RAPID COMMUNICATIONS IN MASS SPECTROMETRY
Rapid Commun. Mass Spectrom. 2002; **16**: 2115–2123
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/rcm.831

Amino acid residue specific stable isotope labeling for quantitative proteomics‡

Haining Zhu[1†], Songqin Pan[1], Sheng Gu[1], E. Morton Bradbury[1,2] and Xian Chen[1*]

[1]BN-2, Bioscience Division, MS M888, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[2]Department of Biological Chemistry, School of Medicine, University of California at Davis, Davis, CA 95616, USA

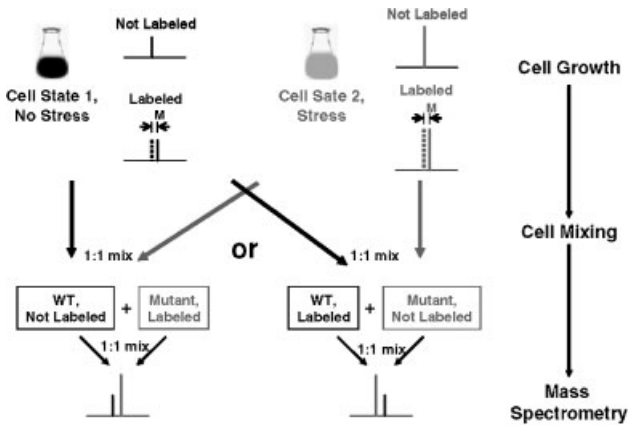
Received 22 July 2002; Revised 11 September 2002; Accepted 12 September 2002

Various stable isotope labeling (SIL) techniques have recently emerged to improve the efficiency and accuracy of protein quantitation by mass spectrometry (MS). We have developed a mass-tagging strategy to incorporate stable isotope tagged amino acids into cellular proteins in a residue-specific manner during cell growth. In this study, we further extend this residue-specific SIL approach to the accurate quantitation of protein abundances in different cell populations. For proteins whose expression levels are the same in cells grown in the normal and labeled media, the relative areas of the normal (light) and labeled (heavy) isotopic peaks are linearly correlated with the cells mixing ratios. This approach was first used to determine the effect of the zinc-responsive transcription factor Zap1 on the yeast proteome. Ten protein spots from a PAGE gel were chosen randomly and their differential protein expression levels in wild-type and *zap1*Δ cells were readily determined by the isotopic ratio. Methionine synthase (Met6) was identified to be up-regulated more than four times in the *zap1*Δ mutant strain whereas the expression level of other nine proteins remained unchanged. Further, we applied this strategy to study the cellular response to radiation in human skin fibroblast cells. Analyzing one protein band randomly selected from SDS-PAGE, the expression level of a novel protein was found to increase two-fold in response to radiation whereas the expression level of a control protein remained unchanged. This strategy is generally applicable using any particular type of amino acid as the labeling precursors for accurate quantitation of protein relative abundances. Published in 2002 by John Wiley & Sons, Ltd.

The tremendous interest in large-scale proteome research has created the need to improve methods for the rapid identification and quantitation of proteins. Currently, mass spectrometry (MS) is the method of choice because of its femtomole sensitivity, high-quality MS and MS/MS data, rapid data collection ability, and high-throughput feature. It has been demonstrated that MS is effective for protein identification.[1,2] In contrast, the MS-based quantitation of protein expression remains a challenging task. Traditionally, proteins are resolved by two-dimensional (2-D) gel electrophoresis and their expression levels are determined from the intensities of stained protein spots. Recently, several techniques coupling stable isotope labeling (SIL) to MS have been introduced for protein quantitation.[3–8] Mass spectrometry can accurately measure the isotopic mass of each isotope peaks, resolve the mass difference between unlabeled and labeled peptides, then quantify their relative abundances for the peptides with the same sequences. These SIL approaches are based on the generation of two pools of proteins from the cells grown under different conditions — one population cultured in normal media, the other in the labeled media containing heavier isotopes such as ^{2}D, ^{13}C, ^{15}N or ^{18}O. These two pools of proteins undergo the same sample handling procedure until the MS measurement. Equal quantities of these two protein pools mixed at the initial steps during sample preparation will minimize differential effects of sample handling, thus allowing an accurate quantitation of the relative protein expression levels between two cell populations.

There are two different strategies to quantify proteins through SIL, either *in vitro* after cell lysis[3–8] or *in vivo* during cell growth.[9–13] *In vitro* SIL approaches usually target cysteine residues,[3] N-terminal[5,14] or C-terminal[6–8,15–17] residues of tryptic peptides. An elegant example is the design of an isotope-coded affinity tag (ICAT) that contains a thiol-reactive group that reacts with cysteine residues, a polyether linker that can be labeled with deuterium, and an affinity tag, e.g. biotin, that allows the rapid purification of the cysteine-containing peptides.[3,4] This strategy has the advantage of being able to analyze the cysteine-containing peptides selectively, thus reducing the complexity of the samples. ICAT has been employed successfully for protein quantitation in some instances.[3,18,19] However, the ICAT approach cannot analyze peptides lacking a cysteine residue. Other approaches

*Correspondence to: X. Chen, BN-2, Bioscience Division, MS M888, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.
E-mail: chen_xian@lanl.gov
†Current address: Department of Molecular and Cellular Biochemistry, College of Medicine, University of Kentucky, Lexington, KY 40536, USA.



Scheme 1. Scheme for the strategy of stable isotope assisted mass spectrometry for accurate quantitation of protein expression levels in two different cell populations. Stress can be radiation, heat shock, mutation of the same cell and others in different studies.

have been developed to label either the N-termini[5,14] or C-termini[6–8,15–17] and can be used for all proteolytic peptides. Acetyl *N*-hydroxysuccinimide or its d_3-acetyl analogue has been used to label the N-terminal residue by an acetylation reaction.[5] At the C-terminal site, ^{18}O can be incorporated into the carboxyl group by carrying out a proteolytic digestion in ^{18}O-enriched water.[6,15–17,20] Alternatively, methanol or methanol-d_3 can be used to label the C-terminal residue through esterification of the carboxyl group.[7] In general, these approaches introduce various isotopic tags through chemical reactions with proteins/peptides *in vitro* after cell lysis. A common disadvantage is the relatively low efficiency and specificity of these chemical modification reactions.

Stable isotope tags can also be incorporated into cellular proteins *in vivo* during cell growth. For example, Oda *et al.* used ^{15}N uniformly labeled medium to label all nitrogen atoms in the whole proteome and applied this strategy to quantitate protein expression and modification.[9] However, this ^{15}N uniform labeling strategy lacks specificity in assigning the labeled peptides because of the variable number of nitrogen atoms in peptides. Thus it becomes difficult to distinguish between a labeled peptide and another unrelated peptide with similar mass. Also, because of the variable number of nitrogen atoms in different peptides, this strategy essentially requires a theoretical isotope abundance calculation for each individual peptide, thus complicating the quantitative measurements. Our laboratory has been developing a different *in vivo* labeling strategy using stable isotope labeled amino acid precursors.[11–13] In our approach, the heavy isotope labeled amino acid precursors can be incorporated into proteins in a residue-specific manner. This results in a characteristic mass-split pattern that can be readily recognized for those peptides containing the labeled amino acid residue. Thus, the content of particular amino acid residue(s) in the peptide can be readily determined from the mass spectra. This partial amino acid composition information is a valuable addition to the molecular mass for the rapid identification of proteins[11,12] and their modifications[13] in a mixture. In this study, the residue-specific SIL strategy was extended for the accurate quantitation of protein expression and modification levels, as illustrated by Scheme 1. The advantages of this method include: (1) it simplifies the interpretation of mass spectra compared with the uniform labeling; (2) it simplifies the calculation of the ratio of the labeled to unlabeled peptides; (3) it can be used to label different amino acid residues, which allows greater flexibility and increases peptide sequence coverage; (4) it minimizes any differential artifacts between the normal and labeled protein pools because the two different cell populations are counted and mixed before the cell lysis; and (5) it does not require any additional chemical reactions because isotopic tags have been incorporated in the proteome during cell growth.

We applied this approach to quantitate the relative abundances of proteins in yeast and human skin fibroblast cells. In both cases, we can easily recognize differentially expressed proteins and simultaneously accomplish protein identification and accurate protein quantitation.

EXPERIMENTAL

Chemicals

The deuterium-enriched amino acid precursors, L-leucine-5,5,5-d_3-98% (Leu-d_3), L-serine-2,3,3-98%-d_3 (Ser-d_3) and L-tyrosine-3,3-98%-d_2 (Tyr-d_2), were purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA, USA). Unlabeled amino acids, protease inhibitors and α-cyano-4-hydroxycinnamic acid were obtained from Sigma (St. Louis, MO, USA). Sequencing-grade modified trypsin was purchased from Roche Diagnostics (Indianapolis, IN, USA).

Residue-specific labeling of yeast cells

Saccharomyces cerevisiae wild-type strain DY 1457 (*MAT*α *ade6 can1 his3 leu2 trp1 ura3*) and *zap1*Δ strain were generously provided by Dr. David Eide (University of Missouri). The wild-type strain was grown in a synthetic defined (SD) medium. The SD medium consists of 20 g dextrose, 1.7 g yeast nitrogen base without amino acids, 5.0 g ammonium sulfate, and 0.1 g each of adenine, histidine, tryptophan and leucine in 1 L of H_2O. Yeast cells were inoculated into 10 mL SD medium and grown overnight at 30 °C. The overnight yeast culture was diluted to a starting optical density (OD) of 0.02 in 100 mL SD medium containing either naturally occurring Leu or the labeled Leu-d_3. No growth difference was observed between cells in regular and labeled media, i.e. isotope-labeled amino acids did not affect cell growth as a source of nutrition. Cells were harvested during log-phase (OD approximately 1.0). The ODs of cell cultures grown in the normal and the labeled medium were measured. Cells from the normal and the labeled media were mixed at different ratios (1:5, 1:2, 1:1, 2:1 and 5:1) as indicated. The cell pellet was washed twice with Millipure H_2O to remove excess medium and stored at −70 °C until needed.

The cell pellet was suspended in 10 mM Tris HCl (pH 8), 10 mM dithiothreitol, 10 mM ETDA, 100 mM NaCl, 0.1% sodium dodecyl sulfate (SDS) and 5 mg/mL protease inhibitors including leupeptin, pepstatin A and chemostatin A (Sigma, MI, USA). Cells were lysed by 10 cycles of 1-min vortexing with glass beads followed by 1-min incubation on

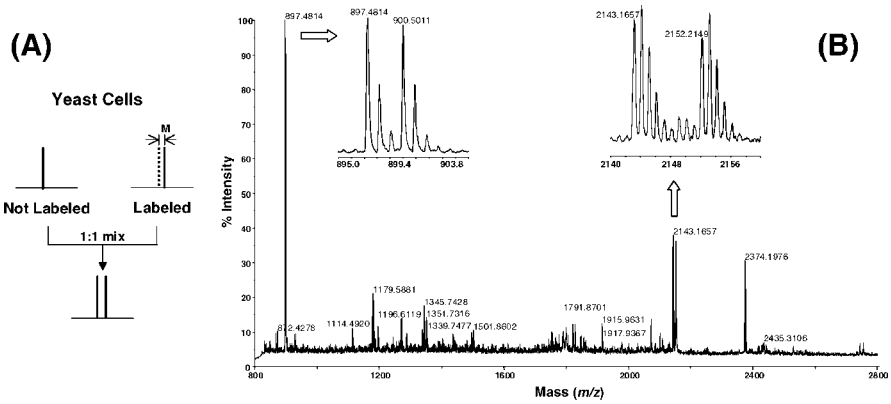


Figure 1. (A) Theoretical spectrum of the light and heavy isotopic peaks when the equal number of cells grown in normal and Leu-d_3 labeled media are mixed. (B) Peptide mass map of the protein band with an apparent MW of 25 kDa. The sample is the lysate of mixture of equal number of cells grown in normal and labeled media. The two insets are the expanded spectra of peptides with *m/z* values of 897.4814 and 2143.1657 showing characteristic 3- and 9-Da splitting, respectively.

ice. The sample was centrifuged at 14 000 rpm for 30 min at 4 °C, the supernatant was collected and subjected to 15% SDS polyacrylamide gel electrophoresis (SDS-PAGE). Protein bands were visualized by Coomassie blue (G-250) staining. Different protein bands with apparent molecular weight 15–60 kDa were sliced, destained with 50% (v/v) acetonitrile 50 mM NH_4HCO_3, then subjected to overnight in-gel trypsin digestion as previously described.[13] Briefly, 15 μL sequencing-grade modified trypsin (10 μg/mL in 50 mM NH_4HCO_3) was added to the lypholized gel, 20 μL 50 mM NH_4HCO_3 were added to cover the rehydrated gel and the reaction was incubated overnight at 37 °C. The gel slices were extracted with 5% acetic acid and 50% acetonitrile/5% acetic acid for 20 min, respectively. The supernatant was collected, lyophilized and then resuspended in a 0.1% TFA solution. The gel extract solution was desalted using C_{18} Zip Tips (Millipore Corp.) and final elution solution was 50% acetonitrile/0.1% TFA.

Wild-type and *zap1*Δ strains under zinc-depleted condition

Yeast strain DY 1457 (wild-type) and *zap1*Δ were grown in chelated synthetic defined (CSD) medium as previously described.[21] The zinc concentration in the medium was maintained at less than 50 nM. The *zap1*Δ strain was grown in normal CSD while the wild-type strain was grown in Leu-d_3 labeled CSD medium. Cells were then harvested as described above. The following two subsets of cells were mixed: (i) equal numbers of wild-type cells grown in unlabeled and labeled media, and (ii) equal numbers of wild-type cells grown in the labeled medium and *zap1*Δ cells grown in the unlabeled medium. The cell pellet was washed twice with Millipure H_2O to remove excess medium and stored at −70 °C until needed.

Cell pellets were lysed as described above and the cell lysates were subjected to 2-D gel electrophoresis: isoelectrical focusing using 17 cm pI 4–7 ReadyStrip IPG strip (BioRad Laboratories) followed by 15% SDS-PAGE. Protein spots on a 2-D gel were visualized by silver staining. 10 protein spots were selected at random, destained with potassium ferricyanide and sodium thiosulfate. Proteins were digested by trypsin; the tryptic peptides were extracted from gel and desalted as described above.

Residue-specific labeling of HSF cells and radiation treatment

The human skin fibroblast (HSF) cells were grown in the α-MEM medium at 37 °C with 5% CO_2 and 90% relative humidity. The medium was supplemented with 10% dialyzed fetal bovine serum (FBS), 100 units/mL penicillin, and 100 μg/mL streptomycin sulfate. For the amino acid residue specific labeling of HSF cells, the selected isotope-enriched amino acid precursor (Leu-d_3, Ser-d_3 or Tyr-d_2) was used in the medium instead of its unlabeled counterpart. At about 80% confluence, the cells grown in the labeled media were irradiated with 500 Rads of γ-ray and then recovered for 4 h under normal conditions. Cells grown in the normal medium were grown without radiation treatment. Cells were harvested with the trypsin-EDTA treatment, and washed three times with a large volume of phosphate-buffered saline (PBS) buffer to remove any trypsin and FBS residuals. Four sets of cells were obtained: (i) cells grown in the unlabeled medium without radiation treatment; (ii) a mixture of equal numbers of cells grown in the Leu-d_3 labeled (with radiation) and normal (without radiation) media; (iii) a mixture of equal numbers of cells grown in the Ser-d_3 labeled (with radiation) and unlabeled (without radiation) media; and (iv) a mixture of equal numbers of cells grown in the Tyr-d_2 labeled (with radiation) and unlabeled (without radiation) media. Cell pellets were stored at −70 °C until needed.

Cells were lysed with the SDS loading buffer in boiling water for 10 min. After centrifugation at 14 000 rpm for 15 min, the supernatant was collected and subjected to 12% SDS-PAGE. The protein bands were visualized by Coomassie blue staining. Five protein bands were sliced from the gel, proteins were digested by trypsin, and the tryptic peptides were extracted from the gel and desalted as described above.

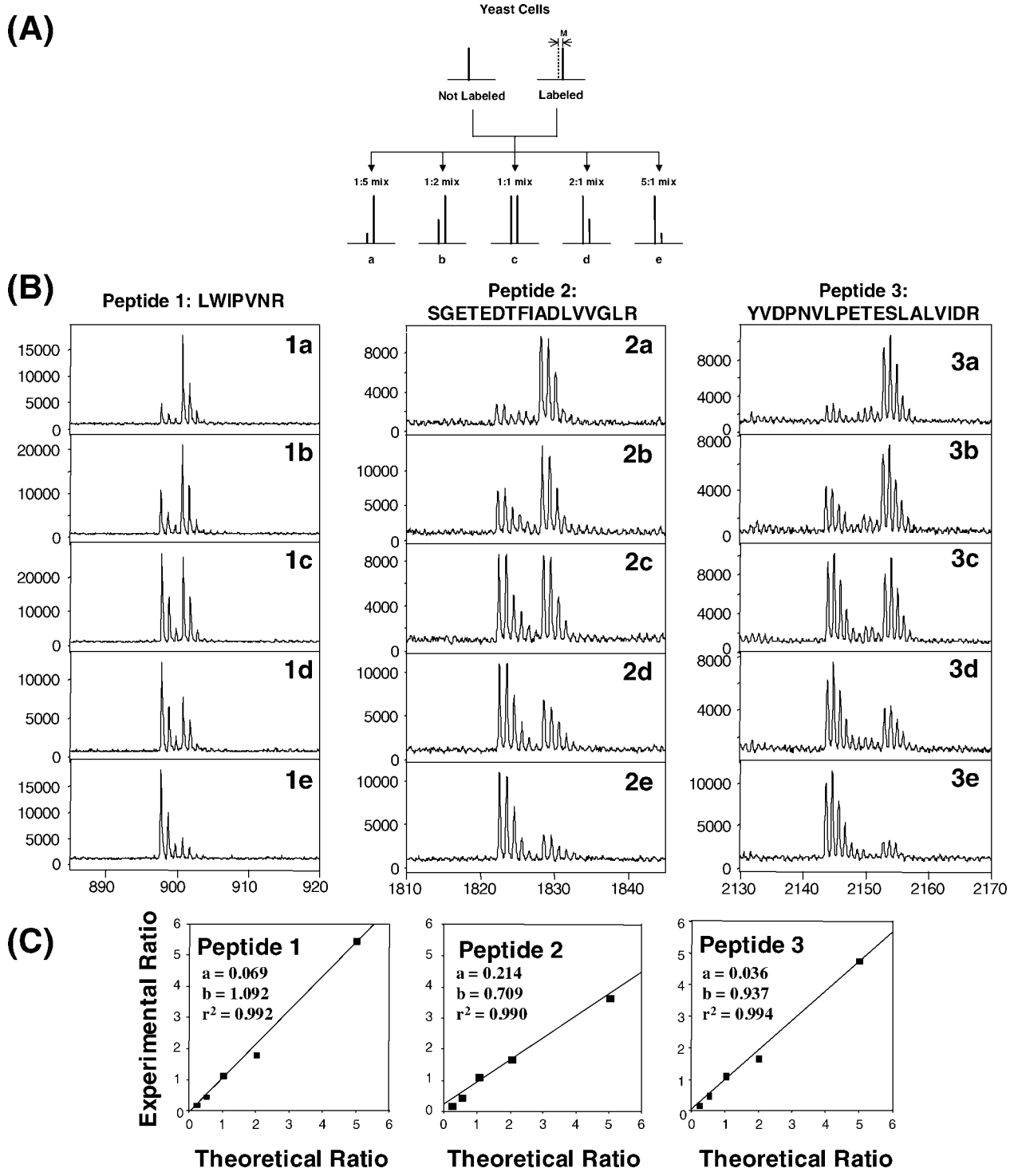


Figure 2. (A) Theoretical spectra expected from the samples with different mixing ratios of cells from normal and Leu-d_3 labeled media. (B) Expanded spectra of peptides 1, 2 and 3. The sequence of each peptide is shown in the top panels. Samples a–e are mixtures of cells grown in normal and leu-d_3 labeled media with ratios of 1:5, 1:2, 1:1, 2:1 and 5:1, respectively. (C) Plot of the experimental ratio values against the theoretical ratio values in peptides 1, 2 and 3. Linear regression ($y = a + bx$) parameters a, b and regression linear coefficiency r^2 for each set of data are listed.

Mass spectrometric measurements

Desalted tryptic peptide sample (0.5 μL) was spotted onto the MALDI target plate. MALDI matrix solution (1 μL, 10 mg/mL α-cyano-4-hydroxycinnamic acid in 50% (v/v) acetonitrile containing 0.1% TFA) was added and the mixture was allowed to air dry. All MS experiments were carried out on a PE Voyager DE-STR MALDI-TOF Biospectrometry workstation equipped with a N_2 laser (337 nm, 3-ns pulse width, 20-Hz repetition rate) in both linear and reflectron mode. The mass spectra of proteolytic digests were acquired in the reflectron mode with delayed extraction (DE). The *m/z* values of proteolytic peptides were calibrated externally with Calmix 1 (PE Biosystems).

Table 1. Ratio of the areas of the light and heavy isotopic peaks in peptides 1–3

Mixing ratio of cells from normal and labeled media	Peak area ratio		
	897.48/900.50	1821.76/1827.80	2143.16/2152.21
Sample a (1:5)	0.21	0.21	0.21
Sample b (1:2)	0.49	0.48	0.52
Sample c (1:1)	1.16	1.13	1.13
Sample d (2:1)	1.80	1.71	1.68
Sample e (5:1)	5.48	3.70	4.78

Liquid chromatography/electrospray ionization mass spectrometry (LC/ESI-MS)

A micro-flow HPLC system (Beckman Coulter) configured with a 1.0×100 mm Aquasil C18 column (Keystone, Bellefonte, PA, USA) was used in LC/MS experiments coupled with an LCQ-DECA ion trap mass spectrometer (Thermo Finnigan, San Jose, CA, USA). The LC flow rate was set at 50 μL/min and all the eluate went into the mass spectrometer without splitting the flow. The ion trap mass spectrometer was operated at Ispray voltage 3.5 kV, and the other parameters were tuned using 500 fmol/μL angiotensin I solution to ensure optimal working conditions. Mobile phase A was 0.1% TFA in water and mobile phase B was 0.1% TFA in acetonitrile. The gradient was kept at 2% B for 2 min, then increased linearly from 2% to 65% B in 20 min, 65% to 85% B in 3 min and kept at 85% B for 3 min, then followed by 85% to 2% B in 3 min.

RESULTS

Ratio of isotopic peaks corresponding to the equally expressed proteins linearly correlates with cell mixing ratio

As shown in Scheme 1, we designed a simple experiment to prove that our strategy is effective. If the two cell populations in Scheme 1 are identical but grown in normal and labeled media, respectively, the mass spectrum of the equal mixture of these cells is expected to show the same intensities of the normal and labeled isotopic peaks as illustrated in Fig. 1(A).

Equal numbers of wild-type yeast cells grown in normal and Leu-d_3 labeled media were mixed and the cell lysate was loaded onto SDS-PAGE followed by Coomassie blue staining. The peptide mass maps (PMMs) of five protein bands in the range 15–60 kDa were obtained using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOFMS). The representative PMM of a band with apparent molecular weight (MW) of 25 kDa is shown in Fig. 1(B). Six peaks were found to have the characteristic 3n Da splitting (n = number of Leu residues in a peptide); expanded views of two of these peaks are shown in the insets of Fig. 1(B). The 3-Da splitting of the peptide signal at 897.4814 shows that it contains one Leu residue; the 9-Da splitting of the peptide signal at 2143.1657 shows that it contains three Leu residues. These externally calibrated *m/z* values were submitted to an MS-FIT[22] search using the SwissProt database for *S. cerevisiae* with 100-ppm error tolerance. One of the proteins identified is phosphoglycerate mutase 1 (MW 27608.8 Da; pI 8.81). The peptide sequence corresponding to 897.4814 is LWIPVNR. The peptide corresponding to 2143.1657 is YVDPNVLPETESLALVIDR. These assignments are consistent with the number of Leu residues revealed by the isotopic mass splitting pattern.

The ratio of the areas of the isotopic peaks of 897.4814 and 900.5011 is measured to be 1.16. The ratio of the isotopic peaks of 2143.1657 and 2152.2149 is 1.13. These ratios directly reflect the relative amount of phosphoglycerate mutase 1 protein expressing in the normal and labeled medium, respectively. Because the same wild-type yeast strain was grown under identical conditions and equal numbers of cells from normal and labeled media were mixed, the theoretical ratio of the isotopic peaks should be at 1 (see Fig. 1(A)). The experimentally determined values 1.13 and 1.16 are reasonably close to the theoretical value of 1 that can be obtained by the calibration with different peptides from the same protein. The experimental derivation between these two peptide signals is 3%. The departure of these two values from the theoretical unity of 1 is due to the systematic uncertainty of cell mixing, the isotope purity of labeled precursors, etc.

We further designed a differential mixing experiment in which the wild-type yeast cells grown in normal and labeled media were mixed at different ratios: 1:5, 1:2, 1:1, 2:1 and 5:1 (mixtures **a–e**, respectively, in Fig. 2(A)). The ratios of the isotopic peak sets are expected to correlate with the corresponding cell mixing ratios, as illustrated in Fig. 2(A).

The cell lysate from each cell mixture was subjected to SDS-PAGE and the PMMs of five protein bands in the range 15–60 kDa were obtained by MALDI-TOFMS. A number of Leu-containing peptides were distinguished by the characteristic 3n Da splitting in the PMM of each protein band. In these protein bands with the same apparent MW but from different mixtures of yeast cells, the ratio of the areas of the light and heavy isotopic peaks changes accordingly. Three representative Leu-containing peptides are shown in Fig. 2(B). Peptides 1 and 3 are from the protein band with apparent MW of 25 kDa identified as phosphoglycerate mutase 1. The sequence of peptide 1 at *m/z* 897.4814 is LWIPVNR and the sequence of peptide 3 at *m/z* 2143.1657 is YVDPNVLPETESLALVIDR. Peptide 2 is from a protein with an apparent MW of 45 kDa; an MS-FIT search of the PMM with the Leu content constraint shows that this peptide at *m/z* 1821.7628 is from enolase 2 (MW 46914.6 Da and pI 5.67) and its sequence is SGETEDTFIADLVVGLR.

In the expanded view of peptides 1, 2 and 3 from cell mixtures **a–e** (Fig. 2(B)), the ratio of the intensity areas of the light to heavy monoisotopic peaks correlates well with the

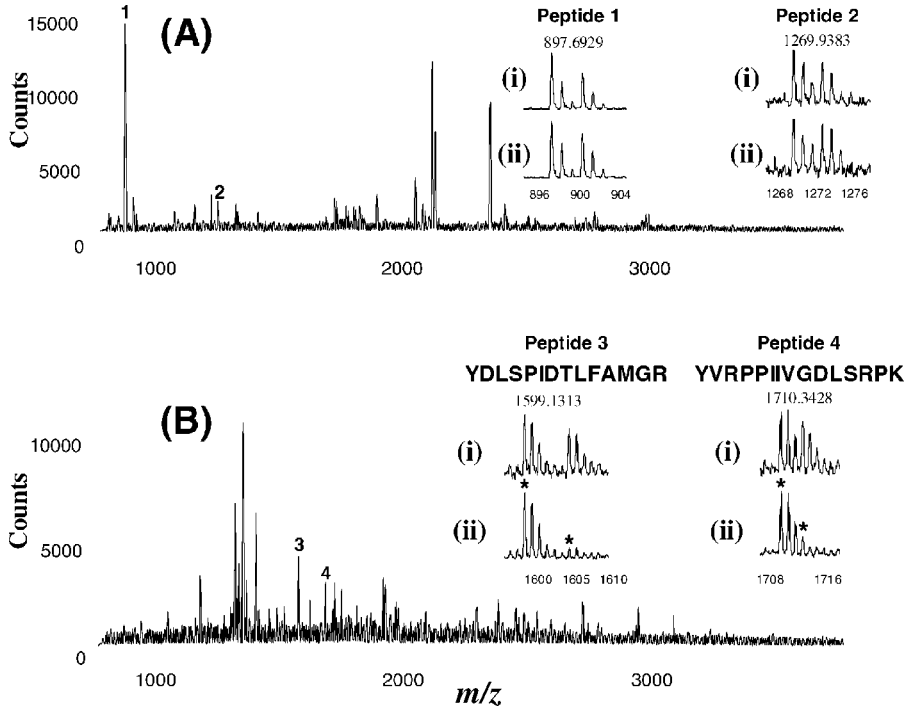


Figure 3. Mass spectrometry of two representative protein spots from 2-D gel electrophoresis. (A) A representative of nine proteins that were equally expressed in the wild-type and *zap1*Δ strains. The expanded mass spectra of peptides 1 and 2 are shown in the inset. (B) One protein, whose expression level was increased approximately four times in the *zap1*Δ strain, was identified to be methionine synthase (Met6). The expanded mass spectra of peptides 3 and 4 are shown in the inset. Mass spectrum (i) for each peptide was obtained from sample (i) that was an equal mixture of wild-type cells grown in normal and labeled media. Mass spectrum (ii) for each peptide was obtained from sample (ii) that was the mixture of an equal number of wild-type cells grown in labeled medium and *zap1*Δ cells grown in normal medium.

mixing ratio of the cells from normal and Leu-d_3 labeled media. The ratios are summarized in Table 1. The experimental ratios are plotted versus the theoretical ratios in Fig. 2(C) and show excellent linear correlations. The linear regression coefficiencies (r^2) for peptides 1–3 are 0.992, 0.990 and 0.994, respectively.

Quantitation of proteins affected by the zinc-responsive transcription factor Zap1 in yeast

Zinc is an essential metal ion for cell growth and plays critical roles in protein structure stabilization and acid-base catalysis.[23,24] Zap1p is a zinc-responsive transcription factor that senses cellular zinc status and increases expression of its target genes in response to zinc deficiency.[21,25,26] DNA microarray studies have revealed that it regulates zinc homeostasis as well as other genes such as iron uptake transporters and alcohol dehydrogenase.[21] Our residue-specific labeling technique for quantitative proteomics was used to measure the relative protein abundances in the wild-type (WT) and the *zap1*Δ strain lacking the Zap1 gene.

WT yeast cells were grown in both normal and Leu-d_3 labeled media; *zap1*Δ mutant cells were grown in the normal medium. Two sets of cells were mixed: (i) equal numbers of WT cells grown in normal and labeled media; and (ii) equal numbers of WT cells grown in the labeled medium and the mutant cells grown in the normal medium. These two sets of cell mixture were lysed and the cell lysate was resolved by 2-D gel electrophoresis. Without any prior knowledge of the protein expression profiles, ten proteins spots were selected at random from the gel of sample (i). The corresponding 10 spots were also picked from the gel of sample (ii). These protein spots were trypsin-digested and the tryptic peptides were extracted and their MALDI-TOF mass spectra were obtained as previously described. Out of the ten protein spots, nine showed no change in expression levels between the WT and *zap1*Δ mutant; the mass spectrum of one representative protein is shown in Fig. 3(A). The mass spectrum of the protein showing different abundances between the WT and *zap1*Δ mutant is given by Fig. 3(B).

The expanded views of the mass spectra of four peptides, 1 and 2 from the protein equally expressed in both cell types, 3 and 4 from the differentially expressed protein, are shown in Fig. 3. Two spectra were obtained for each peptide: the spectrum (i) from sample (i) and the spectrum (ii) from sample (ii). The intensity areas of the light (normal) and heavy (labeled) isotopic peaks in each spectrum (i) are approximately equal as expected. The ratio of the areas of the light and heavy isotopic peaks in each spectrum (ii) gives the relative abundance of corresponding proteins in *zap1*Δ and WT cells. When the areas of the light and heavy isotopic

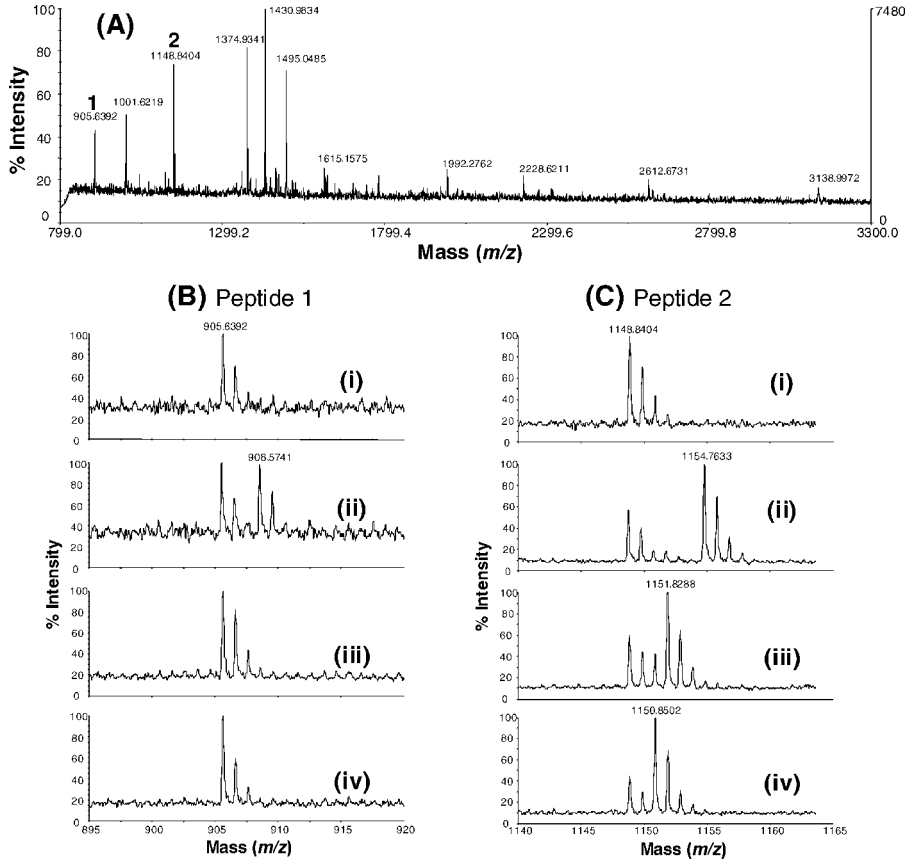


Figure 4. Mass spectrometry of a representative protein band with apparent molecular weight of 25 kDa from SDS-PAGE of HSF cell lysate. (A) PMM of the protein band. (B) Expanded mass spectrum of peptide 1 with *m/z* value of 905.6392. The sequence of peptide 1 is DIIFELR. (C) Expanded mass spectrum of peptide 2 with *m/z* value of 1148.8404. LCQ-MS/MS *de novo* sequencing suggested the sequence is LYRSVNQLR or YLRSVNQLR. Mass spectrum (i) for each peptide was obtained from sample (i) that was cells grown in normal medium. Mass spectrum (ii), (iii) or (iv) for each peptide was obtained from sample (ii), (iii) or (iv) that was an equal mixture of unlabeled cells with Leu-d_3, Ser-d_3 and Tyr-d_2 labeled cells, respectively. Cells grown in labeled medium were irradiated with γ-rays whereas cells grown in normal medium were not irradiated.

peaks are equal, as shown in Fig. 3(A), spectra (ii), we can conclude that the protein expression level does not change between the WT and *zap1*Δ mutant. No further experiments or protein identifications were performed for these equally expressed proteins. When the areas of the light and heavy isotopic peaks are different, it suggests that the protein expression level changes between the WT and *zap1*Δ mutant. Further calculations and protein identifications were only carried out for the differentially expressed proteins.

Using peptide mass mapping, the protein that was differentially expressed between the WT and *zap1*Δ mutant was identified to be methionine synthase (Met6). Peptides 3 and 4 (with *m/z* values 1599.1313 and 1710.3428) should contain two and one leucines because their mass spectra showed characteristic 6 and 3 Da mass splitting, respectively. The MS-FIT search revealed that peptides 3 and 4 are YDLSPIDTLFAMGR and YVRPPIIVGDLSRPK, respectively, which is consistent with the mass splitting pattern. The ratio of the areas of the light and heavy isotopic peaks was calculated to be 4.62 (see Discussion for details on calculation), i.e. methionine synthase is up-regulated more than four times in the *zap1*Δ mutant strain. Met6 is a zinc-containing enzyme catalyzing the methylation of homosysteine.[27] It is reasonable that the expression of Met6 would increase to compensate lack of zinc when the zinc-responsive transcription factor Zap1p is absent, particularly under the limited zinc conditions in this study.

Quantitation of proteins in response to radiation treatment in human skin fibroblast cells

To explore any novel genes that may be involved in radiation-dependent stress responses in human cells,[28] we used residue-specific mass tagging to quantitate differential protein expression between normal and radiated human skin fibroblast cells. As illustrated by Scheme 1, cells grown in the labeled medium were irradiated with γ-rays while cells grown in the normal medium were not. A 1-D SDS-PAGE gel of total cellular proteins was systematically surveyed to identify any induced proteins that may play roles in cellular responses to radiation stress. With several protein bands analyzed, we identified some proteins that were induced by the radiation treatment with many other

proteins unaffected. The overall results of such a study and possible cellular functions of the newly identified inducible genes in radiation response will be presented later. Here, we demonstrate that this residue-specific labeling approach can be effectively employed to quantitate protein relative abundances in human proteomic studies.

Figure 4(A) is the PMM of a protein band with an apparent MW of 25 kDa from a SDS-PAGE gel of HSF cell lysate. The PMM was submitted for MS-FIT search using the NCBInr database for human proteins at 100-ppm error tolerance. Two representative peptides, 1 and 2, were chosen for the detailed analysis and their mass spectra are shown in Figs 4(B) and 4(C), respectively. Spectrum (i) is obtained from sample (i) that is originally from the cells grown in normal medium. Sample (ii) is an equal mixture of cells from normal and Leu-d_3 media. In addition, cells were labeled with two different deuterium-enriched amino acid precursors, Ser-d_3 and Tyr-d_2. Sample (iii) is an equal mixture of cells from normal and Ser-d_3 labeled media, sample (iv) is an equal mixture of cells from normal and Tyr-d_2 labeled media. Their mass spectra are shown as (i), (ii), (iii) and (iv) in Figs 4(B) and 4(C).

Mass spectra of peptide 1 from the multiply labeled samples showed a 3 Da splitting for Leu labeling in Fig. 4B(ii) but no splitting for Ser or Tyr labeling in Fig. 4B(iii) or Fig. 4B(iv), suggesting that this peptide has one Leu residue but no Ser or Tyr residue (the strategy of multiple labeling is described elsewhere; Pan *et al.*, manuscript submitted). An MS-FIT search using the PMM revealed that peptide 1 is from protein BC014207, a protein without a known name and function; its sequence is DIIFELR that is consistent with the multiple labeling experiment. Also, the ratio of the light and heavy isotopic peaks is approximately equal (see Fig. 4B(ii)), suggesting that the expression level of this protein BC014207 is approximately the same in the cells with (Leu-d_3 labeled) and without (unlabeled) radiation treatment.

The mass spectra of peptide 2 at 1148.8404 showed characteristic mass splitting in the multiple labeling experiments: a 6-Da splitting for Leu-d_3 labeling in Fig. 4C(ii), a 3-Da splitting for Ser-d_3 labeling in Fig. 4C(iii) and a 2-Da splitting for Tyr-d_2 in Fig. 4C(iv). The multiple labeling experiments suggest that peptide 2 has two Leu, one Ser and one Tyr residue. However, MS-FIT and MS-Seq searches did not give a valid protein matching with all these parameters. *De novo* sequencing using LC/MS/MS with LCQ DECA ion trap MS suggested that the sequence of peptide 2 is LYRSVNQLR or YLRSVNQLR, which is consistent with the multiple labeling experiments (MS/MS spectrum not shown). However, a Sequest® search using LC/MS/MS data or a BLAST search using the peptide sequence still did not give the protein identity. It is likely that this novel protein is not in the NCBR-nr and Swiss-Prot databases, which has been reported by other researchers.[29] However, regardless of its identity, the relative abundance of this protein X in the irradiated and non-irradiated HSF cells was quantitated by our method. The ratio of the heavy and light isotopic peaks in Fig. 4C(ii) is measured to be 1.97, showing that the expression level of this protein is up-regulated two-fold in the radiation-treated cells grown in the Leu-d_3 labeled medium in comparison with the untreated cells grown in the normal medium. Interestingly, the ratios of the heavy and light isotopic peaks in the Ser-d_3 and Tyr-d_2 labeled samples are found to be 2.11 and 1.85, respectively (see Figs 4C(iii) and 4C(iv)). These three values obtained from Leu-d_3, Ser-d_3 and Tyr-d_2 labeled experiments are consistent with a two-fold increase in the expression of this protein induced by the IR treatment of HSF cells. It also clearly demonstrates the flexibility of this residue-specific labeling approach using various amino acids in studies of cellular processes.

In summary, we found that expression of a novel gene product was significantly induced in HSF cells by γ-ray treatment; as the control, another protein BC014207 in the same band was not affected by the same treatment. This result suggests that this novel protein may be a new component of cellular machinery involved in radiation-related stress responses, such as DNA double-strand break repair. Additional genomic and cloning approaches will be taken to further determine the gene locus and DNA sequence for this novel protein, and its functional roles in radiation responses will also be characterized by genetic and biochemical approaches in future studies.

DISCUSSION

Our strategy of *in vivo* amino acid specific labeling has been used to increase specificity, accuracy and throughput of protein quantification. There are several advantages of this method. Firstly, leucine, which is used most often in this study, is a better choice than cysteine targeted by ICAT and other approaches. Leucine has a higher codon frequency, so therefore we expect higher coverage of leucine-labeled peptides in PMM and quantitation of more proteins. In addition, it is common to obtain MS of two or more Leu-containing peptides from each protein, as shown in Figs 1(B) and 3(B). Thus, we can quantitate the relative abundances of proteins from two sets of mass spectra, i.e. it will be more reliable and accurate. Secondly, a variety of amino acids can be chosen as the labeled precursors using our *in vivo* labeling technique. Other types of amino acids can be alternatively labeled if one specific amino acid does not exist in the peptide of interest. Consequently, more labeled peptide signals can be detected and measured by MS and more cellular proteins can be quantitated. Thirdly, multiple peptides from one protein can be used to normalize the calibration curves, therefore giving more accurate quantitation results. In this study, we have successfully achieved consistent quantitation results using Leu-d_3, Ser-d_3 and Tyr-d_2 in human skin fibroblast cells. In comparison, other approaches, such as ICAT, utilize chemical reactions with one specific type of amino acid and cannot be used if the specific reactive amino acid is not present in the peptide of interest. Therefore, our strategy is more generally applicable and can be flexibly employed in virtually any studies using cell culture systems. Thirdly, the labeled amino acids are incorporated into proteins during cell growth. The proteome is labeled with the isotope mass tags in a residue-specific manner as soon as cells are grown; thus, it does not require any further chemical modifications. Thus, we can avoid the problems caused by side reactions and the relatively low efficiency of the desired reaction. In addition, as we

demonstrated previously,[11–13] and in this study, this amino acid residue specific labeling introduces a characteristic mass splitting pattern that can be easily recognized by MS. Therefore, a large number of spectra can be quickly and accurately analyzed in high throughput. Lastly, equal numbers of cells are mixed in this strategy instead of equal quantities of proteins as required in other approaches. It is relatively easier and more accurate to measure cell numbers rather than quantities of proteins in complex mixtures. Also, two cell populations are mixed during cell harvest before cells are lysed, further eliminating artifacts that may arise during cell lysis.

Though this strategy can be very accurate in theory, experimental errors are inevitable. As previously discussed, the areas of the light and heavy isotopic peaks in Fig. 3, spectra (i) should be equal in theory. In reality, these two peaks have approximately the same areas but the ratio is not exactly equal to 1. The ratio of the light and heavy peaks obtained from spectra (i), R1, can be used as internal standard for calibration. The ratio of the light and heavy peaks directly obtained from spectra (ii), R2, can be used as the relative protein abundance in two cell populations. However, R3 (R3 = R2/R1) is more accurate because it is internally calibrated by R1. R3 takes into account random experimental errors as well as systematic error, such as isotopic interference, thus R3 is more accurate than R2. For example, four steps were taken to calculate the ratio of Met6 abundances in WT and *zap1*Δ mutant cells: firstly, $R_{1,\text{peptide 3}}$ and $R_{1,\text{peptide 4}}$ that were obtained from spectra (i) of peptides 3 and 4; secondly, $R_{2,\text{peptide 3}}$ and $R_{2,\text{peptide 4}}$ were obtained from spectra (ii) of peptides 3 and 4; thirdly, $R_{3,\text{peptide 3}} = R_{2,\text{peptide 3}}/R_{1,\text{peptide 3}}$ and $R_{3,\text{peptide 4}} = R_{2,\text{peptide 4}}/R_{1,\text{peptide 4}}$ were calculated; lastly, the average of $R_{3,\text{peptide 3}}$ and $R_{3,\text{peptide 3}}$ were calculated to be 4.62. By using the internally calibrated R3 value and the average of R3 values calculated for two peptides from the same protein, errors occurring in the experiments can be reduced to a minimum.

There are several limitations of this strategy. Firstly, isotopic tags are incorporated into proteins during cell growth, i.e. samples must be obtained from cultured cells. However, cell culture systems are widely used in basic biological research. Therefore, our strategy is still applicable in many studies. Secondly, Leu-d_3 can only introduce a 3-Da mass shift per Leu residue in the labeled peptides. For many peptides containing only one Leu residue, the monoisotopic distribution peak at [M + 3] may overlap with the signal from the Leu-d_3 labeled peptide, thus resulting in error in measuring protein relative abundance. This interference could be particularly severe for peptides with high mass. However, there are two ways to avoid this potential error. The abundance of the monoisotopic distribution peak at [M + 3] could be calculated and adjusted. Alternatively, the R3 value discussed above, instead of R2, has taken the native [M + 3] isotopic peaks into consideration, thus can be used to reduce the isotopic interference problem by internal calibration.

In conclusion, this *in vivo* amino acid specific labeling strategy can be successfully employed to measure protein relative abundances in two cell populations. It can be internally calibrated to achieve highly accurate quantitation. More importantly, it can label different types of amino acids to quantitate more proteins that cannot be studied by other approaches that target only one specific amino acid.

Acknowledgements

This work is supported by DOE Human Genome Instrumentation (Grant ERW9840) and Los Alamos National Laboratory (LANL) Directed Research Development (LDRD) Grant 200071 (to X.C.) and DOE grant KP1103010 (to X.C. and E.M.B.). H.Z. is supported by the LANL Director's Postdoctoral Fellowship. X.C. is a recipient of a Presidential Early Career Award for Scientists and Engineers (PECASE). We thank Dr. David Eide for generously providing wild-type DY 1457 and *zap1*Δ yeast strains. We thank Dr. Thomas Hunter for discussions and Dr. Karen Larue and Ms. Mary Jo Waltman for assistance in HSF cell culture and radiation treatment. Part of this article was presented at the 50th ASMS annual meeting in June 2002.

REFERENCES

1. Aebersold R, Goodlett DR. *Chem. Rev.* 2001; **101**: 269.
2. Mann M, Hendrickson RC, Pandey A. *Annu. Rev. Biochem.* 2001; **70**: 437.
3. Gygi SP, Rist B, Gerber SA, Turece F, Gelb MH, Aebersold R. *Nat. Biotechnol.* 1999; **17**: 994.
4. Zhou H, Ranish JA, Watts JD, Aebersold R. *Nat. Biotechnol.* 2002; **20**: 512.
5. Geng M, Ji J, Regnier FE. *J. Chromatogr. A* 2000; **870**: 295.
6. Shevchenko A, Chernushevich I, Ens W, Standing KG, Thomson B, Wilm M, Mann M. *Rapid Commun. Mass Spectrom.* 1997; **11**: 1015.
7. Goodlett DR, Keller A, Watts JD, Newitt R, Yi EC, Purvine S, Eng JK, von Haller P, Aebersold R, Kolker E. *Rapid Commun. Mass Spectrom.* 2001; **15**: 1214.
8. Cagney G, Emili A. *Nat. Biotechnol.* 2002; **20**: 163.
9. Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. *Proc. Natl. Acad. Sci. USA* 1999; **96**: 6591.
10. Veenstra TD, Martinovic S, Anderson GA, Pasa-Tolic L, Smith RD. *J. Am. Soc. Mass Spectrom.* 2000; **11**: 78.
11. Chen X, Smith LM, Bradbury EM. *Anal. Chem.* 2000; **72**: 1134.
12. Hunter TC, Yang L, Zhu H, Majidi V, Bradbury EM, Chen X. *Anal. Chem.* 2001; **73**: 4891.
13. Zhu H, Hunter TC, Pan S, Bradbury EM, Chen X. *Anal. Chem.* 2002; **74**: 1687.
14. Münchbach M, Quadroni M, Miotto G, James P. *Anal. Chem.* 2000; **72**: 4047.
15. Schnölzer M, Jedrzejewski P, Lehmann WD. *Electrophoresis* 1996; **17**: 945.
16. Kosaka T, Takazawa T, Nakamura T. *Anal. Chem.* 2000; **72**: 1179.
17. Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C. *Anal. Chem.* 2001; **73**: 2836.
18. Griffin TJ, Gygi SP, Rist B, Aebersold R, Loboda A, Jilkine A, Ens W, Standing KG. *Anal. Chem.* 2001; **73**: 978.
19. Baldwin MA, Medzihradszky KF, Lock CM, Fisher B, Settineri TA, Burlingame AL. *Anal. Chem.* 2001; **73**: 1707.
20. Shevchenko A, Shevchenko A. *Anal. Biochem.* 2001; **296**: 283.
21. Lyons TJ, Gasch AP, Gaither LA, Botstein D, Brown PO, Eide DJ. *Proc. Natl. Acad. Sci. USA* 2000; **97**: 7957.
22. Available: http://prospector.ucsf.edu.
23. Lipscomb WN, Sträter N. *Chem. Rev.* 1996; **96**: 2375.
24. McCall KA, Huang C, Fierke CA. *J. Nutr.* 2000; **130**: 1437S.
25. Zhao H, Eide DJ. *Mol. Cell Biol.* 1997; **17**: 5044.
26. Zhao H, Butler E, Rodgers J, Spizzo T, Duesterhoeft S, Eide D. *J. Biol. Chem.* 1998; **273**: 28713.
27. Matthews RG, Goulding CW. *Curr. Opin. Chem. Biol.* 1997; **1**: 332.
28. Ward JF. *Radiat. Res.* 1994; **138**: S85.
29. Dreger M, Bengtsson L, Schoneberg T, Otto H, Hucho F. *Proc. Natl. Acad. Sci. USA* 2001; **98**: 11943.