ratios less than the difference between our program and XPRESS was not significant. This is understandable since the interference from a relatively small amount of incomplete labeling to a relatively large third isotopic peak is minimal.

The best accuracy by our method was achieved when the *f* value was determined using a LC−MS/MS prerun to analyze a portion of the ^{18}O labeled sample. However, this practice doubles instrument time and sample consumption, which makes it impractical in a large scale multiple dimensional separation or with very small samples. For this reason, the relationship between the peptide sequence and the efficacy of the ^{18}O exchange reaction is currently under investigation. If *f* values can be predicted by the peptide sequences, then the accuracy and the speed of quantitative proteomics by ^{18}O labeling will be much improved.

To further evaluate the robustness of our approach, it was applied to a more complex biological sample. This mixture contained $^{16}O/^{18}O$ labeled DRMMs (i.e., lipid rafts) isolated from HeLa cells differentially transfected with plasmids expressing mutant and wild-type HIV Gag protein, respectively. The process of the retrovirus assembly requires that several viral constituents must localize to the plasma membrane of an infected cell and then assemble into a budding virus particle. Central to this process is the retroviral Gag polyprotein.[33]

To study the Gag trafficking mechanism, a mutant Gag protein with an amino-terminal myristylation site, which has been disrupted by amino acid substitution, was expressed in HeLa cells. The effects of this mutation on protein translocation to the plasma membrane were investigated. Indeed, the abundance of mutant Gag protein in DRMMs was found to be significantly lower compared to that of wild type protein. This finding is in agreement with the role of the previously proposed myristyl switch for the regulation of the Gag membrane binding and the subsequent Western blotting experiment against Gag protein.[28]

Several proteins quantified by our algorithm are likely associated with Gag translocation, e.g., human CD59 protein is up-regulated by 2.0-fold, flotillin-1 is down-regulated by 0.4 times, while flotillin-2 is up-regulated by 2.8 times. It has been reported that CD59 prevents assembly of the membrane attack complex in HIV transfection.[34,35] Flotillins are typical lipid raft associated membrane proteins, and it has been found that the overall lipid composition of native HIV membranes resembles DRMMs.[36] The changes in abundance of cellular proteins associated with DRMM upon the introduction of wild-type versus mutant Gag proteins are likely suggesting they are closely associated with Gag translocation and therefore the virus assembly, which is currently under investigation.

In summary, this work describes an algorithm amenable to automated interpretation and computation of $^{18}O/^{16}O$ ratios from spectra obtained by LC−MS/MS analysis of differentially labeled isotopomeric clusters. The algorithm is capable to accurately calculate $^{18}O/^{16}O$ ratios and eliminate artifacts caused by variable ^{18}O exchange. It uses centroid peak data and is suitable for any high-resolution MS platform. The algorithm has been successfully tested with the quantitation of a BSA protein standard and then applied to a comparative proteomic profiling of a more complex mixture containing $^{16}O/^{18}O$ labeled DRMM proteins isolated from HeLa cells, differentially transfected with plasmids expressing HIV Gag protein and its myristylation-defective N-terminal mutant. The relative abundance of mutant Gag protein in DRMMs was found to be significantly lower compared to that of wild type Gag, which is highly consistent with a well-accepted notion that N-terminal myristylation is crucial for Gag-DRMMs association.[29]

ACKNOWLEDGMENT

This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contracts HHSN261200800001E and NO1-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the United States Government.

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 May 16, 2010. Accepted May 20, 2010.

AC101284C

(33) Freed, E. O. *Virology* **1998**, *251*, 1–15.

(34) Takefman, D. M.; Spear, G. T.; Saifuddin, M.; Wilson, C. A. *J. Virol.* **2002**, *76*, 1999–2002.

(35) Hu, W.; Yu, Q.; Hu, N.; Byrd, D.; Amet, T.; Shikuma, C.; Shiramizu, B.; Halperin, J. A.; Qin, X. *J. Immunol.* **2009**, *184*, 359–368.

(36) Brugger, B.; Glass, B.; Haberkant, P.; Leibrecht, I.; Wieland, F. T.; Krausslich, H. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 2641–2646.