Published in final edited form as:

Proteomics. 2010 April ; 10(7): 1408–1415. doi:10.1002/pmic.200900684.

Quantitative analysis of SILAC data sets using spectral counting

Sarah J. Parker^{1,2}, Brian D. Halligan¹, and Andrew S. Greene^{1,2}

¹Biotechnology and Bioengineering Center, Medical College of Wisconsin, WI, USA

²Department of Physiology, Medical College of Wisconsin, WI, USA

Abstract

We report a new quantitative proteomics approach that combines the best aspects of stable isotope labeling of amino acids in cell culture (SILAC) labeling and spectral counting. The SILAC peptide count ratio analysis (SPeCtRA, <http://proteomics.mcw.edu/visualize>) method relies on MS² spectra rather than ion chromatograms for quantitation and therefore does not require the use of high mass accuracy mass spectrometers. The inclusion of a stable isotope label allows the samples to be combined before sample preparation and analysis, thus avoiding many of the sources of variability that can plague spectral counting. To validate the SPeCtRA method, we have analyzed samples constructed with known ratios of protein abundance. Finally, we used SPeCtRA to compare endothelial cell protein abundances between high (20 mM) and low (11 mM) glucose culture conditions. Our results demonstrate that SPeCtRA is a protein quantification technique that is accurate and sensitive as well as easy to automate and apply to high-throughput analysis of complex biological samples.

1 Introduction

An active area of ongoing research is the development and optimization of strategies for the relative quantification of proteins in MS experiments. The goal of such efforts is to generate methods that provide balance between sensitivity, specificity, and accuracy, with the throughput required to meet the particular requirements of clinical or scientific studies.

Stable isotope labeling of amino acids in cell culture (SILAC) has become a common technique in quantitative proteomics [1]. Since the labeling in SILAC occurs at the protein level [2], samples to be compared can be mixed directly after cell lysis thus avoiding variability due to sample preparation and MS. This is in contrast to chemical labeling techniques such as ¹⁸O- or iTRAQ-labeling, which samples are combined after extraction and labeling [3] or spectral counting in which samples are not combined before analysis.

In the SILAC peptide count ratio analysis (SPeCtRA) approach, following sample mixing and protease digestion, SILAC labeled and unlabeled peptides are analyzed simultaneously by LC coupled to data-dependent MS (LC-MS/MS). In the standard SILAC approach,

quantitative differences between the samples are determined by analyzing the MS parent ion chromatograms (PICs) of labeled and unlabeled peptides [4]. While the quantification of SILAC using PICs has been shown to be accurate and reproducible [5], it has limitations. The requirements for high signal-to-noise ratio, high mass accuracy, and the presence of matching peptide pairs from both samples leads to a fraction of peptides that cannot be quantified by using PICs. This can lead to a reduction in the depth of analysis and a failure to analyze proteins that have significant expression in only one sample [3]. Furthermore, quantitation of PICs can be a tedious and relatively time consuming, hands-on procedure that reduces throughput.

The label-free technique of spectral counting is based on the summing of the number of MS² spectra assigned to each protein identified in a sample, comparing the quantity of observed MS² spectral counts for a given protein between separate samples to calculate the ratio of protein abundance between the two samples [6, 7]. Since spectral counting is very easily automated, it is very suitable for high-throughput proteomic analysis. Spectral counting can provide better proteome coverage (*i.e.* more total proteins quantified) relative to PICs [8] and has been shown to have a slightly better dynamic range [9]. Additionally, spectral counting can be used to determine the statistical significance of presence *versus* absence for proteins that occur only in one sample [10]. With spectral counting, the potential exists for a large degree of variability between any two separate LC-MS/MS runs due to differences in sample preparation, protease digestion, chromatography, ion suppression by co-eluting peptides, or stochastic sampling by the instrument [3]. With SPeCtRA, many of these limitations are minimized.

Recently, Zybaïlov *et al.* combined ¹⁵N isotopic labeling with label-free data analysis and directly compared spectral count ratios to paired peptide ion chromatogram abundance ratios [8]. To minimize the technical variability in spectral count ratios, the authors employed a strategy whereby spectral counts from labeled and unlabeled proteins in a 1:1 mixture were generated from separate searches of the same raw data set that differed only in the amino acid masses used by the search. While this work has been cited extensively as support for the validity and utility of spectral counting, the multi-step data analysis protocol they describe is difficult to carry out and has not been extensively validated or applied to other isotopic labeling strategies such as SILAC. In this report, we build on this work by greatly simplifying and automating the data analysis and extending the approach to SILAC labeling.

The objectives in the current study are: (i) to combine the SILAC and spectral counting strategies to generate a new approach, which we termed “SPeCtRA”, for the relative quantification of proteomic data, and (ii) to provide a thorough characterization of the strengths and weaknesses of the SPeCtRA method. To demonstrate the value of the SPeCtRA method for the analysis of complex biological samples, we evaluated the ability of SPeCtRA to detect relevant differences in protein abundance caused by exposure to 20mM hyperglycemic media. Additionally, we show that the SPeCtRA is able to detect and compensate for the intracellular conversion of isotopically labeled ¹³C₆¹⁵N₄L-Arginine into ¹³C₅¹⁵N₁-Proline [11]. This amino acid conversion has been previously shown to reduce the accuracy of SILAC quantitation and required more involved experimental corrections in SILAC studies [12].

2 Materials and methods

2.1 Primary cell culture and SILAC labeling

SILAC protocols were adapted from manufacturer recommendations and previously published protocols [13]. Vascular endothelial cells (VECs) were isolated from 8wk old male Sprague Dawley rats (Harlan, Madison, Wisconsin) as described previously [14]. Cells were allowed to reach confluence and were subsequently split in a 1:3 ratio into Lysine and Arginine depleted RPMI media (Pierce Biotechnology, Rockford, IL, USA) to which the following were added: $^{13}\text{C}_6\text{L}$ -Lysine (40 mg/mL), $^{13}\text{C}_6^{15}\text{N}_4\text{L}$ -Arginine (200 mg/L) (Pierce Biotechnology) 20% FBS (Gibco, Invitrogen, Carlsbach, CO), 1% antibiotic/antimycotic cocktail, and 0.04% Gentamicin (Gibco, Invitrogen, Carlsbach, CO) (heavy media). A parallel set of cells was split 1:3 into media containing all of the above supplements with the exception of unlabeled $^{12}\text{C}_6\text{L}$ -Lysine and $^{12}\text{C}_6^{14}\text{N}_4\text{L}$ -Arginine (light media). At each passage, one plate of cells from each group was split 1:3 into corresponding heavy or light media. The remaining two plates *per* passage (approximately 2×10^6 cells) were rinsed twice in ice cold PBS (without Ca^{2+} or Mg^{2+}), scraped, and pelleted by centrifugation (10 min at $1000 \times g$). Cell pellets were stored at -80°C until subsequent analysis. Cells were cultured through four passages, resulting in eight total pellets (one heavy and one light for each passage). Cells derived from three separate animals were used in the current study. Animal protocols were approved by the Medical College of Wisconsin Institutional Animal Care and Use Committee. Chemicals used were obtained from Sigma-Aldrich (St. Louis, MO, USA) unless otherwise noted.

2.2 Hyperglycemic culture

Three plates of cells from one animal grown through four passages in light media were paired with three plates of cells grown in heavy media to generate three separate treatment replicates. After reaching approximately 80% confluence, cells grown in light media from each pair were switched to media containing 20mM glucose (high glucose) for 72 h at 37°C , with exchange of fresh high-glucose media every 24 h. For each plate of hyperglycemic cells (HG), a separate plate of SILAC labeled cells was grown in SILAC heavy media (as described above) to serve as a normoglycemic control (NG, 11mM glucose). Following 72 h, cells were scraped and pelleted and stored at -80°C until protein extraction.

2.3 Protein extraction and sample preparation

Cell pellets were re-suspended in isolation buffer (200mM mannitol, 70mM sucrose, 10mM HEPES, 1mM EDTA) containing protease inhibitors and lysed mechanically by passage through a 23 gauge needle. Cell lysate was centrifuged at $15\,000 \times g$ for 15min, and the supernatant was collected as a crude soluble fraction. The protein concentration in the soluble fraction was determined *via* a modified Lowry protein assay (Biorad, Hercules CA, USA) [15]. For analysis of SILAC incorporation across cell passages, 50 μg of total protein from each passage was transferred into a fresh tube, and protein digests were performed using a modified “tube gel” protocol, which has been described previously [16]. Briefly, protein solutions were polymerized in 15% acrylamide (Biorad) in the cap of a 1.5mL microcentrifuge tube. Gel pieces were then fixed, washed, and reduced with 10mM DTT and alkylated with 55mM iodoacetamide. Gel pieces were re-suspended in 200mM

ammonium bicarbonate and sequencing-grade porcine trypsin (Promega, Madison WI, USA) was added in a ratio of 1:50 trypsin to protein. Digestion was allowed to proceed for 18 h at 37°C. Digested peptides were extracted with 100% ACN, and peptide solutions were dried, re-suspended in 0.1% Formic acid (FA), desalted on C18 Porous columns (Perceptive Biosystems, Framingham, MA, USA) dried again, and re-suspended in Buffer A (95% HPLC grade H₂O/5% ACN/0.1% FA). For protein mixtures, proteins derived from heavy and light labeled cells were mixed in ratios of 1:1 (25 µg heavy, 25µg light), 2:1 (25 µg heavy, 12.5 µg light), 3:1 (24 µg heavy, 8 µg light), 5:1 (25 µg heavy, 5 µg light), and 10:1 (25 µg heavy, 2.5 µg light). Similarly, proteins derived from cells exposed to hyperglycemia were mixed 1:1 (*i.e.* 25 µg to 25µg) with proteins derived from heavy labeled normoglycemic cells. Following mixing, proteins were prepared and digested as described above.

2.4 Nanospray LC-MS/MS

Peptide mixtures were analyzed with automated nanospray LC-MS/MS, performed on a linear ion trap XL LTQ mass spectrometer (Thermo Scientific, San Jose, CA, USA). The instrument was run in data-dependent mode cycling between MS and MS/MS scans of the six most abundant ions. An injection of 2 µL of peptides suspended in solvent A was loaded onto a 10 cm column of 5 µm C18 resin (Phenomenix, Cheshire, UK) by a Surveyor 2 autosampler (Agilent Technologies). Peptides were eluted from the solid phase with a 180 min gradient between solvent A and solvent B (95% ACN/5% H₂O/0.1% FA) progressing in the following manner: 0–5 min 100% A; 5–60 min 0%–25% B; 60–90 min 25%–75% B; 90–100 min 75%–100% B; 100–115 min 100% B; 115–125 min 100%–0% B; 125–180 min 0% B at a flow rate of 1 µL/min. Dynamic exclusion was used, excluding redundant ions within a ±1.5 Da mass range observed more than twice over a 30 s time frame. Additional instrument settings were as follows: Default charge state of 2, normalized collision energy 35%, minimum signal required 500, automatic gain control “on”.

2.5 Data analysis

The MCW Automated Proteomics Workflow was used to carry out the data analysis steps. Peak lists were generated from raw LC-MS/MS spectra using the program extract_msn (Version 3, Thermo Scientific). For each experiment, raw LC-MS/MS data were automatically searched two times in separate SEQUEST (Version 27, rev 12 Thermo Scientific) runs against the rodent UniProt database, version 54.0 containing 21 226 proteins and 10 908 928 amino acids. Search results were assigned peptide probability scores and filtered using Epitomize (proteomics.mcw.edu/epitomize). Peptide probability scores were determined with a Bayesian classifier using a probability model built with a search of a decoy database in which decoy sequences were constructed from randomized peptides. In the multiple search method, one search is designed to identify only unlabeled peptides (*i.e.* light), and thus did not specify any modification on lysine or arginine residues. In a second search, the search parameters were altered to include fixed modifications of +10 on arginine residues (¹³C₆¹⁵N₄L-arginine) and +6 on lysine residues (¹³C₆L-lysine) as well as a variable modification of +6 for proline residues, so that only heavy peptides would be identified. In both searches, variable modifications of carbamidomethylation of cysteines and oxidation of methionines were also specified. Additional parameters set in both searches were: trypsin

as the specified enzyme with up to three missed cleavages, precursor mass tolerance of ± 2.5 , fragment ion mass tolerance of 0. The following criteria were used to filter protein lists in order to assure high-confident identifications and scan count (SC) ratios: Peptide and protein probability score of at least 0.95 (*i.e.* 5% or lower global false discovery rate); a minimum of two SCs *per* protein for labeling efficiency experiments, a minimum of $n+1$ SCs, where n = the expected ratio for experiments with known heavy to light mixtures, and a minimum of at least four SCs for experiments comparing protein abundance in HG *versus* NG cells. Data from all technical replicates of each passage or mixture were combined, and total SCs for each protein identified in the light and/or heavy searches were compared. Lists of redundant peptide hits (*i.e.* peptides that could not be unambiguously assigned to only one protein) were collapsed to one representative protein. Peptides lacking lysine or arginine were eliminated from the analysis.

To determine the specificity of SPeCtRA in differentiating labeled from unlabeled peptides, LC-MS/MS data sets from three different samples of unlabeled cells were searched with the SPeCtRA method. Since all of the proteins in this data set were unlabeled (light), false hits were defined as any hits to peptides containing heavy arginine or lysine identified in the heavy search (SC_{heavy}). This was compared with the sum of numbers of arginine or lysine peptides identified by either the light (SC_{light}) or the heavy searches (SC_{heavy}). Thus, isotope false identification rate (IFIR) was calculated as:

$$\text{IFIR} = (SC_{\text{heavy}} / (SC_{\text{heavy}} + SC_{\text{light}})) * 100 \quad (1)$$

To determine the efficiency of labeling in our primary cell line, proteins derived from cells grown only in heavy media were used. Percent labeling of each protein was calculated as:

$$\%_{\text{Heavy}} = SC_{\text{heavy}} / (SC_{\text{heavy}} + SC_{\text{light}}) * 100 \quad (2)$$

The mean percent labeling of all proteins in a given cell passage was calculated, and these values were used to determine the change in labeling efficiency across passages one through four.

For mixtures of known amounts of total heavy and light protein, ratios were calculated with the equation:

$$\text{Ratio}_{H/L} = SC_{\text{Heavy}} / SC_{\text{Light}} \quad (3)$$

Incomplete labeling of proteins was accounted for by calculating an adjusted ratio ($\text{Ratio}_{\text{Adj}}$) with the equation:

$$\text{Ratio}_{\text{Adj}} = \text{Ratio}_{H/L} / \text{CF} * \text{Ratio}_{\text{Expected}} \quad (4)$$

Thus, the observed ratio calculated in Eq. 3 is divided by a correction factor determined by percent labeling in experiment 1, and this value was then multiplied by the expected ratio

(Ratio_{Expected}) for a given mixture (*i.e.* 2 for a 2:1 mixture) to yield a Ratio_{Adj} of heavy to light scans for that protein. For HG *versus* NG studies, log₂ ratio and fold-change were calculated from Ratio_{Adj} for each protein, and variability in the abundance ratio for a given protein between treatment replicates was evaluated by observing the SD. To test the statistical significance of protein spectral abundance ratios, we merged the treatment replicates into a combined data set and performed a maximum likelihood statistical significance or *G*-test on each protein ratio [10]. The significance threshold used was $p < 0.05$. Since each protein ratio was considered to be independent of all other protein ratios, we assert that correction for multiple testing for the statistical significance of individual protein ratios is not necessary.

3 Results

3.1 General characteristics of the SPeCtRA method

We have developed a new strategy for quantitative proteomics, SPeCtRA, that provides an alternative for analyzing SILAC MS data by combining the reduced technical variability of SILAC experimental design with the increased analytical depth and throughput of spectral counting. The overall study design is summarized in Fig. 1. Data from a preliminary experiment revealed that spectral counts generated from our method vary linearly ($R^2 = 0.95$, $p < 0.001$) with actual protein abundance (Supporting Information Fig. S1).

The SPeCtRA technique showed a low IFIR, and differentiation between labeled and unlabeled peptides improved at more stringent peptide probability values (Fig. 2). Given the low IFIR (<1.0 %) of the method for differentiating between heavy and light labeled peptides at a probability score of 0.95 or greater, this value was used as a cut-off for filtering all subsequent peptides and proteins in the current experiment.

To account for the metabolic conversion of stable isotope labeled arginine into labeled proline in our cells [11], we added a variable +6Da modification on proline to our SILAC search parameters. We found a number of peptides that contained isotopically labeled proline, examples of which are shown in Supporting Information data (Supporting Information Table S1 and Fig. S2). Overlay of selected MS/MS fragmentation spectra of heavy and light peptides revealed the expected 10Da mass shift of *y* ions derived from isotopically labeled peptides as well as an additional 6Da shift in peptides identified as containing heavy isotopes of both arginine and proline, an example of which is shown in the data supplement (Supporting Information Fig. 2). Therefore, manual analysis of MS/MS spectral data confirms that the SILAC searches are correctly identifying labeled and unlabeled peptides.

3.2 Determination of labeling efficiency in SILAC grown primary endothelial cells by SPeCtRA

In order to test the accuracy of SPeCtRA for quantification of SILAC labeled proteins, it was first necessary to determine the incorporation of heavy $^{13}\text{C}_6\text{L}$ -lysine and $^{13}\text{C}_6^{15}\text{N}_4\text{L}$ -arginine into proteins synthesized in our cell line. A total of 126, 134, 219, and 146 non-redundant proteins across three biological replicates of passages one, two, three, and four, respectively, were used to determine percent labeling. From this data, we observed that

labeling efficiency reached a plateau of approximately 90% after the second cell passage (Fig. 3A). The average heavy isotope incorporation of 91% for proteins from cells in passage four was used as a correction factor to adjust the observed ratios of scans from heavy and light searches for subsequent mixture experiments (see Eq. 4 in Section 2).

3.3 Accuracy of the SPeCtRA method for detecting known protein abundance differences

To determine the ability of SPeCtRA to detect experimentally produced protein abundance differences, we generated a series of mixtures of protein from cells grown in heavy or light media. Mean adjusted ratios for all proteins in mixtures of 1:1, 2:1, and 3:1 were within 9, 11, and 2% (respectively) of expected protein abundance (Fig. 3B), while observed ratios underestimated five- and tenfold differences in protein abundance by greater than 15%. The standard error of the mean for all ratios was low, which supports good overall precision in our method.

3.4 Biological application of SPeCtRA

To demonstrate that the SPeCtRA method could detect changes in protein abundance in a biological experiment, we analyzed 1:1 mixtures of digested proteins derived from unlabeled cells exposed to 20mM glucose (HG) and cells cultured under NG levels in heavy media. Three technical replicates were performed on three treatment replicates, for a total of nine separate LC-MS/MS runs. A total of 278, 276, and 233 protein abundance ratios were detected in biological replicates one through three, respectively. After filtering to include only proteins with a minimum of four SCs, 81, 68, and 55 proteins remained eligible for comparison. Mean abundance ratios were calculated for 55 proteins found in at least two treatment replicates (Fig. 4), 21 of which showed significant abundance differences between the two glucose conditions. We also tested whether combining biological replicates into one master file of heavy and one master file of light search results would improve the depth of our analysis. *G*-tests were performed for each protein, and statistically significant increases or decreases in abundance were determined by *G* values with *p*-value less than 0.05. Combining treatment replicates resulted in a total of 126 protein abundance ratios (*i.e.* total SCs = 4). Of these, seven additional proteins (*i.e.* $n = 28$) were found to have significantly different abundance (Supporting Information Table S2), and the remaining 98 were not significantly different between HG and NG groups (Supporting Information Table S3). Further, 28 proteins were found exclusively in the HG or NG sample, but not both, all of which were statistically significant (Supporting Information Table S4).

4 Discussion

In this study, we describe a method, SPeCtRA, for the high-throughput quantitation of proteins that combines advantageous aspects of SILAC and spectral counting while overcoming some limitations of both techniques. Zybailov *et al.* [8] previously applied a similar approach to the analysis of ^{15}N labeled proteins. In their study, Zybailov *et al.* showed that SC abundance ratios were highly correlated with ratios derived from peptide ion chromatograms, which supports the equivalency in accuracy of the combined labeling plus spectral counting approach compared with PIC quantification. As an advantage over PIC quantification, isotopic labeling analyzed with spectral counting showed better dynamic

range, lower variability, and greater reproducibility. Although the combined method reported by these authors is similar in concept to that implemented by SPeCtRA, it has neither been used subsequently by other independent authors, nor applied to the analysis of other forms of isotopic labeling. Reasons for this may include lack of data on method validation, lack of a publicly available, and high-throughput software tool for performing the analysis, and the use of the method exclusively for ^{15}N rather than simpler and more commonly used isotopic labeling strategies such as SILAC. In the current study, we have built off the method reported by Zybaylov *et al.*, and applied a modified approach for the analysis of SILAC labeled data sets. Further, in contrast to previous work, we have performed extensive validation, detailed methodological information, and developed easily utilized and freely available software tools (<http://proteomics.mcw.edu/visualize>) to allow simple global application of our method.

The results of the current study corroborate and extend those of Zybaylov and colleagues. Specifically, we show that SPeCtRA accurately differentiates between modified and unmodified peptides, with an IFIR below 1.0% for peptides with high probability scores. Using the SPeCtRA method, we determined a 91% labeling efficiency for proteins synthesized in our primary line of VECs. Further, SPeCtRA detected accurate ratios in known mixtures of heavy and light proteins, and the observed ratios increased linearly with known protein abundance. We also demonstrated that SPeCtRA is capable of quantifying proteins that are differentially expressed in cells cultured in 20mM hyperglycemia, detecting significant abundance ratios as low as 1.8-fold.

There is another recently published SILAC quantification method, MaxQuant [17], which is capable of improving the depth of peptide identification and quantification over traditional PIC analysis. This method has been specifically developed and optimized for use with high mass accuracy instruments such as the LTQ Orbitrap and the LTQ FT ICR and the MASCOT search algorithm, which are not available to all researchers. Further, the MaxQuant method, unlike the SPeCtRA method, does not address the problem of isotopic arginine conversion into proline and how this phenomenon would be accounted for in data analysis. Thus, while MaxQuant may be appropriate for certain experimental situations, SPeCtRA provides a simple and accurate alternative, appropriate for use with a wide range of currently available LC-MS/MS instruments and open source software.

Previous studies with SILAC have been performed with immortalized cell lines grown through five or more doublings [18–20]. We achieved 90% labeling within two cell passages, which failed to increase further with subsequent passages. Our rapid label incorporation may have been achieved by splitting cells 1:3; thus cell abundance would triple rather than double with cell passage. The plateau in labeling was likely due to the presence of 20% undialyzed FBS, which provided a constant source of unlabeled amino acids but is required for growth of VECs in culture. We successfully corrected for incomplete labeling by applying a simple correction factor to the observed abundance ratio for each protein. Similar strategies have been applied to correct for incomplete SILAC labeling as low as 75% [21], and the small difference between adjusted observed ratios and expected ratios in our protein mixing experiments validate this strategy. Overall, these

results support the use of SILAC in quantitative proteomics experiments with primary rat VECs.

The validity of the SPeCtRA method for detecting differences in protein abundance was tested in this study by analyzing a series of labeled and unlabeled protein mixtures. While adjusted ratios closely approximated expected abundance differences for proteins in 1:1, 2:1, and 3:1 mixtures, ratios of 5:1 and 10:1 were slightly underestimated. Reduced accuracy of quantitative techniques at higher fold difference has been observed previously, both in studies of traditional SILAC quantification [5] and label-free spectral counting [6] and is a common limitation in MS-based proteomic quantification [22]. For many experiments quantitative goals are simply to identify proteins that are “up” or “down” regulated between conditions, and thus while SPeCtRA is an inappropriate method for absolute quantification, the accuracy and high throughput make it a valuable tool for relative quantification in proteomic experiments.

In this regard, we have shown that SPeCtRA is capable of detecting significant abundance differences in a biological experiment. We were able to quantify the abundance differences of 154 proteins between HG and NG treated cells and found 28 proteins in both HG and NG treatment groups with significantly different abundance. We also detected 28 additional proteins that were exclusively detected only in HG or NG. It is unclear by our analysis whether the “all-or-none” detection of these proteins is an actual biological event or an artifact of the data preparation or LC-MS/MS methods, and additional experiments are necessary to verify such conclusions. Despite the need for additional validation, this finding represents the potential of SPeCtRA for detecting proteins exclusively expressed in only one experimental condition. While we initially identified many more proteins in our SEQUEST search (415), the accuracy of spectral count ratios is diminished when proteins with fewer than four total scans are used [6], and thus we only included proteins with more than four SCs for quantification. The need for several SCs can be a limitation for quantification of low-abundance proteins with spectral counting. Additional replicates can assist in increasing the number of total scans and thereby the number of proteins with sufficient number of scans to be quantified by spectral counting. We observed this when we combined our treatment replicates into one large data set; we nearly tripled the depth of our analysis (*i.e.* 55 comparisons to 154 comparisons).

In conclusion, SPeCtRA is a sensitive, accurate, and high-throughput technique for use in quantitative proteomic experiments, and provides a valuable alternative to existing experimental strategies. SPeCtRA is particularly applicable for experiments performed on mass spectrometers with high scan rates but lower mass accuracy, and allows for high-throughput quantification of a large number of proteins in a complex sample. Software for SPeCtRA can be found at <http://proteomics.mcw.edu/visualize>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The Authors would like to thank Dani Didier for her help with study planning, as well as Justin Friske for his assistance with isolating and culturing the primary rat endothelial cell line used in these experiments. This work was supported by NHLBI proteomics contract N01-HV-28182 and NIH predoctoral training grant T32 HL007852.

Abbreviations

FA	formic acid
HG	hyperglycemic cell
IFIR	isotope false identification rate
NG	normoglycemic
PIC	parent ion chromatogram
SC	scan count
SILAC	stable isotope labeling of amino acids in cell culture
SPeCtRA	SILAC peptide count ratio analysis
VEC	vascular endothelial cell

References

1. Mann M. Functional and quantitative proteomics using SILAC. *Nat Rev Mol Cell Biol.* 2006; 7:952–958. [PubMed: 17139335]
2. Gruhler, S.; Kratchmarova, I. 2D PAGE: Sample Preparation and Fractionation. Posch, A., editor. Humana Press; New York: 2008. p. 101-111.
3. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem.* 2007; 389:1017–1031. [PubMed: 17668192]
4. Ong, S-E.; Mann, M. Quantitative Proteomics by Mass Spectrometry. Sechi, S., editor. Humana Press; New York: 2007. p. 37-52.
5. Ong SE, Kratchmarova I, Mann M. Properties of ¹³C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J Proteome Res.* 2003; 2:173–181. [PubMed: 12716131]
6. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, et al. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics.* 2005; 4:1487–1502. [PubMed: 15979981]
7. Liu H, Sadygov RG, Yates JR. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem.* 2004; 76:4193–4201. [PubMed: 15253663]
8. Zybaïlov B, Coleman MK, Florens L, Washburn MP. Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. *Anal Chem.* 2005; 77:6218–6224. [PubMed: 16194081]
9. Asara JM, Christofk HR, Freemark LM, Cantley LC. A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. *Proteomics.* 2008; 8:994–999. [PubMed: 18324724]
10. Zhang B, VerBerkmoes NC, Langston MA, Uberbacher E, et al. Detecting differential and correlated protein expression in label-free shotgun proteomics. *J Proteome Res.* 2006; 5:2909–2918. [PubMed: 17081042]
11. Austic RE. Conversion of arginine to proline in the chick. *J of Nutr.* 1973; 103:999–1007. [PubMed: 4716199]

12. Bendall SC, Hughes C, Stewart MH, Doble B, et al. Prevention of amino acid conversion in SILAC experiments with embryonic stem cells. *Mol Cell Proteomics*. 2008; 7:1587–1597. [PubMed: 18487603]
13. Amanchy R, Kalume DE, Pandey A. Stable isotope labeling with amino acids in cell culture (SILAC) for studying dynamics of protein abundance and posttranslational modifications. *Sci STKE*. 2005; 2005:12.
14. Pellitteri-Hahn MC, Warren MC, Didier DN, Winkler EL, et al. Improved mass spectrometric proteomic profiling of the secretome of rat vascular endothelial cells. *J Proteome Res*. 2006; 5:2861–2864. [PubMed: 17022658]
15. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the folin phenol reagent. *J Biol Chem*. 1951; 193:265–275. [PubMed: 14907713]
16. Yu H, Wakim B, Li M, Halligan B, et al. Quantifying raft proteins in neonatal mouse brain by 'tube-gel' protein digestion label-free shotgun proteomics. *Proteome Sci*. 2007; 5:17. [PubMed: 17892558]
17. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotech*. 2008; 26:1367–1372.
18. Graumann J, Hubner NC, Kim JB, Ko K, et al. SILAC-labeling and proteome quantitation of mouse embryonic stem cells to a depth of 5111 proteins. *Mol Cell Proteomics*. 2008; 7:672–683. [PubMed: 18045802]
19. Blagoev B, Kratchmarova I, Ong S-E, Nielsen M, et al. A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat Biotech*. 2003; 21:315–318.
20. Ibarrola N, Kalume DE, Gronborg M, Iwahori A, Pandey A. A proteomic approach for quantitation of phosphorylation using stable isotope labeling in cell culture. *Anal Chem*. 2003; 75:6043–6049. [PubMed: 14615979]
21. Gruhler A, Schulze WX, Matthiesen R, Mann M, Jensen ON. stable isotope labeling of *Arabidopsis thaliana* cells and quantitative proteomics by mass spectrometry. *Mol Cell Proteomics*. 2005; 4:1697–1709. [PubMed: 16088002]
22. Smith JR, Olivier M, Greene AS. Relative quantification of peptide phosphorylation in a complex mixture using ¹⁸O labeling. *Physiol Genomics*. 2007; 31:357–363. [PubMed: 17684036]

Figure 1.

Schematic of the SPeCtRA method and overall experimental design. Three separate experiments were conducted in the current study. In experiment 1, proteins from cells grown only in $^{13}\text{C}_6$ - $^{15}\text{N}_4$ -Arginine+ $^{13}\text{C}_6$ Lysine containing media (Heavy) were digested and analyzed with LC-MS/MS. In experiment 2, proteins from cells grown in heavy and light media were mixed to generate samples with known abundance differences, which were subsequently analyzed with LC-MS/MS. In experiment 3, proteins from cells grown in either heavy normoglycemic media (11mM) or light hyperglycemic media (20mM glucose) were mixed in 1:1 ratio, digested, and analyzed with LC-MS/MS. All LC-MS/MS data were then searched with the SILAC search protocol.

Figure 2.

Ability of the SPeCtRA method to differentiate between labeled and unlabeled peptides. Three samples of unlabeled cells were searched with SPeCtRA, and the IFIR of the method (expressed as the percentage of falsely identified labeled peptides (SC_{heavy}) relative to total number of identified peptide scans (SC_{total})) was plotted at increasing peptide probability scores.

Figure 3.

Ability of SPeCtRA to identify differences in protein abundance across cell passages and in experimentally produced protein mixtures. (A) Incorporation of isotopically labeled arginine and lysine across cell passages as determined by spectral counting of SILAC labeled and unlabeled peptides. (B) Observed spectral abundance ratio between known mixtures of isotopically labeled proteins and unlabeled proteins. Error bars correspond to one SD.

Figure 4.

Differences in protein abundance between hyperglycemic and normoglycemic cell treatments. Protein abundance ratios for 55 proteins derived from both the unlabeled high glucose (20mM) treated cells, and the isotopically labeled low glucose (11mM) treated cells. Solid bars represent non-significant abundance ratios. Open bars represent significant abundance ratios. Significant differences were determined by *G*-scores. $p < 0.05$. Error bars correspond to one SD.