METHODS IN MOLECULAR BIOLOGY™ 359

Quantitative Proteomics by Mass Spectrometry

Edited by
Salvatore Sechi

HUMANA PRESS

Quantitative Proteomics by Mass Spectrometry

METHODS IN MOLECULAR BIOLOGY™

John M. Walker, SERIES EDITOR

393. Plant Secondary Metabolites, edited by *Harinder P. S. Makkar, P. Siddhuraju, and Klaus Becker, 2007*

392. Molecular Motors: *Methods and Protocols*, edited by *Ann O. Sperry, 2007*

391. MRSA Protocols, edited by *Yinduo Ji, 2007*

390. Protein Targeting Protocols, Second Edition, edited by *Mark van der Giezen, 2007*

389. *Pichia* Protocols, Second Edition, edited by *James M. Cregg, 2007*

388. Baculovirus and Insect Cell Expression Protocols, Second Edition, edited by *David W. Murhammer, 2007*

387. Serial Analysis of Gene Expression (SAGE): *Digital Gene Expression Profiling*, edited by *Kare Lehmann Nielsen, 2007*

386. Peptide Characterization and Application Protocols, edited by *Gregg B. Fields, 2007*

385. Microchip-Based Assay Systems: *Methods and Applications*, edited by *Pierre N. Floriano, 2007*

384. Capillary Electrophoresis: *Methods and Protocols*, edited by *Philippe Schmitt-Kopplin, 2007*

383. Cancer Genomics and Proteomics: *Methods and Protocols*, edited by *Paul B. Fisher, 2007*

382. Microarrays, Second Edition: *Volume 2, Applications and Data Analysis*, edited by *Jang B. Rampal, 2007*

381. Microarrays, Second Edition: *Volume 1, Synthesis Methods*, edited by *Jang B. Rampal, 2007*

380. Immunological Tolerance: *Methods and Protocols*, edited by *Paul J. Fairchild, 2007*

379. Glycovirology Protocols, edited by *Richard J. Sugrue, 2007*

378. Monoclonal Antibodies: *Methods and Protocols*, edited by *Maher Albitar, 2007*

377. Microarray Data Analysis: *Methods and Applications*, edited by *Michael J. Korenberg, 2007*

376. Linkage Disequilibrium and Association Mapping: *Analysis and Application*, edited by *Andrew R. Collins, 2007*

375. In Vitro Transcription and Translation Protocols: *Second Edition*, edited by *Guido Grandi, 2007*

374. Quantum Dots: *Applications in Biology*, edited by *Marcel Bruchez and Charles Z. Hotz, 2007*

373. Pyrosequencing® Protocols, edited by *Sharon Marsh, 2007*

372. Mitochondria: *Practical Protocols*, edited by *Dario Leister and Johannes Herrmann, 2007*

371. Biological Aging: *Methods and Protocols*, edited by *Trygve O. Tollefsbol, 2007*

370. Adhesion Protein Protocols, *Second Edition*, edited by *Amanda S. Coutts, 2007*

369. Electron Microscopy*: Methods and Protocols, Second Edition*, edited by *John Kuo, 2007*

368. Cryopreservation and Freeze-Drying Protocols, *Second Edition*, edited by *John G. Day and Glyn Stacey, 2007*

367. Mass Spectrometry Data Analysis in Proteomics, edited by *Rune Matthiesen, 2007*

366. Cardiac Gene Expression: *Methods and Protocols*, edited by *Jun Zhang and Gregg Rokosh, 2007*

365. Protein Phosphatase Protocols, edited by *Greg Moorhead, 2007*

364. Macromolecular Crystallography Protocols: *Volume 2, Structure Determination*, edited by *Sylvie Doublié, 2007*

363. Macromolecular Crystallography Protocols: *Volume 1, Preparation and Crystallization of Macromolecules*, edited by *Sylvie Doublié, 2007*

362. Circadian Rhythms: *Methods and Protocols*, edited by *Ezio Rosato, 2007*

361. Target Discovery and Validation Reviews and Protocols: *Emerging Molecular Targets and Treatment Options, Volume 2*, edited by *Mouldy Sioud, 2007*

360. Target Discovery and Validation Reviews and Protocols: *Emerging Strategies for Targets and Biomarker Discovery, Volume 1*, edited by *Mouldy Sioud, 2007*

359. Quantitative Proteomics by Mass Spectrometry, edited by *Salvatore Sechi, 2007*

358. Metabolomics: *Methods and Protocols*, edited by *Wolfram Weckwerth, 2007*

357. Cardiovascular Proteomics: *Methods and Protocols*, edited by *Fernando Vivanco, 2006*

356. High-Content Screening: *A Powerful Approach to Systems Cell Biology and Drug Discovery*, edited by *D. Lansing Taylor, Jeffrey Haskins, and Ken Guiliano, and 2007*

355. Plant Proteomics: *Methods and Protocols*, edited by *Hervé Thiellement, Michel Zivy, Catherine Damerval, and Valerie Mechin, 2007*

354. Plant–Pathogen Interactions: *Methods and Protocols*, edited by *Pamela C. Ronald, 2006*

353. Protocols for Nucleic Acid Analysis by Nonradioactive Probes, *Second Edition*, edited by *Elena Hilario and John Mackay, 2006*

352. Protein Engineering Protocols, edited by *Kristian Müller and Katja Arndt, 2006*

351. *C. elegans:* *Methods and Applications*, edited by *Kevin Strange, 2006*

350. Protein Folding Protocols, edited by *Yawen Bai and Ruth Nussinov 2007*

349. YAC Protocols, *Second Edition*, edited by *Alasdair MacKenzie, 2006*

348. Nuclear Transfer Protocols: *Cell Reprogramming and Transgenesis*, edited by *Paul J. Verma and Alan Trounson, 2006*

347. Glycobiology Protocols, edited by *Inka Brockhausen, 2006*

346. *Dictyostelium discoideum* Protocols, edited by *Ludwig Eichinger and Francisco Rivero, 2006*

345. Diagnostic Bacteriology Protocols, *Second Edition*, edited by *Louise O'Connor, 2006*

344. *Agrobacterium* Protocols, *Second Edition: Volume 2*, edited by *Kan Wang, 2006*

METHODS IN MOLECULAR BIOLOGY™

Quantitative Proteomics by Mass Spectrometry

Edited by

Salvatore Sechi

National Institute of Diabetes
and Digestive and Kidney Diseases
Bethesda, MD

HUMANA PRESS TOTOWA, NEW JERSEY

999 Riverview Drive, Suite 208
Totowa, New Jersey 07512

www.humanapress.com

This publication is printed on acid-free paper.
ANSI Z39.48-1984 (American Standards Institute)

Permanence of Paper for Printed Library Materials.

Production Editor: Jennifer Hackworth

Cover design by Patricia F. Cleary

Cover illustration: Fig. 4B, Chapter 13, "Quantitative Proteomic Analysis of Phosphotyrosine-Mediated Cellular Signaling Networks," by Christine C. Wu and Michael J. MacCoss.

For additional copies, pricing for bulk purchases, and/or information about other Humana titles, contact Humana at the above address or at any of the following numbers: Tel.: 973-256-1699; Fax: 973-256-8341; E-mail: orders@humanapr.com; or visit our Website: www.humanapress.com

Printed in the United States of America. 10 9 8 7 6 5 4 3 2 1

ISSN 1064-3745

E-ISBN 1-59259-255-6

Library of Congress Cataloging-in-Publication Data

Quantitative proteomics by mass spectrometry / edited by Salvatore Sechi.
p. ; cm. -- (Methods in molecular biology, ISSN 1064-3745 ; 359)
Includes bibliographical references and index.
ISBN-13: 978-1-58829-571-2
ISBN-10: 1-58829-571-0 (alk. paper)
1. Mass spectrometry--Technique. 2. Proteomics--Methodology. I. Sechi, Salvatore, 1960- II. Series: Methods in molecular biology (Clifton, N.J.) ; v. 359.
[DNLM: 1. Proteomics--methods. 2. Spectrum Analysis, Mass--methods. W1 ME9616J v.359 2007 / QU 58.5 Q16 2007]
QP519.9.M3Q36 2007
572'.36--dc22 2006015412

Preface

An Overview of Quantitative Proteomics by Mass Spectrometry

The dynamic nature of the proteome and its complexity undoubtfully present huge technological challenges. Recent developments in mass spectrometry (MS) dramatically improved the throughput in protein identification and quantification. Although major advances have been made and these technologies already offer a great opportunity for better understanding human diseases and for identifying biomarkers, it is apparent that we still need to put major efforts to improving the correct methods. For example, the coverage of the proteome for all organisms is still very limited and our understanding of the dynamic processes of posttranslational modifications, at the most, is still rudimentary. *Quantitative Proteomics by Mass Spectrometry* describes in detail the methods and protocols used for many of the most significant recent developments in this field. MS has played a major role in proteomics and it is becoming an essential tool for studying complex biological systems and diseases. Some instruments have reached attomole and even zeptomole sensitivity. Newborns today can be screened for almost 50 diseases, mainly using MS, for less than $2 per disease.

Most of the quantitative proteomics approaches by MS utilize isotopic labels as a reference for either relative or absolute quantitation. These labels can be introduced in vivo, for example, growing an organism in a media enriched with specific isotopes. Several chapters in this book describe this approach. Ong and Mann describe the stable isotope labeling by amino acids in cell culture approach where cells are labeled by growing them in the presence of isotopically labeled amino acids. Wu and MacCoss describe a procedure for labeling mammalian organisms and use the tissues as standard reference in Multidimensional Protein Identification Technology (MudPIT) analysis. Sato et al. describe a procedure where isotopically labeled cells are used as a standard in the analysis of various tissues.

An isotopic label can also be introduced by proteolyzing a protein with trypsin in the presence of ^{18}O-water. Fenselau and Yao describe this approach and Liu et al. couple the labeling with ^{18}O-water to cysteine containing peptide enrichment for high-throughput quantitative proteomics.

Another way to introduce an isotopic label is to react the protein or the protein mixture with a reagent that contains isotopes. Excellent targets for this reaction are cysteines because they are a very reactive nucleophile. Ranish et

al. describe the isotope coded affinity tags (ICAT) approach for quantifying protein complexes, and Turko and Sechi describe the use of acrylamide as a simple tool in quantitative proteomics. Several other amino acids can be the targets of specific reactions with istopically labeled reagents. For example, Regnier describes a coding strategy involving the labeling of both amine and carboxyl groups.

Substantial effort is ongoing in the characterization of posttranslational modifications and perhaps the major advancements have been in the characterization of the phosphoproteome. The procedure described by Zhang et al. couples affinity enrichment of phosphopeptides to stable-isotope labeling and perhaps this is one of the most comprehensive approaches to characterize the phosphoproteome that has been developed. Labeling with isotopes has been used mostly for determining the relative quantities of proteins, isotopically labeled reagents and peptides can also be used for determining the absolute quantities of specific peptides and proteins. For example, Gerber describes how to determine the absolute quantity of a specific protein and its phosphorylation state and Lu et al. describe the use of an isotopically labeled reagent that targets cysteines and that can be used for absolute quantitation.

Tandem MS has been widely used for the detection of inborn errors of metabolism. This is perhaps one of the most apparent applications of MS to disease detection and Turecek et al. describe a procedure for the determinations of enzyme activities that could potentially be used for large-scale screening of newborns. Quantitation can also be achieved without labeling with isotopes and Roy and Becker describe this methodology. The challenge here is to be able to have a highly reproducible system and excellent software for correcting experimental variations that are usually intrinsic in a proteomic experiment.

The methodologies described here are among the leading technologies in quantitative proteomics used today. Their application to complex biological systems and human diseases is becoming a reality. Although we are a long way from a comprehensive understanding of the proteome, considering the pace of recent developments we can be optimistic that MS will indeed play a key role in deciphering the complexity of cellular networks and in the development of patient-tailored medicine.

Salvatore Sechi, PhD

Contents

Contributors

Ruedi Aebersold • *Institute of Molecular Systems Biology, ETH Zurich, Switzerland and Faculty of Science, University of Zurich, Switzerland*

Christopher H. Becker • *Biomarker Discovery Sciences, PPD, Inc., Menlo Park, CA*

Patricia Bottari • *Departments of Chemistry and Biochemistry, University of Washington, Seattle, WA*

Marjorie Brand • *Ottawa Health Research Institute, Molecular Medicine Program, Ottawa, Ontario, Canada*

David G. Camp, II • *Environmental Molecular Sciences Laboratory and Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA*

Catherine Fenselau • *Department of Chemistry and Biochemistry, University of Maryland, College Park, MD*

Michael H. Gelb • *Departments of Chemistry and Biochemistry, University of Washington, Settle, WA*

Scott A. Gerber • *Department of Genetics, Norris Cotton Cancer Center and Dartmouth Medical School, Hanover, NH*

Steven P. Gygi • *Department of Cell Biology, Harvard Medical School, Boston, MA*

Yasushi Ishihama • *Eisai Co., Ltd., Tsukuba, Ibaraki, Japan*

Arminja N. Kettenbach • *Department of Cell Biology, Harvard Medical School, Boston, MA*

Tao Liu • *Environmental Molecular Sciences Laboratory and Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA*

Yu Lu • *Departments of Chemistry and Biochemistry, University of Washington, Settle, WA*

Michael J. MacCoss • *Department of Genome Sciences, University of Washington, Seattle, WA*

Matthias Mann • *Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Martinsried, Germany*

Yoshiya Oda • *Eisai Co., Ltd., Tsukuba, Ibaraki, Japan*

Shao-En Ong • *Proteomics and Biomarker Discovery, The Broad Institute of MIT and Harvard, Cambridge, MA*

Wei-Jun Qian • *Environmental Molecular Sciences Laboratory and Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA*

Jeffrey A. Ranish • *Institute for Systems Biology, Seattle, WA*
Fred E. Regnier • *Department of Chemistry and Cancer Center, Purdue University, West Lafayette, IN*
Sushmita Mimi Roy • *Biomarker Discovery Sciences, PPD, Inc., Menlo Park, CA*
John Rush • *Cell Signaling Technology, Beverly, MA*
Toshitaka Sato • *Eisai Co., Ltd., Tsukuba, Ibaraki, Japan*
Salvatore Sechi • *National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD*
C. Ronald Scott • *Department of Pediatrics, University of Washington, Seattle, WA*
Richard D. Smith • *Environmental Molecular Sciences Laboratory and Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA*
Illarion V. Turko • *Center for Advanced Research in Biotechnology, NIST/UMBI, Rockville, MD*
František Tureček • *Department of Chemistry, University of Washington Seattle, WA*
Forest M. White • *Biological Engineering Division, MIT, Cambridge, MA*
Alejandro Wolf-Yadlin • *Biological Engineering Division, MIT, Cambridge, MA*
Christine C. Wu • *Department of Pharmacology, University of Colorado Health Sciences Center, Aurora, CO*
Xudong Yao • *Department of Chemistry, University of Connecticut, Storrs, CT*
Yi Zhang • *Biological Engineering Division, MIT, Cambridge, MA*

1

Acrylamide—A Cysteine Alkylating Reagent for Quantitative Proteomics

Illarion V. Turko and Salvatore Sechi

Summary

Mass spectrometry-based relative quantification of proteins is often achieved by the labeling of two samples with isotopically light and heavy reagents. The intensities of the ions with different masses, but same chemical properties, can be reliably used for determining relative quantities. Several strategies of labeling with various weakness and strength and degrees of complexity have been described. In this chapter, we describe a simple and inexpensive protein-labeling procedure based on the use of acrylamide and deuterated acrylamide as a cysteine alkylating reagent. Gel electrophoresis is one of the most commonly used techniques for analyzing/visualizing proteins, thus, we emphasize the use of acrylamide as a labeling procedure for quantifying proteins isolated by one- and two-dimensional polyacrylamide gel electrophoresis.

Key Words: Quantitation; proteomics; acrylamide; cysteine alkylation; gel electrophoresis; mass spectrometry; isotopic labeling; quantitative proteomics.

1. Introduction

Quantitative proteome profiling using mass spectrometry (MS) and protein labeling with stable isotopes can provide the relative abundance of many individual proteins within two different samples ***(1,2)***. The isotopic label can be introduced in vivo or in vitro ***(2)***. For example, a label can be easily introduced in vivo in bacteria or yeast by growing them in media enriched with specific isotopes ***(3)***. A label can also be introduced by performing a tryptic digest in the presence of ^{18}O-water ***(4–6)***. Several approaches using isotopically labeled reagents that react with specific amino acids or the protein N-terminus have also been developed ***(2)***.

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

Cysteine is a strong nucleophile that is readily modified by a variety of reagents ***(7)*** and has been used as a target for attaching isotopic labels by several reagents including ICAT ***(8)*** and acrylamide ***(7)***. The use of acrylamide and deuterated acrylamide (D_3-acrylamide) in quantitative proteomics was first introduced by Sechi ***(9)*** and was later combined with two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) ***(10–12)***. Acrylamide together with D_3-acrylamide has been used for quantifying proteins isolated from various biological samples ***(9–14)***. The alkylation reaction with acrylamide when optimized was found to be highly specific ***(7)*** and the accuracy and dynamic range in proteomic experiments were evaluated and further validated in several studies ***(9–14)***.

In a typical quantitative proteomic experiment that involves acrylamide as alkylating reagent, one sample is labeled with acrylamide and the other sample is labeled with D_3-acrylamide (**Fig. 1**). D_3-acrylamide has three deuteriums instead of three protons, and, thus, is three mass units heavier than acrylamide. Acrylamide and D_3-acrylamide have the same chemical properties and equal alkylation efficiency. After mixing the acrylamide- and D_3-acrylamide-labeled samples in a 1:1 ratio, the mixture can be loaded on a one-dimensional (1D)- or 2D-PAGE separation system. The bands (1D) or spots (2D) can be excised, digested with an enzyme (e.g., trypsin), and analyzed by MS. The isotopic distributions of the cysteine-containing peptides as shown in the mass spectra are the result of the overlap of two isotopic envelopes that differ by three mass units (the difference between acrylamide and D_3-acrylamide). Quantification is accomplished by comparing the intensities of the normal and deuterated mass components of the cysteine-containing peptide(s) (**Fig. 1**). Simultaneous protein identification can usually be accomplished by peptide mass finger printing ***(9)*** or by obtaining the tandem mass spectra of a few peptides.

Cysteine alkylation is an important step in proteomics for improving the protein identification process and/or for obviating heterogeneous alkylation that might occur from the reaction with unpolymerized acrylamide during electrophoresis ***(7)***. Free thiols are relatively strong nucleophiles and have a tendency to be easily modified by several chemicals, thus, it is important to stabilize them with a well-controlled reaction, such as the alkylation with acrylamide or other well-characterized alkylating reagents prior to electrophoresis. Considering the importance of stabilizing the cysteines by alkylation, the quantitation that can be obtained by alkylating with isotopically labeled acrylamide could also be considered a “free” added benefit.

The analysis of samples using 1D-PAGE is quite simple and fast. It can be used for studying/characterizing specific proteins or relatively simple protein

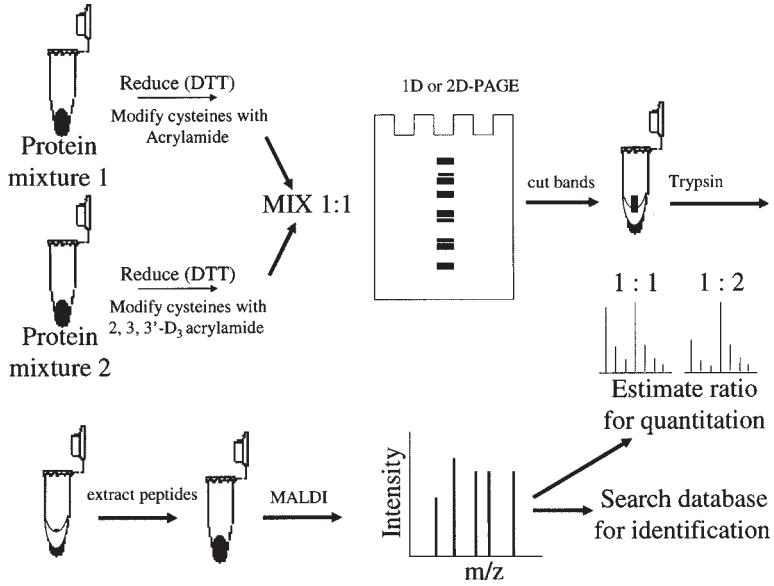


Fig. 1. Diagram depicting the general scheme for relative quantification. The cysteines from two protein mixtures in which the relative protein quantification needs to be determined are separately reduced and alkylated using acrylamide and D_3-acrylamide. The two samples are then mixed in a 1:1 ratio and the proteins are separated using gel electrophoresis. The proteins of interest are excised from the gel, digested with trypsin, and the tryptic peptide maps obtained by MALDI-TOF. The isotopic envelopes of the cysteine-containing peptides are used to determine the relative quantities of the proteins in the two samples. A schematic representation of the isotopic envelope for a cysteine-containing peptide in the case when the protein is present in equal amounts and in a 1:2 ratio in the two samples is shown. The same mass spectrum is used for identifying or confirming the protein. Reproduced with permission from **ref. *9***.

mixtures. For example, characterization of protein complexes isolated by immunoprecipitation can be easily achieved by 1D-PAGE and MS ***(15)***. For resolving/studying complex protein mixtures, like whole protein lysates, the use of 2D-PAGE or other separation procedure coupled to 1D-PAGE is recommended. Here, we will not describe the coupling of other separation methods (e.g., ion exchange chromatography) to 1D-PAGE, and we will limit the description to the last steps of these procedures (i.e., 1D-PAGE and MS) that is common to most of them. Analysis by 2D-PAGE is definitely more complex and delicate than 1D-PAGE. One of the reasons for this is that it can be easily

hampered by the presence of lipids and salts in the sample; thus, an additional chloroform/methanol precipitation of the sample is often required for ensuring a good quality separation. In this chapter, we illustrate the quantitation of bovine serum albumin (BSA) isolated and separated by 1D-PAGE, and the quantitation of several mitochondrial proteins isolated from the heart of diabetic and normal rats separated by 2D-PAGE. The alkylation with acrylamide and D_3-acrylamide reveals itself to be a simple and inexpensive proteomic tool for determining the relative quantity of proteins. An intrinsic benefit of this quantitative proteomic approach is the stabilization of the cysteines.

2. Materials

2.1. Quantitation of Proteins Separated by 1D-PAGE

1. BSA stock solution: BSA (Sigma-Aldrich, St. Louis, MO) and other protein standards solutions are obtained by weighing the protein crystals and dissolving them in water.
2. Reducing sample buffer for 1D-PAGE: for preparing this solution, 25 m*M* dithiothreitol (DTT) is added to NuPAGE sample buffer (Invitrogen, Carlsbad, CA). NuPAGE buffer consist of 1.09 *M* glycerol, 141 m*M* Tris-base, 106 m*M* Tris-HCl, 73 m*M* lithium dodecylsulfate, 0.51 m*M* ethylenediaminetetraacetic acid (EDTA), 0.22 m*M* Coomassie G-250, and 0.175 m*M* phenol red, pH 8.5. A concentrated 2X or 4X sample buffer could be used and should be appropriately diluted with the protein sample to be analyzed or water.
3. Alkylating stock solutions: a 4 *M* acrylamide (Bio-Rad, Hercules, CA) and 4 *M* 2,3,3'-D_3-acrylamide (Cambridge Isotope Laboratories, Inc., Woburn, MA) stock solutions are prepared in water (*see* **Note 1**). Acrylamide and D_3-acrylamide are highly toxic and light sensitive. Stock solutions can be stored at –20°C. Extreme caution is needed when weighting and handling acrylamide crystals. These crystals should be handled in a chemical hood.
4. 1D-PAGE buffer: precast 10% acrylamide Nu-PAGE gel (Invitrogen) or other gel of appropriate acrylamide concentration are run using a MES buffer (Invitrogen) that consist of 50 m*M* 2-(*N*-morpholino) propane sulfonic acid, 50 m*M* Tris-base, 3.5 m*M* sodium dodecylsulfate (SDS), and 1 m*M* EDTA, pH 7.3.

2.2. Quantitation of Proteins Separated by 2D-PAGE

2.2.1. Preparation, Reduction, and Alkylation of Soluble Mitochondrial Proteins

1. Male Sprague-Dawley rats (median body weight 250 g; Harlan, Indianapolis, IN) are maintained on standard rat chow and tap water *ad libitum.*
2. STZ solution: 15 mg/mL STZ (Sigma-Aldrich) in 0.1 *M* citrate buffer, pH 4.5. This solution should be prepared fresh before injecting the rats.

3. Extraction buffer: 5 m*M* HEPES-KOH, pH 7.4, 0.25 *M* mannitol, 0.1 m*M* EDTA, and 0.1% BSA.
4. Percoll solution: a 30% Percoll (Amersham Biosciences, Piscataway, NJ) solution is prepared in extraction buffer following the manufacturer instruction.
5. Solubilizing buffer: 2 *M* thiourea, 7 *M* urea, and 4% CHAPS in 20 m*M* Tris-HCl, pH 8.5.
6. Acrylamide solutions: 4 *M* acrylamide and 4 *M* D_3-acrylamide solutions are prepared in solubilizing buffer.

2.2.2. Chloroform/Methanol Precipitation

1. Methanol, chloroform, and water used for the precipitation are HPLC grade. Methanol and chloroform are toxic and should be handled with appropriate care in a chemical hood (*see* **Note 2**).

2.2.3. 2D-PAGE

1. Immobilized pH gradient strips pH 3.0–10.0 (BioRad) are prepared following the manufacturer instructions.
2. 2D-PAGE sample buffer: 2 *M* thiourea, 7 *M* urea, 4% CHAPS, 0.5% Triton X-100, 0.2% Biolytes 3-10, and some bromophenol blue.
3. Equilibration buffer: 50 m*M* Tris-HCl, pH 8.8, with 6 *M* urea, 30% glycerol, 2% SDS, and some bromphenol blue.
4. Electrode buffer: 25 m*M* Tris-HCl, pH 8.3, with 192 m*M* glycine and 0.1% SDS.

2.3. 1D- and 2D-PAGE Staining

1. Coomassie solutions: 0.1% Coomassie blue R-250 in 50% methanol/10% acetic acid. Coomassie Simply Blue™ solution (Invitrogen) or other Coomassie solution might also be used.

2.4. In-Gel Digestion and Peptide Extraction of Proteins Isolated by 1D- and 2D-PAGE

1. Destaining solutions: 25 m*M* NH_4HCO_3/50% CH_3CN; 50% CH_3CN/0.1% FC_3COOH.
2. Trypsin digestion solution: 5 μg/mL of sequencing grade-modified trypsin (Promega, Madison, WI) in 25 m*M* NH_4HCO_3, pH 8.0 (no need to adjust pH).
3. Siliconized 0.65-mL microcentrifuge tubes (Corning Inc., Corning, NY) or other nonstick tubes (*see* **Note 3**).

2.5. MALDI-MS

1. MALDI matrix solution: 10 mg/mL α-cyano-4-hydroxycinnamic acid (Sigma-Aldrich) in 50% CH_3CN/0.1% FC_3COOH (*see* **Note 4**). The solution should be

vortexed for 5 min and centrifuged at 15,000*g* for 2–3 min to remove any undissolved matrix.

3. Methods

Depending on the complexity of the protein sample, a 1D- or 2D-PAGE separation might be appropriate. The type of protein separation to be used determines how the sample should be handled. Here, we use similar but not identical procedures to demonstrate an accurate relative quantification of proteins labeled with acrylamide and D_3-acrylamide, which are then separated by 1D-PAGE (**Figs. 2** and **3**) or 2D-PAGE (**Figs. 4** and **5**).

3.1. Quantification of Proteins Isolated by 1D-PAGE

3.1.1. Sample Reduction and Alkylation

1. Prior to cysteine alkylation it is important to ensure that all cystines are reduced to cysteines. For this purpose, the solution containing the protein of interest, in this case a specific amount of BSA, is mixed with reducing sample buffer and incubated at 100°C for 5 min (*see* **Note 5**). The sample should be cooled down to room temperature and briefly centrifuged to spin down the condensate.
2. For determining the relative quantity of a protein, two samples are reduced and alkylated separately using two different alkylating reagents. In the example shown in **Figs. 2** and **3**, the two samples to be compared, containing various amount of BSA, were alkylated with acrylamide or D_3-acrylamide.
   One sample is alkylated by adding 1/10 (v/v) of acrylamide and the other sample is alkylated adding 1/10 (v/v) of D_3-acrylamide (e.g., to 8–20 µL of reduced sample, 2 µL of alkylating solution are added). The solutions are mixed by gently pipetting several times and then left in the dark at room temperature for 1 h (*see* **Note 5**).
3. After completion of the alkylation reaction described in **step 2**, the two samples to be compared (one alkylated with acrylamide and the other with D_3-acrylamide) are mixed in a 1:1 ratio (v:v).

3.1.2. 1D-PAGE

1. Samples to be quantified are loaded on a 10% acrylamide NuPAGE gel and run using an Invitrogen MES buffer following the manufacturer instructions.
2. After the run is completed, the gel is stained with a Coomassie solution.

3.2. Quantitation of Proteins Separated by 2D-PAGE

3.2.1. Preparation, Reduction, and Alkylation of Soluble Mitochondrial Proteins

In this example we use alkylation with acrylamide and D_3-acryalmide to identify and quantify changes in the protein profile caused by diabetes in heart mitochondria from STZ-treated rats.

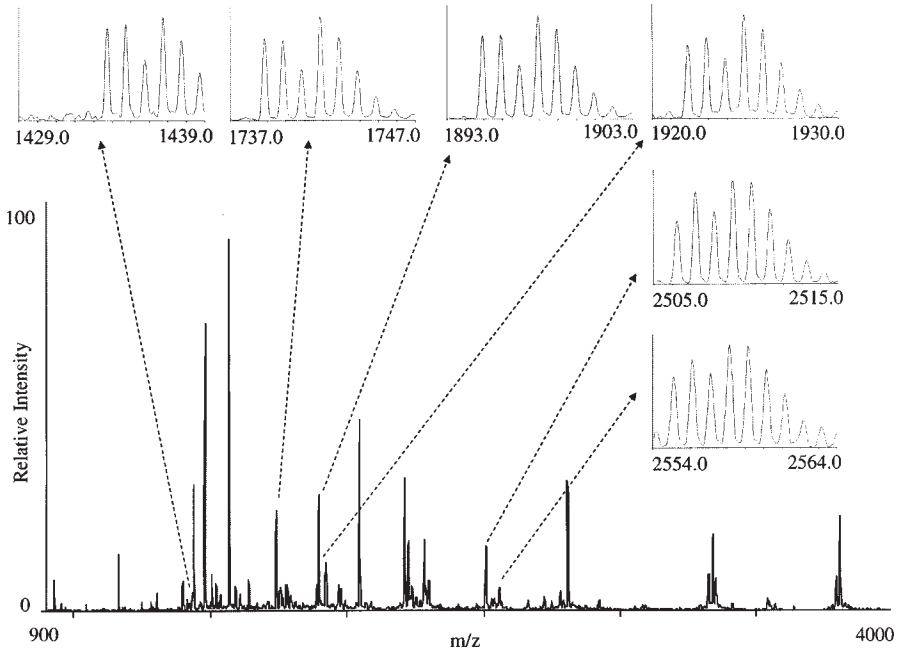


Fig. 2. MALDI-TOF spectrum of the tryptic digest of acrylamide-labeled bovine serum albumin (BSA). Two solutions each containing 250 femtomoles of BSA were reduced and alkylated separately with acrylamide or D_3-acrylamide and mixed with a 1:1 ratio. The sample was then analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and the band corresponding to BSA was excised and digested with trypsin. The MALDI-TOF spectrum of this tryptic digest is shown. The six panels show zoom images of the isotopic envelopes of six cysteine-containing peptides. Reproduced with permission from **ref. *9***.

1. To trigger insulin-dependent diabetes mellitus, rats are injected intraperitoneally with STZ at a dose of 60 mg of STZ per kg of body weight. Control animals receive the corresponding volume of citrate buffer. Diabetic rats and aged matched control rats are sacrificed 1 or 4 wk after STZ injection ***(13)***.
2. Homogenize individual rat hearts in extraction buffer. Large cell debris and nuclei should be first pelleted by centrifugation at 4000*g* for 20 min and discarded. Mitochondria are then pelleted from the supernatant by centrifugation at 12,000*g* for 20 min. To further purify mitochondria, various density gradients could be used. In this particular experiment, mitochondria are suspended in extraction buffer, loaded on the top of the Percoll solution, and centrifuged for 30 min at 95,000*g*. Percoll-purified mitochondria are collected from the lower part of the dense, brownish-yellow band, washed twice with extraction buffer, sonicated, and centrifuged at 100,000*g* for 1 h. The high-speed supernatant fraction includes soluble mitochondrial proteins.

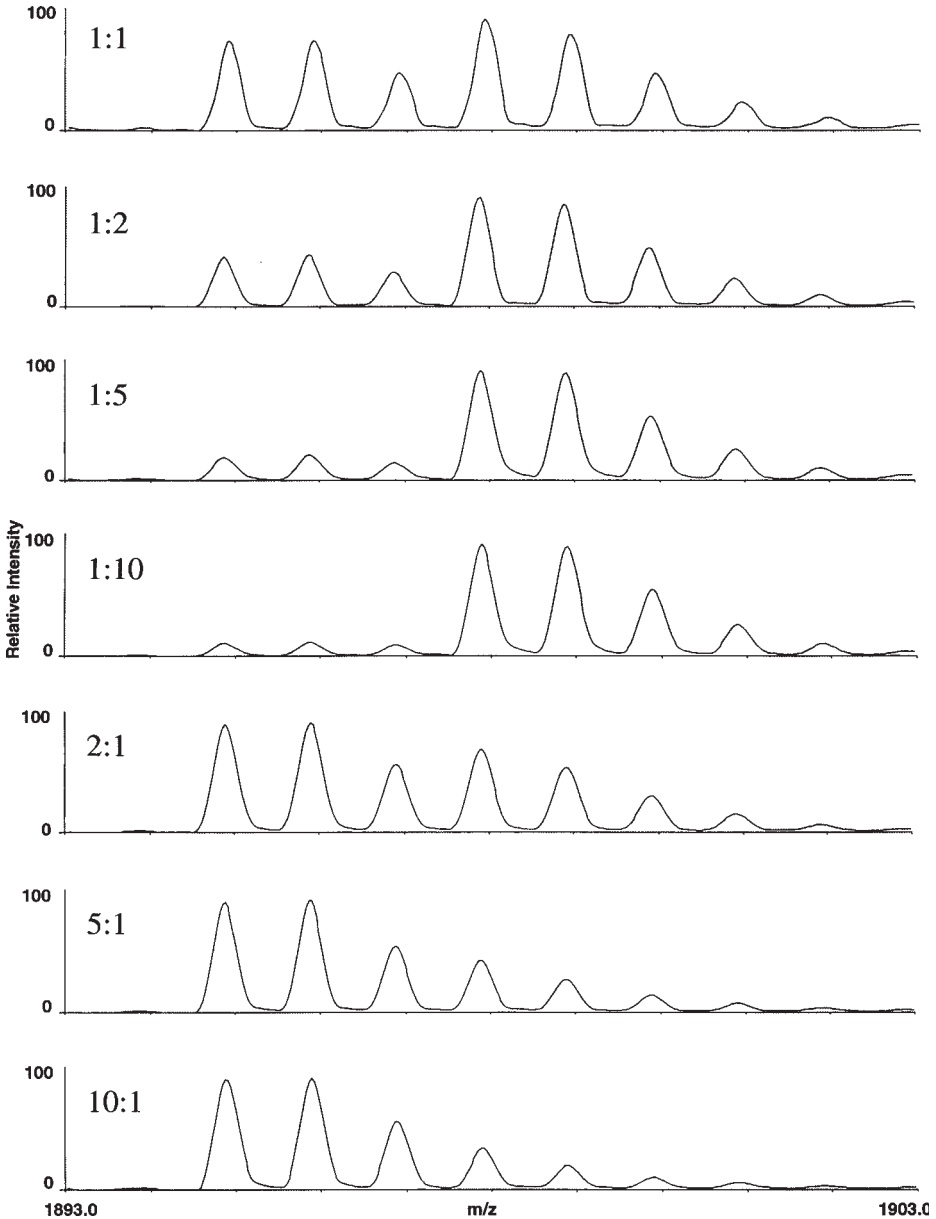


Fig. 3. Isotopic envelopes for one cysteine-containing peptide from the MALDI-TOF spectra of acrylamide-labeled bovine serum albumin (BSA). Two hundred fifty picomoles of BSA treated with acrylamide and 250 picomoles of BSA treated with D_3-acrylamide were mixed at different ratios (1:1, 1:2, 1:5, 1:10, 2:1, 5:1, and 10:1) and run on a one-dimensional polyacrylamide gel electrophoresis. The bands corresponding to

3. Use solubilizing buffer with 40 m*M* DTT for preparing samples containing 1 mg/mL of soluble mitochondrial proteins and incubate the solutions at room temperature for 1 h.
4. Mix 90 µL of reduced mitochondrial proteins from the control heart with 10 µL of acrylamide solution, and mix 90 µL of reduced mitochondrial proteins from the diabetic heart with 10 µL of D_3-acrylamide solution. Gently pipet the two samples to ensure good mixing and leave the samples in the dark at room temperature for 5 h (*see* **Note 6**).
5. In 1.5-mL conical microcentrifuge tubes, mix mitochondrial proteins treated with acrylamide and mitochondrial proteins treated with D_3-acrylamide in a 1:1 ratio (v:v). To ensure good mixing, pipet the sample several times and then proceed to the chloroform/methanol precipitation.

3.2.2. Chloroform/Methanol Precipitation

The main purposes of the chloroform/methanol protein precipitation are to delipidate and desalt the sample. This is a critical step for improving the quality of the 2D-PAGE separation. The following procedure is described for a 150-µL protein sample. Smaller samples might be adjusted to 150 µL with water and larger samples can be divided in several microcentrifuge tubes.

1. Add 600 µL of methanol, 150 µL of chloroform, and 450 µL of H_2O to a protein sample in the listed order.
2. Briefly vortex the tubes, then centrifuge the tubes using a microcentrifuge at 16,000*g* for 10 min at room temperature. Expect to see separation into two phases with a precipitated protein layer on the interface. Carefully remove and discard the upper phase, making sure to keep the protein layer untouched.
3. Add another 600 µL of methanol and centrifuge at 16,000*g* for 10 min at room temperature. Expect to see a single liquid phase and a solid protein pellet at the bottom of tube. Carefully remove the liquid phase and let the protein pellet air-dry for 3–5 min.
4. On completion, the expected amount of precipitated acrylamide-labeled mitochondrial proteins should be approx 135 µg. This amount is suitable for a single 2D-PAGE separation using 7-cm immobilized pH gradient strips.

3.2.3. 2D-PAGE

1. Dissolve protein pellets in 125 µL of 2D-PAGE sample buffer and rehydrate a 7-cm immobilized pH gradient strip (pH 3.0–10.0) with this sample. During loading a sample in the focusing tray channel, take care not to introduce any bubbles

---

Fig. 3 *(continued from opposite page)* BSA were excised, digested with trypsin, and the MALDI-TOF spectra of the tryptic peptides obtained. The isotopic envelopes, for all ratios, of one cysteine-containing peptide, are shown. Reproduced with permission from **ref. 9**.

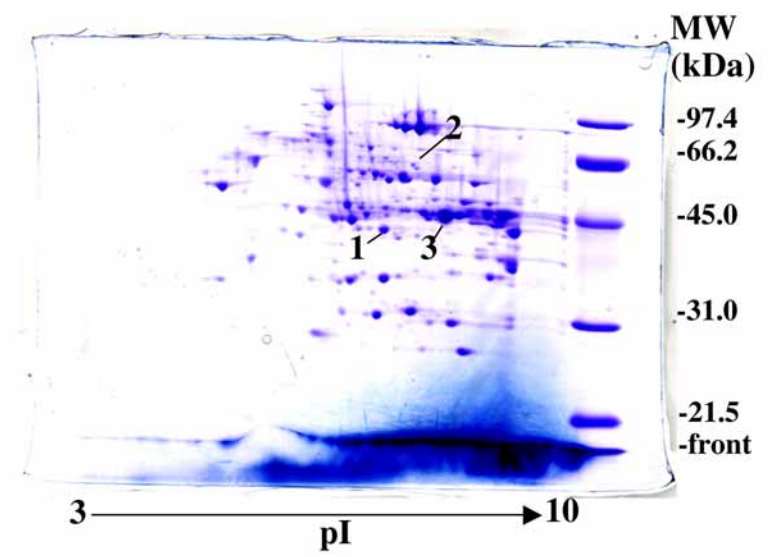


Fig. 4. Two-dimensional polyacrylamide gel electrophoresis pattern of acrylamide-labeled soluble mitochondrial proteins from rat heart. Soluble mitochondrial proteins from normal and 4 wk diabetic rat heart were treated with normal and deuterated acrylamide, respectively, and mixed at a 1:1 ratio. Many proteins were identified and quantified. The selected proteins which quantification is shown in the **Fig. 5** are numbered: 1, acyl CoA dehydrogenase, short chain; 2, catalase; 3, creatine kinase. Reproduced with permission from **ref.** ***12***.

which may interfere with the contact of the sample and immobilized pH gradient strip. Place the strip gel side down toward the sample and apply mineral oil to each channel containing a sample and strip. Make sure the entire strip is covered. After rehydration, conduct isoelectric focusing at 250 V for 15 min, linearly increase voltage to a maximum of 4000 V over a 2-h period, and then run to accumulate a total of 20,000 V/h.

2. For the second dimension (*see* **Note 7**), equilibrate the strips for 15 min in equilibration buffer. Embed the strips in 0.7% (w/v) agarose on the top of the appropriate polyacrylamide slab gels and run SDS-PAGE separation using the electrode buffer.
3. After the run is completed, stain the gels with Coomassie solution for 10 min at room temperature and destain it with 10% acetic acid/20% methanol.

3.3. In-Gel Digestion and Peptide Extraction

1. Excise Coomassie-stained protein bands/spots of interest from a polyacrylamide gel. Cut each gel piece into small particles (~1 mm^2) using a scalpel, and place into a 0.65-mL siliconized tube (*see* **Note 8**).

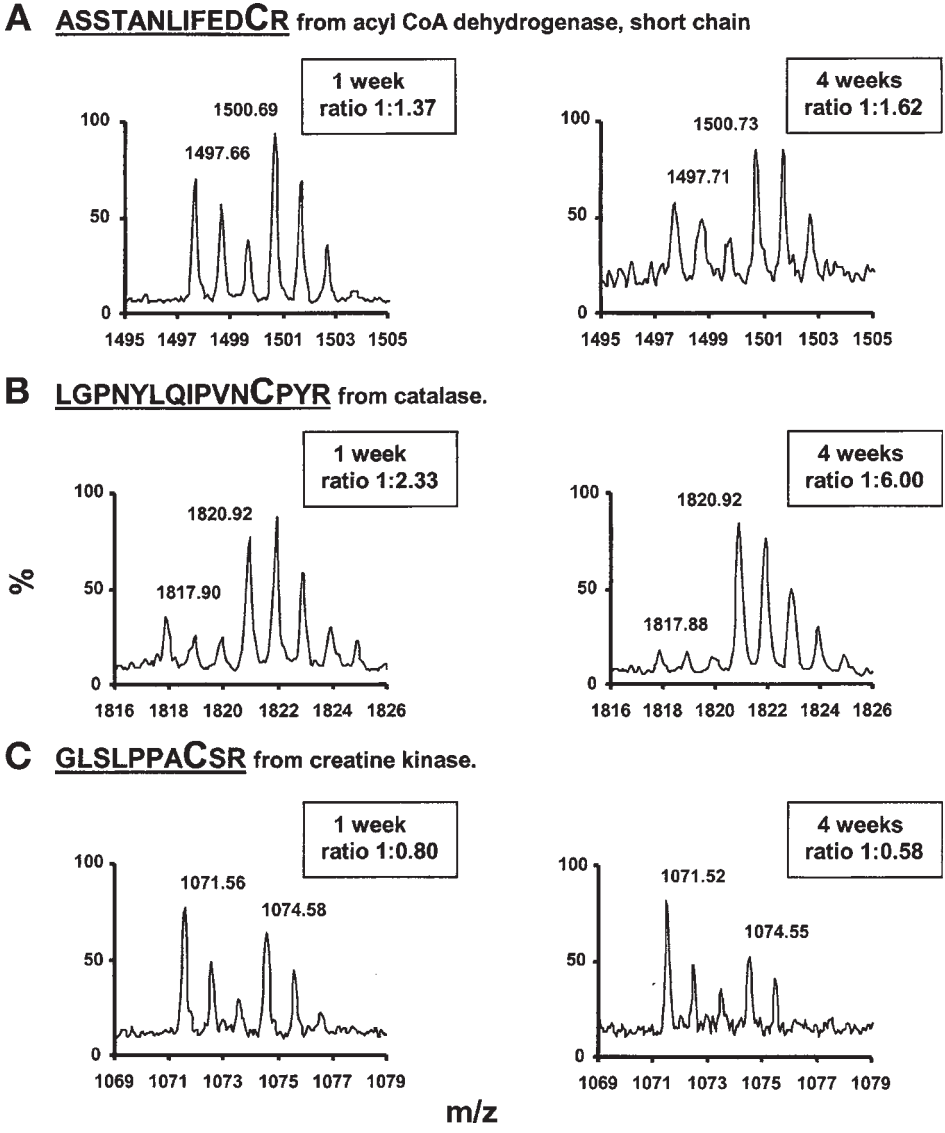


Fig. 5. Changed protein abundance after one of 4 wk of diabetes. Soluble mitochondrial proteins from normal and diabetic rat heart were treated with acrylamide and D_3-acrylamide, respectively, and mixed at a 1:1 ratio. Representative isotopic envelopes for selected cysteine-containing peptides are shown for 1- and 4-wk diabetic samples. Reproduced with permission from **ref.** ***12***.

2. For removing the Coomassie from the gel pieces, wash several times with 25 m*M* NH_4HCO_3/50% CH_3CN. Every 10–15 min or when the solution is blue, centrifuge, discard supernatant, and add more solution to the gel pieces (*see* **Note 9**).
3. After the gel pieces are destained, dry them by putting the tubes in a speed-vac apparatus (Savant, Holbrook, NY) for 10–20 min.
4. Add 5 µL of trypsin solution to the gel pieces and then add sufficient 25 m*M* NH_4HCO_3 to rehydrate the gel pieces (*see* **Note 10**).
5. Incubate at 37°C overnight (~10–16 h).
6. To extract the peptides, add 20–100 µL of 50% CH_3CN/0.1% FC_3COOH in each tube (the extraction solution should cover the gel pieces), incubate for 10 min, and occasionally vortex or use a mixer. Using a gel-loading tip, transfer the liquid containing the extracted peptides to a new siliconized tube. The extraction can be repeated a second time and the gel pieces dehydrate by using 10–50 µL of CH_3CN. If more than one extraction is made, combine the solution of the first extraction with the second extraction and add the CH_3CN wash.
7. Dry the solution containing the peptide extract using a speed-vac. These samples are stable for several months at –20°C.

3.4. MALDI-MS

1. Dissolve a dry peptide sample in 2 µL of 50% CH_3CN/0.1% FC_3COOH. Mix 0.5 µL of the matrix solution with 0.5 µL of the peptide solution on the sample target (and let it air-dry [*see* **Note 11**]).
2. Insert sample target into the mass spectrometer and analyze the sample as recommended by the manufacturer.
3. Spectra shown in **Figs. 2**, **3**, and **5** were acquired using a Voyager-DE STR mass spectrometer (Applied Biosystems, Framingham, MA) operated in the delayed extraction and reflector mode (*see* **Note 12**).

3.5. Protein Identification and Quantification

1. All MALDI mass spectra should be internally calibrated using the monoisotopic masses of the autolysis peptides of trypsin (e.g., *m/z* at 842.51 and 2211.10) (*see* **Note 13**).
2. To identify proteins from peptide mass fingerprinting data, use MASCOT software (www.matrixscience.com) or Profound (http://prowl.rockefeller.edu) (*see* **Note 14**) assuming that all cysteines were modified with acrylamide. A second search can be done assuming that all proteins are modified with D_3-acrylamide.
3. To achieve quantification, the manual analysis is required because at this stage the software we used could not recognize the distribution of cysteine-containing peptides when labeled with a mixture of acrylamide and D_3-acrylamides. Divide the peak height of the monoisotopic peak of the D_3-acrylamide-labeled peptide by the peak height of the monoisotopic peak of the acrylamide-labeled form of the peptide (*see* **Note 15**).

4. Notes

1. It might be difficult to properly weigh milligram quantities in a chemical hood with good ventilation; thus, it is advisable to put the acrylamide crystals in a closed vial and determine the weight outside of the hood. The appropriate volume of buffer can then be added.
2. Do not use serological plastic pipets to handle chloroform because they will "melt." Borosilicate glass pipets and tubes, as well as polypropylene tubes, are chloroform resistant.
3. Some batches of siliconized tubes were not well washed by the manufacturer and still had some residues of reagents used in the siliconization process that lead to artifacts in the mass spectra. Depending on the tube used and the batch, it might be useful to wash the tubes with 50% CH_3CN/0.1% FC_3COOH prior to use.
4. The MALDI matrix solution should be prepared on the same day that it is used. In some cases we found that the quality of the matrix was not sufficiently good for MS. In these cases, it is possible to further purify α-cyano-4-hydroxycinnamic acid by recrystallization in 40% ethanol. However, often it is possible to find from Sigma or other vendors α-cyano-4-hydroxycinnamic acid of MS quality. The amount of F_3CCOOH and CH_3CN can be varied and optimized for specific molecular weight range.
5. A simple way to have the appropriate final concentration of buffer is to use 2X reducing sample buffer and mix it 1:1 with the sample to be analyzed.
   When the protein concentration is in a range of 1–2 mg/mL, the amount of DTT present in the reducing sample buffer is usually sufficient for completely reducing all cystines. However, for more concentrated protein samples, the amount of DTT should be increased to warrant a complete reduction. A 50-*M* excess of DTT over cysteine residues is sufficient for ensuring complete reduction. It is also important to note that this is assuming that the proteins are completely denatured and that the final pH is appropriate. In some cases, the sample might modify the pH of the sample buffer. For example, a very acidic sample might bring the pH to 7.0. In these cases, a more concentrated sample buffer with stronger buffering capacity might be used (e.g., 4X sample buffer). From our experience, the reduction and alkylation is optimal at pH 8.0–9.0. An easy and approximate way to check the pH of the sample is to use 1 μL of sample on pH paper.
6. It seems that the alkylation reaction in the buffer used for 2D-PAGE proceeds slower than in 1D-PAGE and that longer incubation time might be needed when preparing samples for 2D-PAGE analysis. Although acrylamide, in the conditions described in **Subheadings 3.1.** and **3.2.**, has no strong tendency to react with other amino acids, it is advisable to not overextend the alkylation reaction time. In some cases, no major difference was observed when extending the reaction to overnight, but in this case it would be appropriate to verify that indeed no other amino acids besides cysteines are getting modified by acrylamide.

7. Immobilized pH gradient strips after isoelectrofocusing can be stored for several months at –80°C before the second dimension.
8. The small gel particle size facilitates the removal of SDS and Coomassie, and improves trypsin access to the gel.
9. At this point, the visual criterion should be used to determine whether or not additional washes should be performed. In some cases, washing with 40% ethanol might also help for completely removing the Coomassie stain. However, the major interfering compound is SDS, not Coomassie. Usually three washing steps are enough to remove SDS. Although some gel pieces may remain slightly blue, it is all right to proceed further.
10. Do not use too much trypsin or the mass spectra will be dominated by trypsin autolysis products. Usually 5 μL of trypsin solution contains a sufficient amount of trypsin (i.e., 0.025 μg) for proteolyzing most samples isolated by 1D- or 2D-PAGE. However, for very concentrated or very large spots, a higher amount of trypsin might be used.
11. Load the matrix solution first. If the sample has a high concentration of peptides a sample more diluted with matrix could be prepared (e.g., 0.5 μL sample + 2–5 μL of matrix solution).
12. Often 100 laser shots are sufficient for obtaining good quality spectra. However, for obtaining higher-quality spectra, more spectra can be collected and averaged. For example, in the spectra shown in **Figs. 2** and **3**, five spectra each generated with 100 laser shots were averaged.
13. These two peptides (842.51 and 2211.10) usually give the most intense ions in the spectra amongst the autolysis trypsin products. In addition, the *m/z* 2284.6 from trypsin could be used for internal calibration if necessary. The autolysis products differ depending on the trypsin manufacturer and the digestion conditions. It is therefore important to do the digest also on a piece of gel not containing any protein for confirmation of the peptides that are derived from trypsin.
14. In several cases it might be useful to identify a protein using more than one software platform. If an instrument with MS–MS capability is available, it might also be useful to further confirm uncertain identification by obtaining the MS/MS spectra of several peptides.
15. Depending from the mass of the peptide, there might be less or more overlap between the isotopic envelope of the peptide labeled with acrylamide and the envelope of the peptide labeled with D_3-acrylamide. If there is significant overlap, the relative quantity could be determined more accurately using the M+1 or M+2 isotopes. Ideally there should be no overlap between the isotopes used for the quantitation. However, several companies and investigators have been developing software for automatically deconvoluting and determining the areas under the isotopic envelopes, so better alternative for analyzing these spectra might be available soon. Such software would make this procedure even more accurate. It should also be noted that in most cases the relative quantity could be determined within a 20% error by just determining the peak heights.

References

1. Zhang, H., Yan, W., and Aebersold, R. (2004) Chemical probes and tandem mass spectrometry: a strategy for the quantitative analysis of proteomes and subproteomes. *Curr. Opin. Chem. Biol.* **8,** 66–75.
2. Sechi, S. and Oda, Y. (2003) Quantitative proteomics using mass spectrometry. *Curr. Opin. Chem. Biol.* **7,** 70–77.
3. Oda, Y., Huang, K., Cross, F. R., Cowburn, D., and Chait, B. T. (1999) Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. USA* **96,** 6591–6596.
4. Yao, X., Afonso, C., and Fenselau, C. (2003) Dissection of proteolytic ^{18}O labeling: endoprotease-catalyzed ^{16}O-to-^{18}O exchange of truncated peptide substrates. *J. Proteome Res.* **2,** 147–152.
5. Mirgorodskaya, O. A., Kozmin, Y. P., Titov, M. I., Korner, R., Sonksen, C. P., and Roepstorff, P. (2000) Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)O-labeled internal standards. *Rapid Commun. Mass Spectrom.* **14,** 1226–1232.
6. Yao, X., Freas, A., Ramirez, J., Demirev, P. A., and Fenselau, C. (2001) Proteolytic ^{18}O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.* **73,** 2836–2842.
7. Sechi, S. and Chait, B. T. (1998) Modification of cysteine residue by alkylation. A tool in peptide mapping and protein identification. *Anal. Chem.* **70,** 5150–5158.
8. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17,** 994–999.
9. Sechi, S. (2002) A method to identify and simultaneously determine the relative quantities of proteins isolated by gel electrophoresis. *Rapid Commun. Mass Spectrom.* **16,** 1416–1424.
10. Gehanne, S., Cecconi, D., Carboni, L., Righetti, P. G., Domenici, E., and Hamdan M. (2002) Quantitative analysis of two-dimensional gel-separated proteins using isotopically marked alkylating agents and matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **16,** 1692–1698.
11. Cahill, M. A., Wozny, W., Schwall, G., et al. (2003) Analysis of relative isotopologue abundances for quantitative profiling of complex protein mixtures labeled with the acrylamide/D_3-acrylamide alkylation tag system. *Rapid Commun. Mass Spectrom.* **17,** 1283–1290.
12. Turko, I. V. and Murad, F. (2003) Quantitative protein profiling in heart mitochondria from diabetic rats. *J. Biol. Chem.* **278,** 35,844–35,849.
13. Turko, I. V., Li, L., Aulak, K. S., Stuehr, D. J., Chang, J.-Y., and Murad, F. (2003) Protein tyrosine nitration in the mitochondria from diabetic mouse heart. Implication to dysfunctional mitochondria in diabetes. *J. Biol. Chem.* **278,** 33,972–33,977.

14. Thompson, L., Turko, I., and Murad, F. (2005) Mass spectrometry-based relative quantification of human neutrophil peptides 1, 2, and 3 from biological samples. *Mol. Immunol.* **43,** 1485–1489.
15. Meetei, A. R., Sechi, S., Wallisch, M., et al. (2003) A multiprotein nuclear complex connects Fanconi anemia and Bloom syndrome. *Mol. Cell. Biol.* **23,** 3417–3426.

2

Using Stable Isotope Tagging and Mass Spectrometry to Characterize Protein Complexes and to Detect Changes in Their Composition

Jeffrey A. Ranish, Marjorie Brand, and Ruedi Aebersold

Summary

One of the primary goals of proteomics is the description of the composition, dynamics, and connections of the multiprotein modules that catalyze a wide range of biological functions in cells. Mass spectrometry (MS) has proven to be an extremely powerful tool for characterizing the composition of purified complexes. However, because MS is not a quantitative technique, the usefulness of the data is limited. For example, without quantitative measurements, it is difficult to detect dynamic changes in complex composition, and it can be difficult to distinguish *bona fide* complex components from nonspecifically copurifying proteins. In this chapter, we describe a strategy for characterizing the composition of protein complexes and their dynamic changes in composition by combining affinity purification approaches with stable isotope tagging and MS. The use of software tools for statistical analysis of the data is also described.

Key Words: Mass spectrometry; stable isotope tagging; ICAT reagents; quantification; protein complex; dynamics; affinity purification; SEQUEST; Peptide Prophet; Protein Prophet; ASAPratio.

1. Introduction

Mass spectrometric analysis of purified protein complexes is an extremely powerful tool for identification of protein components and their posttranslational modifications ***(1)***, and the development of rapid methods for protein complex purification such as tandem affinity tagging ***(2)***, has enabled the isolation of numerous complexes for mass spectrometric analysis ***(3–6)***. However, even with the best purification protocols, it is often difficult to purify a

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

complex to homogeneity. Without proper controls, this can lead to the identification of nonspecifically copurifying proteins. Furthermore, protein complexes are dynamic. Individual subunits have a wide range of affinities for the complex, and composition can change depending on the status of the cell. Because mass spectrometry (MS) is not an inherently quantitative technique, it is difficult to detect these changes.

The development of stable isotope tagging approaches permits quantification of the relative levels of proteins in two or more samples ***(7)***. Peptides derived from proteins, which are differentially labeled with stable isotopes, can be distinguished by a characteristic mass shift in the mass spectrometer. Importantly, the isotopically labeled peptides are virtually chemically identical and thus generate similar specific signal intensities in a mass spectrometer. Therefore, the relative levels of the isotopically labeled peptides can be determined by comparing the signal intensities of sibling peptides. Application of isotope tagging and MS to the analysis of protein complexes can guide the identification of *bona fide* complex components by comparing the relative abundances of peptides derived from a specific purification and a control purification in which the complex of interest is not enriched ***(8–11)*** **Fig. 1**. It can also be used to detect changes in the composition of complexes by comparing the relative abundances of peptides derived from complexes isolated from cells exposed to different growth conditions ***(9–12)***. Here, we describe methods for using stable isotope tagging and MS to characterize the composition of protein complexes and to detect changes in their composition. There are now several approaches available for isotope tagging, including isotope-coded affinity tagging reagents (ICAT) and related approaches, oxygen 18 (18O) labeling, and SILAC (*see* **ref.** ***8*** for review). In this chapter we concentrate on the ICAT-labeling technique because, like 18O labeling, it is a postisolation approach, and these approaches are the most general. In addition, they are compatible with labeling tissues.

2. Materials

2.1. Characterization of Protein Complexes

2.1.1. Preparation of Yeast Nuclear Extracts

1. YPD (for cell growth): 10 g yeast extract, 20 g peptone, and water to 950 mL; autoclave and add 50 mL 40% glucose before use.
2. YPD/S: 20 g yeast extract, 40 g peptone, 40 g glucose, 364 g sorbitol, and water to 2 L. Prepare fresh on the day extracts are being prepared.
3. YPD/S (4°C): 20 g yeast extract, 40 g peptone, 40 g glucose, 364 g sorbitol, and water to 2 L. Prepare fresh on the day extracts are being prepared. Store at 4°C.
4. 1 *M* sorbitol (4°C): 182 g sorbitol and water to 1 L. Prepare fresh. Store at 4°C.

5. Zymolyase (ICN Biomedicals, Irvine, CA): dissolve at 6 mg/mL in 50 m*M* Tris-HCl, pH 8.0 with 2X concentrated protease inhibitors. Incubate 10 min on ice before using. This material does not dissolve well, so keep in suspension as well as possible. Zymolyase is reportedly contaminated with proteases, so extra care is needed to wash spheroplasts.
6. 400 mL Buffer A: 18% polysucrose 400 (or Ficoll 400), 10 m*M* Tris acetate pH 7.5, 20 m*M* potassium acetate, 5 m*M* magnesium acetate, 1 m*M* ethylenediaminetetraacetic acid (EDTA). The polysucrose takes many hours to dissolve and is frequently stirred overnight. Add 0.5 m*M* spermidine, 0.15 m*M* spermine, 3 m*M* dithiothreitol (DTT), and protease inhibitors before use.
7. Buffer B: 100 m*M* Tris acetate, 50 m*M* potassium acetate, 10 m*M* magnesium sulfate, 20% glycerol, and 2 m*M* EDTA. Adjust pH to 7.9 with KOH. Store buffer at 4°C. 3 m*M* DTT and protease inhibitors are added before use.
8. Buffer C: 20 m*M* HEPES, 10 m*M* magnesium sulfate, 1 m*M* EGTA, and 20% glycerol. Adjust pH to 7.6 with KOH. Store buffer at 4°C. 3 m*M* DTT and protease inhibitors are added before use.
9. 1.5 L Buffer C + 75 m*M* ammonium sulfate: store buffer at 4°C. 3 m*M* DTT and protease inhibitors are added before use.
10. Protease inhibitors:
    a. 16 mg/mL Phenylmethyl Sulfonyl Fluonide (PMSF) 0.1 *M* 100X in ethanol. Store at –20°C.
    b. 32 mg/mL 100X benzamidine in water. Store at –20°C.
    c. 0.15 mg/mL leupeptin (500X) in ethanol. Store at –70°C for less than 6 mo.
    d. 0.28 mg/mL pepstatin (200X) in methanol. Store at –20°C.
    e. 5 mg/mL chymostatin (2500X) in dimethyl sulfoxide. Store at –20°C.

2.1.2. Immunopurification of Protein Complexes

1. Buffer IP: 20 m*M* HEPES, pH 7.6, 100 m*M* potassium acetate, 1 m*M* EDTA, 5 m*M* magnesium acetate and 0.05% NP-40. Add protease inhibitors before use.
2. Buffer EL: 20 m*M* HEPES, pH 7.6, 100 m*M* potassium acetate, 1 m*M* EDTA, 5 m*M* magnesium acetate 0.0025% NP-40.
3. FLAG-M2 agarose (Sigma-Aldrich, St. Louis, MO).
4. 3X FLAG peptide (Sigma-Aldrich).

2.1.3. Stable Isotope-Labeling With Cleavable ICAT Reagents

1. Microcon 10 concentrators (Millipore, Billerica, MA).
2. TE 8.3: 20 m*M* Tris-HCl, pH 8.3, and 1 m*M* EDTA.
3. Tributylphosphine (TBP) (Sigma-Aldrich): TBP is toxic, so work in the hood. Prepare a 0.2 *M* 5% solution in 1-propanol (high performance liquid chromatography [HPLC] grade) (*see* **Note 1**).
4. Cleavable ICAT reagents (Applied Biosystems, Foster City, CA). The reagents are light sensitive. Perform labeling reactions in tubes supplied by vendor.

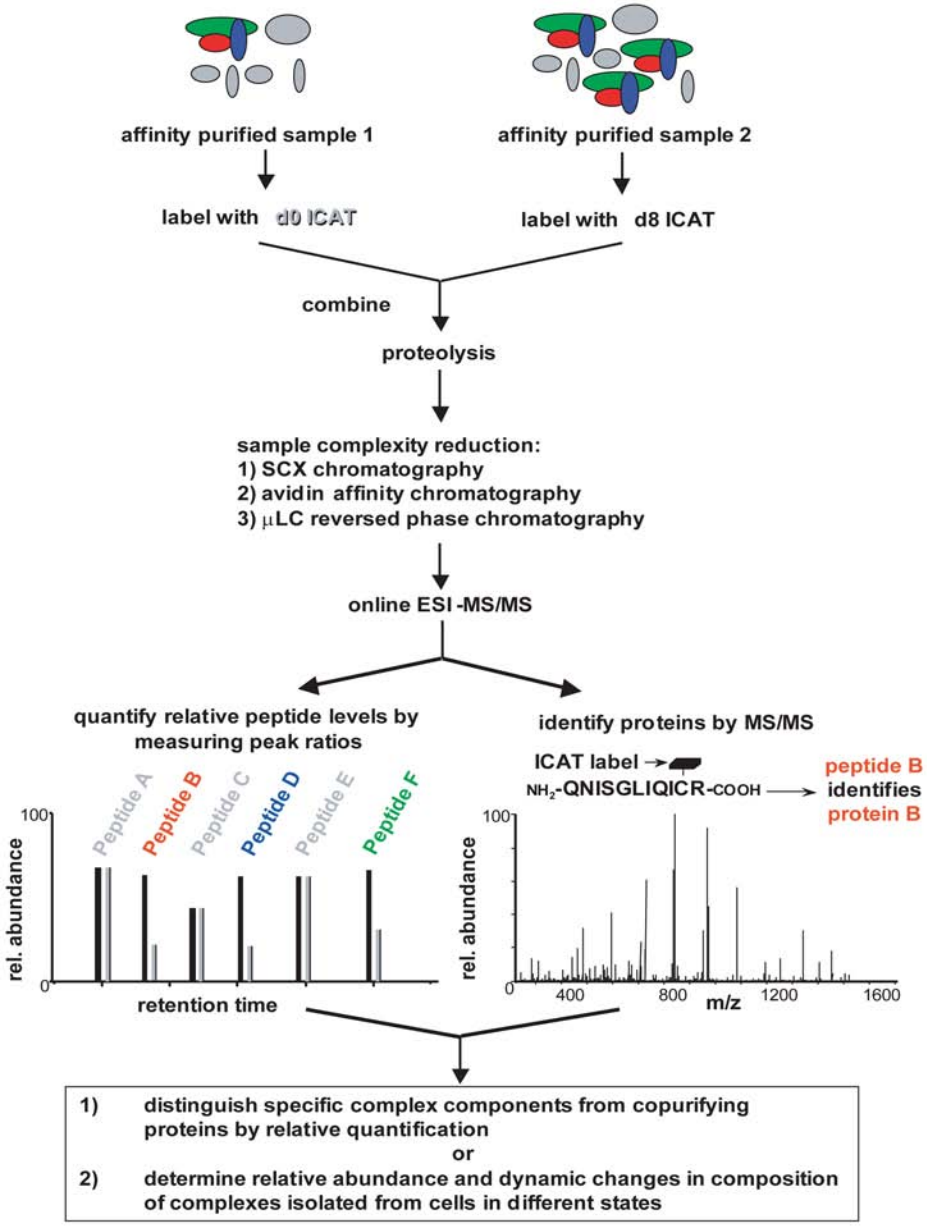


Fig. 1. Schematic representation of the quantitative proteomics approach for the analysis of affinity-purified macromolecular complexes. To distinguish specific complex components from copurifying proteins, a control purification (sample 1) is performed in which the complex of interest is not enriched. To detect quantitative changes in the abundance and composition of a complex isolated from cells in different states,

2.1.4. Proteolysis, Fractionation, and Purification of ICAT-Labeled Peptides

1. Endoproteinase LysC, sequencing grade (Roche Diagnostics Corporation, Indianapolis, IN). Resuspend in 0.4% acetic acid at 0.25 mg/mL.
2. Trypsin, sequencing grade modified (Promega Corporation, Madison, WI). Resuspend in buffer supplied by vendor at 1 mg/mL.

2.1.5. Strong Cation Exchange Fractionation of Peptides

1. Strong cation exchange (SCX) buffer: 5 m*M* KH_2PO_4, pH to 3.0 with 10% phosphoric acid, then add acetonitrile to 25% of final volume.
2. SCX buffer + 1 *M* KCl: 5 m*M* KH_2PO_4, 1 *M* KCl, pH to 3.0 with 10% phosphoric acid, then add acetonitrile to 25% of final volume.
3. 200-µL SCX cartridges (Applied Biosystems) (*see* **Note 2**).

2.1.6. Purifying ICAT-Labeled Peptides

1. 10X phosphate buffered saline (PBS) (Fisher Scientific International Inc., Hampton, NH): 1.37 *M* sodium chloride, 0.027 *M* potassium chloride, and 0.119 *M* phosphate buffer. Dilute 10X PBS one-fifth and one-tenth for 2X and 1X solutions, respectively.
2. Monomeric avidin cartridge (Applied Biosystems) (*see* **Note 3**).
3. Avidin elution buffer: 30% acetonitrile and 0.4% trifluoroacetic acid (TFA). Prepare fresh.
4. Avidin wash buffer: 50 m*M* ammonium bicarbonate and 20% methanol. Prepare fresh.
5. Glass collection vials (Waters, Milford, MA; cat. no. WAT025054).

Fig. 1 *(continued from opposite page)* the samples are prepared identically. Affinity-purified proteins from specific and control purifications are reduced, labeled with either the isotopically heavy or normal version of the ICAT reagent, and combined. After proteolysis, sample complexity is reduced in three sequential chromatographic steps, followed by electrospray ionization-MS/MS analysis. During this process, peptides pairs are quantified by measuring their peak ratios as they coelute from the C18 column into the mass spectrometer. In every other scan, peptides are selected for fragmentation. The resulting MS/MS spectra are used to search sequence databases using SEQUEST to identify the peptides and, thus, the proteins from which they originated. Depending on the experiment, the relative quantification can be used to distinguish specific complex components from copurifying proteins, or to detect changes in the abundance and composition of complexes isolated from cells in different states. Reproduced with permission from **ref**. ***8***.

2.1.7. Removing Biotin From ICAT-Labeled Peptides

1. Cleaving reagent A (Applied Biosystems) contains concentrated TFA. TFA is toxic, so work in the hood.
2. Cleaving reagent B (Applied Biosystems) contains a scavenger that reduces side reactions during the cleaving reaction.

2.1.8. MS Analysis

1. 75 µ ID X 360 OD micron-fused silica capillary tubing (Polymicro Technologies, Cedar Hill, TX).
2. Pressure bomb (Mass Evolution, Houston, TX).
3. Magic C18 resin (Michrom Bioresources, Inc., Auburn, CA), 5 µ, 200 A.
4. HPLC buffer A: 0.4% acetic acid and 0.005% heptafluorobutyric acid.
5. HPLC buffer B: 100% acetic acid, 0.4% acetonitrile, and 0.005% heptafluorobutyric acid.
6. LCQ ion trap mass spectrometer with nanoelectrospray ionization source (Thermofinnigan, Waltham, MA).

2.1.9. Data Analysis

1. SEQUEST software (Thermofinnigan).
2. Protein Prophet, Peptide Prophet, Xpress, ASAPratio (http://sashimi.sourceforge.net/).

2.2. Characterizing Changes in Complex Composition

2.2.1. Preparation of Murine Erythroleukemia Cell Nuclear Extracts

1. Cell culture medium: RPMI 1640, 10% fetal bovine serum, 100 U/mL penicillin, 0.1 mg/mL streptomycin, and 2 m*M* glutamine.
2. Buffer A: 10 m*M* HEPES, pH 7.9, 1.5 m*M* $MgCl_2$, and 10 m*M* KCl. Store at 4°C. Add 0.5 m*M* DTT and protease inhibitors cocktail EDTA-free (Roche) before use.
3. Buffer B: 20 m*M* HEPES pH 7.9, 1.5 m*M* $MgCl_2$, 600 m*M* KCl, 25% glycerol, and 0.2 m*M* EDTA. Store at 4°C. Add 0.5 m*M* DTT and protease inhibitors cocktail EDTA-free (Roche) before use.
4. Buffer C: 20 m*M* HEPES pH 7.9, 5 m*M* $MgCl_2$, 100 m*M* KCl, and 20% glycerol. Store at 4°C. Add 0.3 m*M* DTT and protease inhibitors cocktail EDTA-free (Roche) before use.

2.2.2. Immunopurification of Protein Complexes

1. p18-Specific rabbit polyclonal antibodies (Santa Cruz, Santa Cruz, CA, sc-477).
2. Normal rabbit immunoglobulin G (IgG) (Santa Cruz, sc-2027).
3. Protein A-sepharose resin (Pharmacia, New York, NY).
4. Dimethylpimelimidate (DMP) (Sigma-Aldrich).
5. Crosslink buffer A: 3 *M* NaCl and 50 m*M* Na borate, pH 9.0.

6. Crosslink buffer B: 3 *M* NaCl and 200 m*M* Na borate, pH 9.0.
7. 0.2 *M* ethanolamine, pH 8.0.
8. IP buffer: 25 m*M* Tris-HCl, pH 7.9, 5 m*M* $MgCl_2$, 10% glycerol, and 0.1% NP-40. Store at 4°C. Add 0.3 m*M* DTT and protease inhibitors cocktail EDTA-free (Roche) before use.
9. Elution buffer: 6 *M* urea. Prepare fresh.

*2.2.3. Stable Isotope Labeling With ICAT Reagents (*see ***Subheading 2.1.3.****)*

1. Nanosep 3K centrifugal devices (Pall Corporation, New York, NY).

2.2.4. Preparation of Control Peptides

1. 1-cc MCX cartridges (Waters).

3. Methods

We describe methods for characterizing the composition of protein complexes and for detecting changes in the composition of complexes using stable isotope tagging and MS. In both approaches, single-step affinity purifications can be used for complex isolation. Owing to the high resolving power of liquid chromatography–mass spectrometry (LC–MS) techniques, the copurifying contaminants are not detrimental. In fact, they are actually useful for calibrating a common ratio. In addition, the potential for sample losses and for the dissociation of weakly interacting factors is minimized by using single-step affinity purifications. In the method for characterizing the composition of protein complexes, specific complex components are distinguished from nonspecific, copurifying proteins by comparing the relative abundance of isotopically labeled peptide pairs derived from affinity purification of the complex of interest and a control purification that is performed in parallel ***(8,9)***. In the example presented here, a previously uncharacterized yeast protein is tagged with a FLAG epitope at its' chromosomal locus, and extracts are prepared from this strain and from an untagged strain ***(10)***. Immobilized anti-FLAG antibodies are used to purify the FLAG-tagged protein along with associated proteins. In an attempt to preserve complex interactions, the antibody resin is washed with a buffer containing low concentrations of salt and detergent (100 m*M* potassium acetate, 0.05% NP-40). Proteins are eluted from the resin by competition with a triple FLAG-containing peptide and prepared for labeling with ICAT reagents by concentration and buffer exchange with filtration devices. After labeling with ICAT reagents, the protein mixtures are combined and digested with endoproteinase Lys-C and trypsin. Peptides are fractionated on strong cation exchange columns and ICAT-labeled peptides are then isolated by avidin affinity chromatography. Labeled peptides are analyzed by microcapillary

reversed-phase liquid chromatography (μLC), electrospray ionization, tandem mass spectrometry (MS/MS), and peptides are identified by sequence database searching using the search algorithm SEQUEST ***(13)***. The relative abundances of identified ICAT-labeled peptide pairs are determined from the ratio of the peptides' signal intensities using Xpress ***(14)*** or ASAPratio software ***(15)***. *Bona fide* components of the complex (or complexes) are identified by their increased abundance in the specific purification compared with the nonspecific purification. The significance of each peptide and protein identification is estimated using the software tools, Peptide Prophet ***(16)*** and Protein Prophet ***(17)***. ASAPratio provides a statistical assessment to help distinguish potential complex components with significant abundance changes from the population of nonspecific proteins.

In the method for detecting changes in complex composition, the transcription factor p18NF-E2/MafK is immunopurified from murine erythroleukemia (MEL) cell extracts, derived from either proliferating or differentiating cells, using an immobilized p18-specific antibody ***(12)***. After washing the antibody resin, the bound proteins are eluted by incubation with 6 *M* urea for 2 h at 37°C. Control immunopurifications are performed in parallel using normal rabbit IgG. Eluted proteins from the p18 immunopurifications are prepared for isotopic tagging and MS analysis as previously described. Proteins from the control purification are directly analyzed by MS without isotopic tagging (*see* **Note 4**). Protein identification and relative abundance ratios are determined using SEQUEST ***(13)*** and Xpress ***(14)***, or ASAPratio ***(15)***, respectively. Only proteins that are enriched in the p18-specific purifications and are not detected in the control fractions (as tested by MS and/or Western blot) are considered p18-interacting proteins. The significance of each peptide and protein identification was estimated using Peptide Prophet ***(16)*** and Protein Prophet ***(17)***.

3.1. Characterizing the Composition of Protein Complexes

3.1.1. Preparation of Yeast Nuclear Extracts

(http://www.fhcrc.org/labs/hahn/methods/biochem_meth/)

Day 1

1. The *Saccharomyces cerevisiae* strains JRY14 (TFB5-FLAG) and BWG1-7a (untagged) are grown in YPD medium to an OD_{600} of three at 30°C. For wild-type cells, approx 2.5 mL of a saturated overnight culture inoculated per liter at 5:30 pm gives A600 of approx three at 9:00 AM. Grow 3 L of cells.
2. Harvest cells in 1-L bottles (4500*g* for 10 min, i.e., 4000 rpm in a Beckman J6-HC centrifuge). Drain excess media as well as possible and weigh cells. Expected yield is 20–35 g cells. If cells are overgrown, zymolyase will work poorly in spheroplasting cells.

3. Resuspend cell pellets in 35 mL 50 m*M* Tris-HCl, pH 7.5 and 30 m*M* DTT. Usually this can be done by gently shaking the centrifuge bottles. Leave cells in 1-L bottles. Incubate at 30°C for 15 min.
4. Pellet cells (4500*g* for 8 min) and resuspend in 20 mL YPD/S. Add 2–3 mL 2 *M* sorbitol and an equal volume 6 mg/mL zymolyase solution. The amount required can vary from approx 12–18 mg depending on the yeast strain. Incubate at 30°C with occasional gentle mixing (*see* **Note 5**).
5. Check progress of spheroplasting every 15 min. To check, mix 4 μL of cells with 4 μL 1% SDS on a glass slide. Observe the number of cell ghosts under the microscope. Incubate cells until about 80% spheroplasts are obtained. This can take anywhere from 30 min to 2.5 h. If cells are spheroplasting slowly after 1 h, an extra 1–2 mL of zymolyase can be added. However, if cells were overgrown (A600 >5), they may never spheroplast. Spheroplasting is also somewhat strain dependent.
6. After spheroplasting has reached about 80%, add 100 mL YPD/S (room temperature) and pellet cells (4500*g* for 12 min).
7. Resuspend cells in 250 mL YPD/S (room temperature) and incubate at 30°C for 30 min to allow cells to recover. The resuspension of spheroplasts works best if a small volume (~50 mL) of YPD/S is first added and cells are resuspended using a baking spatula. Then add the remaining YPD/S.
8. Pellet cells (4500*g* for 12 min) and resuspend in 200 mL cold YPD/S (4°C). Resuspend as in **step 7**. Keep everything cold from this point on. Cells can be kept on ice for an hour or so if other cells are still spheroplasting.
9. Repeat **step 8**.
10. Pellet cells (4500*g* for 12 min) and resuspend in 250-mL cold 1 *M* sorbitol.
11. Pellet cells (4500*g* for 12 min) and drain sorbitol media as well as possible (careful—sometimes the spheroplast pellet is not very tight). Resuspend in 100-mL buffer A at 4°C.
12. Dounce the spheroplasts three times using a B-type pestle. Transfer dounced cells to 250-mL centrifuge bottles.
13. Spin at 4100*g* for 8 min (5000 rpm in a GSA rotor). Transfer supernatant to new centrifuge bottles. Do not worry about the slimy loose pellet that also transfers. Repeat.
14. Spin supernatant at 4100*g* for 5 min. Transfer supernatant to a new centrifuge bottle. Repeat. By the last (fourth) spin, the slimy nonpelleted material should be nearly gone and the pellets firm.
15. Transfer supernatant to 50-mL centrifuge tubes and pellet crude nuclei. Spin 20,200*g* for 30 min (13,000 rpm in an SS34 rotor). Remove supernatant by decanting and remove the remaining supernatant by inverting the tubes.
16. Resuspend crude nuclear pellets with a small spatula in 10-mL buffer B and transfer to 50-mL screwcap tubes. The prep can be stopped at this point. Quick freeze and store resuspended nuclear pellets at –70°C.

Day 2

1. Thaw nuclei on ice and measure volume. Add 3 *M* ammonium sulfate (pH 7.5) to

0.5 *M* final concentration (1/5 original volume of nuclei), and immediately mix and incubate on roller in cold room for 30 min. After 10 min, break up any lumps with a glass rod. This step lyses nuclei.
2. Transfer to SW28 thick-walled ultracentrifuge tubes and spin at 141,000*g* for 90 min at 4°C in an ultracentrifuge.
3. Carefully remove supernatant with a 5-mL pipet (and Pasteur pipet if necessary) being careful to avoid the pellet. Do not worry about the white floating material. Transfer to a 50-mL screwcap tube.
4. Add 0.35 g solid ammonium sulfate per milliliter supernatant and immediately incubate on cold room roller for 30 min. The ammonium sulfate can be added all at once if a number of preps are being done. However, it is best if ammonium sulfate is added slowly while stirring supernatant in a beaker. The pH should remain greater than 7.0 (it almost always does) but should be checked. Adjust pH with 1 *M* NaOH, if necessary.
5. Transfer to thick-walled ultracentrifuge tubes and spin in SW28 at 18,000*g* for 20 min at 4°C. Remove supernatant by dumping and respin pellets at 18,000*g* for 4 min. Carefully remove all remaining supernatant with a Pasteur pipet.
6. Resuspend pellets in buffer C containing DTT and protease inhibitors. Depending on protein pellet size, resuspend in 0.4–1.5 mL buffer. This can be done with a small dounce homogenizer or a blue pipet tip depending on the amount of protein. Extracts can be frozen on dry ice and stored at –70°C at this point.
7. Dialyze nuclear extracts against 500 mL buffer C + 75 m*M* ammonium sulfate at 4°C. Exchange buffer after 2 and 4 h.
8. Aliquot extract and store at –70°C.
9. Measure the protein concentration using a Bradford assay. It may be difficult to get reproducible measurements of protein concentration using the Bio-Rad assay. This modified method works well. Dilute extract one-quarter in 0.1% SDS. Add 1–2 μL of diluted extract to 0.8 mL water in a 13 × 100-mm disposable test tube. Add 1 μL 0.1% SDS to protein standards. Add 0.2 mL dye reagent. After 10 min, read absorbance at A595. Extracts should be 25–50 mg/mL in protein. Protocol available online at http://www.Fhcrc.org_labs_hahn_methods_biochem_meth_.

3.1.2. Immunopurification of Protein Complexes

1. 12.5-mg nuclear extract from JRY14 (TFB5-FLAG) and BWG1-7a is diluted to 5 mL in buffer IP, incubated at 22°C for 10 min, and then centrifuged at 3000*g* for 2 min. The supernatant is retained.
2. 2 mL of a 50% slurry of FLAG-M2 agarose beads is prepared by washing with 20 mL buffer IP in a 10-mL Bio-Rad (Hercules, CA) column, followed by washing with approx 4 mL 0.1 *M* glycine pH 3.5 over 15 min. Next, the beads are equilibrated with 10 mL buffer IP. Be careful not to overexpose the beads to acidic glycine.
3. Transfer 1 mL of the 50% slurry of FLAG-M2 beads to a 15-mL tube. Gently pellet the beads by centrifuging at 1000*g* for 5 min. Remove the supernatant, and add the clarified extracts. Extracts are incubated with the beads for 2 h at 4°C

with gentle agitation, after which the beads are pelleted at 1000*g* for 5 min and the supernatants are removed. Save the supernatants for monitoring binding efficiency.

4. The beads are washed twice with 12 mL buffer IP, and then twice with 12 mL buffer EL, by incubating them for 5 min at 4°C with gentle agitation.
5. Beads are transferred to mini spin columns (Bio-Rad), the supernatant is discarded, and proteins are eluted by incubating beads in 0.5 mL buffer EL containing 0.1 mg/mL 3X FLAG peptide for 30 min at 22°C. The supernatant is collected and the elution step is repeated. The beads are washed with 0.5 mL buffer EL and the wash is combined with the previous elutions. Save approximately one-fortieth of the eluates for analysis by Western blotting and silver-stained sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

3.1.3. Stable Isotope Labeling With ICAT Reagents

1. Eluted protein samples (each containing approx 30 μg of total protein) are concentrated to approx 50 μL in Microcon 10 devices by centrifuging at 3000*g* in the cold room. The buffer is exchanged by addition of 500 μL of TE 8.3 containing 50 m*M* NaCl, and the volume is reduced to approx 25 μL. Save an aliquot of the concentrated samples for analysis by SDS-PAGE (*see* **Note 6**).
2. SDS is added to 0.3%, and the samples are boiled for 5 min.
3. Proteins are reduced with 5 m*M* TBP at 37°C for 30 min, and then diluted with 125 μL TE 8.3 containing 7.2 *M* urea.
4. Isotopically heavy or light ICAT reagents are added to 1.5 m*M* (*see* **Note 7**). There are 175 nmol of reagent per tube. Briefly spin two tubes of isotopically heavy and two tubes of isotopically light ICAT reagent in a microcentrifuge to bring the reagent to the bottom of the tubes. Resuspend one tube of heavy and one tube of the light reagent in 17.5 μL methanol (10 nmol/μL). Add each protein solution to one tube of heavy or light ICAT reagent. To reach 1.5 m*M* ICAT concentration, add an additional 5 μL of the appropriate reagent to each tube. Vortex the samples thoroughly to resuspend the ICAT reagent, and incubate for 90 min at 22°C.
5. Reactions are quenched by addition of 10 m*M* β-mercaptoethanol or DTT for 20 min at 37°C. Optional: save an aliquot (~1/40) of each sample to monitor labeling efficiency by SDS-PAGE analysis.

3.1.4. Proteolysis and SDS-PAGE Analysis

1. Samples are combined, and proteins are digested by addition of endoproteinase Lys-C (1:100 w/w) at 37°C for 3 h. SDS and urea concentrations are reduced to 0.01% and 1.2 *M*, respectively, by addition of TE 8.3, and samples are digested with trypsin (1:20 w/w) overnight at 37°C. Optional: save an aliquot (~1/80) to monitor digestion efficiency by SDS-PAGE analysis.
2. Before performing SCX fractionation, it is advisable to analyze the samples by SDS-PAGE and silver staining to monitor digestion efficiency. Analyze an aliquot of the starting samples, the ICAT-labeled samples, and the combined

digested sample. The lane containing the digested sample should contain very little full-length proteins. Trypsin migrates at 24 kDa and it may be visible. If the sample is not completely digested, add more trypsin (1:20 w/w), and incubate at 37°C for 3 h.

3.1.5. SCX Fractionation of Peptides

This step removes SDS and trypsin, and permits reduction of sample complexity.

1. Peptides are diluted with an equal volume of SCX buffer and the pH is adjusted to 3.0 with 10% trifluoroacetic acid (TFA). 10% TFA is added in 1–5 μL increments and the pH is monitored by spotting 1 μL of sample onto pH paper.
2. SCX cartridges are prepared by washing with 3 mL SCX buffer containing 1 *M* KCl, and equilibrated with 3 mL SCX buffer (*see* **Note 2**). To wash SCX cartridges, fill a 1–2.5 mL Hamilton syringe with the appropriate solution, remove air bubbles, insert the syringe needle into the needle port adapter, and slowly depress the plunger. For washing and equilibrating, inject the solution so that two to three drops per second flow from the outlet.
3. Peptides are slowly loaded onto equilibrated SCX cartridges (~1 drop/s). Save the flow through. Peptides are successively eluted with 0.75 mL SCX buffer containing 40, 200, 350, and 600 m*M* KCl into Eppendorf tubes (*see* **Note 8**).
4. To clean the SCX cartridge, wash with 2 mL SCX buffer containing 1 *M* KCl. To store the column, wash with 2 mL SCX buffer. For long-term storage include 0.1% sodium azide in the wash. Cartridges can be reused approx 20 times depending on the complexity of the samples.

3.1.6. Purifying ICAT-Labeled Peptides

1. Reduce the acetonitrile concentration by drying the samples under reduced pressure to approx 350 μL, and dilute the samples twofold with 2X PBS (pH 7.2). Adjust the pH to approx 7.0 with approx 10 μL 1 *M* ammonium bicarbonate.
2. Prepare a monomeric avidin cartridge by washing with 2 mL avidin elution buffer, followed by 2 mL 2X PBS (pH 7.2) (~2 drops/s) (*see* **Note 3**).
3. Slowly load peptides onto avidin cartridges (~1 drop/s). Save the flow through.
4. Wash the cartridge with 2 mL 2X PBS (pH 7.2), followed by 1 mL 1X PBS (pH 7.2) to reduce the salt concentration.
5. Wash the cartridge with 1 mL avidin wash buffer. This step removes nonspecifically bound peptides.
6. Wash with 1 mL Milli-Q water.
7. Peptides are eluted by slowly injecting 800 μL avidin elution buffer (1 drop/s). Allow the first 50 μL to go to waste. Collect the remaining 750 μL in a glass vial.
8. To purify additional samples, repeat the procedure beginning at **step 2**.
9. To store the cartridge, wash with 2 mL avidin elution buffer, followed by 2 mL 2X PBS (pH 7.2). Include 0.1% sodium azide for long-term storage.

3.1.7. Removing Biotin From ICAT-Labeled Peptides

1. Purified peptides are dried under reduced pressure.
2. Prepare the final cleaving reagent by combining cleaving reagent A and cleaving reagent B in a 95:5 ratio. Approximately 90 µL of final cleaving reagent are needed for each fraction.
3. Vortex to mix, then centrifuge for a few seconds to bring the solution to the bottom of the tube.
4. Add approx 90 µL of freshly prepared cleaving reagent to each sample tube.
5. Vortex to mix, then centrifuge for a few seconds to bring the solution to the bottom of the tube.
6. Incubate for 2 h at 37°C.
7. Centrifuge the tube for a few seconds to bring the solution to the bottom of the tube.
8. Evaporate the sample to dryness in a centrifugal vacuum concentrator (~30–60 min).

3.1.8. MS Analysis

1. Peptides are resuspended in 10% acetonitrile and 0.1% TFA.
2. Peptides are pressure loaded onto in-house prepared 10 cm × 75-µm fused silica microcapillary columns packed with 5 µ Magic C18 beads and equilibrated with HPLC buffer A containing 10% acetonitrile using a pressure cell at 1000 psi helium (*see* **Note 9**).
3. Columns are placed in-line with a mass spectrometer that is capable of acquiring tandem mass spectra. Peptides are resolved by running 80 min gradients from 10–40% HPLC buffer B at 0.3 µL/min, and analyzed by automated data-dependent MS/MS. The mass spectrometer is set to scan from 400–1800 *m/z* followed by one data-dependent MS/MS scan on the most abundant ion. Dynamic exclusion is set to exclude ions that have been selected for MS/MS analysis for 2 min with a mass window of 2 Da.

3.1.9. Data Analysis

1. Peptides are identified by searching MS/MS spectra against an appropriate database using the SEQUEST algorithm. For ion trap data, peptide mass tolerance is set at 3 Da, and average masses are used for the precursor masses and for fragment masses. The mass of cysteine is statically modified by 227.13 Da, which accounts for the mass of the isotopically normal form of ICAT after the cleavage reaction that is added to the peptides. In addition, the mass of cysteine is differentially modified by 9.03 Da to account for peptides modified with the isotopically heavy form of the ICAT reagent. Methionine is also differentially modified with 16 Da to account for oxidized methionine residues.

2. Next, the search results are analyzed with the programs Peptide Prophet and Protein Prophet ***(16,17)***. These programs estimate the probability of each peptide and protein identification using a statistical model that is based on a number of criteria, including Sequest search scores and the number of tryptic termini of peptides. We typically select proteins with probability values ≥ 0.9 for further analysis.
3. The relative abundance ratios of correctly identified ICAT-labeled peptide pairs are determined using XPRESS ***(14)*** or ASAPratio ***(15)*** programs. Ratios are calculated by comparing the summed ion intensity of each peptide over its elution peak. To adjust for any systematic error from sample handling, ASAPratio normalizes the abundance ratios. This is done by generating a ratio distribution from the logarithm (base 10) of all peptide ratios, fitting it with a normal distribution, and setting the most common ratio to 1.
4. Specific complex components are distinguished from nonspecifically, copurifying proteins by inspection of the distribution of all abundance ratios. ASAPratio calculates a *p*-value for each protein from the distribution that can be used to distinguish true interactors from the background of copurifying proteins (*see* **Note 10**).

3.2. Characterizing Changes in Complex Composition

3.2.1. Preparation of Mouse MEL Cell Nuclear Extracts

1. MEL cells are grown in spinner flasks in RPMI cell culture medium up to a concentration of 1.5×10^6 cell/mL, in the absence (proliferating) or presence (differentiated state) of 2% dimethyl sulfoxide. Grow 8 L of each differentiation state.
2. Nuclear extracts from proliferating and differentiating cells are prepared separately.
3. Cells are harvested in 500-mL conical tubes (500*g* for 10 min at 4°C, i.e., 1500 rpm in a Beckman Allegra X-15R, rotor SX 4750), and washed with five packed cell volumes (PCVs) of ice-cold PBS buffer.
4. Cells are resuspended in five PCVs of ice-cold buffer A, and incubated on ice for 10 min to allow swelling.
5. Cells are pelleted (500*g* for 10 min), resuspended in two PCVs of ice-cold buffer A, and lysed on ice by 10 dounces using a B-type pestle (Kimble/Kontes, Vineland, NJ). Cell lysis is checked under a microscope using Trypan blue.
6. Nuclei are pelleted at 25,000*g* for 30 min at 4°C (Beckman Coulter, Inc., Fullerton, CA; rotor JA25.50) and the supernatants are discarded.
7. Nuclei are resuspended in one nuclear pellet volume of ice-cold buffer B, and nuclear proteins are extracted on ice by 10 dounces using a B-type pestle (Kimble/Kontes), followed by a 30-min incubation on a roller in the cold room.
8. Nuclear extracts are recovered after centrifugation at 25,000*g* for 30 min at 4°C (Beckman Coulter, Inc.; rotor JA25.50), and dialyzed twice against 50 vol of buffer C for 2 and 12 h, respectively.
9. After dialysis, nuclear extracts are further centrifuged (17,000*g* for 15 min at 4°C) to eliminate proteins that precipitated during dialysis. The supernatants are then recovered and NP-40 is added up to 0.1% final volume.

10. Nuclear extracts are aliquoted, snap frozen in liquid nitrogen, and stored at –80°C.
11. The total protein concentration is measured using a Bradford assay, and should be 1–5 mg/mL for the differentiated extract and 5–10 mg/mL for the proliferating extract.

3.2.2. Immunopurification of Protein Complexes

3.2.2.1. Crosslink of p18-Specific Antibodies on Protein A Sepharose

1. The pH of the antibody solution (containing 1.6 mg of p18-specific antibodies) is adjusted to 9.0 with NaOH. NaCl is added to this solution up to a final concentration of 3 *M*. Check the pH again and eventually readjust to 9.0 with NaOH.
2. 1.6 mL of protein A-Sepharose resin is washed with 10 vol H_2O and equilibrated with 10 vol of IP buffer containing 100 m*M* KCl.
3. Mix the antibody solution with the protein A-Sepharose resin for 1 h at room temperature with rotation.
4. The beads are washed twice with 10 vol of crosslink buffer A and resuspended in 10 vol of crosslink buffer B. Save an aliquot for Coomassie-stained SDS-PAGE analysis of the crosslink.
5. Add dimethylpimelimidate (DMP) to bring the final concentration to 20 m*M* and mix for 30 min at room temperature with rotation. Save an aliquot for Coomassie-stained SDS-PAGE analysis of the crosslink.
6. The reaction is stopped by washing the beads once in 0.2 *M* ethanolamine and incubating in 0.2 *M* ethanolamine for 2 h at room temperature with rotation.
7. The antibody-bound beads are equilibrated with IP buffer containing 100 m*M* KCl by washing three times with 10 bed volumes.
8. The efficiency of the crosslink is verified by the absence of the antibody heavy chain in the samples after crosslink as analyzed on the Coomassie gel.

3.2.2.2. Crosslink of Normal Rabbit IgG on Protein A-Sepharose

1. Normal rabbit IgG (200 μg) is crosslinked to protein A-Sepharose (200 μL) resin with 20 m*M* final of dimethylpimelimidate, as described in **Subheading 3.2.2.1.**

3.2.2.3. Immunopurification Using p18-Specific Antibodies

1. 8 mL of nuclear extract from proliferating and differentiating cells (containing equal amounts of p18) are separately incubated with the p18-specific antibody-bound resin (0.8 mL each) at 4°C for 12 h with rotation. Beads are recovered by centrifugation at 500*g* for 5 min and the supernatants (unbound proteins) are kept for analysis by Western blot.
2. Antibody-bound proteins are washed twice with 10 bed volumes of ice-cold IP buffer containing 300 m*M* KCl, and equilibrated similarly with IP buffer containing 100 m*M* KCl. All washes are kept for analysis by Western blot.
3. Bound proteins are then eluted by incubation with one bed volume of preheated 6 *M* urea for 2 h at 37°C with rotation. Approximately one-fortieth of the eluates are saved for analysis by Western blot and silver-stained SDS-PAGE.

3.2.2.4. Control Immunopurification Using Normal Rabbit IgG

1. 1 mL of nuclear extract from proliferating and differentiating cells (containing equal amounts of p18) are separately incubated with the normal rabbit IgG resin (0.1 mL each) at 4°C for 12 h with rotation.
2. Samples are then treated as described in **Subheading 3.2.2.3.**, and proteins are eluted in 100 µL 6 *M* urea.

3.2.3. Stable Isotope Labeling With ICAT Reagents

1. Eluted proteins from the p18 purifications (each containing approx 3 µg of total protein) are concentrated to 25 µL each in Nanosep 3 K centrifugal devices (*see* **Note 6**).
2. The solutions are adjusted to 20 m*M* Tris-HCl, pH 8.3 and 1 m*M* EDTA. Keep an aliquot for analysis by silver-stained SDS-PAGE.
3. The proteins are denatured by adding SDS to 0.3% and heating to 37°C (*see* **Note 11**).
4. The samples are reduced by addition of TBP to 5 m*M* and incubation at 37°C for 30 min. Samples are then diluted with 125 µL TE 8.3 buffer containing 54 mg of urea in order to obtain final concentrations of 0.05% SDS and 6 *M* urea.
5. One tube of heavy and one tube of light ICAT reagent are briefly spun in a microfuge to pellet the reagent. Protein solutions are added separately to each tube of ICAT reagent and after vortexing thoroughly, the tubes are incubated with shaking for 90 min at 22°C.
6. Reactions are quenched by addition of 10 m*M* β-mercatopethanol or DTT for 20 min at 37°C. Optional: save an aliquot (~1/40) of each sample to monitor labeling efficiency by SDS-PAGE analysis. Please refer to **Subheadings 3.1.4.–3.1.6.** for the remainder of the protocol.

3.2.4. Preparation of Control Peptides

A mixed bed cation exchange column is used to remove small molecules, such as salts, SDS and urea, prior to MS.

1. The solutions containing eluted proteins from the control purifications are adjusted to 20 m*M* Tris-HCl, pH 8.3 and 1 m*M* EDTA. Keep an aliquot for analysis by silver-stained SDS-PAGE.
2. Proteins are reduced by addition of DTT to 2 m*M* and incubation at 37°C for 30 min.
3. Reduced cysteines are alkylated by addition of iodoacetamide to 10 m*M*, and samples are incubated at 22°C for 20 min in the dark.
4. Proteins are digested by addition of endoproteinase Lys-C (1:100 w/w) at 37°C for 3 h. The urea concentration is reduced to 1.2 *M*, by addition of TE 8.3, and samples are digested with trypsin (1:20 w/w) overnight at 37°C. (Optional: save an aliquot [~1/40] to monitor digestion efficiency by SDS-PAGE analysis.)
5. The pH of the samples is adjusted to approx 3.0 by addition of 10% TFA in small aliquots (~1–5 µL) Check by spotting approx 1 µL to pH paper.

6. Prepare a 1-cc MCX mixed bed cation exchange cartridge by wetting with 2 mL methanol, followed by washing with 5 mL 80% acetonitrile, 0.1% TFA, 2 mL water, 2 mL 10% ammonium hydroxide, 90% methanol, and 2 mL water.
7. Equilibrate with 2 mL 0.1% TFA.
8. Load sample slowly to the cartridge.
9. Wash with 5 mL 80% acetonitrile and 0.1% TFA.
10. Wash with 2 mL water.
11. Elute with 1 mL 10% ammonium hydroxide and 90% methanol prepared fresh.
12. Dry samples in speed-vac (avoid drying acidic samples at the same time), and resuspend samples in 0.5 % acetonitrile and 0.1 % TFA. For information on MS analysis and data analysis, please *see* **Subheadings 3.1.8.** and **3.1.9.**, respectively.

Data was analyzed as described in **Subheading 3.1.9.**, and a histogram of the measured abundance ratios was plotted. A ratio of 1.2 was assessed as indicative of a significant enrichment either before or after differentiation. This value was chosen after examining the abundance ratios obtained for p18 and for the nonspecific proteins (the ones present in both the p18 fractions and the control fractions), which served as internal controls, and displayed an average ratio < 1.2 (*see* **Note 4**). It should be noted that owing to the variation of the signal-to-noise ratios between different peptides, in some cases the abundance ratios should only be regarded as tendency of enrichment and not as absolute enrichment values between different proteins.

4. Notes

1. Tris (2-carboxyethyl) phosphine hydrochloride (TCEP·HCl; Pierce, Rockford, IL) can be used instead of TBP. It is water soluble. A 0.2 *M* solution is prepared in 100 m*M* Tris-HCl, pH 8.3.
2. SCX microcolumns can be prepared in cases where there is a small amount of sample (i.e., ~10 μg). Western Analytical Products, Murrieta, CA sells 1 × 10-mm SCX cartridges.
3. When working with small amounts of sample (i.e., ~10 μg starting material) immobilized monomeric avidin (Pierce) can be used by packing it into a microcolumn from Western Analytical Products or into a Pasteur pipet blocked with glass wool.
4. With the use of iTRAQ reagents (Applied Biosystems), it is now possible to compare the relative abundance ratios of up to four samples in one MS experiment. This is very useful in experiments directed at detecting changes in complex composition because the proteins from control experiments can be isotopically tagged and analyzed along with the proteins from the specific purifications. Thus, in one experiment it is possible to accurately distinguish specific complex components from nonspecific proteins, and to detect dynamic changes in complex composition. Mass spectrometers that can detect fragment ions between 114 and 117 *m/z* are required for quantification of iTRAQ-labeled peptides.

5. Recombinant lyticase (Sigma-Aldrich) can be used instead of zymolyase. It is supposed to contain lower levels of proteases than zymolyase. When using lyticase, **step 9** of the extract prep in **Subheading 3.1.1.** can be omitted.
6. A fraction of the sample can be analyzed on a silver-stained SDS-polyacrylamide gel along with a titration of a standard protein. This is useful for assessing the quantity, quality, and complexity of the sample.
7. In cases where larger amounts of protein are being labeled, it is important to use enough ICAT reagent to label all of the cysteine residues in the sample. To estimate the total amount of cysteine in the sample, measure the protein concentration of the sample, calculate the number of moles of protein assuming an average molecular weight of 50 kDa, and multiply this number by 6 (approximately six cysteine residues per protein). ICAT can be added in a fivefold molar excess over the total moles of cysteine.
8. The number of fractions collected depends on the complexity and quantity of the sample.
9. Many laboratories use autosamplers (i.e., Famos autosampler) in place of pressure bombs for loading samples onto C18 columns.
10. When analyzing the data, it is important to consider the following factors: (1) in a typical experiment, the accuracy of quantification is around 20–25% and (2) for low abundance peptides or peptides with low ionization efficiency, the measured enrichment values may be affected by signal to noise issues.
11. If urea is used for eluting proteins, the sample should not be boiled before reducing and labeling. Urea can carbamylate proteins at high temperatures. Incubation at 37°C for 30 min is sufficient for denaturing the proteins.

Acknowledgments

This work was supported by NCI grant CA84698, NIH Research Resource Center grant RR11823, by Federal funds from the National Heart, Lung, and Blood Institute, NIH, under contract number NO1-HV-28179 to R.A., and by NIH postdoctoral fellowship GM19884 to J.A.R. Partial funding for this work came via a gift from Merck and Co. to the Institute for Systems Biology.

References

1. Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422,** 198–207.
2. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B., A. (1999) Generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17,** 1030–1032.
3. Gavin, A. C., Bosche, M., Krause, R., et al. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415,** 141–147.
4. Ho, Y., Gruhler, A., Heilbut, A., et al. (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* **415,** 180–183.

5. Krogan, N. J., Peng, W. T., Cagney, G., et al. (2004) High-definition macromolecular composition of yeast RNA-processing complexes. *Mol. Cell.* **13,** 225–239.
6. Butland, G., Peregrin-Alvarez, J. M., Li, J., et al. (2005) Interaction network containing conserved and essential protein complexes in Escherichia coli. *Nature* **433,** 531–537.
7. Julka, S. and Regnier, F. (2004) Quantification in proteomics through stable isotope coding: a review. *J. Proteome Res.* **3,** 350–363.
8. Ranish, J. A., Yi, E. C., Leslie, D. M., et al. (2003) The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* **33,** 349–355.
9. Blagoev, B., Kratchmarova, I., Ong, S. E., Nielsen, M., Foster, L. J., and Mann, M. A. (2003) Proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* **21,** 315–318.
10. Ranish, J. A., Hahn, S., Lu, Y., et al. (2004) Identification of TFB5, a new component of general transcription and DNA repair factor IIH. *Nat. Genet.* **36,** 707–713.
11. Flory, M. R., Carson, A. R., Muller, E. G., and Aebersold, R. (2004) An SMC-domain protein in fission yeast links telomeres to the meiotic centrosome. *Mol. Cell* **16,** 619–630.
12. Brand, M., Ranish, J. A., Kummer, N. T., et al. (2004) Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics. *Nat. Struct. Mol. Biol.* **11,** 73–80.
13. Eng, J. K., McCormack, A. L., and Yates, J. R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in protein databases. *J. Am. Soc. Mass Spectrom.* **5,** 976–989.
14. Han, D. K., Eng, J., Zhou, H., and Aebersold, R. (2001) Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.* **19,** 946–951.
15. Li, X. J., Zhang, H., Ranish, J. A., and Aebersold, R. (2003) Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. *Anal. Chem.* **75,** 6648–6657.
16. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74,** 5383–5392.
17. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75,** 4646–4658.

3

Stable Isotope Labeling by Amino Acids in Cell Culture for Quantitative Proteomics

Shao-En Ong and Matthias Mann

Summary

Mass spectrometry (MS)-based quantitative proteomics is an increasingly popular approach to study changes in protein abundances in biological samples. Stable isotope labeling by amino acids in cell culture (SILAC), one of the more widely used methods for quantitative proteomics, is a metabolic-labeling strategy that encodes whole cellular proteomes. Cells are grown in a culture medium where the natural form of an amino acid is replaced with a stable isotope form, such as arginine bearing six ^{13}C atoms. Incorporation of the "heavy" amino acid occurs through cell growth, protein synthesis, and turnover. SILAC allows "light" and "heavy" proteomes to be distinguished by MS while avoiding any chemical derivatization and associated purification. In this chapter, we provide detailed SILAC protocols and explain how to incorporate SILAC into any experiment.

Key Words: SILAC; stable isotope labeling; mass spectrometry; proteomics; gene expression.

1. Introduction

Several key advances in mass spectrometry (MS)-based proteomics over the past decade have provided entirely new perspectives in protein analysis and greatly accelerated the study of biological systems. These advances include the development of nano-flow liquid chromatography (LC) coupled to automated mass spectrometric analyses, faster and more sensitive mass spectrometers, and new methods for quantitation of protein abundances with MS ***(1)***.

MS is not inherently quantitative, as peptide ionization efficiency cannot be predicted. Therefore, more accurate quantitation in MS is facilitated through

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

the use of stable isotope-labeled standards. Stable isotope labels such as ^{13}C, ^{2}H, ^{15}N, and ^{18}O are incorporated in place of the natural abundance isotope in the "heavy" standard. This increases the mass of the labeled species, but keeps the chemical structure and hence other properties, as similar to the analyte of interest as possible (**Fig. 1**). This approach to quantitation is well established in small molecule quantitation in the pharmaceutical industry.

Stable isotope labeling by amino acids in cell culture (SILAC) is a metabolic-labeling strategy in quantitative proteomics to label whole proteomes. This is done through incorporation of stable isotope-labeled amino acids, like L-arginine containing six ^{13}C, through natural protein turnover and cell growth ***(2–4)***. Cells are cultured in two separate medium formulations, the "light" medium with amino acids of natural isotope abundance while the "heavy" medium contains the SILAC amino acid of choice (*see* **Note 1**). In contrast to radioactive labeling in biology, such as commonly used pulse-chase experiments with ^{32}P or ^{35}S-Met, SILAC seeks to replace the labeled amino acid within the proteome completely (*see* **Fig. 2** and **Note 2**). Furthermore, with radioactivity, detection with scintillation counters or film registers the signal from the radioactive material only; full incorporation is neither necessary to detect signal nor to make relative comparisons. SILAC depends on MS for readout, thus even a small percentage of unlabeled amino acid will be detectable as a population of unlabeled peptides. When proteins or peptides from unlabeled and labeled samples are analyzed by MS, they are separated by a residue-specific mass difference corresponding to the number of stable isotope labels (arg, lys), as well as the number of residues of the labeling amino acid incorporated (such as one or two args in a peptide). This is a significant advantage over the ^{15}N metabolic-labeling approach as it is possible to identify a "light" and "heavy" peptide pair by mass difference alone and without having to first identify the peptide. The SILAC method is simple and robust, and labels entire proteomes without chemical derivatization or sample handling steps. Labeled cells can be mixed at the stage of whole cells, making the SILAC approach ideally suited to quantitative proteomics experiments that study subproteomes like cellular organelles, as complex purification protocols can be performed without any detrimental effect on quantitative accuracy. The method was first demonstrated for the quantitation of protein abundance changes during a time course of myoblast differentiation in mouse C2C12 cells ***(2)***. It was extended to other formats, including global protein profiling ***(5,6)***, functional assays in protein–protein interaction ***(7)***, identification of proteins enriched in particular cellular structures ***(8,9)***, and multiplexed analyses with triple encoding SILAC to simultaneously compare three cellular states ***(10–12)***. SILAC has also been applied to study protein turnover ***(13,14)***, which is a

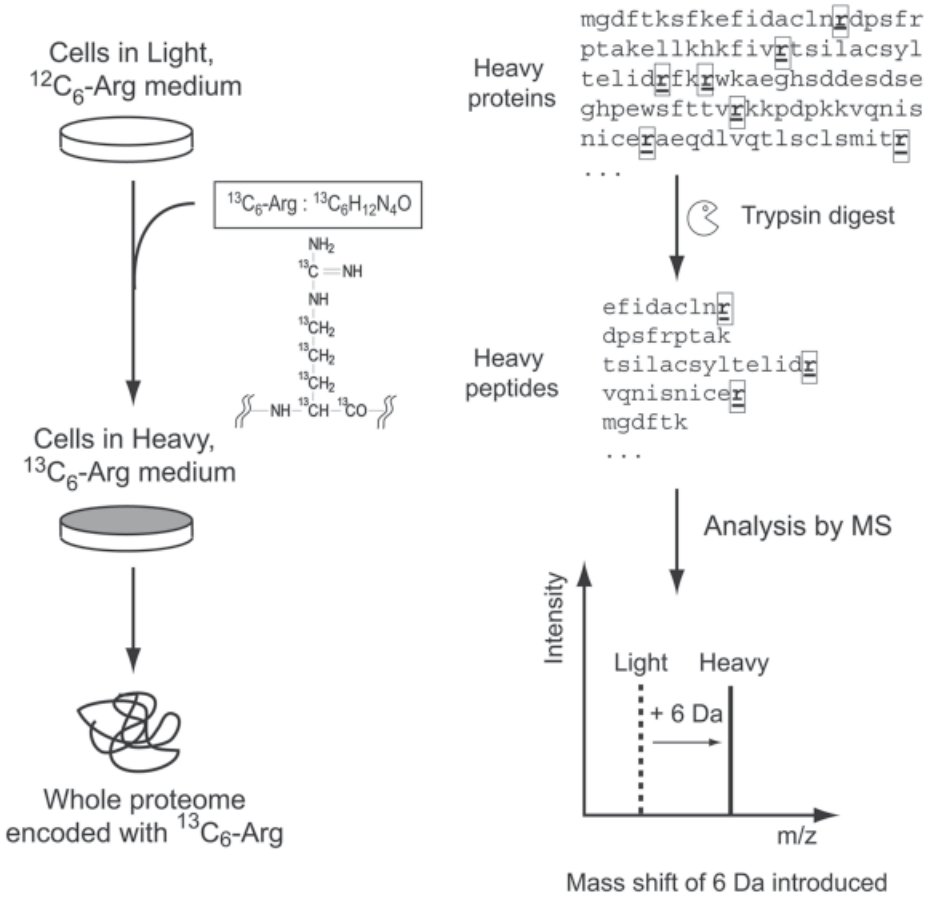


Fig. 1. Encoding whole proteomes using stable isotope labeling by amino acids in cell culture (SILAC). Cells growing in normal Dulbecco's modified Eagle's medium (DMEM) medium are SILAC labeled by culturing them in DMEM medium containing a $^{13}C_6$ stable isotope-labeled version of arginine in place of normal arginine (**left panel**). After five cell doublings, every instance of arginine in each protein in these cells are now "heavy," $^{13}C_6$-Arg. After subsequent digestion of the isolated proteins to peptides by trypsin, arg-containing peptides are now six Da heavier than their nonlabeled, "light" counterparts (multiple arg containing peptides will have n * six Da, where *n* is the number of args) (**right panel**).

unique capability of metabolic-labeling strategies. Furthermore, SILAC labeling with tyrosine ***(15)***, methionine ***(16)***, and complete labeling with arginine and lysine ***(17,18)*** has proven to be especially useful for the identification and quantitation of protein posttranslational modifications (*see* **Note 1**).

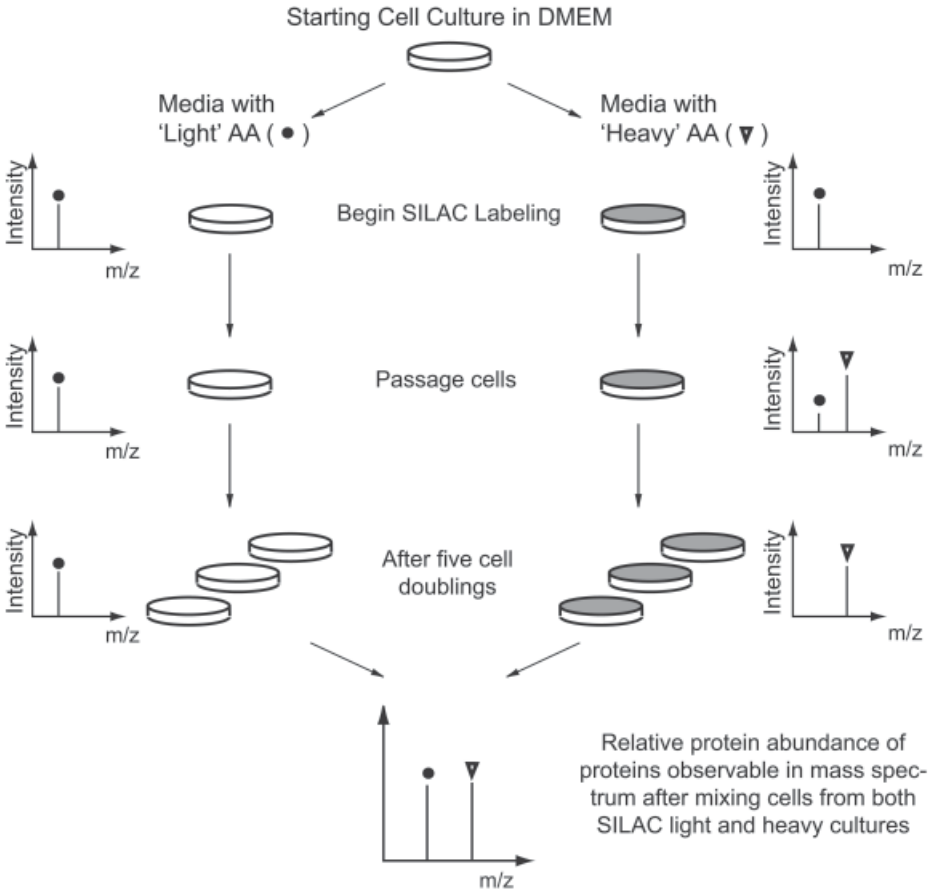


Fig. 2. Adaptation of cells to allow full incorporation of the SILAC amino acid. Starting from a single dish in normal Dulbecco's modified Eagle's medium, cells are split into two separate dishes, containing "light" (●) and "heavy" (▼) media, respectively. Within a few hours, protein turnover and synthesis will result in incorporation of the "heavy" amino acid (*see* **Subheading 3.2.** and **Fig. 1**). After five cell doublings, virtually all proteins have incorporated the "heavy" amino acid. As this point, the two cell populations are distinguishable in mass spectrometry and analysis of the mixture of proteins from the two pools will allow direct quantitative measurements of protein abundances.

The combination of high quantitative accuracy, flexibility in experimental design and application, ease of use, and robustness makes SILAC a uniquely powerful tool for quantitative proteomics. The rapidly increasing number of applications of SILAC in cell culture systems and even the quantitation of tis-

sues ***(19)*** are testaments to the power of this approach for quantitative proteomics.

2. Materials

2.1. Preparation of SILAC-Labeling Medium

1. Here we describe the SILAC method using Dulbecco's modified Eagle's medium (DMEM) and RPMI-1640, two widely used medium formulations, as examples. By removing arginine, lysine, and methionine from standard formulations of DMEM (Invitrogen, cat. no. 11885-084) and RPMI (Invitrogen, cat. no. 11875-085), we can decide which of these amino acids to replace in the medium in their stable isotope-labeled forms. The depleted media were obtained as custom synthesized media products (Gibco-Invitrogen, Carlsbad, CA).
2. Stable isotope-labeled amino acids: L-arginine-$^{13}C_6$ hydrochloride (Cambridge Isotope Labs, Andover, MA; cat. no. CLM-2265), L-arginine-$^{13}C_6$,$^{15}N_4$ hydrochloride (Sigma-Isotec, St. Louis, MO; cat. no. 608033), L-lysine-$^{13}C_6{}^{15}N_2$ hydrochloride (Sigma-Isotec, no. 608041), and L-methionine-(methyl-$^{13}C^2H_3$) (Sigma-Isotec, no. 299154).
3. Filter flasks for sterile filtration of media (Nalgene Nunc, Rochester, NY).
4. 10% (v/v) dialyzed fetal bovine serum (Gibco-Invitrogen) was supplemented to both "light" and "heavy" media.
5. 1% (v/v) antibiotics (penicillin and streptomycin) and glutamine (Gibco-Invitrogen) was supplemented.

2.2. Cell Culture, Lysis, and Gel Electrophoresis

1. The human adenocarcinoma cell line, HeLa, with a doubling time of approx 24 h (ATCC, Manassas, VA) (*see* **Note 3**).
2. 0.25% trypsin solution with ethylenediamine tetraacetic acid (EDTA) for dissociating attached cells in cell culture was from Gibco-Invitrogen.
3. Cell lysis buffer: 6 *M* urea and 2 *M* thiourea prepared in 50 m*M* Tris (Sigma-Aldrich) and stored at room temperature.
4. Cell scrapers for lysing cells (Sarstedt, Newton, NC).
5. Estimation of protein amounts with Bradford assay (Bio-Rad, Hercules, CA).
6. Dithiothreitol (DTT; Sigma-Aldrich, D9163) and iodoacetamide (Sigma-Aldrich, cat. no. I1149) for reduction and alkylation of proteins, respectively.
7. Polyacrylamide gel electrophoresis analysis was performed using the NuPAGE® Novex 10% Bis-Tris gel system with the 2-morpholinethane sulfonic acid (MES) buffer system (Invitrogen).
8. Colloidal Coomassie stain for visualizing proteins (SimplyBlue, Invitrogen).

2.3. Trypsin Digestion and Nano-Flow LC–MS

1. Gel destaining solution: 50% absolute ethanol in 50 m*M* ammonium bicarbonate (Sigma-Aldrich). Stable at room temperature.
2. Washing buffer: absolute ethanol.

3. Equilibration buffer: 50 m*M* ammonium bicarbonate (Sigma-Aldrich). Stable at room temperature.
4. For enzymatic digestion of gel slices, trypsin (sequencing grade-modified trypsin; Promega, Madison, WI) at 12.5 ng/μL with 50 m*M* ammonium bicarbonate was prepared as needed and stored on ice.
5. Extraction solution 1: 30% acetonitrile in 0.1% trifluoroacetic acid. Stable at room temperature for months.
6. Extraction solution 2: acetonitrile.
7. LC buffers:
   a. Buffer A: 0.1% acetic acid.
   b. Buffer B: 80% acetonitrile in 0.1% acetic acid.

2.4. Mass Spectrometric Data Analysis and Quantitation

1. MS instrument vendor supplied software to generate database search engine compatible MS/MS peak list files (mascot.dll with QSTAR XL [ABI-SCIEX], http://www.matrixscience.com/help/instruments_analyst.html).
2. Search engine to analyze MS data: Mascot v1.9 or higher (Matrix Science LLC, London).
3. Quantitation software: MSQuant (v1.4 or higher) (http://msquant.sourceforge.net) ***(20)***.

3. Methods

3.1. Preparation of SILAC-Labeling Medium

Any defined medium formulation with known sources of amino acids can be adapted for SILAC labeling. Simply leave out the amino acid for labeling from a medium formulation so that the "light" and "heavy" amino acid stocks supplied during medium preparation is the only available source of amino acid to the growing cells. Although we describe SILAC in mammalian cell culture here, the approach can be extended into other cell culture systems, such as yeast ***(21)***, bacteria, plants ***(22)***, and whole organisms. The following example describes preparation of SILAC-labeling medium for labeling with arginine-$^{13}C_6$ (*see* **Notes 1** and **4**). We recommend obtaining amino acid depleted media formulations compatible with all the SILAC-labeling strategies (arg, lys or met alone, arg and lys), allowing flexibility in experimental design and media usage. For instance, when labeling with arginine, simply supplement normal isotope abundance lysine and methionine to the media.

Cells may have to be adapted to grow in SILAC media because of slight differences in medium formulations. In particular, dialyzed serum or altered amino acid concentrations can alter growth rates in some cell types. An overview of the adaptation procedure is outlined in **Fig. 2**.

1. Stable isotope-labeled amino acids like L-arginine-$^{13}C_6$ (Cambridge Isotope Labs), L-lysine-$^{13}C_6{}^{15}N_2$, and L-methionine-(methyl-$^{13}C^2H_3$) (Sigma-Isotec) are dissolved as concentrated stock solutions in either phosphate buffered saline or in the medium lacking the amino acid. These were made as 1000X stocks for use in DMEM where possible, depending on solubility of the individual amino acid. **Table 1** lists the concentrations of SILAC amino acid stock solutions made, *see* **Note 5**.
2. Amino acids that are depleted in the formulation but not used in labeling are replaced with the natural isotope abundance form. Here, although labeling with arginine-$^{13}C_6$, normal lysine and methionine are supplemented to the medium.
3. Divide the medium into two portions—one aliquot for SILAC "light" and the other for SILAC "heavy."
4. Supplement the "light" and "heavy" forms of arginine to the respective aliquots.
5. Filter the media using filter flasks.
6. Dialyzed fetal bovine serum (Gibco-Invitrogen) was supplemented to both "light" and "heavy" media at 10%. Dialyzing serum removes low molecular weight (<10 kDa) components and this removes amino acids and growth factors (*see* **Note 2**).
7. Antibiotics and glutamine (Gibco-Invitrogen) were added as 100X supplements to both media.

3.2. Adaptation of Cells to SILAC-Labeling Media

1. Cells growing in normal DMEM were passaged into two separate dishes containing "light" or "heavy" SILAC media (*see* **Fig. 2**). Trypsin was used to detach adherent cells from the dish. Continued maintenance and passaging of cells was performed in the respective SILAC-labeling medium. Cells are passaged in a low ratio (10–20% of cells in the confluent dish into the next passage) in early passages in order to reduce the amount of SILAC media used and also to ensure that cells are allowed to undergo sufficient doublings for incorporation of the label. After a minimum of five cell doublings, cells are expanded into the number of dishes required for the experiment (*see* **Notes 6–8**).
2. When labeling with arginine in SILAC, arginine may be metabolically converted to proline in some cell lines (HeLa is an example) (*see* **Note 9)**. In order to minimize the arg–pro conversion, we recommend performing a titration of arginine concentration with all new cell lines and empirically determining the concentration at which metabolic conversion is minimized. The arginine concentration determined to avoid the arg–pro conversion in our HeLa cell stocks is one-fourth of DMEM concentration (21 mg/L). It is usually sufficient to perform the titration once for each cell stock, as the value should hold for all frozen cell stocks of the same line. Importantly, note that the change in concentration of any of the amino acids used should be made in both "light" and "heavy" media, effectively nullifying any effect that the change in concentration may have on both cell populations.

Table 1
Concentrations of SILAC Amino Acid Stock Solutions for DMEM and RPMI-1640

Amino acid	Final concentration (DMEM, Invitrogen 11885-084	Final concentration (RPMI, Invitrogen 118875-085	Conc. Prepared	Working stock for DMEN	Working stock for RPMI
Arginine	84 mg/L	240 mg/L	84 g/L	1000X	350X
$^{13}C_6$-Arginine (*see* **Note 4**)	21.85 mg/L (HeLa, 1/4th *see* **Note 9** and text)	– (HeLa, 1/4th *see* **Note 9** and text)	87.4mg/mL	4000X	–
Lysine	146 mg/L	40 mg/L	146 g/L	1000X	3650X
Methionine	30 mg/L	15 mg/L	30 g/L	1000X	2000X

3.3. Application of the Differential Treatment to the SILAC Cells

1. Proteins from cells adapted to "light" and "heavy" SILAC media are distinguishable by MS. The perturbation to the cells is applied to the cells and both treated and untreated cells harvested for analysis (*see* **Note 10**). The decision to treat either the "light" or "heavy" cell population is entirely up to the investigator. It can be very useful to repeat experiments and perform a reverse labeling, particularly in some experimental designs ***(20)***. Briefly, if "heavy" cells are treated in the first experiment, "light" cells are treated and "heavy" cells become the control in the repeated experiment. Ratio changes that indicate a perturbation-specific change are inversed in the repeated experiment and, thus, fully recapitulate and validate the finding.

3.4. Harvest of Cells, Lysis, and Estimation of Total Protein

1. Cells can be harvested using any standard protocol in tissue culture. In this example using HeLa cells, aspirate the medium from the cells and rinse the attached cells in warm phosphate buffered saline to remove remaining serum proteins. This is done twice to reduce the amount of serum albumin (a major component of serum) detectable in subsequent MS analyses. Scrape the cells in a small volume of lysis buffer (750 µL for a 14-cm dish), vortex frequently to ensure cell lysis, and obtain a small aliquot of each cell lysate for a Bradford assay to estimate protein concentration.
2. Collect another set of unmixed "light" and "heavy" cell lysates and freeze these (*see* **Note 11**). Analysis of these unmixed lysates may be necessary to ensure that the cells have fully incorporated the SILAC amino acid. These unmixed protein samples may also be used to calculate a correction factor for the mixing if incomplete incorporation has occurred (*see* **Notes 2** and **6**).
3. After having determined the protein concentrations of the lysates, dilute them with lysis buffer to equalize the protein concentration between the "heavy" and "light" lysate, if necessary. This will simplify subsequent mixing.
4. Mix the lysates in the desired proportions. A one-to-one mixing ratio is often desirable, as it is much easier to identify these pairs in MS analyses. Changes in protein abundances present themselves as fold changes in peptide intensities between "light" and "heavy" peptides (*see* **Note 12** and **Fig. 3**).
5. The disulfide bonds of cysteine in proteins are reduced with DTT (1 µg DTT:50 µg protein) and alkylated with iodoacetamide (5 µg iodoacetamide:50 µg protein) prior to gel separation.
6. Load the sample on the NuPage gel and stain the gel to visualize the proteins. Divide the entire lane into equal slices, digest with trypsin, and process each gel slice for MS analysis (*see* **Note 13**).

3.5. MS Analyses, Protein Identification, and Quantitation

1. Analysis of the peptide mixtures by LC–MS (not described in this protocol) will generate several large data files. In our example, the data extraction script sup-

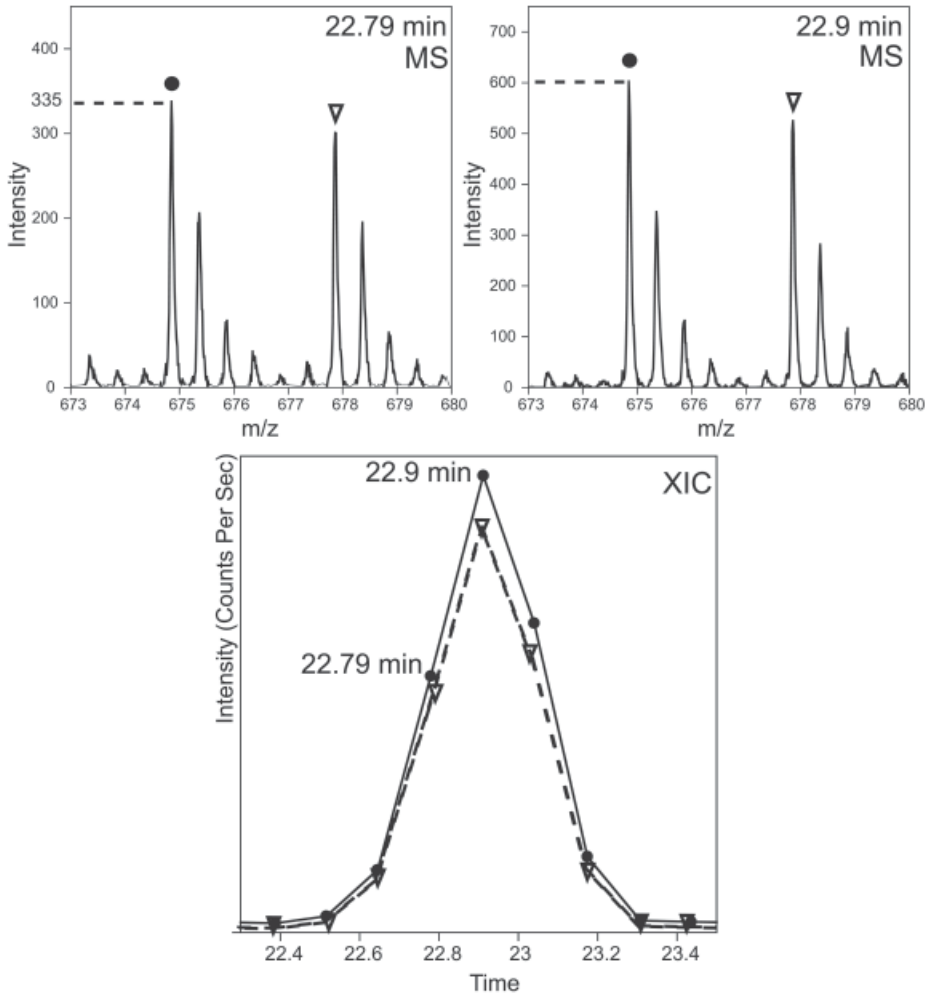


Fig. 3. Quantitation of SILAC-labeled peptide pairs. The top panels show a "light" (●) and "heavy" (▼) peptide pair for a doubly charged peptide labeled with $^{13}C_6$-arginine. The mass separation between the pair is 6 Da. The top panels show individual mass spectrometry (MS) scans of the peptide pair. The bottom panel is the extracted ion chromatogram of the two peptides—"light" (solid line) and "heavy" (dashed line) generated by plotting the intensities determined from the top panel over chromatographic time. The fold abundance ratio can be determined in two ways: determining the ratio of the intensities of each peptide from individual MS spectrums (**top panel**) or by determining the ion chromatograms of the "light" and "heavy" peptides as it elutes from the reversed-phase column, and then determining the ratio of the areas under these curves (**bottom panel**).

plied by Matrix Science (mascot.dll) generates a concatenated, formatted peak list file comprising of all the tandem MS/MS experiments in a single LC–MS run for database searching with the Mascot search program (v1.9 or higher).
2. The resulting "Peptide Summary Report" (saved in .html format) and the raw QSTAR data file (.wiff) are the inputs required by MSQuant (v1.4).
3. Peptides are quantitated by obtaining a ratio between the "light" and "heavy pair. Ratios for peptides from a protein are averaged to give the protein ratio. **Figure 3** illustrates how quantitative information is derived from the MS spectrum.
4. Peptide and protein quantitation is performed by MSQuant in an automated fashion. The software allows the user to manually inspect each peptide pair, validate the quantitative data, and select appropriate spectra to include in determination of the quantitative ratio (*see* **Note 14**) ***(1,23)***.
5. Normalization of the quantitative dataset may be necessary if the original mixing ratio is not close to one-to-one (*see* **Note 15**).

3.6. Conclusion

As shown in this chapter, SILAC is a uniquely versatile quantitative proteomics technology. Application is straightforward and only requires some initial attention to cell culture conditions. We expect the utility of SILAC to grow as many laboratories are currently adapting it to study a wide variety of questions in cell biology, cell signaling, and biomedicine. At the same time, the technology is being reduced to "kit form" by reagent companies. Currently, the main impediment to widespread use is the availability of high performance mass spectrometric instrumentation and especially quantitation software. When all these building blocks are in place, SILAC may become a powerful adjunct to the ubiquitous microarray platforms currently used to address "systems biological" questions.

4. Notes

1. The choice of labeling amino acid is important. Most crucially, the "heavy" amino acid should provide at least a 4-Da separation of "heavy" and "light" peptides (*see* **Fig. 1**). This reduces the overlapping of "light" and "heavy" peptide isotope clusters and inaccurate quantitation (for example, with a 4-Da separation, isotopic distributions of peptides of higher charge states [$z > 4$] will overlap). If peptide clusters overlap, then deconvolution of the isotopic distributions would be necessary to obtain accurate quantitative data. Arginine and lysine are useful for labeling amino acids because trypsin cleaves after these residues and trypsin is a very common proteolytic enzyme used in proteomics ***(24)***. In addition, using both arginine and lysine will label essentially all tryptic peptides, except the carboxyl terminal peptide of the protein. Therefore, because almost all peptides can be quantified, this labeling strategy is very useful for the quantitation of posttranslationally modified peptides ***(17,21)***. It can also be advantageous to label

the posttranslationally modified residues directly, like tyrosine-$^{13}C_9$ in the case of tyrosine phosphorylation ***(15)*** or methionine-(methyl-$^{13}C^2H_3$), which labels the methyl group in several methylated residues directly ***(16)***. The type of stable isotope label is also important as deuterated molecules elute earlier than their nondeuterated counterparts in reversed-phase chromatography ***(25)***. Although using the more expensive ^{13}C- and ^{15}N-labeled amino acids will reduce this separation between "light" and "heavy" peptides, these amino acids can also lead to slight elution shifts, which in high accuracy experiments can dominate the quantitation error. Noncoeluting peptides are more accurately quantified by their extracted ion chromatograms (*see* **Fig. 3**).

2. Some cell lines do not grow well in dialyzed serum because of the loss of low molecular weight growth factors, therefore, all cell lines should be tested before a SILAC-labeling experiment. One possible workaround is to grow these cells with a combination of undialyzed serum (2.5%) and dialyzed serum (7.5%), although this leads to a small percentage of "light" peptides in the "heavy" cell state ***(26)***. Determining the degree of incorporation in a separate analysis of the "heavy" sample alone can provide a correction factor.
3. HeLa, the human adenocarcinoma cell line, is used in this protocol, but any mammalian cell line that can grow in dialyzed serum can also be used. The adaptation of the protocol to other culture systems should adhere to the fundamental principles of the SILAC mammalian culture system.
4. We expect that this methodology will be soon be commercially available from several vendors, and may therefore be packaged under different trade names. Regardless, these packages would come in a simple ready-to-use format, complete with stable isotope labeled-amino acid, a common media formulation like DMEM, and dialyzed serum. Although these packages are excellent for new users interested in testing SILAC in their experimental system, convenience does come at a premium. We have found that it can be more economical to prepare SILAC medium from components purchased separately, if larger experiments or routine use is planned.
5. SILAC "heavy" amino acids are not sterile filtered as stock solutions to avoid losses. Normal isotope abundance, "light" amino acids were prepared in the same manner, but differences in the molecular weight of the amino acid were noted so that molar equivalents of "light" and "heavy" amino acid were prepared in stock solutions. Stock solutions of normal amino acids were filtered through a 0.22-µm syringe filter.
6. A combination of protein turnover, new cell division, and protein synthesis contributes to the incorporation of the SILAC amino acid. Even if one does not take into account protein turnover, the maximum remaining unlabeled protein is $(1/2^n)$th of the total protein, where n is the number of cell divisions. Therefore, after five cell doublings, one should minimally have $1–(1/2^5)$ or approx 97% of protein in the labeled form.
7. Depending on the doubling time of the cells, it may be possible to achieve the five cell doublings while expanding the cell culture to the required number of

dishes. As an illustration, cultivating a single, confluent 10-cm dish into 10 confluent 14-cm dishes (an increase in 30X culture surface area) will achieve the requisite number of doublings for full incorporation without discarding any cells. Therefore, by passaging cells and using just a fraction of the original cell pool (a 1-in-10 split), the adaptation can be achieved on a much smaller and convenient scale (10% of the confluent 10-cm dish into a confluent 14-cm dish is equivalent to the previous example).

8. Cells can be continually passaged in SILAC-labeling media if necessary. The cost of maintaining cells in "light" and "heavy" media may be high but would save another adaptation period to the labeling media if a second round of experiments is necessary.
9. The arg–pro interconversion is a result of an oversupply of arginine in the medium for the particular cell line ***(27)***. Medium formulations have historically been developed and evaluated by their ability to support cell growth. General medium formulations like DMEM or RPMI-1640 are intended to support a large number of cell lines and, thus, the concentration of any amino acid in the medium may be far higher than required for normal growth. The titration series reduces the amount of arginine used in steps, and through MS analyses, peptides that contain proline and arginine are inspected to see if they contain a $^{13}C_6$-Arg form, as well as a $^{1}C_6$-Arg and $^{13}C_5$-Pro-containing form. This can be done by searching the MS/MS data using both $^{13}C_6$-Arg and $^{13}C_5$-Pro as variable modifications in Mascot or by visually looking for a coeluting satellite peak of mass 5 Da higher than the 1^3C_6-Arg-containing peptide. The usable arg concentration will usually be a range of values rather than a narrow setpoint. Note however, that using too low an arg concentration may result in a conversion of proline to arginine from the proline present in the media, and potentially retardation in the growth rate of cells.
10. Any differential treatment can be applied after full incorporation of the SILAC label has taken place. Some examples include drug or growth factor treatment, comparison of differentiated cells to undifferentiated cells, immunoprecipitation of cell lysates to a different bait, or exposure to an external stimuli, such as ultraviolet treatment. If the differential treatment is a prolonged cultivation step, such as differentiation of cells into a different cell type (such as myoblast to myotube differentiation), care should be taken to continually maintain the control state and use the same number of cells, or the same amount of protein, in the final comparison.
11. Analyses of unmixed cell lysates from a SILAC experiment serve as a good quality control for the experiment and can be essential for troubleshooting. This will easily pinpoint potential problems in the experimental design and is strongly recommended in a new SILAC experiment.
12. Most often, protein abundances between abundant, housekeeping proteins will remain unchanged (1:1 ratio) if the SILAC experiment is designed to track subtle changes in protein abundance. The high accuracy and precision possible with SILAC quantitative analyses ***(27)*** make it possible and highly desirable to per-

form such experiments. Indeed, comparing two very different cell populations with the SILAC (or any other) approach will yield a long list of quantitative changes and the arduous task of validating each of these remains a challenge.

13. The one-dimensional gel separation coupled to nano-flow LC–MS is a popular approach for proteomic analysis of samples. This system couples the high resolving powers of sodium dodecyl sulfate-polyacrylamide gel electrophoresis at the protein level and the C18 reversed-phase separation at the peptide level (reviewed in **ref.** ***28***). MS-based quantitation of SILAC samples can be performed with any MS analytical workflow, including complex mixture analysis with in-solution digests using one dimensional-LC–MS or multidimensional LC–MS, MALDI-based workflows, and even the top-down approach with whole protein analyses ***(29)***.
14. In order to accurately quantify a peptide pair, overlapping isotopic peaks that contribute to the signal of either member of the pair must be avoided or corrected for. The isotopic distributions of the "light" and "heavy" pair should not overlap. The signal-to-noise ratio of both members of a pair should be sufficiently high to obtain good quantitative information. Furthermore, several spectra (ideally > nine to model a Gaussian peak) of the peptide pair should be acquired across the eluting peak in order to accurately plot their respective ion chromatograms. Software for automated peptide and protein quantitation are useful for proteomic analyses of larger datasets (RelEx ***[30]***, ASAPRatio ***[31]***, and MSQuant [http://msquant.sourceforge.net] are examples of such software).
15. The ratios of these "unchanging" proteins are centered about the mixing ratio and can sometimes be a useful normalization factor. For instance, if the protein concentration of the lysates determined by Bradford assay was inaccurate, and the lysates were mixed in a 1:0.8 ratio, then the large proportion of unchanging components would show the ratio of 0.8 (heavy-light). Dividing the quantitative ratios in the entire dataset by the normalized mixing ratio will then yield the correct fold change.

References

1. Ong, S. E. and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* **1,** 252–262.
2. Ong, S. E., Blagoev, B., Kratchmarova, I., et al. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell Proteomics* **1,** 376–386.
3. Zhu, H., Pan, S., Gu, S., Bradbury, E. M., and Chen, X. (2002) Amino acid residue specific stable isotope labeling for quantitative proteomics. *Rapid Commun. Mass Spectrom.* **16,** 2115–2123.
4. Jiang, H. and English, A. M. (2002) Quantitative analysis of the yeast proteome by incorporation of isotopically labeled leucine. *J. Proteome Res.* **1,** 345–350.
5. Everley, P. A., Krijgsveld, J., Zetter, B. R., and Gygi, S. P. (2004) Quantitative cancer proteomics: stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research. *Mol. Cell Proteomics* **3,** 729–735.
6. Gu, S., Du, Y., Chen, J., et al. (2004) Large-scale quantitative proteomic study of

PUMA-induced apoptosis using two-dimensional liquid chromatography-mass spectrometry coupled with amino acid-coded mass tagging. *J. Proteome Res.* **3,** 1191–1200.
7. Blagoev, B., Kratchmarova, I., Ong, S. E., Nielsen, M., Foster, L. J., and Mann, M. (2003) A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* **21,** 315–318.
8. de Hoog, C. L., Foster, L. J., and Mann, M. (2004) RNA and RNA binding proteins participate in early stages of cell spreading through spreading initiation centers. *Cell* **117,** 649–662.
9. Foster, L. J., De Hoog, C. L., and Mann, M. (2003) Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proc. Natl. Acad. Sci. USA* **100,** 5813–5818.
10. Andersen, J. S., Lam, Y. W., Leung, A. K., et al. (2005) Nucleolar proteome dynamics. *Nature* **433,** 77–83.
11. Blagoev, B., Ong, S. E., Kratchmarova, I., and Mann, M. (2004) Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* **22,** 1139–1145.
12. Kratchmarova, I., Blagoev, B., Haack-Sorensen, M., Kassem, M., and Mann, M. (2005) Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. *Science* **308,** 1472–1477.
13. Pratt, J. M., Petty, J., Riba-Garcia, I., et al. (2002) Dynamics of protein turnover, a missing dimension in proteomics. *Mol. Cell Proteomics* **1,** 579–591.
14. Doherty, M. K., Whitehead, C., McCormack, H., Gaskell, S. J., and Beynon, R. J. (2005) Proteome dynamics in complex organisms: using stable isotopes to monitor individual protein turnover rates. *Proteomics* **5,** 522–533.
15. Ibarrola, N., Molina, H., Iwahori, A., and Pandey, A. (2004) A novel proteomic approach for specific identification of tyrosine kinase substrates using [13C]tyrosine. *J. Biol. Chem.* **279,** 15,805–15,813.
16. Ong, S. E., Mittler, G., and Mann, M. (2004) Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nat. Methods* **1,** 119–126.
17. Ibarrola, N., Kalume, D. E., Gronborg, M., Iwahori, A., and Pandey, A. (2003) A proteomic approach for quantitation of phosphorylation using stable isotope labeling in cell culture. *Anal. Chem.* **75,** 6043–6049.
18. Ballif, B. A., Roux, P. P., Gerber, S. A., MacKeigan, J. P., Blenis, J., and Gygi, S. P. (2005) Quantitative phosphorylation profiling of the ERK/p90 ribosomal S6 kinase-signaling cassette and its targets, the tuberous sclerosis tumor suppressors. *Proc. Natl. Acad. Sci. USA* **102,** 667–672.
19. Ishihama, Y., Sato, T., Tabata, T., et al. (2005) Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. *Nat. Biotechnol.* **23,** 617–621.
20. Schulze, W. X. and Mann, M. (2004) A novel proteomic screen for peptide-protein interactions. *J. Biol. Chem.* **279,** 10,756–10,764.
21. Gruhler, A., Olsen, J. V., Mohammed, S., et al. (2005) Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell Proteomics* **4,** 310–327.

22. Gruhler, A., Schulze, W. X., Matthiesen, R., Mann, M., and Jensen, O. N. (2005) Stable isotope labeling of Arabidopsis thaliana cells and quantitative proteomics by mass spectrometry. *Mol. Cell Proteomics* **4,** 1697–1709.
23. MacCoss, M. J. and Matthews, D. E. (2005) Quantitative MS for proteomics: teaching a new dog old tricks. *Anal. Chem.* **77,** 294A–302A.
24. Olsen, J. V., Ong, S. E., and Mann, M. (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell Proteomics* **3,** 608–614.
25. Zhang, R. and Regnier, F. E. (2002) Minimizing resolution of isotopically coded peptides in comparative proteomics. *J. Proteome Res.* **1,** 139–147.
26. Gehrmann, M. L., Hathout, Y., and Fenselau, C. (2004) Evaluation of metabolic labeling for comparative proteomics in breast cancer cells. *J. Proteome Res.* **3,** 1063–1068.
27. Ong, S. E., Kratchmarova, I., and Mann, M. (2003) Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J. Proteome Res.* **2,** 173–181.
28. Ishihama, Y. (2005) Proteomic LC-MS systems using nanoscale liquid chromatography with tandem mass spectrometry. *J. Chromatogr. A.* **1067,** 73–83.
29. Meng, F., Forbes, A. J., Miller, L. M., and Kelleher, N. L. (2005) Detection and localization of protein modifications by high resolution tandem mass spectrometry. *Mass Spectrom. Rev.* **24,** 126–134.
30. MacCoss, M. J., Wu, C. C., Liu, H., Sadygov, R., and Yates, J. R., 3rd (2003) A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. *Anal. Chem.* **75,** 6912–6921.
31. Li, X. J., Zhang, H., Ranish, J. A., and Aebersold, R. (2003) Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. *Anal. Chem.* **75,** 6648–6657.

4

Quantitative Proteomics of Mouse Brain and Specific Protein-Interaction Studies Using Stable Isotope Labeling

Toshitaka Sato, Yasushi Ishihama, and Yoshiya Oda

Summary

We describe a new method for quantitative tissue proteomics using culture-derived isotope tags (CDIT), which are cells grown in stable isotope-enriched medium and added to each tissue sample to provide internal standards. After protein identification by mass spectrometry (MS), each peak derived from tissue protein is quantified relative to the corresponding CDIT peak. The amounts of each peak in different tissue samples can be compared relative to CDIT. Even if the corresponding peak from CDIT can not be detected, a peak with a similar scan number, but different sequence on liquid chromatography (LC)–MS, can be used to obtain semiquantitative values. Absolute quantification is possible by determining the protein amount in CDIT in advance using unlabeled synthetic peptides; this is less costly than other methods, such as AQUA.

For identification of specific components in a protein complex, target proteins are enriched or isolated by affinity techniques using bait-conjugated matrix, but many nonspecific binders are often found. Stable isotope labeling strategies have proven particularly advantageous for the discrimination of proteins specifically associated with the target population from nonspecifically, copurified contaminants. We also describe a protocol for efficient in-gel digestion and high-performance nano-LC column preparation, which makes it possible to quantify larger numbers of proteins.

Key Words: Tissue proteome; in vivo labeling; metabolic labeling; stable isotope labeling; relative quantification; absolute quantification; protein interaction; protein complex; culture-derived isotope tag.

1. Introduction

In our original publication about in vivo labeling of proteins using heavy isotope incorporation, protein expression profiles were quantified through the

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

use of whole cell stable isotope labeling ***(1)***. The basic premise of this approach involves growing cells on a medium lacking an essential nutrient and supplementing the medium with a stable isotope-labeled compound (reference culture). The advantages of this technique include minimal sample manipulation. After the labeling, reference and sample cultures are mixed. Subsequent procedures, such as cell lysis, protein extraction, sample work-up, and separation on electrophoretic gels or liquid chromatography (LC), are thus identical for the two cultures, eliminating the possibility that variability in these steps may affect the calculated ratio of proteins of interest ***(2)***. However, in vivo labeling of proteins in living cells with stable isotopes requires that the metabolic pathways be accessible to the label, and some samples, particularly animal tissues, are very difficult to label with stable isotopes. Recently, Wu et al. reported a method to label mammalian organisms by long-term feeding with a diet enriched in stable isotope (in vivo labeling) ***(3)***. The drawbacks of their method are that it takes a long time (44 d) to get a labeled rat, the diet is expensive, and some tissues, such as brain, can not be completely labeled with stable isotopes. We have developed another quantitative approach to the mammalian proteome using in vivo labeling ***(4)***. Mouse neuroblastoma Neuro2A cells were cultured in ^{13}C-labeled leucine-rich medium. We named these cells culture-derived isotope tags (CDIT), and used them as comprehensive internal standards to normalize the variations of sample preparation and analysis (**Fig. 1**). The CDIT strategy was applied to quantify the mouse brain proteome.

Absolute concentrations of proteins in the sample are also important, in addition to the relative concentrations between different stages. Conventional methods, such as antibodies, enzymatic assays, and staining with dyes have been used to measure the protein amounts. Another technique for absolute quantitation is to use mass spectrometry (MS) after spiking known amounts of isotopically labeled analytes as an internal standard; this is called isotope dilution ***(5)*** or AQUA ***(6)***. One of the proteolytic peptides of a particular target protein is synthesized using isotope-labeled reagents and the absolute amount is measured. Because isotope-labeled synthetic peptides are added to unlabeled protein mixtures in the isotope dilution method, results are directly influenced by the recovery rate in purification steps, such as immunoprecipitation and in digestion steps. Recently Shevchenko et al. reported that the AQUA approach does not work for in-gel-digested proteins because of the lower recovery in the digestion or extraction step ***(7)***. In addition, synthesis of isotope-labeled peptides is difficult to apply to the whole proteome, because peptide synthesis usually requires 10-fold excess amounts of reagents (expensive isotopically labeled reagents in this case), and the synthetic scale is μmol scale, whereas MS requires only fmol scale. We also developed a novel strategy for absolute quantification, in which quantified synthetic unlabeled peptides and labeled

cultured cells are used ***(4)***. This approach was combined with CDIT and applied to mouse whole brain proteins (**Fig. 2**).

Recent progress toward defining the complete proteomes of various eukaryotic organisms has generated a need for efficient procedures to determine the functions of newly identified proteins. Proteins often interact with each other or with nonprotein molecules, such as DNA, to form transient or stable complexes, which mediate biological activities. Identification of the direct physical interactions of protein complexes provides an important clue to the functional units of cellular molecular machinery, and the study of protein interactions has emerged as a valuable method for finding novel components of signaling pathways. The yeast two-hybrid screening is a genetic selection procedure that is designed to detect binary interactions, but this approach only records interactions between pairs of genes, misses interactions stabilized by more than two partners, and does not necessarily reflect the physiological environment. Another approach for the analysis of protein complexes involves affinity purification using a bait molecule for complex isolation or enrichment. A tag is introduced onto a target protein, and the tagged protein is purified from a whole cell lysate, together with associated proteins, which are subsequently characterized by MS. Affinity purification appears to be the most efficient and gentle discriminatory separation technique for the retrieval of protein complexes. In particular, well-characterized small affinity tags such as FLAG, His6, or glutathione-*S*-transferase have been widely used in affinity purification of recombinant proteins or protein complexes. MS can rapidly and reliably identify the binding partners in a protein complex, and therefore the current limiting step appears to be the finding of specific proteins, rather than protein identification. An affinity approach may fail to purify proteins of low abundance because of nonspecific contamination during purification, and therefore, it remains difficult to distinguish specific from nonspecific interactions. The use of more stringent washes to reduce contaminating proteins may also affect the binding of weaker interactors. For example, regulatory subunits, which are often not tightly associated with the enzymes they regulate, may be lost during the purification procedure. Each protein has unique properties, which can be exploited for its purification, but it is difficult to design a procedure valid for all cases. Thus, comprehensive analysis of macromolecular complexes has been hindered by the lack of a general and efficient purification strategy. In biochemical studies to identify drug-binding proteins (drug target proteins), compound-conjugated affinity matrix reagents have played an important role; however, the affinity and specificity of synthetic small molecules for their protein targets are rather low in many cases. Thus, nonspecific interactions often lead to difficulty in specifying the primary binding partners of a synthetic compound. The stable isotope labeling strategies described here

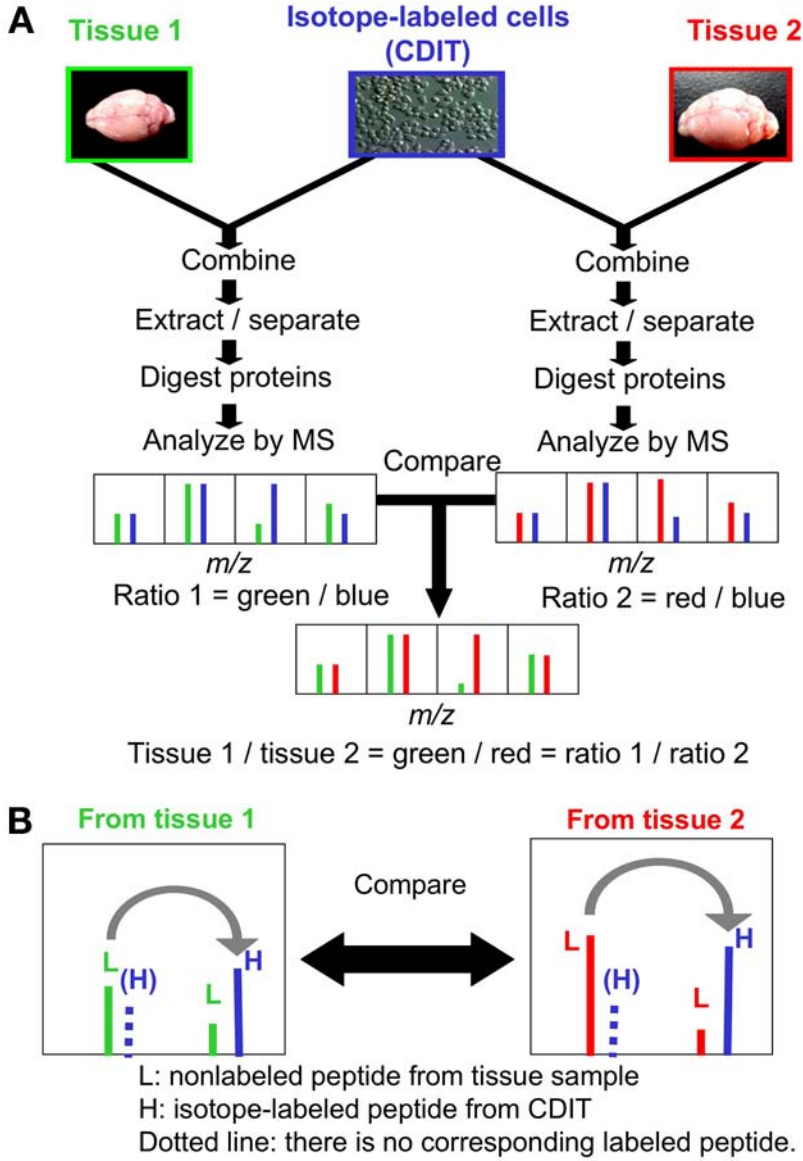


Fig. 1. **(A)** Quantitative tissue proteome analysis using stable isotope-labeled cultured cells as global internal standards. Tissue samples 1 and 2 are mixed with cultured cells early in the process to obviate variations during sample preparation. After protein extraction and separation, digested proteins are analyzed by mass spectrometry to identify and quantify proteins. The ratio between the two isotopic distributions (one from the tissue sample and one from cultured cells labeled with isotopes) can then be determined from the mass spectra. Changes of protein levels in two tissue samples are estimated by

provide the researcher with powerful new ways to approach these problems (**Fig. 3**). Functionally important specific interactions can be picked out of the background binding through the detection of isotope ratios on MS, avoiding a trade-off between false-positive binding and the ability to detect weak components ***(8–10)***. This approach provides a powerful new tool for the characterization of a wide range of drug–protein complexes and macromolecular complexes.

2. Materials

2.1. Preparation of Stable Isotope Labeled Cells

Any type of labeling method can be used for CDIT. Here we present one example.

1. Culture medium: custom-made L-leucine-deficient RPMI-1640 (Invitrogen, Grand Island, NT) supplemented with U-$^{13}C \times$ six labeled l-leucine (Cambridge Isotope Laboratories), 10% fetal bovine serum (Morgate, Australia), 1% penicillin (5000 U/mL)-streptomycin (5 mg/mL) solution (Invitrogen) (*see* **Note 1**).
2. Cell lines: HCT116 (human colorectal cancer cell line) and Neuro 2A (mouse neuroblastoma cell line).

2.2. Protein Extraction

1. Trypsin (0.05%)-ethylenediaminetetraacetic acid (EDTA) solution (Invitrogen).
2. Protease inhibitor cocktail (Roche Diagnostics GmbH, Mannheim, Germany).
3. Cell extraction buffer: 10 m*M* HEPES-KOH (pH 7.4), 10 m*M* KCl, 1 m*M* dithiothreitol (DTT), 1 m*M* $MgCl_2$, 1 m*M* NaF, 1 m*M* $NaVO_3$, protease inhibitor cocktail.

2.3. Preparation of Protein Mixtures From Mouse Brain for Quantification

1. Cell extraction buffer: 10 m*M* HEPES-KOH (pH 7.4), 10 m*M* KCl, 1 m*M* DTT, 1 m*M* $MgCl_2$, 1 m*M* NaF, 1 m*M* $NaVO_3$, protease inhibitor cocktail.
2. Fractionation buffer: 0.32 *M* sucrose, 1 m*M* sodium hydrogen carbonate-protease inhibitor cocktail, and 1.2 *M* sucrose solution.

Fig. 1. *(continued from opposite page)* calculating the ratio of the two ratios, ratio 1/ratio 2, a procedure which cancels out the internal standards (cultured cells). **(B)** Method for semi-quantifying a protein found in a tissue sample, but not in the cells cultured with stable isotopes.

The ratio of a target peptide, which does not have a corresponding labeled peak in cultured cells, is obtained by using the peak ratio to an isotope-labeled, cultured-cell-derived peptide of different sequence, but with the closest (ideally the same) retention time in LC–MS.

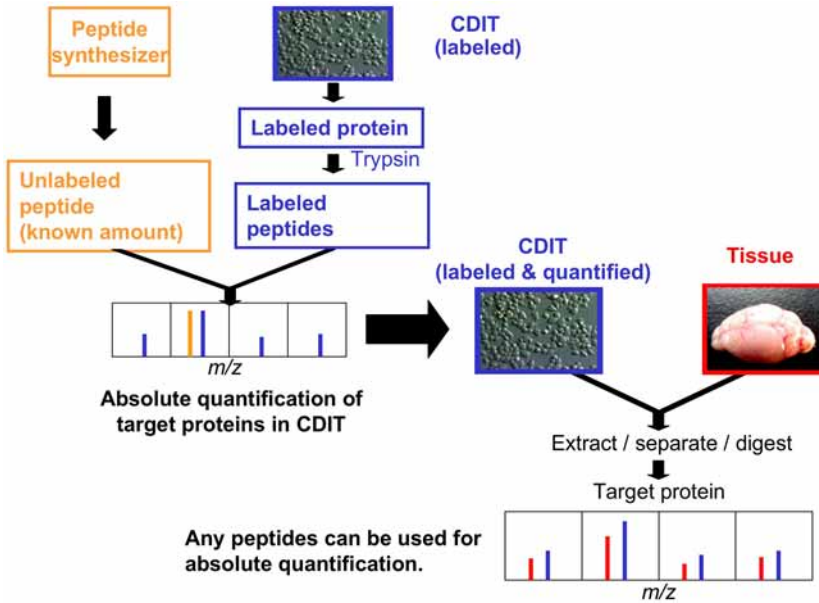


Fig. 2. Absolute quantification using amplified isotope double dilution. An unlabeled synthetic peptide is used as an internal standard for a target protein expressed in CDIT. This reverse approach is cost-effective, because the scale of conventional peptide synthesis is nanomole to micromole, but mass spectrometry requires only femtomole to picomole levels of peptides. Also, peptide synthesis requires a large excess of reagents, which should therefore be unlabeled on cost grounds. On the other hand, all proteins in CDIT are labeled very efficiently with stable isotopes. Another advantage of this method is that once a target protein in CDIT is quantified based on the known amount of synthetic peptide, all digested peptides derived from that protein are available to obtain the absolute amount of the target protein in tissue samples. This improves the chance of finding a paired target peptide and internal standard, and more reliable results can be obtained by calculating the average of the ratios of all pairs.

2.4. Affinity Column Preparation for Protein Complex Analysis

1. NHS-activated Sepharose™ 4 Fast Flow (Amersham Bioscience, Uppsala, Sweden) (*see* **Note 2**) or Affi-gel 10 (Bio-Rad, Hercules, CA).
2. Coupling buffer: buffer without primary amine, such as 0.1 *M* $NaHCO_3$, HEPES, or MOPS at pH 6.0–8.0. For compound coupling, tetrahydrofuran/methanol/water = 1/1/1 (v/v/v) plus a small amount (ca. 0.1%) of triethylamine.
3. Protein ligand, antibody, enzyme, or compounds, for example.
4. Ethanolamine and Tris-HCl.

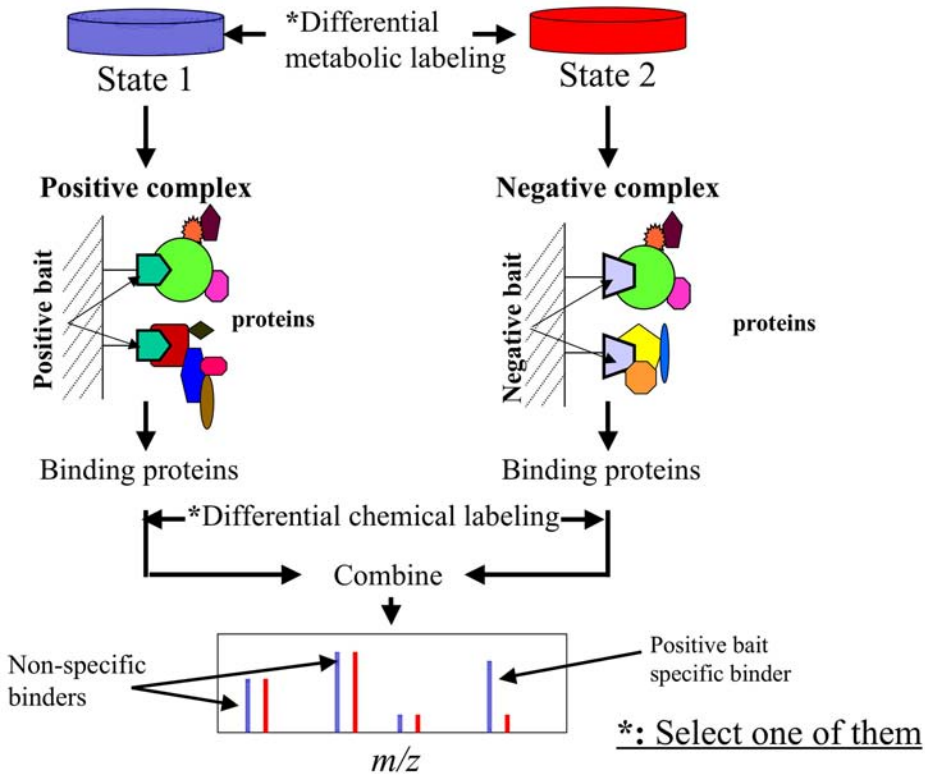


Fig. 3. Strategy for the identification of specific binding proteins from protein complexes using stable isotope labeling. To distinguish specific complex components from copurifying proteins, a control purification (negative complex, i.e., point-mutated bait) is carried out in which the complex of interest is not enriched. To discriminate nonspecific binding partners of a complex isolated from cells, the samples are prepared identically. Stable isotope labeling is performed metabolically at the cell culture level, or by chemically attaching isotopically heavy or light tags after affinity purifications. After combining two binding samples, followed by proteolysis, peak ratios are calculated to determine the specificity of each component of a complex.

2.5. In-Gel Digestion

1. Reducing buffer: approx 1 g DTT (Sigma-Aldrich, St. Louis, MO) in 100 mL of 0.1 *M* Tris-HCl–6 *M* guanidine-HCl, pH 8.5 (*see* **Note 3**). Store at –80°C in aliquots. Do not repeat the freeze–thaw process.
2. Alkylating buffer: approx 1 g of acrylamide (Bio-Rad) in 100 mL of 0.1 *M* Tris-HCl–6 *M* guanidine-HCl, pH 8.5 (*see* **Note 3**). Store at –80°C after aliquots. Do not repeat the freeze–thaw process.

3. Washing buffer: methanol/water/acetic acid = 50/40/10 (v/v/v). Use high-performance liquid chromatography (HPLC) grade or equivalent grade.
4. Washing buffer: methanol/water/acetic acid = 50/40/10 (v/v/v). Use HPLC grade or equivalent grade.
5. Digestion buffer: 0.1% of 5-cyclohexyl-1-pentyl-β-D-maltoside (CYMAL-5) (Anatrace, Maumee, OH) in 50 m*M* ammonium bicarbonate buffer (pH 8.0) including an appropriate amount of TPCK-treated, sequencing-grade modified trypsin (Promega, Madison, WI) (*see* **Note 4**).
6. Extracting buffer 1: acetonitrile/water/trifluoroacetic acid (TFA) (50/50/0.1 [v/v/v]). Use HPLC grade or equivalent grade.
7. Extracting buffer 2: acetonitrile/water/TFA (75/5/0.1 [v/v/v]). Use HPLC grade or equivalent grade.

2.6. Sample Preparation for LC–MS

1. Empore Disk C18 (3M, cat. no. 2215).
2. Needle and plunger.
3. Methanol.
4. Buffer A: acetonitrile/water/TFA (5/95/0.1 [v/v/v]). Use HPLC grade or equivalent grade.
5. Buffer A: acetonitrile/water/TFA (80/20/0.1 [v/v/v]). Use HPLC grade or equivalent grade.
6. 1-mL plastic syringe.

2.7. Nano-Column Preparation

1. Capillary puller (Sutter Instrument, P-2000).
2. Fused silica capillary (e.g., Polymicro Technology).
3. Methanol.
4. Innova Quartz capillary adaptor (Phoenix; cat. no. 3-2302S).
5. 3-μm C18 silica particles (e.g., ReproSil-Pur 120 C18-AQ; cat. no. r13.aq).
6. Air-pressure pump (Nikkyo Technos, Tokyo, Japan); other similar products are also available (*see* http://www.proxeon.com or http://www.brechbuehler.ch/usa/).
7. Slurry solution: prepare 3-μm C18 silica particles, 10–30 mg, in 100 μL of methanol. Can be stored for months at room temperature.

2.8. Nano LC–MS System

1. Custom flow-splitter: tee connector (Upchurch, cat. no. P775) and glass pipet (Supelco, cat. no. P0549-1PAK).
2. HPLC mobile phase A: 0.5% acetic acid. Prepare 1 L in HPLC grade water (e.g., from Millipore system, Milli-Q Biocel A10). Can be stored for months at room temperature.
3. HPLC mobile phase B: 0.5% acetic acid and 80% acetonitrile. Prepare 1 L in HPLC grade water. Can be stored for months at room temperature.
4. Equipment:

a. Electrospray tandem mass spectrometer, e.g., quadrupole TOF hybrid: QSTAR Pulsar (AB/MDS-Sciex) or Ultima (Micromass).
b. Linear iontrap: QTrap (AB/MDS-Sciex).
c. Iontrap: LCQ Deca (Finnigan) or Esquire (Bruker).
d. Triple quadrupole: API4000 (MDS-Sciex) or Quantum (Finnigan).
e. Database search software, e.g., Mascot.

3. Methods

3.1. Preparation of Stable Isotope-Labeled Cells

1. Warm culture medium, trypsin (0.05%)-EDTA, and phosphate-buffered saline (PBS) in a water bath at 37°C.
2. Place 20 mL of culture medium in a 75-cm^3 culture flask.
3. Transfer 1×10^6 cells to the culture flask.
4. Incubate the flask at 37°C in an incubator at 5% CO_2.
5. Cells at 70–80% confluency (subconfluent) should be passaged. Remove the culture medium by aspiration and wash cells with PBS. Add 5 mL of trypsin (0.05%)-EDTA solution to the flask and incubate at 37°C for 5 min.
6. Add 5 mL of culture medium to inactivate trypsin.
7. Transfer the suspension to a 15-mL centrifuge tube and centrifuge at 200*g* for 5 min at room temperature.
8. Aspirate the supernatant, then resuspend the cells in 5 mL of culture medium.
9. Place 20 mL of culture medium in a 75-cm^3 culture flask and transfer 1 mL of resuspended cells to the culture flask.
10. Incubate the flask at 37°C in an incubator at 5% CO_2.
11. Repeat **steps 7–10** at least three times to label proteins of the cells with U-^{13}C × six labeled L-leucine.

3.2. Protein Extraction

1. Grow labeled cells in 10 150-mm diameter tissue culture dishes to the desired cell density. A 150-mm diameter tissue culture dish of confluent cells of a typical adherent cell line usually contains about 1 to 5×10^7 cells per dish.
2. Aspirate the medium, wash cells gently with PBS, and aspirate the wash solution.
3. Harvest cells by scraping with Teflon cell scraper (*see* **Note 5**).
4. Transfer cells to a 50-mL centrifuge tube, wash with PBS, and centrifuge at 200*g* for 5 min at 4°C.
5. Add 20 mL of cell extraction buffer to the cell pellet, suspend thoroughly, and incubate on ice for 15 min. Put a drop of the cell suspension on a glass slide and check under a microscope. The cells should be swollen, but should not have burst.
6. Transfer the cell suspension to a precooled Dounce tissue homogenizer and homogenize the cells on ice. Check the homogenized cells under a microscope after every 10 strokes. Stop when more than 90% of the cells have burst.
7. Transfer the homogenate to a 50-mL centrifuge tube and centrifuge at 500*g* for 5 min at 4°C. The pellet from this step (nuclear fraction) contains unbroken cells and nuclei.

8. Remove the supernatant carefully, transfer it to centrifuge tubes, and centrifuge at 100,000*g* for 1 h at 4°C.
9. Transfer the supernatant (soluble fraction) to a tube. Suspend the pellet (membrane fraction) thoroughly in 5 mL of cell extraction buffer.
10. These fractions are stored at –80°C until use.

3.3. Preparation of Protein Mixtures From Mouse Brain for Quantification

1. All mice should be treated ethically according to the rules of the institute.
2. Sacrifice the mice to be compared by rapid decapitation under anesthesia, then remove each brain and weigh it.
3. Combine each experimental brain with labeled Neuro2A cells (1×10^8 cells/g wet brain) and suspend the mixture in 0.32 *M* sucrose solution containing 1 m*M* sodium hydrogen carbonate and protease inhibitor cocktail (10 mL/g wet brain).
4. Homogenize the suspension in a Teflon Potter-type homogenizer and centrifuge at 710*g* for 10 min to remove the nuclear fraction.
5. Centrifuge the supernatant at 13,800*g* for 10 min to separate the soluble fraction and insoluble materials. Resuspend the pellets in 0.32 *M* sucrose solution, and layer the suspension on 1.2 *M* sucrose solution, then centrifuge at 82,500*g* for 2 h to separate cytosol, trafficking and secretion-related organella, and mitochondrial fractions.
6. Suspend the organella thoroughly in cell extraction buffer containing 1% CHAPS and 8 *M* urea. Extract protein for 1 h at 4°C, then centrifuge at 100,000*g* for 1 h at 4°C and collect the supernatant.
7. For whole brain analysis, suspend frozen mouse brain with added Neuro2A cells in cell extraction buffer (10 mL buffer/g wet brain) and homogenize on ice. Centrifuge the homegenate at 100,000*g* for 1 h at 4°C. The supernatant and the pellet are collected as the soluble fraction and membrane-nuclear fraction, respectively.
8. Suspend the pellet fraction thoroughly in 5 mL of cell extraction buffer containing 1% CHAPS and 8 *M* urea. Extract protein for 1 h at 4°C, centrifuge at 100,000*g* for 1 h at 4°C, and collect the supernatant.
9. The extracted proteins are stored at –80°C until use.

3.4. Affinity Column Preparation for Protein Complex Analysis

Determination of specificity in protein complex analysis to distinguish nonspecific binders is achieved by comparison of peak intensities between bait-positive and bait-negative complexes. Therefore, the strategy in principle is the same as that for relative quantification using stable isotope labeling, e.g., for elucidation of biomarkers in disease models vs normal models.

1. Dissolve the ligand to be coupled in a suitable coupling buffer.
2. Wash NHS-activated Sepharose™ 4 Fast Flow with 10–15 gel volumes of cold 1 m*M* HCl on a Buchner funnel.

3. Equilibrate the washed gel with coupling buffer and mix with the coupling solution. The recommended ratio of volumes, coupling solution to gel, is 1:2 (*see* **Note 6**).
4. Continue gentle agitation of the gel slurry on a rocker, shaker, or wheel for 2–4 h at room temperature or overnight at 4°C.
5. After the coupling is completed, any nonreacted groups on the gel should be blocked by leaving the gel to stand in Tris-buffer or ethanolamine for a few hours.
6. To wash the gel after coupling, use a method which alternates two different buffers (high and low pH, respectively).
7. The coupled affinity gel is now ready for use.

3.5. In-Gel Digestion

To identify more proteins after gel-based separation, efficient in-gel digestion is a critical step ***(11–13)***.

1. Recommended gels are thin (typically 0.5 to 1.0-mm thickness) and short (typically less than 10-cm length), because lower gel volume results in a higher recovery from the gel.
2. Negative staining, such as the zinc-imidazole method (Bio-Rad), is recommended because negative staining provides better peptide recovery or affords a better signal-to-noise ratio in the MS spectra than silver staining.
3. A whole lane on a sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel is excised and the gel is cut into 3 to 5-mm slices. Then each gel strip is cut into 1 × 1-mm pieces to increase the surface volume (*see* **Note 7**).
4. Put gel pieces from same gel-strip into a 0.6-mL Eppendorf tube or a 0.8-mL well of a 96-well plate.
5. Add 0.2 mL of reducing buffer and sonicate tubes/plates for 10 min.
6. Discard reducing buffer.
7. Add 0.2 mL of alkylating buffer and sonicate tubes/plates for 10 min.
8. Discard alkylating buffer.
9. Add 0.2 mL of washing buffer and sonicate tubes/plates for a period of from 5 min to several hours.
10. Discard washing buffer.
11. Repeat **steps 9** and **10** three to five times.
12. Add 0.2 mL of 50 m*M* ammonium bicarbonate buffer to neutralize gel pieces and sonicate tubes/plates for 5 min.
13. Discard ammonium bicarbonate buffer.
14. Add 0.2 mL of acetonitrile to dehydrate gel pieces and sonicate tubes/plates for 5 min.
15. Discard acetonitrile.
16. Dry gel pieces thoroughly in a speed-vac evaporator (Thermo Savant, NY) (*see* **Note 8**).
17. Add a minimum volume of digestion buffer (*see* **Note 8**).
18. Add a sufficient volume of 50 m*M* ammonium bicarbonate buffer so that gel pieces recover to their original size.

19. Incubate samples at 37°C for a period from 2 h to overnight.
20. Add extracting buffer 1 in an amount of 10 times the volume of gel pieces, and sonicate tubes/plates for 15 min.
21. Transfer solvent to new tubes/plates.
22. Add extracting buffer 2 in an amount of five times the volume of gel pieces, and sonicate tubes/plates for 15 min.
23. Transfer solvent to tubes/plates of the same specification.
24. Dry samples in a Speed-Vac evaporator.

3.6. Sample Preparation for LC–MS

Although they can be achieved in the same step, desalting, filtering, and concentration serve different purposes. Salts can interfere with the action of reversed-phase materials, decreasing the column resolution. Particles can easily become irreversibly lodged in analytical columns, resulting in complete sample loss. Dilute samples can take unacceptably long times to load onto the analytical column. These points certainly apply to in-gel-digested samples, but are absolutely critical for in-solution-digested samples. A general procedure for desalting, filtering, and concentration is as follows ***(14)***:

1. Place an Empore Disk C18 (3M, cat. no. 2215) on a flat, clean surface such as a disposable plastic Petri dish.
2. Wet the membrane using 20 μL of methanol and keep it wet throughout the protocol.
3. Punch out a small disk using a blunt-tipped hypodermic needle. The disk sticks in the needle and can be transferred into a pipet tip.
4. Push the disk out of the needle and fix it in the tapering part of the pipet tip by using a piece of fused silica or tubing fitting inside the needle (*see* **Note 9**).
5. Acidify the digested peptide sample using 2% TFA. For a 10 m*M* Tris-HCl (pH 8.0) buffered solution, one-tenth volume is sufficient. Check with pH strips on an equal amount of digestion buffer.
6. Condition a C_{18}-StageTip by introducing 5 μL of methanol from the top and pressing the liquid through at 50 μL/min using a 1-mL plastic syringe fitted with a pipet tip trimmed at both ends to serve as an adapter between the syringe and StageTip.
7. Equilibrate the StageTip using 10 μL of buffer A at 50 μL/min.
8. Load sample at 20 μL/min.
9. Wash with 10 μL of buffer A at 50 μL/min.
10. Elute using 2–5 μL of buffer B at 10 μL/min, and dilute 10 times by adding buffer A.
11. Directly load onto a precolumn or an analytical column of an LC–MS system for analysis. Reduce the volume if required by evaporation in a vacuum centrifuge; however, do not dry completely.

3.7. Nano-Column Preparation

In order to achieve higher sensitivity, we are currently using columns with 50 to 100-μm inner diameter at the flow rate of 100–200 nL/min. NanoHPLC columns can be self-packed without compromising separation efficiency. Instead of a capillary with frit we use tapered fused-silica capillaries without any frit, which also work as electrospray emitters. In such a "stone-bridge" column, where the opening size is two- to fivefold larger than the average particle size, particles at the end of the column arch over the opening and these self-assembled particles work as a frit. Here, we describe the procedure to use our "stone-bridge" columns with self-assembled particle frits using 3-μm particles and an 8-μm opening capillary ***(15)***.

1. Prepare a fused silica capillary 40 cm in length, and burn the middle point of the capillary to remove the outer coating.
2. Pull the fused silica capillary using a P-2000 to make tapered needles (opening size: ca. 6 μm. Tapered needles are also commercially available (e.g., PicoTip™ cat. no. FS360-75 or 100-8-N-20; New Objective, Woburn, MA).
3. Immerse a tapered needle for 1 s into a glass reservoir filled with a slurry of C18 silica particles (3 μm)* in methanol (10–30 mg/100 μL).
4. Pack the particles using a disposal syringe with an Innova Quartz capillary adaptor (Phoenix, AZ) filled with methanol. This step is to avoid contamination of small particles from ferrules, for instance, as well as avoid destroying the opening structure during the following packing process.
5. Place the slurry solution in an air-pressure pump connected to a helium gas bomb and connect the capillary to the pump.
6. Apply a pressure up to 50 bar to pack the C18 materials.
7. After packing is completed (the column reaches the desired length), the column is connected to an HPLC pump and washed with HPLC buffers B and A at 100–150 bar. The bed will shrink 5–10% in this step.
8. The column can be stored in HPLC-grade water.

3.8. Nano LC–MS System

Peptides eluting from nanoLC columns are ionized via electrospray ionization and transferred with high efficiency into a mass spectrometer for analysis by single and tandem (MS/MS) mass spectrometry. The combination of efficient peptide separation and sensitive identification of individual peptides makes nanoLC–MS/MS an extremely powerful technique for the characterization of complex mixtures of peptides. We recommend careful optimization of all parameters for nanoelectrospray of a peptide mixture.

1. Prepare the mass spectrometer for measurement according to the instructions of the manufacturer and calibrate the machine using a standard.

2. Place the column in front of the orifice.
3. Set the potential between the column and the entrance of the mass spectrometer to give a stable spray (usually between 1500 and 2500 V).
4. Set the data-dependent acquisition method in the MS software according to the manufacturer's instructions. The *m/z* range is selected to encompass potential ions of interest. A limited *m/z* range adapted to the abilities of the mass spectrometer may, however, be advantageous to increase the scan time, improve sensitivity, or to eliminate background noise in the low *m/z* region. The use of high-purity solvents helps to reduce background noise and thereby increases the overall sensitivity of the nanoLC MS/MS experiment.
5. Set the pump for a standard 80 min gradient to:

Buffer B (%)	Time (min)
5	0
10	5
30	65
100	70
100	80
5	80.1

6. Start the acquisition and the gradient. Stop the acquisition after the gradient has finished eluting from the column.
7. Generate a peak list from the raw data file using a peak-picking script (e.g., Mascot.dll for Qstar).
8. Submit the peak list file to a database searching engine (e.g., Mascot) for protein identification.
9. Save and store the output html file from Mascot.

3.9. Differential Measurement Using CDIT

Metabolic labeling for quantitative proteomics has been developed using CDIT in various states. Because the incorporation of stable isotopes into proteins is accomplished at the initial stage of the sample preparation using isotopically modified medium, metabolic-labeling approaches have a significant advantage over chemical-labeling approaches, in which variations among samples can be introduced during preparation. However, samples from tissues are difficult to label metabolically. In this chapter, we describe a new approach for quantitative tissue proteomics using cultured cells as a source of internal standards for tissue samples, as illustrated in **Fig. 1**.

1. Analyze the two samples of interest by LC–MS followed by database searching to obtain protein identification results (*see* **Note 10**).
2. Extract Leu-containing peptides from the search results.
3. Integrate the peak areas of the Leu-containing peptides and their isotope pairs (*see* **Note 11**). Quantify each peak relative to its corresponding isotope peak to

obtain the ratio of the unlabeled peptide to the labeled peptide from Neuro2A cells (Ratio A = tissue A/Neuro2A) (*see* **Note 12**).
4. Quantify ratios from another tissue sample relative to Neuro2A cells using the same procedure (Ratio B = tissue B/Neuro2A).
5. Calculate ratios of tissue A to tissue B by dividing ratio A by ratio B.

3.10. Absolute Quantification Using CDIT

1. Select amino acid sequences of tryptic peptides, considering the amino acid used for isotope labeling, ultraviolet (UV) absorbance, mass, and hydrophobicity (tyrosine-containing peptides are preferable in terms of purification and quantitation by LC–UV at 280 nm, and proper hydrophobicity and mass are necessary for LC retention and MS detection, respectively) (*see* **Note 13**).
2. Synthesize peptides with a peptide synthesizer according to the instructions of the manufacturer.
3. Purify and quantify peptides using HPLC with UV detection at 280 nm based on the molar extinction of tyrosine-containing peptides ***(16)***.
4. Prepare CDIT (e.g., Neuro2A cells with $^{13}C_6$ Leu labeling).
5. Spike synthesized peptides into CDIT cells and extract proteins using ultrasonication.
6. Dissolve proteins in Tris buffer (pH 8.0) and 8 *M* urea, then reduce, alkylate, and digest with Lys-C. Then dilute with 50 m*M* ammonium bicarbonate buffer (pH 9.0) and digest with trypsin.
7. Perform LC–MS analyses to obtain the ratio of labeled peptides to the unlabeled peptides. The spiked amounts are adjusted to obtain a ratio in the range of 0.1–10 (*see* **Notes 10** and **12**).
8. Quantify protein amounts in Neuro2A cells from the spiked amounts and the peak ratio between Neuro2A-labeled peptides and unlabeled synthetic peptides.
9. Spike Neuro2A cells into the sample (e.g., mouse brain proteins) and perform relative quantitation using LC–MS.
10. Calculate the absolute amounts of mouse brain proteins from the peak ratio between Neuro2A-labeled peptides and unlabeled brain peptides in the mass spectra and the spiked amounts of Neuro2A proteins.

3.11. Determination of Specificity for Protein-Interaction Studies

1. Analyze the two samples of interest by LC–MS followed by database searching to obtain protein identification results (*see* **Note 10**).
2. Extract Leu-containing peptides from the search results.
3. Integrate the peak areas of the Leu-containing peptides and their isotope pairs. Quantify each peak relative to its corresponding isotope peak to obtain the ratio of the unlabeled peptide to the labeled peptide (Ratio A = unlabeled A/labeled A). If unlabeled peptides are from bait-positive complex, choose higher ratio values, which are specific binders to bait (*see* **Note 12**).

4. Notes

1. Dialyzed serum is usually used instead of normal serum ***(17–19)***. The use of normal serum in culture medium contaminates the pool of stable isotope-labeled amino acids and leads to the incorporation of unlabeled and labeled amino acids into synthesized proteins, because normal serum contains large amounts of free (unlabeled) amino acids. Thus, the use of dialyzed serum without free amino acids allows more efficient incorporation of labeled amino acids. However, dialyzed serum is devoid of low-molecular-weight components and impairs the growth of cells that require these components. As an alternative, it is reported that $^{13}C_6$-Arg and $^{13}C_6$, $^{15}N_2$-Lys are efficiently incorporated (~94%) with normal serum ***(20)***. In the case of $^{13}C_6$-Leu, however, we identified unlabeled proteins in cells after repeated passage, and we also observed that around 10% of proteins were unlabeled in commercially available dialyzed serum culture medium, which results in almost the same level of unlabeled proteins as that of normal serum culture. This potentially leads to a lower accuracy, though separate analysis of a labeled state sample to determine the amount of unlabeled protein present provides a correction factor for quantification. When one peptide contains multiple labeled amino acids, this correction factor should be carefully calculated. For instance, 10% unlabeled amino acid contamination in labeled culture medium gives ratios of 10 (unlabeled):90 (labeled) in peptide containing one labeled amino acid, 1 (2 unlabeled):81 (2 labeled) in peptide containing two labeled amino acids, and 0.1 (3 unlabeled):72.9 (3 labeled) in peptide containing three labeled amino acids.
2. Immobilization should be attempted through the least critical region of the ligand. Although NHS-activated Sepharose™ 4 Fast Flow immobilizes ligands containing $–NH_2$ groups, ligands containing –SH and –OH groups can be immobilized onto gel with Thiopropyl Sepharose™ 6B and Epoxy-activated Sepharose™ 6B (Amersham Bioscience), respectively.
3. Denaturing conditions for in-gel alkylation are very important for cysteine-rich proteins. Instead of 6 *M* guanidine-HCl, 8 *M* urea can be used.
4. Adding CYMAL-5 (or *n*-octyl glucoside at the same concentration) is very effective for in-gel digestion, probably because an increase in the solubility of the protein enhances the efficiency of the digestion, CYMAL-5 serves to solubilize the digested peptides, and CYMAL-5 prevents adsorption of peptides on the tube wall and pipet tip. However, over-addition of CYMAL-5 (or *n*-octyl glucoside) interferes with mass spectrometric analysis. The maximum amount of CYMAL-5 (or *n*-octyl glucoside) is up to 0.1% concentration × 10 μL volume. The amount of trypsin is usually up to 50 ng per tube per well of plate.
5. Commercially available kits can be used to prepare cell lysate and to isolate cell organella. For example, Pierce Biotechnology Inc. (Rockford, IL) provides M-PER (mammalian cell extraction), Mem-PER (membrane protein enrichment), NE-PER (nuclear and cytoplasmic extraction) and a mitochondrial isolation kit. For details, *see* the instructions of these kits.

6. Be aware that NHS-ester groups are hydrolyzed at higher pH. Equilibrated gel and coupling solution should be mixed rapidly.
7. TGel pieces that are too small are easy to lose during in-gel digestion procedures, and may clog pipet tips and nanoLC columns.
8. A dehydration–rehydration process is used to drive the entry of protease into the gel matrix for in-gel digestion. The yields of in-gel digestion in dry gels are generally about 30–50% higher than in semi-dry gels. Also, in the rehydration process, trypsin may not sufficiently penetrate into the gels when digestion buffer still remains in the tube per well of plate after the gels have recovered to the original size, which reduces the digestion efficiency. The added volume of digestion buffer is usually 1–10 μL, but is highly dependent on the total volume of gel strips.
9. StageTips can be stored dry at room temperature.
10. When LCQ/LTQ (Thermoelectron) is used, choose the profile mode during MS scanning (MS/MS scanning is usually performed in the centroid mode). The profile mode gives a better mass chromatogram, which makes it easier to calculate peak areas.
11. Labeled peptides with different sequences eluting at the same retention times can be used when the corresponding labeled peptides are not found.
12. To calculate the peak area of a target peak automatically, use MSQuant (http://msquant.sourceforge.net/) or Xome (http://xome.hydra.mki.co.jp/en/).
13. To measure the absolute amount of a synthetic peptide is the most difficult point. In general, amino acid analysis is used to obtain the absolute amount of peptides, but this analysis requires large amounts of materials, and synthetic peptides contain water, salts, and impurities. Therefore, it is really difficult to determine accurately the absolute amount.

References

1. Oda, Y., Huang, K., Cross, F. R., Cowburn, D., and Chait, B. T. (1999) Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. USA* **96,** 6591–6596.
2. Sechi, S. and Oda, Y. (2003) Quantitative proteomics using mass spectrometry. *Curr. Opin. Chem. Biol.* **7,** 70–77.
3. Wu, C. C., MacCoss, M. J., Howell, K. E., Matthews, D. E., and Yates, J. R. 3rd. (2004) Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal. Chem.* **76,** 4951–4959.
4. Ishihama, Y., Sato, T., Tabata, T., et al. (2005) Quantitative mouse brain proteomics using culture-derived isotope tags asinternal standards. *Nat. Biotechnol*. **23,** 617–621.
5. Barr, J. R., Maggio, V. L., Patterson, D., et al. (1996) Isotope dilution—mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. *Clin. Chem.* **42,** 1676–1682.
6. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. USA* **100,** 6940–6945.

7. Havlis, J. and Shevchenko, A. (2004) Absolute quantification of proteins in solutions and in polyacrylamide gels by mass spectrometry. *Anal. Chem.* **76,** 3029–3036.
8. Oda, Y., Owa, T., Sato, T., et al. (2003) Quantitative chemical proteomics for identifying candidate drug targets. *Anal. Chem.* **75,** 2159–2165.
9. Ranish, J. A., Yi, E. C., Leslie, D. M., et al. (2003) The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* **33,** 349–355.
10. Blagoev, B., Kratchmarova, I., Ong, S. E., Nielsen, M., Foster, L. J., and Mann, M. (2003) A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* **21,** 315–318.
11. Katayama, H., Nagasu, T., and Oda, Y. (2001) Improvement of in-gel digestion protocol for peptide mass fingerprinting by matrix assisted laser desorption/ionization-time of flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **15,** 1416–1421.
12. Katayama, H., Sato K., Takeuchi M., Deguchi-Tawarada M., Oda Y., and Nagasu, T. (2003) Optimization of in-gel digestion system in combination with thin gel separation and negative staining in 96-well plate format. *Rapid Commun. Mass Spectrom.* **17,** 1071–1078.
13. Katayama, H., Tabata, T., Ishihama, Y., Sato, T., Oda, Y., and Nagasu, T. (2004) Efficient in-gel digestion procedure using 5-cyclohexyl-1-pentyl-beta-D-maltoside as an additive for gel-based membrane proteomics. *Rapid Commun. Mass Spectrom.* **18,** 2388–2394.
14. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75,** 663–670.
15. Ishihama, Y., Rappsilber, J., Andersen, J. S., and Mann, M. (2002) Microcolumns with self-assembled particle frits for proteomics. *J. Chromatogr. A* **979,** 233–239.
16. Gill, S. C. and von Hipple, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182,** 319–326.
17. Ong, S. E., Blagoev, B., Kratchmarova, I., et al. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as simple and accurate approach to expression proteomics. *Mol. Cell Proteomics* **1,** 376–386.
18. Ong, S. E., Kratchmarova, I., and Mann, M. (2003) Properties of ^{13}C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J. Proteome Res.* **2,** 173–181.
19. Everley, P. A., Krijgsveld, J., Zetter, B. R., and Gygi, S. P. (2004) Quantitative cancer proteomics: stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research. *Mol. Cell Proteomics* **3,** 729–735.
20. Hathout, Y., Flippin, J., Fan, C., Liu, P., and Csaky, K. (2005) Metabolic labeling of human primary retinal pigment epithelial cells for accurate comparative proteomics. *J. Proteome Res*. **4,** 620–627.

5

The Absolute Quantification Strategy

Application to Phosphorylation Profiling of Human Separase Serine 1126

Scott A. Gerber, Arminja N. Kettenbach, John Rush, and Steven P. Gygi

Summary

The absolute quantification (AQUA) strategy provides a means to determine the precise protein or modified protein levels directly from cells or tissues. The technique is based on two major principles: stable isotope dilution theory and the use of synthetic peptides containing such stable isotopes to exactly mimic native counterparts after proteolysis. These peptides can be synthesized with modifications such as phosphorylation, methylation, and acetylation to allow for the direct, quantitative analysis of posttranslationally modified proteins. In this chapter, we discuss the development of an AQUA method and demonstrate its usefulness in the measurement of endogenous levels of the human protein separase at a functionally relevant phosphorylation site, serine 1126.

Key Words: Quantitative proteomic; quantitation; proteomics; mass spectrometry; isotopic labeling; absolute quantification; phosphoprotein.

1. Introduction

Quantitative mass spectrometry has been greatly enhanced by the use of stable isotope-labeled internal standards ***(1–4)***. Because synthetic biomolecules enriched in ^{18}O, ^{13}C, ^{2}H, and ^{15}N are otherwise chemically and physically indistinguishable from their native counterparts, they are ideal internal standards for use in mass spectrometric assays. Given a minimum threshold of isotope enrichment, separation and detection of these isotopomers via mass spectrom-

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

etry (MS) is straightforward. For the quantitative analysis of the levels of proteins and posttranslationally modified proteins, isotope-labeled tryptic peptides can be synthesized to mimic their native equivalents. We have previously described this approach as the "AQUA" strategy, as it is capable of returning absolute quantitative numbers (e.g., copies/cell, ng/mL, and so on) instead of a relative ratio ***(5–8)***. In this chapter, we will present an overview of the strategy with a focus on the quantification of the extent of cell cycle-dependent phosphorylation of separase from HeLa cells.

2. Materials

2.1. Peptide Synthesis

1. Fmoc-protected heavy-isotope amino acid monomers (Cambridge Isotope Laboratories, Inc., Andover, MA; *see* **Note 1**).
2. Dimethyl sulfoxide (DMSO) and *N*,*N*-dimethylformamide (American Bioanalytical, Natick, MA).
3. Small glass reaction vessels with sidearms for Symphony peptide synthesis instrument (Protein Technologies, Tucson, AZ; cat. no. SMP-010135).

2.2. Cell Culture and Lysis

1. Dulbecco's modified Eagle's medium (DMEM) (Gibco/Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (HyClone, Ogden, UT), penicillin/streptomycin (Gibco/Invitrogen) (penicillin at 100 U/mL, streptomycin at 100 μg/mL).
2. Phosphate buffered saline (PBS), pH 7.4 (Gibco/Invitrogen)
3. Trypsin-ethylenediaminetetraacetic acid (EDTA) (0.25% trypsin with EDTA-4Na) (Gibco/Invitrogen).
4. Thymidine: dissolve 2.4 g of thymidine HCl (Sigma, St. Louis, MO) in 50 mL water to make a 100X stock. Incubation at 37°C may be required for complete solubilization. Sterile filter (0.2 μm), aliquot, and store at –20°C.
5. Nocodazole: add 2 mL DMSO to a 2 mg vial of nocodazole (Sigma) to make a 10,000X stock. Aliquot and store at –20°C. Immediately prior to use, warm to 37°C, and predilute 1:100 in sterile media for the estimated total amount to be used.
6. 100X Phosphatase inhibitor cocktail A. **Warning**: always wear nitrile gloves when handling DMSO-solvated phosphatase inhibitors, as they are extremely toxic and can penetrate skin and latex more easily than nitrile gloves. Microcystin-LR (dissolve 0.5 mg vial in 1 mL DMSO for 10,000X stock), (–)-*p*-bromotetramisole oxalate (dissolve 9.3 mg in 1 mL DMSO for 1000X stock), and cantharidin (dissolve 10 mg in 1 mL DMSO for a 10,000X stock) (Sigma-Aldrich). Store concentrated stocks at –80°C. Dilute to 100X working stock in DMSO (e.g., 10 μL microcystin-LR, 10 μL cantharidin, and 100 μL (–)-*p*-bromotetramisole oxalate concentrated stocks in 880 μL DMSO) and store at –20°C until use.

7. 100X Phosphatase inhibitor cocktail B: 100 m*M* sodium fluoride, 100 m*M* β-glycerophosphate, 120 m*M* sodium molybdate, and 400 m*M* sodium tartrate (Sigma-Aldrich). Store in aliquots at –20°C until use.
8. Lysis buffer: 50 m*M* Tris-HCl, pH 8.0, 150 m*M* sodium chloride, 4 m*M* magnesium chloride, 1 m*M* EDTA, 0.5% Triton X-100 in Milli-Q grade water.
9. Protease inhibitors (Mini-Complete, EDTA-free; Roche, Penzburg, Germany).
10. Hemacytometer and phase-contrast microscope.
11. Bench-top centrifuge.
12. 500-mL liquid nitrogen.
13. Microtip sonicator.

2.3. Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

1. Sample buffer: LDS 4X sample buffer (Invitrogen).
2. Gels: Novex NuPage 4–12% bis-Tris gels, 1.5-mm width, 10-well.
3. MOPS–sodium dodecyl sulfate (SDS) running buffer (Invitrogen).
4. Molecular weight markers: Mark12 unstained markers (Invitrogen).
5. Coomassie blue stain: 0.5% (w/v) Coomassie blue R-250 (Pierce, Rockville, MD), 10% acetic acid, and 30% methanol in water.
6. Destain: 10% acetic acid and 30% methanol in water.

2.4. In-Gel Digestion With Trypsin

1. In-gel destain: 50 m*M* ammonium bicarbonate (no pH adjustment) in 50% high performance liquid chromatography (HPLC)-MS grade acetonitrile/50% HPLC-MS grade water (both solvents, Honeywell Burdick and Jackson, Muskegon, MI).
2. Digestion buffer: 500 m*M* ammonium bicarbonate (no pH adjustment) in purified water as a 10X stock. Prepare dilutions fresh for use. Store at room temperature.
3. Trypsin stock solution: reconstitute a 20-μg vial of modified sequencing grade trypsin (Promega, Milwaukee, WI) with 80 μL of the supplied reconstitution buffer to make a 12.5X stock. Store at –20°C for up to 1 mo.

2.5. Liquid Chromatography

1. FAMOS autosampler (LC Packings, Palo Alto, CA) with inert PEEKsil valve and Peltier chiller.
2. Agilent 1100 binary HPLC pump (Agilent, Palo Alto, CA) with degasser.
3. Microtees, 100- and 175-μm I.D. PEEK tubing (Upchurch, Bellingham, WA).
4. Reverse-phase C_{18} column (1 × 150-mm) (Vydac, Hisperia, CA).
5. Aqueous buffer A: 0.4% ultra high-purity acetic acid (JT Baker, Phillipsburg, NJ), 0.005% heptafluorobutyric acid (HFBA) (Pierce, Rockford, IL), 5% HPLC-MS grade acetonitrile in HPLC-MS grade water.
6. Organic buffer B: 0.4% ultra high-purity acetic acid (JT Baker, Phillipsburg, NJ), 0.005% HFBA (Pierce), 5% HPLC-MS grade water in HPLC-MS grade acetonitrile.
7. Peptide loading buffer/autosampler chase solution: 4% high-purity formic acid (EM Science, Gibbstown, NJ), 8% HPLC-MS grade acetonitrile in HPLC-MS grade water.

2.6. Mass Spectrometry

1. Quantum triple quadrupole mass spectrometer (ThermoElectron, San Jose, CA).
2. Grade 5.0 nitrogen and argon gases.

2.7. Data Processing

1. QualBrowser software (supplied with mass spectrometer).

3. Methods

The AQUA strategy is capable of profiling protein posttranslational modifications directly from sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) separated whole cell lysates ***(5,6,9,10)***. In this sense, one can think of an AQUA experiment as analogous to a quantitative Western blot. The use of an isotopically labeled internal standard peptide provides a recognition marker in the reverse-phase chromatographic, as well as tandem MS fragmentation space. Only signals generated in the corresponding light channel at the same retention time as the chemically identical heavy peptide are unique to the native peptide formed by trypsin digestion. As both phosphorylated and unphosphorylated AQUA peptides can be synthesized, the extent of phosphorylation at a given residue can be established by monitoring the native levels of both in a single experiment.

The overall AQUA process is described in **Fig. 1**. An AQUA method is first developed by synthesizing and standardizing an appropriate AQUA peptide, with one residue enriched in stable isotopes. This purified standard is then interrogated by tandem mass spectrometry to select an informative and sensitive fragment ion for use in a selected reaction monitoring (SRM) experiment. A biologically relevant whole cell lysate is then fractionated by gel electrophoresis, and the region corresponding to the migration of the protein of interest is excised and digested in the presence of the internal standard peptide. Finally, the digest is extracted from the gel and transferred to an LC–SRM setup for analysis.

Human separase is a cysteine endoproteinase that acts on its substrate cohesin in a cell-cycle dependant manner ***(11)***. Cohesin is a multisubunit protein complex that holds sister chromatids together during the early stages of mitosis. When certain criteria have been established that ensure fidelity of chromosome segregation into progeny cells, separase is activated and proteolyzes cohesin, allowing chromatids to separate and thereby marking the initiation of cell division.

One mechanism that prevents separase from being prematurely activated is inhibitory phosphorylation. We have previously shown that separase activity is blocked by phosphorylation at serine 1126 ***(5,6)***. As an example method of development for the AQUA strategy, we describe here in detail the process for

monitoring endogenous, cell cycle-dependant phosphorylation of human separase at serine 1126.

3.1. Synthesis of AQUA Peptides

1. For accurate quantification, the isotope clusters of the biological peptide and the heavy-isotope AQUA peptide should not overlap. This is achieved by synthesizing AQUA peptides with at least six heavy isotopes, and ideally this additional mass originates from just one heavy isotope amino acid residue. Most AQUA peptides contain heavy isotopes from Leu, which adds seven to the peptide mass (6 ^{13}C and 1 ^{15}N), or Pro or Val, which add six to the mass (5 ^{13}C and 1 ^{15}N). When necessary, more expensive heavy-isotope monomers are incorporated, for example, Asn (+6), Phe (+10), or Tyr (+10). In unusual cases AQUA peptides contain two heavy-isotope monomers, for example, any combination of two Ala (+4) and Gly (+3) residues (*see* **Note 1**).
2. Peptides are made following the same synthesis strategies and approaches as for their ordinary peptide counterparts, except that at one particular residue, an expensive heavy-isotope monomer is substituted for the usual light-isotope monomer. Much of the strategy unique to AQUA peptide synthesis is devoted to making cost-effective use of this expensive heavy-isotope monomer. One approach is to "trick" a conventional peptide synthesis instrument into making a smaller than usual amount of peptide (*see* **Note 2**). Another strategy is to use a highly specialized peptide synthesis instrument designed for small volume (<0.1 mL) reagent delivery (*see* **Note 3**).
4. For the Rainin/PTI Symphony instrument, dissolve heavy-isotope monomers in DMSO to make 0.1 *M* stocks. For example, add 2.78 mL DMSO to 0.1 g of U-^{13}C-^{15}N-L-Leu-N-Fmoc. Let the monomer sit at room temperature for 5 min with gentle agitation, and examine the vial to be sure the monomer has dissolved. If peptides are made at 5-μmole scale, 11 peptides can be made from the dissolved heavy-isotope monomer stock. Unused stock can be stored at –20°C for 6 mo (*see* **Note 4**).
5. Program the Symphony instrument using cycles for a 25-μmol scale synthesis, the smallest scale allowed on the Symphony. This delivers 1.25 mL of each synthesis reagent to the reaction vessel. However, instead of 25 μmol of preloaded resin, put 5-μmol resin in a small, glass, sidearmed reaction vessel (*see* **Note 5**). This reduces the amount of heavy-isotope monomer needed for synthesis fivefold.
6. At the heavy-isotope position, use a special Symphony cycle that deprotects and washes the peptide resin, but does not couple the monomer. Start peptide synthesis but program a pause for that reaction vessel at the heavy-isotope position. When the pause is reached, remove the cap from the reaction vessel sidearm and add 0.25 mL freshly made activator solution, then add 0.25 mL monomer. In Manual Operations, select the Mix operation (with drain on) for 120 min to couple the monomer to peptide resin. When coupling is finished, in Manual Operations wash the resin three times and then resume synthesis, which should wash the resin further and continue automated synthesis of the remainder of the peptide.

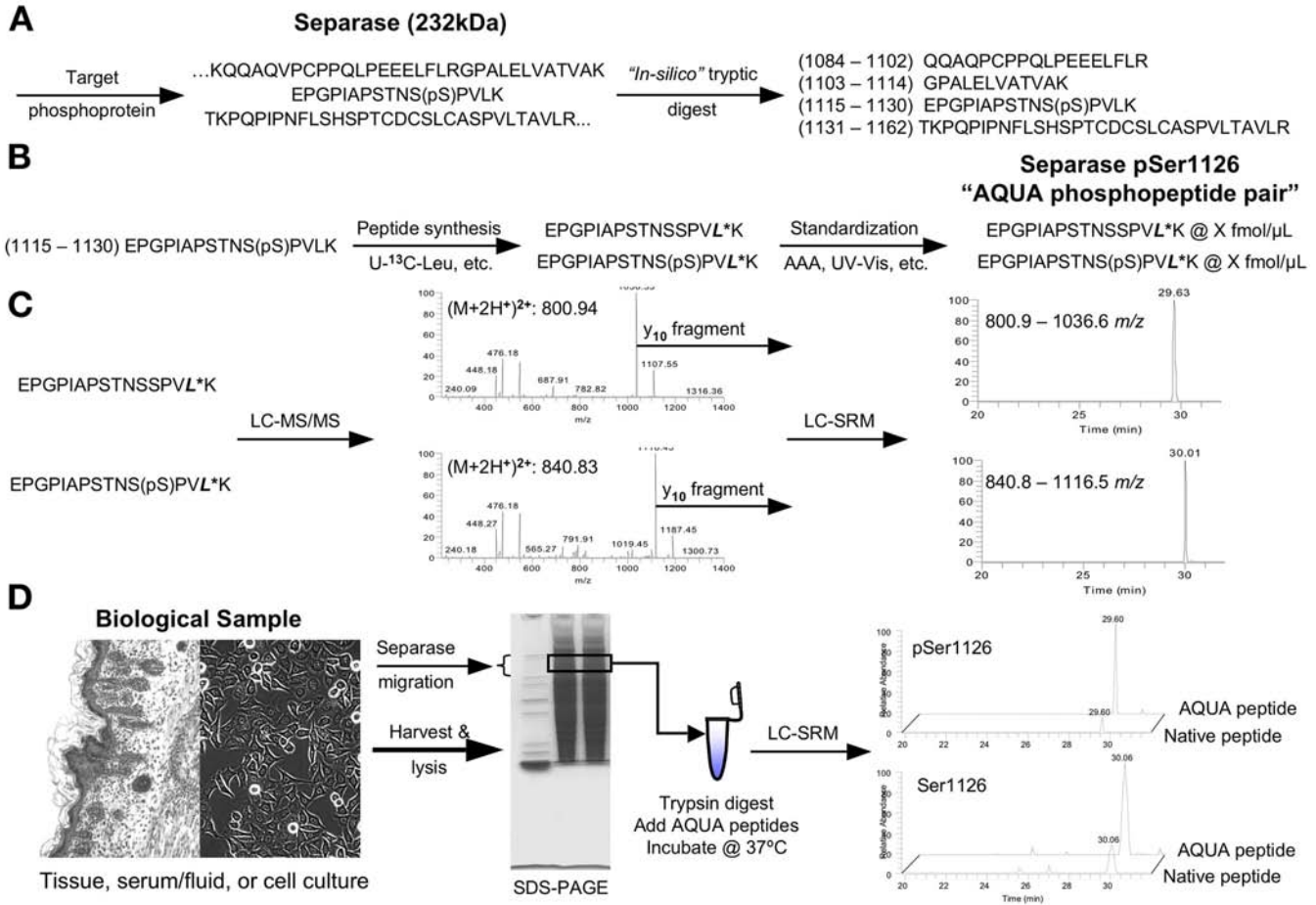

A
Separase (232kDa)
Target phosphoprotein
…KQQAQVPCPPQLPEEELFLRGPALELVATVAK
EPGPIAPSTNS(pS)PVLK
TKPQPIPNFLSHSPTCDCSLCASPVLTAVLR...
"In-silico" tryptic digest
(1084 – 1102) QQAQPCPPQLPEEELFLR
(1103 – 1114) GPALELVATVAK
(1115 – 1130) EPGPIAPSTNS(pS)PVLK
(1131 – 1162) TKPQPIPNFLSHSPTCDCSLCASPVLTAVLR
B
Separase pSer1126
"AQUA phosphopeptide pair"
(1115 – 1130) EPGPIAPSTNS(pS)PVLK
Peptide synthesis
U-13C-Leu, etc.
EPGPIAPSTNSSPVL*K
EPGPIAPSTNS(pS)PVL*K
Standardization
AAA, UV-Vis, etc.
EPGPIAPSTNSSPVL*K @ X fmol/µL
EPGPIAPSTNS(pS)PVL*K @ X fmol/µL
C
EPGPIAPSTNSSPVL*K
EPGPIAPSTNS(pS)PVL*K
LC-MS/MS
(M+2H+)2+: 800.94
y10 fragment
(M+2H+)2+: 840.83
y10 fragment
LC-SRM
800.9 – 1036.6 m/z
840.8 – 1116.5 m/z
Time (min)
D
Biological Sample
Tissue, serum/fluid, or cell culture
Separase migration
Harvest & lysis
SDS-PAGE
Trypsin digest
Add AQUA peptides
Incubate @ 37°C
LC-SRM
pSer1126
Ser1126
AQUA peptide
Native peptide
Time (min)

Fig. 1

7. For the Intavis MultiPep instrument, make 0.23 *M* heavy-isotope monomer stocks in *N*,*N*-dimethylformamide. Store unused stocks as described in **Subheading 3.1.4.**
8. Write a special double-coupling cycle for the MultiPep instrument that delivers 17.5 μL monomer per coupling and uses 25 μL per cycle (as excess volume of amino acid) to prime the delivery line. If only one of the 192 peptides uses the heavy-isotope monomer during one particular synthesis cycle, the total monomer volume consumed is 85 μL; however, if all 192 peptides use the monomer during one cycle, the total volume consumed is 6770 or 35.3 μL per residue. Under these conditions, the heavy-isotope monomer is coupled at twofold molar excess, but it is double-coupled to help drive the reaction to completion and to compensate for any minor reagent delivery errors. Ordinary amino acids are coupled as 17.5 μL of 0.6 *M* stocks (fivefold molar excess) with 50 μL per cycle for priming.
10. At the end of synthesis, peptides are cleaved from peptide resins in an automated manner (Symphony) or manually (MultiPep) and worked up by ether precipitation and washing from trifluoroacetic acid (TFA)-based cleavage solutions. Aliquots of redissolved peptides are analyzed by reversed-phase HPLC and MALDI-TOF mass spectrometry, and the remainder of the peptide is lyophilized and stored at –20°C.
11. AQUA peptide stocks are prepared by dissolving about 1 mg of dry peptide in 0.1 mL of 0.1% TFA and purifying the peptide by reversed-phase HPLC on a 4.5-mm ID column. Fractions are collected using a fraction collector with an automated peak slope detector, so the collector advances as significant changes in absorbance at 214 nm are detected. Adjustments are made to the HPLC gradient according to crude peptide purity, so that the gradient and re-equilibration time for high-quality crude peptides can be as short as 15 min. Each purified peptide is reanalyzed by MALDI-TOF MS and stored without further handling at –20°C in 0.1% TFA and the acetonitrile concentration at which the peptide eluted.

Fig. 1. *(continued from opposite page)* Overview of the absolute quantification (AQUA) strategy. **(A)** The primary sequence of a protein of interest is digested *in silico* to tryptic peptides to evaluate their candidacy for the AQUA process. A section of the sequence surrounding separase serine 1126 is demonstrated as an example. **(B)** The sequences corresponding to the phosphorylated and unphosphorylated tryptic peptides containing serine 1126 are synthesized with an isotopically enriched leucine residue and standardized to a known concentration. **(C)** The resultant AQUA peptide pair is subjected to tandem mass spectrometry (MS/MS) and an informative and sensitive ion fragment is chosen for use in a selected reaction monitoring (SRM) experiment. The actual MS/MS spectra and liquid chromatography (LC)–SRM transitions are shown for the separase pSer1126 AQUA peptide pair. **(D)** A biologically relevant sample is analyzed by lysis, rapid and simple fractionation on an SDS-PAGE gel, excision of the region containing the protein (or proteins) of interest, digestion with a specific protease to generate native peptides, and LC–SRM analysis.

12. Purified peptide stocks are quantified by acid hydrolysis and amino acid analysis using the Waters PicoTag® kit. Each sample is hydrolyzed, derivatized, and HPLC analyzed in duplicate. To be considered a valid measure of peptide concentration, the duplicate analyses must give concentrations that are within 10% of each other, the amino acid monomer amounts must fall in the linear range defined by analysis of a low-concentration and high-concentration amino acid calibration standard, and the measured amino acid composition must fit the expected composition reasonably well.

3.2. HeLa Cell Synchronization

1. The following protocol is a template, designed to yield at least 10^7 cells for four individual cell-cycle timepoints: interphase, prometaphase, anaphase, and telophase. This can be adjusted according to the desired number of cells and/or number of timepoints. The synchronization starts with one confluent 15-cm plate of HeLa cells (roughly $12–15 \times 10^6$ cells), which are plated into four separate dishes, each synchronized independently (*see* **Note 6**).
2. All media, arresting agents, and PBS are warmed to 37°C prior to use.
3. One confluent 15-cm dish of HeLa cells is washed twice with 10 mL of PBS, then trypsinized for 1–3 min with 3 mL of trypsin-EDTA solution at 37°C. The cells are then harvested by washing the plate first with 10 mL and then 5 mL of media, and collecting the media washes containing cells in a sterile 50-mL conical vial. Mix the collected cells thoroughly, and then establish an accurate cell count using a hemacytometer.
4. Plate 3×10^6 HeLa cells on four separate 15-m dishes in a total of 20 mL of DMEM. After 12 h, add 200 µL of thymidine stock to each plate, tilt back-and-forth repeatedly to mix, and return the cells to 37°C incubator.
5. After 18 h, remove media containing thymidine, and wash each plate twice with 10 mL of PBS. Add 20 mL of fresh media to each plate and return the plates to 37°C.
6. 8 h after thymidine release, add 200 µL of thymidine stock to each plate. Return the plates to 37°C.
7. After an additional 18 h, wash each plate twice with 10 mL of PBS. Add 20 mL of fresh media to each plate and return the plates to 37°C.
8. 4 h after the second thymidine release, add 200 µL of the media-diluted nocodazole stock to three of the plates, which are then returned to 37°C. The interphase timepoint is collected by washing the fourth plate twice with PBS, trypsinizing, and harvesting and collected as described in **step 3**. The cells are then washed to remove serum proteins by centrifugation at 1000*g* for 5 min. The media is aspirated to waste and the cells are resuspended in 50 mL PBS and centrifuged again. The PBS is aspirated to waste, then the cells are resuspended in 10 mL of PBS, transferred to a 15-mL conical vial, and a final cell count is established prior to centrifuging a final time. The PBS is removed, and the cell pellet is snap–frozen in liquid nitrogen and stored at –80°C until lysis.
9. After an additional 15 h, mitotic cells from the other three plates are collected by mitotic shake-off. While holding the plate level, rap sharply on an edge of the

plate, while taking care not to spill the contents. After 15 or 20 sharp raps, remove the media to a sterile 50-mL conical vial without further washing the plate. Count cells and split into one- and two-thirds. One-third of the cells are washed as described in **step 8** to remove serum proteins and collected as prometaphase time point. The other two-thirds are released from nocodazole arrest by washing the cells twice as described in **step 5** for a thymidine release, substituting media in place of PBS. After the second wash, resuspend the cells in 40 mL of media, and replate 20 mL of cell suspension to separate clean dishes.

10. After 30 min from nocodazole release, collect the anaphase timepoint.
11. After 120 min from nocodazole release, collect the telophase timepoint.

3.3. Cell Lysis, SDS-PAGE, and In-Gel Digestion

1. Chill 8 mL lysis buffer thoroughly on ice. Add 80 µL of each phosphatase inhibitor working stock solution to the chilled lysis buffer. Immediately prior to thawing the frozen cell pellets, add one tablet of protease inhibitors and sonicate briefly to dissolve.
2. Partially thaw the frozen pellets in a beaker of room temperature water by swirling for 2–3 min. Add 1 mL lysis buffer per 5×10^6 cells—this yields a final protein concentration of roughly 2 mg/mL (about 2 mL lysis buffer total). Sonicate the pellet at 15% power, constant emission, for 30 s, taking care not to allow the tip to get too close to the upper level of liquid as this can cause foaming. Best results are obtained by placing the microtip probe immediately above the cell pellet. Return each lysate to ice for 1 min, with intermittent swirling to ensure even heat dissipation. Repeat the sonication twice to completely lyse cells.
3. Transfer the lysate to duplicate 1.5-mL Eppendorf tubes (~1 mL in each) and centrifuge at 13,000 rpm (16,000 rcF) in a microfuge at 4°C for 20 min.
4. Carefully remove the supernatant to clean Eppendorf tubes. At this point, samples can be aliquoted for archival at –80°C. To do so, snap–freeze each lysate tube in liquid nitrogen before storage.
5. For SDS-PAGE, prepare 2X sample buffer by diluting the 4X LDS-SDS sample buffer 1:1 in purified water and adding DTT to a final concentration of 10 m*M*. Mix the 2X sample buffer with lysate in a 1:1 ratio (200 µL of each). Boil the samples for 3 min and allow to cool to room temperature.
6. Load 40 µL into each well. For each sample or timepoint, load a total of four wells. This yields a final load of each sample of roughly 200 µg on-gel (*see* **Note 7**).
7. On 4–12% bis-Tris gels, using the Invitrogen molecular weight markers, separase runs at a lower molecular weight than expected. Previous reports indicate that the full-length protein runs at or slightly above a 200-kDa (myosin) marker, and an autolytic N-terminal cleavage product (which contains serine 1126) migrates to about 155 kDa ***(6,11,12)***. Using this PAGE system, full-length separase runs at roughly 175 kDa, and the N-terminal fragment runs at roughly 140 kDa. This was confirmed by Western blotting using two different anti-separase antibodies ***(13)***. To ensure collection of both portions of the protein, the entire region between the 120- and 200-kDa markers should be excised.

8. Run the gel for 30 min at 80 V. Turn up the voltage to 120 V until the dye front is roughly half-way down the gel. For this assay, it is only important to separate the 120- to 200-kDa region from all other molecular weights. Partial electrophoresis of the gel reduces the amount of polyacrylamide excised for in-gel digestion (and therefore the trypsin digestion and extraction solution volumes).
9. Remove gel from the loading cassette and stain for 30 min in Coomassie blue dye. Decant the Coomassie stain to the appropriate waste container and add destain solution and a folded Kimwipe. Allow to destain until the molecular weight markers are clearly visible.
10. Clean a fresh scalpel blade with a water-soaked Kimwipe. Be sure to wear powder-free gloves. Cut horizontally across the gel from one 200-kDa marker to the other. Repeat across the gel just above the 120-kDa marker. Cut away excess gel between lanes, dice stained regions into 1-mm cubes, and collect all regions from the same time point into a 1.5-mL Eppendorf tube. *See* **Fig. 1D** for an example of the gel. Repeat for as many time points and gels as necessary.
11. Add 1-mL in-gel destain solution to the pieces and vortex to mix. Incubate at 37°C for 3 h, remove supernatant to waste, and add an additional 1 mL of in-gel destain solution plus 150 µL HPLC-MS grade purified water to each tube. Vortex to mix and incubate at 37°C for an additional 2 h or until completely clear, mixing occasionally.
12. Remove the destain solution to waste. Add 1 mL of acetonitrile to each tube and vortex to mix. Allow tubes to stand for 10 min, followed by vortexing to mix and removal of the acetonitrile. Replace with fresh acetonitrile. Using the tip of a pipet, attempt to disrupt and dislodge the gel aggregate into individual pieces. Allow to stand for an additional 10 min, followed by removal of solvent. Dry in a vacuum centrifuge for 15 min to completely remove traces of acetonitrile.
13. Dilute stock AQUA peptides from stated concentration to 200 fmol/µL in 500 m*M* ammonium bicarbonate. Perform this dilution immediately prior to use. Dilute enough to add 5 µL of this solution to the digest in the next step.
14. Prepare 735 µL of a 50 m*M* solution of ammonium bicarbonate from the concentrated stock and allow to chill completely on ice. If frozen, allow the trypsin stock solution to thaw on ice. Place tubes containing dried gel pieces on ice. Add 60 µL of trypsin stock and 5 µL of the AQUA peptide dilution to the 50 m*M* ammonium bicarbonate solution and vortex well to mix. Add 200 µL of the working trypsin solution to each tube and allow to stand on ice for 20–30 min, or until completely rehydrated. If necessary, add additional 50 m*M* ammonium bicarbonate solution to maintain a slight excess of liquid over the gel pieces, leaving them on ice during the rehydration process (*see* **Note 8**).
15. Once fully rehydrated, allow the samples to digest at 37°C for at least 12 h. After digestion, add 200 µL of peptide extraction solution to each tube, vortex to mix, and allow to stand for 20 min. Centrifuge the tubes in a bench-top microfuge at 13,000 rpm (16,000 rcF) for 30 s, then remove the supernatant to a clean 500-µL Eppendorf tube. Add an additional 150 µL of extraction solution and repeat the vortexing, standing, and centrifugation. Remove the supernatant and combine

with the first extraction. Dry the extraction solutions completely in a vacuum centrifuge. The samples can then be stored at –20°C until ready for analysis.

3.4. Selected Reaction Monitoring Experiment Via Tandem MS

The LC–MS setup described here is intended as a working example only. Changes to the setup may be required for *ad-hoc* instrumentation in different laboratories.

1. Outfit the FAMOS with a 25-µL PEEK loop (250-µm inner diameter × 50-cm length). Place a microtee immediately before the autosampler inlet, and attach a 75-µm ID × 75-cm length piece of fused silica capillary to the microtee to supply the necessary backpressure. Use 100-µm inner diameter PEEK postautosampler to connect the column to the inlet of the mass spectrometer. To minimize analyte band broadening post-column, use as little tubing as possible to connect the LC column to the MS source. If the API inlet has internal fused silica tubing or a stainless steel transfer needle to the electrospray inlet, use as small of an inner diameter as allowed by the manufacturer.
2. Affix nitrogen and argon gas lines to the appropriate inlets. Pressurize per manufacturer's directions.
3. It is important to recognize that individual peptides may require unique tuning parameters. For the separase peptides described here, we were able to optimize the instrument to obtain a maximum signal for both the phosphorylated and unphosphorylated peptides under the same conditions. It should also be noted that, in our hands, nanospray sources require more careful, peptide-dependant tuning optimization than do microspray or standard probe interfaces. Peptide tuning should be performed using identical LC format, but with the column removed and replaced with a standard loop injector mounted through the multiport valve on the front of the mass spectrometer. We typically dilute our AQUA peptides to 1–5 µ*M* in a buffer composition similar to the organic content in their LC elution profile. As an example, we estimate that these separase peptides elute at between 18 and 22% buffer B. We therefore perform all of our tuning with the separase AQUA peptides diluted in 20% buffer B 80% buffer A, to a concentration of 2 µ*M*. If necessary, calibrate the instrument prior to tuning.
4. SRM optimization can be performed either while infusing or in the final LC format. It is quicker to optimize the SRM while infusing, although this expends much more labeled peptide (which may be expensive). When optimizing signal, important parameters to consider are inlet sheath and auxiliary gas flow rates, tube lens and capillary offset voltages, *m/z* maxima for Q1 and Q3 resolution settings (typically 0.7 and 1.0 FWHM, respectively), and Q2 collision gas pressure and energy. In infusion mode, this is performed on-the-fly; when using an LC format, iterative adjustments are evaluated by replicate injections (*see* **Note 9**).
5. Establish that the LC–SRM is working correctly in the LC format if infusion is used in **step 4**. Our LC program is as follows: 0–1 min, 0% B; 1–12 min, 0–40% B; 12–13 min, 40–100% B; 13–18 min, 100% B; 18–19 min, 100–0% B; 19–25

min, 0% B. Use a flow rate such that the pressure at the split is maintained between 90 and 100 bar. Remember to include SRM channels for the native peptide species, as well as the AQUA peptides of interest in the final SRM method.

6. Resuspend each time point in 15 μL of peptide loading buffer and vortex thoroughly to mix. Centrifuge the samples to pellet debris at 13,000 rpm (16,000 rcF) for 15 min in a bench-top microfuge. Using a microbore gel-loader pipet tip, carefully remove the liquid from each timepoint, and place directly at the bottom of a deactivated glass limited volume insert. Place the insert in a 2-mL autosampler vial and seal the vial.
7. Perform the LC-SRM by injecting 12 μL of each sample.
8. Data analysis is performed by plotting the total ion chromatogram of each SRM channel and integrating each peak within identical time limits on data processing software supplied by the instrument manufacturer. Example raw data is depicted in **Fig. 2**. The extent of phosphorylation is expressed as the ratio of phosphorylated analyte to AQUA peptide divided by the sum of the ratios of phosphorylated analyte to AQUA peptide and unphosphorylated analyte to AQUA peptide.

4. Notes

1. Fmoc-protected heavy-isotope amino acid monomers can be purchased from other commercial sources, such as the Isotec Division of Sigma-Aldrich. The monomers we use most frequently are all from Cambridge Isotope Laboratories: L-Leu-N-Fmoc (U-$^{13}C_6$, ^{15}N), cat. no. CNLM-4345; L-Pro-N-Fmoc (U-$^{13}C_5$, ^{15}N), cat. no. CNLM-4347; and L-Val-N-Fmoc (U-$^{13}C_5$, ^{15}N), cat. no. CNLM-4348. For peptides with sequences that do not contain Leu, Pro, or Val, we use L-Asn-N-Fmoc (U-$^{13}C_4$, $^{15}N_2$), cat. no. CNLM-6193; L-Phe-N-Fmoc (U-$^{13}C_9$, ^{15}N), cat. no. CNLM-4362; and L-Tyr-N-Fmoc (U-$^{13}C_9$, ^{15}N), cat. no. CNLM-4349. As a last resort we will incorporate two heavy-isotope monomers of Ala or Gly in any combination per peptide: L-Ala-N-Fmoc (U-$^{13}C_3$, ^{15}N), cat. no. CNLM-4355 and L-Gly-N-Fmoc (U-$^{13}C_2$, ^{15}N), cat. no. CNLM-4357.
2. Many laboratories have access to conventional peptide synthesis instruments through either departmental resources or core facilities, e.g., the Rainin/Protein Technologies Symphony instrument. This instrument makes 12 peptides at once at 25- to 100-μmol scale, which corresponds to about 40–175 mg of a 15-residue peptide. Because amino acid monomers are usually coupled at a fivefold molar excess, a 25-μmole scale synthesis would require 125 μmol of heavy-isotope monomer or 45 mg of heavy-isotope L-Leu-N-Fmoc. Therefore, the lowest synthesis scale possible with these instruments would be wasteful with respect to both peptide amount and usage of heavy-isotope monomer. This is easily addressed by (1) using the instrument to synthesize just 5 μmol of peptide and (2) manually delivering heavy-isotope monomer to peptide resin through a reaction vessel sidearm.
3. A new development in peptide synthesis instrumentation is the availability of high-throughput, small-scale synthesis instruments, for example, the Intavis MultiPep (*see* http://www.intavis.com/multipep.html). These highly specialized

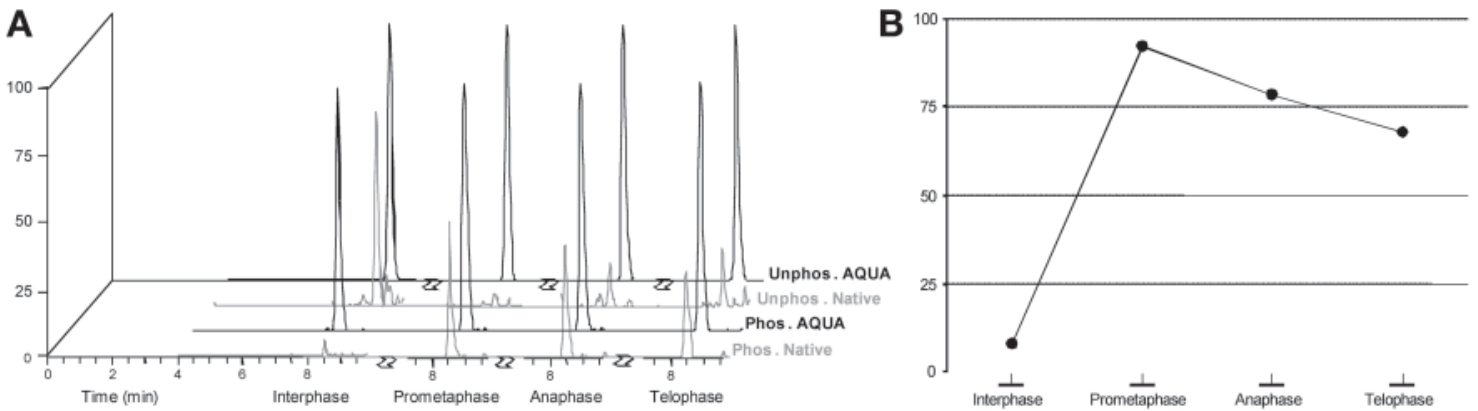


Fig. 2. Liquid chromatography-selected reaction monitoring (LC–SRM) analysis of cell-cycle, stage-specific phosphorylation of separase at serine 1126. **(A)** Absolute quantification (AQUA) and native LC–SRM channels for the phosphorylated and unphosphorylated tryptic peptides containing serine 1126 at interphase, prometaphase, anaphase, and telophase. The individual ion chromatograms are normalized to the intensity of each AQUA peptide. Note the relative decrease in intensity of the phosphorylated peptide signal after release from nocodazole, and the concomitant increase in signal for its unphosphorylated cognate. **(B)** The peak areas for each species in **A** is transformed into percent phosphorylation values at each time point.

instruments are essentially low-volume liquid handlers and not yet widely used. The MultiPep can make up to 192 peptides at once at 1- to 5-μmol scale in a microtiter-plate format. We use this instrument to make AQUA peptides at 2-μmol scale, which corresponds to about 3.5 mg of a 15-residue peptide. In contrast to the approach previously described for conventional instruments, all reagents are used in proportion to the peptide synthesis resin and peptides are synthesized in a fully automated manner.

4. Remove monomer stocks from the freezer at least 1 h before use, to avoid contaminating the stock with condensation.
5. The lower limit for the Symphony synthesis scale is set by the rather large size of the reaction vessel (10 mL). Our experience is that if resin amounts below 5 μmol are used with these reaction vials, much of the resin is lost during synthesis, cleavage, and handling. On the other hand, the much smaller volume of the MultiPep reaction wells (0.6 mL) gives good peptide yields for scales as low as 2 μmol. The waste associated with priming the MultiPep delivery line (vs wasteless manual addition of heavy-isotope monomer with the Symphony) has to be weighed against the benefit of fully unattended, automated synthesis. The Symphony uses 25 μmol of heavy-isotope monomer to make 5 μmol of peptides (fivefold) vs 8 μmol monomer (plus waste) for 2 μmol peptides on the MultiPep (at least fourfold).
6. The synchronization consists of two sequential thymidine blocks to generate coarse synchronicity entering G_2/M. The cells are then arrested in prometaphase with nocodazole, which can be washed out to allow progression through mitosis. The synchronization process requires manipulation of the cells over the course of 3 d; prepare a workflow schedule in advance to avoid conflicts.
7. Load molecular weight markers on both sides of the gel. This is important, as the markers act as rulers to determine the regions of gel to be excised. As the gel can smile, lean, or list to one side or the other, using markers at each side can ameliorate these problems.
8. It is important to establish a trypsin titration curve for the system of interest ***(14)***. A constant amount of AQUA peptide can be added to serial dilutions of trypsin to determine a concentration that will adequately digest the sample. Note that the required amount of trypsin in an actual experiment may be greater than that for purified protein because of the presence of many other, possibly more abundant, proteins in the excised section of gel. In this regard, adding the AQUA peptide to the digest to mimic extraction efficiency is also important to avoid underestimating the amount of target.
9. In choosing a fragment ion to monitor, we have generally found that intense, singly charged fragments of higher *m/z* than that of the parent yield the cleanest SRM traces, as they reduce noise contributions from singly charged species in complex samples. In addition, a wider dynamic range between the analyte and AQUA peptide can be achieved when the monitored fragment ion contains the labeled residue. Of course, the exact nature and complexity of the sample to be probed should dictate the SRM in the final analysis. For example, it has occurred

that a low-intensity SRM transition (at one-tenth the signal relative to a more intense fragment transition) for purified AQUA peptides yielded greater signal-to-noise in the final AQUA experiment resulting from less interference in that particular SRM channel. It is of course possible to monitor multiple fragment ions per parent (multiple reaction monitoring [MRM]). However, when starting with the most intense MS/MS fragment in a single-channel SRM experiment, the maximum absolute signal improvement in a two-channel SRM can be not more than twice the single-channel result. Also, each additional channel contains idiosyncratic chemical noise that must be evaluated, as well as reduces the duty cycle of the SRM series. In our hands, it is much more informative to split a sample in half and perform two separate AQUA experiments that monitor different, single-channel SRM transitions than to analyze a single sample using an MRM.

Acknowledgments

This work was supported by grants from the NIH (HG3456 and GM67945).

References

1. de Leenheer, A. P. and Thienpont, L. M. (1992) Applications of isotope-dilution mass-spectrometry in clinical-chemistry, pharmacokinetics, and toxicology. *Mass Spectrom. Reviews* **11,** 249–307.
2. Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422,** 198–207.
3. Tao, W. A. and Aebersold, R. (2003) Advances in quantitative proteomics via stable isotope tagging and mass spectrometry. *Curr. Opin. Biotechnol.* **14,** 110–118.
4. Sechi, S. and Oda, Y. (2003) Quantitative proteomics using mass spectrometry. *Curr. Opin. Chem. Biol.* **7,** 70–77.
5. Gerber, S. A., Rush, J., Stemmann, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. USA* **100,** 6940–6945.
6. Stemmann, O., Zou, H., Gerber, S. A., Gygi, S. P., and Kirschner, M. W. (2001) Dual inhibition of sister chromatid separation at metaphase. *Cell* **107,** 715–726.
7. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17,** 994–999.
8. Ong, S. E., Blagoev, B., Kratchmarova, I., et al. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1,** 376–386
9. Kirkpatrick, D., Gerber, S. A., and Gygi, S. P. (2005) The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* **35,** 265–273.
10. Broenstrup, M. (2005) Absolute quantification strategies in proteomics based on mass spectrometry. *Expert Rev. Proteomics* **1,** 503–512.
11. Zou, H., McGarry, T. J., Bernal, T., and Kirschner, M. W. (1999) Identification of a vertebrate sister-chromatid separation involved in transformation and tumorigenesis. *Science* **285,** 418–422.

12. Waizenegger, I., Gimenez-Abian, J. F., Wernic, D., and Peters, J. M. (2002) Regulation of human separase by securin binding and autocleavage. *Cell* **12,** 1368–1378.
13. Chestukhin, A. and DeCaprio, J. A. (2003) Western blot screening for monoclonal antibodies against human separase. *J. Immunol. Methods* **274,** 105–113.
14. Havlis, J. and Shevchenko, A. (2004) Absolute quantification of proteins in solutions and in polyacrylamide gels by mass spectrometry. *Anal. Chem.* **76,** 3029–3036.

6

Quantification of Proteins and Metabolites by Mass Spectrometry Without Isotopic Labeling

Sushmita Mimi Roy and Christopher H. Becker

Summary

We demonstrate the quantification capability and robustness of a new integrated liquid chromatography–mass spectrometry (LC–MS) approach for large-scale profiling of proteins and metabolites. This approach to determine differential expression relies on linearity of signal vs molecular concentration using electrospray ionization LC–MS, reproducibility of sample processing, a novel normalization strategy and associated data analysis software. No isotopic tagging or spiking of internal standards is required. The method is general and applicable to the proteome and metabolome from all biological fluids and tissues. Small or large numbers of samples can be profiled in a single experiment. Differential profiling of 6000 molecular ions per sample by one-dimensional chromatography LC–MS and 30,000 molecular ions per sample by two-dimensional chromatography LC–MS is demonstrated using rheumatoid arthritis patient samples compared with control samples. A new approach to peptide identification is described that involves building libraries of previously identified peptides, circumventing the need to acquire MS/MS data during profiling. Robustness of the platform was tested by repeating sample preparation and LC–MS differential expression analysis after 10 mo, using independent serum aliquots stored at –80°C. To the best of our knowledge, this is the first demonstration of long-term robustness of a platform for quantitative proteomics and metabolomics.

Key Words: Proteomics; metabolomics; quantification; quantitative mass spectrometry; protein identification; differential profiling; differential expression.

1. Introduction

It is anticipated that the ability to quantify changes in the proteome and metabolome will lead to the discovery of useful biomarkers, a greater under-

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

standing of underlying biology and, thus, allow more efficient drug discovery ***(1,2)***. The usefulness of such technology strongly relies on the ability to perform reliable and quantitative differential profiling, as well as properly identify proteins and metabolites. Differential quantification of thousands of peptides and metabolites must be highly sensitive while covering a large dynamic range, and be amenable to analysis of a variety of biological fluids or tissues. In addition, the methods must be sufficiently robust and stable to enable comparisons over a statistically and biologically significant number of samples.

We present a method to differentially quantify polypeptides and metabolites in up to hundreds of samples. Furthermore, the approach uses no isotopic tagging or labeling, nor the spiking of standards. This method relies on the long-term reproducibility of signal and the linearity of signal vs molecular concentration. The approach is undirected, comprehensive, and equally applicable to proteins and metabolites, as well as to all body fluids and tissues. Molecular ion intensities are directly measured for each sample and are globally normalized to all other samples in the study.

Data analysis has been made possible by developing and employing a computer application, the MassView™ software, which deisotopes and tracks molecular ions, determines monoisotopic mass from a resolved isotopic envelope, corrects any shifts in retention times between liquid chromatography–mass spectrometry (LC–MS) runs, and performs normalization by employing signals of molecules that do not change concentration from sample to sample ***(3)***. For each sample, approx 6000 molecular ions are currently differentially quantified in the human serum proteome using one-dimensional (1D) separation. More than 2000 molecular ions in the serum metabolome are tracked and quantified. These comparisons have been made successfully in single-study sizes of over several hundred samples.

This approach has recently been applied successfully to relative quantification of protein phosphorylation stoichiometry ***(4)***. Although the 1D LC–MS profile shows clear differences between the disease and control cohort, more information can be gained by extending the dynamic range of the proteome measurements ***(5,6)***. We have recently developed a method for quantification of approx 30,000 molecular ions by employing two-dimensional (2D) chromatography. Access to a greater dynamic range is achieved by off-line and orthogonal fractionation of the digested serum proteins into approximately eight fractions using strong cation exchange (SCX) chromatography. Each of the resulting fractions is individually subjected to (on-line) 1D LC–MS. These methods are also described.

Research in the field of quantitative MS has thus far led to technologies that enable differential profiling (*see* Chapters 2–4). Stable isotope tagging methods for relative quantification have been extended to serve quantification of

peptides by subjecting a specific amino acid in the proteome sample to mass labeling chemistry by tagging or metabolic incorporation ***(7,8)***. Although these approaches have been used successfully, disadvantages include reagent expense, additional processing steps, and required presence of a specific amino acid ***(9)***. Because metabolites lack a common tagging site, a general labeling approach cannot be used for metabolomics. It has been speculated that a comprehensive set of spiked standards can be synthesized for quantification ***(10)***. However, the complexity of the sample makes spiking of large numbers of known standards a challenging approach.

In addition to differential expression quantification by 1D and 2D LC–MS, we present a new approach to large-scale identification of peptides/proteins and metabolites, regardless of their degree of differential expression. This approach is based on building a library of observed biomolecules for any given species and sample type, where the identity of the peptide is stored along with its accurate *m/z*, charge state, and retention time. Having identical sample processing and LC for the separate profiling and identification processes is key. After quantification of differential profiling data, the library is used to match the *m/z*, charge state, and retention time of profiled peaks. The same human serum peptide library, for example, may be used every time a human serum differential profiling study is conducted. For molecular ions that are changing significantly, but are not yet successfully matched in this process, additional tandem MS (MS/MS) experiments are conducted. The new directed identification results are added to the library, making the library more comprehensive with time. This method has been successfully used to identify significantly changing proteins and metabolites in serum, cerebrospinal fluid, and urine as well as tissues from human, primate, dog, rat, and mouse.

In this chapter, we describe our quantification and identification methods, and show results from 1D and 2D LC coupled to MS. Examples are shown comparing sera from individuals with rheumatoid arthritis (RA) to individuals from a control group. Finally, the reproducibility and robustness of the entire platform is demonstrated by repeating an entire study using 1D LC–MS differential profiling of separate frozen serum aliquots 10 mo after the first study. We suggest that such experiments be carried out to validate a given quantitative MS platform prior to its use in studies of biological significance.

2. Materials

2.1. Serum Samples

1. Pooled human serum for proteomics was purchased from Sigma-Aldrich (St. Louis, MO).
2. For the RA, serum was collected from patients diagnosed with different degrees of RA, as well as individuals with no symptoms of RA serving as controls. The

handling of these biological materials must be performed in accordance with US Department of Health and Human Services guidelines for level 2 laboratory biosafety as found in **ref. *11***.

2.2. Sample Preparation

1. The 5-kDa molecular weight cut-off spin filter was the model Centricon Plus-20 (Millipore Corp., Bedford, MA).
2. For depletion of the top six abundant proteins (serum albumin, immunoglobulin [Ig] G, IgA, transferrin, haptoglobin, and antitrypsin) the antibody-based Multiple Affinity Removal System (Agilent, Palo Alto, CA) has been used effectively. In earlier studies, affinity depletion was by use of small molecule-coated beads purchased from Prometic Biosciences (Cambridge, UK) for human serum albumin and IgG removal.
3. Modified trypsin (Promega Corp., Madison, WI) was used at 1% weight equivalence of the proteins. Sample protein concentrations are determined via the Bradford assay (Pierce Biotechnology, Inc., Rockford, IL).
4. For desalting, a C-18 SPE cartridge (Sep-Pak, Waters Corporation, Milford, MA) was used. The washing was performed with 6 mL of 0.1% formic acid in water and the elution with 1.5 mL of a solution containing 90% acetonitrile and 0.1% formic acid.
5. Desalting steps were automated using the RapidTrace system from Caliper Life Sciences, Inc. (Hopkinton, MA).
6. For SCX chromatography, a Keystone BioBasic, 250 × 2.1-mm column from Thermo Electron Corp. (Waltham, MA) was used.
7. Antibody column for depletion of abundant proteins from human serum was the MARS column from Agilent.
8. 0.2-μm nylon filter was a Spin-X centrifuge tube from Fisher Scientific (Hampton, NH).
9. 25 m*M* phosphate buffer, pH 6.0 is prepared with a 1000-mL beaker with 0.8 g dibasic sodium phosphate heptahydrate and 3.0 g monobasic sodium phosphate monohydrate. Volume is brought up to 950 mL with high-performance liquid chromatography (HPLC)-grade water while stirring. With a pH meter, adjust pH to 6.0 with droplets of 6 *N* NaOH while stirring.
10. 6 *M* guanidine hydrochloride: 576 g guanidine hydrochloride plus 12.1 g Tris base (2-amino-2-[hydroxymethyl]-1,3-propanediol) in a 1000-mL beaker, with the addition of HPLC water to the 950-mL level, with stirring for dissolution. Using a pH meter, adjust pH to 8.3 by titration with droplets of concentrated HCl while stirring. Bring the volume to 1000 mL with additional HPLC-grade water. Filter. Shelf life is 1 mo at room temperature.
11. 10 m*M* dithiothreitol (DTT): 0.16 g of DTT (white powder) is transferred to a 1.5-mL tube to which 1.0 mL of HPLC-grade water is added, then vortexed to dissolve completely. Store at –20°C when not in use. Shelf life is 3 mo frozen.
12. 25 m*M* iodoacetic acid: 0.19 g of iodoacetic acid (yellow powder) is transferred to a 1.5-mL tube to which 1 mL of 1 *M* NaOH is then added and vortexed to dissolve completely. Make fresh and discard after use. Shelf life is 1 d.

13. 50 m*M* $(NH_4)_2CO_3$ buffer at pH 8.3: transfer 3.95 g of $(NH_4)_2CO_3$ (white powder) to a 1000-mL beaker, then add HPLC-grade water to the 950-mL mark, stirring to dissolve. Then with a pH meter, titrate solution to pH 8.3 by addition of drops of 28% ammonium hydroxide while stirring. Bring the total volume to 1000 mL by addition of HPLC-grade water. Filter the solution. Store at room temperature. Shelf life is 2 wk.
14. Preparation of SCX buffers A and B begins with preparation of a concentrated potassium phosphate solution. This concentrated solution is made in a 1000-mL bottle with 54.4 g of potassium phosphate monobasic, plus HPLC-grade water and 3.0 mL of concentrated HCL. Mix well, and add additional water to 1000 mL. Filter.
    a. For buffer A, take one part of this concentrated solution and add to five parts HPLC-grade acetonitrile plus 14 parts HPLC-grade water, and mix well. Store at room temperature with a 2-wk shelf life.
    b. For buffer B, take one part of the concentrated phosphate solution and add to five parts acetonitrile, 10 parts water, and 4 parts of filtered 2.5 *M* KCl; mix well. Store at room temperature with a 2-wk shelf life.
15. All other general reagents were purchased either from Fisher Scientific or VWR Scientific (West Chester, PA).

2.3. LC–MS

1. A binary Agilent capillary 1100 series HPLC was used for separation on-line with the MS.
2. Reverse-phase capillary columns (C_{18}; 320 μm × 15 cm) were used for LC–MS (Micro-Tech Scientific, Inc., Vista, CA).
3. Electrospray ionization-time-of-flight (ESI-TOF™) mass spectrometer used for profiling was a microTOF™ (from Bruker Daltonics, Billerica, MA) or LCT™ (from Micromass, Waters Corp).
4. A quadrupole TOF mass spectrometer (microQTOF™ was used for identification by MS/MS (from Micromass, Waters Corp).
5. All peptide identification searches were conducted using Mascot Search Engine (Matrix Science, London, UK).

3. Methods

Figure 1 gives an overview of the LC–MS platform developed and a typical workflow for quantitative proteomics without tagging or spiking. This includes sample preparation, LC–MS analysis, data analysis for differential quantification, statistics (based on study design), and identification of components (tracked molecular ions) by library construction as well as directed MS/MS.

3.1. Sample Preparation (1D LC–MS)

1. 1 mL serum was fractioned into serum proteome and serum metabolome using a 5-kDa molecular weight cut-off spin filter.

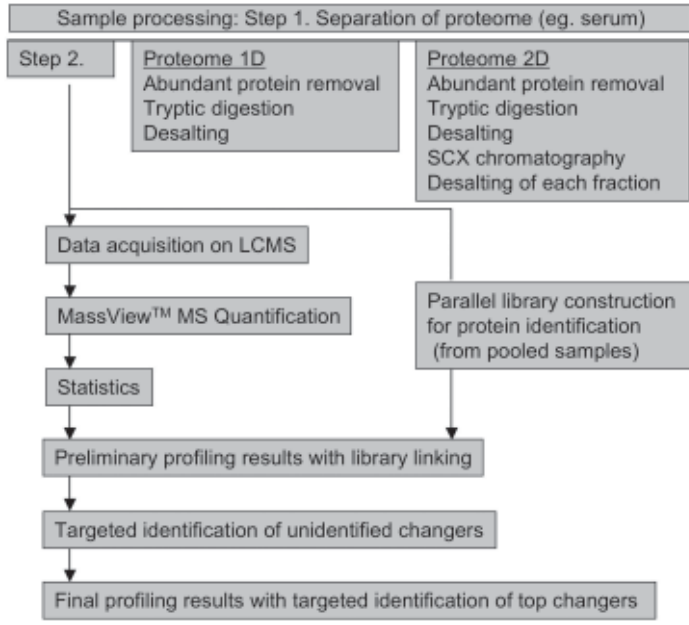


Fig. 1. Differential profiling workflow. The major steps of sample processing, data acquisition, quantification, and identification are shown for one-dimensional and two-dimensional liquid chromatography–mass spectrometry.

2. 25 µL of the high molecular-weight fraction (serum proteome) was diluted by addition of 425 µL of 25 m*M* phosphate buffer (pH 6.0) before it was applied to affinity beads for human serum albumin and IgG removal. After a 10-min incubation in a 0.2-µm nylon filter tube, the sample was spun to recover depleted serum. Alternatively, in most recent experiments, the use of affinity beads for albumin and IgG removal has been supplanted by use of an antibody column for the most abundant proteins in serum; for humans this is albumin, IgG, IgA, α-1-antitrypsin, transferrin, and haptoglobin.
3. The albumin- and IgG-depleted serum proteome was denatured by addition of 1 mL of 6 *M* guanidine hydrochloride, reduced by addition of 10 µL of 10 m*M* DTT and alkylated with 25 µL of 25 m*M* iodoacetic acid at room temperature.
4. The denaturant and reduction–alkylation reagents were removed from the mixtures by buffer exchange against 50 m*M* $(NH_4)_2CO_3$ at pH 8.3 using a 5-kDa molecular weight cut-off spin filter.
5. Modified trypsin was then added to the mixtures with incubation at 37°C for 14 h. A total of approx 20 µg of processed peptides in a 20-µL volume of 0.1% formic acid water was injected into the LC–MS. This amount of material is equivalent to that in about 1.5 µL of serum starting fluid.

6. Samples from each of the two cohorts were processed and analyzed in pairs. The order of each pair was randomized over the study in order to minimize skewing because of sample processing or instrument variability over time. For example, for cohorts A and B, the run order was AB, BA, AB, BA and so on. Pairs were matched by minimizing the sex and age of candidates.

3.2. LC–MS

1. All LC–MS profiles were run at a flow rate of 8 μL/min.
2. Gradient elution of the proteome sample was achieved using 100% solvent A (0.1% formic acid in H_2O) to 40% solvent B (0.1% formic acid in acetonitrile) over 100 min.
3. The throughput for proteome analysis was 50 samples per week per instrument.

3.3. Quantification

1. All data analysis uses the MassView software developed in-house ***(3,12–14)***; this software is not available for sale at this time, but similar software packages are being developed, for example, the MarkerView™ Software (MDS Sciex, Concord, Ontario, Canada). Data stored as a list of peaks for every scan undergoes baseline subtraction, smoothing, and deisotoping. As part of deisotoping, an isotopic pattern assignment is made along with charge state and accurate mass determination. (*See* **ref. *3*** for further information on the software methods employed.)
2. Baseline correction, smoothing, deisotoping, and application of a modest threshold ensure that all the signals being tracked have substantial ion counts.
3. A chromatographic peak is then built by linking together a series of consecutive scans that contain a signal above threshold at a given *m/z* with an uncertainty window of +/– 0.05–0.10 Da.
4. A list of deisotoped peaks is obtained for a given LC–MS run, each peak distinguished by its characteristic monoisotopic *m/z*, retention time, charge state, and maximum intensity.
5. Retention times of components for each file are mapped onto a reference file using a nonlinear mathematical function that allows for minor local shifts, contractions, and expansions in the chromatographic time base. The algorithm was earlier developed for speech recognition ***(15)***.
6. A global intensity normalization is performed by choosing one file as a reference and normalizing all other files one at a time. The single normalization constant for each file is taken as the median of the ratios of intensities for all components between the file in question and the reference file. Thus, the method relies on employing signals of molecules that do not substantially change concentration from sample to sample. In this way, differences in sample concentrations and/or instrument response over time are taken into account.
7. Peaks within user-adjustable *m/z* and retention time windows are then correlated between all samples by a process called “component building.” Typically a com-

ponent (tracked and quantified molecular ion) that is observed in a minimum of 25% of all samples is qualified for final differential quantification. This number can be modified depending on the study requirement.

8. Finally a list of all molecular ions profiled in the study is obtained, each associated with its characteristic monoisotopic *m/z*, retention time, charge state, and relative intensity in all samples quantified.

3.4. Validation of Sample Processing and LC–MS Reproducibility

A 1D LC–MS proteome sample obtained from injecting 20 μL (approx 20 μg) of processed peptides in 0.1% formic acid water derived from the equivalent of a starting 1.5 μL of serum, after depletion of serum albumin and IgG, displays over 6000 molecular ions (not counting isotopes). Because this method relies on reproducibility, sample processing variability and the variability in LC–MS (from chromatographic separation, sample injection volumes, ion suppression, transmission, and detection in the mass spectrometer) should both be measured (**Fig. 2**). This type of measurement is an essential part of setting up a quantitative proteomics platform and should be followed as described.

1. To measure LC–MS variance independent of sample preparation, 20 human serum samples were pooled after independently being prepared, realiquoted, and analyzed on the LC–MS platform. A median coefficient of variance (CV) of 15.1% was measured for the approx 5000 (deisotoped) molecular ions measured per run in each of these 20 samples. The distribution of CVs is shown in **Fig. 2A**.
2. To measure the variability in the platform independently of biological variability (i.e., variability in sample preparation plus LC–MS measurement), 20 aliquots of the same, pooled human serum were individually prepared in parallel and individually subjected to LC–MS analysis and differential quantification. A median CV of 25.4% was measured for the entire platform (**Fig. 2B**).
3. Using the already obtained values for variance of the total platform, and that of the LC–MS or instrument alone, the median CV of sample preparation for the proteome was calculated to be 20.4% ***(12)***, based on a sum of the square of the means.
4. These experiments have been used to formulate an error model for the platform ***(12)*** and to calculate the biological variability within a normal and a diseased group ***(13)***.
5. If the median CVs for the LC–MS are significantly higher than those demonstrated here, this indicates lack of stability of one or more aspects of the LC–MS set-up. Stability of the HPLC pump, column performance as well as stability of the mass spectrometer signal intensity over 20 or more sample injections, are crucial.
6. If the median CV for sample preparation is significantly higher than demonstrated here, this indicates greater need for standardizing sample collection methods, operating procedures, reagent preparation methods, and any automating steps, as well as training laboratory technicians to operate identically.

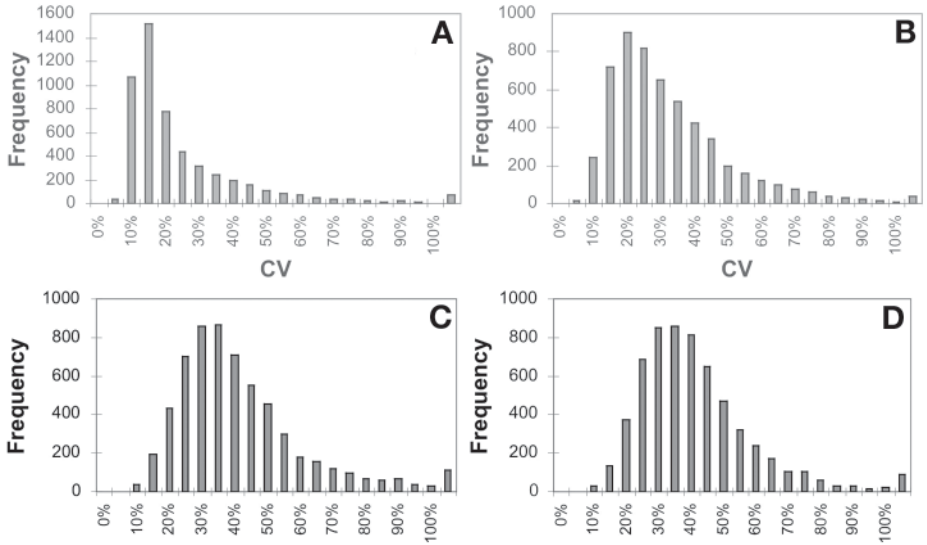


Fig. 2. The frequency distribution of coefficients of variation (CVs) for **(A)** serum pooled after processing before injection into the liquid chromatography–mass spectrometry (LC–MS), **(B)** individually processed serum with no pooling, and a typical one-dimensional LC–MS study with **(C)** normal and **(D)** diseased populations. Each panel reports the data from 20 samples and about 5000 molecular ions per sample (not counting isotopes). For case **(A)** the median CV is 15.1%, whereas for case **(B)** it is 25.4%. It can then be calculated that the median CV for sample preparation of the proteome is about 20%. Histograms of CVs for the proteome from healthy individuals **(C)** and individuals diagnosed with rheumatoid arthritis **(D)** result in a median CV of 33.8 and 35.0%, respectively.

7. These median CVs can be further reduced as evidenced by very recent work in our laboratory.

3.5. Statistics

1. Differential quantification is performed by comparing the intensity of the components between groups or pairs of individuals using a standard two-sided t-test or a nonparametric test, as appropriate. For three or more group comparisons, ANOVA tests may be used ***(16)***.
2. A standard deviation can easily be calculated for each component. The variance of this measurement is defined as the square of the standard deviation, and the coefficient of variance is the ratio of the standard deviation divided by the mean value for a given group ***(16)***.
3. The significance of any observed change can be determined by its *p*-value and significantly changing molecules that are not already identified in previous library construction can be marked for directed identification efforts.

4. A variety of additional data mining algorithms can be applied as appropriate for the study.

3.6. Differential Profiling by 1D LC–MS: Example

SurroMed is undertaking a longitudinal RA study that will measure disease progression by metabolic and proteomic profiling using this mass spectrometric method. Some initial results comparing a subset of individuals enrolled in this study who are healthy (controls) to those diagnosed with RA are presented. This method may be followed for differential profiling in any biological context.

1. Sera from 19 selected RA patients and 19 age- and sex-matched healthy controls were prepared as described before in **Subheading 2.2.**
2. The MassView software platform was able to distinguish isotopic patterns, build peak lists for a given sample, build components by connecting these peaks between all 38 samples, and quantify each observed component (deisotoped molecular ion) in all samples being studied.
3. More than 6000 components/molecular ions were quantified in the proteome for each of these 38 samples. Each molecular ion was quantified by using its signal intensity at the maximum of elution.
4. The distribution of CVs for the proteome for both the control and RA groups is shown in **Figs. 2C** and **D**, respectively. These CVs have contributions from biological variation, as well as processing and instrument variation. The median CV of the distribution is 35.5% for the RA group and 34.6% for the 19 normal individuals.

3.7. 2D Sample Preparation

1. For 2D LC–MS, 50 μL of serum was depleted of human serum albumin (HSA) and IgG and digested, following the methods of **Subheading 3.1.**
2. Depleted, digested serum was fractionated off-line into eight fractions on a SCX column. The gradient of SCX buffers A and B was adjusted so that each fraction would contain a similar quantity of protein. The number of fractions collected may be varied depending on the scope of the project and material availability.
3. All fractions were lyophilized to dryness.
4. Each fraction was redissolved in 5 mL of water acidified to a final concentration of 1.0% formic acid and desalted with a C-18 SPE cartridge, eluting with 90% acetonitrile-water-0.1% formic acid. These samples were then dried again.

3.8. 2D LC–MS

1. The final dimension of the 2D LC–MS set-up is identical to the 1D LC set-up, using on-line reverse-phase capillary HPLC coupled with ESI-TOF read-out. Samples for each fraction from all individuals were analyzed.
2. Desalted samples from the first chromatographic dimension were each dissolved in 40 μL of 0.1% formic acid, and 20 μL was then injected into the LC–MS at approx 20 μg per injection.

3. For 2D LC, the order in which fractions were run was carefully considered. The first fraction of all samples were run back-to-back on the LC–MS to ensure minimal deviation in recorded *m/z*, retention times, and detector sensitivity within a fraction. Remaining fractions were run in a similar fashion. Run order within a given fraction was determined as described for 1D LC–MS, with pairing between individual samples from the two cohorts (**Subheading 3.1.**, **step 6**).

3.9. 2D Quantification

1. For 2D LC, data from all samples of a given fraction are grouped and treated as a 1D study. All peaks are differentially quantified, fraction by fraction, with all the data from each fraction being treated independently of the other fractions.
2. After this fraction-wise quantification, the list of all components for the eight fractions within a given sample are collated.
3. Next, those molecular components between neighboring fractions merge if they have essentially the same *m/z* and reverse-phase chromatographic retention time (recognizing two such fraction-based components as a single component). Any components eluted in adjacent fractions within 0.05 *m/z*, the same charge state, and 1 min retention time are considered the same component and their intensities are added.

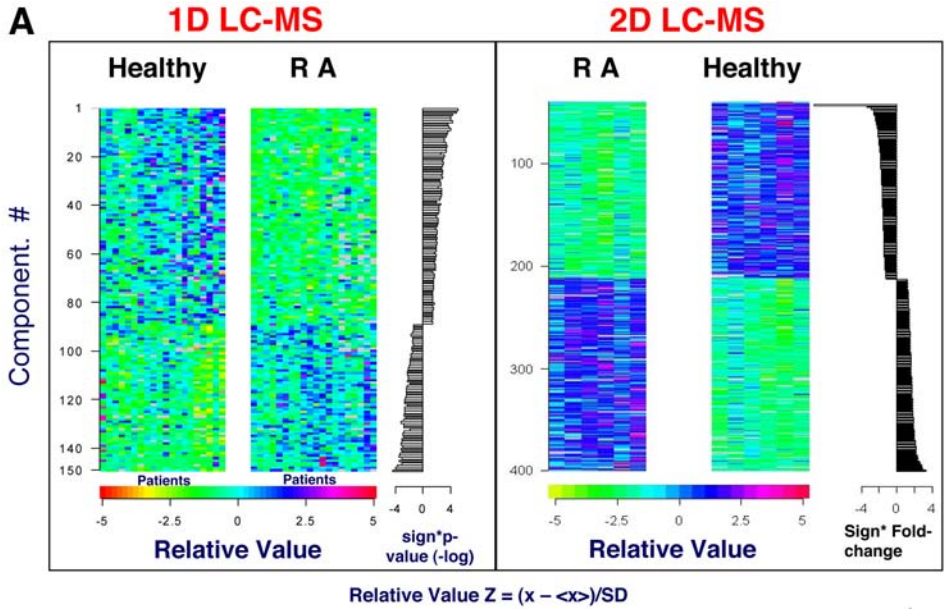
3.10. Differential Profiling by 2D LC–MS: Example

1. To compare the profiles of the RA and control group using 2D LC–MS, six samples were picked from each group and subjected to this mass spectrometric platform.
2. The great increase in information content of 2D LC–MS compared with 1D LC–MS is shown in **Fig. 3A**. In this figure, 38 individual profiles were used in the 1D study and 12 individual profiles in the 2D study. Although many more samples were profiled in the 1D study, the number of significant changes observed is far greater in the 2D study. This shows the advantage of using a larger amount of starting serum and adding an orthogonal chromatography step. The same MassView software was used for both analyses.
3. The extent of significant changes observed in the 1D and 2D profile data was also compared (**Fig. 3B**) for the same 12 samples (out of a total of 38 samples profiled by 1D LC–MS). Only 23 significant changes were found ($p < 0.005$) between the two cohorts by 1D LC–MS. In comparison, the 2D method was able to uncover 364 molecular ions that were significantly different between the two groups ($p < 0.005$).

3.11. Other Dimensions of Separation

This quantification method described in **Subheading 3.8.**, can be expanded or modified to incorporate other separations. These could include, for example, alternative 1D chromatography for the 2D LC–MS approach, adding a third dimension, or employing subcellular fractionation when working with tissues.

A



B

Number of Significant Changes: 1D vs. 2D LC-MS		
Method (# subjects)	**1D LC-MS (12)**	**2D LC-MS (12)**
CV (Ave) Rh. Arthritis	**31.3%**	**30.8%**
CV (Ave) Normal	**30.2%**	**29.6%**
CV (Median) Rh. Arthritis	**27.9%**	**26.3%**
CV (Median) Normal	**28.0%**	**25.7%**
No. of Subjects	**12**	**12**
P<0.0001	0	8
P<0.001	1	88
P<0.005	23	364
P<0.01	59	697
P<0.05	290	2651
No. of Components	**2452**	**34,380**

Fig. 3 **(A)** Comparison of proteome difference maps generated using the one-dimensional (1D) liquid chromatography–mass spectrometry (LC–MS) (6000 molecular components, 19 individuals per group) and two-dimensional (2D) LC–MS (34,000 molecular components, six individuals per group). Significant changes ($p < 0.005$) are shown for patients diagnosed with rheumatoid arthritis (RA) and healthy individuals. Cells in one row correspond to the same component. Columns correspond to individuals. Each cell is assigned a color corresponding to its Z-score. The Z-parameter is defined as the dif-

3.12. Protein Identification

For protein identification, a unique strategy is used so that all samples do not require analysis by MS/MS.

1. First, in the case of human serum or plasma, a library consisting of approx 1000 proteins identified by MS/MS was created on the Waters microQTOF ESI-quadrupole-TOF mass spectrometer. This can be done by undirected MS/MS, as the starting run using identical chromatography conditions as used for the profiling work. Data from this run is then searched against a large sequence database (such as RefSeq from the National Center for Biotechnology Information at NIH's National Library of Medicine) for identification matches using standard commercial software such as Mascot. Some number of posttranslational modifications (PTMs) are routinely included in the searches. For microQTOF data, a Mascot score threshold of about 40 is generally used, although other constraints can be applied, such as retention time prediction. In the second run, molecular ions that have been identified during the first run are put on an exclusion list so that MS/MS data is collected on ions as yet unidentified. Subsequent runs are searched against the database and successfully identified ions are continually added to the exclusion list. As more runs are conducted, abundant peptides get appended to the exclusion list, and peptides of lower concentration become identifiable. The built library contains information on the *m/z*, charge state, retention time, and identity (peptide sequence, protein name, and accession number) of the identified molecular ion.
2. All molecular ions measured in the MS differential profiling study were matched against the library using their accurate mass and chromatographic retention times. Any component from differential profiling results found within 50 or 100 mDa and +/– 1 min with the same charge state in the library is considered a match, the matching or linking range being flexible depending on the experiment.

Fig. 3 *(continued from opposite page)* ference between the individual measure Xi and the average of all measures <Xi>, divided by the standard deviation s; Z= (Xi-<Xi>)/s. Each component (for RA and controls together) is scaled to zero mean and unit variance in order to apply a mapping to the color scale. Normal individuals clearly show a different pattern of their proteomic 1D and 2D profiles from RA patients. In addition, different individuals show slightly different patterns within a given group. The 2D LC–MS platform differentially profiles more low-abundance proteins, identifies more significant changes, and easily differentiates healthy individuals from RA patients. **(B)** Comparison of CVs and significant changes found by 1D LC–MS and 2D LC–MS. Profiling data from the same 12 individuals are compared. In the 1D study, although a total of 38 samples were quantified, this data only considers the 12 samples that were followed up with 2D LC analysis. Maintaining 1D CVs, many more significant changes can be quantified by 2D LC.

Molecular Ion				Peptide Identification by linking to library or directed MSMS					Quantification	
#	m/z	R.T. min	z	gi #	Protein	Peptide	Score	DM (mD)	Fold Change	P used
1	456.75	39.14	2						-1.71	4.29E-04
2	508.56	46.06	3						-1.36	2.83E-04
3	628.27	54.36	3						-1.48	3.21E-04
4	641.26	43.62	2						-1.34	4.20E-04
5	977.95	60.76	2						-1.26	4.88E-04
6	340.48	23.19	3						1.21	4.22E-04
7	387.52	67.40	3						1.34	3.36E-04
8	510.23	23.19	2						1.30	1.04E-04
9	581.92	73.91	3						1.25	7.84E-04
10	370.86	50.11	3						1.51	9.10E-06
11	387.21	26.78	2						1.65	7.63E-06

ID Library → Directed ID

Molecular Ion				Peptide Identification by linking to library or directed MSMS					Quantification	
#	m/z	R.T. min	z	gi #	Protein	Peptide	Score	DM (mD)	Fold Change	P used
1	456.75	39.14	2	4557871	transferrin	YYAVAVVK	5538	-27	-1.71	4.29E-04
2	508.56	46.06	3	4557871	transferrin	LKCDEWSVNSVGK	11028	-41	-1.36	2.83E-04
3	628.27	54.36	3	4557871	transferrin	ADRDQYELLCLDNTR	21362	-66	-1.48	3.21E-04
4	641.26	43.62	2	4557871	transferrin	CDEWSVNSVGK	13874	-30	-1.34	4.20E-04
5	977.95	60.76	2	4557871	transferrin	NLNEKDYELLCLDGTR	5322	-14	-1.26	4.88E-04
6	340.48	23.19	3	9257232	Orosomucoid-1	DKCEPLEK	8090	-47	1.21	4.22E-04
7	387.52	67.40	3	9257232	Orosomucoid-1	WFYIASAFR	8514	-45	1.34	3.36E-04
8	510.23	23.19	2	9257232	Orosomucoid-1	DKCEPLEK	8090	-20	1.30	1.04E-04
9	581.92	73.91	3	9257232	Orosomucoid-1	EQLGEFYEALDCLR	19432	-45	1.25	7.84E-04
10	370.86	50.11	3						1.51	9.10E-06
11	387.21	26.78	2						1.65	7.63E-06

Fig. 4. Identification (ID) using libraries. After obtaining profiling data, matches to the ID library are linked. Unmatched components of interest (low *p*-values) are targeted for ID by directed tandem mass spectrometry. Finally, differentially profiled proteins can be observed. Proteins with more than one peptide showing consistency in fold change, as shown in the figure, are more reliable. There still may be profiled molecular ions that are difficult to identify because of poor fragmentation or the presence of posttranslational modifications.

3. For molecular ions that were not successfully matched against the existing peptide ID library, additional directed MS/MS experiments are conducted. These experiments are prioritized based on the statistical results. **Figure 4** illustrates the sequential and decoupled process of (1) profiling, (2) statistics, and then (3) identification, with a two-step identification process.
4. An important validation or quality check is that when tracking several peptides from the same protein they must show similar fold changes in concentration. An example is shown in **Fig. 4**. Note that there can be valid exceptions to this consistency, especially for peptides that reflect PTMs or in vivo enzymatic cleavages.
5. Not all top-changing (lowest *p*-value) molecules can be easily identified by this method. Unidentified molecular ions may reflect difficulty in obtaining high-

quality fragmentation data, or the complexity of determining unusual PTMs. However, in our experience, if the molecular ion that is changing is of great interest, and shows, for example, a significantly greater fold change than all other components, it can be targeted for searches with additional PTMs, and can also be subjected to *de novo* sequencing.

3.13. Validation of Overall Platform Reproducibility

Technology reproducibility and robustness are key in any quantitative MS approach, especially involving clinical samples. The following describes a large-scale reproducibility study designed to evaluate this platform.

1. The 1D LC–MS study on RA described previously in **Subheading 3.6.** was repeated after 10 mo by using a second set of frozen serum aliquots stored at –80°C and repeating sample processing, LC–MS analysis, and quantification. Identical sample processing and data analysis procedures were used.
2. Results from the original and the repeat study were compared to demonstrate the long-term validity of this platform. **Table 1** shows that the fold changes measured during differential profiling of the RA and control group during the first study are very similar to those measured during the repeat study. All 35 proteins having *p*-values less than 0.05 show identical trends (direction of concentration change, up or down).
3. The data in **Table 1** for the percentage difference was combined into a histogram format; *see* **Fig. 5**. This figure shows that most of the fold changes measured in the repeat study are within 10% of those measured in the original study. Considering that the overall CVs for measuring thousands of peaks is in the range of 20 to 30%, it is very encouraging that fold changes can be reproduced within 20%. Note that the values of fold change for the protein are based on the average fold change from those tryptic peptides with $p < 0.05$.
4. This type of reproducibility study demonstrates the capability of this quantification platform. It is recommended that any differential expression platform be tested in this or a similar manner before conducting clinical studies whose results will be used to make decisions on drug or diagnostic development.

4. Notes

1. All biological samples should be collected under identical conditions ***(17)***. Care should be taken to collect serum and/or plasma as well as cerebrospinal fluid without hemolysis.
2. Several serum and plasma collection tubes contain polymer plugs that result in considerable amounts of polyethelyne glycol (PEG) in the sample. As PEG can interfere with signals from analytes, it is best to develop methods that decrease PEG contamination, for example, washing ultrafiltration devices repeatedly with organic solvents and water prior to use. In fact, differential results in the amount of PEG seen in different samples may be a good way to track and understand inconsistencies in sample preparation. The authors currently prefer tubes spray-coated with ethylenediaminetetraacetic acid or heparin for plasma collection.

Table 1
Percent Change Upon Repeat Analysis[a]

#	Protein Description	gi #	Fold Change S1	Fold Change S2	Trend S1	Trend S2	% Diff. (S2-S1)/S1
1	transferrin	4557871	-1.21	-1.18	-	-	-2.6%
2	serine (or cysteine) proteinase inhibitor, clade A; alpha-1-antitrypsin	21361198	1.32	1.34	+	+	1.5%
3	albumin precursor; PRO0883 protein	4502027	-1.20	-1.25	-	-	4.8%
4	alpha 2 macroglobulin precursor	4557225	-1.34	-1.27	-	-	-5.2%
5	ceruloplasmin (ferroxidase); Ceruloplasmin	4557485	1.30	1.38	+	+	6.0%
6	haptoglobin-related protein; Haptoglobin-related locus	23821019	1.33	1.57	+	+	18.3%
7	orosomucoid 1 precursor; Orosomucoid-1 (alpha-1-acid glycoprotein-1)	9257232	1.48	1.36	+	+	-8.0%
8	haptoglobin	4826762	1.53	1.63	+	+	6.9%
9	alpha-1-antichymotrypsin, precursor; alpha-1-antichymotrypsin	4501843	1.37	1.50	+	+	9.1%
10	similar to Ceruloplasmin precursor (Ferroxidase)	27481320	1.28	1.37	+	+	7.2%
11	inter-alpha (globulin) inhibitor, H2 polypeptide	4504783	-1.24	-1.23	-	-	-0.6%
12	leucine-rich alpha-2-glycoprotein	16418467	1.65	1.69	+	+	2.3%
13	GC1_HUMAN Ig gamma-1 chain C region	121039	1.46	1.45	+	+	-0.6%
14	orosomucoid 2; alpha-1-acid glycoprotein, type 2	4505529	1.35	1.40	+	+	3.5%
15	serine (or cysteine) proteinase inhibitor, clade C, antithrombin III	4502261	-1.42	-1.32	-	-	-6.8%
16	beta globin	4504349	-2.26	-1.77	-	-	-21.7%
17	GC3_HUMAN Ig gamma-3 chain C region (Heavy chain disease protein)	121045	1.50	1.75	+	+	16.7%
18	LAC_HUMAN Ig lambda chain C regions	125946	1.42	1.47	+	+	3.6%
19	apolipoprotein A-II precursor	4502149	-1.71	-1.34	-	-	-21.9%
20	alpha 1 globin [Homo sapiens]	4504347	-1.15	-1.69	-	-	47.0%
21	KV3M_HUMAN IG KAPPA CHAIN V-III REGION HIC PRECURSOR	125819	1.52	1.61	+	+	6.1%
22	Ig gamma-3 chain C region, form LAT - human	321150	-1.13	-1.11	-	-	-1.2%
23	JC-kappa protein - human	1082553	1.31	1.39	+	+	5.8%
24	delta globin	4504351	-2.02	-1.77	-	-	-12.4%
25	isocitrate dehydrogenase 3 (NAD+) alpha precursor	5031777	1.27	1.30	+	+	2.7%
26	G-gamma globin; hemoglobin, gamma G	6715607	-2.06	-1.75	-	-	-15.3%
27	hypothetical protein DKFZp761B2423.1 - human (fragment)	11360219	1.29	1.39	+	+	8.0%
28	retinoblastoma-associated protein RAP140	14150229	1.20	1.18	+	+	-1.4%
29	complement component 4A preproprotein; complement component 4S	14577919	1.33	1.35	+	+	1.2%
30	HMG-BOX transcription factor BBX; x 001 protein	18378731	1.35	1.27	+	+	-6.3%
31	adaptor-related protein complex 3, delta 1 subunit; adaptin, delta	20127438	1.69	1.65	+	+	-2.1%
32	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 33	20336302	-1.45	-1.36	-	-	-5.9%
33	similar to KIAA1902 protein	20535708	1.36	1.57	+	+	15.5%
34	RAB3A interacting protein (rabin3)-like 1	21361440	1.26	1.41	+	+	11.8%
35	CD19 antigen; B-lymphocyte antigen CD19	23397636	-1.30	-1.65	-	-	26.8%

[a]The original study (S1) and repeat study (S2, after 10 mo) show identical trends for all 35 proteins. Fold changes are consistent and percentage difference (% diff) between the original and repeat study values are small. The gi# is the identification number from the National Center for Biotechnology Information (NCBI) RefSeq data base.

Instructions for serum or plasma collection should be followed closely with special attention to the time within which collected sample is spun and separated.

3. Removal of abundant proteins from serum or plasma is considered of major importance. Based on sample availability and cost constraints, it may be practical to remove only the top two abundant proteins instead of the top six, as described in **Subheading 2.2.**, **step 2**.
4. These methods may be used in nano-LC or micro-LC mode. The protein identification library is constructed on an LC–MS system where the HPLC conditions are identical to that used for profiling. Simple standards (such as peptide mixtures) are used to monitor any drifts in retention time.

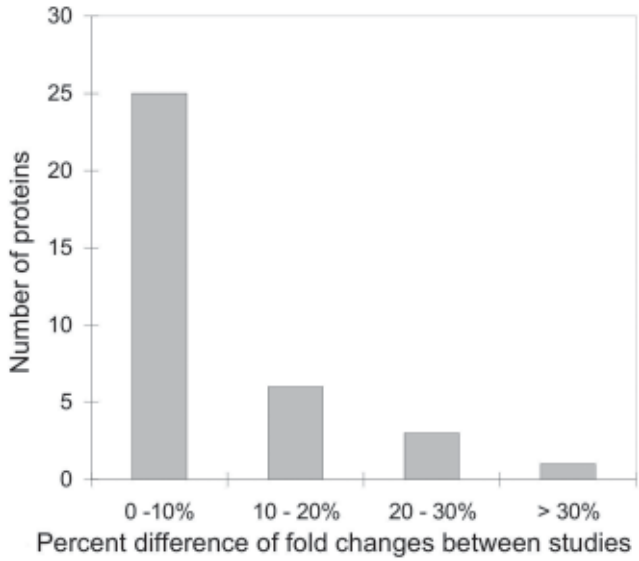


Fig. 5. Histogram of percent difference in fold changes recorded on repeat analysis (1D liquid chromatography–mass spectrometry rheumatoid arthritis studies separated by 10 mo). Data is shown for 35 statistically significant ($p < 0.05$) putative protein biomarkers.

5. It is important to consider chromatographic resolution while planning a 2D LC study. Increasing the number of fractions to large numbers does not increase the dynamic range of profiling when the elution time window of a collected fraction approaches peak widths limited by the chromatographic resolution. 2D chromatographic methods and fraction size should be designed for a given amount and type of sample material, SCX column size, and resolution.
6. In this method, the quantification software is key to differential profiling. Quantification results need to be validated in detail at all stages of software development.
7. It is common to see that the distribution of CVs in a control group is less broad than a disease group, and that CVs from humans are greater than that for laboratory animals, especially mice and rats.
8. It is important to consider the number of samples necessary to profile for reliable statistics. We suggest a minimum of 20 individuals in each group to get relatively reliable statistical power in a 1D study. Statistical power calculations are recommended.
9. Library building and the linking of profiling data to the library require considerable attention to detail. Consistency of fold change between all peptides from the same protein should be checked. Inconsistent peptides must be followed up to

investigate an incorrect link or the presence of a PTM. Singletons, i.e., proteins represented by a single peptide in the library or in the linked data, must be treated with caution for biological interpretation. We suggest following published guidelines for such data.

10. This protocol can be adapted to any biological fluid by starting with a similar amount of protein (after abundant proteins have been depleted). For example, in cerebral spinal fluid, protein concentrations are about 200-fold less than in serum, yet a similar method for denaturation, reduction, alkylation, and digestion of proteins is used. For tissue samples, standard methods of protein extraction may be used prior to the steps described for serum or plasma protocols. This LC–MS quantification and identification strategy can be applied to any proteomic or metabolomic data, irrespective of the biological fluid or sample type.

References

1. Bodovitz, S. and Patterson, S. D. (2003) Protein biomarker strategies. *Drug Discovery World* **Fall 2003**, 67–78.
2. Rubenstein, K. (2003) Post-genomic biomarkers: revolutionizing drug development and diagnostics. D and MD Publications, Westborough, MA.
3. Wang, W., Zhou, H., Lin, H., et al. (2003) Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal. Chem.* **75,** 4818–4826.
4. Steen, H., Jebanathirajah, J. A., Springer, M., and Kirschner, M. W. (2005) Stable isotope-free relative and absolute quantitation of protein phosphorylation stoichiometry by MS. *Proc. Natl. Acad. Sci. USA* **102,** 3948–3953.
5. Anderson, N. L. and Anderson, N. G. (2003) The human plasma proteome: history, character and diagnostic prospects. *Mol. Cell. Proteomics* **2**, 50.
6. Adkins, J. N., Varnum, S. M., Auberry, K. J., et al. (2002) Toward a human blood serum proteome: Analysis by multidimensional separation coupled with mass spectrometry. *Mol. Cell. Proteomics* **1,** 947–955.
7. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol.* **17,** 994–999.
8. Ji, J., Chakraborty, A., Geng, M., et al. (2000) Strategy for qualitative and quantitative analysis in proteomics based on signature peptides. *J. Chromatogr. B* **745,** 197–210.
9. Goshe, M. B. and Smith, R. D. (2003) Stable isotope-coded proteomic mass spectrometry. *Curr. Opin. Biotechnol.* **14,** 101–109.
10. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. USA* **100,** 12,6940–12,6945.
11. US Department of Health and Human Services. *Biosafety in Microbiological and Biomedical Laboratories*, 4th Edition, HHS Publication No. (CDC) 93-8395.
12. Hastings C. A., Norton, S. M., and Roy, S. (2002) New algorithms for processing and peak detection in liquid chromatography/mass spectrometry data. *Rapid Commun. Mass Specrom.* **16,** 462–467.

13. Anderle, M., Roy, S., Lin, H., Becker, C., and Joho, K. (2004) Quantifying reproducibility for differential proteomics: noise analysis for protein liquid chromatography-mass spectrometry in human serum. *Bioinformatics* **20,** 183,575–183,582.
14. Roy, S., Anderle, M., Lin, H., and Becker, C. (2004) Differential expression profiling of serum proteins and metabolites for Biomarker Discovery. *Int. J Mass Spec.* **238,** 163–171.
15. Rabiner, L. and Juang, B. -H. (1993) *Fundamentals of Speech Recognition*, Prentice-Hall Publishers, Englewood Cliffs, NJ.
16. Glantz, S. A. (1997) *Primer of Biostatistics*, McGraw-Hill, New York.
17. Carr, S., Aebersold, R., Baldwin, M., Burlingame, A., Clauser, K., and Nesvizhskii, A. (2004) The need for guidelines in publication of peptide and protein identification data: working group on publication guidelines for peptide and protein identification data. *Mol. Cell. Proteomics* **3,** 531–533.

7

The Use of a Quantitative Cysteinyl-Peptide Enrichment Technology for High-Throughput Quantitative Proteomics

Tao Liu, Wei-Jun Qian, David G. Camp, II, and Richard D. Smith

Summary

Quantitative proteomic measurements are of significant interest in studies aimed at discovering disease biomarkers and providing new insights into biological pathways. A quantitative cysteinyl-peptide enrichment technology (QCET) can be employed to achieve higher efficiency, greater dynamic range, and higher throughput in quantitative proteomic studies based on the use of stable isotope-labeling techniques combined with high-resolution capillary or nano-scale liquid chromatography–mass spectrometry (LC–MS) measurements. The QCET approach involves specific $^{16}O/^{18}O$-labeling of tryptic peptides, high-efficiency enrichment of cysteinyl-peptides, and confident protein identification and quantification using high mass accuracy LC–Fourier transform ion cyclotron resonance mass spectrometry (FTICR) measurements and a previously established database of accurate mass and LC elution time information for the labeled peptides. This methodology has been initially demonstrated by using proteome profiling of naïve and in vitro-differentiated human mammary epithelial cells as an example, which initially resulted in the identification and quantification of 603 proteins in a single LC–FTICR analysis. QCET provides not only highly efficient enrichment of cysteinyl-peptides for more extensive proteome coverage and improved labeling efficiency for better quantitative measurements, but more importantly, a high-throughput strategy suitable for quantitative proteome analysis where extensive or parallel proteomic measurements are required, such as in time course studies of specific pathways and clinical sample analyses for biomarker discovery.

Key Words: Quantitative proteomics; QCET; ^{18}O-labeling; cysteinyl-peptide enrichment; FTICR; AMT.

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

1. Introduction

Quantitative proteomic measurements play a significant role in studies aimed at discovering disease biomarkers and providing new insights into biological pathways. A common strategy for obtaining these measurements is to combine stable isotope-labeling techniques with automated liquid chromatography (LC)–tandem mass spectrometric analyses (MS/MS) ***(1–4)***. Inherent with this strategy, however, are a number of analytical challenges that stem from sample complexity and the wide dynamic range of protein abundances, in addition to the low analysis throughput that results from extensive chromatographic fractionation often needed to minimize MS/MS undersampling issues ***(5)*** and improve overall proteome coverage.

In this chapter, we detail a quantitative proteomics approach—the quantitative cysteinyl-peptide enrichment technology (QCET)—that was developed in response to these challenges. The QCET approach involves specific $^{16}O/^{18}O$-labeling of tryptic peptides, high-efficiency enrichment of cysteinyl-peptides, and confident protein identification and quantification using high mass accuracy LC–Fourier transform ion cyclotron resonance mass spectrometry (FTICR) measurements and a previously established database of accurate mass and elution time information. This technology enables higher efficiency, greater dynamic range, and higher throughput quantitative proteomic analyses than previous quantitation technologies ***(6)***.

This chapter is organized as follows with the materials required for the QCET approach being presented in **Subheading 2.** Following an overview of the technological approach and some of its attractive features, **Subheading 3.** provides stepwise methods that exemplify the QCET approach for protein profiling, using naïve and in vitro-differentiated human mammary epithelial cells (HMEC) as our specific example. **Subheading 4.** concludes with notes that pertain to both materials and methods.

2. Materials

In our example of protein profiling using QCET, naïve and in vitro-differentiated HMEC nontumorigenic strain 184A1 were routinely cultured in DFCI-1 medium (Gibco BRL, Gaithersburg, MD) until 90% confluence was achieved ***(7)***. Cells in eight dishes (1×10^7 cells/dish) were treated with 200 n*M* Phorbol 12-myristate 13-acetate (PMA) for 24 h to provide in vitro-differentiated cells, and cells in another eight dishes were cultured under normal conditions to provide naïve cells. The materials required for this example are listed by activity.

2.1. Protein Digestion and Clean-Up

1. Common phosphate-buffered saline.

2. Cell lysis buffer: 10 m*M* Tris, 150 m*M* NaCl, 1% NP-40 (v/v), 1m*M* $NaVO_3$, 10 m*M* NaF, and protease inhibitor cocktail (Roche, Indianapolis, IN), pH 7.4 (*see* **Note 1**).
3. Bicinchoninic acid protein assay (Pierce, Rockford, IL).
4. Reducing buffer: 50 m*M* Tris-HCl, pH 8.2.
5. Solid high purity urea (Sigma, St. Louis, MO) for denaturing the proteins.
6. 200 m*M* tributylphosphine (TBP) solution (Sigma) for reducing the proteins (*see* **Note 2**).
7. Digestion buffer: 20 m*M* Tris-HCl, pH 8.2.
8. Sequencing grade-modified porcine trypsin (Promega, Madison, WI), freshly dissolved in digestion buffer to a final concentration of 1 μg/μL.
9. 1-mL solid-phase extraction (SPE) C18 column per cell state (Supelco, Bellefonte, PA).
10. SPE conditioning solution: 0.1% trifluoroacetic acid (TFA).
11. SPE washing solution: 5% acetonitrile and 0.1% TFA.
12. SPE eluting solution: 80% acetonitrile and 0.1% TFA.

2.2. Postdigestion ^{16}O- to ^{18}O-Exchange

1. Immobilized trypsin (Applied Biosystems, Foster City, CA), used as supplied by the manufacturer (*see* **Note 3**).
2. 1 *M* NH_4HCO_3 stock solution in ^{18}O water: dissolve 79 mg (solid) NH_4HCO_3 in 1 mL ^{18}O-enriched water (95%; Isotec, Miamisburg, OH).
3. 1 *M* $CaCl_2$ stock solution prepared with regular (^{16}O) water.
4. Thermal mixer (Model Thermomixer R; Eppendorf, Westbury, NY).
5. 60% methanol prepared with regular (^{16}O) water.
6. Handee Mini-Spin column (Pierce).

2.3. Cysteinyl-Peptide Enrichment

1. Thiopropyl Sepharose 6B affinity resin (Amersham Biosciences, Uppsala, Sweden) (*see* **Note 4**).
2. Handee Mini-Spin column kit (Pierce) (*see* **Note 5**).
3. Coupling buffer: 50 m*M* Tris-HCl, pH 7.5 and 21 m*M* ethylenediaminetetraacetic acid (EDTA).
4. Washing buffer: 50 m*M* Tris-HCl, pH 8.0 and 1 m*M* EDTA.
5. 100 m*M* dithiothreitol (DTT) in coupling buffer: made fresh from aliquots of 1 *M* DTT stock stored at –80°C (*see* **Note 6**).
6. 2 *M* NaCl.
7. 80% acetonitrile/0.1% TFA solution.
8. 20 m*M* DTT in washing buffer: made fresh from aliquots of 1 *M* DTT stock stored at –80°C.
9. 1 *M* iodoacetamide solution: made fresh from solid iodoacetamid (Sigma).

2.4. Strong Cation Exchange Fractionation

1. Agilent 1100 series high-performance liquid chromatography system (Agilent, Palo Alto, CA).

2. Polysulfoethyl A 200 × 2.1-mm column (PolyLC, Columbia, MD) preceded by a 10 × 2.1-mm guard column (PolyLC) with a flow rate of 0.2 mL/min.
3. Solvent A: 10 m*M* ammonium formate and 25% acetonitrile, pH 3.0.
4. Solvent B: 500 m*M* ammonium formate and 25% acetonitrile, pH 6.8.

2.5. Capillary LC–MS/MS and LC–FTICR Analyses

1. High-pressure capillary LC system ***(8)***.
2. LTQ ion trap mass spectrometer (ThermoFinnigan, San Jose, CA).
3. Apex III 9.4-Tesla FTICR mass spectrometer (Bruker Daltonics, Billerica, MA).
4. 5-µm Jupiter C18-bonded particles (Phenomenex, Torrence, CA).
5. 65-cm long, 150 µm-id × 360 µm-od fused silica capillary (Polymicro Technologies, Phoenix, AZ).
6. 2-µm retaining stainless steel screen (Valco Instruments Co., Houston, TX).
7. On-line vacuum degasser (Jones Chromatography Inc., Lakewood, CO).
8. Mobile phase A: 0.2% acetic acid and 0.05% TFA in water.
9. Mobile phase B: 0.1% TFA in 90% acetonitrile/10% water.

3. Methods

Quantitative proteomics analysis using the QCET approach involves two stages *(see* **Fig. 1**). In the first stage, proteins from different cell states are extracted, mixed, and digested (to improve the proteome coverage). The cysteinyl-peptides from the tryptic digest are covalently enriched using a thiol-affinity resin (*see* **Fig. 2**) and further fractionated using strong cation exchange (SCX) chromatography. LC–MS/MS analysis of each SCX fraction provides accurate calculated mass and normalized elution time (NET) information for each peptide that serve as an effective "look-up table" of markers or "accurate mass and time tags" (AMT tags) for future peptide identifications ***(9)***.

In the next (typically second) stage, equal masses of protein from two different cell states are separately digested by trypsin under identical conditions, and the tryptic peptides from each sample are labeled with either ^{16}O or ^{18}O by trypsin-catalyzed oxygen exchange in either regular or ^{18}O-enriched water. The highly efficient postdigestion ^{18}O-labeling strategy incorporates two ^{18}O atoms in essentially all tryptic peptides. The $^{16}O/^{18}O$-labeled peptide pairs coelute during LC separations. As a result, errors potentially introduced by differences in electrospray ionization suppression effects are minimized, and a framework for accurate quantitation is established.

The differentially labeled peptide samples are combined and subjected to cysteinyl-peptide enrichment. Enrichment of cysteinyl-peptides by the thiol-specific covalent resin is reproducible, highly efficient, and amenable to automation for high-throughput studies. The reversible capture and release reaction of cysteinyl-peptides has no side reactions, and the enriched cysteinyl-peptide does not have a labile tag, which eliminates the problem of fragment in produc-

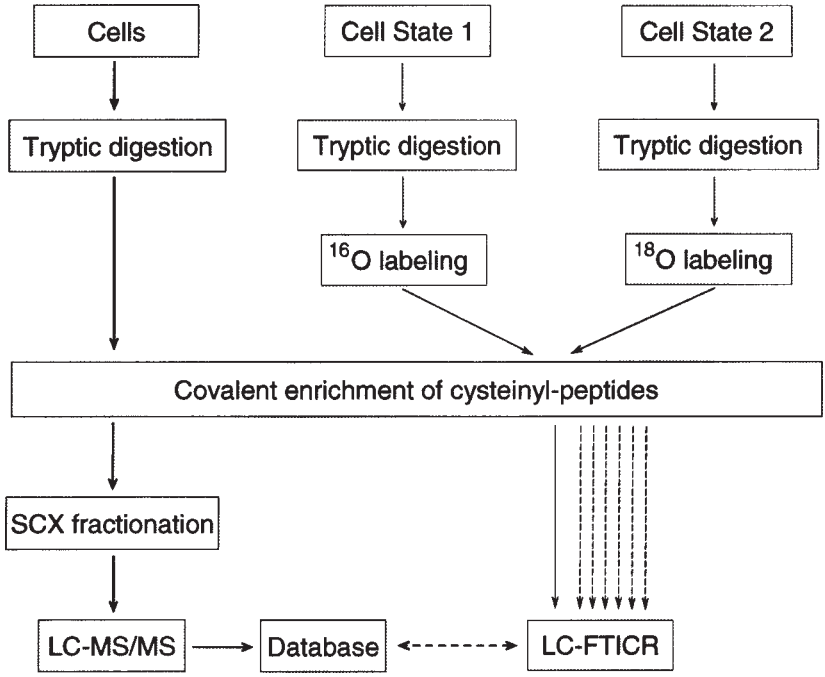


Fig. 1. Strategy for quantification of differential protein expression using quantitative cysteinyl-peptide enrichment technology. In the first stage, proteins from different cell states are mixed and digested by trypsin, followed by cysteinyl-peptide enrichment using thiol-affinity resin. The enriched cysteinyl-peptides are fractionated by strong cation exchange chromatography with each fraction analyzed by liquid chromatography–tandem mass spectrometry (LC–MS/MS). An accurate mass and time (AMT) tag database is generated based on the calculated masses and normalized elution times for all identified peptides. In the second stage, the two protein mixtures representing two different cell states are digested by trypsin separately. The resulting tryptic peptides are labeled by trypsin-catalyzed oxygen exchange using ^{16}O- and ^{18}O-enriched water, respectively. The two samples are combined and cysteinyl-peptides are selectively enriched and analyzed by LC–Fourier transform ion cyclotron resonance mass spectrometry (FTICR) without prefractionation. Peptide features are identified and quantified by matching to the AMT tag database without the need for additional LC–MS/MS analyses. Once an AMT tag database is established for a biological system, the system can be extensively investigated in a high-throughput manner by analyzing samples generated under different conditions using LC–FTICR. (Modified from **ref. *6*** with permission from the American Chemical Society.)

tion from the tags (e.g., isotope-coded affinity tag [ICAT] and ICAT-like reagents) during collision induced dissociation (CID). Use of high efficiency cysteinyl-peptide enrichment along with global analysis has been demonstrated

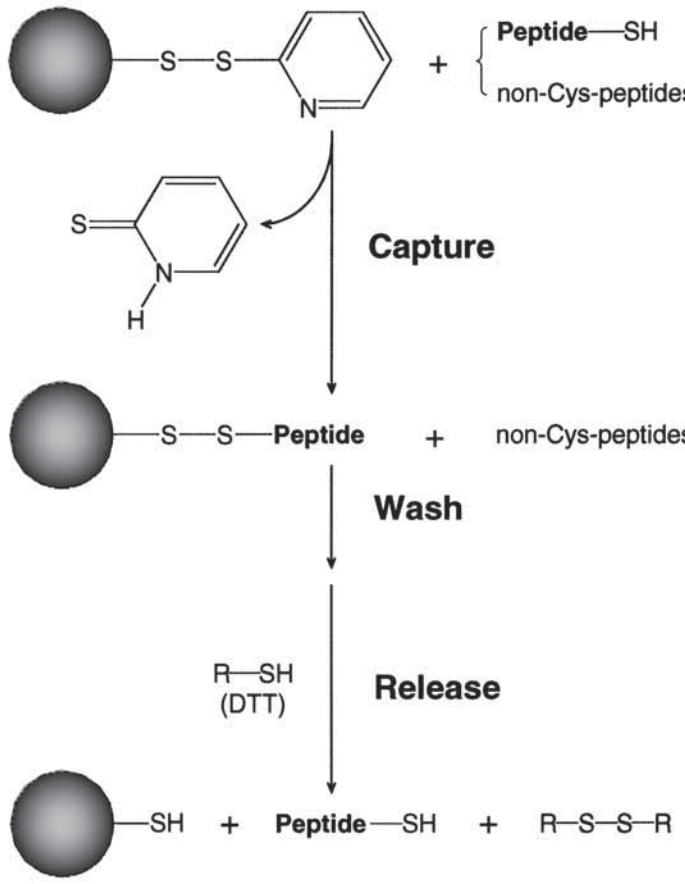


Fig. 2. Enrichment of cysteinyl-peptides using thiol-specific covalent resin. Cysteinyl-peptides were captured by the resin through the formation of a mixed disulfide bond. Stringent washes were applied to remove the nonspecifically, noncovalently bound peptides, after which a low molecular weight reducing reagent (e.g., DTT) was added to release the bound cysteinyl-peptides. (Reproduced from **ref.** ***10*** with permission from John Wiley and Sons, Inc.)

to significantly improve overall proteome coverage ***(10)***, and the significantly reduced sample complexity makes the AMT tag approach even more effective for profiling complicated mammalian systems.

The enriched cysteinyl-peptides are then separated using LC conditions identical to those used in the first stage and analyzed by FTICR, which provides increased sensitivity and dynamic range, in addition to higher throughput proteomic measurements ***(11)***. The greater sensitivity improves identification

of lower abundance peptides, and thus leads to better overall proteome coverage without the need for SCX fractionation, and higher overall throughput. An LC–FTICR detected peptide "feature" is identified when both the measured mass and NET values match those of a particular AMT tag in the database within a defined error tolerance (*see* **Fig 3**). Once peptides are identified, relative abundance differences are quantified based on the MS peak intensities of the $^{16}O/^{18}O$-labeled peptide pairs.

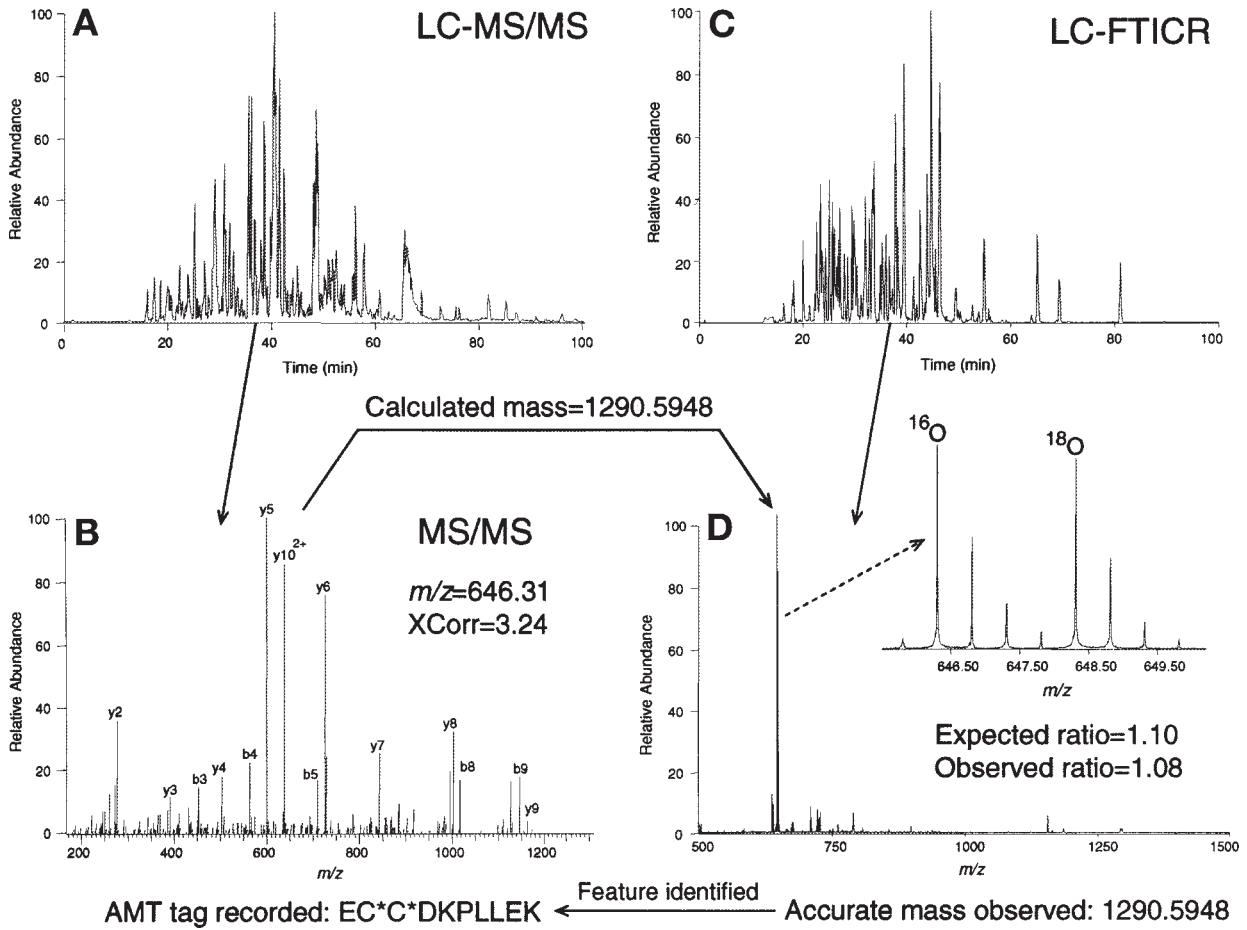
The combination of cysteinyl-peptide enrichment and AMT tag strategy can be readily combined with other labeling methods, such as ^{15}N labeling ***(2)*** and labeling using amino acids in cell culture ***(3)***. Another advantage of this two-stage strategy is that once a database has been established for a particular biological system, time-consuming LC–MS/MS analyses are replaced in subsequent studies by high-throughput LC–FTICR analyses that enable extensive investigations of that biological system. In practice, the QCET approach for high-throughput quantitative proteome profiling can be effectively used with any high resolution MS (e.g., quadrupole time-of-flight).

In our initial demonstration of HMEC protein profiling using QCET ***(6)***, 603 proteins were identified and quantified in a single LC–FTICR analysis, with a number of proteins displaying either up- or down-regulation following PMA treatment (*see* **Fig. 4**). The histograms of mass and NET errors for peptides identified from AMT tag identifications (*see* **Fig. 4B**) display a typical Gaussian distribution. Note that the majority of these identifications are distributed within a mass error of two parts per million (ppm) and a NET error of 2%. Moreover, no peptide pairs are identified as C-terminal open tryptic peptides (peptides lacking a C-terminal lysine or arginine) when applying 5 ppm for mass and 5% for NET criteria. Trypsin-catalyzed ^{18}O-labeling has specificity for only peptides with K or R at the C terminus, thus any peptide pairs matching to C-terminal open partial tryptic peptides represent false-positive hits. This result shows that the 5 ppm mass and 5% NET tolerances provide highly confident peptide identifications.

The stepwise methods employed for this demonstration of QCET protein profiling are detailed below.

3.1. Preparation of HMEC Protein Digests

1. Naïve and PMA-treated HMEC cells are placed in 50-mL Falcon tubes and spun at 1000*g* for 10 min. Each cell pellet is washed three times with ice-cold phosphate-buffered saline.
2. Add cell lysis buffer (0.4 mL/dish) to the cell pellets and lyse the cells using intermittent sonication on ice for 1 min.
3. Centrifuge the lysates for 20 min at 4°C and 14,000*g* to pellet any cellular debris. Collect the supernatants into separate containers and then split the supernatant


A
LC-MS/MS
Relative Abundance
Time (min)
C
LC-FTICR
Relative Abundance
Time (min)
Calculated mass=1290.5948
B
MS/MS
m/z=646.31
XCorr=3.24
y5
y10 2+
y6
y2
y3
b3
y4
b4
b5
y7
y8
b8
b9
y9
m/z
D
16O
18O
646.50
647.50
648.50
649.50
m/z
Expected ratio=1.10
Observed ratio=1.08
Relative Abundance
m/z
Feature identified
AMT tag recorded: EC*C*DKPLLEK
Accurate mass observed: 1290.5948

from each cell state into two aliquots: the first aliquot contains 1.5 mg of protein for generating the AMT tag database and the second contains 100 µg of protein for quantitative analysis. The protein concentrations are approx 10 mg/mL using the bicinchoninic acid protein assay. All aliquots are stored at –80°C until further use.

4. Mix 1.5 mg of protein from naïve cells with 1.5 mg of protein from PMA-treated cells and digest as described in **steps 6** and **7**.
5. Digest 100 µg of protein from each of the two lysates separately as described in **steps 6** and **7**.
6. First, dilute each sample by a factor of two by adding the reducing buffer. Then, add 200 m*M* TBP stock solution and solid urea to final concentrations of 10 m*M* and 8 *M*, respectively. Incubate samples at 37°C for 1 h with gentle mixing.
7. Dilute the reduced protein mixture by a factor of 10 using the digestion buffer. Add sequencing grade-modified porcine trypsin to obtain a trypsin:protein ratio of 1:50 (w/w) and incubate at 37°C for 3 h.
8. Precondition 1-mL SPE C18 columns by slowly passing 3 mL methanol and then 2 mL SPE conditioning buffer through the column.
9. Load each of the tryptic digests onto separate preconditioned SPE C18 columns; pass each sample through, and wash each column with 4 mL of SPE washing buffer.

Fig. 3. *(continued from opposite page)* Experimental steps involved in establishing and using an accurate mass and time (AMT) tag and the identification and quantification of liquid chromatography–Fourier transform ion cyclotron resonance mass spectrometry (LC–FTICR)-detected features using the AMT tag strategy. Protein mixtures A and B, each containing the following five proteins but at different concentrations are prepared using $^{16}O/^{18}O$ labeling and cysteinyl-peptide enrichment: bovine serum albumin (BSA; 1:1, A/B); bovine ribonuclease A (1:6); chicken lysozyme (6:1); chicken ovalbumin (3:1); and rabbit glyceraldehyde-3-phosphate dehydrogenase (1:3). **(A)** Enriched cysteinyl-peptides from the protein mixture are analyzed by LC–tandem mass spectrometry (MS/MS). **(B)** A tryptic peptide from BSA (EC*C*DKPLLEK, where C* represents alkylated cysteine residues) is identified by LC–MS/MS. The calculated mass of this peptide based on its sequence (i.e., 1290.5948 Da) and its observed elution time are recorded in the AMT tag database. **(C)** In the second stage, the same sample is analyzed under the same LC conditions using a FTICR mass spectrometer. **(D)** A doubly charged feature was observed at the same elution time (36 min), having a mass within 1 ppm (i.e., 1290.5948 Da) of the calculated mass of this AMT tag. This feature is then identified as the corresponding peptide from BSA. The $^{16}O/^{18}O$ ratio for this peptide was estimated as 1.08 using the maximum intensities of paired monoisotopic peaks (inset). The expected ratio was calculated from the known amount of BSA present in each mixture. (Modified from **ref. *6*** with permission from the American Chemical Society.)

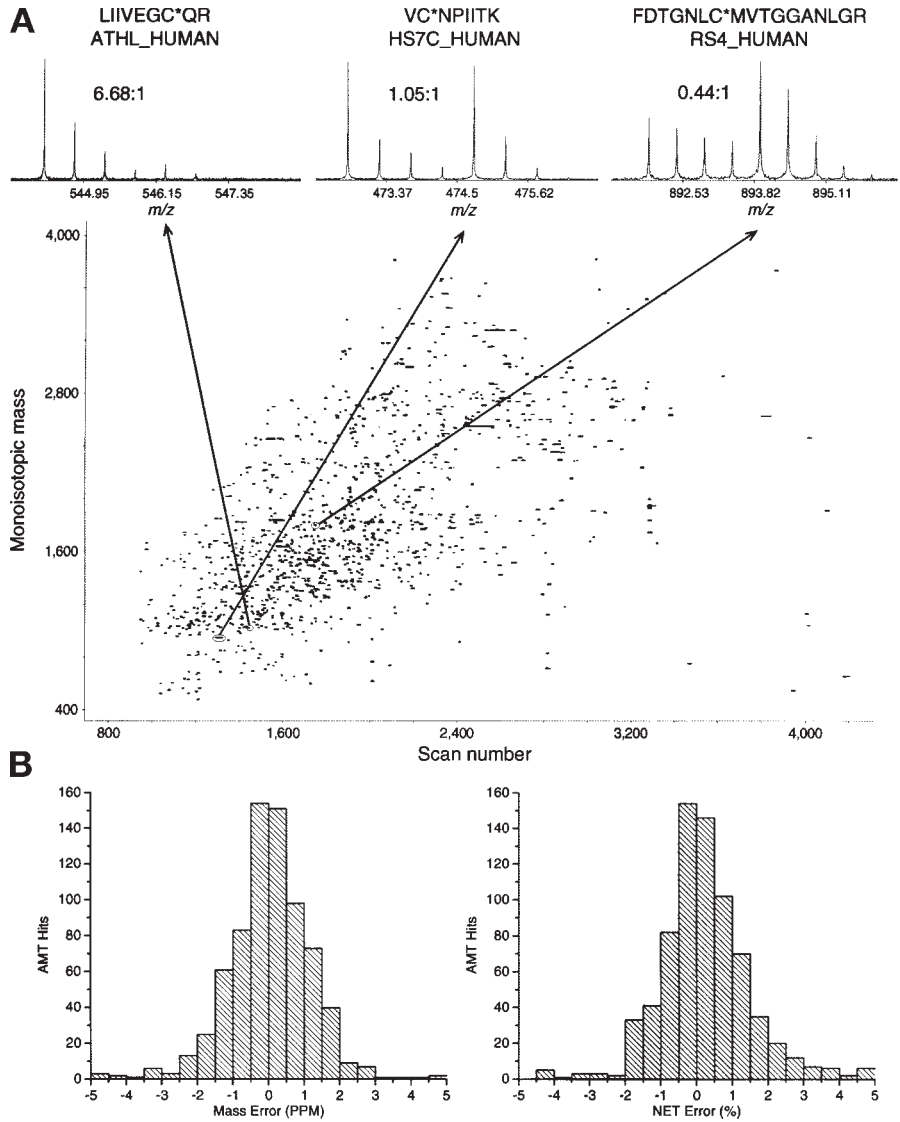


Fig. 4. Quantitative profiling of labeled and enriched cysteinyl-peptides from naive and Phorbol 12-myristate 13-acetate (PMA)-treated human mammary epithelial cells (HMEC)-cells using a single liquid chromatography–Fourier transform ion cyclotron resonance mass spectrometry analysis. (**A**) A two-dimensional display of 1348 peptide pairs from which 935 pairs were identified as unique accurate mass and time (AMT) tags corresponding to 603 proteins identified and quantified. Insets show three samples of peptide pairs with their sequences, corresponding proteins, and the $^{16}O/^{18}O$ ratios. **(B)** The mass error (left) and normalized elution time error (right) distributions of the 935 AMT tag hits. (Reproduced from **ref.** ***6*** with permission from the American Chemical Society.)

10. Elute the peptides from each SPE C18 column with 1 mL of SPE eluting buffer and dry each of the samples under reduced vacuum using a Speed-Vac.
11. Store samples at –80°C until further use.

3.2. Postdigestion ^{16}O- to ^{18}O-Exchange

1. Treat each sample to be labeled with ^{16}O or ^{18}O (100 μg each) separately until **step 8**. Dissolve the dried peptide samples in 100 μL of 50 m*M* NH_4HCO_3, and boil for 10 min using water bath; immediately cool the samples on ice for 5 min to eliminate the residual trypsin activity (*see* **Note 7**).
2. Dry samples completely using a Speed-Vac.
3. Dissolve the dried peptide samples in 20 μL of acetonitrile plus 100 μL of 50 m*M* NH_4HCO_3 in either ^{18}O-enriched water (made freshly from the 1 *M* stock in ^{18}O-enriched water) or regular ^{16}O water, depending on cell type. Peptides from naïve HMEC are to be labeled with ^{16}O and those from the PMA-treated HMEC with ^{18}O.
4. Add 1 μL of 1 *M* $CaCl_2$ and 5 μL of immobilized trypsin resin to each of the peptide samples, then mix with constant shaking for 24 h at 30°C using a thermal mixer (*see* **Note 8**).
5. Centrifuge the samples for 5 min at 15,000*g*, and collect each supernatant in a separate, new microcentrifuge tube (*see* **Note 9**).
6. Add 100 μL of 60% methanol to the remaining pellet of immobilized trypsin resin and thoroughly mix the pellet into solution. Using a pipet, transfer the suspension into an empty Handee Mini-Spin column (with frit) in a 1.7-mL receiving tube; and collect the flow-through by centrifuging at 1000*g*.
7. Repeat **step 6** to wash the immobilized trypsin one more time.
8. Now, combine the supernatants that correspond to the ^{16}O- and ^{18}O-labeled samples from **step 6** and the flow-through from **step 7** (i.e., mix the naïve HMEC-^{16}O peptide sample with the corresponding PMA-treated HMEC-^{18}O peptide sample).
9. Dry the combined sample using a speed-vac and store at –80°C until further use (*see* **Note 10**).

3.3. Cysteinyl-Peptide Enrichment by Thiopropyl Sepharose 6B

1. Degas the coupling and washing buffers and 100 mL of water for 20 min to prevent oxidation of the thiol content.
2. With the exception of the sample to be used to generate the AMT tag database, dissolve the $^{16}O/^{18}O$-labeled sample in 20 μL of coupling buffer, and add 1 μL of 100 m*M* DTT (5 m*M*) to the sample; incubate at 37°C for 1 h to reduce any possible mixed disulfide formation in the sample. Because of the larger size of the peptide sample that is being used to generate the AMT tag database, dissolve this sample in 80 μL of coupling buffer, and add 4 μL of 100 m*M* DTT.
3. Place 5 × 35 mg (5 × 2 μmol disulfide exchange capacity) of dried Thiopropyl Sepharose 6B resin into individual 1.7-mL tubes. Add 1 mL of water to each tube, and rehydrate the resin for 15 min at room temperature. Suspend the resin

well using the pipet and modified 1-mL tips (*see* **Note 11**), and then place in a rack for approx 10 min. The final volume of rehydrated resin should be around 100 µL.

4. Remove and discard 0.5 mL of supernatant (water) off the top of each tube and resuspend the resin with the solution remaining in the tube. Use a pipet and the modified 1-mL tip to carefully transfer each suspension to a Handee Mini-Spin column (without the bottom cap on). Place each spin column in a 2-mL receiving tube and centrifuge at 1000*g* for 30 s to remove the water.
5. Add 0.5 mL of water to the spin column, which will readily resuspend the resin when water is added. Spin at 1000*g* for 30 s to remove the water. Repeat this washing step five more times.
6. Wash the resin 10 times each with 0.5 mL of coupling buffer in the spin column. After the last wash, tightly put the bottom caps on.
7. Dilute the reduced samples to 100 µL by adding coupling buffer to a final DTT concentration of 1 m*M*. Dilute the peptide sample for generating the AMT tag database to 400 µL and then split the sample into four 100-µL aliquots.
8. Add each 100 µL sample to a spin column that contains approx 100 µL of thiol-affinity resin, and put the top caps on the spin columns. Place each spin column with sample into a 1.7-mL tube (with caps cut off) and shake at medium speed for 1 h at room temperature to allow for cysteinyl-peptide capture by the resin (*see* **Note 12**).
9. Place each spin column into a new 1.7-mL tube with both top and bottom caps removed and spin at 1500*g* for 1 min to collect the unbound portion (noncysteinyl-peptides, which can be retained for global quantitative analysis based on noncysteinyl-peptides).
10. Extensively wash the resin using 0.5 mL (six times) of: washing buffer (six times), 2 *M* NaCl (six times), 80% acetonitrile/0.1% TFA (six times), and washing buffer (six times) as in **steps 5** and **6** above to remove the peptides bound nonspecifically to the resin via ionic interactions or hydrophobic interactions.
11. Put the bottom caps back on the spin columns. Add 100 µL of 20 m*M* DTT to the resin, close the top caps, and incubate the spin columns at room temperature for 30 min with shaking. Collect the released cysteinyl-peptides as flow-through by centrifuging the spin columns at 1500*g* for 1 min with both top and bottom caps removed. Use 2-mL receiving tubes for easy pooling of the flow-through with subsequent eluates.
12. Add another 100 µL of 20 m*M* DTT to the resin and mix using pipet tips (*see* **Note 13**). After incubating at room temperature for 10 min with shaking, place the spin columns into the same collection tubes and centrifuge at 1500*g* for 1 min to collect the released cysteinyl-peptides. Repeat this step once.
13. Repeat **step 12**, but this time use 80% acetonitrile instead of 20 m*M* DTT.
14. Immediately alkylate the released cysteinyl-peptides by adding 32 µL (80 m*M*) of 1 *M* iodoacetamide solution to the pooled eluates (a total of 400 µL) and incubate for 1 h at room temperature in the dark.
15. Dilute the alkylated samples to 2 mL by adding washing buffer, followed by SPE C18 clean up as described in **Subheading 3.1.**

16. Dry the SPE C18 eluates using a speed-vac and store at –80°C until further use.

3.4. SCX Fractionation of Enriched Cysteinyl-Peptides

1. Reconstitute the enriched cysteinyl-peptides from the large peptide sample (for generating the AMT tag database) with 900 µL of solvent A and inject onto the SCX column with a flow rate of 0.2 mL/min.
2. Once loaded, maintain an isocratic gradient with 100% solvent A for 10 min. Separate peptides by applying a gradient from 0 to 50% B over 40 min, followed by a gradient of 50–100% B over 10 min. Hold the gradient isocratically at 100% solvent B for an additional 10 min.
3. Collect a total of 35 fractions using the automatic fraction collector on the high-performance liquid chromatography. Dry each fraction and store at –80°C until time for analysis.

3.5. Capillary LC–MS/MS and LC–FTICR Analyses

1. In our laboratory, peptide samples are analyzed using a custom-built capillary LC system ***(8)*** coupled online to either a LTQ ion trap mass spectrometer (for generating the AMT tag database) or to an Apex III 9.4-Tesla FTICR mass spectrometer equipped with an electrospray ionization (ESI) interface for subsequent measurements.
2. The reversed-phase capillary column is prepared by slurry packing 5-µm Jupiter C18 bonded particles into a 65-cm long, 150-µm inner diameter (id) fused silica capillary.
3. Mobile phases are degassed on-line using a vacuum degasser. After injecting 10 µL of peptide sample onto the reversed-phase capillary column, hold the mobile phase at 100% A for 20 min. Attain exponential gradient elution by increasing the mobile-phase composition to approx 70% B over 150 min, using a stainless steel mixing chamber. Apply this same gradient for both LC–MS/MS and LC–FTICR analyses.
4. The ion trap and the FTICR mass spectrometers are operated under normal conditions. Details of the LC–MS/MS and LC–FTICR operating parameters employed for this demonstration are available in **refs.** ***10*** and ***12***.

3.6. Data Analysis

In our laboratory, data analysis and processing steps are automated and proceed as follows:

1. For LC–MS/MS analyses, peptides are identified using SEQUEST and static modification of cysteine residues (+57 Da) to search the MS/MS spectra against a normal nonredundant human International Protein Index database (consisting of 47,306 protein entries at the time of our analysis; available online at http://www.ebi.ac.uk/IPI) and its sequence-reversed version.
2. The false-positive rate of the SEQUEST search results is evaluated using a reversed protein sequence database ***(13)***. From this evaluation, the following set

of criteria are determined for filtering the raw SEQUEST data with an overall confidence of greater than 95% ($\Delta Cn \geq 0.1$): for the 1+ charge state, $Xcorr \geq 1.5$ for fully tryptic peptides, and $Xcorr \geq 3.1$ for partially tryptic peptides; for the 2+ charge state, $Xcorr \geq 1.9$ for fully tryptic peptides and $Xcorr \geq 3.8$ for partially tryptic peptides; and for the 3+ charge state, $Xcorr \geq 2.9$ for fully tryptic peptides and $Xcorr \geq 4.5$ for partially tryptic peptides. Nontryptic peptides are not included.

3. Peptides that meet the criteria in **step 2** are included in the AMT tag database. The peptide retention times from each LC–MS/MS analysis are normalized to a range of 0 to 1 by using a predictive peptide LC–NET model and linear regression ***(14)***. An average NET value and NET standard deviation are assigned to each identified peptide, provided the same peptide was observed in multiple runs. Both the calculated accurate monoisotopic mass and the NET of the identified peptides are included in the AMT tag database.
4. The initial analysis of raw LC–FTICR data involves a mass transformation or deisotoping step using ICR2LS, an analysis tool based on the THRASH algorithm ***(15)***. ICR2LS data analysis generates a text file report for each LC–FTICR data set that includes both the monoisotopic masses and the corresponding intensities for all detected species for each spectrum.
5. Following ICR2LS analysis, data are processed to yield a two-dimensional mass and LC elution time data set. Automated data processing steps include filtering data, finding features (i.e., a peak with a unique mass and elution time signature) and pairs of features, computing abundance ratios for pairs of features, normalizing LC elution times, and matching the accurately measured masses and NET values of each feature to the corresponding AMT tag in the database to identify peptide sequences. The peptide sequences of a given feature or pair of features are assigned when the measured mass and NET match the calculated mass and NET in the AMT tag database within 5 ppm mass error and 5% NET error.
6. The abundance ratios ($^{18}O/^{16}O$) for labeled peptide pairs are computed by using an equation similar to that previously reported ***(4)***.

$$R\left(\frac{^{18}\mathrm{O}}{^{16}\mathrm{O}}\right) = \frac{I_4 - \frac{M_4}{M_0} I_0 + I_2\left(1 - \frac{M_2}{M_0}\right) - \left(1 - \frac{M_2}{M_0}\right)\frac{M_2}{M_0} I_0}{I_0} \qquad (1)$$

where I_0, I_2, and I_4 are the measured intensities for the monoisotopic peak for the peptide without ^{18}O label, the peak with mass 2 Da higher, and the peak with 4 Da higher mass, respectively. M_0, M_2, and M_4 are the predicted relative abundances for the monoisotopic peak for the peptide, the peak with mass 2 Da higher, and the peak with mass 4 Da higher, respectively. The

$\frac{M_2}{M_0}$ and $\frac{M_4}{M_0}$ ratios are estimated based on the following two equations (**Eq. 2** and **3**) ***(16)*** where *Mr* represents peptide molecular weight.

$$\frac{M_2}{M_0} = 3 \times 10^{-7} Mr^{1.9241} \quad (2)$$

$$\frac{M_4}{M_0} = 2 \times 10^{-12} Mr^{3.2684} \quad (3)$$

7. Ratios from multiple observations of the same peptide across different analyses are averaged to give one ratio per peptide. All quantified peptides are rolled up to nonredundant protein groups using ProteinProphet ***(17)*** (*see* **Note 14**), and the abundance ratio for each protein group is calculated by averaging the ratio of multiple unique peptides stemming from the same protein group.

4. Notes

1. Unless otherwise stated, all solutions should be prepared in deionized water that has a resistivity of 18.2 MΩ-cm and total organic content of less than 5 ppb (parts per billion). This standard is referred as "water" or regular "(^{16}O) water" in the text.
2. TBP is received in sealed Ampere vials. Once opened, TBP may decline in quality. Thus, it is best used if stored at 4°C under nitrogen gas for no longer than 1 wk.
3. Immobilized trypsin is in the form of a fine resin and is received as a white suspension. Mix well or shake briefly before use.
4. Thiopropyl Sepharose 6B resin is received as dried powder. Once opened, it is best stored at 4–8°C in a desiccator.
5. The Handee Mini-Spin column kit contains an empty spin column (0.5 mL capacity, with frit), top cap, and bottom cap. Make sure the bottom cap is fitting tightly with the bottom of the spin column to prevent leakage during incubation with shaking.
6. Because of the frequent use of DTT in the cysteinyl-peptide enrichment experiment, it is convenient to make a 1 *M* DTT stock solution in water and then split it into 10-μL aliquots. This stock solution can be stored at –80°C for more than 3 mo without decline in quality.
7. Failure to eliminate residual trypsin activity will result in ^{16}O and ^{18}O back-exchange. Alternatively, immobilized trypsin can be used for protein digestion, but additional sample handling may required (e.g., washing the immobilized trypsin resin following the digestion).
8. White precipitates are observed once the $CaCl_2$ and immobilized trypsin are added. It is necessary to shake the suspension at high speed (e.g., 1300 rpm) to keep the immobilized trypsin homogeneously distributed to improve the reaction efficiency.

9. Keep corresponding ^{16}O- and ^{18}O-labeled samples separate in different tubes until immobilized trypsin resin is filtered using the empty spin column. We recommend mixing differentially labeled samples from the same biological system as the last step to prevent trypsin-catalyzed ^{16}O- and ^{18}O back-exchange. If a precipitate is still observed in any of the final combined samples, we recommended centrifuging again and transferring the supernatant to a new tube.
10. There will be a light yellowish or whitish layer on the bottom of tubes after completely dried, but it can be readily dissolved in the coupling buffer in the cysteinyl-peptide enrichment step, or in other common buffers, such as 25 m*M* NH_4HCO_3.
11. The resin tends to get sticky at the beginning of the rehydration and, thus, forms chunks that make pipetting difficult using regular tips. Cut 5 mm off the end of the tip to make the hole larger for easier mixing and transfer.
12. We recommend shaking the spin column at no less than 800 rpm to keep the resin homogeneously distributed. However, the use of a speed of higher than 1000 rpm may cause spillage of the solution through the top cap, resulting in sample loss. Also, we recommend checking the bottom caps after shaking for 5 min. Finding a potential leakage problem earlier minimizes the chance of failure in the cysteinyl-peptide enrichment experiment.
13. Once incubated with DTT, the resin tends to get packed and sticky after the low-speed centrifugation step because its structure is changed with the elution of cysteinyl-peptides. Use pipet tips to completely mix the freshly added DTT with the resin.
14. All peptides that pass the filtering criteria are given the identical probability score of 1, and entered into the Protein Prophet program only for clustering analysis to generate a final nonredundant list of proteins or protein groups.

Acknowledgments

The authors would like to thank Drs. Wan-Nan U. Chen and Brian Thrall for providing naïve and PMA-treated HMEC cells. The research discussed in this chapter is supported by the NIH National Center for Research Resources (RR18522) located in the Environmental Molecular Science Laboratory (a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory). Pacific Northwest National Laboratory is a multiprogram national laboratory operated by Battelle Memorial Institute for the US Department of Energy under contract DE-AC05-76RLO-1830.

References

1. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17,** 994–999.
2. Veenstra, T. D., Martinovic, S., Anderson, G. A., Pasa-Tolic, L., and Smith, R. D. (2000) Proteome analysis using selective incorporation of isotopically labeled amino acids. *J. Am. Soc. Mass Spectrom.* **11,** 78–82.

3. Ong, S. E., Kratchmarova, I., and Mann, M. (2003) Properties of ^{13}C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J. Proteome Res.* **2,** 173–181.
4. Yao, X. D., Freas, A., Ramirez, J., Demirev, P. A., and Fenselau, C. (2001) Proteolytic ^{18}O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.* **73,** 2836–2842.
5. Tabb, D. L., MacCoss, M. J., Wu, C. C., Anderson, S. D., and Yates, J. R. 3rd. (2003) Similarity among tandem mass spectra from proteomic experiments: detection, significance, and utility. *Anal. Chem.* **75,** 2470–2477.
6. Liu, T., Qian, W. J., Strittmatter, E. F., et al. (2004) High-throughput comparative proteome analysis using a quantitative cysteinyl-peptide enrichment technology. *Anal. Chem.* **76,** 5345–5353.
7. Chen, W. N., Woodbury, R. L., Kathmann, L. E., et al. (2004) Induced autocrine signaling through the epidermal growth factor receptor contributes to the response of mammary epithelial cells to tumor necrosis factor alpha. *J. Biol. Chem.* **279,** 18,488–18,496.
8. Shen, Y., Zhao, R., Belov, M. E., et al. (2001) Packed capillary reversed-phase liquid chromatography with high-performance electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry for proteomics. *Anal. Chem.* **73,** 1766–1775.
9. Smith, R. D., Anderson, G. A., Lipton, M. S., et al. (2002) An accurate mass tag strategy for quantitative and high throughput proteome measurements. *Proteomics* **2,** 513–523.
10. Liu, T., Qian, W. J., Chen, W. N., et al. (2005) Improved proteome coverage by using high efficiency cysteinyl peptide enrichment: the human mammary epithelial cell proteome. *Proteomics* **5,** 1263–1273.
11. Qian, W. J., Camp, D. G. 2nd, and Smith, R. D. (2004) High throughput proteomics using Fourier transform ion cyclotron resonance (FTICR) mass spectrometry. *Expert Rev Proteomics* **1,** 89–97.
12. Belov, M. E., Anderson, G. A., Wingerd, M. A., et al. (2004) An automated high performance capillary liquid chromatography-Fourier transform ion cyclotron resonance mass spectrometer for high-throughput proteomics. *J. Am. Soc. Mass. Spectrom.* **15,** 212–232.
13. Qian, W. J., Liu, T., Monroe, M. E., et al. (2005) Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: The human proteome. *J. Proteome Res.* **4,** 53–62.
14. Petritis, K., Kangas, L. J., Ferguson, P. L., et al. (2003) Use of artificial neural networks for the prediction of peptide liquid chromatography elution times in proteome analyses. *Anal. Chem.* **75,** 1039–1048.
15. Horn, D. M., Zubarev, R. A., and McLafferty, F. W. (2000) Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. *J. Amer. Soc. Mass Spectrom.* **11,** 320–332.
16. Johnson, K. L. and Muddiman, D. C. (2004) A method for calculating $^{16}O/^{18}O$ peptide ion ratios for the relative quantification of proteomes. *J. Am. Soc. Mass Spectrom.* **15,** 437–445.

17. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75,** 4646–4658.

8

An Isotope Coding Strategy for Proteomics Involving Both Amine and Carboxyl Group Labeling

Fred E. Regnier

Summary

A stable isotope coding strategy is described for the analysis of all types of tryptic peptides, including those that are N-terminally blocked and from the C-terminus of proteins. The method exploits differential derivatization of amine and carboxyl groups generated during proteolysis as a means of coding. Carboxyl groups produced during proteolysis incorporate ^{18}O from $H_2{}^{18}O$. Peptides from the C-terminus of proteins were not labeled with ^{18}O unless they contained a basic C-terminal amino acid. Primary amines form controls, and experimental samples were differentially acylated after proteolysis with either $^{1}H_3$- or $^{2}H_3$-*N*-acetoxysuccinamide. When these two types of labeling were combined, unique coding patterns were achieved for peptides arising from the C-termini and blocked N-termini of proteins.

Key Words: Proteomics; isotope coding; C-terminal; N-terminal; TACT; ^{18}O coding.

1. Introduction

In vitro quantification based on stable isotope labeling requires that proteins or peptides be derivatized with an isotopically coded labeling agent. This can be done in several ways. One is by derivatizing a functional on an amino acid side chain of the protein or peptide. This is the basis for the popular ICAT method in which sulfhydryl groups on cysteine residues are coded ***(1,2)***. Among the many advantages of this approach are that differentially coded proteome samples can be mixed and digested simultaneously, eliminating differences in proteolysis between samples. Another is that the coding agent contains biotin and only cysteine-containing peptides are selected from a proteome

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

digest. Greater than 80% of the peptides in proteolytic digests are eliminated through affinity selection while still sampling a major portion of the proteome ***(3)***. A limitation of side chain derivatization is that a protein might not have the specific amino acid being derivatized ***(4)*** or it is present in such low abundance that protein identification could be based on a single peptide. There is also the problem that unless a posttranslationally modified peptide contains the derivatizable side chain can it be selected and studied ***(5)***.

Another coding strategy is to derivatize either amino ***(6)*** or carboxyl groups ***(7)*** arising from proteolysis of the proteome. Peptides thus formed contain a basic amino acid at their C-terminus and a primary amine at their N-terminus, even in the case of peptides with posttranslational modifications (PTMs). Primary amine groups formed during proteolysis have been stable isotope coded in a number of ways ranging from acylation and reductive amination to dansylation and isocyanate derivatization.

Stable isotope coding of carboxyl groups is achieved exclusively by incorporation of ^{18}O from $H_2{}^{18}O$ into the C-terminus of peptides. When proteolysis is achieved with a serine protease, a reversible covalent linkage is formed between peptide products and the active site serine residue. Even after proteolysis, peptides still continue to be covalently bound and released from serine proteases. The reversibility of this reaction makes it possible to incorporate two moles of ^{18}O into the C-terminal carboxyl group of a peptide.

A problem with all these methods is that they are not truly global. Each has cases in which a class of peptides will fail to be derivatized. Limitations of amino acid side chain derivatization have already been noted in this respect. When the N-terminus of a protein is acylated, as is the case with many serum proteins, stable isotope coding methods that target α-amino groups cannot label the amino terminal peptide of the protein. The problem with ^{18}O coding is that it never codes the C-terminal peptide from a protein unless the C-terminal amino acid is either lysine or arginine. The same is true if the C-terminus of the protein is derivatized in some way. **Figure 1A** shows the general structure of the three types of tryptic peptides derived from none end-blocked proteins. The C-terminal peptide does not contain any basic amino acids, whereas all other peptides contain either an arginine or lysine residue at the C-terminus. **Figure 1C** shows three classes of tryptic peptides obtained from either an N-terminally or C-terminally blocked protein. A peptide with a methyl ester at the C-terminus is shown as an example of a C-terminally blocked protein. Peptides from N-terminally blocked proteins are indicated as being derivatized with an “acyl” group. The other amino acids that could be at the C-terminus of a peptide are indicated by $R_{1–20}$. No missed cleavages or adjacent basic amino acid-containing peptides are shown.

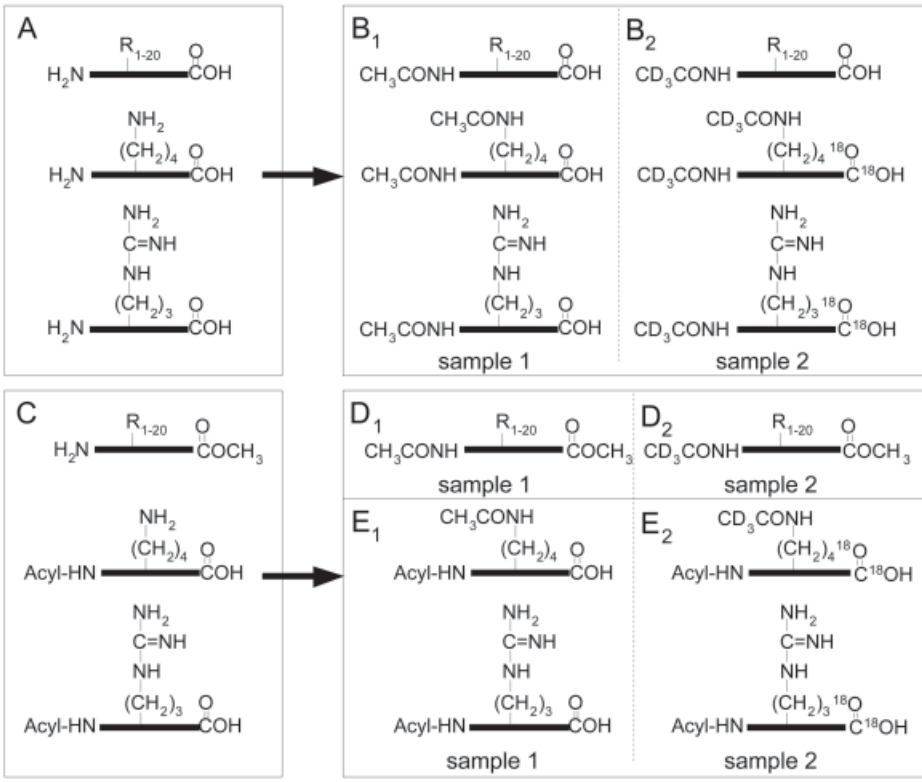


Fig. 1. A scheme for tagging amino and carboxyl termini of peptides simultaneously. The amino acid side chains in this illustration are on the C-terminal amino acid of the peptide. The bold solid line is the rest of the peptide backbone. **(A)** illustrates the general types of peptides obtained from unblocked proteins. **(B1)** shows how these peptides would be derivatized when a proteome sample was tryptic digested in $H_2{}^{16}O$ and derivatized with *N*-($^1H_3$)acetoxy succinimide. **(B2)** shows how peptides from a second sample would be derivatized when they were tryptic digested in $H_2{}^{18}O$ and derivatized with *N*-($^2H_3$)acetoxy succinimide. **(C)** illustrates peptides that would be obtained from proteins that are either N-terminally or C-terminally blocked. A methyl ester is used as an example of a C-terminal block. Differential labeling of the C-terminally blocked peptides from two samples is illustrated in **(D1)** and **(D2)**. Differential labeling of the N-terminally blocked peptides from these samples is seen in **(E1)** and **(E2)**. Note that in the case of terminally blocked peptides, only one end of the peptide is labeled. After differential labeling of the samples they are mixed and analyzed by liquid chromatography–mass spectrometry.

These problems can be circumvented by labeling tryptic peptides with acetate on their primary amino groups and ^{18}O in the carboxyl groups of basic amino acids at their C-termini. The products of the differential labeling process are seen in **Fig. 1**. Peptides from the A and C panels labeled with the light form of acetate are seen in panels B_1, C_1, and E_1, respectively. Labeling occurs in all cases except those that are N-terminally blocked with an "acyl" group. Carboxyl groups at the C-termini of peptides are not labeled in this case because proteolysis was carried out in $H_2{}^{16}O$.

When proteolysis is carried out in $H_2{}^{18}O$ and primary amines are labeled with trideuteroacetate the labeling pattern in panels B_2, D_2, and E_2 is achieved. This method of simultaneously tagging amino and carboxyl termini is referred to as the TACT method ***(8)***.

2. Materials

2.1. Reagents

1. The reference proteins bovine cytochrome c, chicken lysozyme, and turkey lysozyme (Sigma, St. Louis, MO).
2. Sequence grade trypsin (Promega, Madison, WI).
3. Reference peptides (Bachem, Torrance, CA).
4. Oxygen-18 enriched water (95–98 atom% ^{18}O) (Isotech, Miamisburg, OH).
5. C^2H_3COOH (Isotech).
6. *N*-hydroxysuccinimide (Aldrich, Milwaukee, WI).
7. *N*-hydroxylamine, iodoacetic acid, cysteine, dithiothreitol, *N*-tosyl-L-lysyl chloromethyl ketone, and all the reagents for trypsin digestion were purchased from Sigma.
8. The MALDI matrix α-cyno-4 hydroxycinnamic acid (Aldrich).

2.2. Solvents

1. Acetonitrile (ACN) (Sigma).
2. Reagent grade trifluoroacetic acid (TFA) (Sigma).
3. Hexane (Aldrich).

2.3. Equipment

1. MALDI-mass spectrometry (MS). MALDI-MS was performed using a Voyger DE-RP BioSpectrometry Workstation (PE Biosystems, Framingham, MA). The matrix used to analyze the peptides was a solution containing 10 mg/mL of α-cyno-4 hydroxycinnamic acid in 50% ACN, 50% water, and 0.1% TFA. Dried fractions collected from RPC separations were reconstituted in 20 µL of the matrix solution and 1 µL of the sample was spotted onto the plate for analysis. All peptides were analyzed in the reflective, positive ion mode by delayed extraction.
2. Electrospray ionization–MS. Samples were analyzed by an electrospray ionization source on a QSTAR hybrid Q-TOF MS (Applied Biosystems/PE SCIEX,

Foster City, CA) equipped with an Integral Liquid Chromatography Work Station (Applied Biosystems). Samples were dissolved in ACN-H_2O (50:50) containing 0.1% formic acid, and infusion injected with a syringe pump flowing at 10–30 μL/min. The QSTAR was operated above 8000 in resolution with a mass accuracy of 10–30 parts per million using an external calibration maintained for 24 h. MS/MS sequencing was achieved by selecting the parent ion in the quadruple using the unit-resolution selection. The instrument was run in the positive TOF mode (*m/z* 300–2000). The typical ion spray voltage used was 5000 V with a collision energy of 30–60 eV to fragment peptides. Nitrogen gas was used for the collision gas and typical pressures in the collision cell during MS/MS were from 4 to 6 $\times 10^{-6}$ torr. Sequences were obtained from the MS/MS spectrum through manual interpretation and with the MASCOT software (Matrix Sciences, London, UK).

3. Liquid chromatography. RPC separations were achieved using either a BioCAD Liquid Chromatography Works Station or an Integral Liquid Chromatography Work Station (Applied Biosystems) fitted with a 2 × 250 mm or 4.6 × 250 C_{18} PepMap (Applied Biosystems) column packed with 5 μm particle diameter reversed-phase sorbent material.

3. Methods

3.1. Coding Agent Synthesis

The synthesis of *N*-($^1H_3$)acetoxysuccinamide and *N*-($^2H_3$)acetoxysuccinamide were achieved in an identical manner. The deuterated reagent was synthesized by substituting ($^2H_3$)acetic anhydride for ($^1H_3$)acetic anhydride in the protocol. These reagents were synthesized according to the literature ***([9])***.

1. A solution of 4.0 g (34.8 mmol) of *N*-hydroxysuccinimide in 10.7 g (105 mmol) of ($^1H_3$)acetic anhydride was stirred for 15 h at room temperature. White crystals began to deposit in 10 min.
2. The reaction was terminated by filtration of the crystalline residue.
3. The crystals were then washed with hexane to remove residual acetate and acetic anhydride and dried in vacuum. A 5.43 g (100%) yield of product was obtained, melting point 133–134°C.

3.2. Proteolysis

1. Control and experimental proteome samples with total protein concentrations in the range of 0.1 to 1 mg/mL were individually reduced and alkylated in 0.1 *M* ammonium bicarbonate buffer (pH 8.2) containing 6.2 *M* urea and 10 m*M* dithiothreitol. The difference in the treatment of these two samples was that the proteolysis step was carried out in $^{18}H_2O$ with the experimental sample.
2. After a 2-h incubation at 37°C, iodoacetic acid was added to each of these samples to a final concentration of 30 m*M*, and incubated in darkness on ice for two more hours.

3. Cysteine was then added to both the control and experimental sample reaction mixtures to a final concentration of 40 m*M*, and the reaction was allowed to proceed at room temperature for 30 min.
4. After dilution with 0.1 *M* ammonium bicarbonate buffer to a final urea concentration of 0.8 *M*, sequence grade trypsin (2% [w/w] enzyme to that of protein based 1 mg/mL of protein) was added and incubated for 24 h at 37°C.
5. Proteolysis was terminated by the addition of *N*-tosyl-L-lysyl chloromethyl ketone in slight molar excess to the trypsin concentration.

3.3. Peptide Derivatization With N-Acetoxysuccinimide Coding Agent

Both affinity chromatography-selected samples on nonaffinity selected samples were treated in the same manner in this step of the analytical process. Experimental samples were treated in exactly the same manner as control samples except that *N*-($^2H_3$)acetoxysuccinamide was used as the isotopic coding agent instead of *N*-($^1H_3$)acetoxysuccinamide in the control sample.

1. A fivefold molar excess of *N*-($^1H_3$)acetoxysuccinamide was added to the tryptic digest of the control sample and incubated at room temperature for 4 h.
2. *N*-hydroxylamine was then added to adjust to pH 12.0. Incubation with hydroxylamine for 10 min was used to hydrolyze esters that might have been formed during the acylation reaction, especially tyrosine esters.
3. Equal aliquots of this sample and the isotopically coded experimental sample described below were mixed and then separated on a C_{18} RPC column prior to mass spectral analysis.

3.4. RPC of Isotopically Labeled Peptides

1. Derivatized peptides were fractionated using a 250 × 4.6-mm C_{18} PepMap (Applied Biosystems) reversed-phase column with a 60 min gradient ranging from 100% solvent A (0.1% TFA with 1% ACN/99% water) to 80% solvent B (0.1% TFA with 90% ACN/10% water). Peptide elution was monitored at 214 nm.
2. RPC fractions where either transferred directly to a mass spectrometer through an electrospray interface or collected and examined by MALDI-MS.
3. Collected fractions were evaporated to dryness and resuspended in MALDI matrix before MALDI analysis.

4. Notes

As opposed to other forms of stable isotope coding, coded peptides from the TACT method vary in mass difference between the heavy and light forms of peptides. The notes below provide a set of rules for interpretation.

1. C-terminal peptides from nonblocked proteins. Peptides derived from the C-terminus of proteins that are not blocked and do not have a basic amino acid at their C-terminus will only be labeled with d_0/d_3 acetate at the N-terminus. (Recall

that ^{18}O is not incorporated into peptides derived from the C-terminus of proteins.) This means the difference in ion clusters seen in the mass spectra will by 3 amu.

2. N-terminal peptides from nonblocked proteins. These peptides will have both a free α-amino group and a basic amino acid at the C-terminus that can incorporate two moles of ^{18}O. After labeling with $^{2}H_3$-acetate, heavy peptide isotopomers with an arginine at their C-terminus will contain three ^{2}H and two ^{18}O. This means they will have 3 + 4 = 7 amu higher molecular weight that the light form of the peptide.

   Peptides with a lysine at their C-terminus are also acetylated on their ε-amino group so the difference in the molecular weight of heavy and light forms will be 3 + 3 + 4 = 10 amu.
3. Peptides from the interior of proteins. Miscleavages by trypsin can occasionally generate peptides that have two lysine residues. In this case, the mass difference in the heavy and light forms of peptides would be 3 + 3 + 3 + 4 = 13 amu.
4. Peptides from the blocked C-termini of proteins. Because the C-terminal peptides arising from proteins do not incorporate ^{18}O anyway, stable isotopelabeling patterns can not differentiate between unblocked and blocked C-termini of proteins.
5. Peptides from the blocked N-termini of proteins. A block at the N-terminus of a protein yields a peptide without a free α-amino group. When these peptides have an arginine at their C-terminus, they are only labeled at their C-terminus with ^{18}O. This means the difference in mass between the heavy and light forms of peptides will be +4. Peptides that contain a lysine residue at the C-terminus will be labeled by d_3-acetate and ^{18}O. This means the heavy and light forms differ in mass by 3 + 4 = 7 amu. This is the same difference as for an internal peptide with an arginine at the C-terminus.
6. Interpretation rules can be broken down into four parts. The first is that peptide clusters differing by 3 amu always come from the C-terminus of a protein. The second is that peptide clusters differing by 4 amu always come from the N-terminus of a protein and contain arginine at their C-terminus. A third is peptide clusters varying by more than 7 amu always come from the interior of a protein and have one or more lysine residues. And the fourth is that peptide clusters varying by 7 amu can either come from the interior of a protein and carry arginine at their C-terminus, or be from a N-terminally blocked protein and have a lysine residue at their C-terminus. Differentiation between these two is easily achieved by collision-induced dissociation analysis in MS/MS analysis. Peptides from N-terminally blocked proteins will have all the heavy isotopes in the C-terminal lysine residue. In contrast, internal peptides with arginine at the C-terminus will be heavy isotope labeled at both ends of the molecule. This is very apparent in the "b" and "y" ions in tandem mass spectra.
7. Based on the knowledge that the TACT method provides global labeling of all peptides, all differentially code peptide mixtures should appear in mass spectra as doublet clusters according to the rules outlined in **step 6**. It is therefore surprising when single clusters of ions are found in mass spectra. (Singlet clusters

refers to the normal envelope of M, M+1, and M+2 ions seen in a mass spectrum.) Singlet clusters give very important clues about differences in biological systems. Singlets may occur in the following ways. One would be when there is a profound regulatory change that causes huge changes in the concentration of a protein between the two samples. In this case, other peptides from the same protein should show the same large change in relative concentration. A second reason would be because of a PTM. A PTM will change both the molecular weight and probably the chromatographic properties enough that peptides will not coelute from the RPC column and appear in the expected 3, 4, 7, 10, or 13 amu spacing between doublet clusters. Finally there is the possibility of a single amino acid polymorphism ***(10)***. Genetic differences, i.e., mutations, between the organisms providing the two samples being used in the differential analysis can be another reason that the expected 3, 4, 7, 10, or 13 amu spacing between doublet clusters will not be seen. Peptides will appear as singlets for the same reason as with a PTM.

Acknowledgments

The author gratefully acknowledges financial support from NIH grants R01 GM 59996 and RO1 AG025362 during the development of the methods described in this chapter.

References

1. Gygi, S. P., Rist, B., Griffin, T. J., and Eng, J. (2002) Proteome analysis of low abundance proteins using multidimensional chromatography and isotope-coded affinity tags. Aebersold, R. *J. Proteome. Res.* **1,** 47–54.
2. Han, D. K., Eng, J., Zhou, H., and Aebersold, R. (2001) Quantitative profiling of differentiation-induced microsomal proteins using isotope coded affinity tags and mass spectrometry. *Nat. Biotechnol.***19,** 946–951
3. Gulna, T., Purvine, S. O., Yi, E. C., et al. (2003) Quantitative proteomic analysis indicates increased synthesis of a quinolone by Pseudomonas aeruginosa isolates from cystic fibrosis pathways. *Proc. Natl. Acad. Sci. USA* **100,** 2771–2776.
4. Wang, S., Zhang, X., and Regnier, F. E. (2002) Quantitative proteomics strategy involving the selection of peptides containing both cysteine and histidine from tryptic digests of cell lysates. *J. Chromatogr. A* **949,** 153–162.
5. Rabilloud, T. (2002) Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains. *Proteomics* **2,** 3–10.
6. Mahmoud, H. and Righetti, P. G. (2002) Modern strategies for protein quantification in proteome analysis: advantages and limitations. *Mass Spectrom. Rev.* **21,** 287–302.
7. Regnier, F. E., Riggs, L., Zhang, R., et al. (2002) Comparative Proteomics based on stable isotope labeling and affinity selection. *J. Mass Spectrom*. **37,** 133–145.
8. Schevchenko, A. and Schevchenko, A. (2001) Evaluation of the efficiency of in-gel digestion of proteins by peptide isotope labeling and MALDI mass spectrometry. *Anal. Biochem.* **296,** 279–283.

9. Liu, P. and Regnier, F. (2002) An isotope coding strategy for proteomics involving both amine and carboxylic group labeling. *J. Proteome Res.* **1,** 443–450.
10. Ji, J., Chakraborty, A., Geng, M., et al. (2000) Strategy for qualitative and qauntitative analysis in proteomics based on signature peptides. *J. Chromatogr. B* **745,** 197–210.
11. Liu, P. and Regnier, F. (2003) Recognizing single amino acid polymorphism in proteins. *Anal. Chem.* **75,** 4956–4963.

9

Proteolytic Labeling With ^{18}O for Comparative Proteomics Studies

Preparation of ^{18}O-Labeled Peptides and the $^{18}O/^{16}O$ Peptide Mixture

Catherine Fenselau and Xudong Yao

Summary

The method reported here uses proteolytic catalysis to introduce two ^{18}O atoms into the carboxyl termini of peptides in mixtures, and is intended to be part of the work-flow in comparative proteomics strategies. Proteins are first cleaved with trypsin in water, and subsequently the peptide products are dried and labeled by incubation with trypsin in ^{18}O-enriched water. One important aspect of this two-step procedure is that peptides, and not proteins, are dried and redissolved in $H_2{}^{18}O$ for the labeling reaction. Incorporation can exceed 95% if it is carried out in water that is sufficiently enriched with $H_2{}^{18}O$. The byproduct of the reaction is water. The use of catalytic enzyme immobilized on beads facilitates its removal and termination of the exchange. In differential proteomic studies, heavy isotope-labeled peptides are combined with peptides carrying ^{16}O for isotope ratio measurements by mass spectrometry.

Key Words: Comparative proteomics; isotope labeling; proteolytic catalysis; $^{18}O/^{16}O$ labeling; mass spectrometry; differential analysis; non-gel proteomics; peptide analysis; isotope ratios; immobilized trypsin; serine proteases.

1. Introduction

The use of mass spectrometry (MS) in combination with bioinformatics has elevated proteomics to the level of a new paradigm. As the inventory stage implemented by this paradigm matures, scientists are focusing on the enormous potential of comparative studies or differential proteomics. The histori-

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

cally important capability of MS to provide measurements of isotope ratios has been newly exploited to provide reliable analyses of pairs or multiples of peptides carrying heavy and light labels. Although these labels can be introduced metabolically into cells growing in culture, protein samples from animals or clinical specimens are most practically labeled by chemical or enzymatic reactions. Advantage has been taken of the mechanism of serine proteases to introduce ^{18}O atoms into the carboxyl termini of peptides. The introduction of labeled oxygen in the hydrolysis of proteins was reported more than 25 yr ago ***(1)***. The ability of trypsin and related enzymes to rebind the peptide products and to introduce a second atom of ^{18}O by microreversibility was documented in 1996 ***(2)***. This mechanism is shown in **Fig. 1** and forms the basis for the labeling method reported in this chapter.

The reaction shown in **Fig. 1** allows protein labeling to be separated from peptide cleavage. This adds greatly to the versatility of the method. It means that after proteins are cleaved under optimal conditions, the peptide products are dried and redissolved in labeled water, eliminating the need to resolubilize proteins. The two-step strategy ***(3)*** means, for example, that proteins can be first digested with Lys-C endoproteinase, which is active at a higher urea concentration, and then the peptide products can be cleaved and labeled with trypsin. Peptides labeled with ^{18}O are combined with peptides carrying ^{16}O, prepared in parallel from another protein sample, and the isotope ratios are measured by MS. This is a global labeling strategy. Only the peptide carrying the original carboxyl terminus without a terminal arginine or lysine is not labeled. It should be pointed out that this labeling method is effective with both lysine- and arginine-terminated peptides, as well as both long and short peptides. Different peptides are labeled at different rates. Thus, the exchange is extended with the addition of a second batch of immobilized trypsin. A major advantage is that the $^{18}O/^{16}O$ peptide pairs coelute from reverse-phase high-performance liquid chromatography (HPLC) columns. A spectrum of such a peptide pair is shown in **Fig. 2**, along with a theoretical distribution of natural isotopes. The ratio of the abundance of the ^{18}O-labeled peptide to the ^{16}O-labeled peptide can be calculated from peak heights of the monoisotopic species, corrected for ^{13}C contributions and single ^{18}O incorporation by the formula:

$$\text{Ratio} = \{I_4 - (M_4/M_0)I_0 - (M_2/M_0)[I_2 - (M_2/M_0)I_0] + [I_2 - (M_2/M_0)I_0]\} / I_0$$

In a controlled study using Glu-C endoprotease the ratio of isotopic peaks was shown to correlate linearly with independently determined peptide ratios, with a correlation coefficient of 0.997 ± 0.238 ***(4)***. The precision of the liquid chromatography (LC)–MS measurement has been shown to be less than ± 10%, using either the previously mentioned equation or the simpler ratio of peak heights I_4/I_0. Variation in sample preparation and biological variation can con-

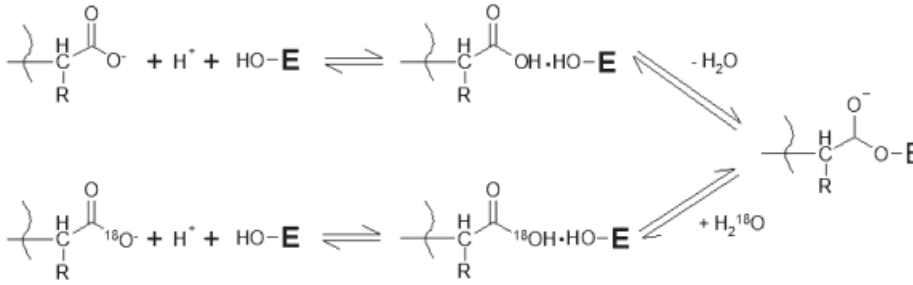


Fig. 1. Mechanism for introduction of ^{18}O into peptides, catalyzed by a serine protease. Introduction of two atoms of ^{18}O is achieved when the reaction is repeated.

tribute additional uncertainty to the measurement (*see* **Notes 1–5**).The method has been used with most types of mass spectrometer analyzers, including Fourier Transform analyzers ***(3,5)*** and low-resolution ion traps ***(6–8)***, and is suitable for high-throughput protocols ***(7)***. Both MALDI and electrospray ionization (ESI) have been used successfully. Because the isotope pair is located at the carboxyl terminus of each peptide, all fragment ions that contain the carboxyl terminus will be marked by isotope doublets. This facilitates interpretation of tandem mass spectra. An example is shown in **Fig. 3**.

2. Materials

The more enriched the isotope-labeled water, the more completely labeled the peptides will be.

1. $H_2{}^{18}O$ (>97% enrichment; Isotec Inc., Miamisburg, OH).
2. $H_2{}^{16}O$ (HPLC grade or deionized).
3. Stock solution (0.5 µg/µL) of modified trypsin in $H_2{}^{16}O$. Into a vial of 100 µg of trypsin (treated with reductive methylation and L-[tosylamido-2-phenyl] ethyl chloromethyl ketone [TPCK]; Trypsin Gold, mass spectrometry grade, Promega, Madison, WI), add 200 µL of $H_2{}^{16}O$. Transfer aliquots of the trypsin solution in desired volumes into nonstick Eppendorf tubes. Fast freeze the enzyme stock solutions and store at –80°C (*see* **Note 6**).
4. Digestion buffer: 20 m*M* Tris-HCl, 50 m*M* NaCl, and 50 m*M* $CaCl_2$. Mix 2 mL of 100 m*M* Tris-HCl stock solution (pH 8.0; Sigma-Aldrich, St. Louis, MO), 5 mL of 100 m*M* $CaCl_2$, 100 µL of 5 *M* NaCl stock solution and 2.9 mL of $H_2{}^{16}O$.
5. Prewet microcentrifugal filter. Add 100 µL of a solution of the digestion buffer and acetonitrile (8:2 v/v) to a microcentrifugal filter (UFC30VV25; Millipore, Bedford, MA) and spin the filter unit at 4000*g* for 1 min.
6. Prewashed immobilized trypsin beads. Add 100 µL of a solution of the digestion buffer and acetonitrile (8:2 v/v) on top of a prewet microcentrifugal filter and then add 40 µL of the immobilized trypsin resin in suspension (Applied Biosystems, Framingham, MA or Stratagene, Inc., La Jolla, CA). Mix well by reversing the

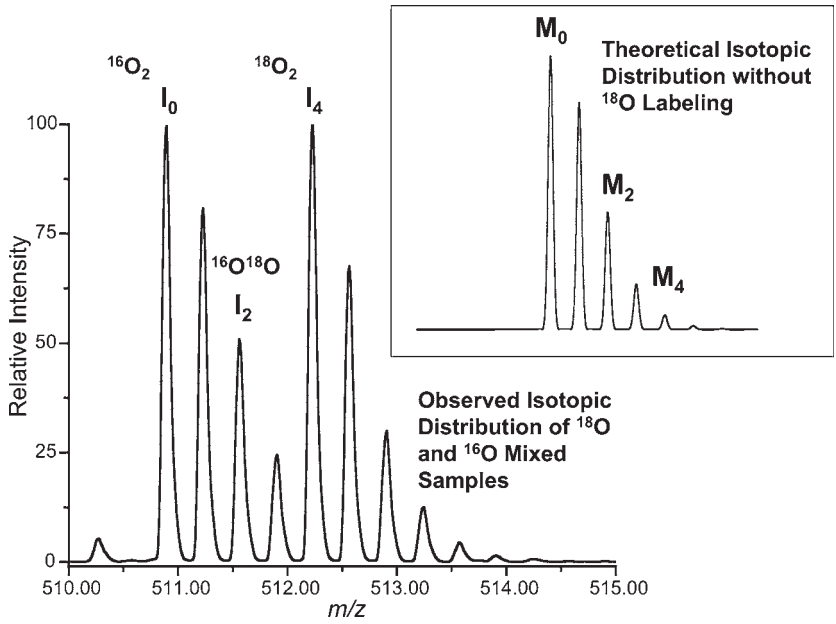


Fig. 2. Partial mass spectrum of a pair of triply charged peptides carrying ^{18}O or ^{16}O labels. The inset shows the theoretical isotope distribution, reflecting mainly the natural abundance of ^{13}C.

unit 10 times and spin at 4000*g* for 1 min. Wash the immobilized trypsin beads with 100 µL portions of a solution of the digestion buffer and acetonitrile (8:2 v/v) for four more times. Add 40 µL of digestion solution to resuspend the immobilized trypsin.

7. Anhydrous acetonitrile (Sigma-Aldrich).
8. Trifluoroacetic acid (TFA) solution (10% v/v). Add 10 µL of TFA (Sequanal grade, Pierce, Milwaukee, WI) to 10 mL of $H_2{}^{16}O$.
9. TFA solution (1% v/v). Add 100 µL of TFA (Sequanal grade) to 10 mL of $H_2{}^{16}O$.

3. Methods

In this strategy, proteins are cleaved first and then the peptides are labeled with either ^{18}O or ^{16}O. Peptides from one tissue or cell sample are labeled with ^{18}O, and peptides from a second sample are labeled with ^{16}O. The peptides are mixed for further fractionation (*see* **Notes 7–9**) and isotope ratio measurements by LC–MS, LC-ESI-MS/MS, and LC-MALDI-MS/MS.

1. Dissolve a mixture of reduced and alkylated proteins (100 µg, desalted) in a 0.5-mL nonstick tube, by adding 5 µL acetonitrile and 20 µL of the digestion solution (*see* **Note 10**).

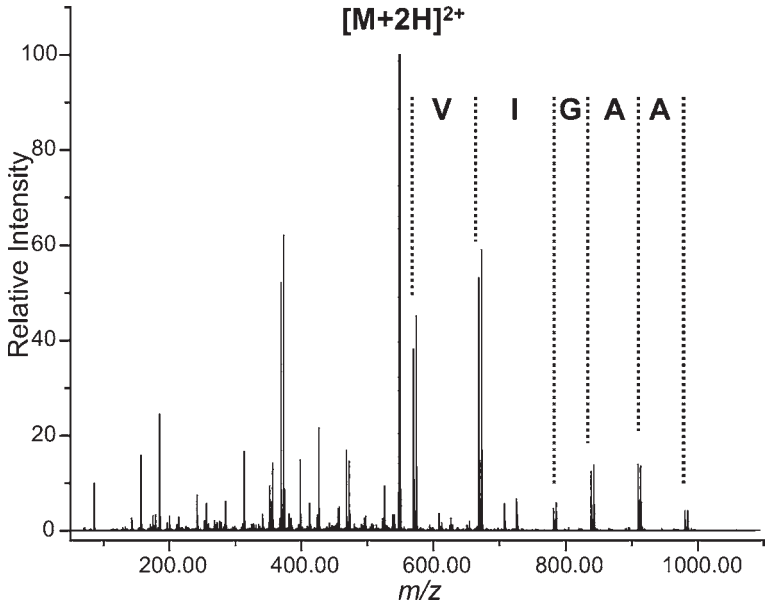


Fig. 3. Tandem mass spectrum of a pair of peptides carrying ^{18}O or ^{16}O labels at their carboxyl termini. Pairs of fragment ions containing the C-termini (*y* ions) are indicated, from which a partial sequence has been deduced.

2. Add 4 µL trypsin stock solution (0.5 µg/µL) in $H_2{}^{16}O$.
3. Spin the tube at 1000*g* for 0.5 min and incubate the digestion solution without shaking and at room temperature for 6 h.
4. Add 20 µL more of the digestion buffer.
5. Mix the prewashed immobilized trypsin beads by reversing the microcentrifugal filter unit 10 times.
6. Add 10 µL of prewashed immobilized trypsin beads to the digestion solution.
7. Dry the digestion mixture by vacuum centrifugation (SpeedVac, Thermo Savant, Holbrook, NY) at room temperature.
8. Add 40 µL of anhydrous acetonitrile, slightly tap the tube (to mix), and then incubate for 5 min. Dry the mixture by vacuum centrifugation at room temperature one more time.
9. Dissolve the dried mixture in 10 µL of anhydrous acetonitrile.
10. Add 40 µL of $H_2{}^{18}O$ and gently vortex the tube.
11. Incubate the suspension with gentle rotation overnight.
12. Acidify the reaction suspension with a minimal volume of 10% (v/v) TFA to pH < 3.0 (*see* **Notes 11–13**).
13. Transfer the reaction suspension onto a prewet microcentrifugal filter and spin at 4000*g* for 1 min (*see* **Note 14**).

14. Wash the reaction tube with 50 μL of a solution of 1% TFA:acetonitrile (8:2 v/v) and transfer the washing solution to the filter and spin at 4000*g* for 1 min.
15. Repeat **step 14** once. Combine all the filtrates from **steps 13–15**.
16. Dry the reaction products by vacuum centrifugation at room temperature.
17. Add 200 μL of anhydrous acetonitrile and dry the reaction products by vacuum centrifugation at room temperature one more time.
18. Store the ^{18}O-labeled peptide mixture for future analyses.
19. Prepare ^{16}O-labeled peptides by following the same procedure in $H_2{}^{16}O$.
20. Dissolve and combine the $^{16}O/^{18}O$-labeled peptide mixtures for mass spectrometric analysis.

4. Notes

1. This labeling method is reported to be compatible with smaller sample quantities than others ***(8)***. Formation of the enzyme/peptide adduct is driven by high concentrations of immobilized trypsin, despite low concentrations of target peptides. The large molar excess of water drives the exchange reaction to completion.
2. In the current literature, the dynamic range has been limited to about 1;20 for $^{16}O/^{18}O$ or $^{18}O/^{16}O$ ***(9,10)***.
3. Accuracy for ratios of peptide pairs with a significantly low concentration of the ^{18}O-peptide may be increased by using an inverse labeling strategy ***(6)***.
4. Isotope ratios are more difficult to determine accurately in mass spectra obtained in tandem experiments. Ratios should be measured in conventional scans.
5. Multiple peptide ratios should not be averaged to obtain the protein ratio, because peptides are labeled at varying rates ***(3)***. Rather, the ratio that involves the peptide with the highest incorporation of ^{18}O is generally the most meaningful.
6. Glu-C endoproteinase ***(4)*** and chymotrypsin ***(3)*** have been shown to catalyze the labeling reaction.
7. This differential labeling method can be combined with any affinity separation method. Bonenfant et al. have successfully combined it with metal-affinity chromatography ***(9)***.
8. Proteolytic labeling can be combined with glycolytic labeling, which marks the sites of attachment of *N*-glycosylated peptides ***(4)***.
9. Because the isotope labels are introduced only after proteins have been isolated and cleaved, care must be taken that both protein pools are processed analogously.
10. Anhydrous dimethyl sulfoxide can be used as an alternative to anhydrous acetonitrile as organic solvent to dissolve peptides for the exchange reaction. In this case 0.5% (v/v) of 2-mercaptoethanol should be included in the exchange suspension to prevent oxidation of peptides. Dimethyl sulfoxide should not be used for the digestion reaction.
11. Clay and Murphy have demonstrated that ^{18}O labels in a carboxylic acid group are relatively stable in the pH range 2.0 to 8.0, and that the labels are stable in acetonitrile gradients in reverse-phase HPLC ***(11)***. Aspartate and glutamate side chains will not be exchanged under the conditions of **Subheading 3.** used because they

do not evaporate during vacuum centrifugation and provide buffering capacity after reconstitution with labeling water.

12. Nonvolatile buffers should be avoided because they are not compatible with MS.
13. Staes and colleagues have carried out the postdigestion exchange at pH 4.5 ***(12)***, and Zang et al. ***(8)*** used pH 6.75, reporting that the exchange occurs more rapidly.
14. The use of immobilized trypsin allows high ratios of enzyme to peptide and minimizes trypsin autolysis. Most importantly, immobilized trypsin can be readily removed from the reaction (*see* **Subheading 3**
.), thus precluding catalysis of back exchange. A common practice is to reduce the pH of the exchange reaction solution to slow enzyme-catalyzed back exchange ***(11)***.

Acknowledgments

The authors acknowledge support for the development of this method from the National Institutes of Health, GM21248.

References

1. Rose, K., Simona, M. G., Offord, R. D., Prior, C. P., Otto, B., and Thatcher, D. R. (1983) A new mass spectrometric C-terminal sequencing technique finds a similarity between gamma-interferon and alpha 2-interferon and identifies a proteolytically clipped gamma-interferon that retains full antiviral activity. *Biochem. J.* **215,** 273–277.
2. Schnolzer, M., Jedrzejewski, P., and Lehmann, W. D. (1996) Protease-catalyzed incorporation of ^{18}O into peptide fragments and its application for protein sequencing by electrospray and matrix-assisted laser desorption /ionization mass spectrometry. *Electrophoresis* **17,** 945–953.
3. Yao, X., Afonso, C., and Fenselau, C. (2003) Dissection of proteolytic ^{18}O labeling: endoprotease-catalyzed ^{16}O to ^{18}O exchange of truncated peptide substrates. *J. Proteome Res.* **2,** 747–752.
4. Reynolds, K. J., Yao, X., and Fenselau, C. (2002) Proteolytic ^{18}O labeling for comparative proteomics: evaluation of endoprotease Glu-C as the catalytic agent. *J. Proteome Res.* **1,** 27–33.
5. Yao, X., Freas, A., Ramirez, J., Demirev, P. A., and Fenselau, C. (2001) Proteolytic ^{18}O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.* **73,** 2836–2842.
6. Wang, Y. K., Ma, Z., Quinn, D. F., and Fu, E. W. (2001) Inverse ^{18}O labeling mass spectrometry for the rapid identification of marker/target proteins. *Anal. Chem.* **73,** 3742–3750.
7. Heller, M., Mattou, H., Menzel, C., and Yao, X. (2003) Trypsin catalyzed ^{16}O to ^{18}O exchange for comparative proteomics: tandem mass spectrometry comparison using MALDI-TOF, ESI-QTOF, and ESI-ion trap mass spectrometers. *J. Am. Soc. Mass Spectrom.* **14,** 704–718.
8. Zang, L., Toy, D. P., Hancock, W. S., Sgroi, D. C., and Karger, B. L. (2004) Proteomic analysis of ductal carcinoma of the breast using laser capture microdissection, LC-MS and $^{16}O/^{18}O$ isotopic labeling. *J. Proteome Res.* **3,** 604–612.

9. Bonenfant, D., Schmelzle, T., Jacinto, E., et al. (2003) Quantitation of changes in protein phosphorylation: a simple method based on stable isotope labeling and mass spectrometry. *Proc. Nat. Acad. Sci. USA*. **100,** 880–885.
10. Brown, K. J and Fenselau, C. (2004) Investigation of doxorubicin resistance in MCF-7 breast cancer cells using shot-gun comparative proteomics with proteolytic ^{18}O labeling. *J. Proteome Res.* **3,** 455–462.
11. Clay, K. L. and Murphy, R. C. (1980) Back exchange of ^{18}O-labeled amino acids by erythrocytes: the possible role of amino acid transport. *Biochem. Biophys. Res. Commun*. **95,** 1205–1210.
12. Staes, A., Demol, H., Damme, J. V., Martens, L., Vandekerckhove, J., and Gevaert, K. (2003) Global differential non-gel proteomics by quantitative and stable labeling of tryptic peptides with oxygen-18. *J. Proteome Res.* **3,** 786–791.

10

Tandem Mass Spectrometry in the Detection of Inborn Errors of Metabolism for Newborn Screening

František Tureček, C. Ronald Scott, and Michael H. Gelb

Summary

Tandem mass spectrometry has been used for determinations of enzyme activities in biological samples. Activities in rehydrated dried blood spots of lysosomal enzymes glucocerebrosidase, acid sphingomyelinase, galactocerebroside β-galactosidase, acid-α-galactosidase, acid α-glucosidase, and α-D-iduronidase are measured simultaneously by multiple-reaction monitoring of ion dissociations from cations produced by electrospray ionization of enzymatic products. Simple and inexpensive assay protocols are described that are readily adopted for handling multiple samples in 96-well microtiter plates, employing simple separation steps, and using less than or equal to 3 μmol of synthetic or commercially available substrates, and less than 25 nmol of internal standards per analysis. The assays have the potential of being used for large-scale screening of newborns for the detection of inborn errors of metabolism.

Key Words: Lysosomal storage diseases; mucopolysaccharidosis; enzyme assays; tandem mass spectrometry; multiple reaction monitoring; clinical chemistry; biochemical diagnosis.

1. Introduction

Mass spectrometry (MS) is a highly sensitive method that has been widely used for the selective detection and quantitation of various metabolites that are present at trace levels in body fluids and manifest genetic diseases ***(1)***. In the traditional analytical approach to metabolite analysis, the compounds of interest are extracted from the biological material (urine, blood serum, and so on) ***(2)*** and separated by a gas chromatograph or a high-performance liquid chromatography that is coupled on-line to the mass spectrometer. The use of gas chromatography often requires that the metabolites be chemically derivatized

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

prior to analysis to increase their volatility ***(3)***. The mass spectrometer ionizes the metabolite molecules and provides information on the mass-to-charge ratios (m/z) of the ions formed, in addition to their chemical composition and structure. Owing to the known and often unambiguous relationship between gas-phase ions and their precursor analyte molecules, ion detection and analysis serves to identify and quantify the analytes. Molecular structure is typically inferred from ion dissociations occurring in the mass spectrometer that can be directly induced by ionization of the analyte molecules, or by collisional activation of stable ions produced from the analyte by a soft ionization method such as electrospray ionization (ESI) ***(4)***. In the latter case, ions are separated by their mass-to-charge ratios in the mass spectrometer, the ions of interest are selected, activated by collisions with gas, and one or several of their dissociation products are monitored, as described in **Figs. 1–3**. The group of techniques using dissociations of mass-selected ions is referred to as tandem mass spectrometry (MS/MS) ***(5)***.

In this chapter we describe a novel approach of using MS/MS for the detection of inborn errors of metabolism in patients. In particular, we focus on two groups of lysosomal storage diseases (LSD) that are caused by defective enzymes in the lysosome ***(6)***. The first group includes diseases caused by deficiencies of lysosomal enzymes involved in catabolic degradation of sphingolipids ***(7)***. The other group is mucopolysacharidoses that are caused by deficiencies of enzymes that catalyze degradation of glycosaminoglycans, such as heparan sulfate and dermatan sulfate ***(8)***. We note that the methods reported here have evolved from our previous studies of mass spectrometric methods for diagnosing enzyme deficiencies in cultured cells using affinity capture-elution ESI MS (ACESIMS) ***(9,10)*** that addressed several LSD ***(11,12)*** and congenital disorders of glycosylation ***(13)***. The advantage of MS/MS over other mass spectrometry-based approaches is that it (1) provides highly selective and efficient separation of ions from low-abundance analytes in the mass spectrometer, (2) increases overall selectivity and sensitivity, (3) works with rehydrated dried blood spots (DBS) from newborn screening cards as biological sample, and, last but not least, (4) uses simpler and less expensive substrates and procedures for routine use in clinical practice.

The use of MS/MS for enzyme assays in DBS is illustrated in **Fig. 1**. The sample, which is a 2–5 mm diameter punch of a DBS on a screening card, is incubated with a buffer containing a nonendogenous substrate conjugate that incorporates a structure moiety that is similar to that in the natural substrate and is recognized by the enzyme. Action of the enzyme forms the enzymatic product (P) whose molecular mass is different from that of the substrate conjugate. ESI produces gas-phase ions (Ionized P) by protonation, alkali metal ion attachment, or deprotonation of P that are directly related to its molecular mass

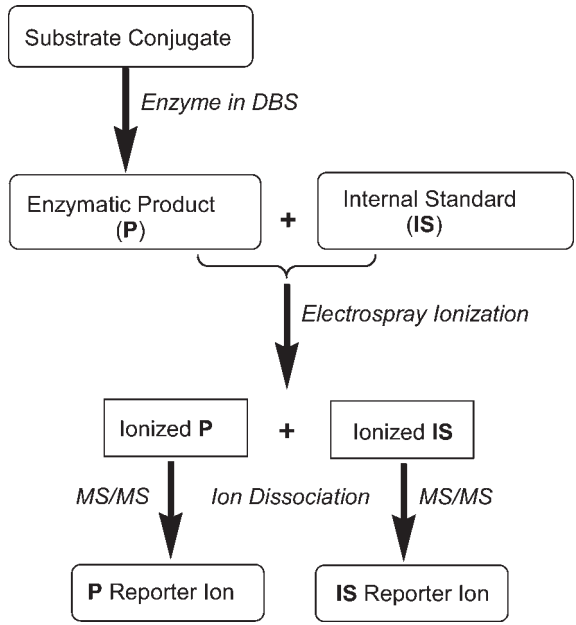


Fig. 1. Flow chart of quantitative analysis of enzyme activity in dried blood spots by electrospray ionization–tandem mass spectrometry.

and can be readily separated in the mass spectrometer from those of the substrate conjugate. Internal standard (IS) is added to P that upon ESI forms gas-phase ions that are chemically similar to ionized P, but differ in mass. Collisional activation in the tandem mass spectrometer produces fragment ions from ionized P and IS that are unequivocally related to P and IS and are used to monitor the ionized P and IS intensities. Because ion intensities from electrospray are proportional to analyte concentrations ***(4)***, the measured relative intensities of P and IS reporter fragment ions are proportional to P and IS concentrations. The proportionalities (response factors in ESI–MS/MS) are determined from calibration curves for P and IS. In case IS is an isotopomer of P, their ESI–MS/MS response factors are nearly identical and no external calibration is necessary.

Although ESI–MS/MS is suitable for the sensitive analysis of biomolecules in complex mixtures, ESI can be suppressed by the relatively large amounts of nonvolatile buffer and detergent components present in the DBS reaction mixtures. To remove the interfering components we developed a simple solid-phase extraction methods using silica gel or C_{18}-coated silica gel.

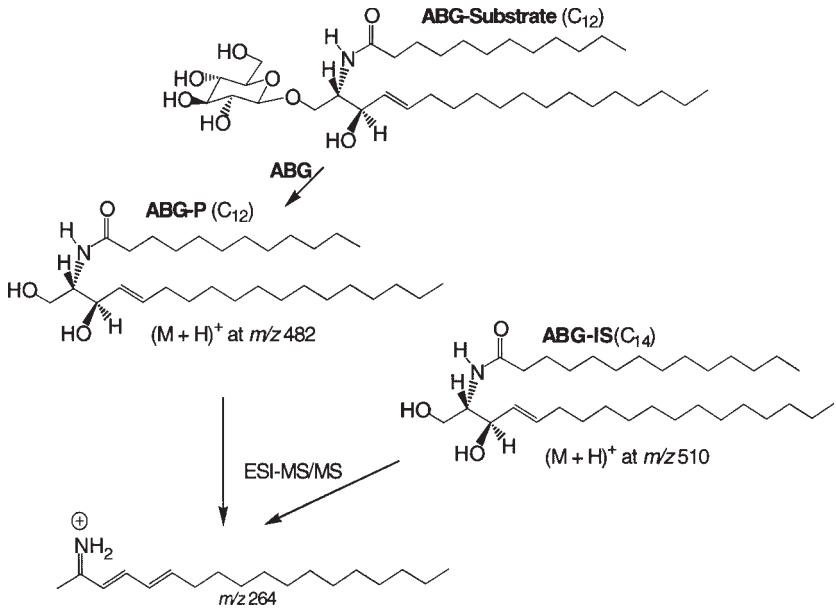


Fig. 2. Chemical, ionization, and ion dissociation processes in assaying glucocerebrosidase activity by electrospray ionization–tandem mass spectrometry.

The MS/MS analysis is carried out on a tandem quadrupole mass spectrometer, abbreviated as Q1-q2-Q3, which is an instrument commonly used in clinical laboratories. Q1-q2-Q3 mass spectrometers are currently available from ABI-Sciex, (Foster City, CA), ThermoElectron (Finnigan, Waltham, MA), or Waters (Micromass, Milford, MA). Precursor analyte ions that were produced by ESI are selected by the first quadrupole mass filter (Q1), allowed to dissociate in a collision multipole (q2), and their dissociation products are detected after passing through the second quadrupole mass filter (Q3). It should be noted that other types of tandem mass spectrometers could be used for MS/MS measurements, e.g., quadrupole-time of flight (Q-TOF), quadrupole-ion trap (Q-Trap), or TOF-TOF.

Figure 2 illustrates the principle of an MS/MS analysis of glucocerebrosidase (ABG) activity in DBS ***(14)***. ABG deficiency causes accumulation of its substrate, glucosylceramide, in macrophages (Gaucher disease). The enzyme is assayed by a glucosylceramide (ABG-substrate) containing an unnatural C_{12} acyl-containing ceramide. The deglucosylated C_{12}-ceramide, ABG-P, is ionized by ESI together with an added internal standard which is a C_{14}-ceramide, ABG-IS, giving precursor ions at *m*/*z* 482 and 510, respectively. Each of these ions is selected by Q1 and collisionally dissociated in q2 to pro-

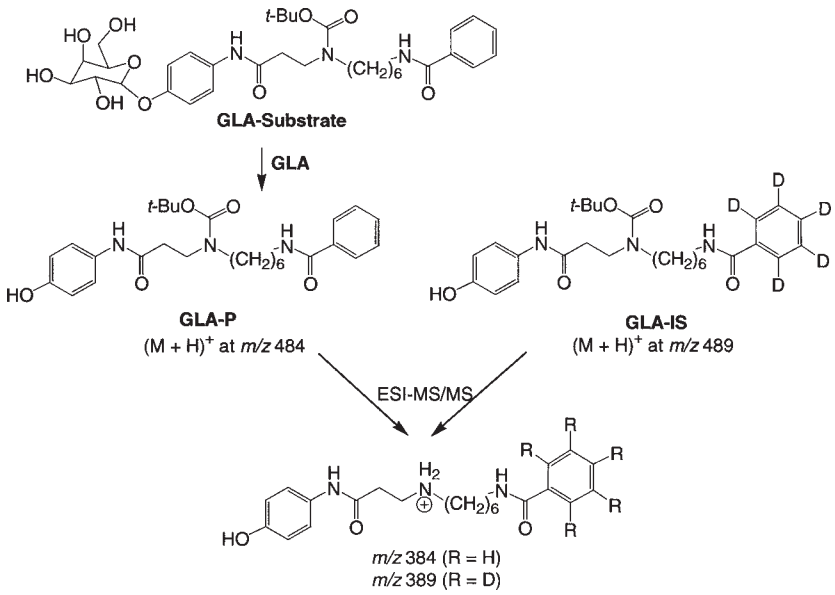


Fig. 3. Chemical, ionization, and ion dissociation processes in assaying acid-α-galactosidase activity by electrospray ionization–tandem mass spectrometry.

duce a common fragment ion at *m*/*z* 264. Q3 is set to transmit *m*/*z* 264, whereas Q1 is scanned over a mass window including *m*/*z* 482 and *m*/*z* 510 to produce a precursor ion scan. The ion intensities in the detection channels corresponding to *m*/*z* 482 → *m*/*z* 264 and *m*/*z* 510 → *m*/*z* 264 are then used to quantify the ion currents carried by ionized ABG-P and ionized ABG-IS and, thus, to determine the concentration of ABG-P relative to ABG-IS.

The same method is used to monitor ceramides produced by lysozomal enzymes acid sphingomyelinase (ASM-P and ASM-IS with C_6 and C_4 fatty acid chains, respectively) for the detection of Niemann-Pick type A/B disease, and galactocerebroside β-galactosidase (GALC-P and GALC-IS with C_8 and C_{10} fatty acid chains, respectively) for the detection of Krabbe disease ***(15)***. All of these ceramides give the common *m*/*z* 264 ion on ESI–MS/MS and can be monitored in one scan spanning the precursor ion *m*/*z* values, thus allowing the three enzymes to be assayed in one analysis.

Isotope labeling is used to monitor P from assays of acid-α-galactosidase (GLA) for the detection of Fabry disease, and lysosomal acid α-glucosidase (GAA) for the detection of Pompe disease, as shown in **Fig. 3**. GLA-P produced by enzymatic action on GLA-substrate is ionized by ESI together with GLA-IS. The product ion intensities from GLA-P (*m*/*z* 384) and GLA-IS (*m*/*z* 389) are measured in the multiple reaction monitoring (MRM) mode ***(16)*** that

alternatively records fragment ion intensities in the *m*/*z* 484 → *m*/*z* 384 and *m*/*z* 489 → *m*/*z* 389 channels. GAA is assayed in the same MRM mode by using homologous P and IS that are one CH_2 group (14 Da) heavier than those for GLA. The measured activities are summarized in **Table 1** ***(14)***.

MRM is also employed to assay the enzyme α-L-iduronidase (IDUA) ***(17)*** which is crucial for the degradation of glycosaminoglycans and whose deficiency causes mucopolysaccharidosis type I (Hurler disease) ***(8)***. The molecular structures and ion *m*/*z* ratios used for the ESI–MS/MS assay are summarized in **Fig. 4**. The measured IDUA activities in rehydrated DBS are available as supporting information to **ref.** ***17***.

2. Materials

2.1. Equipment

1. Pipetman (AccuTek, San Diego, CA; http://www.accuteklab.com).
2. 96-well microtiter plates (Greiner Bio-One, Longwood, FL; http://www.gbo.com/bioscience; cat. no. 655001).
3. Aluminum foil tape (Hillas Packaging, Inc., Fort Worth, TX; cat. no. 3M 425 3 X60).
4. 96-well polypropylene microtiter plate (E&K Scientific, Inc., Santa Clara, CA; cat. no. 21201).
5. Megatiter 96-well polypropylene deep-well plates (2.2-mL well volume) (Continental Lab Products, San Diego, CA; cat. no. 2045-Megatiter plates).
6. Vacuum manifold (Millipore, Billerica, MA; http://www.millipore.com; cat. no. MAVM0960R or MAVM0960T).
7. Teflon-lined covers (Cap Mats, E&K Scientific, cat. no. EK99116).
8. Well plate for IDUA assays (Innovative Microplate, Billerica, MA; http://www.innovativemicroplate.com; cat. no. F20005).
9. Silicagel, 230–400 mesh (Merck Chemicals, Darmstadt, Germany).

2.2. Dried Blood Spots

1. All DBS were obtained from anonymous donors through Seattle Children's Hospital and Regional Center (Dr. C. R. Scott) or from the Laboratory for Neurochemistry, Buenos Aires, Argentina (Dr. Mariana Blanco) and handled in compliance with the Institutional Review Board review. LSD- and mucopolysacharidoses-I-affected patients had been previously diagnosed with the particular disorder using established clinical and biochemical procedures. DBS from nonaffected individuals were obtained from adults (18–55 yr of age) and infants (3rd to 15th d postpartum).
2. DBS on report cards were kept at ambient temperature during shipment (less than 10 d) and then stored at –20°C in zip-lock plastic bags (one bag inside of a second bag). Zip-lock bags were kept in a sealed plastic box containing $CaSO_4$ granules as a desiccant. DBS for assays of ABG, ASM, GALC, GAA, and GLA

Table 1
Lysozomal Enzyme Activities in DBS[a,b]

Affected patients				Heterozygotes		Healthy adults				Healthy infants			
Enzyme	*n*	Maximum activity	Percent of mean of healthy adults[c]	*n*	Minimum activity	*n*	Max. activity	Min	Median	*n*	Max. activity	Min.	Median
ABG	6	0.18[d]	5	5	0.81	48	8.48	0.94	3.36	32	9.60	0.89	3.31
ASM	5	0.32	14	5	0.40	48	8.53	1.05	2.12	32	11.3	0.92	4.41
GALC	9	0.20	19	0		17	1.92	0.60	0.72	15	1.53	0.42	1.01
GAA	5	0.33	14	5	0.89	48	4.02	1.19	2.19	32	7.33	0.93	3.12
GLA	5	0.17	9	5	0.34	48	3.40	0.97	1.91	32	5.65	0.77	2.23

[a]For all assays except GALC, dried blood spots (DBS) were obtained between 2000 and 2003. For GALC assays, all DBS were obtained in 2003. All DBS were stored in sealed plastic bags at 4°C.

[b]All activities in μmol/h/(L blood).

[c]Maximum enzyme activity observed among the affected patients divided by the mean activity measured for healthy adults.

[d]Corrected by subtracting blank activities. The blank values (activity [SD]) were as follows: ABG: 0.12 (0.09), ASM: 0.10 (0.04), GALC: 0.21 (0.07), GAA: 0.06 (0.02), GLA: 0.15 (0.02), each from eight independent assays.

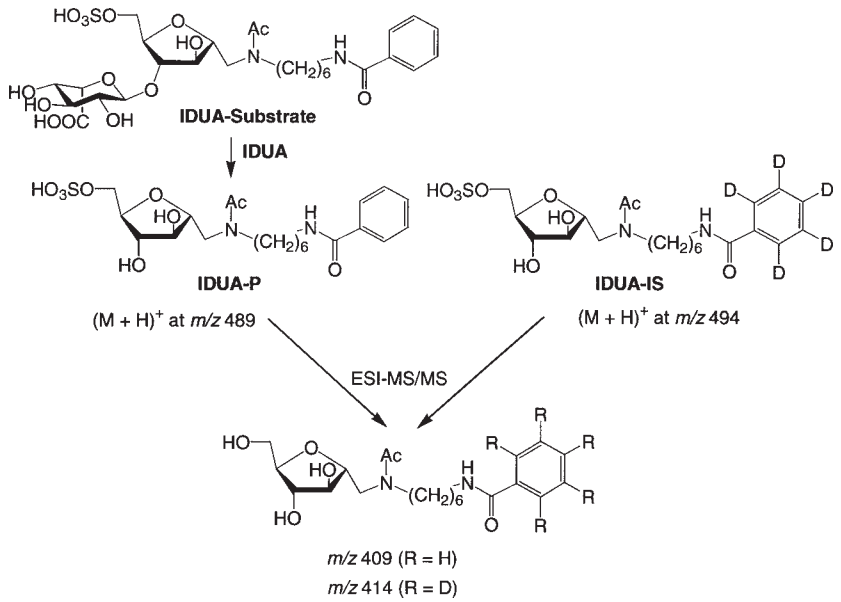


Fig. 4. Chemical, ionization, and ion dissociation processes in assaying α-L-iduronidase activity by electrospray ionization–tandem mass spectrometry.

enzymes can be stored at 4°C for up to 70 d with loss of approx 10% of enzyme activity.

2.3. Substrate Conjugates

1. All glycosylceramide substrates and ceramide IS for assaying ABG (C_{12} substrate, C_{14} IS), ASM (C_6 substrate, C_4 IS), and GALC (C_8 substrate, C_{10} IS) were purchased from Avanti Polar Lipids (Alabaster, AL; http://www.avantilipids.com/index) and used as received.
2. GAA and GLA substrates and IS were synthesized as described in supporting information for **ref.** ***14*** and are available from Dr. M. H. Gelb (gelb@chem.washington.edu) on request.
3. The substrate for IDUA was prepared by D. Wang on the scale of 1–2 g from commercially available heparin by a seven-step synthesis. The IDUA substrate and IS are available from M. H. Gelb on request.

2.4. Detergents, Inhibitors, Additives, and Buffers

1. Sodium taurocholate, oleic acid, and Triton X-100 were purchased from Sigma-Aldrich (Milwaukee, WI; http://www.sigma-aldrich.com) and used as received.
2. Acarbose was purchased from Toronto Research Chemicals, Inc. (North York, Ontario, Canada; http://www.trc-canada.com; cat. no. A123500) and used as received.

3. *N*-acetyl-galactosamine was purchased from Sigma-Aldrich.
4. DBS elution buffer for ABG, ASM, GALC, GAA, and GLA: 20 m*M* sodium phosphate (NaH_2PO_4, pH 7.0).
5. DBS elution buffer for IDUA: 50 m*M* sodium formate and 0.04 m*M* D-saccharic acid-1,4-lactone (Sigma, St. Louis, MO), pH 2.8.
6. ABG incubation buffer: to prepare 0.62 *M* citrate-phosphate buffer (pH 4.95), dissolve NaH_2PO_4 in water to 0.62 *M* and add solid trisodium citrate to 0.31 *M* followed by adjusting to pH 4.95 ± 0.05 with 6 *M* HCl.
7. ASM incubation buffer: 0.92 *M* sodium acetate buffer, pH 5.5.
8. GALC incubation buffer: to prepare 0.3 *M* citrate-phosphate buffer (pH 4.42), dissolve NaH_2PO_4 monohydrate in water to 0.3 *M* and add solid trisodium citrate to 0.15 *M* followed by adjusting to pH 4.42 with 6 *M* HCl.
9. GAA incubation buffer: to prepare 0.3 *M* citrate-phosphate buffer (pH 3.9), dissolve NaH_2PO_4 monohydrate in water to 0.3 *M* and add solid trisodium citrate to 0.15 *M* followed by adjusting to pH 3.9 with 6 *M* HCl.
10. GLA incubation buffer: 1.7 *M* sodium acetate, pH 4.6.
11. 85 m*M* Glycine carbonate buffer (pH 10.5): from 6.4 g/L glycine and 27.5 g/L sodium carbonate.

2.5. Stock Solutions

1. Stock solutions (3.0 m*M*) of substrates for ABG, ASM, and GALC were prepared in methanol and stored in Teflon septum-lined screw cap vials at –20°C. GLA and GAA substrate stock solutions were made at 10 m*M* in methanol.
2. Stock solutions of ceramide products and IS were prepared by weighing the lipid (typically 4.0 ± 0.1 mg), dissolving to 4 m*M* in methanol, and then diluting to 0.01 m*M* in methanol.
3. Stock solutions of GLA and GAA products and IS were prepared at 10 m*M* in methanol and diluted to 0.1 m*M* with water:methanol (2:1).
4. Stock solution of IDUA substrate was made at 1 m*M* in water.
5. All stock solutions were stored in screw-cap vials at –20°C.
6. All solvent manipulations and volume measurements were carried out with Pipetman.

2.6. Assay Solutions

The substrate and IS concentrations in assay solutions were optimized for carrying out the assays either with 2-mm DBS punches for each enzyme, or with an enzyme extract from a single 5-mm DBS punch. The following solutions and procedures are for using an enzyme extract from a 5-mm DBS punch corresponding to approx 10 µL of blood that typically contains 75,000–180,000 leukocytes for adults and newborns (first day, full term), respectively ***(18)***.

2.6.1. ABG Assay Solution

1. Add ABG substrate solution (3 m*M* in methanol, 200 µL), ABG IS stock solution (0.05 m*M* in methanol, 480 µL), and sodium taurocholate solution (12% w/v in water, 240 µL) to a 5-mL vial.

2. Place the vial in a desiccator attached to a vacuum pump and remove the solvent to give a white residue (~30 min at room temperature).
3. Add 1.80 mL of the ABG incubation buffer.
4. Vortex the vial until all residue is dissolved, and any emulsion was broken by centrifugation. The final assay solution (1.80 mL volume) contains 0.33 m*M* (0.6 μmol) ABG substrate, 13.3 μ*M* (24 nmol) ABG-IS, 1.6% sodium taurocholate, and 0.62 *M* citrate-phosphate, pH 4.95.

2.6.2. ASM Assay Solution

1. Mix ASM substrate stock solution (3 m*M* in methanol, 200 μL), ASM-IS solution (0.05 m*M* in methanol, 240 μL), and sodium taurocholate (12% w/v in water, 15 μL) in a 5-mL vial.
2. Place the vial in a desiccator attached to a vacuum pump and remove the solvent to give a white residue (~30 min at room temperature).
3. Add 1.80 mL of ASM incubation buffer.
4. Vortex the vial until all residue is dissolved and any emulsion was broken by centrifugation. The final assay solution (1.80 mL volume) contains 0.33 m*M* (0.6 μmol) ASM substrate, 6.67 μ*M* (12 nmol) ASM-IS, 0.1% sodium taurocholate, and 0.92 *M* sodium acetate, pH 5.5.

2.6.3. GALC Assay Solution

1. Mix GALC substrate stock solution (3 m*M* in methanol, 200 μL), ASM-IS solution (0.05 m*M* in methanol, 240 μL), 240 μL of an aqueous solution of sodium taurocholate (12% w/v), and 1.2% oleic acid in a 5-mL vial.
2. Place the vial in a desiccator attached to a vacuum pump and remove the solvent to give a white residue (~30 min at room temperature).
3. Add 1.80 mL of the GALC incubation buffer.
4. Vortex the vial until all residue is dissolved and any emulsion was broken by centrifugation. The final assay solution (1.80 mL volume) contains 0.33 m*M* (0.6 μmol) GALC substrate, 6.67 μ*M* (12 nmol) GALC-IS, 1.6% sodium taurocholate, 0.16% oleic acid, and 0.3 *M* citrate-phosphate, pH 4.42.

2.6.4. GAA Assay Solution

1. Mix GAA substrate stock solution (10 m*M* in methanol, 120 μL), GAA-IS solution (0.1 m*M* in methanol, 120 μL), acarbose solution (8 m*M* in water, 30 μL) (*see* **Note 1**), and Triton X-100 (1.8 μL) in a 5-mL vial.
2. Place the vial in a desiccator attached to a vacuum pump and remove the solvent to give a white residue (~30 min at room temperature).
3. Add 1.80 mL of the GAA incubation buffer.
4. Vortex the vial until the residue dissolved and any emulsion was broken by centrifugation. The final assay solution (1.80 mL volume) contains 0.33 m*M* (1.2 μmol) GAA substrate, 6.67 μ*M* (12 nmol) GAA-IS, 0.13 m*M* acarbose, and 0.3 *M* citrate-phosphate, pH 3.9.

2.6.5. GLA Assay Solution

1. Mix GLA substrate stock solution (10 m*M* in methanol, 300 µL), GLA-IS solution (0.1 m*M* in methanol, 120 µL), sodium taurocholate solution (12% w/v in water, 15 µL), and *N*-acetyl-galactosamine solution (1 *M* in water, 105 µL) in a 5-mL vial.
2. Place the vial in a desiccator attached to a vacuum pump and remove the solvent to give a white residue (~30 min at room temperature).
3. Add the GLA incubation buffer (300 µL) and vortex the vial until all residue is dissolved and any emulsion was broken by centrifugation. The final assay solution (300 µL volume) contains 10 m*M* (3 µmol) GLA substrate, 40 µ*M* (12 nmol) GLA-IS, 0.6% sodium taurocholate, 350 m*M* *N*-acetyl-galactosamine, and 1.85 *M* sodium acetate, pH 4.6.

3. Methods

3.1. Enzyme Assays

3.1.1. DBS Extraction

1. To a single well of a 96-well microtiter plate, add a 5-mm DBS disk (containing ~10 µL of blood) followed by 80 µL of DBS elution buffer.
2. Seal the plate with aluminum foil tape.
3. Incubate for 1 h at 37°C with orbital shaking (250 rpm) in a thermostated air shaker.

3.1.2. Incubation and Work Up for ABG, ASM, GALC, GAA, and GLA

1. To each of 5 wells of a 96-well polypropylene microtiter plate (E&K Scientific) add 10 µL of the DBS extract solution, followed by 15 µL of ASM, GALC, ABG, GAA assay cocktails, and 2.5 µL of GLA assay cocktail, one cocktail per well.
2. Seal the plate with aluminum foil tape.
3. Incubate at 37°C for 24 h with orbital shaking at 150 rpm.
4. Quench the reactions by addition of 100 µL of methanol:chloroform (2:1) to each (*see* **Note 2**).
5. Pump up and down five times the solution in each well with a Pipetman P200.
6. Combine all five solutions into one well of a Megatiter plate.
7. Add 400 µL each of chloroform (*see* **Note 2**) and water to the combined mixture.
8. Place the plate in a thermostated air shaker with orbital shaking (250 rpm).
9. Allow the mixture to separate into two layers (~1 h).
10. Using a plate loader, load 100 mg of silica gel to each of the wells of a 96-well filter plate (Innovative Microplate, Chicopee, MA; cat. no. F20005) (*see* **Note 3**).
11. Using a Pipetman, remove 400 µL of the bottom (chloroform) layer made in **step 9** and apply to a single well made in **step 10**.
12. Place the filter plate in a Millipore vacuum manifold containing a 96-well Megatiter polypropylene deep-well plate as the filtrate receiver.
13. Filter the chloroform solution through the silica gel by suction.

14. Charge the well with 700 µL of methanol:chloroform (9:1) (*see* **Note 2**).
15. Elute the P and IS from silica gel with methanol:chloroform (9:1) by suction.
16. Evaporate the solvent from the receiver well in a vacuum desiccator (~10 min at room temperature).
17. Take up the resulting residue in 200 µL of 5 m*M* ammonium formate in methanol:chloroform (3:1).
18. Analyze the solution by tandem mass spectrometry or store at –20°C.

3.1.3. Incubation and Work Up for IDUA

1. Obtain a 3-mm DBS punch by use of a standard leather punch and place it in a well of a 96-well plate (*see* **Note 4**).
2. Soak the DBS in 40 µL of the IDUA elution buffer and cap the plate with a Teflon-lined cover.
3. Shake the plate gently on an orbital shaker for 10 min at room temperature.
4. To the same well add 20 µL of 1 m*M* IDUA substrate stock solution.
5. Cap the plate and gently shake at 37°C for 24 h in a thermostated air shaker.
6. Quench the reaction by addition of 200 µL of glycine-carbonate buffer.
7. Add, 20 µL of 10 µ*M* IDUA-IS solution in water (*see* **Note 5**).
8. Mix the solution in the well by pipetting the liquid up and down a few times (a multichannel pipetor can be used when running several IDUA assays in parallel).
9. Transfer the liquid to a well of a 96-well filter plate containing C_{18} silica (for preparation, *see* **Note 6**).
10. Attach the filter plate to a vacuum manifold system that was attached to a water aspirator.
11. Add 400 µL of 50 mL/L methanol in water and wash the C_{18} silica solid phase by suction.
12. Attach the manifold to a Megatiter deep-well receiver plate and use suction to elute IDUA-P and IDUA-IS with one 400 µL portion of 500 mL/L methanol in water.
13. Evaporate the solvents from the receiver under reduced pressure by use of a vacuum desiccator (~1 h at room temperature).
14. To each well add 70 µL of 5 m*M* ammonium formate in methanol-chloroform, and infuse the sample solutions in the mass spectrometer.

3.2. Mass Spectrometric Analysis

1. ESI–MS/MS analyses were performed on a Sciex API-III tandem quadrupole tandem mass spectrometer equipped with an electrospray ion source and operated in positive ion mode (*see* **Note 7**).
2. Sample solutions were loaded in a gas-tight syringe and flow-injected by a syringe pump (Harvard Apparatus, Holliston, MA) through a 100 µm-internal diameter fused silica capillary at 2–3.3 µL/min flow rates (*see* **Note 8**).
3. The capillary was flushed with 4 × 50 µL of acetonitrile after each sample infusion to prevent cross-contamination.

4. The ESI source was operated with the following parameters: needle voltage, 4300 V; counter electrode, 650 V; nebulizer gas pressure, 35 psi; nebulizer gas flow rate, 0.6 L/min of N_2; curtain gas flow rate, 1.8 L/min of N_2; skimmer voltage, 60 V; q0 rod offset voltage, 30 V; Q1 rod offset voltage, 28 V; q2 rod offset voltage, 13 V (*see* **Note 9**); Q3 rod offset voltage, 5 V; collision gas (Ar) number density, 2.25×10^{14} molecules/cm^3 corresponding to 0.007 torr at 25°C.
5. The data acquisition parameters in the MRM mode were as follows: dwell time, 3–100 ms and pause time, 0.05 ms. Ion intensities from 200 to 250 scans were summed in each MRM channel over approx 2 min and used for quantification.
6. The molar amount of each P (mol_P) is calculated from the ion abundance ratio of ionized P (I_P[sample]) to ionized IS (II_S) minus the same ratio for a blank, multiplied by the amount of added IS ($molI_S$), and divided by the response factor ratio of P to IS (R_P).

$$mol_P = \frac{I_P\left(sample\right) - I_P\left(blank\right)}{I_{IS}} \times \frac{mol_{IS}}{R_P}$$

   RP is determined separately by infusing a mixture of P and IS of known concentrations and measuring the ratio of the relevant peak intensities in the MS/MS mass spectrum.
7. The enzyme activity is calculated from the blank-corrected amount of P from a DBS sample that was divided by the incubation time and the blood volume, and is reported as μmol/h/(L blood) (*see* **Note 10**).

4. Notes

1. Acarbose is used to effectively inhibit the enzyme maltase glucoamylase (K_i = 0.14 μ*M*) that is present in neutrophils ***(6)*** and would interfere with the determination of GAA activity in Pompe patients ***(14)***.
2. Chloroform can be replaced by the less toxic ethyl acetate. An optimized procedure using ethyl acetate is expected in early 2006. Contact M. H. Gelb for more information.
3. The products would be in the top layer if ethyl acetate was used.
4. An F96 MaxiSorp Nunc-Immuno Plate from Nalge Nunc International, Rochester, NY (cat. no. 442404).
5. All solutions were stored at –20°C and can be freeze-thawed multiple times.
6. The C_{18}-filled filter plate used C_{18}-silica bulk media (Sigma-Aldrich; cat. no. 377635) that was slurried in 10 mL of dichloromethane, and 1 mL of slurry was added to each well. Solvent was removed by suction on a filter manifold, and the solid phase was washed with 3 mL of methanol followed by 3 mL of 50 mL/L methanol in water.
7. Optimum infusion flow rates may differ and need to be optimized for the mass spectrometer being used.
8. Other tandem quadrupole mass spectrometers equipped with electrospray ionizers can be used to the same end, e.g., ABI-Sciex API 2000, 3000, 4000, or 5000 series (http://www.appliedbiosystems.com), Waters Quattro Premier and LCT

(http://www.waters.com/WatersDivision) Premier, or ThermoElectron TSQ Quantum (http://www.thermo.com). *See* http://www.directindustry.com for instrument listing.

9. The difference between the q2 and Q1 voltages determines the ion collision energy in the laboratory frame of reference. Collision energies between 15 and 25 V were used for ion dissociations to optimize the yields of product ions for MRM on the Sciex API-III.
10. Blank-corrected enzyme activities for ABG, ASM, GALC, GAA, and GLA in extracts from 5-mm DBS are given in **Table 1**.

Acknowledgments

This research was supported by a grant from the National Institutes of Diabetes and Digestive and Kidney Diseases (DK 67859). Thanks to Dr. Yijun Li and Mr. Ding Wang for performing the measurements, Joan Keutzer, Helmut Kallwass, Kate Zhang, and Knut Brockmann for helpful discussions, and to Dr. Martin Sadilek for technical assistance with mass spectrometry. Some of the DBS from patients with Krabbe and Pompe disease were kindly supplied by the Hunter's Hope Foundation (Orchard Park, NY) and by Dr. D. Millington (Duke University, Durham, NC), respectively.

References

1. Koeberl, D. D., Young, S. P., Gregersen, N. S., et al. (2003) Rare disorders of metabolism with elevated butyryl- and isobutyryl-carnitine detected by tandem mass spectrometry. *Pediatr. Res.* **54,** 219–223.
2. Stevens, R. D., Hillman, S. L., Worthy, S., Sanders, D., and Millington, D. S. (2000) Assay for free and total carnitine in human plasma using tandem mass spectrometry. *Clin. Chem.* **46,** 727–729.
3. Fauler, G., Leis, H. J., Huber, E., et al. (1997) Determination of homovanillic acid and vanillylmandelic acid in neuroblastoma screening by stable isotope dilution GC-MS. *J. Mass Spectrom.* **32,** 507–514.
4. Cole, R. B. (ed.) (1997) *Electrospray Ionization Mass Spectrometry. Fundamentals, Instrumentation and Applications*, Wiley-Interscience, New York.
5. Busch, K. L., Glish, G. L., and McLuckey, S. A. (eds.) (1988) *Mass Spectrometry/Mass Spectrometry: Techniques and Applications of Tandem Mass Spectrometry*, VCH Publishers, New York.
6. Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D. (eds.) (2001) Lysosomal Storage Disorders. In: *The Metabolic & Molecular Bases of Inherited Disease*, 8th Ed., McGraw-Hill, New York, Part 16, pp. 3371–3896.
7. Sandhoff, K., Kolter, T., and Harzer, K. (2001) Sphingolipid activator proteins. In: *The Metabolic & Molecular Bases of Inherited Disease*, 8th Ed., (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds.) McGraw-Hill, New York, pp. 3371–3375.

8. Neufeld, E. F. and Muenzer, J. (2001) The mucopolysaccharidoses. In: *The Metabolic & Molecular Bases of Inherited Disease*, 8th Ed., (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds.) McGraw-Hill, New York, pp. 3421–3426.
9. Gerber, S. A., Scott, C. R., Tureček, F., and Gelb, M. H. (1999) Analysis of rates of multiple enzymes in cell lysates by electrospray ionization mass spectrometry. *J. Am. Chem. Soc.* **121,** 1102–1104.
10. Tureček, F. (2002) Mass spectrometry in coupling with affinity capture-release and isotope-coded affinity tags for quantitative protein analysis. *J. Mass Spectrom.* **37,** 1–14.
11. Gerber, S.A., Scott, C. R., Tureček, F., and Gelb, M. H. (2001) Direct profiling of multiple enzyme activities in human cell lysates by affinity chromatography/electrospray ionization mass spectrometry. *Anal. Chem.* **73,** 1651–1657.
12. Zhou, X.-F., Tureček, F., Scott, C. R., and Gelb, M. H. (2001) Quantitation of cellular acid sphingomyelinase and galactocerebroside β-galactosidase activities by electrospray ionization mass spectrometry. *Clin. Chem.* **45,** 874–881.
13. Li, Y., Ogata, Y., Freeze, H. H., Scott, C. R., Tureček, F., and Gelb, M. H. (2003) Affinity capture and elution/electrospray ionization mass spectrometry assay of phosphomannomutase and phosphomannose isomerase for the multiplex analysis of congenital disorders of glycosylation yypes Ia and Ib. *Anal. Chem.* **75,** 42–48.
14. Li, Y., Scott, C. R., Chamoles, N. A., et al. (2004) Direct multiplex assay of lysosomal enzymes in dried blood spots for newborn screening. *Clin. Chem.* **50,** 1785–1796.
15. Li, Y., Brockmann, K., Tureček, F., Scott, C. R., and Gelb, M. H. (2004) Tandem mass spectrometry for the direct assay of enzymes in dried blood spots: application to newborn screening for Krabbe disease. *Clin. Chem.* **50,** 638–640.
16. For the nomenclature of mass spectrometric techniques, *see* the Mass Spectrometry Terms and Definitions Project at http://www.msterms.com/wiki/index.
17. Wang, D., Eadala, B., Sadilek, M., et al. (2005) Tandem mass spectrometric analysis of dried blood spots for screening of mucopolysaccharidosis I in newborns. *Clin. Chem.* **51,** 898–900.
18. Dacie, J. V. and Lewis, S. M. (1984) *Practical Haematology*, 6th Ed., Churchill Livingstone, Edinburgh, Scotland, pp. 9.

11

Absolute Quantification of Specific Proteins in Complex Mixtures Using Visible Isotope-Coded Affinity Tags

Yu Lu, Patricia Bottari, Ruedi Aebersold, František Tureček, and Michael H. Gelb

Summary

There is intense interest in determining the absolute abundance of specific proteins in complex mixtures, for example in the area of disease biomarker discovery. We have developed a set of protein-tagging reagents called visible isotope-coded affinity tags (VICAT reagents) that contain a protein-tagging reagent for reaction with cysteine SH groups, a visible probe for monitoring the chromatographic behavior of the target peptides, a photo-releasable biotin affinity tag for selective capture and release of tagged peptides, and a heavy isotope tag for differentiating analyte from internal standards. These reagents are used together with isoelectric focusing and reverse-phase microbore chromatography/electrospray ionization/tandem mass spectrometry to determine the absolute abundance of a set of target proteins in a complex mixture, such as a cell lysate. VICAT reagents should also be useful for detecting low abundance proteins in biological fluids such as serum, and for the detection of posttranslational protein modifications and different splice variants.

Key Words: Mass spectrometry; proteomics; protein abundance; disease biomarkers; protein-tagging reagents.

1. Introduction

Modern mass spectrometric methods provide an extremely powerful set of tools for detection of biomolecules, including peptides derived from protease-digested proteins. More information is obtained if the abundance of proteins is measured in addition to identifying the proteins. Introduction of stable isotopes into peptides using isotope-coded affinity tags or related reagents or

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

methods provides a way to quantify the abundance of peptides derived from proteolysis of complex protein mixtures present in biological samples, such as blood and tissue or cell homogenates ***(1–3)***. Such survey proteomic methods tend to detect the most abundant components in the complex mixture.

There are many cases where one wants to determine the absolute abundance of a specific set of target proteins, for example in the exploration of specific proteins as candidate disease markers whose identities have been obtained by, for example, hypothesis-driven experiments or by microarray mRNA studies. We recently developed visible isotope-coded affinity tags (VICAT reagents) as a set of protein- or peptide-tagging reagents that allow specific target proteins to be identified and quantified in an absolute sense in complex protein mixtures ***(4,5)***. Methods based on the use of VICAT reagents have the demonstrated capability to detect proteins that are expressed even at very low levels in complex biological samples ("needle-in-the-haystack" problem).

The VICAT method makes use of the set of three reagents shown in **Fig. 1** and the scheme outlined in **Fig. 2**. Proteins in a complex biological sample, such as serum or a cell lysate, are denatured and reduced to rupture any disulfide bonds. All protein cysteine residues are covalently tagged by treatment with the protein-tagging reagent $VICAT_{SH}$ (via its SH-selective iodoacetamido group). The mixture of tagged proteins is digested with trypsin to generate a mixture of tagged and untagged peptides. A known amount of internal standard (IS) is added to the sample. The IS is prepared by treating a cysteine-containing peptide (unique to the protein of interest and prepared by standard solid-phase peptide synthesis) with the IS reagent ^{14}C-$VICAT_{SH}$(+6). The IS is chemically identical to the tagged tryptic peptide derived from the protein of interest but is isotopically distinguished by the presence of carbon-13 and nitrogen-15 in the isotope tag linker (**Fig. 1**). Thus, IS and sample-derived tagged peptides are separately detected during electrospray ionization–tandem mass spectrometry (ESI–MS/MS) because of the 6-Da mass shift. The IS also contains carbon-14 of known specific radioactivity so that a precise amount of IS can be added to the biological sample. The trypsin digest is also spiked with the IEF marker. The IEF marker, prepared by treatment of the synthetic peptide with the IEF marker reagent, is chemically similar to the sampled-derived and IS-tagged peptides except that the IEF marker contains a shorter, two-carbon linker (**Fig. 1**). The IEF marker serves two purposes. First it is "visible" after separation of peptides by IEF owing to its carbon-14 label. Thus, the precise position of the IEF marker and the IS- and sample-derived tagged peptides, which comigrate with the IEF marker, can readily be determined by following the radiolabel. The second function of the IEF marker is to serve as a carrier to suppress nonspecific loss of the sample-derived tagged peptide during sample processing (for example, because of peptide absorption to vessel

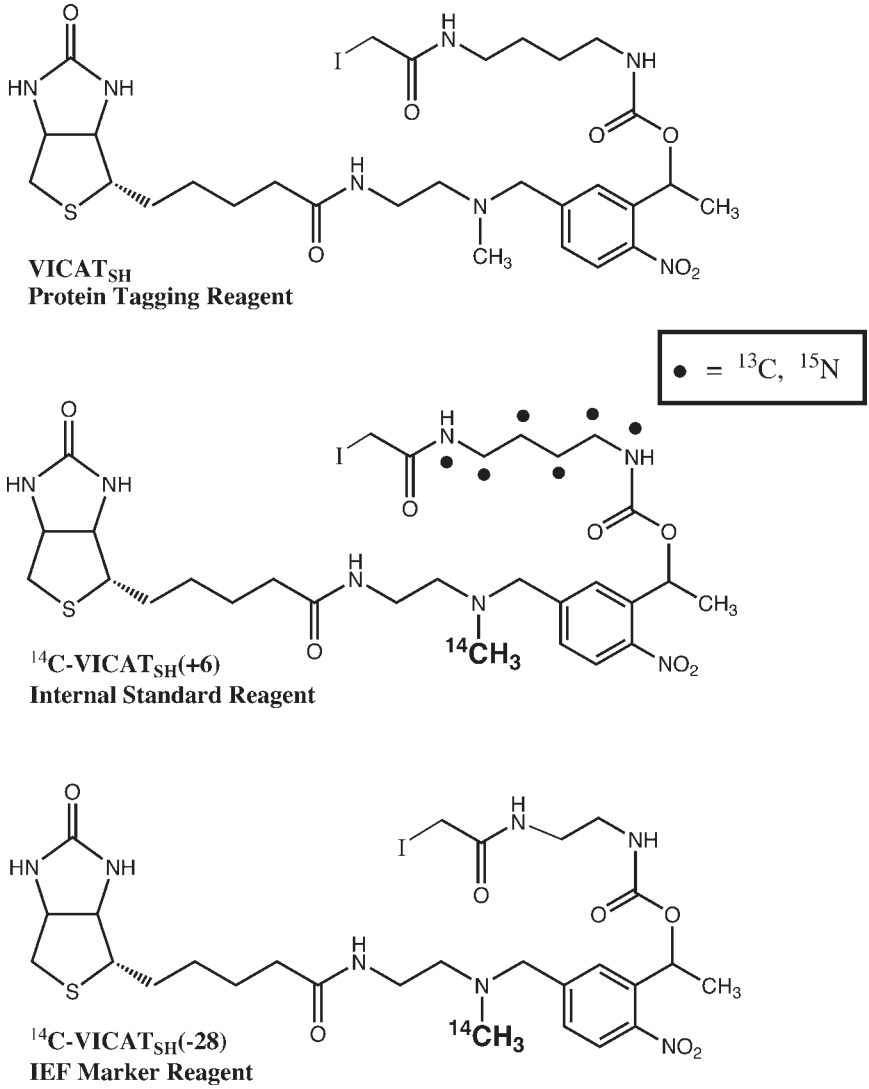


Fig. 1. Structure of the VICAT$_{SH}$ reagents. **(**Reproduced from **ref. *4*** with permission.)

walls, a factor that is often overlooked in peptide detection schemes). The amount of IEF marker added to the sample is typically two- to three-orders of magnitude larger than the amount of IS- and sample-derived tagged peptides and, thus, the IEF marker serves as an effective carrier. Because of the shorter linker in the IEF marker reagent, the IEF marker is 28 Da lighter than the sample-derived tagged peptide (34 Da lighter than the IS) and, thus, is readily

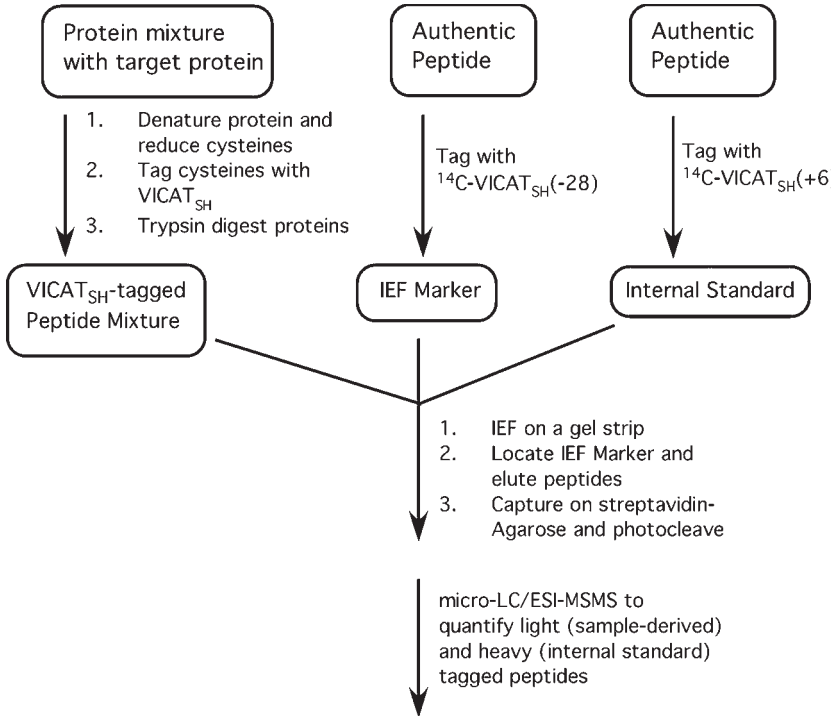


Fig. 2. Outline of the method of absolute protein quantification using the VICAT method. (Reproduced from **ref. *4*** with permission.)

excluded from detection during ESI–MS/MS. Although both the IS and IEF marker reagents contain carbon-14 (**Fig. 1**), the radioisotope serves different functions.

The trypsin digestion is subjected to IEF on a commercial gel strip containing an immobilized pH gradient (the same gel strip that is used as the first dimension of a standard two-dimensional protein gel). The IEF method has several nice features. First, it separates peptides according to their isoelectric points (pI), a parameter nearly independent of the peptide retention time on a reverse-phase high-performance liquid chromatography (HPLC) column. Thus, the combination of IEF and reverse-phase micro-LC provides a powerful two-dimensional peptide separation scheme, which is probably essential for finding low abundance peptides in highly complex peptide mixtures, such as those derived from cell lysates or serum. Second, peptides are focused into sharp bands (typically 1–2 mm in width) during IEF, which provides for high-resolution peptide fractionation. Third, the pI value of a peptide is highly predict-

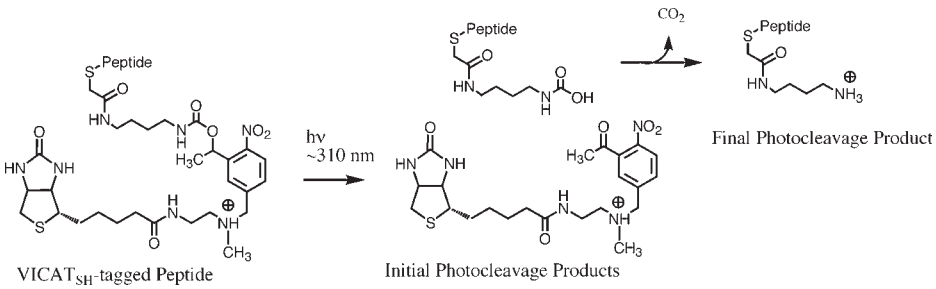


Fig. 3. Photocleavage of the peptides tagged with $VICAT_{SH}$ reagents to yield the peptide bearing a small isotope tag. (Reproduced from **ref. *4*** with permission.)

able based on its amino acid composition. This enables optimal choice of the pH range of the IEF gel strip used for the desired peptide separation. Finally, the capacity of commercial IEF gel strips (up to 3–5 mg of digested protein) is much higher than other techniques used, such as sodium dodecyl sulfate-polyacrylamide gel electrophoresis. This effectively increases the sensitivity of the VICAT method by allowing a larger amount of sample protein to be analyzed.

Following IEF, the appropriate region of the gel strip is cut from the strip and tagged peptides are eluted and captured using solid-phase streptavidin via the biotinyl moiety of the VICAT tag. This provides additional purification of tagged peptides and also allows buffer ampholytes from the IEF step to be removed. Captured peptides are then released into solution by photocleavage (via the photocleavable linker in the VICAT tag). Photocleavage has two nice features. It allows for additional purification of tagged peptides. It also leads to removal of the major portion of the tag including the radiolabel, so that the micro-LC retention time and ESI–MS/MS fragmentation are dominated by structural features of the peptide rather than of the tag. After photocleavage, the tagged peptides retain the isotope tag (**Fig. 3**) so that sample-derived, IS and IEF marker peptides are distinguished during ESI–MS/MS.

Finally, the peptide mixture after photocleavage is subjected to combined micro-LC (C18 reverse-phase)/ESI-MS/MS and selected reaction monitoring (SRM) is used for selective detection and quantification of the sample-derived and IS isotope-tagged peptides, whereas the more abundant tagged peptide derived from the IEF marker is excluded from detection. Because the absolute amount of IS is precisely know, the absolute amount of the sample-derived peptide is obtained by comparing the area of the appropriate ion peaks to those of the IS.

Additional useful features of the VICAT method are discussed in **Subheading 4**.

2. Materials

2.1. VICAT Reagents and Modified Peptides

1. Detailed synthetic methods for the preparation of the reagents ^{14}C-VICAT$_{SH}$(+6) and ^{14}C-VICAT$_{SH}$(–28) have been reported ***(5)***. These reagents may also be obtained from M. H. Gelb. Stock solutions of ^{14}C-VICAT$_{SH}$(–16) and ^{14}C-VICAT$_{SH}$(–28) are prepared in HPLC-grade CH_3CN and stored in glass vials with Teflon septum-lined screw caps at –20°C in the dark.
2. HPLC solvent A is high-purity water (Milli-Q; Millipore, Billerica, MA) or other HPLC-quality water) containing 0.08% (v/v) HPLC-grade trifluoroacetic acid (TFA).
3. HPLC solvent B is HPLC-grade acetonitrile (CH_3CN) containing 0.08% (v/v) HPLC-grade TFA. HPLC solvents A and B should be filtered through 0.2 to 0.5-µ nylon-66 membranes prior to use.
4. The C18-reverse phase HPLC column is a Vydac T238EV54, 0.46 × 20-cm column (The Separations Group, Hesperia, CA), but any C18 reverse-phase column suitable for peptide purification should be adequate.
5. 250 m*M* dithiothreitol (Sigma, St. Louis, MO) in water should be stored at –20°C.
6. 1.0 *M* Tris-HCl, pH 8.3 buffer (Sigma) should be stored at 4°C.
7. 250 m*M* iodoacetamide (Sigma) in 50% CH_3CN should be stored at –20°C.
8. Synthetic peptides are made by commercial solid-phase peptide synthesis and purified by HPLC on a C18-reverse phase column using standard methods with water/CH_3CN gradients (containing typically 0.08% TFA). HPLC-purified peptides are concentrated to dryness by lyophilization or by use of a Speed-Vac (Thermo Electron Corporation, West Palm Beach, FL).

2.2. Sample Preparation, VICAT Tagging, and Trypsin Digestion

1. 25.4 m*M* VICAT$_{SH}$, reagent (*see* **Subheading 2.1.**, **step 1**) in CH_3CN in a glass vial capped with a Teflon-septum lined screw cap (stored at –20°C in a light-tight box).
2. Cell lysis buffer: 50 m*M* Tris-HCl, pH 8.3, 6 *M* urea, 2% (w/v) CHAPS, and 5 m*M* EDTA (all reagents from Sigma) (stored at 4°C).
3. 250 m*M* TCEP (Pierce Chemicals, Rockford, IL) in high-purity water (stored at –20°C).
4. Bradford dye binding protein assay reagent (Bio-Rad, Hercules, CA).
5. Bovine serum albumin (Sigma).
6. 100X IPG buffer (Amersham Biotech, Piscataway, NJ; cat. no. 17-6000-87).
7. 1% (w/v) Bromophenol blue (Sigma) in ethanol.
8. 6 *M* Urea and 2% (w/v) CHAPS in high-purity water (reagents from Sigma).
9. Trypsin (modified, sequencing grade, Promega) in trypsin storage buffer, stored at –80°C (Promega, Madison, WI).

2.3. Preparative IEF and IEF Strip Elution of Peptides

1. 13-cm Commercial IPG strips (linear pH range 3.0 to 10.0; Amersham Biotech).
2. Multiphor II IEF device and programmable power supply (Amersham Biotech.).

3. Mineral oil for IEF (Amersham Biotech.).
4. Absorbent paper (Whatmann 3MM or the equivalent).
5. Plastic wrap (Saran Wrap or the equivalent).
6. Reagent grade 1% (w/v) NH_4OH in 20% (v/v) CH_3CN in high-purity water.
7. 80% (v/v) CH_3CN and 0.1% (v/v) TFA in high-purity water.
8. 20% (v/v) CH_3CN in high-purity water.
9. Speed-Vac concentrator (Savant Instruments or the equivalent).

2.4. Affinity Capture and Photocleavage

1. Streptavidin-agarose (1:1 slurry as supplied by the manufacturer; Sigma, cat. no. S-1638).
2. Phosphate-buffered saline (PBS): 144 mg/L KH_2PO_4, 9000 mg/L NaCl, 795 mg/L Na_2HPO_4.
3. 20% (v/v) CH_3CN and 5 m*M* Tris-HCl, pH 8.3.
4. β-mercaptoethanol.
5. Vibrating mixer platform (Eppendorf Thermomixer or the equivalent).
6. Ultraviolet (UV) lamp (BLAK-RAY longwave model B-100AP, 100 watts).
7. Speed-Vac concentrator (Thermo Electron Corporation or the equivalent).

2.5. Micro-LC/ESI–MS/MS

1. 20% (v/v) CH_3CN in high-purity water.
2. Micro-LC column (10 cm × 75 μm ID, 5 μm C-18 magic beads, 100 Å pores, Michrom BioResources, Auburn, CA, column packed as described by **ref. *6***).
3. Micro-LC solvent A (high purity water with 0.1% [v/v] reagent-grade formic acid.
4. Micro-LC solven B (100% CH_3CN).
5. Ion trap mass spectrometer (Thermo Finnigan, San Jose, CA, model LCQ Deca XP).

3. Methods

The VICAT method for absolute quantification of a specific target protein in a complex biological sample is described for the detection of a specific cysteine-containing, tryptic peptide derived from human group V secreted phospholipase A_2 (hGV). The biological test sample is a lysate made from insect cells (*Spodoptera frugiperda*, *Sf*9 cells) because these cells lack the human protein and serve as a negative control to which various amounts of hGV can be added. This serves as an important validation sample because the amount of hGV detected by the VICAT method can be compared with the amount of recombinant hGV added to the *Sf*9 cell lysate.

3.1. Preparation of IS

1. The IS is prepared as follows. The hGV-derived peptide SYNPQYQYFPNILCS (100 μL of 1 m*M* in water) in a 1.5-mL polypropylene microfuge tube is mixed with 20 μL of CH_3CN, 2 μL of 1 *M* Tris-HCl, pH 8.3, and ^{14}C-VICAT$_{SH}$(+6) (3

μL of 14 m*M* in CH_3CN) and the mixture is stirred for 2 h (or overnight) at room temperature in the dark (by wrapping the tube with aluminum foil) (*see* **Note 1**).
2. Then, 13 μL of 250 m*M* dithiothreitol in water is added, the tube is incubated for 30 min in the dark, then 40 μL of 250 m*M* iodoacetamide in 50% CH_3CN is added, and the tube was incubated in the dark for 15 min (*see* **Note 2**).
3. The IS is purified by HPLC (Vydac T238EV54, 0.46 × 20 cm column) with HPLC solvent A and B: 0–10 min, 0–25% B; 10–50 min, 25–45% B; 50–60 min, 45–80% B at 0.7 mL/min with monitoring at 240 nm (elution at 35% B).
4. The HPLC fraction containing the IS is collected into a 1.5-mL polypropylene microfuge tube.
5. An aliquot of IS is submitted to scintillation counting, and the dpm of carbon-14 together with the specific radioactivity of the ^{14}C-VICAT$_{SH}$(+6) reagent is used to calculate the molar concentration of IS in the stock solution.
6. The structure of the IS is confirmed by submitting an aliquot of the stock solution to ESI–MS. The observed peak for the $(M+2H^+)^{2+}$ ion is 1239.3.
7. The solution of IS is stored at –20°C in the dark.

3.2. Preparation of the IEF Marker

1. The IEF marker is made in the same way as the IS using 100 μL of 1 m*M* peptide stock, 30 μL CH_3CN, 2 μL of 1 *M* Tris-HCl, pH 8.3 and 1 μCi of ^{14}C-VICAT$_{SH}$(–28) (30 Ci/mol).
2. The IEF marker reaction mixture is treated with dithiothreitol and iodoacetamide as described in **step 2** for the IS.
3. The IEF marker is purified by HPLC and a stock solution is prepared as described in **step 3** for the IS.
4. The IEF marker stock solution is submitted to scintillation counting and ESI-MS/MS as described in **steps 4–6** for the IS.
5. The IEF marker stock solution is stored at –20°C in the dark.

3.3. Preparation of the Cell Lysate, VICAT$_{SH}$ Tagging, and Trypsin Digestion

1. This procedure starts with *Sf*9 cells that had been previously washed with PBS and stored at –80°C. The procedure has also been carried out with washed human lung macrophages ***(4)***. In a 1.5-mL polypropylene microfuge tube containing 2 million *Sf*9 cells 200 μL of ice-cold cell lysis buffer is added, and cells are lyzed by mixing with a vortex mixer (several 10 to 20-s bursts with intermittent cooling on ice).
2. The sample is centrifuged at 4°C at approx 12,000*g*, the supernatant is transferred to a new microfuge tube, and the supernatant is centrifuged again and transferred to a new tube.
3. The protein concentration in the cell lysate is determined using a small aliquot and the Bradford dye binding assay (Bio-Rad) using bovine serum albumin as a standard (protein assay carried out as described by the manufacturer).
4. A 50-μL aliquot (100 μg protein) is transferred to a new microfuge tube, and the desired amount of recombinant hGV ***(7)*** is added (*see* **Note 3**).

5. The IS is added to the cell lysate (*see* **Notes 4** and **5**).
6. TCEP (0.5 µL of 250 m*M* stock solution) is added to give 2.5 m*M*, and the sample is incubated at 37°C for 30 min (*see* **Note 6**).
7. $VICAT_{SH}$ (8 µL of the 25.4 m*M* stock solution) is added to give 3.5 m*M*, and the sample is incubated in the dark at room temperature for 3 h (*see* **Note 7**).
8. The sample is diluted threefold with high-purity water, 6.5 µg of trypsin is added, and the sample is incubated for 15 h at 37°C in the dark (foil-wrapped tube) (*see* **Note 8**).
9. The sample is concentrated back to its original volume (50 µL) in a Speed-Vac.
10. IEF marker (~8900 dpm, 130 pmol) is added (*see* **Note 5**) follwed by 2.5 µL of 100X IPG buffer and 1 µL of 1% bromophenol blue stock solution.
11. Finally, 6 *M* urea and 2% CHAPS solution is added to bring the sample volume to 250 µL.
12. The sample can be stored at –20°C or submitted immediately to IEF.

3.4. Preparative IEF of Tagged Peptides

1. The commercial 13-cm IPG strip (linear pH range 3.0 to 10.0) is rehydrated with sample in the rehydration tray provided with the Multiphor II system.
2. The rehydrated IPG strip is submitted to IEF using the Multiphor II device as described by the manufacturer using the following power supply program: 0–100 V in 1 min, 100 V for 1 h, 100–500 V in 1 min, 500 V for 1 h, 500–3500 V in 8 h, 3500 V for 10 h. The IEF unit is covered with aluminum foil to block light.
3. The IPG strip is removed with tweezers, and excess mineral oil is removed by wiping the plastic backing with a tissue and tapping the edge several times on a tissue (avoid contacting the gel-side of the strip with the tissue).
4. The strip is laid gel-side up on a piece of Whatmann 3MM paper, the strip is covered with a sheet of plastic (Saran Wrap), and the plastic sheet is taped to the paper. The plastic wrap above the desired region of the strip is marked with a pen (using a ruler as a guide) (*see* **Note 9**).
5. The desired region of the wrapped IPG strip is cut into eight approx 1-mm wide pieces (*see* **Note 10**) using scissors (cutting through the paper, the IPG strip and the plastic wrap), and individual IPG pieces are separated from the paper and plastic wrap with tweezers and transferred to individual 1.5-mL polypropylene microfuge tubes.

3.5. Elution of Peptides From the IPG Strips

1. Gel slices are subjected to continuous shaking on a vibrating platform in 200 µL of 1% NH_4OH in 20% CH_3CN for 1 h, and the liquid phases are transferred to new 1.5-mL polypropylene microfuge tubes.
2. The extracts are concentrated to dryness (Speed-Vac).
   Gel slices are shaken for 1 h in 150 µL 80% CH_3CN/0.1% TFA water and extracts combined.
4. The samples are concentrated to dryness (speed-vac).
5. 100 µL of 20% CH_3CN is added to each tube, and 5 to 10-µL aliquots are submitted to scintillation counting (*see* **Note 11**).

3.6. Affinity Capture and Release

1. Streptavidin-agarose (50 µL of a 1:1 slurry as supplied by the manufacturer) is placed in a 1.5-mL polypropylene microfuge tube, and the gel is pelleted by brief centrifugation.
2. The gel is washed two times with 1-mL portions of PBS.
3. After removal of the final wash, 1 mL of PBS is added, followed by 200 µL of peptide sample (from two IEF combined eluant fractions, in some cases only a single IEF eluant fraction is obtained, *see* **Note 11**).
4. The sample is gently rocked for 1 h at room temperature.
5. The gel is washed two times with 1-mL portions of water.
6. 200 µL of 20% CH_3CN/5 m*M* Tris-HCl, pH 8.3 and 2 µL of β-mercaptoethanol (*see* **Note 12**) are added to the gel pellet.
7. The tube is shaken for 16 min with a vibrating mixer platform with a UV lamp bulb (BLAK-RAY longwave model B-100AP, 100 W) held 6 cm from the top of the open tube (*see* **Note 13**).
8. The gel is pelleted, and the supernatant is transferred to a new 1.5-mL polypropylene microfuge tube
9. The sample is concentrated to dryness (Speed-Vac).

3.7. Microbore Reverse-Phase Liquid Phase Chromatography Combined With ESI–MS/MS

1. The dried sample is dissolved in 10 µL of 20% CH_3CN.
2. An 8-µL aliquot is subjected to C18 micro-LC/ESI–MS/MS using a mico-LC system capable of delivery of column eluant directly into the LCQ ESI–MS/MS source (*see* **ref. *8*** for a typical setup). The sample is loaded onto a pre-column (2-cm capillary, 5-µm C-18 magic beads, 100 Å pores; Michrom BioResources, Auburn, CA) at a flow rate of 120 µL/min.
3. After sample loading, the pre-column is washed with micro-LC solvent A at 120 µL/min.
4. Then, the column gradient is run: 0–5 min, 5–20% micro-LC solvent B in micro-LC solvent A; 5–25 min, 20–40% B; 25–35 min, 40–80% B; 35–60 min, 80% B at a flow rate of 150 nL/min.
5. Eluting peptides are analyzed in the ion trap mass spectrometer using SRM (as indicated by the manufacturer) of 10 fragment ions (*see* **Note 14**) from each of the two doubly charged, precursor ions (light- and heavy-tagged peptides) (*see* **Note 14**). SRM parameters are given in **Table 1**. The *b*- and *y*-ion fragments monitored for the hGV-derived peptide (10 for the light-tagged peptide and the analogous 10 for the heavy-tagged peptide) are listed in **Table 2**.
6. For each of the 10 fragment ions monitored, the ratio of light to heavy was obtained using the ion peak integration parameters and integration mass ranges listed in **Tables 3** and **4**. Results are summarized in **Table 5**. Among the 10 ratios, the lowest and highest were discarded, and the remaining eight were averaged. This average ratio was multipled by the moles of IS added to the original sample to give the moles of hGV protein in the sample (*see* **Note 15**).

Table 1
Instrument Settings for SRM[a]

Positive electrospray, spray voltage:	2.2 kV
Ion injection AGC:	1×10^7 ions
Maximum ion injection time:	200 ms
Scan rate:	6 microscans/scan
Two SRM events (SRM1 and SRM2) were carried out in alternation as follows:	
SRM1:	
Precursor ion m/z:	983.5 Da
Isolation window width:	3.5 Da
Normalized collision energy:	30%
Activation Q:	0.25
Activation time:	30.0 ms
SRM2:	
Precursor ion m/z:	986.5 Da
Other parameters:	as for SRM1

[a]Instrument parameters were automatically tuned and calibrated according to the procedure provided by Thermo Electron. AGC, automatic gain control.

3.8. Concluding Remarks

A number of advantages of the VICAT method for the absolute quantification of specific target proteins in complex protein mixtures have been summarized ***(4)***. A few additional points are given here. The iodoacetamido group of the VICAT reagents leads to selective capture of cysteine-containng peptides, which leads to an approx 10-fold reduction in the complexity of the initial peptide mixture. In some cases, it may not be possible to find a signature tryptic peptide in the protein of interest that contains a cysteine. We have shown that amino groups in tryptic peptides (N-terminal and lysine side chain) can be quantitatively converted with commercially available reagents to $-NHCOCH_2SH$ groups so that VICAT reagents can be used to tag tryptic peptides that do not contain cysteine ***(5)***.

The detection of peptides derived from low abundance proteins in serum is particularly useful for the quantification of potential biomarkers. For example, immunochemical detection of prostate-specific antigen in serum is routinely carried out at the low nanogram per milliliter level. However, serum analysis by MS suffers from the presence of abundant proteins (~40 mg protein per milliliter). To detect 1 ng of a molecular weight 50,000 protein in 1 mL of serum by the VICAT method where 3 mg of trypsin-digested protein can be loaded on the IEF strip, the ESI–MS/MS must be capable of detecting 1.5 fmol of peptide, or 150

Table 2
Monitored Fragments for Each SRM[a]

SRM1					SRM2
Light fragment (L)	m/z	m/z range	heavy fragment (H)	m/z	m/z range
y_2^+(L)	337.1	8	y_2^+(H)	343.1	8
y_3^+(L)	450.1	8	y_3^+(H)	456.1	8
y_6^+(L)	774.3	8	y_6^+(H)	780.3	8
y_{12}^{2+}(L)	800.8	6	y_{12}^{2+}(H)	803.8	6
y_{13}^{2+}(L)	857.8	6	y_{13}^{2+}(H)	860.8	6
y_7^+(L)	921.4	8	y_7^+(H)	927.4	8
y_8^+(L)	1084.5	8	y_8^+(H)	1090.5	8
y_9^+(L)	1212.5	8	y_9^+(H)	1218.5	8
y_{10}^+(L)	1375.6	8	y_{10}^+(H)	1381.6	8
y_{12}^+(L)	1600.7	8	y_{12}^+(H)	1606.7	8

[a]SRM1 is for the light-tagged peptide SerTyrAsnProGlnTyrGlnTyrPheProAsnIleLeuCys (S-$CH_2CONH(CH_2)_4NH_2$)Ser and SRM2 is for the heavy-tagged peptide SerTyrAsnProGlnTyrGlnTyrPheProAsnIleLeuCys (S-$CH_2CO^{15}NH(^{13}CH_2)_4{}^{15}NH_2$)Ser.

Table 3
Data Processing

Area integration and comparison was carried out using LCQuan in the Finnigan Xcalibur v1.3 software bundle using the following parameters:

Peak identification algorithm: ICIS, highest peak in specified time window.

ICIS peak integration parameters:

Smoothing points:	1
Baseline window:	40
Area noise factor:	5
Peak noise factor:	10

amole if the overall yield of trypsin-generated and isotope-tagged peptide is 10%. This sensitivity is achievable with state-of-the-art tandem quadrupole instruments, such as the triple quadrupole mass spectrometer from ABI-Sciex, but not with the ion trap instrument described in this chapter. We have recently found that tens of milligrams of serum protein can be first digested with trypsin and then subjected to HPLC on a C18 reverse-phase semi-preparative column (1 ×

Table 4
Integration Mass Range for All Fragments

SRM1		SRM2	
Fragment	m/z range	Fragment	m/z range
y_2^+(L)	336.6–339.1	y_2^+(H)	342.6–345.1
y_3^+(L)	449.6–452.1	y_3^+(H)	455.6–458.1
y_6^+(L)	773.8–776.3	y6$^+$(H)	779.8–782.3
y_{12}^{2+}(L)	800.5–801.8	y_{12}^{2+}(H)	803.5–804.8
y_{13}^{2+}(L)	857.5–867.8	y_{13}^{2+}(H)	860.5–861.8
y_7^+(L)	920.9–923.4	y_7^+(H)	926.9–929.4
y_8^+(L)	1084.0–1086.5	y_8^+(H)	1090.0–1092.5
y_9^+(L)	1212.0–1214.5	y_9^+(H)	1218.0–1220.5
y_{10}^+(L)	1375.1–1377.6	y_{10}^+(H)	1381.1–1383.6
y_{12}^+(L)	1600.2–1602.7	y_{12}^+(H)	1606.2–1608.7

25 cm) (Zu, Y. and Gelb, M.H., to be published). The portion of the eluant containing the target peptide of interest can be liberally collected (we have found that the retention of a variety of tryptic peptides is not altered when the peptide is coinjected with a serum protein digest) and submitted to tagging with $VICAT_{SH}$. The rest of the procedure follows that which is outlined in this chapter. In this way, we have been able to detect low nanogram amounts of proteins in 1 mL of serum. This would not be possible with other methods that rely on sodium dodecyl sulfate-polyacrylamide gel electrophoresis as the first dimension of sample separation because typically 100 μg of protein can be loaded onto a lane of a protein gel.

It should also be mentioned that the VICAT method of target protein quantification is relatively low throughput compared with immunochemical methods that are typically used in clinical laboratories. We envision that the VICAT method will be particularly useful in hypothesis-derived biomarker discovery. For example, the exploration of say 100 candidate biomarkers by the VICAT method requires only that 100 different peptides be prepared by solid-phase synthesis on a sub- to low-milligram scale. The development of 100 ELISA assays from scratch is prohibitively expensive and time consuming. Once an acceptable biomarker is discovered, a robust and rapid immunochemical method can be developed for clinical use.

4. Notes

1. To avoid undesirable cleavage of the photocleavable linker in the VICAT reagents, all steps with VICAT reagents were carried out under fluorescent tube room light and away from direct sunlight.

Table 5
hGV Analysis in Sf9 Cell Lysate

Fragment	Peak area	Peak area ratio (light/heavy)	Signal-to-noise
y_2^+(L)	4693014	0.644	506
y_2^+(H)	7285601		640
y_3^+(L)	9786995	0.632 (omitted)	624
y_3^+(H)	15489135		785
y_6^+(L)	91872727	0.7	3172
y_6^+(H)	131230308		2414
y_{12}^{2+}(L)	111205120	0.723	10518
y_{12}^{2+}(H)	153797709		18179
y_{13}^{2+}(L)	75243665	1.198 (omitted)	1619
y_{13}^{2+}(H)	62816693		64552
y_7^+(L)	68430536	0.873	1345
y_7^+(H)	78425108		1731
y_8^+(L)	39667363	0.733	1107
y_8^+(H)	54125354		8612
y_9^+(L)	36940331	0.809	1132
y_{10}^+(L)	26939516	0.688	971
y_{10}^+(H)	39177251		1359
y_{12}^+(L)	23355174	0.739	1016
y_{12}^+(H)	31586726		1220
		Average[a] = 0.738625	

[a]The average is the average peak area ratio with the highest and lowest ratio omitted.

2. This reductive/alkylation step serves to cap the SH group of any remaining unreacted peptide because contamination of the biological sample with unreacted peptide would lead to false-positive detection of the protein of interest.
3. The addition of recombinant hGV in this case is to validate the method. Normally one is interested in determining the absolute amount of endogenous target protein in the biological sample, and authentic target protein would not be added. The concentration of hGV in the stock solution is determined by submitting approx 10 µg for amino acid analysis (Protein Chemistry Laboratory, Department of Biochemistry and Biophysics, Texas A&M University). The nmole amount of hGV is taken as the average of the individual nmole values obtained by dividing the nmole of each amino acid determined by the number of residues determined from the amino acid sequence.
4. The amount of IS added should be in the same ballpark as the amount of sample-derived peptide present in the sample. The latter may not be known *a priori*, so

the amount of IS may have to be adjusted in subsequent runs. If the amount of IS and sample-derived tagged peptide differ by more than approx 100, there could be error introduced into the quantification of the sample-derived tagged peptide because of nonlinearity in the dynamic range of the ESI–MS/MS response.

5. More than one IS and IEF marker can be added to the sample. Multiple regions of the IEF gel strip, corresponding to the multiple IEF markers, can be cut in the case that the quantification of multiple target multiple peptides is desired.
6. TCEP is added to ensure that all cysteines are in the reduced state. Excess TCEP beyond what is recommended above should be avoided because TCEP reacts slowly with the iodoacetamido group of $VICAT_{SH}$ reagents.
7. If it is assumed that the average molecular weight of the proteins in the sample is 50,000, that each protein has 10 cysteines, and the amount of SH groups per 100 μg of sample protein is 20 nmol. Thus, the 200 nmol of $VICAT_{SH}$ used is sufficient to tag all SH groups. If more than 100 μg of sample protein is used, the amount of $VICAT_{SH}$ reagent should be increased correspondingly.
8. The amount of trypsin can be appropriately increased if more than 100 μg of sample protein is analyzed.
9. To estimate the migration position of the IEF marker, the pI of the marker is calculated using the amino acid sequence of the peptide component of the marker with an extra histidine residue (pK_a of the histidine side chain is close to that of the tertiary ammonium of $VICAT_{SH}$) assuming the pH varies linearly along the IEF strip.
10. Once the IEF gel-strip position of the IEF marker is known from the first run, the strip can be cut into six pieces instead of eight.
11. Typically, greater than 80% of the dpm along the entire IPG strip is found in one to two tubes (i.e., ~1–2 mm of the IPG strip).
12. β-mercaptoethanol is added to prevent photo-oxidation of the peptides during photocleavage.
13. To ensure reproducible photocleavage, it is recommended that the UV lamp intensity be periodically measured. This can be done with an inexpensive UV detector (i.e., model UV340, Mannix, Lynbrook, NY).
14. It is best to monitor by MRM the major *b*- and *y*-ions derived from the sample-derived and IS peptides. After photocleavage, these peptides will contain the isotope tag ($-CH_2CONH(CH_2)_4NH_2$ or the heavy isotope-substitute tag). The major *b*- and *y*-ions will not be known *a priori*. A portion of the IS can be captured on streptavidin-agarose and subjected to photocleavage using the methods described in **Subheading 3.6.** The tagged peptide can then be analyzed by ESI–MS in full scan mode so that the major *b*- and *y*-ions can be selected for subsequent analysis of protein sample-derived tagged peptide and IS.
15. **Figure 4** shows the detection of the IS and the hGV-derived tagged peptide by micro-LC/ESI–MS/MS. The bottom panel in **Fig. 4** shows that the absolute amount of hGV detected by the VICAT method correlates well with the expected amount based on the amount of hGV added to the *Sf*9 cell lysate. **Figure 5** shows the absolute quantification of hGV in human lung macrophages.

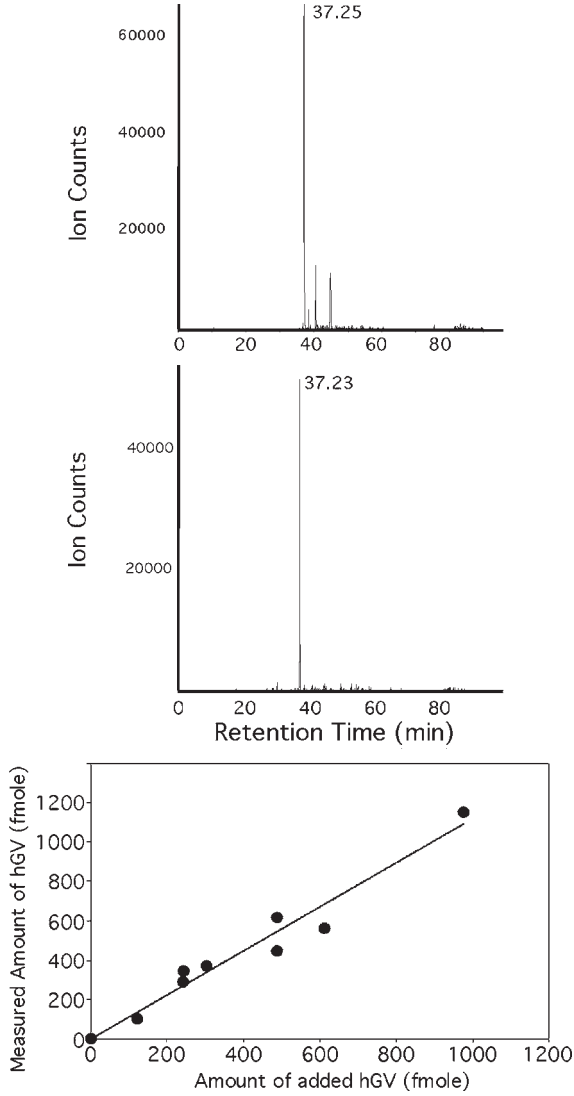


Fig. 4. Quantification of hGV protein in *Sf*9 cell lysates. The top panel shows the ion chromatography for the Y_{10}^{+} ion derived from the light-tagged peptide, and the middle panel shows the same for the heavy-tagged peptide. These peptides are seen to coelute from the micoLC column. The lower panel shows that the amount of hGV detected by the VICAT method is the amount expected based on the amount of recombinant hGV added to the *Sf*9 cell lysate (the slope of the line is unity). (Reproduced from **ref.** ***4*** with permission.)

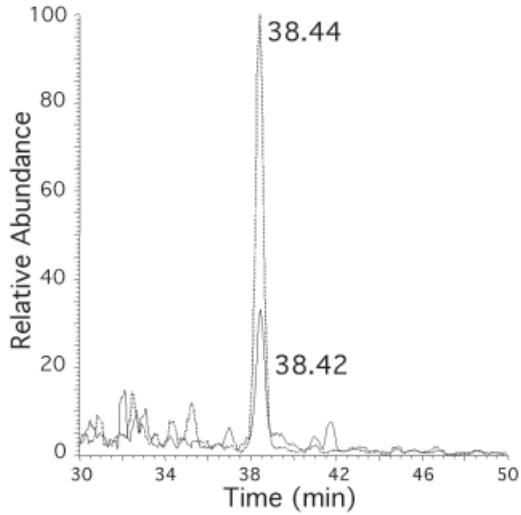


Fig. 5. Detection of the hGV-derived tagged peptide and internal standard in human lung macrophages. (Reproduced from **ref.** *4* with permission.)

Acknowledgments

This work was funded by grants from the National Institutes of Health (DK67859 to F.T. and M.H.G., and HL36235 to M.H.G.) and by federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, under contract no. N01-HV-28179 (R.A).

References

1. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17,** 994–999.
2. Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422,** 198–207.
3. Tao, W. A. and Aebersold, R. (2003) Advances in quantitative proteomics via stable isotope tagging and mass spectrometry. *Curr. Opin. Biotechnol.* **14,** 110–118.
4. Lu, Y., Bottari, P., Turecek, F., Aebersold, R., and Gelb, M. H. (2004) Absolute quantification of specific proteins in complex mixtures using visible isotope-coded affinity tags. *Anal. Chem.* **76,** 4104–4111.
5. Bottari, P., Turecek, F., Aebersold, R., and Gelb, M. H. (2003) Design and synthesis of visible isotope-coded affinity tags for the absolute quantification of specific proteins in complex mixtures. *Bioconj. Chem.* **15,** 380–388.

6. Lee, H., Yi, E. C., Wen, B., et al. (2004) Optimization of reversed-phase microcapillary liquid chromatography for quantitative proteomics. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **803 ,** 101–110.
7. Singer, A. G., Ghomashchi, F., Le Calvez, C., et al. (2002) Interfacial kinetic and binding properties of the complete set of human and mouse groups I, II, V, X, and XII secreted phospholipases A2. *J. Biol. Chem.* **277,** 48,535–48,549.
8. Yi, E. C., Lee, H., Aebersold, R., and Goodlett, D. R. (2003) A microcapillary trap cartridge-microcapillary high-performance liquid chromatography electrospray ionization emitter device capable of peptide tandem mass spectrometry at the attomole level on an ion trap mass spectrometer with automated routine operation. *Rapid Commun. Mass Spectrom.* **17,** 2093–2098.

12

Computational Analysis of Quantitative Proteomics Data Using Stable Isotope Labeling

Michael J. MacCoss and Christine C. Wu

Summary

Over the last few years, new proteomics methods have been developed for making quantitative comparisons using stable isotope labeling. Although these methods have paved the way for quantitative proteomics, the analysis of these data is often the rate-limiting step. In fact, many analyzes are still carried out manually, which adds a level of subjectivity to the data that will vary between laboratories and even analysts. In this chapter, we have attempted to summarize several of the key steps necessary for an individual to automate the analysis of quantitative proteomics data. The approach is straightfoward to implement for an individual with moderate programming experience and used to process proteomics data in an objective manner.

Key Words: Computer software; stable isotope labeling; quantitative proteomics; mass spectrometry; isotopomer ratios; automated analyses.

1. Introduction

Over the last decade, mass spectrometry (MS) has become the premier analytical method for the quantitative analysis of individual analytes in complex biological matrices. To ensure high precision and accuracy throughout the measurement, many quantitative MS methods measure the analyte of interest relative to a corresponding stable isotope-labeled analog of the same molecule. The use of this internal standard (IS) mimics losses and drifts that occur during the sample preparation and analytical measurement process, because the compound and IS experience similar biases and errors. Because the mass spectrometer measures a ratio of the ion currents produced by the sample and its respective IS and not the mole ratio directly, the ion current ratio is converted

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

to a mole ratio using a calibration curve from the response of samples of "known" abundance relative to the labeled IS.

Quantitative proteomics is further complicated because there are thousands of protein levels that need to be measured within a single or limited number of measurements. To address this problem, many elegant approaches have been reported for the incorporation of "heavy'" stable isotope atoms into proteins or peptides—creating a global IS for every protein in a complex mixture. Several of these methods are described in detail in other chapters of this text and have made quantitative protein analyses using MS possible. This labeled IS can then be added to multiple different unlabeled samples and used as a common reference standard for relative comparisons between samples. The peptides containing the natural abundance atoms and enriched "heavy" atoms from these proteins are then identified by liquid chromatography–tandem mass spectrometry (LC–MS/MS), and software can be used to evaluate the relative abundance between the unlabeled and labeled peptide pairs.

Several computer programs have been described for the analysis of quantitative proteomics data using stable isotope labeling ***(1–3)***. Most programs developed for analyzing quantitative LC–MS/MS data first derive extracted ion chromatograms from the precursor ion scans for the peptide pairs and then compute relative abundances from the background subtracted areas of the two extracted ion chromatograms. This process is complicated by both instrument and chemical noise, differences between the isotope distributions of the unlabeled and labeled peptide, and in some cases chromatographic separation between the unlabeled and labeled peptide forms. Furthermore, how can we account for differences in measurement response between the two peptide forms without a standard curve for each and every peptide measured during the analysis? Quantitative analysis software should be capable of assessing the quality of the quantitative data, eliminate outliers, and estimate the protein ratio from multiple peptides in an objective manner without user intervention.

This chapter describes a step-wise approach for the computational analysis of quantitative proteomics data using stable isotope labeling. We divide the analysis into six separate steps (**Fig. 1**) that are interchangeable with alternative strategies. Many of the steps in the analysis are implemented as components of the freely distributed computer program RelEx (http://fields.scripps.edu/relex/) ***(2)***. Nevertheless, as many of the steps are generalized, we hope to provide a sufficient description of the approach so that an individual with moderate programming experience could reproduce the process for their own needs.

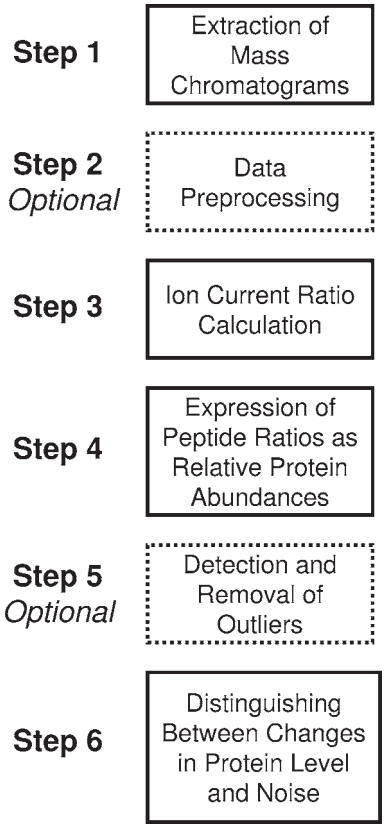


Fig. 1. General six-step scheme for the computational analysis of quantitative proteomics data. **Steps 2** and **4** are optional and can be eliminated depending on the experiment and the quality of the data.

2. Methods

2.1. Extraction of Ion Chromatograms

Nearly all analyses of quantitative data from LC–MS are performed on ion chromatograms instead of individual mass spectra. Because analytes elute over time, the combined use of data from the entire elution profile provides greater ion statisitics and thus greater precision. Ion chromatograms facilitate the assessment and subtraction of background—an essential step to minimize sys-

tematic errors in complex mixtures. Additionally, because some stable isotope labels will result in a chromatographic separation between unlabeled and labeled peptide pairs, the ratio between the separated peptide pairs will vary dramatically from scan to scan, whereas the ratio of the respective chromatographic areas will be affected minimally. Thus, the first step in calculating the relative abundance between an unlabeled and stable isotope-labeled peptide identified by LC–MS/MS is the extraction of ion chromatograms.

The peptide sequences identified by MS/MS and database searching are used to calculate the elemental composition and predict the isotope distribution for the unlabeled and labeled peptide pairs. The predicted isotope distributions are used to define *m/z* ranges for the extraction of the respective unlabeled and labeled ion chromatograms. These ion chromatograms are normally extracted from the MS scans (precursor scans) within a time window where the MS/MS spectrum was acquired that identified the respective peptide sequence. Software to extract ion chromatograms should not be restricted to any specific stable isotope labeling approach and can be easily adopted for any quantitative proteomics labeling technique.

The accurate conversion of measured ratios in the mass spectrometer to mole ratios is complicated and care must be taken during the ion chromatogram extraction to minimize the introduction of systematic errors because a polyatomic molecule containing enriched atoms will result in a combinatorial distribution of isotopomers depending on the enrichment and number of "labeled" atoms. Additionally, the isotope distributions *will be different* between the unlabeled and labeled isotope distributions. This difference between isotope distributions is particularly problematic for analytes containing ^{13}C-enriched molecules. Because a majority of the isotope distribution from an organic molecule is a result of natural abundance ^{13}C, a peptide that contains multiple enriched ^{13}C atoms will have a substantial difference between the isotope distributions of the unlabeled and labeled peptides (**Fig. 2**). These effects could potentially cause a large error in the mole ratio measurement depending on which isotope peak or peaks were used to determine the ratio (*see* **Note 1**).

Although most quantitative proteomics approaches obtain quantitative data from the MS scans, most traditional quantitative MS analyses use MS/MS to acquire data for selected analyses to minimize chemical interferences and improve sensitivity. These analyses are difficult to extend to proteomics because we often do not know the identities of the analytes being quantified unless the experiments are targeted toward individual peptides unique to specific proteins ***(4)*** (*see* **Note 2**).

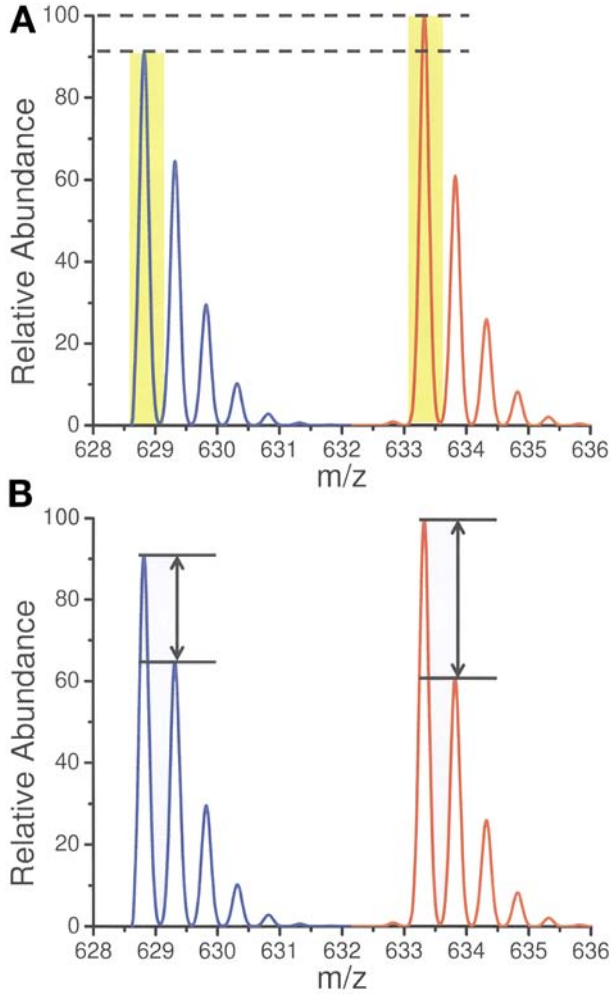


Fig. 2. The isotope distribution from a peptide containing only natural abundance isotopes and the isotope distribution of a peptide enriched in ^{13}C will be different. **(A)** The difference in intensity of the base isotope peak (most intense isotope peak) is significant for the peptide YAGILDC$_{ICAT}$FK with the ICAT tag containing only natural abundance isotopes and the ICAT tag enriched with 9 x ^{13}C at greater than 99.9% atom percent excess. **(B)** The difference in intensity of the base isotope peak is because there are now nine less carbon atoms contributing 1.09% to the M+1 isotope. Because of mass balance, the sum of the two isotope distributions are the same but the intensity of each isotope peak will be different. The same thing happens with enriched isotopes of ^{15}N and ^{2}H, however, their effect on the two isotope distributions is less pronounced because they have much lower natural abundance.

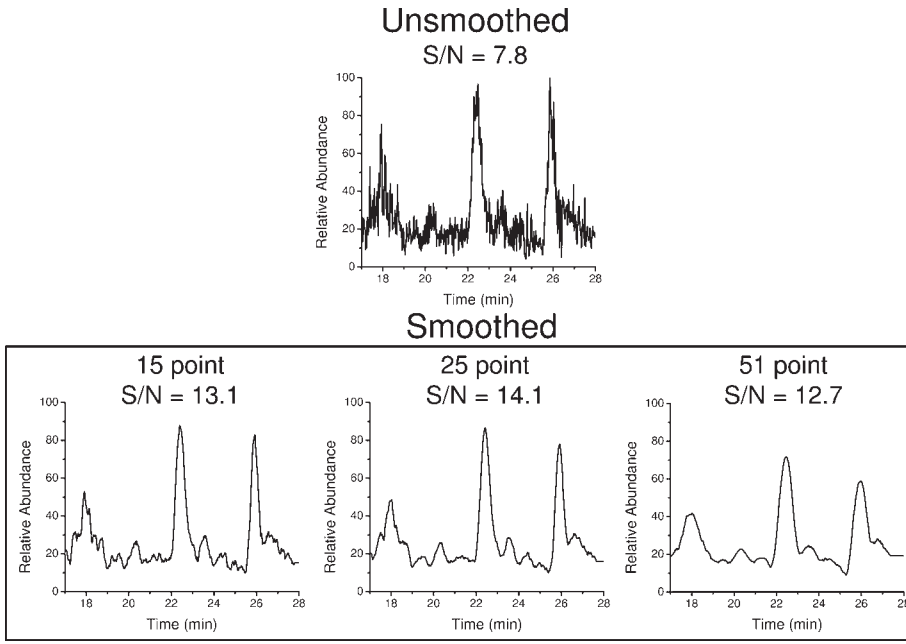


Fig. 3. Effect of Savitsky–Golay smoothing on the signal to noise and distortion of extracted ion chromatograms. Smoothing greatly improves the quality of raw mass spectrometry data until the smoothing window becomes greater the full width at half maximum of the peak. When the smoothing window is too large, the resolution begins to deteriorate and the peaks shrink and broaden.

2.2. Data Preprocessing—Smoothing and Noise Reduction

MS data is composed of a combination of both signal and noise. Unless the number of ions contributing to the measured signal is large, the automated analysis of the measured chromatographic peak will be complicated because of the random errors in the measurement, and the data will result in an irregular signal that is difficult to process. Because most noise is of greater frequency than the signal, most mass spectrometer data systems and quantitative proteomics data analysis software use some smoothing capability as a preprocessing step to aid in the detection of the chromatographic peak and assessment of the peak area(s). Although there are numerous approaches for smoothing data to reduce noise, we will focus on one of the most frequently implemented techniques—the method of Savitzky and Golay ***(5)***.

The basis of the Savitsky–Golay method involves the fitting of a small subset of the data to a polynomial using a least squares regression. In this approach,

the least squares value of a given point is calculated as a weighted combination of itself and *m* points on either side of it. Although the implementation of this approach may initially appear complicated, in practice, the Savitsky–Golay convolution method is as easy to implement as a weighted moving average. This implementation is further simplified by Gorry who reported a simple and general procedure for calculating the convolution weights at all positions ***(6)***—including the endpoints that were truncated in the initial development. Furthermore, Gorry reported an extension of the Savitsky–Golay approach for smoothing data of unequally spaced data ***(7)***, as characteristic with ion-trap mass spectrometers using automatic gain control (*see* **Note 3**).

Figure 2 demonstrates the effect of different sized Savitsky–Golay smoothing windows on the raw extracted ion chromatograms. A quadratic smooth was implemented exactly as described by Gorry ***(6)*** using different window sizes. In general, as the window size is increased the noise is reduced, but after a certain window size the improvement over smaller window sizes becomes insignificant. However, when the window becomes too large, peaks become distorted and there is a loss of resolution along the time axis—the peaks become lower and broader. Thus, although a larger window will further reduce the noise, a window that is too large will unfortunately also decrease the signal-to-noise ratio (S/N) and significantly distort the resulting peak shape (*see* **Note 4**).

2.3. Ion Current Ratio Calculation

Software to automatically derive the area under a chromatographic peak is often the most challenging aspect of any quantitative MS analysis. These programs must determine where the peak starts and ends, while also determining the contribution from background on which the peak is superimposed. The estimate of peak onset and background requires that the peaks be well defined. However, with proteomics data, the chromatograms are often crowded and an algorithm may not have a sufficiently large region of pure baseline to appropriately subtract the true background. Furthermore, although the desired output is the background subtracted ion-current ratio, most algorithms treat the two ion-chromatograms separately instead of as a pair. For chromatograms that are detected and integrated separately, the precision and accuracy of the ratio will ultimately be limited by the software's ability to handle the ion chromatogram with the poorest S/N. Although the objective and reproducible assessment of peak area is reliable for high S/N chromatograms, peaks of even modest S/N are difficult to integrate reproducibly. Any objective criteria that can be used to evaluate peak locations are often dependent on peak shape and will likely differ substantially between the thousands of peptides identified in a μLC/μLC/MS/MS run.

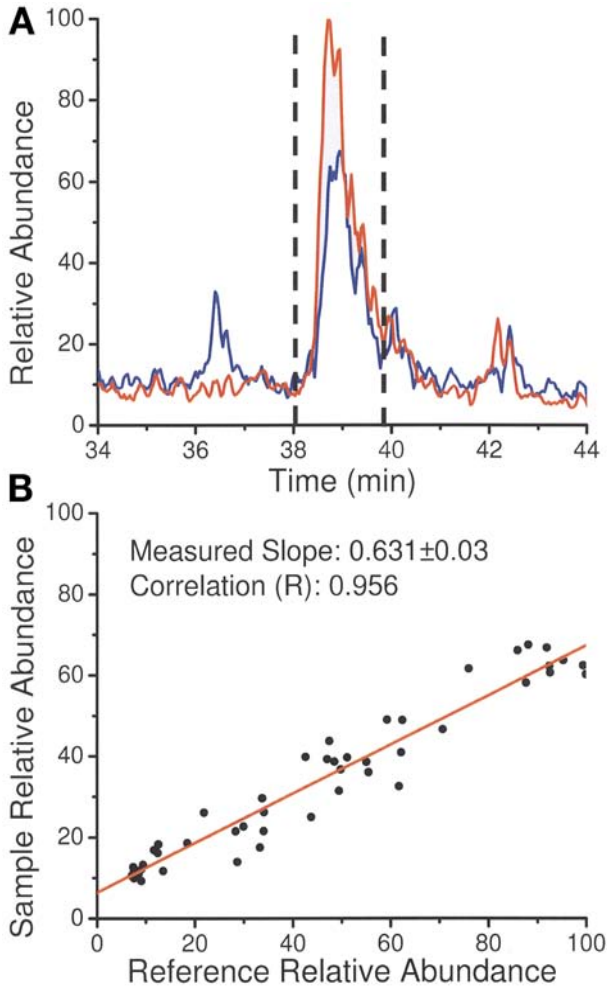


Fig. 4. Use of a least squares regression for the calculation of the ion-current ratio between two isotopomers. The intensity of the ^{15}N-enriched reference chromatogram (gray) is plotted vs the intensity of the unlabeled sample chromatogram (black). The best fit line through the data is calculated and the slope of the line represents the background subtracted ratio.

Thorne et al. ***(8)*** described an elegant approach for the calculation of background subtracted ratio between two isotopomers. This algorithm uses a least squares regression to evaluate the background subtracted mass spectrometer ion-current ratio from two extracted ion chromatograms. In a single calculation, the slope of the regression provides a measure of the background sub-

tracted ratio, the intercept provides a measure of the ratio of the two backgrounds, and the correlation coefficient provides a measure of the ratio quality (**Fig. 4**). Using this approach, both chromatograms are handled simultaneously, and the quality of the ratio is independent of the of the chromatogram peak shape and is only marginally effected by the algorithm's chosen start and stop points. Furthermore, methods that use traditional peak integration where the two chromatograms are treated independently will always be limited by the ability of the software to detect and integrate the extracted ion chromatogram of the lowest S/N. In contrast, because the least squares regression algorithm considers both chromatograms simultaneously, the peak detection only needs to be performed on the isotopomer with the greatest intensity.

Because both the unlabeled and labeled chromatographic traces will have error, the best fit line cannot be determined using a simple linear regression. In a simple linear regression, a line is found by minimizing the sum of the squares of the vertical distance from each point to the line. However, when error is present on both the *x*- and *y*-axis, then the regression should minimize the sum of the squares of the perpendicular distance from the line to each point. A simple and fast implementation of a linear regression when the data has error on both axises has previously been described by York ***(9)*** (*see* **Note 5**).

2.4. Expression of Peptide Ratios as Relative Protein Abundances

The most common and obvious treatment of quantitative proteomics data is to average the ratios of the individual peptides that map back to an individual protein sequence. An important, but often overlooked, consequence of quantitation on the peptide level is that any extrapolation of that information back to the protein level may or may not be entirely appropriate. Because many proteins can give rise to the same peptide sequence postdigestion, the resulting ratio will be a weighted average of the individual protein forms giving rise to that peptide. Even if the peptide is unique to a specific protein loci, the resulting peptides will provide a weighted average of different localized forms, posttranslationally modified isoforms, and so on. For example, if a peptide's relative abundance decreases between two conditions, it could be because the protein it was derived from decreased in abundance or because the stoichiometry of the unmodified protein form decreased as a greater fraction of the protein obtained a posttranslational modification.

Figure 5 provides an illustrated example for an unlabeled vs ^{15}N-enriched peptide ratio measured for several different peptides that map to a single gene sequence. In a relatively simplistic situation where a protein only exists in three different forms, the interpretation of the peptide data would be complicated. In this example, the average of the measured unmodified peptide ratios would result in a very noisy measurement with minimal change from unity. Although

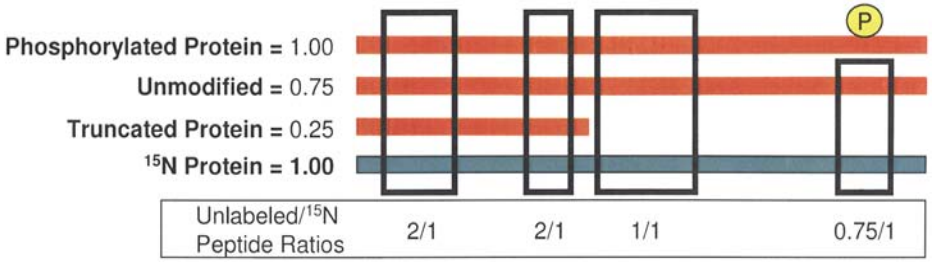


Fig. 5. The limitation of making protein measurements on the peptide level. A single protein loci giving rise to multiple protein forms will be difficult to interpret on the peptide level. This difficulty in interpreting the results is particularly complicated if there is only a couple peptide ratios measured as the interpretation of the resulting protein ratio will depend heavily on the relative location of the respective peptide.

the individual peptide ratios can provide data that is biologically meaningful, the mean of peptides spanning different protein isoforms mapping to the same gene loci may be less legitimate. This should be a reminder of the importance of obtaining quantitative proteomics data with high-sequence coverage across every protein. Effort needs to be made to quantify peptides throughout the entire protein sequence and we should all be skeptical of single peptide ratios.

We are not saying that the weighted relative abundance of the individual peptides will never be biologically informative. We and others have reported changes in protein abundance using relative abundance measurements on the peptide level; however, the caveat that the measured peptide ratio may not be an accurate reflection of the protein ratio should never be forgotten.

2.5. Detection and Removal of Outliers

An outlier in a dataset is an observation that is inconsistent with the remainder of the data. The data points that deviate from the others result in a significant increase in the standard deviation and complicates our ability to identify differences between datasets. Thus, we need to identify and remove these outliers before performing further analyses. In proteomics, the multiple peptide measurements provide a measure of uncertainty for the relative protein abundance measured by quantitative proteomics ***(10)***. Outlier peptide ratios could be a result of an incorrect sequence identification, a posttranslationally modified peptide, poor S/N chromatogram, etc. Detection of these outliers in an automated and objective manner is essential for the estimate of the relative protein abundance from the average of peptide ratios mapping to a single protein locus.

There are two common and equally appropriate approaches to the removal of outlier ratios measured by quantitative MS. Both the Dixon's Q-test ***(11,12)*** and the Grubbs test for outliers are commonly used for the rejection of outliers in a dataset. Both the quantitative proteomics packages RelEx ***(2)*** and ASAPRatio ***(3)*** have implemented a Dixon's Q-test for the removal of individual peptide outliers prior to the calculation of the protein ratio.

2.6. Assignment of Significance

Most quantitative proteomics experiments assign significant changes in relative protein abundance if the mean peptide ratios exceed an empirically derived threshold ratio. The use of a threshold ratio to extrapolate conclusions is inappropriate because the precision is not constant and is a function of the number of ions acquired during the analysis and the ratio that was measured ***(13)***. An important and often overlooked consideration of any quantification measurement by isotope dilution MS is that the precision and accuracy is at a maximum for ion-current ratios near 1:1. Although counterintuitive, quantitative measurements for large differences are actually less accurate and precise than small differences ***(13,14)***. Additionally, because quantitative proteomics measurements are complicated from multiple different protein isoforms giving rise to different peptides (*see* **Subheading 2.4.**) a measure of significance needs to be measured on a protein-by-protein basis.

Thus, to identify differences we advise measuring two ratios where each sample is measured relative to a single standard instead of measuring differences within a single ratio. For two samples or conditions with the identical stable isotope labeled-internal protein standard, statistical differences between means for the same protein can be established using a simple t-test. For multiple analyses, the significance can be estimated using analysis of variance (ANOVA). An added benefit of using two ratios is that any systematic error should be present in both analyses and should cancel ***(10)*** (*see* **Note 6**).

3. Notes

1. This problem is minimized by summing the entire isotope distribution or determining the fraction of the total distribution that a single ion represents to derive a correction factor in the calculation of the mass spectrometer ion-intensity ratio.
2. Venable et al. ***(15)*** reported a novel proteomics approach that makes use of modern fast scanning tandem mass spectrometers to acquire tandem mass spectra serially over preselected *m/z* ranges throughout the chromatographic separation. These data are used to qualitatively identify the peptide sequences and quantitatively measure the relative abundance between the unlabeled and labeled peptide pairs. Because the tandem mass spectra are acquired multiple times as peptides

elute from the chromatographic column, these data are used to produce reconstructed ion chromatograms from tandem mass spectra. Although the data acquisition and generation of ion chromatograms is unique, all subsequent and downstream data processing is performed normally.

3. An alternative to using the modified Gorry approach for polynomial smoothing of irregularly spaced data ***(7)*** is to interpolate the irregularly spaced peak intensities into regularly spaced intervals.
4. To minimize the peak distortion, the width of the smoothing window should be less than 90% of the peak full width at half maximum height ***(16)***.
5. Recently, Pan et al. presented a thorough comparison of the regression approach with traditional peak integration and reported superior quantitative accuracy and dynamic range ***(17)***. Thus, there are several advantages to this approach and, to date, the only program that has implemented the least squared regression for calculating MS ion-current ratios in proteomics is RelEx ***(2)***.
6. The use of a ratio of two ratios for relative quantitation improves the accuracy by greater than 30% when compared with using a single ratio alone ***(2)***.

Acknowledgment

The authors gratefully acknowledge the criticism and feedback from the users of our software. Financial support was provided in part from a Basic Research Grant from Merck Research Laboratories and National Institutes of Health grants P41-RR011823, P30-AG013280, and K22-AI059076.

References

1. Han, D. K., Eng, J., Zhou, H., and Aebersold, R. (2001) Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.* **19,** 946–951.
2. MacCoss, M. J., Wu, C. C., Liu, H., Sadygov, R., and Yates, J. R., III (2003) A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. *Anal. Chem.* **75,** 6912–6921.
3. Li, X. J., Zhang, H., Ranish, J. A., and Aebersold, R. (2003) Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. *Anal. Chem.* **75,** 6648–6657.
4. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. USA* **100,** 6940–6945.
5. Savitzky, A. and Golay, M. J. E. (1964) Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36,** 1627–1639.
6. Gorry, P. A. (1990) General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method. *Anal. Chem.* **62,** 570–573.
7. Gorry, P. A. (1991) General least-squares smoothing and differentiation of nonuniformly spaced data by the convolution method. *Anal. Chem.* **63,** 534–536.

8. Thorne, G. C., Gaskell, S. J., and Payne, P. A. (1984) Approaches to the improvement of quantitative precision in selected ion monitoring: high resolution applications. *Biomed. Mass Spectrom.* **11,** 415–420.
9. York, D. (1966) Least-squares fitting of a straight line. *Can. J. Phys.* **44,** 1079–1086.
10. MacCoss, M. J. and Matthews, D. E. (2005) Quantitative MS for proteomics: teaching a new dog old tricks. *Anal. Chem.* **77,** 294A–302A.
11. Dean, R. B. and Dixon, W. J. (1951) Simplified statistics for small numbers of observations. *Anal. Chem.* **23,** 636–638.
12. Rorabacher, D. B. (1991) Statistical treatment for rejection of deviant values: Critical values of Dixon's "Q" parameter and related subrange ratios at the 95% confidence level. *Anal. Chem.* **63,** 139–146.
13. MacCoss, M. J., Toth, M. J., and Matthews, D. E. (2001) Evaluation and optimization of ion-current ratio measurements by selected-ion-monitoring mass spectrometry. *Anal. Chem.* **73,** 2976–2984.
14. Peterson, D. W. and Hayes, J. M. (1978) *Contemporary Topics in Analytical and Clinical Chemistry, Vol. 3.*, (Hercules, D. M., Hieftje, G. M., and Evenson, M. A., eds.), Plenum Press, New York.
15. Venable, J. D., Dong, M. Q., Wohlschlegel, J., Dillin, A., and Yates, J. R. (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat. Methods* **1,** 39–45.
16. Alfassi, Z. B., Boger, Z., and Ronen, Y. (2005) *Statistical Treatment of Analytical Data*. CRC Press, Boca Raton, FL.
17. Pan, C., Hurst, G., McDonald, W. H., et al. (2005) Novel algorithms for determination of peptide abundance ratios in quantitative proteomics with stable isotope labeling. *Proceedings of the 53rd ASMS Conference on Mass Spectrometry and Allied Topics.* San Antonio, TX.

13

Quantitative Proteomic Analysis of Mammalian Organisms Using Metabolically Labeled Tissues

Christine C. Wu and Michael J. MacCoss

Summary

Metabolic labeling of mammalian organisms with stable isotopes can be used to provide tissue-specific internal standards for use in quantitative proteomic analyses. This method provides an alternative and complementary strategy to covalent modification approaches using isotope-coded mass tags. This chapter will focus on the generation of the isotope-labeled tissues, the analysis of the sample using Multidimensional Protein Identification Technology, and the computational analysis of the mass spectrometric data acquired.

Key Words: Stable isotope-labeled tissue; quantitative proteomics; atomic enrichment; internal standard; MudPIT.

1. Introduction

The capability to quantitatively compare differences in protein level between samples is an essential component of proteomics technology. These analyses represent a complex technical challenge because although a mass spectrometer's response is usually linear for most peptide analytes ***(1)***, there are numerous sample preparation steps each with different degrees of efficiency. To account for this sample preparation variability and improve the overall measurement precision, most proteomics laboratories have applied methods that introduce heavy stable isotopes into one sample to be compared against a second sample containing only natural abundance isotopes ***(2)***. These two samples are then mixed and the labeled and unlabeled proteins act as mutual internal standards (IS) for one another. Assuming that the two samples are labeled and mixed

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

prior to any sample preparation, the two forms should undergo similar losses and biases.

Metabolic labeling strategies facilitate these comparisons by producing appropriate IS for all proteins in a given cell type, tissue, or organism. The use of metabolic labeling has the added advantage of producing standards that can (and usually are) mixed prior to any sample preparation. Thus, samples can be mixed directly as intact cells, homogenized tissues, or even intact organisms. Metabolic labeling with stable isotopes for the purpose of producing IS for quantification is becoming routine and has been applied to a variety of organisms ranging from bacteria ***(3)*** and yeast ***(4)***, to flies and worms ***(5)***, and even mammals ***(6)***. The focus of this chapter is on the labeling of mammals, in particular rats and mice. However, these methods can be applied to any organism that the diet can be defined using stable isotope-enriched material.

This chapter summarizes the methodology available to date for the production of metabolically labeled IS for comparative proteomic analyses of mammalian tissue. We will describe the labeling of rodents with ^{15}N-enriched diets to produce tissue-specific IS and the proteomic analysis of the samples using Multidimensional Protein Identification Technology (MudPIT). Modifications have been made to many of the original labeling steps ***(6)*** to simplify the overall process, improve the efficiency of the labeling, and reduce the cost of producing the IS.

2. Materials

2.1. ^{15}N-Enriched Rodent Diet

1. Algal whole cells: ^{15}N-enriched *Spirulina* (>99%) (Spectra Stable Isotopes, Columbia, MD).
2. Protein-free custom research diet in powder form (Harlan Teklad, Indianapolis, IN): 631.8 g/kg sucrose, 200.0 g/kg corn starch, 54.6 g/kg corn oil, 66.5 g/kg cellulose, 10.0 g/kg vitamin mix no. 40060, 0.01 g/kg ethoxyquin, 13.37 g/kg mineral mix no. 79055, 23.72 g/kg $CaHPO_4$, and 0.038 g/kg $CaCO_3$.
3. Nonstick surface (e.g., aluminum foil, wax paper).
4. Mortar and pestle.
5. Razor blade.
6. Food dehydrator.

2.2. Tissue Harvest

1. Dissection tools (scalpels, scissors, forceps).
2. Glass Petri dishes.
3. Phosphate-buffered saline: prepare 10X stock with 80 g NaCl, 2.0 g KCl, 14.4 g Na_2HPO_4, 2.4 g KH_2PO_4 in 1 L of Milli-Q water (adjust to pH 7.4 with HCl if necessary). Prepare working solution by dilution of one part with nine parts water.
4. Cryotubes for storage of tissue.
5. Liquid nitrogen.

2.3. Tissue Preparation

1. Homogenization buffer: 100 m*M* K_2HPO_4/KH_2PO_4, pH 6.7, 250 m*M* sucrose, and 5 m*M* $MgCl_2$.
2. Dounce homogenizer.
3. Lowry H Protein Assay (Bio-Rad, Hercules, CA).

2.4. Protein Sample Digestion

1. Trypsin digestion buffer: 0.2% RapiGest SF (Waters, Milford, MA) dissolved in 50 m*M* NH_4HCO_3 (pH 8.0) and 1 m*M* $CaCl_2$.
2. Insulin syringe.
3. 500 m*M* dithiothreitol.
4. 500 m*M* iodoacetamide.
5. Modified trypsin reconstituted in 0.01% acetic acid (1 μg/μL) (Promega, Madison, WI).

2.5. Multidimensional Liquid Chromatography

1. Fused silica capillary tubing: 100 μm inner diameter (Polymicro, Phoenix, AZ).
2. Sutter P-2000 laser puller (Sutter Instruments, Novato, CA).
3. 5-μm Luna C18 chromatography material (Phenomenex, Ventura, CA).
4. 5-μm Partisphere strong cation exchange chromatography material (Whatman, Clifton, NJ).
5. A homemade pressure bomb interfaced with a high pressure helium gas cylinder. The pressure bomb is similar to the one described previously by Yates et al. ***(7)***.
6. High-performance liquid chromatography binary pump and autosampler.
7. Tandem mass spectrometer with a nanospray ion source.

2.6. Software for the Analysis of Quantitative Proteomics Data

1. SEQUEST: *see* http://www.sequest.org.
2. DTASelect: *see* http://fields.scripps.edu/DTASelect/.
3. RelEx: *see* http://fields.scripps.edu/relex/.

3. Methods

3.1. Preparation of Labeled Diet

The consistency of the ^{15}N-*Spirulina* is variable depending on the lot from the manufacturer. Therefore, the stable isotope-labeled algae cells are first ground into a fine powder using a mortar and pestle. The uniform algal powder is then mixed with protein-free powder diet in a 1:3 ratio to produce a diet of approx 20% total protein. The diet can be modified further with the addition of corn oil to accommodate higher fat requirements for fertilization, gestation, and lactation (*see* **Note 1**).

Mixed diet in powder form can be used as is or formulated into pellets by mixing 15 mL H_2O per 100 g diet and kneading the wet powder into dough.

The dough is then rolled out onto a nonstick surface and cut into sections. These pellets are dried in a food dehydrator set at 140–150°C for 24 h or until solid. The pellets should be green-black in color.

3.2. Animals (see Note 2)

Total number of animals labeled should reflect sample requirements for the proteomic analysis. Animals should be purchased a few days prior to weaning so that pups can acclimate to the animal facility and labeling can be initiated upon weaning. Requirements for gender specificity should reflect the proteomic experiment.

3.3. Labeling

1. General labeling scheme: the general scheme for ^{15}N labeling a rat or mouse is shown in **Fig. 1**. The animals are fed an isotope-enriched diet for a predetermined period of time depending on the tissue and required enrichment. After sacrificing the animal, the tissues are characterized for enrichment and morphology prior to use as an IS.
2. *Ad libitum* feeding: labeled diet in pellet form can be made freely accessible in feeder container in rodent cage. Although labeling can also be performed on a more regular schedule (e.g., every 6 h ***[6]***) to reduce the consumption and losses of the stable isotope-labeled diet, this restricted feeding will result in a lower animal mass (*see* **Fig. 2**). This slower growth results in a lower percent body fat and decreased success in gestation (*see* **Note 1**).
3. Labeling is initiated at weaning when the animals can eat solid food (~3 wk for mice and rats). Animals should be monitored daily. Labeling is terminated when the animal is euthanized and tissue is harvested. The labeling time is species and tissue specific. The time required for labeling protein will reflect the tissue specific average protein half-life.

3.4. Tissue Harvest

Animals are euthanized using IACUC-approved procedures specific to experimental protocol. Tissues should be collected immediately on ice and rinsed in phosphate-buffered saline. For long-term storage of frozen tissue, tissue should be sliced into 1-mm slices, snap frozen in cryotubes in liquid nitrogen, and transferred to a –80°C freezer (*see* **Note 3**).

3.5. Characterization of the Labeled IS

1. Tissues are harvested at appropriate time points after the initiation of labeling to ensure that the ^{15}N enrichment of protein in the respective tissues is greater than 90%. These time points are organism and tissue specific (*see* **Note 4**). An example of the ^{15}N amino acid enrichment from protein in selected rat tissues after labeling for 44 d starting at 3 wk of age (after weaning) is shown in **Fig. 3**.

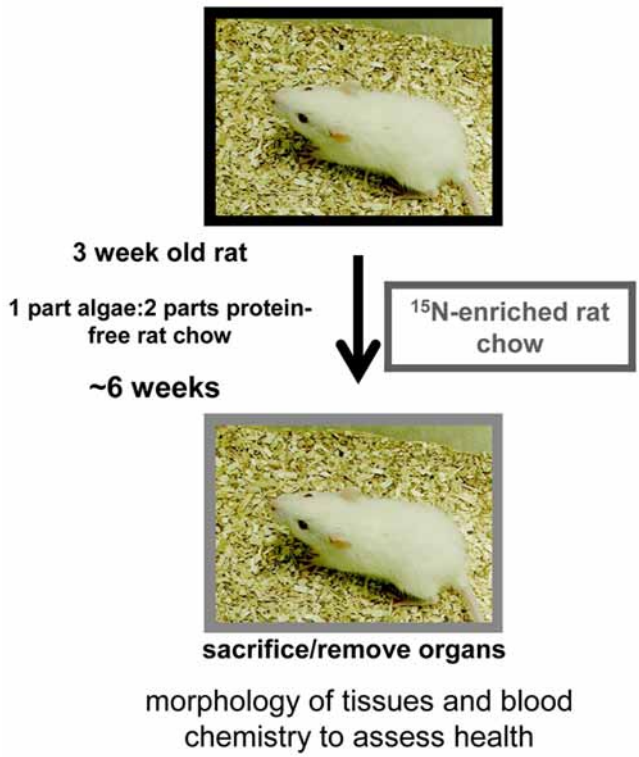


Fig. 1. General scheme for the metabolic labeling of rodents with a ^{15}N-enriched diet for use as a stable isotope-labeled internal standard.

2. Intact tissues are also characterized visually to ensure no physical abnormality and electron micrographs are taken of tissue sections to verify proper ultrastructure. All comparisons are made on a littermate fed an identical diet containing only natural abundance isotopes. An example of the tissue comparisons is illustrated in **Fig. 4**. We also use histology and standard chemical panel for blood chemistry measurements to validate animal health.

3.6. Protein Sample Preparation

1. Tissues are hand homogenized in homogenization buffer at a ratio of 1 g tissue in 5 mL buffer using 10–30 strokes in a Teflon homogenizer until homogeneous (varies with tissue).
2. Homogenates are centrifuged at 3000*g* for 15 min to pellet tissue aggregates, unbroken cells, and nuclei.
3. The protein concentration of the labeled postnuclear supernatant is assessed using a standard protein assay.

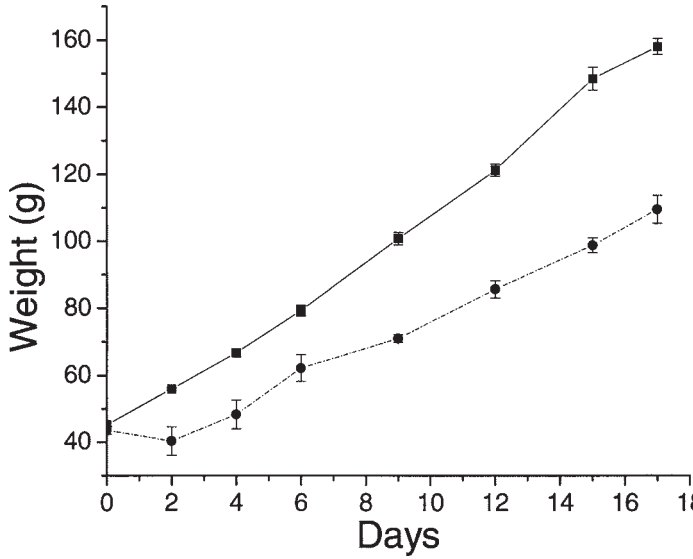


Fig. 2. The effect of *ad libitum* vs restricted feeding on rat growth. The solid line is for an unrestricted *ad libitum* diet and the dashed line is for rats that are fed four times daily (6 AM, noon, 6 PM, and midnight). Data are mean ± SD for $n = 3$ rats.

4. The labeled standard is then mixed in a 1:1 ratio (w:w) with multiple different protein samples for relative quantitation. To avoid systematic errors and minimize any potential effect from the unusual rodent diet, the IS must be added to a minimum of two samples: one control sample and one alternate condition/timepoint/and so on. (**Fig. 5**) ***(2,6,8)***.
5. The postnuclear supernatant of the mixed sample is diluted twofold with 100 m*M* K_2HPO_4/KH_2PO_4, pH 6.7 and centrifuged at 100,000*g* in an ultracentrifuge to separate the soluble proteins (supernatant) from the cellular membranes (pellet). Additional fractionation (e.g., subcellular or biochemical fractionation) can be performed on the mixed sample and any losses can be accounted for by the IS.

3.7. Proteomic Analysis Using MudPIT

1. Membrane pellets (100 µg) are resuspended in 100 µL digestion buffer by drawing three times through an insulin syringe and incubating on ice for 15 min. The membrane suspension is vigorous vortexed and the pellet broken to create a fine protein suspension. The sample is reduced with the addition of 500 m*M* dithiothreitol to a final concentration of 15 m*M* and incubation at 55°C for 15 min. The sample is alkylated with the addition of 45 m*M* iodoacetamide to a final concentration of 50 m*M* and incubation at room temperature in the dark for 15 min. Modified trypsin is added at a 1:50 enzyme:substrate ratio and the sample is incubated at 37°C for 4 h. The reaction is quenched with the addition of HCl to a

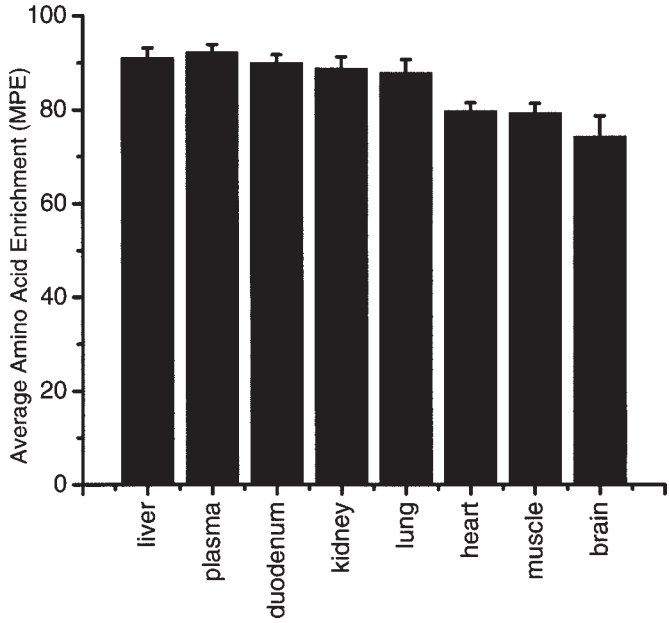


Fig. 3. Measured ^{15}N amino acid enrichments in labeled tissue. The *M+i/M* isotopomer ratios were measured in both unlabeled and ^{15}N-enriched tissues by gas chromatography–mass spectrometry where *M* is *m/z* of the fragment for the natural abundance amino acid, and *i* is the number of nitrogens in the respective amino acid. Enrichments were calculated and expressed in mole percent excess. The mean ± SD is shown for 14 amino acids in eight different tissues.

final concentration of 500 m*M* and incubated at 37°C for 45 min to hydrolyze the RapiGest SF in the digestion buffer. The protein sample is then centrifuged at 15,000*g* for 5 min and the supernatant stored at –80°C until analysis by mass spectrometry (MS).

2. Soluble proteins (100 µg) are precipitated using MeOH/$CHCl_3$ extraction as described previously ***(9)*** and resuspended in 100 µL trypsin digestion buffer. The solution is vigorously vortexed/sonicated to solublize the protein pellet. After solubilization the protein is reduced, alkylated, and digested as described in **Subheading 3.7., step 1** for insoluble membrane pellets.
3. The resulting peptides are analyzed using multidimensional microcapillary liquid chromatography–tandem MS (µLC/µLC/MS/MS). The mass spectrometer is configured to acquire a single "survey" mass spectrum followed by multiple tandem mass spectra in a data-dependent manner. Details of the multidimensional protein identification technology have been described in detail elsewhere ***(10–12)*** and is beyond the scope of this chapter.

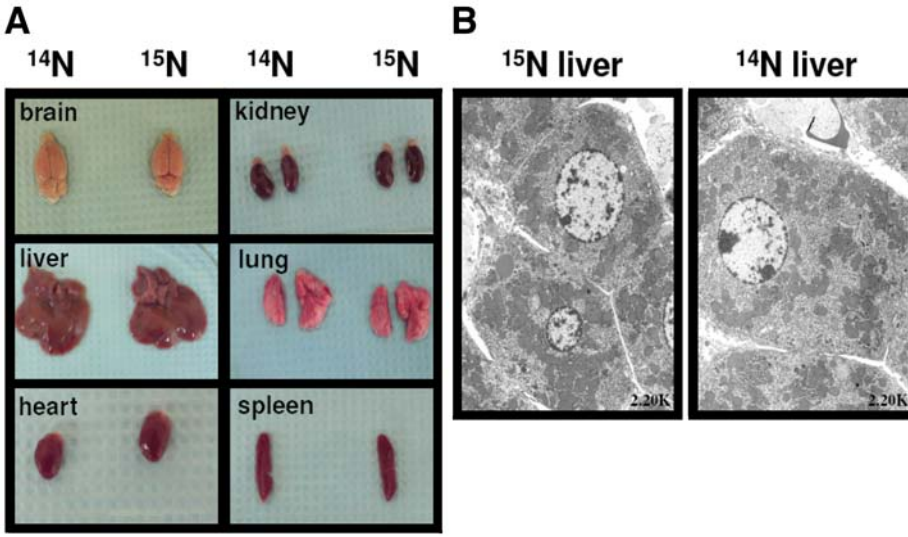


Fig. 4. Confirmation of animal health using multiple approaches on different levels. **(A)** Selected ^{15}N-enriched tissues are visually compared with tissues from a control littermate fed an identical diet containing only natural abundance isotopes to identify any potential abnormalities. **(B)** Cellular ultrastructure is also checked for abnormalities using electron microscopy.

3.8. Computational Analysis Using SEQUEST, DTASelect, and RelEx

1. Database Search: tandem mass spectra must be searched twice (sample and reference) using the program SEQUEST ***(13)*** to create two separate sets of output files stored in two separate directories. The first database search (sample-unlabeled mass parameters) uses a sequest.params file containing only a single static modification of +57 *m/z* on cysteine from carbamidomethylation using iodoacetamide. The second database search (reference-^{15}N-corrected mass parameters) uses a sequest.params file containing a static modification on each amino acid shifting the average mass to account for the enriched nitrogen atoms (*see* **Notes 4** and **5**).
2. DTASelect: the program DTASelect ***(14)*** is used to filter the peptide sequences identified by SEQUEST and select peptide sequences exceeding a limiting threshold of requirements to minimize the false discovery rate to less than 5% (estimated from the number of proteins obtained from the decoy protein sequences; **Note 5**), assemble the peptides into protein identifications, and remove redundant protein identifications.
3. RelEx: for each peptide exceeding the DTASelect criteria in both the sample and reference searches, ion chromatograms are extracted from the Xcaliber data file for the unlabeled and ^{15}N-labeled peptide isotope distributions using the program

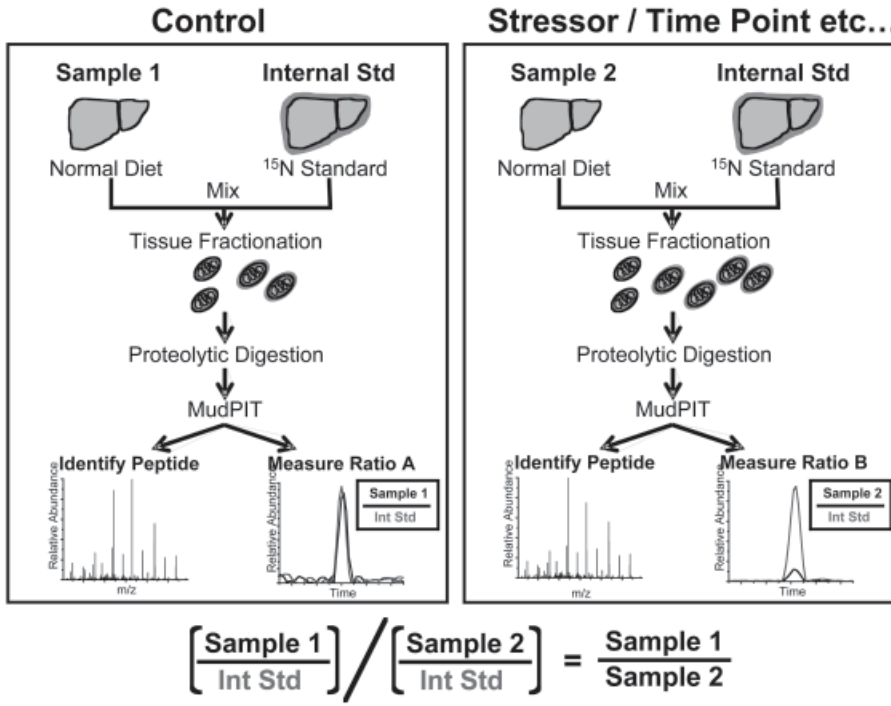


Fig. 5. General approach for measuring relative protein differences using metabolically labeled internal standards (IS). Each sample is mixed with an identical ^{15}N-enriched IS prior to sample preparation. Differences are evaluated using a ratio of two ratios to minimize systematic errors and to facilitate relative abundance measurements between greater than two samples.

EXTRACT-CHRO (an auxiliary program of RelEx ***[8]***). EXTRACT-CHRO requires an accurate measure of the ^{15}N enrichment to assess the correct *m/*z ranges to extract the ion chromatograms. The relative abundances between the two chromatograms are determined and outliers removed using the freely available software RelEx. For details on how to obtain RelEx through an institutional software transfer agreement, *see* http://fields.scripps.edu/relex. Specific details of the analysis of quantitative data using stable isotope labeling can be found in another chapter in this volume (*see* Chapter 12).

4. Notes

1. Percent fat content of specialized diet can be adjusted for adult females used for mating. Typical fat content of diet is 5%. For reproduction, the dietary fat content was doubled to 10% by addition of corn oil. This adjustment of fat content was required to attain body fat levels required for fertility, gestation, and lactation.

2. These experiments have been carried out successfully in rats, mice, and ground squirrels.
3. Most organellar fractionations are optimized for fresh tissue. This should be considered when designing quantitative proteomic experiments using the metabolic labeling approach. Fractionation strategy will vary depending on the sample targeted. A crude fractionation strategy is described here for general applications. Refined protocols are available for subcellular fractionations of specific organelles.
4. Perhaps one of the most important aspects of any quantitative analysis using a stable isotope-labeled IS is the accurate knowledge of the enrichment of the respective labeled atoms. This enrichment is essential to (1) estimate the average amino acid mass shift for each amino acid to be used in the database search, and (2) estimate the *m/z* range to extract the natural abundance and ^{15}N-enriched peptide ion chromatograms. An average enrichment of the labeled atoms from the entire protein mixture can be obtained on the amino acid level ***(6)*** or can be estimated for each protein using the peptide isotope distributions using high-resolution MS data ***(15)***.
5. All tandem mass spectra should be searched against the most recent fasta database containing all known and hypothetical protein sequences for the given species. These databases should then be concatenated to a decoy database to facilitate the adjustment of the threshold cutoffs to achieve a preselected false discovery rate.

Acknowledgments

The authors gratefully acknowledge financial support from the National Institute of Health grants K22-AI059076 (C. C. W) and P41-RR011823 (M. J. M).

References

1. Wang, W., Zhou, H., Lin, H., et al. (2003) Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal. Chem.* **75,** 4818–4826.
2. MacCoss, M. J. and Matthews, D. E. (2005) Quantitative MS for proteomics: teaching a new dog old tricks. *Anal. Chem.* **77,** 294A–302A.
3. Conrads, T. P., Alving, K., Veenstra, T. D., et al. (2001) Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and ^{15}N-metabolic labeling. *Anal. Chem.* **73,** 2132–2139.
4. Oda, Y., Huang, K., Cross, F. R., Cowburn, D., and Chait, B. T. (1999) Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. USA* **96,** 6591–6596.
5. Krijgsveld, J., Ketting, R. F., Mahmoudi, T., et al. (2003) Metabolic labeling of C. elegans and D. melanogaster for quantitative proteomics. *Nat. Biotechnol.* **8,** 927–931.

6. Wu, C. C., MacCoss, M. J., Howell, K. E., Matthews, D. E., and Yates, J. R., III (2004) Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal. Chem.* **76,** 4951–4959.
7. Yates III, J. R., McCormack, A. L., Link, A. J., Schieltz, D. M., Eng, J. K., and Hays, L. (1996) Future prospects for the analysis of complex biological systems using micro-column liquid chromatography-electrospray tandem mass spectrometry. *Analyst* **121,** 65R–76R.
8. MacCoss, M. J., Wu, C. C., Liu, H., Sadygov, R., and Yates, J. R., III (2003) A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. *Anal. Chem.* **75,** 6912–6921.
9. Wessel, D. and Flugge, U. I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* **138,** 141–143.
10. Washburn, M. P., Wolters, D., and Yates, J. R., III (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19,** 242–247.
11. MacCoss, M. J., Wu, C. C., and Yates, J. R., III (2002) Probability-based validation of protein identifications using a modified SEQUEST algorithm. *Anal. Chem.* **74,** 5593–5599.
12. Wu, C. C., MacCoss, M. J., Howell, K. E., and Yates, J. R. (2003) A method for the comprehensive proteomic analysis of membrane proteins. *Nat. Biotechnol.* **21,** 532–538.
13. Eng, J. K., McCormack, A. L., and Yates III, J. R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5,** 976–989.
14. Tabb, D. L., McDonald, W. H., and Yates III, J. R. (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **1,** 21–26.
15. MacCoss, M. J., Wu, C. C., Matthews, D. E., and Yates, J. R., 3rd. (2005) Measurement of the isotope enrichment of stable isotope-labeled proteins using high-resolution mass spectra of peptides. *Anal. Chem.* **77,** 7646–7653.

14

Quantitative Proteomic Analysis of Phosphotyrosine-Mediated Cellular Signaling Networks

Yi Zhang, Alejandro Wolf-Yadlin, Forest M. White

Summary

Receptor tyrosine kinases receive extracellular cues, such as ligand binding, and transmit this information to the cell through both autophosphorylation and phosphorylation of tyrosine residues on selected substrates, stimulating a variety of signal transduction pathways. Quantitative features, including intensity, timing, and duration of phosphorylation of particular residues, may play a role in determining cellular response, but experimental data required for analysis of these features have not previously been available. We have recently developed a methodology enabling the simultaneous quantification of tyrosine phosphorylation of specific residues on dozens of key proteins in a time-resolved manner, downstream of receptor tyrosine kinase activation. In this chapter, we present a detailed description of this mass spectrometry-based method, including conditions for cell culture and stimulation, sample preparation for stable isotope labeling and peptide immunoprecipitation, immobilized metal affinity chromatography–liquid chromatography–tandem mass spectrometry analysis of affinity-enriched tyrosine phosphorylated peptides, and analysis of the resulting MS data.

Key Words: Signal transduction; tyrosine phosphorylation; immobilized metal affinity chromatography; liquid chromatography; mass spectrometry.

1. Introduction

Protein phosphorylation-mediated cellular signaling networks regulate information flow within the cell through activation and inhibition of kinases and phosphatases leading to dynamic, reversible protein posttranslational modification. Information contained within these signaling networks, and therefore biological response to an initial stimulus, can vary depending on three key

From: *Methods in Molecular Biology, vol. 359: Quantitative Proteomics by Mass Spectrometry*
Edited by S. Sechi © Humana Press Inc., Totowa, NJ

factors: intensity, timing, and duration of the signal. Quantifying these factors has been limited by another problem, in that it has been challenging to determine the sites of protein posttranslational modification that control how signals are propagating through the system. However, significant advances over the past several years have been made in the field of phosphoproteomics, such that it is now possible to catalog hundreds to thousands of protein phosphorylation sites from a given sample ***(1,2)***. These advances have now made it possible to focus on analyzing the temporal dynamics of protein phosphorylation events following a perturbation to the system. In one example, stable isotope labeling in cell culture (SILAC) ***(3)*** has been combined with phosphopeptide enrichment through two rounds of affinity chromatography. This approach was used to quantify intensity changes for hundreds of phosphorylation sites on yeast proteins following α-factor stimulation ***(4)***. Although this study only looked at a single time point following stimulation, others have used SILAC to quantify changes in total protein tyrosine phosphorylation at selected time points after epidermal growth factor (EGF) stimulation ***(5)***.

In this chapter, we describe a method by which protein tyrosine phosphorylation levels may be quantified on multiple samples in a single analysis with site-specific resolution. In the method, biological samples are lysed by a protein denaturant, proteins are enzymatically digested to peptides and labeled with an amine-specific stable isotope-coded reagent (iTRAQ) ***(6)***. Following labeling and mixing of the samples, tyrosine phosphorylated peptides are immunoprecipitated with a pan-specific anti-phosphotyrosine antibody ***(7)***. After elution from the antibody, phosphorylated peptides are further enriched through immobilized metal affinity chromatography (IMAC) ***(8)***, passed to a reverse-phase column, and analyzed by liquid chromatography–tandem mass spectrometry (LC–MS/MS). We have applied this method to analyze the temporal dynamics of tyrosine phosphorylation following EGF stimulation, generating temporal phosphorylation profiles for 78 tyrosine phosphorylation sites on 58 proteins in a single IMAC-LC–MS/MS analysis ***(9)***. Although the specifics of the method apply directly to the analysis of epidermal growth factor receptor (EGFR) signaling in human mammary epithelial cells, the general approach should be applicable to a broad variety of systems.

2. Materials

2.1. Cell Culture and Lysis

1. DFCI-1 medium: 1:1 (v/v) MEM-α: Ham-F12 (without sodium bicarbonate) supplemented with 2 m*M* L-glutamine, 100 U/mL penicillin, 100 μg/mL streptomycin, 1% (v/v) fetal bovine serum, 1 μg/mL insulin, 10 μg/mL transferring holo form, 50 μ*M* of freshly made ascorbic acid (all from GIBCO, Carlsbad, CA), 1% (v/v) bovine pituitary extract (Pelfreez Biologicals, Rogers, AZ), 10 m*M* HEPES,

2.8 μ*M* hydrocortisone (dexamethasone), 2 n*M* β-estradiol, 15 n*M* sodium selenite, 0.1 m*M* ethanolamine, 0.1 m*M* *o*-phosphoethanolamine, 10 n*M* triiodothyronine, 1 ng/mL cholera toxin (all from Sigma, St. Louis, MO), 19 m*M* sodium bicarbonate (Mallinckrodt, Hazelwood, MO), 12.5 ng/mL EGFR (Peprotech, Rocky Hill, NJ). The pH should be 7.4 in a 5% CO_2 incubator ***(10)***.
2. Serum-free DFCI-1, with the same components as DFCI-1 medium (**step 1**) minus fetal bovine serum, bovine pituitary extract, and EGF. Supplement with 1 mg/mL of bovine serum albumin (Sigma).
3. Phosphate-buffered saline (PBS) (Gibco).
4. 10X trypsin-EDTA solution (Sigma), diluted to 1X in PBS.
5. Lysis buffer consisting of 8 *M* urea and 1 m*M* activated sodium orthovanadate (Sigma). After preparation, the solution should be buffered to pH 7.0 at room temperature.
6. Cell lifters (Corning, Corring, NY).
7. Micro BCA™ Protein Assay Reagent Kit (Pierce, Rockford, IL).

2.2. Protein Digestion and Sample Preparation

1. Dithiothreitol (Sigma).
2. Iodoacetamide (IAc) (Sigma) (light sensitive).
3. Sequencing grade-modified trypsin (Promega, Fitchburg, WI).
4. Trypsin digestion buffer: 100 m*M* ammonium acetate (Sigma), pH 8.9.
5. C18 Sep-Pak Plus cartridges (Waters, Milford, MA).
6. Glacial acetic acid (Mallinckrodt).
7. 0.1% acetic acid solution (Mallinckrodt).
8. 0.1% acetic acid and 25% acetonitrile (ACN) solution (Mallinckrodt).
9. 0.1% acetic acid and 40% ACN solution (Mallinckrodt).
10. 0.1% acetic acid and 90% ACN solution (Mallinckrodt).

2.3. iTRAQ Stable Isotope Labeling and Phosphotyrosine Peptide Immunoprecipitation

1. 0.5 *M* triethylammonium biocarbonate ($N[Et]_3HCO_3$), pH 8.5 (Sigma).
2. Ethanol (Mallinckrodt).
3. iTRAQ reagent multiplex kit (Applied Biosystems, Foster City, CA).
4. IP buffer: 100 m*M* Tris-HC1 100 m*M* NaCl (both from Sigma), and 0.3% NP40, pH 7.4 (Fluka, Buchs SG, Switzerland).
5. 0.5 *M* Tris-HC1 buffer, pH 8.5.
6. Rinse buffer: 100 m*M* Tris-HC1 and 100 m*M* NaCl, pH 7.4.
7. Elution buffer: 100 m*M* glycine (Sigma), pH 2.5.
8. PY99 Immobilized anti-phosphotyrosine antibody, store at 4°C (Santa Cruz Biotechnology Inc., Santa Cruz, CA).

2.4. IMAC and LC–MS/MS

1. IMAC packing material: POROS 20 MC (Applied Biosystems).
2. 100 m*M* EDTA (Sigma), pH 8.5.

3. 100 m*M* $FeCl_3$ (Sigma).
4. 0.1% acetic acid solution (Mallinckrodt).
5. IMAC organic rinse solution: 25% MeCN, 1% acetic acid, and 100 m*M* NaCl.
6. IMAC elution buffer: 250 m*M* Na_2HPO_4, pH 8.0.
7. High-performance liquid chromatography solvent A: H_2O/acetic acid, 99/1 (v/v).
8. High-performance liquid chromatography solvent B: H_2O/MeCN/acetic acid, 29/70/1 (v/v).
9. Fused silica capillary (360 μm outer diameter [O.D.] × 50 μm inner diameter [I.D.]), (360 μm O.D. × 100 μm I.D.), and (360 μm O.D. × 200 μm I.D.) (Polymicro Technologies, Phoenix, AZ).
10. YMC ODS-A 10 μm packing material (Kanematsu Corp., Tokyo, Japan).
11. YMC ODS-AQ 5 μm packing material (Waters).

3. Methods

3.1. Cell Culture and EGF Stimulation

1. Passage 184A1 human mammary epithelial cells at 80–90% confluence with 1X trypsin-EDTA in order to provide new maintenance cultures on 100-mm tissue cultures plates, and experimental cultures on 150-mm tissue culture plates. To generate an experimental plate approaching confluence after 24 h, incubate the cells obtained from 1.5 100-mm tissue cultures plates (~5×10^6 cells) into one 150-mm tissue culture plate.
2. After reaching confluence in the experimental plate, wash the cells with PBS and incubate with serum-free DFCI-1 for 12 h.
3. After 12 h serum starvation, the cultures are washed with PBS and incubated with 25 n*M* EGF diluted in serum-free DFCI-1 for the desired time. For instance, in **Fig. 1**, four plates of cells were incubated with EGF for 0, 5, 10, or 30 min.

3.2. Preparation of Samples for Phosphotyrosine Peptide Immunoprecipitation

1. Following cell culture and agonist (EGF) incubation, the cultures are placed on ice, washed with PBS, and lysed with 3 mL of lysis buffer.
2. Lysates are collected into 15-mL conical tubes and spun down at 1000*g* for 10 min.
3. A 10-μL aliquot from each sample is taken to perform protein concentration assay, using the bicinchoninic acid assay (micro BCA kit) according to the manufacturer's protocol.
4. The sample lysates are reduced by incubation for 1 h with dithiothreitol at 56°C. Reduction is followed by alkylation by incubation for 1 h in the dark (IAc is sensitive to light) at room temperature with 55 m*M* IAc.
5. Following alkylation, the samples are diluted four times by addition of 10 mL of trypsin digestion buffer, and digested overnight with 40 μg of trypsin (~1:100 trypsin:substrate ratio). Digestion is terminated by acidifying the solution to pH 3.0 with 500–1000 μL of glacial acetic acid.

6. Sep-Pak Plus cartridges are preconditioned by sequentially flowing at a rate of 2 mL/min:
   a. 10 mL of 0.1% acetic acid solution.
   b. 10 mL of 0.1% acetic acid and 90% ACN solution.
   c. 10 mL of 0.1% acetic acid solution.
7. Once the Sep-Pak Plus cartridges had been conditioned, the sample lysates are loaded into different cartridges at a rate of 1 mL/min.
8. Both the hydrophilic and hydrophobic fractions of peptides are recovered by sequential elution with 10 mL of 0.1% acetic acid, 25% ACN solution, and 10 mL of 0.1% acetic acid, 40% ACN solution.
9. The recovered fractions are aliquoted into 1-mL aliquots and concentrated to 100–200 µL volume using a vacuum centrifuge prior to overnight lyophilization. Dried aliquots are stored at –80°C until needed.

3.3. iTRAQ Stable Isotope Labeling and Phosphotyrosine Peptide Immunoprecipitation

1. One aliquot of each sample (control, 5, 10, and 30 min) is dissolved with 30 µL of 0.5 *M* triethylammonium bicarbonate ($N[Et]_3HCO_3$), pH 8.5.
2. Thaw a set of iTRAQ labeling reagents (stored at –80°C) to room temperature. Each of the four iTRAQ reagents is dissolved with 70 µL of ethanol, and transferred into the corresponding sample (for instance, in **Fig. 1**, 114 was added to the control sample, 115 was added to the 5-min sample, 116 was added to the 10-min sample, and 117 was added to the 30-min sample).
3. Mix each sample and centrifuge (2800*g* for 5 min). Reaction proceeds at room temperature for 1 h (*see* **Note 1**).
4. After 1 h, each sample is concentrated to approx 30 µL in a vacuum centrifuge.
5. The four samples are combined, reduced to complete dryness in a vacuum centrifuge, and stored at –80°C.
6. 200 µL IP buffer is used to rinse 10 µg of immobilized PY99 for 5 min at 4°C. Spin down the antibody beads at 2800*g*, and remove the supernatant.
7. Dissolve the iTRAQ-labeled sample in 150 µL of IP buffer and 300 µL of water.
8. Adjust the pH of the iTRAQ-labeled sample to pH 7.4 with 0.5 *M* Tris buffer, pH 8.5, and mix with rinsed antibody. Incubate the sample and antibody mixture overnight at 4°C while rotating (*see* **Note 2**).
9. The antibody beads are spun down at 6000 rpm for 5 min. Remove the supernatant and store at –80°C. The antibody beads are then rinsed three times with 400-µL rinse buffer at 4°C for 5 min. The peptides are eluted with 60 µL of 100 m*M* glycine, pH 2.5 at room temperature for 25 min while rotating. The antibody beads are spun down at 2800*g* for 5 min. The eluted peptides are transferred into a new microcentrifuge tube.

3.4. IMAC and LC–MS/MS

1. Clean and condition an IMAC column (a 10-cm long self-packed IMAC (20MC, Applied Biosystems) capillary column (200 µm I.D., 360 µm O.D. was used to generate the data shown in **Fig. 2**) by passing each of the following solutions

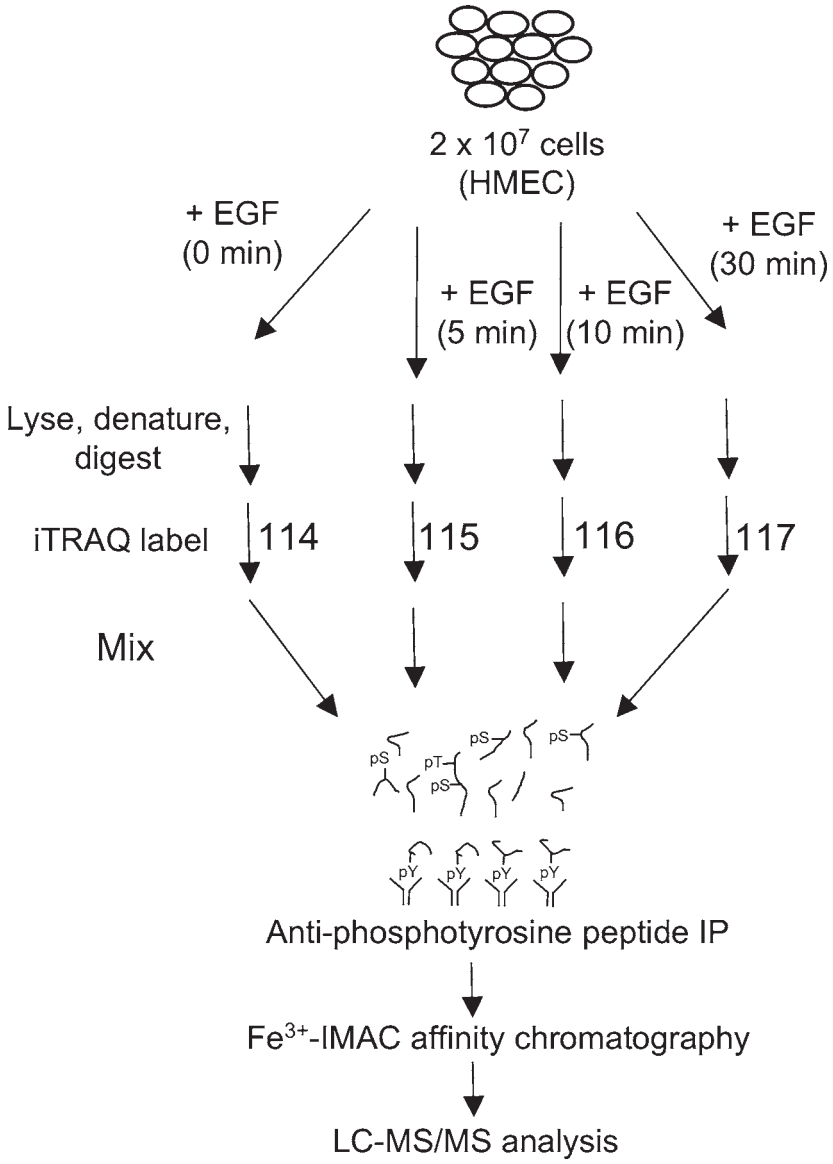


Fig. 1. Schematic representation of mass spectrometry-based approach to analysis of temporal phosphorylation on specific tyrosine residues. Four plates of human mammary epithelial cells were cultured under normal conditions, serum starved for 12 h, and stimulated with 25 n*M* epidermal growth factor for 0, 5, 10, or 30 min. Following cell lysis, proteins were enzymatically digested and the resulting peptide mixture was desalted prior to aliquoting into ten equivalent fractions per stimulation time. For each stimulation time, one aliquot was labeled with one of the four isoforms of the (iTRAQ

through the column at a flow rate of 5 μL/min for 10 min: 100 m*M* EDTA, H_2O, 100 m*M* $FeCl_3$, and 0.1% acetic acid.
2. Peptides eluted from antibody are loaded onto the conditioned IMAC capillary column at 1–2 μL/min.
3. To remove nonspecifically retained peptides, the IMAC column is rinsed with organic rinse solution for 10 min at a flow rate of 5 μL/min.
4. Equilibrate the IMAC column with 0.1% acetic acid for 10 min at a flow rate of 5 μL/min. Connect the IMAC column to a reverse-phase precolumn (we typically use capillary precolumn of dimension 100 μm I.D., 360 μm O.D., packed with 10-cm of YMC ODS-A [10 μm] [Kanematsu Corp.]).
5. Peptides are eluted from the IMAC column onto the reverse-phase capillary precolumn with 50-μL elution buffer. To remove excess phosphate buffer prior to MS analysis, rinse the precolumn with 0.1% acetic acid for 10 min at a flow rate of 1–2 μL/min. After rinsing, connect the precolumn to a reverse-phase analytical column (to generate the data shown in **Fig. 2**, a 10-cm long self-packed C18 [YMC-Waters 5 μm ODS-AQ] analytical capillary column [50 μm I.D., 360 μm O.D.] with an integrated electrospray tip [~1 μm orifice] was used) (*see* **Note 3**).
6. Peptides are eluted from the reverse-phase columns using the following gradient: 10 min from 0 to 15% B, 75 min from 15 to 40% B, and 15 min from 40 to 70% B. Column flow rate should be set to optimum flow rate for the given analytical column depending on column diameter and electrospray emitter tip diameter.
7. As they elute from the column, peptides are directly electrosprayed into a quadrupole time-of-flight mass spectrometer (QSTAR XL Pro, Applied Biosystems). The instrument is operated in information-dependent acquisition mode, in which a full scan mass spectrum is acquired followed by MS/MS spectra of the five most intense peaks with charge state of 2–5. The instrument is set to exclude previously selected peaks for 40 s.
8. To normalize for sample labeling and mixing, a small amount (0.5 μL) of the supernatant from the peptide IP is loaded onto a conditioned IMAC capillary column and analyzed similarly (i.e., repeat **steps 3–7**).

3.5. Data Analysis and Validation

1. MS/MS spectra are extracted (Mascot.dll) and searched against human protein database (NCBI) using ProQuant (Applied Biosystems). Prior to searching, an interrogator database is generated by predigesting the human protein database with trypsin and allowing one missed cleavage and up to six modifications on a single peptide (phosphotyrosine ≤2, phosphoserine ≤1, phosphothreonine ≤1,

---

Fig. 1 *(continued from opposite page)* reagent and then mixed with the labeled aliquots from the other time points. Phosphotyrosine-containing peptides were immunoprecipitated and further enriched by immobilized metal affinity chromatography prior to liquid chromatography–tandem mass spectrometry analysis.

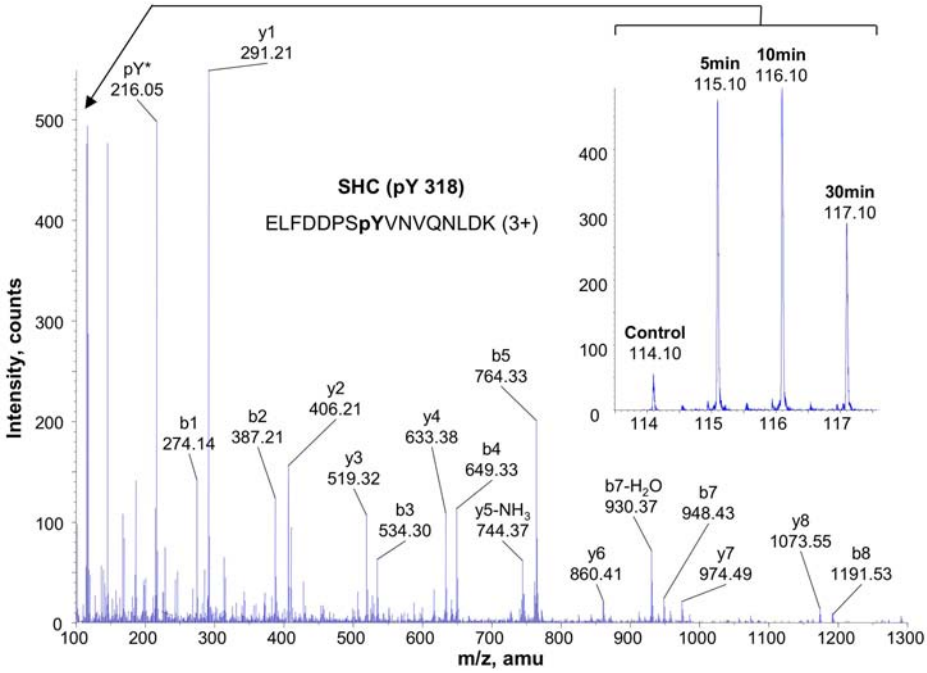


Fig. 2. A representative tandem mass spectrometry (MS/MS) spectrum for one of the tyrosine phosphorylated peptides identified in the analysis of epidermal growth factor-stimulated human mammary epithelial cells. Precursor ion of *m/z* ratio 755.4 with +3 charge state was selected for fragmentation from the full scan mass spectrum. From the resulting MS/MS spectrum, y- and b-type fragment ions enabled peptide identification and phosphorylation site assignment, while peak areas for each of the iTRAQ marker ions (inset, with mass labels and corresponding stimulation time points) enabled quantification of the temporal phosphorylation profiles.

iTRAQ-lysine ≤4, and iTRAQ-tyrosine ≤4). For the database search, mass tolerance is set to 0.15 amu for precursor ions and 0.1 amu for fragment ions.

2. Phosphotyrosine-containing peptides identified from the database search are manually validated by confirming the assignment of y, b, and a-type ions, as well as neutral loss (of H_2O, NH_3, or H_3PO_4 from serine- or threonine-phosphorylated peptides). Peptide sequences are accepted only when all major peaks in the MS/MS spectra can be assigned (*see* **Note 4**).
3. In order to quantify the amount of phosphorylation on a given peptide across the four samples, peak areas for each of the four iTRAQ signature peaks (*m/z*: 114.1, 115.1, 116.1, 117.1) are obtained from the MS/MS spectrum. To account for isotopic overlap, peak areas must be corrected according to the manufacturer's instructions.

4. To correct for sample labeling and mixing, MS/MS data from the IMAC-LC–MS/MS analysis of the supernatant is searched using the same parameters (*see* **step 1**). From the search results, 20 nonphosphorylated peptide hits from abundant proteins are selected and confirmed (*see* **step 2**).
5. In order to normalize the data, one of the iTRAQ marker ions must be selected as the standard, and all quantification is then made relative to the peak area of this ion (for instance, in **Fig. 2** the 5 min sample [the 115 peak] was used as the normalization standard, as this peak had the greatest signal-to-noise ratio and therefore the lowest noise-associated error). The average ratio from the twenty nonphosphorylated peptides reflects the variation in the starting amount of the four samples, and is used to further correct the values of phosphotyrosine-containing peptides.
6. Mean phosphorylation, standard deviation, and *p*-values to estimate statistical significance for differential phosphorylation between the different time points are calculated using Excel. The *p*-values are calculated using a paired, two-tailed student test.
7. To cluster phosphorylation sites with self-similar profiles, a self-organizing map is generated with the Spotfire program. Excel spreadsheets containing the quantification results from the MS data are loaded into Spotfire, the self-organizing map option is selected, and the architecture of the self-organizing map is selected from the options in the window.

4. Notes

1. The recommended peptide amount is 100 μg for each tube of iTRAQ labeling reagent. Incomplete labeling could occur if excess amount of peptides is used. The sample amount should be kept fairly constant for labeling with each of the four iTRAQ isoforms.
2. The quality of phosphotyrosine antibody is crucial for the success of the method. We have observed that many antibodies lose specificity over time even when stored at 4°C. A standard sample should be used to check the quality of the antibody immediately after receipt.
3. Selection of reverse-phase columns will dramatically affect performance, especially with regard to detection limits of the analysis, a crucial point for detecting low-level tyrosine phosphorylated peptides. We typically use custom-made columns in the format originally published by Martin et al. ***(11)***. Electrospray emitter tips are generated on these columns used a Sutter P-2000 laser puller, and typically range from 1 to 2 μ, providing optimal flow rates at 20–50 nL/min.
4. Sample complexity may result in incorrect quantification when using the iTRAQ stable isotope-labeling reagents. This problem manifests when multiple peaks are within the *m/z* isolation window of the selected precursor ion. Under these conditions, multiple peaks are isolated and fragmented simultaneously, with each contributing to the peak areas of the iTRAQ marker ions (*m/z* 114–117). In this case, the relative ratios of the four signature peaks are no longer representative of the relative ratios of the selected precursor ion in the four samples. To avoid such errors in quantification, sample complexity should be decreased through frac-

tionation and the corresponding MS and MS/MS spectrum should be examined for possible contaminants.

Acknowledgments

The authors would like to thank Rebecca Fry for her assistance with Spotfire. This work was supported in part by NIH grants DK42816-015 and DK070172.

References

1. Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., et al. (2002) Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat. Biotechnol.* **20,** 301–305.
2. Beausoleil, S. A., Jedrychowski, M., Schwartz, D., et al. (2004) Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl. Acad. Sci. USA* **101,** 12,130–12,135.
3. Ong, S. E., Blagoev, B., Kratchmarova, I., et al. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell Proteomics* **1,** 376–386.
4. Gruhler, A., Olsen, J. V., Mohammed, S., et al. (2005) Quantitative Phosphoproteomics Applied to the Yeast Pheromone Signaling Pathway. *Mol. Cell Proteomics* **4,** 310–327.
5. Blagoev, B., Ong, S. E., Kratchmarova, I., and Mann, M. (2004) Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* **22,** 1139–1145.
6. Ross, P. L., Huang, Y. N., Marchese, J., et al. (2004) Multiplexed protein quantitation in saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell Proteomics* **3,** 1154–1169.
7. Rush, J., Moritz, A., Lee, K. A., et al. (2004) Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol.* **23,** 94–101.
8. Andersson, L. and Porath, J. (1986) Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. *Anal. Biochem.* **154,** 250–254.
9. Zhang, Y., Wolf-Yadlin, A., Ross, P. L., et al. (2005) Time-Resolved Phosphotyrosine Analysis Reveals Dynamic Modules in EGFR Signaling. *Mol. Cell Proteomics* **4,** 1240–1250.
10. Band, V., Zajchowski, D., Kulesa, V., and Sager, R. (1990) Human papilloma virus DNAs immortalize normal human mammary epithelial cells and reduce their growth factor requirements. *Proc. Natl. Acad. Sci. USA* **87,** 463–467.
11. Martin, S. E., Shabanowitz, J., Hunt, D. F., and Marto, J. A. (2000) Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **72,** 4266–4274.

Index

V

METHODS IN MOLECULAR BIOLOGY™ • 359
SERIES EDITOR: *John M. Walker*

Quantitative Proteomics by Mass Spectrometry

Edited by

Salvatore Sechi
National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD

Quantitative Proteomics by Mass Spectrometry, from the *Methods in Molecular Biology™* series, is a compendium of cutting-edge protocols for quantitative proteomics, and presents the most significant methods used in the field today. The focus on mass spectrometry (MS) is integral, as MS has, and will continue to be, an essential tool in proteomics for studying complex biological systems and human diseases. This volume, written and compiled by leading quantitative proteomic experts, is an indispensable resource in the search for novel biomarkers.

Quantitative Proteomics by Mass Spectrometry presents several innovative MS quantitative procedures, including a variety of methods for introducing isotopic labels and quantifying post-translational modifications. Some of these methods include growing an organism in isotope-enriched media, performing trypsin proteolysis in the presence of ^{18}O-water, reacting protein samples with isotopically labeled reagents, quantifying relative amount of proteins without the use of any isotopic labels. Attention is also given to state-of-the-art techniques for the characterization of the phosphoproteome and tandem MS for detection of inborn errors of metabolism. Specifically, the procedure for determinations of enzymatic activity could be used for large-scale screening of newborns. The protocols in this volume expand both the breadth and depth of readily available methods for quantitative proteomic researchers using MS.

FEATURES

- **Cutting-edge techniques for introducing isotopic labels**
- **Affinity enrichment of phosphopeptides coupled with stable-isotope labeling**
- **Tandem MS for disease detection**
- **Procedures for characterization of the phosphoproteome**
- **Quantification of proteins by MS without isotopic labels**

CONTENTS

Methods in Molecular Biology™ • 359
QUANTITATIVE PROTEOMICS BY MASS SPECTROMETRY
ISBN: 1-58829-571-0 E-ISBN: 1-59259-255-6
ISBN13: 978-1-58829-571-2
ISSN: 1064–3745 humanapress.com

ISBN 1-58829-571-0
9 781588 295712 90000