

Jesus V. Jorrin-Novo
Setsuko Komatsu
Wolfram Weckwerth
Stefanie Wienkoop *Editors*

Plant Proteomics

Methods and Protocols

Second Edition



Humana Press

METHODS IN MOLECULAR BIOLOGY™

Series Editor
John M. Walker
School of Life Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes:
<http://www.springer.com/series/7651>

Plant Proteomics

Methods and Protocols

Second Edition

Edited by

Jesus V. Jorrin-Novo

Department of Biochemistry and Molecular Biology, University of Cordoba, Cordoba, Spain

Setsuko Komatsu

*National Institute of Crop Science, National Agriculture and Food Research Organization,
Tsukuba, Japan*

Wolfram Weckwerth and Stefanie Wienkoop

Department of Ecogenomics and Systems Biology, University of Vienna, Vienna, Austria

 **Humana Press**

Editors

Jesus V. Jorrin-Novo
Department of Biochemistry
and Molecular Biology
University of Cordoba
Cordoba, Spain

Wolfram Weckwerth
Department of Ecogenomics
and Systems Biology
University of Vienna
Vienna, Austria

Setsuko Komatsu
National Institute of Crop Science
National Agriculture and Food Research
Organization
Tsukuba, Japan

Stefanie Wienkoop
Department of Ecogenomics
and Systems Biology
University of Vienna
Vienna, Austria

ISSN 1064-3745 ISSN 1940-6029 (electronic)
ISBN 978-1-62703-630-6 ISBN 978-1-62703-631-3 (eBook)
DOI 10.1007/978-1-62703-631-3
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013945305

© Springer Science+Business Media, LLC 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Humana Press is a brand of Springer
Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This book is the second one published on plant proteomics within this series. The first one, edited by H. Thiellement, M. Zivy, C. Damerval, and V. Mechini, appeared in 2007, 6 years ago, when 2-DE was almost the unique based platform for plant proteomics analysis. Since then, and in a very short period of time, continuous improvements in techniques and protocols for high-throughput proteomics are being made at all workflow stages, from wet (sampling, tissue and cell fractionation, protein extraction, depletion, purification, separation, MS analysis, quantification) to dry lab (experimental design, algorithms for protein identification, bioinformatics tools for data analysis, databases, and repositories). In this book advances made in the field of proteomics in the last 5 years and its application to plant biology and translational research are presented. It will help in exploiting the full potential of proteomics in plant biology research.

It has been written by worldwide recognized scientists, leaders in the field, and includes review papers on workflows, methods, and techniques as well as detailed protocols. It has been organized in nine parts, starting with a general one, in which the field is being reviewed from both a biological and methodological point of view. Part II is devoted to specific methodologies, with emphasis on second-, third-, and fourth-generation techniques (those of gel-free, label- or label-free, imaging, and targeted approaches). Applications of these techniques to the study of experimental model systems, crops, and orphan species (Part III), organs (Part IV), subcellular fractions (Part V), and responses to stresses (Part VI) are included. Parts VII and VIII are devoted to the study and analysis of PTMs, protein interactions, and specific families of proteins. Part IX has been dedicated to the use of proteomics in translational research, with chapters on proteotyping, beverage traceability, and allergens. Finally, the last but not the least, Chapter 53 describes and discusses in detail standards required for a proteomics publication.

Cordoba, Spain
Tsukuba, Japan
Vienna, Austria
Vienna, Austria

Jesus V. Jorrin-Novo
Setsuko Komatsu
Wolfram Weckwerth
Stefanie Wienkoop

Acknowledgement

To my professors and referents Manuel Tena Aldave and Rafael Lopez Valbuena. They not only teached me, but, more important, made me feel free.

Contents

| | |
|-----------------------|-----|
| Preface | v |
| Acknowledgement | vii |
| Contributors | xv |

PART I INTRODUCTORY

| | |
|--|----|
| 1 Plant Proteomics Methods and Protocols..... <i>Jesus V. Jorrin-Novo</i> | 3 |
| 2 From Proteomics to Systems Biology: MAPA, MASS WESTERN, PROMEX, and COVAIN as a User-Oriented Platform..... <i>Wolfram Weckwerth, Stefanie Wienkoop, Wolfgang Hoehnwarter, Volker Egelhofer, and Xiaoliang Sun</i> | 15 |
| 3 Plant Proteomics: From Genome Sequencing to Proteome Databases and Repositories | 29 |
| <i>Katsumi Sakata and Setsuko Komatsu</i> | |
| 4 How to Use 2D Gel Electrophoresis in Plant Proteomics..... <i>Thierry Rabilloud</i> | 43 |
| 5 Standardization of Data Processing and Statistical Analysis in Comparative Plant Proteomics Experiment | 51 |
| <i>Luis Valledor, M. Cristina Romero-Rodríguez, and Jesus V. Jorrin-Novo</i> | |
| 6 The Expanding Universe of Mass Analyzer Configurations for Biological Analysis | 61 |
| <i>Juan J. Calvete</i> | |

PART II EXPERIMENTAL

| | |
|---|-----|
| 7 A Protocol for Protein Extraction from Lipid-Rich Plant Tissues Suitable for Electrophoresis..... <i>Agnieszka Zienkiewicz, Juan David Rejón, Juan de Dios Alché, María Isabel Rodríguez-García, and Antonio Jesús Castro</i> | 85 |
| 8 Making a Protein Extract from Plant Pathogenic Fungi for Gel- and LC-Based Proteomics..... <i>Raquel González Fernández, Inmaculada Redondo, and Jesus V. Jorrin-Novo</i> | 93 |
| 9 Plant Proteomics Methods to Reach Low-Abundance Proteins..... <i>Egisto Boschetti and Pier Giorgio Righetti</i> | 111 |
| 10 Combination of 2DE and LC for Plant Proteomics Analysis..... <i>Sami Irar, Faiçal Brini, Khaled Masmoudi, and Montserrat Pagès</i> | 131 |

| | | |
|----|---|-----|
| 11 | 2DE Analysis of Forest Tree Proteins Using Fluorescent Labels and Multiplexing | 141 |
| | <i>Jenny Renaut, Céline Leclercq, and Sébastien Planchon</i> | |
| 12 | Differential Plant Proteome Analysis by Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) | 155 |
| | <i>Maria J. Martínez-Esteso, Juan Casado-Vela, Susana Sellés-Marchart, María A. Pedreño, and Roque Bru-Martínez</i> | |
| 13 | Global Quantitative Proteomics Using Spectral Counting: An Inexpensive Experimental and Bioinformatics Workflow for Deep Proteome Coverage. | 171 |
| | <i>Tiago S. Balbuena, Diogo Ribeiro Demartini, and Jay J. Thelen</i> | |
| 14 | Combining Chymotrypsin/Trypsin Digestion to Identify Hydrophobic Proteins from Oil Bodies | 185 |
| | <i>Martina Vermachova, Zita Purkrtova, Jiri Santrucek, Pascale Jolivet, Thierry Chardot, and Milan Kodicek</i> | |
| 15 | Mass Western for Absolute Quantification of Target Proteins and Considerations About the Instrument of Choice. | 199 |
| | <i>David Lyon, Wolfram Weckwerth, and Stefanie Wienkoop</i> | |
| 16 | Selected Reaction Monitoring Mass Spectrometry: A Methodology Overview | 209 |
| | <i>H. Alexander Ebhardt</i> | |
| 17 | Mass Spectrometry-Based Imaging of Metabolites and Proteins. | 223 |
| | <i>Manuela Peukert, Michael Becker, Andrea Matros, and Hans-Peter Mock</i> | |
| 18 | Using the Yeast Two-Hybrid System to Identify Protein–Protein Interactions | 241 |
| | <i>Edgar Rodriguez-Negrete, Eduardo R. Bejarano, and Araceli G. Castillo</i> | |
| 19 | Modifications of Wheat Germ Cell-Free System for Functional Proteomics of Plant Membrane Proteins | 259 |
| | <i>Akira Nozawa and Yuzuru Tozawa</i> | |

PART III BIOLOGICAL SYSTEMS

| | | |
|----|--|-----|
| 20 | Arabidopsis Proteomics: A Simple and Standardizable Workflow for Quantitative Proteome Characterization. | 275 |
| | <i>Anja Rödiger, Birgit Agne, Katja Baerenfaller, and Sacha Baginsky</i> | |
| 21 | Analysis of Rice Proteins Using SDS-PAGE Shotgun Proteomics. | 289 |
| | <i>Karlie A. Neilson, Iniga S. George, Samantha J. Emery, Sridevi Muralidharan, Mehdi Mirzaei, and Paul A. Haynes</i> | |
| 22 | <i>Medicago truncatula</i> Proteomics for Systems Biology: Novel Rapid Shotgun LC-MS Approach for Relative Quantification Based on Full-Scan Selective Peptide Extraction (Selpex) | 303 |
| | <i>M. Angeles Castillejo, Christiana Staudinger, Volker Egelhofer, and Stefanie Wienkoop</i> | |
| 23 | Soybean Proteomics | 315 |
| | <i>Zahed Hossain and Setsuko Komatsu</i> | |

- 24 Proteome Analysis of Orphan Plant Species, Fact or Fiction? 333
Sebastien C. Carpentier and Twan America
- 25 An Improved Detergent-Compatible Gel-Fractionation
LC-LTQ-Orbitrap-MS Workflow for Plant and Microbial Proteomics 347
Luis Valledor and Wolfram Weckwerth

PART IV DESCRIPTIVE PROTEOMICS

- 26 Seed Proteomics 361
Ján A. Miernyk
- 27 Back to Osborne. Sequential Protein Extraction and LC-MS Analysis
for the Characterization of the Holm Oak Seed Proteome 379
*M. Cristina Romero-Rodríguez, Ana M. Maldonado-Alconada,
Luis Valledor, and Jesus V. Jorrin-Novo*
- 28 Xylem Sap Proteomics 391
*Thomas Dugé de Bernonville, Cécile Albenne, Matthieu Arlat,
Laurent Hoffmann, Emmanuelle Lauber, and Elisabeth Jamet*
- 29 Suspension-Cultured Plant Cells as a Tool to Analyze
the Extracellular Proteome 407
*Ana B. Sabater-Jara, Lorena Almagro, Sarai Belchi-Navarro,
María J. Martínez-Esteso, Sabry M. Youssef,
Juan Casado-Vela, Juan C. Vera-Urbina, Susana Sellés-Merchart,
Roque Bru-Martínez, and María A. Pedreño*
- 30 Pollen Cultivation and Preparation for Proteomic Studies 435
Heidi Pertl-Obermeyer and Gerhard Obermeyer

PART V SUBCELLULAR PROTEOMICS

- 31 Isolation of the Plant Cytosolic Fraction for Proteomic Analysis 453
*Gonzalo M. Estavillo, Yves Verhertbruggen, Henrik V. Scheller,
Barry J. Pogson, Joshua L. Heazlewood, and Jun Ito*
- 32 Quantitative Proteomic Analysis of Intact Plastids 469
Takeshi Shiraya, Kentaro Kaneko, and Toshiaki Mitsui
- 33 Shotgun Proteomics of Plant Plasma Membrane and Microdomain Proteins
Using Nano-LC-MS/MS 481
*Daisuke Takahashi, Bin Li, Takato Nakayama, Yukio Kawamura,
and Matsuo Uemura*
- 34 Plant Mitochondrial Proteomics 499
*Shaobai Huang, Richard P. Jacoby, A. Harvey Millar,
and Nicolas L. Taylor*
- 35 Separation of the Plant Golgi Apparatus and Endoplasmic Reticulum
by Free-Flow Electrophoresis 527
*Harriet T. Parsons, Susana M. González Fernández-Niño,
and Joshua L. Heazlewood*
- 36 Isolation of Leaf Peroxisomes from Arabidopsis
for Organelle Proteome Analyses 541
Sigrun Reumann and Rajneesh Singh

PART VI COMPARATIVE PROTEOMICS

- 37 Proteomics of Field Samples in Radioactive Chernobyl Area. 555
Katarína Klubicová, Namik M. Rashydov, and Martin Hajduch
- 38 Differential Proteome and Secretome Analysis During Rice–Pathogen Interaction. 563
Yiming Wang, Sang Gon Kim, Jingni Wu, Sun Tae Kim, and Kyu Young Kang
- 39 Protein Extraction and Gel-Based Separation Methods to Analyze Responses to Pathogens in Carnation (*Dianthus caryophyllus* L.). 573
Harold Duban Ardila, Raquel González Fernández, Blanca Ligia Higuera, Inmaculada Redondo, and Sixta Tulia Martínez

PART VII POSTTRANSLATIONAL PROTEOMICS

- 40 StageTip-Based HAMMOC, an Efficient and Inexpensive Phosphopeptide Enrichment Method for Plant Shotgun Phosphoproteomics. 595
Hiroyumi Nakagami
- 41 Simultaneous Identification and Quantification of Nitrosylation Sites by Combination of Biotin Switch and ICAT Labeling. 609
Abasse Fares, Claude Nespoulous, Michel Rossignol, and Jean-Benoit Peltier
- 42 Tandem Metal-Oxide Affinity Chromatography for Enhanced Depth of Phosphoproteome Analysis. 621
Gerold J.M. Beckers, Wolfgang Hoehenwarter, Horst Röhrig, Uwe Conrath, and Wolfram Weckwerth
- 43 N-Glycoprotein Enrichment by Lectin Affinity Chromatography. 633
Eliel Ruiz-May, Carmen Catalá, and Jocelyn K.C. Rose
- 44 Rapid and High-Throughput N-Glycomic Analysis of Plant Glycoproteins. 645
Kentaro Kaneko, Takeshi Shiraya, Toshiaki Mitsui, and Shin-ichiro Nishimura
- 45 Assay for Proteasome-Dependent Protein Degradation and Ubiquitinated Proteins. 655
Takeo Sato, Kaori Sako, and Junji Yamaguchi

PART VIII INTERACTOMICS AND PROTEINOMICS

- 46 Separation of Membrane Protein Complexes by Native LDS-PAGE. 667
Janine Arnold, Alexey Shapiguzov, Geoffrey Fucile, Jean-David Rochaix, Michel Goldschmidt-Clermont, and Lutz Andreas Eichacker
- 47 Identification of Thioredoxin Target Disulfides Using Isotope-Coded Affinity Tags. 677
Per Hägglund, Jakob Bunkenborg, Kenji Maeda, Christine Finnie, and Birte Svensson
- 48 Class III Peroxidases. 687
Sabine Liithje, Claudia-Nicole Meisrimler, David Hopff, Tim Schütze, Jenny Köppe, and Katrin Heino

PART IX TRANSLATIONAL PROTEOMICS

| | | |
|----|--|-----|
| 49 | Proteotyping of Holm Oak (<i>Quercus ilex</i> subsp. <i>ballota</i>) Provenances Through Proteomic Analysis of Acorn Flour | 709 |
| | <i>José Valero Galván, Raquel González Fernández, Luis Valledor, Rafael Ma. Navarro Cerrillo, and Jesus V. Jorrín-Novo</i> | |
| 50 | Immunoproteomics Analysis of Food Allergens | 725 |
| | <i>Rika Nakamura and Reiko Teshima</i> | |
| 51 | Beer and Wort Proteomics | 737 |
| | <i>Takashi Iimure, Makoto Kihara, and Kazuhiko Sato</i> | |
| 52 | Precipitation of Champagne Base Wine Proteins Prior to 2D Electrophoresis | 755 |
| | <i>Clara Cilindre</i> | |
| 53 | The Minimal Information About a Proteomics Experiment (MIAPE) from the Proteomics Standards Initiative | 765 |
| | <i>Salvador Martínez-Bartolomé, Pierre-Alain Binz, and Juan P. Albar</i> | |
| | <i>Index</i> | 781 |

Contributors

- BIRGIT AGNE • *Institut für Biochemie, Abteilung Pflanzenbiochemie, (Biozentrum), Martin-Luther-Universität Halle-Wittenberg, Halle (Saale), Germany*
- JUAN P. ALBAR • *National Center for Biotechnology (CNB-CSIC), Madrid, Spain*
- CÉCILE ALBENNE • *Pôle de Biotechnologie Végétale, LRSV, UMR 5546 Université Paul SabatierToulouse 3/CNRS, Auzerville, Castanet-Tolosan, France*
- JUAN DE DIOS ALCHÉ • *Department of Biochemistry, Cell and Molecular Biology of Plants, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas (CSIC), Granada, Spain*
- LORENA ALMAGRO • *Plant Peroxidases Group, Department of Plant Physiology, Faculty of Biology, University of Murcia, Murcia, Spain*
- TWAN AMERICA • *Plant Research International, Wageningen UR, Wageningen, The Netherlands*
- HAROLD DUBAN ARDILA • *Laboratory Research in Plant Metabolic Activities, Department of Chemistry, National University of Colombia, Bogotá, Colombia*
- MATTHIEU ARLAT • *LIPM, UMR CNRS-INRA 2594-441, Castanet-Tolosan, France*
- JANINE ARNOLD • *Center for Organelle Research (CORE), University of Stavanger, Stavanger, Norway*
- KATJA BAERENFALLER • *Department of Biology, Plant Biotechnology, ETH Zurich, Zürich, Switzerland*
- SACHA BAGINSKY • *Institut für Biochemie, Abteilung Pflanzenbiochemie, Martin-Luther-Universität Halle-Wittenberg, Halle (Saale), Germany*
- TIAGO S. BALBUENA • *Department of Plant Biology, Institute of Biology, University of Campinas, Campinas, SP, Brazil*
- MICHAEL BECKER • *Bruker Daltonik GmbH, Bremen, Germany*
- GEROLD J.M. BECKERS • *Department of Molecular Systems Biology, University of Vienna, Vienna, Austria*
- EDUARDO R. BEJARANO • *Departamento de Biología Celular, Genética y Fisiología, Instituto de Hortofruticultura Subtropical y Mediterránea “La Mayora”, Universidad de Málaga-Consejo Superior de Investigaciones Científicas (IHSM-UMA-CSIC), Universidad de Málaga, Málaga, Spain*
- SARAI BELCHÍ-NAVARRO • *Plant Peroxidases Group, Department of Plant Physiology, Faculty of Biology, University of Murcia, Murcia, Spain*
- PIERRE-ALAIN BINZ • *Swiss-Prot Group, Swiss Institute of Bioinformatics, Geneva, Switzerland*
- EGISTO BOSCHETTI • *EB-Consult, France*
- FAIÇAL BRINI • *Plant Molecular Genetics Laboratory, Centre of Biotechnology of Sfax (CBS), Sfax, Tunisia*
- ROQUE BRU-MARTÍNEZ • *Plant Proteomics and Functional Genomics Group, Department of Agrochemistry and Biochemistry, Faculty of Science, University of Alicante, Alicante, Spain*

- JAKOB BUNKENBORG • *Center of Experimental BioInformatics, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark*
- JUAN J. CALVETE • *Instituto de Biomedicina de Valencia, CSIC, Valencia, Spain*
- SEBASTIEN C. CARPENTIER • *Faculty of Bioscience Engineering, Department of Biosystems, K.U. Leuven, Leuven, Belgium; SYBIOMA: Facility for Systems Biology Based Mass Spectrometry, Leuven, Belgium*
- JUAN CASADO-VELA • *Translational Oncology Unit, Instituto de Investigaciones Biomédicas 'Alberto Sols', Spanish National Research Council (CSIC-UAM), Madrid, Spain*
- M. ANGELES CASTILLEJO • *Department of Molecular Systems Biology, University of Vienna, Vienna, Austria*
- ARACELI G. CASTILLO • *Instituto de Hortofruticultura, Subtropical y Mediterránea, IHSM-UMA-CSIC, Campus Teatinos, Málaga, Spain*
- ANTONIO JESÚS CASTRO • *Department of Biochemistry, Cell and Molecular Biology of Plants, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas (CSIC), Granada, Spain*
- CARMEN CATALÁ • *Boyce Thompson Institute for Plant Research, Department of Plant Biology, Cornell University, Ithaca, NY, USA*
- THIERRY CHARDOT • *Institut Jean-Pierre Bourgin, INRA, UMR1318, Versailles, France*
- CLARA CILINDRE • *Laboratoire d'Oenologie et Chimie Appliquée, Université de Reims Champagne-Ardenne, URVVC UPRES EA 4707, Reims, France; Equipe Effervescence, GSMA CNRS UMR 7331, Université de Reims Champagne-Ardenne, Reims, France*
- UWE CONRATH • *Department of Molecular Systems Biology, University of Vienna, Vienna, Austria*
- THOMAS DUGÉ DE BERNONVILLE • *LIPM, UMR CNRS-INRA 2594-441, Castanet-Tolosan, France*
- DIOGO RIBEIRO DEMARTINI • *Centre de Recherche Public—Gabriel Lippmann, Belvaux, Luxembourg*
- H. ALEXANDER EBHARDT • *Institute of Molecular Systems Biology, Eidgenössische Technische Hochschule (ETH) Zürich, Zürich, Switzerland*
- VOLKER EGELHOFER • *Department of Molecular Systems Biology, University of Vienna, Vienna, Austria*
- LUTZ ANDREAS EICHACKER • *Center for Organelle Research (CORE), University of Stavanger, Stavanger, Norway*
- SAMANTHA J. EMERY • *Department of Chemistry and Biomolecular Sciences, Macquarie University, North Ryde, Australia*
- GONZALO M. ESTAVILLO • *ARC Centre of Excellence in Plant Energy Biology and Research School of Biology, The Australian National University, Canberra, Australia*
- ABASSE FARES • *LPF UR1199 Campus SupAgro/INRA, Montpellier, France*
- CHRISTINE FINNIE • *Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark*
- GEOFFREY FUCILE • *Department of Molecular Biology, University of Geneva, Geneva, Switzerland; Department of Botany and Plant Biology, University of Geneva, Geneva, Switzerland*
- INIGA S. GEORGE • *Department of Chemistry and Biomolecular Sciences, Macquarie University, North Ryde, Australia*
- MICHEL GOLDSCHMIDT-CLERMONT • *Department of Molecular Biology, University of Geneva, Geneva, Switzerland; Department of Botany and Plant Biology, University of Geneva, Geneva, Switzerland*

- RAQUEL GONZÁLEZ FERNÁNDEZ • *Agroforestry and Plant Biochemistry and Proteomics Research Group, Department of Biochemistry and Molecular Biology, University of Córdoba, Agrifood Campus of International Excellence, Córdoba, Spain*
- SUSANA M. GONZÁLEZ FERNÁNDEZ-NIÑO • *Joint BioEnergy Institute and Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
- PER HÄGGLUND • *Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark*
- MARTIN HAJDUCH • *Department of Reproduction and Developmental Biology, Institute of Plant Genetics and Biotechnology, Center of Excellence for White-Green Biotechnology, Slovak Academy of Sciences, Nitra, Slovakia; Institute of Chemistry, Center of Excellence for White-Green Biotechnology, Slovak Academy of Sciences, Nitra, Slovakia*
- PAUL A. HAYNES • *Department of Chemistry and Biomolecular Sciences, Macquarie University, Macquarie Park, Australia*
- JOSHUA L. HEAZLEWOOD • *Joint BioEnergy Institute and Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
- KATRIN HEINO • *Biocentre Klein Flottbek, University of Hamburg, Hamburg, Germany*
- BLANCA LIGIA HIGUERA • *Laboratory Research in Plant Metabolic Activities, Department of Chemistry, National University of Colombia, Bogotá, Colombia*
- WOLFGANG HOEHENWARTER • *Department of Molecular Systems Biology, University of Vienna, Vienna, Austria*
- LAURENT HOFFMANN • *Pôle de Biotechnologie Végétale, LRSV, UMR 5546 Université Paul SabatierToulouse 3/CNRS, Auzerville, Castanet-Tolosan, France*
- DAVID HOPFF • *Biocentre Klein Flottbek, University of Hamburg, Hamburg, Germany*
- ZAHED HOSSAIN • *Department of Botany, West Bengal State University, Kolkata, West Bengal, India*
- SHAOBAI HUANG • *ARC Centre of Excellence in Plant Energy Biology and Centre for Comparative Analysis of Biomolecular Networks (CABiN), The University of Western Australia, Crawley, WA, Australia*
- TAKASHI IIMURE • *Bioresources Research and Development Department, Sapporo Breweries Ltd., Ota, Gunma, Japan*
- SAMI IRAR • *Centre de Recerca en Agrigenomica (CRAG), Campus UAB, Edificio CRAG, Bellaterra, Cerdanyola del Vallès, Barcelona, Spain*
- JUN ITO • *Joint BioEnergy Institute and Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
- RICHARD P. JACOBY • *ARC Centre of Excellence in Plant Energy Biology and Centre for Comparative Analysis of Biomolecular Networks (CABiN), The University of Western Australia, Crawley, WA, Australia*
- ELISABETH JAMET • *Pôle de Biotechnologie Végétale, LRSV, UMR 5546 Université Paul SabatierToulouse 3/CNRS, Auzerville, Castanet-Tolosan, France*
- PASCALE JOLIVET • *Institut Jean-Pierre Bourgin, UMR1318, INRA, Versailles, France*
- JESUS V. JORRIN-NOVO • *Agroforestry and Plant Biochemistry and Proteomics Research Group, Department of Biochemistry and Molecular Biology, University of Córdoba, Agrifood Campus of International Excellence, ceiA3, Córdoba, Spain*
- KENTARO KANEKO • *Department of Applied Biological Chemistry, Faculty of Agriculture, Niigata University, Niigata, Japan*
- KYU YOUNG KANG • *Division of Applied Life Science (BK21 Program), Plant Molecular Biology and Biotechnology Research Center, Gyeongsang National University, Jinju, South Korea*

- YUKIO KAWAMURA • *Cryobiofrontier Research Center, Faculty of Agriculture, Iwate University, Morioka, Iwate, Japan*
- MAKOTO KIHARA • *Bioresources Research and Development Department, Sapporo Breweries Ltd., Ota, Gunma, Japan*
- SANG GON KIM • *Plant Molecular Biology and Biotechnology Research Center, Gyeongsang National University, Jinju, South Korea*
- SUN TAE KIM • *Department of Plant Bioscience, Pusan National University, Miryang, South Korea*
- KATARÍNA KLUBICOVÁ • *Department of Reproduction and Developmental Biology, Institute of Plant Genetics and Biotechnology, Slovak Academy of Sciences, Nitra, Slovakia*
- MILAN KODICEK • *Department of Biochemistry and Microbiology, Institute of Chemical Technology Prague, Prague, Czech Republic*
- SETSUKO KOMATSU • *National Institute of Crop Science, National Agriculture and Food Research Organization, Tsukuba, Japan*
- JENNY KÖPPE • *Biocentre Klein Flottbek, University of Hamburg, Hamburg, Germany*
- EMMANUELLE LAUBER • *LRSV, UMR 5546 UPS/CNRS, Castanet-Tolosan, France*
- CÉLINE LECLERCQ • *Department of Environment and Agrobiotechnologies (EVA), Proteomics Platform, Centre de Recherche Public–Gabriel Lippmann, Belvaux, Luxembourg*
- BIN LI • *United Graduate School of Agricultural Sciences, Iwate University, Morioka, Iwate, Japan*
- SABINE LÜTHJE • *Biocentre Klein Flottbek, University of Hamburg, Hamburg, Germany*
- DAVID LYON • *Department of Molecular Systems Biology, University of Vienna, Vienna, Austria*
- KENJI MAEDA • *Enzyme and Protein Chemistry, Department of Systems Biology, Søltofts Plads, Technical University of Denmark, Kongens Lyngby, Denmark*
- ANA M. Maldonado-Alconada • *Agroforestry and Plant Biochemistry and Proteomics Research Group, Department of Biochemistry and Molecular Biology, University of Cordoba, Cordoba, Spain*
- SIXTA TULIA MARTÍNEZ • *Laboratory Research in Plant Metabolic Activities, Department of Chemistry, National University of Colombia, Bogotá, Colombia*
- SALVADOR MARTÍNEZ-BARTOLOMÉ • *National Center for Biotechnology (CNB-CSIC), Madrid, Spain*
- MARÍA J. MARTÍNEZ-ESTESO • *Plant Proteomics and Functional Genomics Group, Department of Agrochemistry and Biochemistry, Faculty of Science, University of Alicante, Alicante, Spain*
- KHALED MASMOUDI • *Plant Molecular Genetics Laboratory, Centre of Biotechnology of Sfax (CBS), Sfax, Tunisia*
- ANDREA MATROS • *Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany*
- CLAUDIA-NICOLE MEISRIMLER • *Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany*
- JÁN A. MIERNYK • *USDA, Agricultural Research Service, Plant Genetics Research Unit, Department of Biochemistry, Interdisciplinary Plant Group, University of Missouri, Columbia, MO, USA*
- A. HARVEY MILLAR • *ARC Centre of Excellence in Plant Energy Biology and Centre for Comparative Analysis of Biomolecular Networks (CABiN), The University of Western Australia, Crawley, WA, Australia*
- MEHDI MIRZAEI • *Department of Chemistry and Biomolecular Sciences, Macquarie University, North Ryde, NSW, Australia*

TOSHIAKI MITSUI • *Department of Applied Biological Chemistry, Faculty of Agriculture, Niigata University, Niigata, Japan*

HANS-PETER MOCK • *Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany*

SRIDEVI MURALIDHARAN • *Department of Chemistry and Biomolecular Sciences, Macquarie University, North Ryde, NSW, Australia*

HIROFUMI NAKAGAMI • *RIKEN Center for Sustainable Resource Science, Yokohama, Japan*

RIKA NAKAMURA • *National Institute of Health Sciences (NIHS), Setagaya-ku, Tokyo, Japan*

TAKATO NAKAYAMA • *Cryobiofrontier Research Center, Faculty of Agriculture, Iwate University, Morioka, Iwate, Japan*

RAFAEL MA. NAVARRO CERRILLO • *Department of Forestry Engineering, ETSIAM, University of Córdoba, Agrifood Campus of International Excellence, ceiA3, Córdoba, Spain*

KARLIE A. NEILSON • *Department of Chemistry and Biomolecular Sciences, Macquarie University, North Ryde, NSW, Australia*

CLAUDE NESPOULOUS • *LPF UR1199 Campus SupAgro/INRA, Montpellier, France*

SHIN-ICHIRO NISHIMURA • *Graduate School of Science, Frontier Research Center for Post-genomic Science and Technology, Hokkaido University, Sapporo, Japan*

AKIRA NOZAWA • *Cell-Free Science and Technology Research Center, Ehime University, Matsuyama, Ehime, Japan*

GERHARD OBERMEYER • *Molecular Plant Biophysics and Biochemistry, Department of Molecular Biology, University of Salzburg, Salzburg, Austria*

MONTSERRAT PAGÈS • *Centre de Recerca en Agrigenomica (CRAG), Campus UAB, Edifici CRAG, Bellaterra, Cerdanyola del Vallès, Barcelona, Spain*

HARRIET T. PARSONS • *Joint BioEnergy Institute and Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*

MARÍA A. PEDREÑO • *Plant Peroxidases Group, Department of Plant Physiology, Faculty of Biology, University of Murcia, Murcia, Spain*

JEAN-BENOÎT PELTIER • *LPF UR1199 Campus SupAgro/INRA, Montpellier, France*

HEIDI PERTL-OBERMEYER • *Molecular Plant Biophysics and Biochemistry, Department of Molecular Biology, University of Salzburg, Salzburg, Austria*

MANUELA PEUKERT • *Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany*

SÉBASTIEN PLANCHON • *Proteomics Platform, Department of Environment and Agrobiotechnologies (EVA), Centre de Recherche Public—Gabriel Lippmann, Belvaux, Luxembourg*

BARRY J. POGSON • *Centre de Recherche Public—Gabriel Lippmann, Belvaux, Luxembourg*

ZITA PURKRTOVÁ • *Department of Biochemistry and Microbiology, Institute of Chemical Technology Prague, Prague, Czech Republic*

THIERRY RABILLOUD • *Chemistry and Biology of Metals, CNRS UMR 5249, Grenoble, France; Université Joseph Fourier, Grenoble, France; iRTSV/LCBM, CEA Grenoble, Grenoble, France*

NAMIK M. RASHYDOV • *Department of Biophysics and Radiobiology, Institute of Cell Biology and Genetic Engineering, Kyiv, Ukraine*

INMACULADA REDONDO • *Agroforestry and Plant Biochemistry and Proteomics Research Group, Department of Biochemistry and Molecular Biology, University of Córdoba, Agrifood Campus of International Excellence, ceiA3, Córdoba, Spain*

JUAN DAVID REJÓN • *Department of Biochemistry, Cell and Molecular Biology of Plants, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas (CSIC), Granada, Spain*

JENNY RENAUT • *Department of Environment and Agrobiotechnologies (EVA), Proteomics Platform, Centre de Recherche Public—Gabriel Lippmann, Belvaux, Luxembourg*

SIGRUN REUMANN • *Centre for Organelle Research, Faculty of Science and Technology, University of Stavanger, Stavanger, Norway*

PIER GIORGIO RIGHETTI • *Miles Gloriosus Academy, Milano, Italy*

JEAN-DAVID ROCHAIX • *Department of Molecular Biology, University of Geneva, Geneva, Switzerland; Department of Botany and Plant Biology, University of Geneva, Geneva, Switzerland*

ANJA RÖDIGER • *Institut für Biochemie, Abteilung Pflanzenbiochemie, Martin-Luther-Universität Halle-Wittenberg, Halle (Saale), Germany*

MARÍA ISABEL RODRÍGUEZ-GARCÍA • *Department of Biochemistry, Cell and Molecular Biology of Plants, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas (CSIC), Granada, Spain*

EDGAR RODRÍGUEZ-NEGRETE • *Departamento de Biología Celular, Genética y Fisiología, Instituto de Hortofruticultura Subtropical y Mediterránea “La Mayora”, Universidad de Málaga—Consejo Superior de Investigaciones Científicas (IHSU-UMA-CSIC), Universidad de Málaga, Málaga, Spain*

HORST RÖHRIG • *Department of Molecular Systems Biology, University of Vienna, Vienna, Austria*

M. CRISTINA ROMERO-RODRÍGUEZ • *Agroforestry and Plant Biochemistry and Proteomics Research Group, Department of Biochemistry and Molecular Biology, University of Córdoba, Córdoba, Spain; Centro Multidisciplinario de Investigaciones Tecnológicas, Dirección General de Investigación Científica y Tecnológica, Universidad Nacional de Asunción, San Lorenzo, Paraguay*

JOCELYN K.C. ROSE • *Department of Plant Biology, Cornell University, Ithaca, NY, USA*

MICHEL ROSSIGNOL • *LPF UR1199 Campus SupAgro/INRA, Montpellier, France*

ELIEL RUIZ-MAY • *Department of Plant Biology, Cornell University, Ithaca, NY, USA*

ANA B. SABATER-JARA • *Plant Peroxidases Group, Department of Plant Physiology, Faculty of Biology, University of Murcia, Murcia, Spain*

KATSUMI SAKATA • *Maebashi Institute of Technology, Maebashi, Japan*

KAORI SAKO • *Faculty of Science and Graduate School of Life Science, Hokkaido University, Kita-ku, Sapporo, Japan*

JIRI SANTRUCEK • *Department of Biochemistry and Microbiology, Institute of Chemical Technology Prague, Prague, Czech Republic*

KAZUHIRO SATO • *Barley Germplasm Center, Institute of Plant Science and Resources, Okayama University, Kurashiki, Okayama, Japan*

TAKEO SATO • *Faculty of Science and Graduate School of Life Science, Hokkaido University, Kita-ku, Sapporo, Japan*

HENRIK V. SCHELLER • *Joint BioEnergy Institute and Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*

TIM SCHÜTZE • *Biocentre Klein Flottbek, University of Hamburg, Hamburg, Germany*

SUSANA SELLÉS-MARCHART • *Plant Peroxidases Group, Department of Plant Physiology, Faculty of Biology, University of Murcia, Murcia, Spain*

ALEXEY SHAPIGUZOV • *Department of Molecular Biology, University of Geneva, Geneva, Switzerland; Department of Botany and Plant Biology, University of Geneva, Geneva, Switzerland*

- TAKESHI SHIRAYA • *Niigata Agricultural Research Institute, Nagaoka, Japan*
RAJNEESH SINGHAL • *Centre for Organelle Research, Faculty of Science and Technology, University of Stavanger, Stavanger, Norway*
CHRISTIANA STAUDINGER • *Department of Molecular Systems Biology, University of Vienna, Vienna, Austria*
XIAOLIANG SUN • *Department of Molecular Systems Biology, University of Vienna, Vienna, Austria*
BIRTE SVENSSON • *Enzyme and Protein Chemistry, Department of Systems Biology, Soltofts Plads, Technical University of Denmark, Kongens Lyngby, Denmark*
DAISUKE TAKAHASHI • *United Graduate School of Agricultural Sciences, Iwate University, Morioka, Iwate, Japan*
NICOLAS L. TAYLOR • *ARC Centre of Excellence in Plant Energy Biology and Centre for Comparative Analysis of Biomolecular Networks (CABiN), The University of Western Australia, Crawley, WA, Australia*
REIKO TESHIMA • *National Institute of Health Sciences (NIHS), Setagaya-ku, Tokyo, Japan*
JAY J. THELEN • *Department of Biochemistry and Interdisciplinary Plant Group, Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO, USA*
YUZURU TOZAWA • *Cell-Free Science and Technology Research Center, Ehime University, Matsuyama, Ehime, Japan*
MATSUO UEMURA • *Cryobiofrontier Research Center, Faculty of Agriculture, Iwate University, Morioka, Iwate, Japan*
JOSÉ VALERO GALVÁN • *Department of Chemistry-Biology, Biomedical Sciences Institute, Autonomous University of Ciudad Juárez, Chihuahua, México*
LUIS VALLEDOR • *Department of Molecular Systems Biology, University of Vienna, Vienna, Austria*
JUAN C. VERA-URBINA • *Plant Peroxidases Group, Department of Plant Physiology, Faculty of Biology, University of Murcia, Murcia, Spain*
YVES VERHERTBRUGGEN • *Joint BioEnergy Institute and Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
MARTINA VERMACHOVA • *Department of Biochemistry and Microbiology, Institute of Chemical Technology Prague, Prague, Czech Republic*
YIMING WANG • *Plant Molecular Biology and Biotechnology Research Center, Gyeongsang National University, Jinju, South Korea*
WOLFRAM WECKWERTH • *Department of Ecogenomics and Systems Biology, University of Vienna, Vienna, Austria*
STEFANIE WIENKOOP • *Department of Ecogenomics and Systems Biology, University of Vienna, Vienna, Austria*
JINGNI WU • *Division of Applied Life Science (BK21 Program), Gyeongsang National University, Jinju, South Korea*
JUNJI YAMAGUCHI • *Faculty of Science and Graduate School of Life Science, Hokkaido University, Sapporo, Japan*
SABRY M. YOUSSEF • *Plant Peroxidases Group, Department of Plant Physiology, Faculty of Biology, University of Murcia, Murcia, Spain*
AGNIESZKA ZIENKIEWICZ • *Department of Biochemistry, Cell and Molecular Biology of Plants, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas (CSIC), Granada, Spain; Department of Physiology and Molecular Biology of Plants, Nicolaus Copernicus University, Toruń, Poland*

Part I

Introductory

Chapter 1

Plant Proteomics Methods and Protocols

Jesus V. Jorrin-Novo

Abstract

In this first, introductory chapter, it is intended to summarize from a methodological point of view the state of the art in plant proteomics, focusing on mass spectrometry-based strategies. Thus, this chapter is mainly directed at beginners or at those trying to get into the field, rather than at those with real experience or a long trajectory in plant proteomics research. The different alternative workflows, methods, techniques, and protocols from the experimental design to the data analysis will be briefly commented, with cross references to previous monographs and reviews, as well as to the rest of the book chapters. The difficulty of working with proteins, together with the power, limitations, and challenges of the approach will also be briefly discussed.

Proteins, as molecular entities, and the cell proteome, as a whole, are much more complex than what we thought in the past and can be studied in a single experiment. Because of that, fractionation and complementary strategies are required for its study. The MS analysis of complex samples may result in up to 100,000-peptide spectra that cannot be easily analyzed with standard procedures. Therefore, proteomics, more than other -omics, needs a dry lab, time, and an effort in data mining.

As main conclusion, it can be stated that proteomics is in its beginnings. It is starting to make important contributions to a proper gene annotation, identification, and characterization of gene products or protein species and to the knowledge of living organisms, having also an enormous application potential to translational research. However, and despite its great potential, and as in any other experimental approach, it is far from being a Pandora's Box. In the case of plant research, the full potential of proteomics is quite far from being totally exploited, and second-, third-, and fourth-generation proteomics techniques are still of very limited use. Most of the plant proteomics papers so far published belong to the descriptive, subcellular, and comparative proteomics subgroup, mainly using a few experimental model systems—those whose genome has been sequenced—and being from a biological point of view quite descriptive and speculative. From now on we should put more emphasis on the study of posttranslational proteomics and interactomics, and move to targeted, hypothesis-driven approaches. Furthermore, and even more important, we should move to data validation through other -omics or classical biochemical strategies, in an attempt to get a deeper, real, and more accurate view and understanding of cell biology. In the modern Systems Biology concept, proteomics must be considered as a part of a global, multidisciplinary approach. Making biological sense of a proteomics experiment requires a proper experimental design, data validation, interpretation, and publication policy.

Key words Plant proteomics, Proteomics workflows, Descriptive proteomics, Comparative proteomics, Plant systems biology

1 Introduction

The term “proteome” was coined by M. Wilkins and first used during the Siena meeting, back in 1994 [1]. It was an adaptation of the term genome, and simply referred to the PROTEin complement of a genOME. Originally considered, and still being, as just an experimental approach, it could be—at least I do—also considered as a scientific discipline (the classical protein biochemistry in which mass spectrometry and bioinformatics have been incorporated) with focus on the proteome. Thus, the proteome should be understood as the total set of protein species or gene products present in a biological unit (ecosystem, population, organism, organ, tissue, cell, or organelle) at a specific developmental stage and under determined external biotic and abiotic conditions. By using a proteomics approach we aim to know “how,” “where,” “when,” and “what for” are the several hundred thousand of individual protein species produced in a cell, how they do interact with each other and with other molecules to construct the cellular building, how they do work in order to fit in with programmed growth and development, and to interact with their biotic and abiotic environments. Answering all these questions is firstly the objective of Proteomics and secondly of Systems Biology (see Chapter 2).

After the completion of the genome sequence of over 1,000 organisms (4,069, 15,581, and 1,879, complete, incomplete, and targeted projects, respectively; GOLD database, <http://www.genomesonline.org/cgi-bin/GOLD/index.cgi>, January 2013), proteomics is or will become a research priority in any biological project. As proof of that, mass spectrometry-based proteomics, and specifically that of targeted proteomics (called single or multiple reaction monitoring) [3], has been chosen by the journal *Nature Methods* as the Method of the Year 2012. The relevance of proteomics nowadays is supported by a number of international initiatives recently appeared, either general or focused on plants. We should mention, among others, HUPO-iMOP (<http://www.hupo.org/-/http://www.hupo.org/research/imop/->), EuPA-, COST-FA0603 (http://www.costfa0603.org/index.php?option=com_content&view=category&layout=blog&id=1&Itemid=194), INPPO (<http://www.inppo.com/>), and AOAPO (<http://aoapo.org/>). In the last 5 years, plant proteomics special issues have been published in plant sciences and proteomics journals, including *Journal of Proteomics*, *Proteomics*, *Journal of Proteome Research*, *Phytochemistry*, *Journal of Experimental Botany*, and some others. As a very representative example, I must mention the special issue devoted to Plant Proteomics in Europe—COST action [2]. Within the Humana Press Methods in Molecular Biology collection, quite a number of monographs on proteomics have been published, with

the number 355—appearing in 2007—devoted to plants (Plant Proteomics. Methods and Protocols, edited by H. Thiellement, M. Zivi, C. Damerval, and V. Mechlin). The great advances made in the last 5 years within the field justified the reedition of this book in which last advancements, dispersed in the original literature, have been compiled and organized in just a single book.

Proteomics can be thought as a holistic or as a targeted approach, depending on the initial hypothesis, firstly from the most general (proteins are responsible for the phenotypic differences among biological samples) and secondly to the most specific ones (specific proteins or gene products, or groups of proteins according to their physicochemical or biological properties that are involved in the biological phenomenon studied). At the methodological level, proteomics research was originally based on 2-DE protein separation coupled to MS analysis of spots (first generation), moving then to LC-based shotgun strategies (second generation), and later to quantitation approaches including label and label-free variants (third generation). Selected or multiple reaction monitoring (SRM, MRM) hypothesis-driven approaches constitute the last, fourth generation. Innovation in proteomics, implicating a change in the paradigm research within the field, allows researchers to apply the scientific method [3].

Proteomics should be considered as a young discipline, so it is not surprising the daily appearance of some novelty. Advances in proteomics have been made possible due to continuous improvements in protein extraction, purification and separation, mass spectrometry analysis and equipment, developments of bioinformatics tools and algorithms for data analysis, protein identification, quantification and characterization. It resulted in the creation and establishment of quite a number of databases and repositories where information on the proteome of a number of plant species is stored and updated (*see Chapter 3*). The advances in MS, without any doubt, have mainly led to the big progress of the field (*see Chapter 6*). As summarized by JJ Thelen and JA Miernyk [4]

recent improvements in sensitivity, mass accuracy, and fragmentation have led to achievements previously only dreamed of, including whole-proteome identification, and quantification and extensive mapping of specific PTMs (posttranslational modifications); with such capabilities at present, one might conclude that proteomics has already reached its zenith; however, ‘capability’ indicates that the envisioned goals have not yet been achieved.

Proteomics, depending on the research objectives, can be divided in subareas, being the more relevant: descriptive (*see Chapters 26–30*), subcellular (*see Chapters 31–36*), comparative (*see Chapters 37–39*), posttranslational (*see Chapters 40–45*), interactomics (*see Chapters 46 and 47*), proteinomics (*see Chapter 48*), and translational (*see Chapters 49–52*). Most works on plant

proteomics so far published belong to the three first categories, being carried out with model experimental systems whose genome has been sequenced, and that is well annotated or with a good number of genomic or ESTs sequences available, with *Arabidopsis* (see Chapter 20, 31, 36, 41), rice (see Chapter 21, 32, 38), *Medicago truncatula* (see Chapter 22), and soybean (see Chapters 23 and 37) being in the top ranking group. In this regard, proteomics of orphan and unsequenced organisms remains one of the main challenges in proteomics (see Chapters 24 and 39). The interest in the proteome analysis of the unicellular plant algae has just emerged, as they offer the possibility of using single-celled photosynthetic eukaryotes to address biological questions, together with their potential as a source of biofuels [5]. Proteomics as an analytical tool can be used for practical purposes, in areas such as food traceability, substantial equivalence in transgenic, allergens and proteotyping, as the most representative examples (see Chapters 49–52).

During the 1990s and early 2000s, proteomics generated a great expectation. In this sense, B. Marte stated in the March 13th, 2003, *Nature* insight issue on Proteomics [6],

We are just beginning to appreciate the power and limitations of the genomics revolution, yet hard on its heels proteomics promises an even more radical transformation of biological research.

By 2013, 10 years later, we should realize and admit that the reality is quite far from the original expectations. The results gained over the last 15 years have shown that the dynamism, variability, and behavior of proteins are more complex than what was thought. This is concluded considering the number of protein species per gene as a result of alternative splicing, reading frame, and posttranslational modifications, trafficking, turnover, and interactions (protein complexes, rather than individual proteins, are the functional units of the biological machines) [7]. Proteins, if compared to nucleic acids, are molecular entities of the greatest difficulty to work with. Indeed, they are much more diverse in physicochemical properties, greater in number, have a higher dynamic range and, if not enough, there is no PCR for them. As proteins are the major functional component of cells, knowledge of their cellular localization and turnover is crucial to gain an understanding of the biology of multicellular organisms [8]. Because of the reasons mentioned above, it is almost impossible to capture an important fraction of the whole proteome in just a single experiment. For this reason, it is absolutely necessary to fractionate the whole proteome at the subcellular or protein level, in order to study subcellular proteomes or specific groups of physicochemical or functional proteins. It is clear that there is no universal protocol or recipe in proteomics, being necessary to optimize the methods to the experimental system and objectives of the research. Each specific protocol and its variants lead us to a specific fraction of the star filament which is the complete proteome.

Proteomics is making important contributions to the biological knowledge of plants, despite its difficulties and limitations, as it has an enormous applicability to translational research [9]. Moreover, although it is usually not considered, plant systems may represent an invaluable tool for our efforts to understand molecular mechanisms in human diseases. Although *Arabidopsis* and humans diverged 1.6 billion years ago, recent studies have demonstrated a remarkable conservation of protein function and cellular processes between these two distant species [10].

On the long and dark way between genotype and phenotype, proteins are closest to the last, so the information provided goes beyond the possibilities of genomic studies. In addition, there are aspects of cell research that can only be approached by proteomics, such as posttranslational modifications and protein interactions. In these two areas, proteomics is on the frontier of knowledge, and here is where proteomics can make important contributions. Unfortunately, works within this category on plant proteomics are more the exception than the rule. It can be stated that up to now most of the proteomics works published is descriptive and therefore speculative from a biological perspective. We can only get a deeper, realistic view of the cell molecular biology by validating our proteomics data with other techniques together with complementary classical and -omics approaches, in the direction of modern System Biology (*see Chapter 2*). Proteomics will grow in parallel with other areas, nourishing and being nourished by them. The integration of genotyping, pheno/morphotyping and the analysis of the molecular phenotype using metabolomics, proteomics and transcriptomics will reveal a novel understanding of plant biology and its interaction with the environment [11].

This introductory chapter does not aim to deepen the numerous studies on plant biological processes approached by proteomics, from growth and development, to responses to the environment, as it has been periodically reviewed by us and others [12–22]. Instead, and in agreement with the focus of this book, it is intended to summarize from a methodological point of view the state of the art in plant proteomics, focusing on mass spectrometry-based strategies. This chapter is mainly directed at beginners or at those trying to get into the field, rather than at those with real experience or a long trajectory in plant proteomics research.

2 Methodologies and Workflows

In the broadest sense, and from a methodological point of view, proteomics includes techniques utilized in the study of proteins and peptides, from structure to location and function, either *in vivo*, *in vitro* or *in silico* experiments, and whether the protein or peptide samples are extracted from living organisms (intracellular,

extracellular fluids) as products of in vitro synthesis from mRNA or being chemically synthesized.

Proteomics comprises well-established methods of classical protein chemistry/biochemistry used before the term “proteome” was coined, including techniques for protein extraction, purification (gel electrophoresis and liquid chromatography-based methods), labeling, identification (Western blot), and sequencing (EDMAN sequencing). Nonetheless, it is true that these methods are being continuously improved, but also need to be improved and optimized for different experimental systems, as well as for the objectives of the research (descriptive, comparative, PTMs, interactions). Although EDMAN sequencing is still valid, its use has decreased dramatically [23], as it cannot be used in a high throughput way. Western blot can also be useful since it can increase the confident identification of candidate proteins or PTMs [24].

Mass spectrometry is the analytical method that defines proteomics, and differentiates it from classical protein biochemistry, being in the core of this approach. It allows the identification and quantification from hundreds to thousands proteins in a single experiment. Advances in proteomics and its applicability to all areas of the life sciences are continuously driven by the introduction of new mass spectrometric tools with improved performances (*see* Chapter 6) [25]. The ability of the new instrumentation to produce huge volumes of data has emphasized the need for adequate data analysis tools, which are nowadays often considered the main bottleneck for proteomics development [26], together with the availability of well-annotated protein sequence databases or mass spectra atlas. Large volumes of proteomic data have been accumulated in recent years, as with genomic data. These proteomic data have been integrated and organized in databases that enable us to retrieve, leverage and share public data through up-to-date computational technology, such as the latest data management systems and Web-interface techniques (*see* Chapter 3).

Most of the works published on plant proteomics, and presented here in the different chapters, belongs to the group of in vitro analysis of cell extracts or extracellular fluids (*see*, as an example, Chapters 28, 29, 38, 39), with the aim of identifying, quantifying, and characterizing protein species present in a biological unit (cellular compartment, cell, tissue, organ, individual, genotype, population, ecosystem) in specific time-space coordinates (development and growth) and with specific external factors (responses to the environment). The exception to this generality is the new mass spectrometry imaging technique (*see* Chapter 17), where the spatial distribution of metabolites and proteins can be deciphered. In this methodology, tissue sections—instead of a protein/peptide preparation—are subjected to MS, MALDI-TOF analysis. Other in vitro techniques, such as those of the so-called wide genomics approaches, are less represented in the plant proteomics literature, and tend to not be considered *sensu strictus*

as proteomics techniques. A good example of this is the “two-hybrid system,” directed at studying gene product interactions (*see* Chapter 18). Within this so-called wide genomics group, we can include other approaches like the analysis of gene tagged proteins [27] and the *in silico* analysis of the gene sequence [28, 29]. Finally, structural proteomics, aiming at generating protein 3D structures after *in silico*, crystallographic, or spectroscopic analysis, could also be considered within proteomics. Unfortunately we could not, but also we did not intend it, to cover in detail all of them in this book, so we have focused on MS-based proteomics techniques, leaving out methods of wide genomics or structural proteomics.

The workflow of a standard MS-based proteomics experiment includes all or most of the following steps: experimental design, sampling, tissue/cell or organelle preparation, protein extraction/fractionation/purification, labeling/modification, separation, MS analysis, protein identification, statistical analysis of data, validation of identification, protein inference, quantification, and data analysis, management and storage. The different strategies for a proteomics experiment are the result of different protocol combinations in a specific sequence, and can be named and grouped into different categories.

MS analysis can be performed with either proteins (protein centric) or peptides resulting from its protease (usually trypsin, but may be others if lack of trypsin cleavage sites, *see* Chapter 14) digestion (peptide centric), being them called as top down (proteins) or bottom up (peptides) proteomics, respectively. We left out of this book the top-down strategy, but readers are referred to some publications [30, 31]. In both top and bottom approaches, MS analysis can be carried out with total protein extracts (in top down) or digested peptide products (in bottom up) Alternatively, it can be made after a protein separation step, in most of the cases using electrophoresis (one or two dimensional gel electrophoresis) and, to a lesser extent, conventional liquid chromatography, known as gel-based or LC-based (gel-free) techniques, respectively. The simplest technique in terms of sample manipulation and number of steps is the so called MudPIT, in which peptides derived from total protein extracts digestion are subjected to MS analysis after LC (nano or micro) separation [31]. In terms of proteome coverage and number of proteins identified, a combination of 1-DE, protein digestion of specific bands or group of bands, coupled to on line nLC-ESI-MS, has proved to be the most powerful technique (*see* Chapters 21, 25, 27, 29).

Depending on the quantification strategy used (MS-based quantification), we may refer to relative quantitation based on label (DIGE, ICAT, iTRAQ, SILAC) and label-free (peak area or spectral counting) strategies, being them illustrated by Chapters 11, 12, 32 (label), and 13, 20, 22 (label-free). Some of the labeling techniques, like ICAT, have been used to monitor oxidoreduction of protein thiols processes (*see* Chapter 47). Recently, the absolute

quantification of proteins based on stable isotope-labeled integral standard peptides and liquid chromatography coupled to selective reaction monitoring has been reported (*see Chapter 15*).

Finally, the techniques used can be grouped according to the MS equipment, which results in different combinations of ionizer and analyzers, with MALDI sources usually coupled to TOF analyzers and ESI to quadrupoles, although lately a number of hybrids have been commercialized (*see Chapter 6*).

The most appropriate workflow and protocols to be used depend on the biological system and must be optimized for them (e.g., species, organ, tissue, cell), as well as they depend on the objectives of the research (descriptive, comparative, PTMs, interactions, targeted proteomics) (*see Chapters 23, 34, 38, 39*). Alternative workflows, strategies and protocols are complementary to each other; all of them together contributing to a deeper proteome coverage and biological knowledge (*see Chapters 8, 10, 34*). No specific technique can be considered better than another. Despite there is no single mass analyzer capable of performing all applications required in proteomic research, knowledge of the analytical capabilities and performances of the different mass analyzer configurations available in the market, proteomic services, or in the laboratories of colleagues, is a key element to take full advantage of this powerful tool for biological research (*see Chapter 6*).

I would advise, when approaching proteomics for the first time, to start with the simplest alternative, which tends to be 1-D electrophoresis (proven to be very valuable in the analysis of simple proteomes, *see Chapters 28 and 29*), and then moving to the most complicated, including 2-DE and gel-free. 1-D has provided good results in Holm oak proteotyping and variability studies by analyzing the profile of acorn flour (*see Chapter 49*). In our group, we apply the “funnel strategy,” which means starting with simple techniques for high number of samples, and depending on the results moving to more complicated strategies with reduced number of samples. Opinions and views of classical techniques, like 2-DE, determined as being outdated and poorly efficient, must be discarded (*see Chapters 4 and 11*). In the case of plant proteomics, most of the papers published belong to the descriptive, comparative, or subcellular proteomics categories, using 2-DE-based strategies (*Chapter 11*). For some purposes, like the analysis of protein complexes, electrophoretic-based techniques, such as the Blue Native PAGE or the Clear Native PAGE, have provided excellent results (*see Chapter 46*). Even though LC-based separation techniques of peptides (bottom-up, MudPIT) or proteins (top-down, *see Chapter 10*), second- and third-generation techniques for quantitative proteomics, including both label (DIGE, ICAT, iTRAQ, SILAC) and label-free protocols, are still the exception rather than the rule, this tendency is being changed since the last 2 years (*see Chapters 12, 13, 20, 22, 32, 34*).

The first steps prior to MS analysis (biological system dependent), although not always realized, are key to the success of the experiment, despite the fact we use to consider them as trivial. An adequate and correct experimental design, together with the statistical analysis of the data, is requested. Surprisingly, this tends to be ignored or not properly considered, even in current literature (*see* Chapter 5).

Only when a protein is extracted, solubilized, and visualized, it can be identified and quantified. For this reason, attention must be paid to procedures for protein extraction, especially in the case of recalcitrant proteins (highly hydrophobic, with extreme pI or M_r) or plant tissues rich in nonprotein material, such as phenolics, lipids, salts, and polysaccharides (*see* Chapters 7, 8, 33). Apart from this, it will be easy for most biochemists to understand and perform by themselves these preliminary steps, while mass spectrometry and data analysis (identification, quantification, and characterization) require a real, deep expertise. This, together with the cost of the mass spectrometers and associated bioinformatic packages, justifies the use of proteomics services. It would be impossible, however, to exploit the huge amount of information generated by a mass spectrometer having no knowledge of the experimental system, the protein extract and solution, and of how the last was obtained from the previous.

Despite the technological achievements in proteomics, only a tiny fraction of the cell proteome has been characterized so far, and only for a few biological systems (human, fruit fly, *Arabidopsis*, rice, *Chlamydomonas*). Even for these organisms, the function of quite a number of proteins remains to be investigated. Proteomics techniques have a number of limitations, such as sensitivity, resolution, and speed of data capture. They also face a number of challenges, such as deeper proteome coverage, proteomics of unsequenced “orphan” organisms, top-down proteomics, protein quantification, PTMs or interactomics, among others. Most of these limitations and challenges reflect the difficulty of working with the biological diversity of proteins and their range of physico-chemical properties (*see* Chapter 27). Compared to other biological systems, plants present a number of characteristics that makes difficult the obtention of a good protein extract (*see* Chapter 9). On the other side, plants share with the other systems the high protein dynamic range that makes impossible to detect minor proteins. To face this problem, a number of approaches may be proposed, including subcellular or protein fractionation (*see* Chapters 10, 26, 27, 31–36, 38), protein depletion by using antibodies, or the most recent technique developed: the combinatorial peptide ligand libraries or CPLL, described in detail by its inventor Prof. Righetti in Chapter 9. For specific areas, like subcellular proteomics, difficulties are associated with the isolation of relatively pure samples from plant material free of contamination (*see* Chapters 31–36). In this regard, proteomics has shown us that the static view of proteins

fixed to a subcellular fraction is far from being real, with protein trafficking being one of the issues to be studied in depth. Membrane proteins, and especially microdomains, are quite recalcitrant and elusive to standard protocols (*see Chapter 33*). For PTM analysis, an enrichment step prior to MS analysis tends to be required, depending the success of the experiment on the type of modification (lability of the bound), stoichiometry and behavior of the modified protein species in MS (*see Chapters 40–45*). While the *in vitro* analysis of PTMs is straightforward, the *in vivo* studies and the identification of the specific modified peptides are quite challenging (*see Chapters 41 and 42*).

The chapters have been written by worldwide recognized scientists, leaders in the field, and include reviews and original papers. They have been organized in nine sections, starting with a general one in which the field is being reviewed, both from a biological and methodological perspective. Section II is devoted to specific methodologies, with emphasis on second-, third-, and fourth-generation techniques (those of gel-free, label or label-free, imaging, and targeted approaches). Applications of these techniques to the study of experimental model systems, crops, and orphan species (section III), organs (section IV), subcellular fractions (section V), and responses to stresses (section VI) are also included. Sections VII and VIII are devoted to the study and analysis of PTMs, protein interactions, and specific families of proteins. Section IX has been dedicated to the use of proteomics in translational research, with chapters on proteotyping, beverage traceability, and allergens.

The last, but not least, Chapter 53 has been included due to it deals with the standards required in proteomics publications. It includes standards for proteomics data representation, as well as guidelines that set the minimum information to be included when reporting a proteomics experiment (MIAPE). These standards are also listed in the “instruction to authors” of the journals of the field. As editor and referee I have noticed with surprise the high number of manuscripts submitted from authors that ignore such standards, and even the number of published papers that do not fit in them.

References

1. Bini L, Pallini V, Hochstrasser DF et al (2007) From genome to proteome: back to the future. *Proteomics* 7:1561–1563, Editorial
2. Renaud J, Jorrin-Novo J, Panis B, Svensson B, guest editors (2011) Plant proteomics in Europe. *J Proteomics* 74:1161–1164
3. Picotti P, Bodenmiller B, Aebersold R (2013) Proteomics meets the scientific method. *Nat Methods* 1:24–27
4. Thelen JJ, Miernyk JA (2012) The proteomic future: where mass spectrometry should be taking us. *Biochem J* 444:169–181
5. Uhrig G, Moorhead GB (2013) Plant proteomics: current status and future prospects. *J Proteomics* (On line advanced) doi:10.1016/j.jprot.2013.01.018
6. Marte B (2003) Introduction proteomics. *Nature* 422:191
7. Jungblut PR, Schluter H (2011) Towards the analysis of protein species: an overview about strategies and methods. *Amino Acids* 41:219–222
8. Petricka JJ, Schauer MA, Megraw M et al (2012) The protein expression landscape of

- the *Arabidopsis* root. *Proc Natl Acad Sci U S A* 109:6811–6818
9. Agrawal GK, Pedreschi R, Barkla BJ et al (2012) Translational plant proteomics: a perspective. *J Proteomics* 75:4588–4601
 10. Xu XM, Moller SG (2011) The value of *Arabidopsis* research in understanding human disease states. *Curr Opin Biotechnol* 22:300–307
 11. Weckwerth W (2011) Green systems biology: from single genomes, proteomes and metabolomes to ecosystems research and biotechnology. *J Proteomics* 75:284–305
 12. Cánovas FM, Dumas-Gaudot E, Recorbet G (2004) Plant proteome analysis. *Proteomics* 4:285–298
 13. Rossignol M, Peltier JB, Mock HP (2006) Plant proteome analysis: a 2004–2006 update. *Proteomics* 6:5529–5548
 14. Jorrin JV, Rubiales D, Dumas-Gaudot E (2006) Proteomics: a promising approach to study biotic stresses in legumes. A review. *Euphytica* 147:37–47
 15. Rispa N, Dita MA, González verdejo N et al (2007) Plant resistance to parasitic plants: molecular approaches to an old foe. *New Phytol* 173:703–712
 16. Jorrin Novo J, Maldonado AM, Castillejo MA (2007) Plant proteome analysis: a 2006 update. *Proteomics* 7:2947–2962
 17. Jorrin Novo J, Maldonado AM, Echevarría Zomeño S et al (2009) Plant proteomics update (2007–2008). Second generation proteomic techniques, an appropriate experimental design and data analysis to fulfill MIAPE standards, increase plant proteome coverage and biological knowledge. *J Proteomics* 72:285–314
 18. Whitelegge JP, Komatsu S, Jorrin-Novo JV (2011) Diverse facets of plant proteomics. *Phytochemistry* 72:961–962, Editorial
 19. Abril N, Gion J-M, Kerner R et al (2011) Proteomics research on forest trees, the most recalcitrant and orphan plant species. *Phytochemistry* 72:1219–1242
 20. Cox J, Heeren RMA, James P et al (2011) Facing challenges in proteomics today and in the coming decade: report of roundtable discussions at the 4th EuPA scientific meeting, Portugal, Estoril 2010. *J Proteomics* 75:4–17
 21. Valledor L, Jorrin JV (2011) Back to the basics: maximizing the information obtained by quantitative two dimensional gel electrophoresis analyses by an appropriate experimental design and statistical analyses. *J Proteomics* 74:1–18
 22. González-Fernández R, Jorrin-Novo JV (2012) Contribution of proteomics to the study of plant pathogenic fungi. *J Proteome Res* 11:3–16
 23. Salamanca G, Rodriguez R, Quiralte J et al (2010) Pectin methylesterases of pollen tissue, a major allergen in olive tree. *FEBS J* 277:2729–2739
 24. Liu CW, Hsu Y-K, Cheng Y-H et al (2012) Proteomic analysis of salt-responsive ubiquitin-related proteins in rice roots. *Rapid Commun Mass Spectrom* 26:1649–1660
 25. Bantscheff M, Lemeer S, Savitski MM et al (2012) Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal Bioanal Chem* 404:939–965
 26. Cappadona S, Baker PR, Cutillas PR et al (2012) Current challenges in software solutions for mass spectrometry-based quantitative proteomics. *Amino Acids* 43:1087–1108
 27. Ghaemmaghami S, Huh W-K, Bower K et al (2003) Global analysis of protein expression in yeast. *Nature* 425:737–741
 28. Zhang QC, Petrey D, Deng L et al (2012) Structure-based prediction of protein–protein. *Nature*. doi:[10.1038/nature11503](https://doi.org/10.1038/nature11503)
 29. Kaundal R, Saini R, Zhao PX (2010) Combining machine learning and homology-based approaches to accurately predict subcellular localization in *Arabidopsis*. *Plant Physiol* 154:36–54
 30. Tran JC, Zamdborg L, Ahlf DR (2011) Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature* 480:254–258
 31. Wu C, Tran JC, Zamdborg L (2012) A protease for “middle-down” proteomics. *Nat Methods* 9:822

Chapter 2

From Proteomics to Systems Biology: MAPA, MASS WESTERN, PROMEX, and COVAIN as a User-Oriented Platform

Wolfram Weckwerth, Stefanie Wienkoop, Wolfgang Hoehenwarter, Volker Egelhofer, and Xiaoliang Sun

Abstract

Genome sequencing and systems biology are revolutionizing life sciences. Proteomics emerged as a fundamental technique of this novel research area as it is the basis for gene function analysis and modeling of dynamic protein networks. Here a complete proteomics platform suited for functional genomics and systems biology is presented. The strategy includes MAPA (mass accuracy precursor alignment; <http://www.univie.ac.at/mosys/software.html>) as a rapid exploratory analysis step; MASS WESTERN for targeted proteomics; COVAIN (<http://www.univie.ac.at/mosys/software.html>) for multivariate statistical analysis, data integration, and data mining; and PROMEX (<http://www.univie.ac.at/mosys/databases.html>) as a database module for proteogenomics and proteotypic peptides for targeted analysis. Moreover, the presented platform can also be utilized to integrate metabolomics and transcriptomics data for the analysis of metabolite–protein–transcript correlations and time course analysis using COVAIN. Examples for the integration of MAPA and MASS WESTERN data, proteogenomic and metabolic modeling approaches for functional genomics, phosphoproteomics by integration of MOAC (metal-oxide affinity chromatography) with MAPA, and the integration of metabolomics, transcriptomics, proteomics, and physiological data using this platform are presented. All software and step-by-step tutorials for data processing and data mining can be downloaded from <http://www.univie.ac.at/mosys/software.html>.

Key words MAPA, MOAC, COVAIN, PROMEX, MASS WESTERN

1 Introduction

Proteins are the active part of life, and life is centered on proteins. They perform and control all processes from gene regulation to physiology leading to a balance of a thermodynamically open and susceptible system. Eventually the system collapses and imbalance leads to death. Even this process is controlled by proteins involved in programmed cell death (PCD). As a consequence of the discovery of the chemical-physical principles of life in the middle of the

nineteenth century and the structural elucidation of DNA, technologies for genome sequencing such as next generation sequencing (NGS) have revolutionized biological sciences comparable to the early times of physics in the beginning of the twentieth century. However, genome information still remains static information and molecular processes from transcription to translation to metabolic networks result in the exponential amplification of compounds in a multicellular organism. This complexity is neither measurable nor predictable with technologies nowadays available [1]. Systems biology aims to generate and predict dynamic models starting from the genome sequence. In combination with genome-scale measurements of transcripts, proteins, and metabolites these models can be verified and optimized leading to the iterative knowledge cycle of systems biology.

Genome sequencing and the principal role of proteins as key regulators of life makes proteomics technology to be the backbone of systems biology approaches to reveal the regulatory principles of an organism.

Here, techniques for quantitative proteomics are most decisive [2]. In the following sections we will discuss a platform (Fig. 1) which is suited for high-throughput quantitative exploratory proteomics (MAPA [3]) and targeted proteomics (MASS WESTERN [4]) combined with multivariate statistical data mining (COVAIN [5–7]). A core module of this platform is a mass spectral reference database (PROMEX [8, 9]), on the one hand storing data from the exploratory phase, and on the other providing a resource for targeted proteomics. COVAIN is a data mining toolbox for multivariate statistics and modeling approaches also able to integrate different kinds of molecular data such as proteomics, metabolomics, and transcriptomics data [5–7].

In the following sections these tools are discussed and COVAIN is used to integrate and statistically analyze MAPA and MASS WESTERN data of two different growth conditions of *Chlamydomonas reinhardtii*.

Furthermore, a proteogenomic approach based on proteomics and metabolomics data as well as metabolic modeling for *Chlamydomonas reinhardtii* is presented.

In another section examples for the combination of MAPA with metal-oxide affinity chromatography (MOAC) of phosphoproteins are presented demonstrating the convenient application of the complete workflow to phosphoproteomics and signaling.

2 Strategies for Quantitative Proteomics in the Era of Systems Biology

To capture the dynamics of a biological system, its components need to be identified and quantified. Proteomics is confronted with a task that appears unsolvable. Both in plant proteomics and

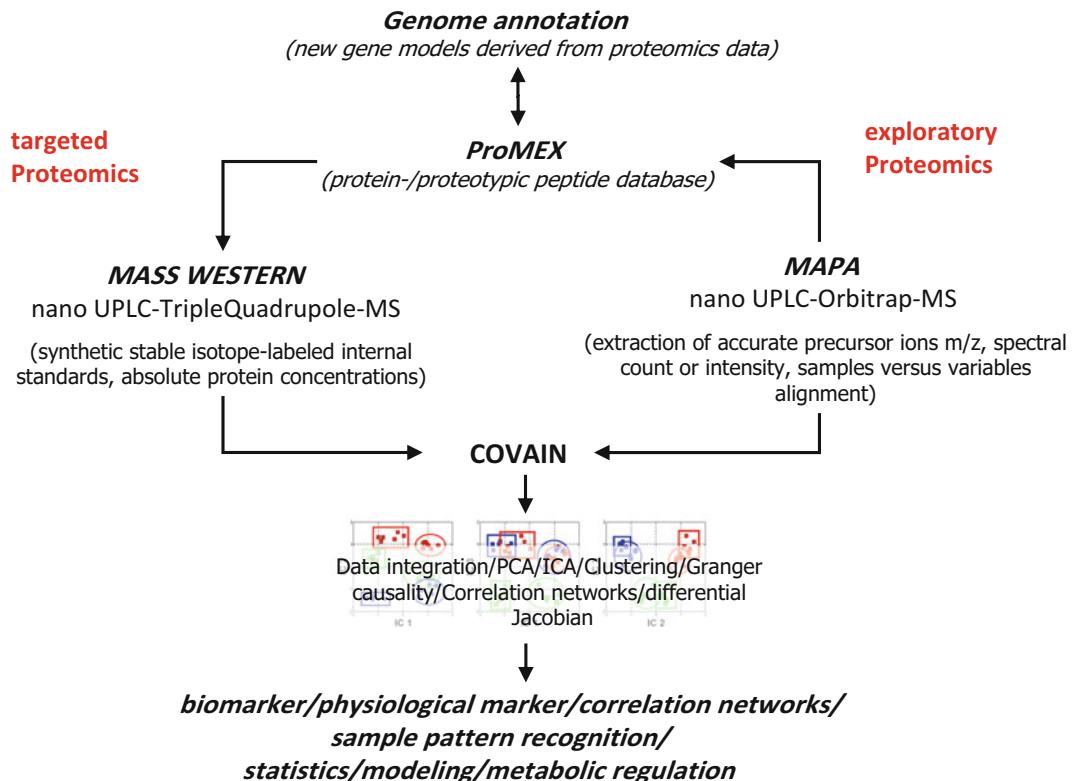


Fig. 1 Overview of a proteomic-toolbox for systems biology. A central proteome/peptide spectral database (ProMEX, www.promexdb.org [8, 9]) serves as a basis for the selection of proteotypic peptides, suitable for the targeted MASS WESTERN analysis [4] of complex proteome samples (see Fig. 3). The MAPA method (MAPA=Mass Accuracy Precursor Alignment [3]) allows for the detection and quantification of all changes of tryptic peptides m/z -ratios in the LC-MS shotgun analyses independent of database search. Quantitative proteomics data are aligned by the ProtMAX algorithm and statistically analyzed using the COVAIN toolbox [5]. Proteomic data enable the suggestion of new gene models not found via computational prediction from the raw genome sequence alone [20]. All these data as well as metadata (experiment, parameters of analysis, and database search) are stored in the database ProMEX. Furthermore, raw spectra from MAPA analysis can be searched against PROMEX [8, 9]

also in human proteomics following numbers are being anticipated: Assuming approximately 20,000–30,000 annotated genes of a genome, after consideration of splice-variances and posttranslational modifications such as phosphorylation or glycosylation, several hundreds or thousands of possible protein species per annotated gene may be reached. It is, however, not assumed that all possible protein isoforms are active at the same time. Nevertheless, existing technologies are confronted with enormous challenges due to high number and dynamic concentration range of all proteins of a single steady state. For gaining a holistic overview of the dynamics of a continuously transient biological system it is important to analyze many different steady states, time series, diverse genotypes, and their phenotypic plasticity. In modern biology, a

simple comparison between states A versus B is no longer adequate to perform functional modeling. Instead, we need high sample throughput technologies able to identify and quantify as many as possible proteins of the system. Thus, there is a need for continuous method development and improvement.

Nevertheless, there are obvious strategies to increase sample throughput as well as number of protein detections. One strategy is based on the unbiased or untargeted identification and quantification of as many proteins possible. In the following this method is called exploratory protein analysis and is based on recent developments in proteomics technology called shotgun proteomics. This is a high sample throughput method in which complex protein samples from tissue or cells are directly proteolytically cleaved into small peptides and subsequently analyzed via liquid chromatography coupled to mass spectrometry. These peptides can be directly quantified without identification using a novel approach called MAPA (mass accuracy precursor alignment) [3]. This allows the detection of all peptides which are quantitatively changed independent whether they are identified by database search. The pros and cons of such an approach over the classical shotgun proteomics technique are discussed in the following chapters. Eventually, peptides are identified via their fragment fingerprint against genomic/predicted proteomic databases (see last paragraph “genome annotation”). Protein identification is then based on the reconstruction from these identified peptides. Compared to classical methods in proteomics, shotgun proteomics is characterized by a very high protein identification rate and thus qualified to establish huge qualitative and quantitative proteome maps for genome-sequenced but also non-sequenced organisms [3, 10]. Another strategy uses those proteome catalogues for the design of “proteotypic” peptides of a protein for a targeted analysis. We call this technique MASS WESTERN because it resembles the Western blot technology with the ability to detect proteins in complex matrices after several preparatory and pre-purification steps for instance using an SDS-PAGE step [4]. The MASS WESTERN has several advantages over a classical antibody-based approach for targeted protein analysis such as higher specificity, simultaneous analysis of several proteins in a complex sample and, of course, no need to produce antibodies against the target protein. In the next paragraphs we introduce these techniques associated with multivariate statistics and a proteomic database.

3 MAPA (*Mass Accuracy Precursor Alignment*) and ProtMAX

Genomic databases and their corresponding computer-predicted proteomic databases became essential for proteome science. They form the bases for protein identification of shotgun proteomics

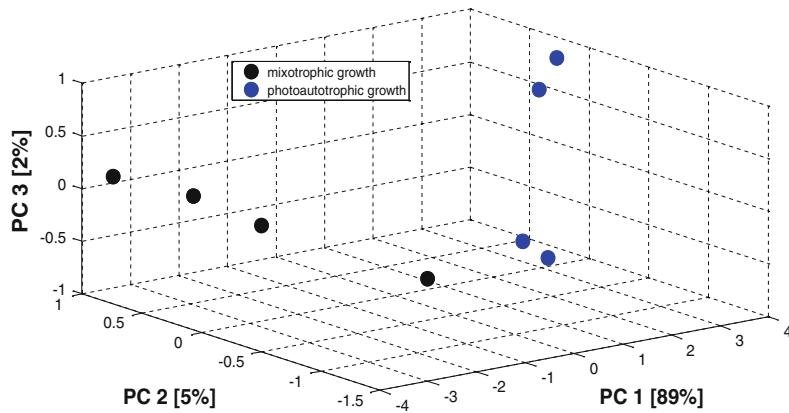
analyses. To enable confident protein identifications, ideally their complete sequence information needs to be available or otherwise “de novo” interpretation of mass spectrometric raw data is necessary. Cutting-edge mass spectrometry enables high mass accuracy analyses. This allows for de novo-sequencing of peptides and proteins, respectively. The exploratory analysis can now be divided into two possible strategies, the database dependent and independent interpretation. The general goal is the identification of as many proteins as possible out of a complex protein mixture. With the database-dependent approach only unambiguously identified sequences are considered. Consequently, peptides not identified by database-search most probably due to polymorphisms, low MS/MS quality, or posttranslational modifications will be lost in this strategy. However, these peptides can be very decisive with respect to sample classification [3]. In contrast, the database-independent approach called MAPA is based upon an algorithm that groups all peptide precursor ions with the same mass to charge ratio (m/z) derived from high mass accuracy mass spectrometric raw data in a data matrix without initial database search. At the same time, the frequency of observed m/z -fragments according to the concentration of peptides is counted (spectral count) or ion intensity and added to the data matrix. Furthermore, the retention time is considered for binning the correct m/z -precursors [11]. This algorithm is called ProtMAX and can be downloaded from <http://www.univie.ac.at/mosys/software.html>. The resulting data matrix can be analyzed statistically—for instance using COVAIN [5]—to rank peptide precursor ions according to their impact on sample classification. COVAIN can be downloaded from <http://www.univie.ac.at/mosys/software.html>. In Fig. 2 an example is shown for different growth conditions—mixotrophic versus photoautotrophic—of the unicellular green alga *Chlamydomonas reinhardtii*. Interesting candidates can then be identified via database search or de novo interpretation according to Hoehenwarter et al. [3]. This allows also for the identification of rarely detectable forms of protein modifications and polymorphisms (Fig. 2e). For instance, the combination of metal oxide affinity chromatography (MOAC) with MAPA enables the unbiased quantification of phosphoproteins [12–14]. In summary, although the untargeted analysis only allows for relative quantification it enables a rapid overview of what is changed at a systems response level. Moreover, this strategy also delivers basic data for the targeted analysis—the MASS WESTERN.

4 MASS WESTERN and ProMEX

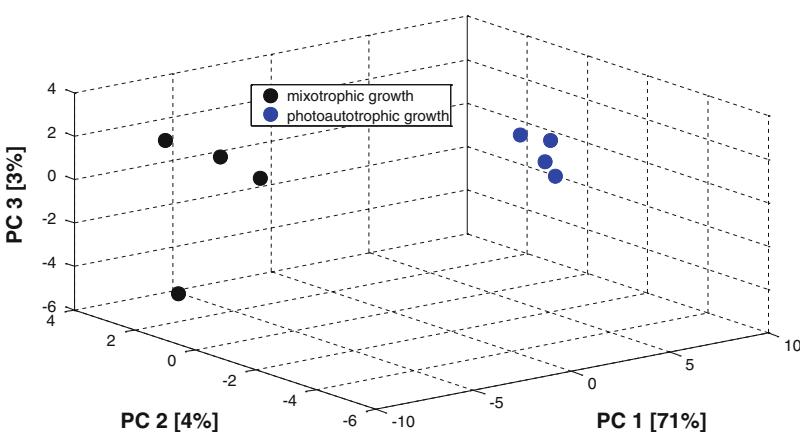
The best known method for a targeted analysis of specific proteins out of a complex sample is based on antibodies (e.g. Western Blot, Fig. 3). Besides time-consuming and extensive

a

MASS WESTERN

**b**

MAPA

**c**

MAPA + MASS WESTERN

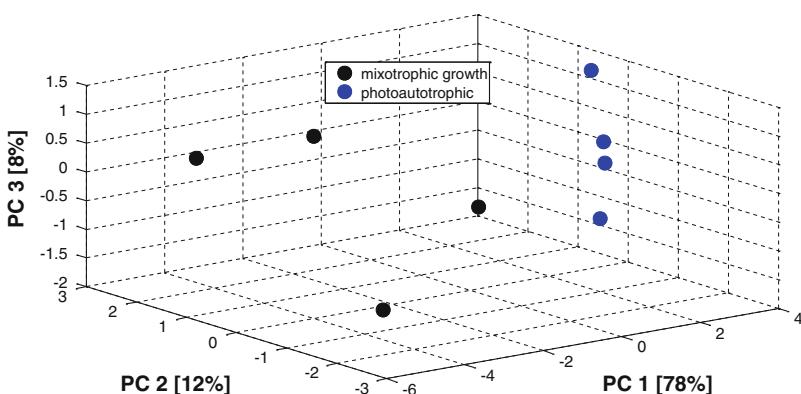
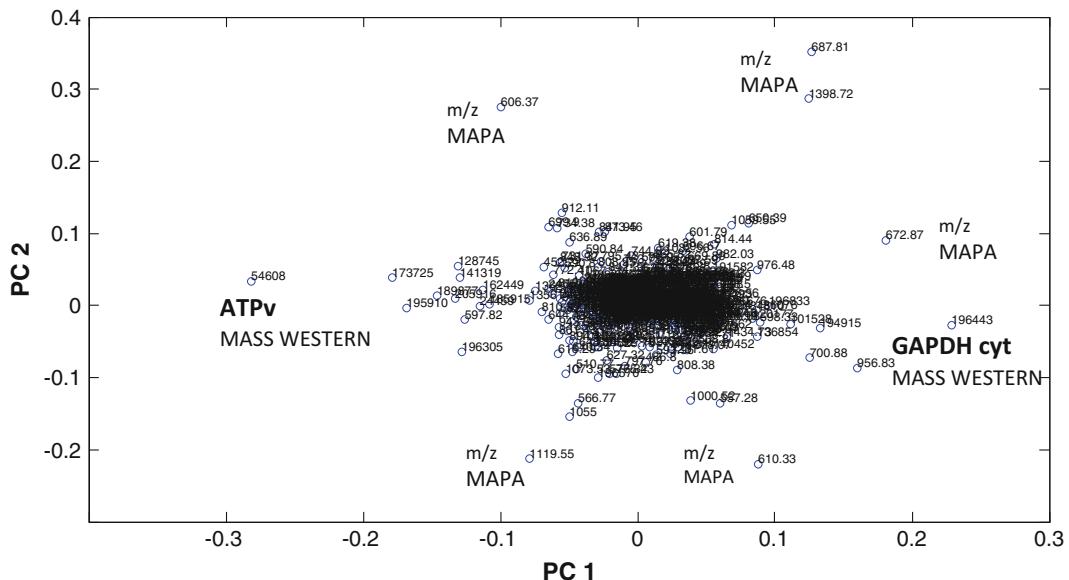


Fig. 2 (a) COVAIN-based principal components analysis (PCA) and sample pattern recognition for a MAPA protein data matrix of two growth conditions of *Chlamydomonas reinhardtii*—mixotrophic and photoautotrophic [16]. Visible is the clear separation of the two growth types. The separation in the principal component analysis is due to the differences in tryptic peptide m/z -ratio abundances. (b) COVAIN-based principal components

d

LOADINGS MAPA + MASS WESTERN

**e**

Precursor ion:
m/z 1199.6112 (3+)

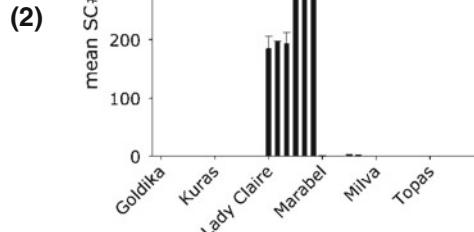
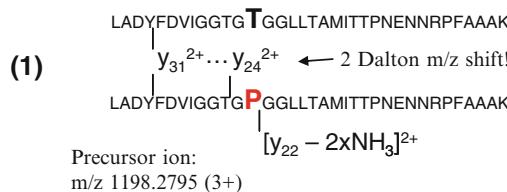


Fig. 2 (continued) analysis (PCA) and sample pattern recognition for a MASS WESTERN protein data matrix of two growth conditions of *Chlamydomonas reinhardtii*—mixotrophic and photoautotrophic [16]. Visible is the clear separation of the two growth types. The separation in the principal components analysis is due to the differences in protein abundances identified by proteotypic peptides. (c, d) MAPA and MASS WESTERN data were combined using COVAIN. First, each single data matrix is uploaded in COVAIN. Then the different data sets are combined in one data matrix by COVAIN-supported selection of samples versus variables alignments. This new data matrix is z-transformed and analyzed by principal components analysis (PCA). All steps are performed by COVAIN. The two growth conditions of *Chlamydomonas reinhardtii*—mixotrophic and photoautotrophic [16]—are clearly separated in (c). In the corresponding loadings plot of (d) either proteins (MASS WESTERN) or tryptic peptides (MAPA) are assigned for sample classification. This procedure leads to rapid identification of protein markers out of millions of variables [3]. (e) Identification of a protein polymorphism using the MAPA approach [3]. (1) De novo sequencing and detection of an amino acid exchange threonine versus proline. (2) This specific protein is exclusively detected in a potato “Lady Claire” cultivar

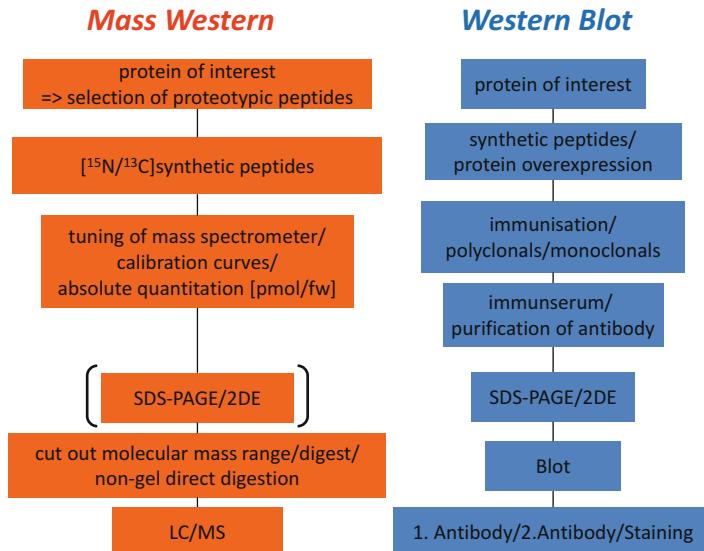


Fig. 3 MASS WESTERN strategy. Proteotypic peptides are initially selected from a proteome analysis. Those peptides serve as model for synthetic internal stable isotope standards and are introduced for absolute quantification using triple quadrupole mass analyzer [4]. For comparison the typical strategy of a Western blot using specific antibodies is shown

production of protein-specific antibodies it is very difficult to distinguish between proteins of high homology leading to so-called cross-reactivity. In addition with this technique absolute quantification of exact concentrations is almost impossible. We developed a strategy based on mass spectrometry, which enables to distinguish and to absolutely quantify protein isoforms or whole pathways out of complex samples using stable isotope-labeled synthetic peptides. Principles of the technique were first employed in the late 1980s [15]. Due to its similarity with the Western Blot it is also called MASS WESTERN (Fig. 3). The MASS WESTERN allows for a high sample throughput and the possibility to analyze many proteins (>100) within a single analysis depending on how many proteotypic peptides per protein are used. Proteome data measured in different cell states can be integrated in a proteome database (e.g. ProMEX, Fig. 1). They comprise proteotypic peptide libraries as base for the MASS WESTERN method design (Figs. 1 and 3). It allows for the detection of entire metabolic pathways such as Calvin Cycle, Glycolysis, and citric acid cycle in parallel [16]. These data are important for the mathematical reconstruction and prediction of metabolic pathways and their regulatory mechanisms and hence relevant for new insights in systems biology.

5 Proteomics and Sample Pattern Recognition in Systems Biology

Systems biology is aiming at a holistic overview on all regulatory processes and reactions (phenotypic plasticity) of a biological system in response to environmental perturbations. Resolution of these processes improves with the amount of data available. Consequently, integration of protein-, metabolite- and transcript data enhances the resolution. Especially, with untargeted protein analysis but also by integrating MASS WESTERN analyzes (Figs. 1 and 2), huge amounts of data are generated. Therefore statistical techniques are necessary for comprehensive data mining and the extraction of biological relevant information. One of the most important methods for data mining and data visualization is a pattern recognition strategy based on supervised and unsupervised multivariate statistics such as Principal Components Analysis (PCA) (Fig. 2a). Through pattern recognition, conclusions about biologically active regulatory processes and proteins can be drawn. In Fig. 2 MAPA and MASS WESTERN data of the same sample of *Chlamydomonas reinhardtii* grown under either mixotrophic or photoautotrophic conditions are shown. In Fig. 2C the integration of MAPA and MASS WESTERN data is shown leading to a clear separation of the two growth conditions. Data integration and principal components analysis was performed with COVAIN [5]. Each single data set was uploaded and then integrated into one data matrix by COVAIN. The data were z-transformed to allow the comparison of MAPA and MASS WESTERN values and visualized by principal components analysis using all the functions of COVAIN. In Fig. 2d the loadings of the integrated data are shown demonstrating that both MASS WESTERN data specific for single proteins as well as tryptic peptide precursors identified by the MAPA algorithm have an impact on sample classification. All the necessary software and manuals for a step-by-step description for all of these data processing and data mining procedures can be downloaded at <http://www.univie.ac.at/mosys/software.html>.

6 Combining MAPA and MOAC for Quantitative Analysis of Phosphoproteins, Signaling Cascades, and Novel Protein Kinase Targets

In Fig. 4 a convenient procedure is described for the unbiased enrichment of phosphoproteins using different variants of a metal-oxide-affinity chromatography MOAC [17–19]. We call this procedure tandem MOAC because the enrichment of phosphoproteins is subsequently combined with the enrichment of phosphopeptides. The procedure was recently published and led to the identification of novel targets of the mitogen-activated

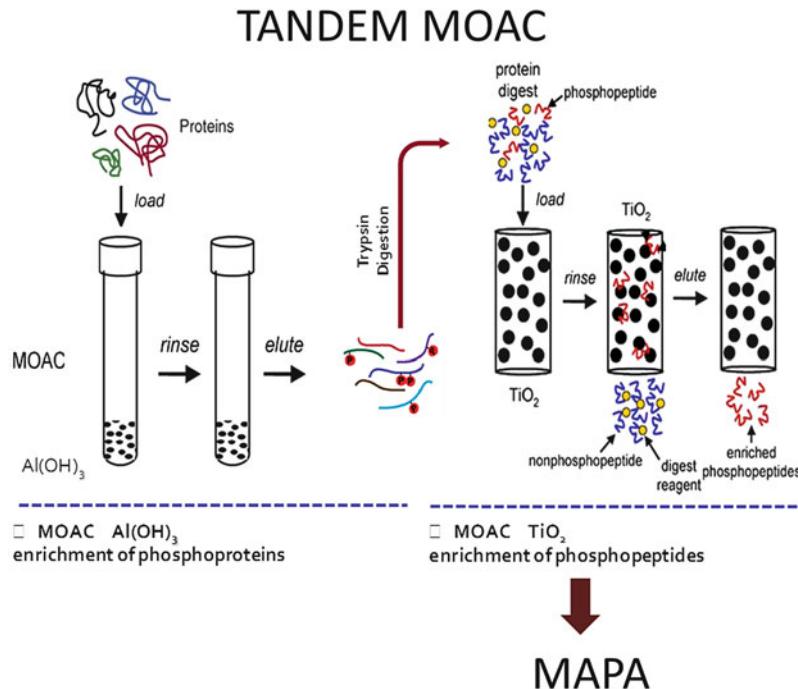


Fig. 4 TANDEM MOAC procedure combined with MAPA [14]. Phosphoproteins are enriched first using alumina-based MOAC [17, 18]. The enriched phosphoprotein fraction is digested and phosphopeptides are enriched using TiO_2 -MOAC. This phosphopeptide fraction is analyzed by LC-MS and MAPA and subsequently with multivariate statistics such as PCA or ICA using COVAIN to identify changed abundances of specific phosphorylation sites in proteins [12]

protein kinase (MPK) signaling pathway [14]. An essential requirement for the confidential identification of novel protein kinase targets as a regulatory response to environmental perturbations is the accurate quantification of the phosphoproteins/phosphopeptides. The reason is that protein phosphorylation—though low abundant—is prevalent and distinguishing changes is only possible with quantification of the respective phosphorylation site of the protein as a result of a switched-on signaling cascade. Therefore we combined MOAC with MAPA as a quantitative phosphoproteomics approach to identify novel protein kinase targets switched on in signaling processes. As discussed above the tandem MOAC procedure was combined with MAPA to identify new MPK substrates [14]. In another study novel phosphoproteins were identified as a response to phytohormone signaling using the combination of MOAC and MAPA [12]. Another study used MOAC and MAPA for the identification of phosphoproteins during pollen tube growth [13].

7 Genome Annotation: Shotgun Proteomics Complements Shotgun Genomics

The number of novel sequenced genomes and new genome projects—both prokaryotic, eukaryotic as well as the “metagenome” of communities of organisms—is exponentially increasing. The high amount of data generated displays the enormous challenge for bioinformatics. A classical approach is the pure computer-assisted annotation of predicted Open Reading Frames (ORF). Newly developed techniques consider experimental data such as proteomic-based data but also metabolomic data for improved genome annotation [20]. A high-throughput method for high protein identification rate is “shotgun proteomics”-analysis described above. Along the principles of shotgun genomics technology, proteins can be reconstructed from tryptic peptides stemming from a digest of whole proteomes. Shotgun proteomics is characterized by a very high protein identification rate and generates huge qualitative proteome catalogues from model organisms. In a recent publication we used shotgun proteomics data for a projection of all identified proteins into a functional genome annotation and subsequent metabolic reconstruction of the unicellular green algae *Chlamydomonas reinhardtii*, a recently sequenced model organism for photosynthesis and CO₂-neutral biomass production also called the “green yeast” [20]. This way, predicted gene models can be confirmed by proteomic data [20]. Furthermore, many protein data that are not predicted by existing gene models may point to new gene models not detectable by computer-based in silico analyses only. Recently, we performed a comprehensive shotgun proteomics analysis of various growth conditions of *Chlamydomonas reinhardtii*, integrated all the proteomics raw data and searched this dataset against different *Chlamydomonas* genome annotation databases of the last 5 years [21]. The result is that the use of different genome annotation databases of the same organism has a strong impact on the functional interpretation of the proteomics data as well as on quantitative proteomics [21]. Therefore it is highly recommended that databases are integrated or 6-frame-translations are used for peptide identification. Furthermore, the differences between the database searches point to critical genomic regions and single genes for gene function analysis and annotation problems [21].

8 Conclusion

The presented proteomics and data integration platform is applicable to many different research fields such as proteomic investigations, metabolic modeling, proteogenomics, and phosphoproteomics.

Novel statistical tools implemented in COVAIN allow the exploration of new relationships within the data. The complete software including step-by-step-tutorials is available at <http://www.univie.ac.at/mosys/software.html>. The platform allows the rapid and convenient analysis of proteomics data and provides many more features which were not be discussed here due to space limitations but can be explored by the user himself. We encourage our colleagues to explore the toolbox and also give feedback for improvements which can be implemented in future.

References

1. Weckwerth W (2011) Unpredictability of metabolism: the key role of metabolomics science in combination with next-generation genome sequencing. *Anal Bioanal Chem* 400:1967–1978
2. Weckwerth W (2003) Metabolomics in systems biology. *Annu Rev Plant Biol* 54:669–689
3. Hoehenwarter W, van Dongen JT, Wienkoop S et al (2008) A rapid approach for phenotype-screening and database independent detection of cSNP/protein polymorphism using mass accuracy precursor alignment. *Proteomics* 8:4214–4425
4. Lehmann U, Wienkoop S, Tschoep H et al (2008) If the antibody fails: a mass western approach. *Plant J* 55:1039–1046
5. Sun X, Weckwerth W (2012) COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics* 8:81–93
6. Doerfler H, Lyon D, Nägele T et al (2013) Granger causality in integrated GC-MS and LC-MS metabolomics data reveals the interface of primary and secondary metabolism. *Metabolomics* 9(3):564–574
7. Mari A, Lyon D, Fragner L et al (2013) Phytochemical composition of *Potentilla anserina* L. analyzed by an integrative GC-MS and LC-MS metabolomics platform. *Metabolomics* 9:599–607
8. Hummel J, Niemann M, Wienkoop S et al (2007) ProMEX: a mass spectral reference database for proteins and protein phosphorylation sites. *BMC Bioinformatics* 8:216
9. Wienkoop S, Staudinger C, Hoehenwarter W et al (2012) ProMEX: a mass spectral reference database for plant proteomics. *Front Plant Sci* 3:125
10. Hoehenwarter W, Larhlimi A, Hummel J et al (2011) MAPA distinguishes genotype-specific variability of highly similar regulatory protein isoforms in potato tuber. *J Proteome Res* 10:2979–2991
11. Egelhofer V, Hoehenwarter W, Lyon D, Weckwerth W, Wienkoop S (2013) Using ProtMAX to create high mass accuracy precursor alignments from label-free quantitative mass spectrometry data generated in large shotgun proteomics experiments. *Nat Protoc* 8:595–601
12. Chen Y, Hoehenwarter W, Weckwerth W (2010) Comparative analysis of phytohormone-responsive phosphoproteins in *Arabidopsis thaliana* using TiO₂-phosphopeptide enrichment and mass accuracy precursor alignment. *Plant J* 63:1–17
13. Chen Y, Liu P, Hoehenwarter W et al (2012) Proteomic and phosphoproteomic analysis of *picea wilsonii* pollen development under nutrient limitation. *J Proteome Res* 11:4180–4190
14. Hoehenwarter W, Thomas M, Nukarinen E et al (2013) Identification of Novel *in vivo* MAP Kinase Substrates in *Arabidopsis thaliana* Through Use of Tandem Metal Oxide Affinity Chromatography. *Mol Cell Proteomics* 12(2):369–380
15. Desiderio DM, Kai M (1983) Preparation of stable isotope-incorporated peptide internal standards for field desorption mass-spectrometry quantification of peptides in biologic tissue. *Biomed Mass Spectrom* 10:471–479
16. Wienkoop S, May P, Kempa S et al (2010) Targeted proteomics for *Chlamydomonas reinhardtii* combined with rapid subcellular protein fractionation, metabolomics and metabolic flux analyses. *Mol Biosyst* 6:1018–1031
17. Wolschin F, Weckwerth W (2005) Combining metal oxide affinity chromatography (MOAC) and selective mass spectrometry for robust identification of *in vivo* protein phosphorylation sites. *Plant Methods* 1:9
18. Wolschin F, Wienkoop S, Weckwerth W (2005) Enrichment of phosphorylated proteins and peptides from complex mixtures using metal oxide/hydroxide affinity chromatography (MOAC). *Proteomics* 5:4389–4397
19. Larsen MR, Thingholm TE, Jensen ON et al (2005) Highly selective enrichment of phos-

- phorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* 4:873–886
20. May P, Wienkoop S, Kempa S et al (2008) Metabolomics-and proteomics-assisted genome annotation and analysis of the draft metabolic network of *Chlamydomonas reinhardtii*. *Genetics* 179:157–166
21. Valledor L, Recuenco-Munoz L, Egelhofer V et al (2012) The different proteomes of *Chlamydomonas reinhardtii*. *J Proteomics* 75:5883–5887

Chapter 3

Plant Proteomics: From Genome Sequencing to Proteome Databases and Repositories

Katsumi Sakata and Setsuko Komatsu

Abstract

Proteomic approaches are useful for the identification of functional proteins. These have been enhanced not only by the development of proteomic techniques but also in concert with genome sequencing. In this chapter, 30 databases and Web sites relating to plant proteomics are reviewed and recent technologies relating to data collection and annotation are surveyed.

Key words Plant genome, Proteome, Database, 2-DE, MS, Shotgun proteomics, Annotation

Abbreviations

| | |
|--------------|---|
| EST | Expressed sequence tag |
| MS | Mass spectrometry |
| <i>n</i> -DE | <i>n</i> -Dimensional electrophoresis |
| SDS-PAGE | Sodium dodecyl sulfate-polyacrylamide gel electrophoresis |

1 Introduction

Comprehensive approaches for biomarker searching and functional analysis include proteomics, transcriptomics, and metabolomics. Proteomic approaches allow us to directly analyze the proteins that exist in a living body just at the emergence of biological functions [1]. The identification of proteins has been dramatically enhanced by mass spectrometers and homology searches against large-scale genome sequence data [2]. Furthermore, many proteins have been discovered in relation to the differentiation and growth of plants, as well as proteins differentially expressed depending on the environmental conditions, based on proteomic techniques that can comprehensively analyze protein interactions [3, 4]. As with genomic data, large volumes of proteomic data have been accumulated in recent years coupled with the development of high-throughput analysis methodologies.

Table 1
Large-scale plant genome sequencing projects [5] and representative proteome databases

| | Order | Organism | Proteome database |
|---------|--------------|--|---|
| Dicot | Malpighiales | California poplar, Cassava(P) | |
| | Fabales | Soybean, <i>Medicago</i> , Lotus(P) | Proteomics of Oilseeds, Soybean Proteome Databases, SoyKB, ProMEX |
| | Brassicaceae | <i>Arabidopsis</i> , Papaya(P) | PPDB, AtProteome, ProMEX, SUBA, PhosPhAt, AT_ CHLORO, AtPID, plprot, TAIR |
| | Rhamnales | Wine grape | |
| | Solanaceae | Tomato(P), Potato(P) | plprot |
| Monocot | Poales | Rice, Sorghum, Corn, Purple false brome(P) | PPDB, Rice Proteome Database, OryzaPG-DB, plprot, DIPOS, PRIN |

The phylogenetic tree among the orders is based on the Angiosperm Phylogeny Group classification [6]. The letter “P” attached to an organism name means that large-scale sequencing analysis of the genome is still in progress [5]

These proteomic data also need to be integrated and organized in databases that enable us to retrieve, leverage, and share public data through up-to-date computational technology such as the latest data management systems and Web-interface techniques.

This overview of the large-scale genome sequencing studies and the status of proteomic data repositories will provide a guide for datasets currently available and also those that should be prepared in future studies. We positioned large-scale sequencing projects [5] and major proteomic databases on a phylogenetic tree of flowering plants that was based on the biological classification by the Angiosperm Phylogeny Group [6] (Table 1). In the dicot plants, large-scale genome sequences have been analyzed across various orders of the biological classification. In the monocot plants large-scale genome sequences are localized in the order Poales, which includes the major cereal grain species. Many proteomic databases have been developed for the model dicot plant *Arabidopsis thaliana* such as AtProteome [7], ProMEX [8], and SUBA [9]. For monocot plants, rice databases such as the Rice Proteome Database [10] and OryzaPG-DB [11] have been developed. The PPDB [12, 13] covers two major species in dicots and monocots, *A. thaliana* and maize (*Zea mays*). The Proteomics of Oilseeds [14] and Soybean Proteome Database [15] store proteomic data primarily based on 2-DE in legumes. The plprot [16] database specializes in plant plastids and covers *A. thaliana*, tobacco, and rice. AtPID [17] contains large-scale protein-protein interaction data. Such molecular interaction databases have been

increasing, now including resources such as DIPOS and PRIN for rice [18, 19]. Furthermore, several databases have been organizing omics including transcriptome data and metabolomic pathways, including ProMEX [8], the Soybean Proteome Database [15], and SoyKB [20]. In this chapter, we also introduce the Web-based proteomic tools and prediction programs GelMap [21], iLoc-Plant [22], MRMAid [23], Musite [24], PeptideAtlas SRM Experiment Library (PASSEL) [25], PredPlantPTSI [26], ProteoRed MIAPE [27], and a general database relating to plant proteomics, Proteomics Identifications (PRIDE) database [28].

Based on the efforts of the Human Proteome Organization Proteomics Standards Initiative (HUPO PSI-MI), molecular interaction data have been standardized and a common query interface (PSI Common QUery InterfaCe, PSICQUIC, <http://code.google.com/p/psicquic/>) was developed [29]. This service allows the querying of nearly 30 molecular interaction databases all together. Interfaces such as these are desirable in the field of plant proteomics.

Genomic and proteomic studies have an interdependent relationship such that the development of one enhances development of the other. In the last decade, the availability of nucleotide sequences has accelerated the identification of proteins obtained by mass spectrometry (MS) and/or protein sequencers [30]. This acceleration has been greatly enhanced by the availability of the genomic sequences that have replaced EST sequences in recent years. The interdependency in the reverse direction might be referred to as proteogenomics [31–34]. It aims at improving genome annotation by using proteomic information based on MS. The information based on proteogenomics is also applicable to several analyses including frameshifts of coding sequences and posttranslational modifications such as N-terminal methionine excision, signal peptides, and proteolysis [34, 35]. The proteogenomics approach has been applied to *A. thaliana* [36].

2 Plant Proteome Databases and Repositories

Here we summarize the representative databases and Web sites relating to plant proteomics. Each section is presented in alphabetical order. The URLs are tabulated in Table 2.

2.1 *Arabidopsis Thaliana* Databases

2.1.1 AT_CHLORO

This database was developed for chloroplast proteome data from *A. thaliana*. The proteome datasets were extracted from *Arabidopsis* leaves. LC-MS/MS-based analysis was used to identify ~1,300 proteins from more than 10,000 unique peptide sequences. The partitioning of each protein in the three chloroplast compartments was validated by using a semiquantitative proteomics approach (spectral count). The chloroplasts were purified by Percoll density gradients and SDS-PAGE [37].

Table 2
URLs for plant proteome databases and Web sites

| <i>Arabidopsis thaliana databases</i> | |
|---------------------------------------|--|
| AT_CHLORO | www.grenoble.prabi.fr/at_chloro/ |
| AtPID | www.megabionet.org/atpid/webfile/ |
| AtProteome | fgcz-atproteome.unizh.ch/ |
| pep2pro | fgcz-pep2pro.uzh.ch/ |
| PhosPhAt | phosphat.mimpimp-golm.mpg.de/db.html |
| SUBA | suba.plantenergy.uwa.edu.au/ |
| TAIR | www.arabidopsis.org/ |
| <i>Rice databases</i> | |
| DIPOS | csb.shu.edu.cn/dipos |
| OryzaPG-DB | oryzapg.iab.keio.ac.jp/ |
| PRIN | bis.zju.edu.cn/prin/ |
| Rice proteome database | gene64.dna.affrc.go.jp/RPD/ |
| RiceRBP | www.bioinformatics2.wsu.edu/cgi-bin/RiceRBP/home.pl |
| <i>Other databases</i> | |
| GabiPD | www.gabipd.org/ |
| Medicago PhosphoProtein DB | phospho.medicago.wisc.edu |
| P3DB | p3db.org/ |
| plprot | www.plprot.ethz.ch/ |
| PPDB | ppdb.tc.cornell.edu/ |
| PRIDE | www.ebi.ac.uk/pride/ |
| ProMEX | promex.pph.univie.ac.at/promex/ |
| Proteomics of Oilseeds | www.oilseedproteomics.missouri.edu/ |
| Seed Proteome Web Portal | www.seed-proteome.com/ |
| Soybean Proteome Database | proteome.dc.affrc.go.jp/Soybean/ |
| SoyKB | soykb.org/ |
| <i>Web-based tools</i> | |
| GelMap | www.gelmap.de/ |
| iLoc-Plant | www.jci-bioinfo.cn/iLoc-Plant |
| MRMaid | www.mrmaid.info/ |
| Musite | musite.net/ |
| PASSEL | www.peptideatlas.org/passel/ |
| PredPlantPTS1 | ppp.gobics.de/ |
| ProteoRed MIAPE web toolkit | www.proteored.org/MIAPE/ |

2.1.2 AtPID

AtPID includes predicted protein–protein interactions for *A. thaliana* [17]. Interactions are predicted from ortholog interactions, microarray profiles, GO analyses [38], conserved domains, and genomic contexts. The database contains ~28,000 protein–protein interaction pairs with ~23,000 pairs generated from prediction methods. The remaining ~5,000 pairs were manually created from the literature and/or from enzyme complexes in KEGG [39].

2.1.3 AtProteome

This database contains an organ-specific proteome map for *A. thaliana*. The protein identification information is displayed by proteogenomic mapping of the peptides onto the genome. Additional information on the identification of proteins is linked, such as the amino acid sequence of the first splice variant of the protein and the detected peptides sorted by their position in the protein [7].

2.1.4 pep2pro

This database is a further development of AtProteome and provides proteome information on *A. thaliana* [40]. The Web site shows the protein identification information by proteogenomic mapping of the peptides onto the genome. The TAIR9 dataset is provided as the default dataset. The pep2pro data analysis pipeline also handles data export to the PRIDE database [28] and data retrieval by the MASCP Gator (<http://gator.masc-proteomics.org/>).

2.1.5 PhosPhAt

PhosPhAt is a database of phosphorylation sites in *A. thaliana*. The database contains ~1,200 defined tryptic peptides matching ~1,000 distinct proteins. Phosphorylation sites are marked as “defined” if the precise location of the phosphorylated amino acid has been unambiguously determined by mass spectrometric analysis [41, 42].

2.1.6 SUBA

This database focuses on protein localization in *A. thaliana*. It stores more than 6,700 nonredundant proteins observed in ten distinct subcellular locations. Various experimental methods were used for the localization, such as chimeric fluorescent fusion proteins, MS, literature references, and location prediction software based on amino acid sequences [9].

2.1.7 TAIR

A comprehensive genome database for *A. thaliana*, The *Arabidopsis* Information Resource (TAIR), also provides data repositories for *Arabidopsis* proteomics resources: (1) primary protein sequences, (2) protein domains, (3) protein structures including 3-D structure images, (4) protein–protein interactions, (5) biochemical properties including enzymes and biochemical pathways [43].

2.2 Rice Databases

2.2.1 DIPOS

DIPOS provides information on interacting proteins in rice (*Oryza sativa*), where the interactions are predicted using two computational methods, interologs and domain-based methods [18]. An interolog is a conserved interaction between a pair of proteins that have interacting homologs in another organism [44]. The database stores nearly 15 M pairwise interactions among 27,746 proteins. Each interaction is assigned a confidence score, and biological explanations of pathways and interactions are also provided.

2.2.2 *OryzaPG-DB*

Proteins stored in this database were extracted from rice (*Oryza sativa*) and MS analysis was conducted based on the shotgun proteomics approach [11]. The proteins contained in the database were compared with protein data from full-length cDNA sequence databases such as RAP-DB [45]. This approach should enable detection of definitively functional proteins compared to the similar Rice Proteome Database [10], which is based on conventional protein sequencing. Nearly 3,200 genes are covered by the peptides identified by searching the product ion spectra against the protein, cDNA, transcript, and genome databases.

2.2.3 *PRIN*

This Web site provides a prediction of protein–protein interactions in *Oryza sativa* [19]. It is based on a method known as interologs [44], and six model organisms where large-scale protein–protein interaction experiments have been applied: yeast, worm, fruit fly, human, *E. coli*, and *Arabidopsis*. Some 76,585 nonredundant rice protein interaction pairs have been predicted among 5,049 rice proteins.

2.2.4 *Rice Proteome Database*

This database contains proteome data from rice (*Oryza sativa* cv. Nipponbare) based on 2-DE techniques. The database stores more than 20 reference maps based on 2-DE of proteins from rice tissues and subcellular compartments. The reference maps comprise more than 10,000 identified proteins showing tissue and subcellular localization, corresponding to ~4,100 separate protein entries in the database. Amino acid sequences were determined by protein sequencers and MS, which were up-to-date technologies at the time the database was developed [10].

2.2.5 *RiceRBP*

This database contains 257 experimentally identified RNA-binding proteins (RBPs) in rice. Many of these have not previously been predicted to be RBPs. For each identified protein, information is provided on transcript and protein sequences, predicted protein domains, details of the experimental identification, and whether antibodies have been generated for public use [46].

2.3 Other Databases

2.3.1 *GabiPD*

This database provides integrated plant “omics” data, and was developed as part of the German initiative for Genome Analysis of the Plant Biological System (GABI) [47]. Data from different “omics” are integrated and interactively connected. 2-D electrophoresis gel images were collected from different tissues of *A. thaliana* and *Brassica napus*. Stored data relating to phosphorylation have links with external data in the PhosPhAt database through related data in the Gene GreenCard database.

2.3.2 *Medicago PhosphoProtein Database*

This database stores phosphoprotein, phosphopeptide, and phosphosite data specific to *Medicago truncatula*, the model system for legume biology. It includes 3,457 phosphopeptides that contain

3,404 nonredundant sites of phosphorylation on 829 proteins. Through a Web-based interface, users can browse identified proteins or search for proteins of interest, and also conduct BLAST searches of the database using peptide sequences and phosphorylation motifs as queries [48].

2.3.3 *P3DB*

This database stores plant protein phosphorylation site data, organizing information on 32,963 nonredundant sites collected from 23 experimental studies of six plant species. The data can be searched for a protein of interest using an integrated BLAST module to query similar sequences with known phosphorylation sites [49].

2.3.4 *plprot*

The *plprot* is a plastid proteome database that provides information about the proteomes of chloroplasts, etioplasts, and undifferentiated plastids. The database stores more than 2,000 proteins. The basic module integrates a homology search and comparative information on the proteomes of different plastid types. Data from *A. thaliana*, tobacco, and rice are contained in the database [16].

2.3.5 *PPDB*

PPDB is a Plant Proteome DataBase for *A. thaliana* and maize (*Zea mays*). *PPDB* was developed for plant plastids and expanded to the whole plant proteome. The name of the database was changed accordingly, using the same abbreviation, from Plastid PDB to Plant PDB. The current database has large-scale proteomic data in diverse forms: (1) 5,000 identified proteins both in *Arabidopsis* and maize, (2) 80 published *Arabidopsis* proteome datasets from subcellular compartments or organs linked to each locus, and (3) 1,500 *Arabidopsis* proteins manually assigned subcellular locations. Information from MS-based identification and posttranslational modification is available for each identified protein [12, 13].

2.3.6 *PRIDE*

The PRIDE database provides standardized MS proteomics data. It is one of the main repositories of proteomics data that have been generated by MS approaches. Recently, it has become a database that various journals in the field are supporting and even mandating deposition of proteomics data in. Datasets are stored in PRIDE without modification or reanalysis, and the research community can access the original results obtained by the research group [28].

2.3.7 *ProMEX*

ProMEX is a database of MS/MS reference spectra (mostly Orbitrap precursor ion mass data) from *plants and microbes* [8, 42, 60]. The data were generated based on liquid chromatography coupled to ion trap MS (LC-IT-MS). This current release 2.9/2012 contains 51,793 tryptic peptide product ion spectra entries of 27,886 different peptide sequence entries from *Medicago truncatula*, *Chlamydomonas reinhardtii*, *Bradyrhizobium japonicum*, *Arabidopsis thaliana*, *Phaseolus vulgaris*, *Lotus japonicus*, *Lotus corniculatus*,

Lycopersicon esculentum, *Solanum tuberosum*, *Nicotiana tabacum*, *Sinorhizobium meliloti*, *Glycine max*, and *Zea mays*. Furthermore a search algorithm is implemented which allows to search single spectra and mzXML-LC-MS/MS runs against the database. Protein data are linked to “omes” such as metabolites, pathways, and transcripts. Peptide identification was based on peptide mass fingerprinting [50].

2.3.8 Proteomics of Oilseeds

This database stores reference maps of soybean (*Glycine max* cv. Maverick) proteins based on 2-DE. The samples were collected during seed filling in the plant and analyzed at 2, 3, 4, 5, and 6 weeks after flowering. It contains expression profiles for 679 protein spots, from which 422 proteins representing 216 nonredundant proteins were identified [14].

2.3.9 Seed Proteome Web Portal

This Web site provides information both on quantitative seed proteomic data and on seed-related protocols. As a proteomic database, it gives access to 475 different *Arabidopsis* seed proteins annotated from 2-DE maps, including quantitative data according to the accumulation profile of each protein during the germination process. The Web site also provides protocols that the authors have routinely used for *Arabidopsis* seed proteome studies, such as procedures for sample preparation, electrophoresis coupled with gel analysis in 2-D electrophoresis, and protein identification by mass spectrometry [51].

2.3.10 Soybean Proteome Database

The current version of this database contains 23 reference maps of soybean (*G. max* cv. Enrei) proteins based on 2-DE [15]. The samples were collected from several organs, tissues, and organelles. The reference maps include 8,262 detected proteins and 672 identified proteins, or proteins for which a sequence or a peptide peak has been determined. An omics aspect is also included as a table linked to temporal expression profiles that reveals relationships among 106 mRNAs, 51 proteins, and 89 metabolites that vary over time under flooding stress (Fig. 1). The Web interface is representative of proteome databases based on 2-DE, and was developed using the Make2DDB II environment [52], which serves a standardized search function for the stored protein spots based on the accession number, description of the protein, and isoelectric point/molecular weight range. The database focuses on the seedling stage in soybean, 0–7 days after seedling emergence, and in this it differs from the Proteomics of Oilseeds [14].

2.3.11 SoyKB

SoyKB abbreviates Soybean Knowledge Base and stores multiple “omes” including temporal profiles of proteins detected from soybean (*G. max* cv. Williams 82) [20]. The proteomic data are available for seeds, roots, and root hairs and for multiple conditions.

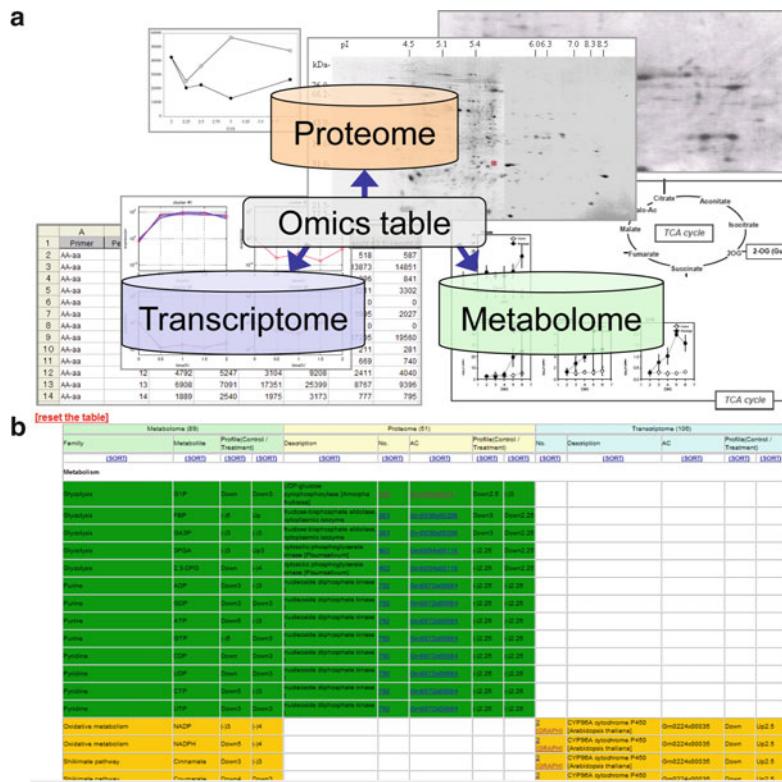


Fig. 1 A representative omics aspect in the Soybean Proteome Database [15]. **(a)** Proteome, transcriptome, and metabolome data are associated by an omics table. Proteomic data include 2-DE images and temporal expression profiles in the seedling stage in soybean. Transcriptomic and metabolomic data also include temporal profiles of the entities. Metabolites are mapped on metabolic pathways, as are associated mRNAs and proteins. **(b)** An example omics table in the database. The table indicates significant relationships between mRNAs, proteins, and metabolites with cells of the same color

Protein sequences and structures are linked to information about the genes such as gene models. Furthermore, metabolomics data from the SoyMetDB database [53] are incorporated into SoyKB.

2.4 Web-Based Proteomic Tools and Prediction Programs

2.4.1 GelMap

The Web site provides a tool for spot visualization on gel images [21]. Users can upload gel images and coordinates/information spreadsheets. The Web site also gives access to functional annotation of identified proteins defined by the user, and annotation of several proteins per analyzed protein “spot” according to MS data. Previously, this type of gel image-based database had to be opened through a server computer that users set up by themselves using a software tool such as the Make 2D-DB II package [52]. With GelMap a user can open the gel-based image data through the GelMap Web site without having to use a server.

2.4.2 iLoc-Plant

This Web site provides a tool for prediction of subcellular localization of plant proteins with single or multiple sites. A new prediction

method, the “multi-labeled learning” approach, was developed that can handle systems containing both single- and multiple-location plant proteins. An overall success rate of 71 % was demonstrated in the authors’ report [22].

2.4.3 *MRMaid*

Selected reaction monitoring (SRM), also called multiple reaction monitoring (MRM), is a tool for targeted quantitative proteomics [23]. This Web site helps users design SRM assays by suggesting peptides and product ions to monitor based on millions of experimental spectra from the PRIDE database [28]. By using data from the public repository PRIDE, MRMaid covers an increasing number of species as the coverage of PRIDE grows. Transitions by the Web site for 25 *A. thaliana* proteins were evaluated experimentally, and found capable of quantifying 23 of these proteins.

2.4.4 *Musite*

Musite is a Web site developed to predict phosphorylation sites based solely on protein sequences. It can be also downloaded as a stand-alone tool. Phosphorylation data from *A. thaliana*, *B. napus*, *G. max*, *M. truncatula*, *O. sativa*, and *Z. mays* were collected for cross-species testing. It was reported that the model for *A. thaliana* could be extended to other organisms, and the overall plant model from Musite was reported to be better than other plant-specific prediction tools in prediction accuracy [24].

2.4.5 *PASSEL*

The PeptideAtlas project has developed an analysis pipeline to identify peptides by tandem mass spectrometry (MS/MS), statistically validate the identifications, and map the identified sequences to the genomes of eukaryotic organisms. The PASSEL is a component of the PeptideAtlas project and a proteomic data repository for the collection and representation of data from SRM measurements. Users can submit, disseminate, and reuse SRM experimental results from analysis of biological samples [25].

2.4.6 *PredPlantPTS1*

Several computational approaches have been developed to predict peroxisomal proteins carrying the peroxisome targeting signal type 1 (PTS1). The Web site PredPlantPTS1 enables plant-specific prediction of PTS1 proteins based on a machine learning approach that is able to predict PTS1 proteins for higher plants (spermatophytes) with high accuracy [26].

2.4.7 *ProteoRed MIAPE Web Toolkit*

This Web site is a bioinformatics tool and performs several functions related to proteomic data standards, the HUPO-PSI’s (Proteomics Standards Initiative) standard data formats and Minimum Information About a Proteomics Experiment (MIAPE) guidelines: (1) verifying that reports fulfill the minimum information requirements of the corresponding MIAPE modules, highlighting inconsistencies or missing information; (2) converting several XML-based data standards directly into human-readable MIAPE reports stored

within the ProteoRed MIAPE repository; and (3) performing the reverse operation, allowing users to export from MIAPE reports into XML files for computational processing, data sharing, or public database submission [27].

3 Data Collection and Annotation of Proteins

We would like to illustrate the two major proteomics workflows: (i) protein extraction, electrophoresis (1-DE or 2-DE), cutting out bands (by 1-DE) or spots (by 2-DE), and MS and (ii) protein extraction and MS. The type-(i) approach is conventional, while the type-(ii) approach has been developed in more recent years and is also referred to as “shotgun proteomics” [54].

In proteomic analyses based on 2-DE (the type-(i) approach), the refined and extracted proteins are separated in the range of isoelectric point 3–10, in the first dimension, and molecular weight 10–100 kDa, in the second dimension [55]. A spot is cut out from the gel and the mass spectrum of the peptide is measured after reductive alkylation, enzymatic digestion, and desalination. Based on the obtained mass list and software searches against protein and nucleic acid databases, we detect the amino acid sequences matching the corresponding peptide mass spectrum data to identify the protein. To improve sensitivity, as well as the conventional Coomassie Brilliant Blue dye, fluorescent staining is also used extensively [2]. In this approach, we can use a gaseous phase protein sequencer instead of MS to determine the amino acid sequence.

The improvement of analysis precision and software development for MS have made possible a proteomics analysis not reliant on electrocataphoresis (the type-(ii) approach) [56, 57]. In this new approach, the refined or the extracted protein is digested with enzymes and the mixture of peptides is separated by liquid chromatography (LC) based on their hydrophobicity. The peptides eluted from the LC column are analyzed by MS, such as MS/MS, which is directly connected with the LC. The first MS separates each peptide ion and the second MS decomposes the peptide into fragments and determines the corresponding sequence from the fragmentation pattern. Software identifies the protein based on a homology search against reference databases. There are two derivations of this approach, labeling proteins before the analysis and analysis without labeling.

At the annotation stage of identifying proteins, we should give attention to the stability of the protein identifiers. A recent report revealed significant differences in the identifiers among the main protein databases: the International Protein Index (IPI), the UniProt Knowledgebase (UniProtKB), the National Center for Biotechnological Information nr database (NCBI nr), and Ensembl [58]. In the report, it was noted that some entries submitted to a

database are deleted afterwards in several months or a couple of years. The report demonstrated that UniProtKB was more stable than IPI and NCBI gi numbers. In recent years, the data deposited in publicly available proteomic databases have been increasing. There has been concern that automatic curation of repository data is dependent on analyses that could cause misannotation. Both database developers and users should be aware of these potential issues. *A posteriori* detection of such data encountered in typical proteomics datasets was also studied [59].

4 Conclusion

In recent years, a large amount of proteomic information including data related to biological functions has been created through proteomic research interlocked with genomic research. The proteomic databases we have introduced here have made such information publicly available. In plants, only a few species have complete genomic information but proteomic techniques can also be applied to species without genomic information. Proteomics techniques are useful to clarify the biological mechanisms underlying important plant traits, and proteomic databases will contribute to the establishment of technology to regulate and enhance such biological mechanisms and improve plant productivity.

References

1. Komatsu S, Konishi H, Shen S et al (2003) Rice proteomics: a step toward functional analysis of the rice genome. *Mol Cell Proteomics* 2:2–10
2. Komatsu S, Yano H (2006) Update and challenges on proteomics in rice. *Proteomics* 6:4057–4068
3. Salekdeh GH, Komatsu S (2007) Crop proteomics: aim at sustainable agriculture of tomorrow. *Proteomics* 7:2976–2996
4. Komatsu S, Ahsan N (2009) Soybean proteomics and its application to functional analysis. *J Proteomics* 72:325–336
5. Plant genomes central: genome projects in progress. <http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html>. Accessed 10 Feb 2012
6. The Angiosperm Phylogeny Group (2009) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot J Linn Soc* 161:105–121
7. Baerenfaller K, Grossmann J, Grobeij MA et al (2008) Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. *Science* 320:938–941
8. Hummel J, Niemann M, Wienkoop S et al (2007) ProMEX: a mass spectral reference database for proteins and protein phosphorylation sites. *BMC Bioinformatics* 8:216
9. Heazlewood JL, Verboom RE, Toni Filippini J et al (2006) SUBA: the *Arabidopsis* subcellular database. *Nucleic Acids Res* 35:D213–D218
10. Komatsu S, Kojima K, Suzuki K et al (2004) Rice Proteome Database based on two-dimensional polyacrylamide gel electrophoresis: its status in 2003. *Nucleic Acids Res* 32:D388–D392
11. Helmy M, Tomita M, Ishihama Y (2011) OryzaPG-DB: rice proteome database based on shotgun proteogenomics. *BMC Plant Biol* 11:63
12. Friso G, Giacomelli L, Ytterberg AJ et al (2004) In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. *Plant Cell* 16:478–499
13. Sun Q, Zybailev B, Majeran W et al (2008) PPDB, the plant proteomics database at Cornell. *Nucleic Acids Res* 37:D969–D974

14. Hajduch M, Ganapathy A, Stein JW et al (2005) A systematic proteomic study of seed filling in soybean. Establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. *Plant Physiol* 137:1397–1419
15. Sakata K, Ohyanagi H, Nobori H et al (2009) Soybean Proteome Database: a data resource for plant differential omics. *J Proteome Res* 8:3539–3548
16. Kleffmann T, Hirsch-Hoffmann M, Gruissem W et al (2006) plprot: a comprehensive proteome database for different plastid types. *Plant Cell Physiol* 47:432–436
17. Cui J, Li P, Li G et al (2008) AtPID: *Arabidopsis thaliana* protein interactome database: an integrative platform for plant systems biology. *Nucleic Acids Res* 36:D999–D1008
18. Sapkota A, Liu X, Zhao X-M et al (2011) DIPOS: database of interacting proteins in *Oryza sativa*. *Mol Biosyst* 7:2615–2621
19. Gu H, Zhu P, Jiao Y et al (2011) PRIN: a predicted rice interactome network. *BMC Bioinformatics* 12(1):161
20. Joshi T, Patil K, Fitzpatrick MR et al (2012) Soybean Knowledge Base (SoyKB): a web resource for soybean translational genomics. *BMC Genomics* 13(Suppl 1):S15
21. Senkler M, Braun H-P (2012) Functional annotation of 2D protein maps: the GelMap portal. *Front Plant Sci* 3:87
22. Wu Z-C, Xiao X, Chou K-C (2011) iLoc-Plant: a multi-label classifier for predicting the subcellular localization of plant proteins with both single and multiple sites. *Mol Biosyst* 7:3287–3297
23. Fan J, Mohareb F, Jones AME et al (2012) MRMaid: the SRM assay design tool for *Arabidopsis* and other species. *Front Plant Sci* 3:164
24. Yao Q, Gao J, Bollinger C et al (2012) Predicting and analyzing protein phosphorylation sites in plants using Musite. *Front Plant Sci* 3:186
25. Farrah T, Deutsch EW, Kreisberg R et al (2012) PASSEL: the PeptideAtlas SRM experiment library. *Proteomics* 12:1170–1175
26. Reumann S, Buchwald D, Lingner T (2012) PredPlantPTSI: a web server for the prediction of plant peroxisomal proteins. *Front Plant Sci* 3:194
27. Medina-Aunon JA, Martinez-Bartolome S, Lopez-Garcia MA (2011) The ProteoRed MIAPE web toolkit: a user-friendly framework to connect and share proteomics standards. *Mol Cell Proteomics* 10:M111.008334
28. Martens L, Hermjakob H, Jones P et al (2005) PRIDE: the proteomics identifications database. *Proteomics* 5:3537–3545
29. Orchard S (2012) Molecular interaction databases. *Proteomics* 12:1656–1662
30. Wijk KJ (2001) Challenges and prospects of plant proteomics. *Plant Physiol* 126:501–508
31. Gupta N, Tanner S, Jaitly N et al (2007) Whole proteome analysis of post-translational modifications: applications of mass-spectrometry for proteogenomic annotation. *Genome Res* 17:1362–1377
32. Ansong C, Purvine SO, Adkins JN et al (2008) Proteogenomics: needs and roles to be filled by proteomics in genome annotation. *Brief Funct Genomic Proteomic* 7:50–62
33. Jaffe JD, Berg HC, Church GM (2004) Proteogenomic mapping as a complementary method to perform genome annotation. *Proteomics* 4:59–77
34. Gupta N, Benhamida J, Bhargava V et al (2008) Comparative proteogenomics: combining mass spectrometry and comparative genomics to analyze multiple genomes. *Genome Res* 18:1133–1142
35. Proteogenomics. <http://en.wikipedia.org/wiki/Proteogenomics>
36. Castellana NE, Payne SH, Shen Z et al (2008) Discovery and revision of *Arabidopsis* genes by proteogenomics. *Proc Natl Acad Sci U S A* 105:21034–21038
37. Ferro M, Brugiere S, Salvi D et al (2010) AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol Cell Proteomics* 9:1063–1084
38. Barrell D, Dimme E, Huntle RP et al (2009) The GOA database in 2009—an integrated gene ontology annotation resource. *Nucleic Acids Res* 37:D396–D403
39. Kanehisa M, Goto S, Sato Y et al (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40:D109–D114
40. Hirsch-Hoffmann M, Gruissem W, Baerenfaller K (2012) pep2pro: the high-throughput proteomics data processing, analysis, and visualization tool. *Front. Plant Sci* 3:123
41. Heazlewood JL, Durek P, Hummel J et al (2008) PhosPhAt: a database of phosphorylation sites in *Arabidopsis thaliana* and a plant-specific phosphorylation site predictor. *Nucleic Acids Res* 36:D1015–D1021
42. Weckwerth W, Baginsky S, Wijk KV (2008) The multinational *Arabidopsis* steering subcommittee for proteomics assembles the largest proteome database resource for plant systems biology. *J Proteome Res* 7:4209–4210
43. Lamesch P, Berardini TZ, Li D et al (2011) The *Arabidopsis* Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res* 40(D1):D1202–D1210. doi:[10.1093/nar/gkr1090](https://doi.org/10.1093/nar/gkr1090)
44. Walhout AJ, Sordella R, Lu X et al (2000) Protein interaction mapping in *C. elegans*

- using proteins involved in vulval development. *Science* 287:116–122
45. Ohyanagi H, Tanaka T, Sakai H et al (2006) The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. *japonica* genome information. *Nucleic Acids Res* 34:D741–D744
46. Doroshenk KA, Crofts AJ, Morris RT et al (2012) RiceRBP: a resource for experimentally identified RNA binding proteins in *Oryza sativa*. *Front Plant Sci* 3:90
47. Usadel B, Schwacke R, Nagel A et al (2012) GabiPD: the GABI Primary Database integrates plant proteomic data with gene-centric information. *Front Plant Sci* 3:154
48. Rose CM, Venkateshwaran M, Grimsrud PA et al (2012) Medicago PhosphoProtein Database: a repository for *Medicago truncatula* phosphoprotein data. *Front Plant Sci* 3:122
49. Yao Q, Bollinger C, Gao J et al (2012) P3DB: an integrated database for plant protein phosphorylation. *Front Plant Sci* 3:206
50. Wolski W, Lalowski M, Martus P et al (2005) Transformation and other factors of the peptide mass spectrometry pairwise peak-list comparison process. *BMC Bioinformatics* 6:285
51. Galland M, Job D, Rajjou L (2012) The seed proteome web portal. *Front Plant Sci* 3:98
52. Mostaguir K, Hoogland C, Binz P-A et al (2003) The Make 2D-DB II package: conversion of federated two-dimensional gel electrophoresis databases into a relational format and interconnection of distributed databases. *Proteomics* 3:1441–1444
53. Joshi T, Yao Q, Franklin LD et al (2010) SoyMetDB: the soybean metabolome database. *Proceedings of IEEE International Conference on Bioinformatics & Biomedicine (BIBM 2010)*, Hong Kong, pp 203–208
54. Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19:242–247
55. O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 250:4007–4021
56. Komatsu S, Wada T, Abaléa Y et al (2009) Analysis of plasma membrane proteome in soybean and application to flooding stress response. *J Proteome Res* 8:4487–4499
57. Nouri M-Z, Komatsu S (2010) Comparative analysis of soybean plasma membrane proteins under osmotic stress using gel-based and LC MS/MS-based proteomics approaches. *Proteomics* 10:1930–1945
58. Griss J, Cote RG, Gerner C et al (2011) Published and perished? The influence of the searched protein database on the long-term storage of proteomics data. *Mol Cell Proteomics* 10:M111.008490
59. Foster JM, Degroeve S, Gatto L et al (2011) A posteriori quality control for the curation and reuse of public proteomics data. *Proteomics* 11:2182–2194
60. Wienkoop S, Staudinger C, Hoehenwarter W et al (2012) ProMEX: a mass spectral reference database for plant proteomics. *Front Plant Sci* 3:125

Chapter 4

How to Use 2D Gel Electrophoresis in Plant Proteomics

Thierry Rabilloud

Abstract

Two-dimensional electrophoresis has nurtured the birth of proteomics. It is however no longer the exclusive setup used in proteomics, with the development of shotgun proteomics techniques that appear more fancy and fashionable nowadays.

Nevertheless, 2D gel-based proteomics still has valuable features, and sometimes unique ones, which make it often an attractive choice when a proteomics strategy must be selected. These features are detailed in this chapter, as is the rationale for selecting or not 2D gel-based proteomics as a proteomic strategy.

Key words Two-dimensional electrophoresis, Proteomics, Blotting, Posttranslational modifications

1 Introduction

At the beginning of the twenty-first century, it is fashionable to describe 2D gel-based proteomics as an outdated and poorly efficient technique. This feeling is mainly due to the fascination of many scientists in proteomics for big lists, and in this case the well-documented undersampling of 2D gel-based proteomics is dearly resented, as well as its almost complete inability to analyze transmembrane proteins [1, 2], despite of lot of effort devoted to this particular issue (e.g. in [3, 4] for plant samples). This undersampling has been documented mostly on animal samples [5, 6], but it is obvious that it will take place on plant samples as well.

In fact, all proteomic techniques undersample complex samples by a ratio of 1:10 (1,000 proteins analyzed on 2D gels among the 10,000 present in a cell, 20,000 peptides analyzed by shotgun proteomics among the 200,000 derived from a cell sample). However, as 2D gel-based proteomics is the only setup that analyzes complete proteins, the consequences of undersampling are clear and heavy. If a protein is not seen, it is lost forever. In addition, the undersampling rules in 2D gel-based proteomics are well known. Hydrophobic proteins, high molecular weight proteins, and rare proteins do not show up.

In shotgun techniques, the undersampling rules are also known. However, they apply to peptides, and how they translate in terms of protein identification and even more in terms of protein quantification is complex, and linked to the complexity of the protein inference problem [7–9].

On top of this strong undersampling problem comes the often unrecognized fact that 2D gel-based proteomics is a low-yield process [10, 11] and the fact that plant samples are among the most difficult samples to prepare for an analysis by 2D gels [12], due to the variety and amounts of interfering, nonprotein compounds present in many plant cell types and tissues.

All these negative sides explain the trend from 2D gel-based proteomics to shotgun proteomics, as exemplified by the study of cadmium response in plant; gel-based in 2006 [13] and shotgun-based in 2011 [14].

However, 2D gel-based proteomics is still very popular in plant proteomics, as shown by recent publications in the field (e.g. [15, 16], but among more than 150 publications in 2011). There are several built-in reasons for this sustained interest in this technique, which are named reproducibility, robustness, efficiency, ability to analyze intact proteins, adequacy to study posttranslational modifications, easy interface with powerful biochemical techniques. These different points will be developed below.

2 Reproducibility and Robustness

It is not an overstatement to write that 2D gel-based proteomics is still, and by far, the most reproducible and robust proteomic setup [17]. A further proof of this statement lies in the much higher expectations from proteomic journals, in terms of sample numbers and experimental power, applied to 2D gel-based proteomics compared to shotgun proteomics [18, 19]. Still today, what is common practice in 2D gel-based proteomics, i.e. quantitative comparison of two or more biological conditions with several individual biological replicates for each condition, is still rarely found in publications using shotgun proteomics.

This reproducibility is further testified by the interlaboratory reproducibility of 2D maps [20, 21], while recent test of shotgun techniques on the same sample (yeast) revealed a much higher standard deviation between laboratories [22, 23].

Ironically enough, the high robustness of 2D gel-based proteomics comes in part from the terrible reproducibility of isoelectric focusing in its early days. Put bluntly, only the gels run simultaneously were comparable. This led the proteomic pioneers to develop devices to run several gels in parallel [24, 25]. When immobilized pH gradients were finally introduced in 2D gel electrophoresis with solid protocols [26], the reproducibility and

robustness of the technique made a quantum leap that is still enjoyed today. Moreover, this parallelization of the technique is also a very strong advantage when multiple comparisons, and not only binary ones, must be performed [27].

3 Efficiency of 2D Gel-Based Proteomics

Nowadays, Edman sequencing for protein identification has been almost completely replaced by mass spectrometry-based protein identification, which is much more sensitive and much faster. However, mass spectrometers are expensive and delicate instruments, and the machine time is often considered as precious.

In this frame, shotgun proteomic techniques often represent a waste of machine time. In many comparative experiments, which represent most of the proteomic experiments, most of the proteins do not change between the various conditions tested. Thus, the mass spectrometer is analyzing again and again a very high number of peptides from proteins that have no interest in the biological question studied.

Oppositely, 2D gel-based proteomics makes an optimal use of mass spectrometer time, because most of what is carried out within the mass spectrometer in shotgun proteomics is carried out upfront at the 2D gel stage in 2D gel-based proteomics. In fact, the modern gel staining techniques [28–30], allow to perform a sensitive and linear detection of the protein spots on the gels. Then, image analysis [31, 32] coupled with statistical analysis [33], allow to perform a quantitative analysis of all protein spots and to determine which spots are worth identifying. These spots are then the only ones that need to be processed by the mass spectrometer.

When financial resources are scarce, which is often the case in academic research, this allows to build a “hub and spokes” model, where a central hub hosting the mass spectrometry platform is able to serve very efficiently several biology-oriented laboratories, each using 2D gel electrophoresis and image analysis to select the relevant proteins within their respective research projects.

With plant proteomics usually less funded than medicine-oriented proteomics, and this situation even harder for the proteomics of traditional crops, this efficiency factor cannot be neglected in the research landscape.

4 Interface with Biochemical Techniques

Within this section, the focus will be shifted from classical, quantitative proteomics, where the name of the game is to find differences in expression on total extracts, to more focused proteomics experiments where the interest is centered on subclasses of proteins. In many

cases, the protein subclasses are defined by a common biochemical feature, and quite often this biochemical feature is a posttranslational modification, for which an affinity reagent exists.

When such an affinity reagent exists, two types of strategies can be devised to perform proteomic studies targeted for the proteins recognized by the affinity reagent.

In the first strategy, the affinity reagent is used as a pre-proteomic preparative tool, aiming at selecting the proteins or peptides of interest in the complex starting extract. The selected proteins/peptides are then analyzed by a classical proteomic setup.

This strategy is the only one that can be used with shotgun proteomics, but it can also be used with 2D gel-based proteomics as well.

In the second strategy, the affinity reagent is used as an analytical reagent, after the separation stage in the proteomic setup, in order to detect selectively the proteins of interest. This strategy works extremely well with 2D gel-based proteomics, thanks to the easy interface provided by the blotting process between 2D gels and affinity reagents.

Thus, the final performance of the whole process will largely depend on the efficiency of the affinity reagent under the preparative and analytical setups, and it must be kept in mind that for several reasons, the affinity reagent is always used in a large excess over the biological extract.

Then two major cases can be distinguished:

1. The affinity reagent is not a protein

Typical examples of that configuration are represented by boronic acid for sugars and either IMAC or metal oxides for phosphate. These nonproteinaceous reagents are best used in the preparative mode, and work wonderfully with shotgun proteomics [34].

2. The affinity reagent is a protein

In this case, the suitability of the reagent as a preparative tool will depend not only on its selectivity, but also on the existence of mild elution conditions that are able to elute all the analytes without eluting out the affinity reagent.

Lectins represent such a happy case and have been used as a preparative reagent in 2D gel-based proteomics [35] and in shotgun proteomics [36–38].

They have also been used as a post-2D gel selective detection reagent [39].

Conversely, antibodies represent a case where elution without polluting the extract with antibodies is very difficult. They are therefore much better used as analytical reagents on blots rather than preparative reagents. Thus, 2D gel-based proteomics is a very efficient strategy to use in conjunction with antibodies, and this strategy has

been used to identify several posttranslational modifications, such as citrullinated proteins[40], carbonylated proteins [41], proteins containing nitrotyrosine [42, 43] or hydroxynonenal adducts [44, 45].

Although very successful, this strategy has a major caveat, i.e. the fact that a minor and modified spot detected with the antibody may comigrate with a major, unrelated spot which will be the one identified by the downstream mass spectrometry process, resulting in a major mistake. The severity of this problem increases with spot crowding and thus with the complexity of the biological extract. There is however an easy counter, at least for acidic and neutral proteins, which is the use of narrow range pH gradients that greatly decrease the probability of comigration. Such narrow pH range gels have been used under a variety of circumstances [46–49], and have proven very efficient, especially when used in conjunction with prefractionation techniques [50].

5 Analysis of Intact Proteins and of Posttranslational Modifications

The more proteomics progresses, the more it stresses the importance of posttranslational modifications in biology (e.g. in [51, 52]). However, not all posttranslational modifications can be easily studied through the use of an affinity reagent, and some, as methylations, are not easily detected except on abundant proteins [53].

For the study of these difficult, and sometimes unknown, modifications, 2D gel-based proteomics is a platform of choice, either through selective and transient labeling [54], or by using the fact that many modifications induce a pH shift and thus a separation of the modified protein from the bulk of the unmodified one. This can be used in the field of protein cleavage [55, 56], but also as a preparative tool to isolate the modified spot, thereby leading to an easier identification of the modifications. Such an approach has been used for the identification of thiol oxidations [57–59].

More generally, the separating ability of 2D gels, coupled with their ability to be used as a micropreparative tool [60], now allow to set very comprehensive maps of protein modifications [61, 62]. In this area, the use of ultra narrow pH gradient [47], with their almost infinite resolving power [63, 64] should be extremely useful, by providing a very thorough deconvolution of the protein spots into individual protein species.

6 Conclusions

Nowadays, 2D gel-based proteomics is usually selected on the basis of its robustness and relatively low cost, among the proteomic techniques. It can be reasonably anticipated, however, that the uses directed toward the identification of posttranslational modifications,

and especially the determination of the real combination of modifications present on a single protein species, will develop considerably in the future.

References

1. Santoni V, Molloy M, Rabilloud T (2000) Membrane proteins and proteomics: un amour impossible? *Electrophoresis* 21:1054–1070
2. Rabilloud T (2009) Membrane proteins and proteomics: love is possible, but so difficult. *Electrophoresis* 30(Suppl 1):S174–S180
3. Santoni V, Rabilloud T, Doumas P et al (1999) Towards the recovery of hydrophobic proteins on two-dimensional electrophoresis gels. *Electrophoresis* 20:705–711
4. Santoni V, Kieffer S, Desclaux D et al (2000) Membrane proteomics: use of additive main effects with multiplicative interaction model to classify plasma membrane proteins according to their solubility and electrophoretic properties. *Electrophoresis* 21:3329–3344
5. Petrank J, Ivanek R, Toman O et al (2008) Deja vu in proteomics. A hit parade of repeatedly identified differentially expressed proteins. *Proteomics* 8:1744–1749
6. Wang P, Bouwman FG, Mariman EC (2009) Generally detected proteins in comparative proteomics—a matter of cellular stress response? *Proteomics* 9:2955–2966
7. Nesvizhskii AI, Aebersold R (2005) Interpretation of shotgun proteomic data: the protein inference problem. *Mol Cell Proteomics* 4:1419–1440
8. Li YF, Arnold RJ, Li Y et al (2009) A bayesian approach to protein inference problem in shotgun proteomics. *J Comput Biol* 16:1183–1193
9. Grobri MA, Qeli E, Brunner E et al (2009) Deterministic protein inference for shotgun proteomics data provides new insights into *Arabidopsis* pollen development and function. *Genome Res* 19:1786–1800
10. Choe LH, Lee KH (2003) Quantitative and qualitative measure of intralaboratory two-dimensional protein gel reproducibility and the effects of sample preparation, sample load, and image analysis. *Electrophoresis* 24:3500–3507
11. Zhou SB, Bailey MJ, Dunn MJ et al (2005) A quantitative investigation into the losses of proteins at different stages of a two-dimensional gel electrophoresis procedure. *Proteomics* 5:2739–2747
12. Saravanan RS, Rose JK (2004) A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics* 4:2522–2532
13. Sarry JE, Kuhn L, Ducruix C et al (2006) The early responses of *Arabidopsis thaliana* cells to cadmium exposure explored by protein and metabolite profiling analyses. *Proteomics* 6:2180–2198
14. Villiers F, Ducruix C, Hugouevoux V et al (2011) Investigating the plant response to cadmium exposure by proteomic and metabolomic approaches. *Proteomics* 11:1650–1663
15. Tanou G, Job C, Belghazi M et al (2011) Proteomic signatures uncover hydrogen peroxide and nitric oxide cross-talk signaling network in citrus plants. *J Proteome Res* 9:5994–6006
16. Catusse J, Meinhard J, Job C et al (2011) Proteomics reveals potential biomarkers of seed vigor in sugarbeet. *Proteomics* 11:1569–1580
17. Hackett M (2008) Science, marketing and wishful thinking in quantitative proteomics. *Proteomics* 8:4618–4623
18. Celis JE (2004) Gel-based proteomics: what does MCP expect? *Mol Cell Proteomics* 3:949
19. Wilkins MR, Appel RD, Van Eyk JE et al (2006) Guidelines for the next 10 years of proteomics. *Proteomics* 6:4–8
20. Corbett JM, Dunn MJ, Posch A et al (1994) Positional reproducibility of protein spots in 2-dimensional polyacrylamide-gel electrophoresis using immobilized pH gradient isoelectric-focusing in the first dimension—an interlaboratory comparison. *Electrophoresis* 15:1205–1211
21. Blomberg A, Blomberg L, Norbeck J et al (1995) Interlaboratory reproducibility of yeast protein-patterns analyzed by immobilized Ph gradient 2-dimensional gel-electrophoresis. *Electrophoresis* 16:1935–1945
22. Paulovich AG, Billheimer D, Ham AJ et al (2010) Interlaboratory study characterizing a yeast performance standard for benchmarking LC-MS platform performance. *Mol Cell Proteomics* 9:242–254
23. Tabb DL, Vega-Montoto L, Rudnick PA et al (2010) Repeatability and reproducibility in

- proteomic identifications by liquid chromatography-tandem mass spectrometry. *J Proteome Res* 9:761–776
- 24. Anderson NG, Anderson NL (1978) Analytical techniques for cell fractions. 21. 2-Dimensional analysis of serum and tissue proteins—multiple isoelectric-focusing. *Anal Biochem* 85:331–340
 - 25. Anderson NL, Anderson NG (1978) Analytical techniques for cell fractions .22. 2-Dimensional analysis of serum and tissue proteins—multiple gradient-slab gel-electrophoresis. *Anal Biochem* 85:341–354
 - 26. Gorg A, Postel W, Gunther S (1988) Methodology of Ipg-Dalt for the analysis of cell lysates and tissue proteins. *Electrophoresis* 9:628
 - 27. Natera SHA, Guerreiro N, Djordjevic MA (2000) Proteome analysis of differentially displayed proteins as a tool for the investigation of symbiosis. *Mol Plant Microbe Interact* 13:995–1009
 - 28. Miller I, Crawford J, Gianazza E (2006) Protein stains for proteomic applications: which, when, why? *Proteomics* 6:5385–5408
 - 29. Patton WF (2000) A thousand points of light: the application of fluorescence detection technologies to two-dimensional gel electrophoresis and proteomics. *Electrophoresis* 21:1123–1144
 - 30. Minden JS, Dowd SR, Meyer HE et al (2009) Difference gel electrophoresis. *Electrophoresis* 30:S156–S161
 - 31. Bandow JE, Baker JD, Berth M et al (2008) Improved image analysis workflow for 2-D gels enables large-scale 2-D gel-based proteomics studies—COPD biomarker discovery study. *Proteomics* 8:3030–3041
 - 32. Stessl M, Noe CR, Lachmann B (2009) Influence of image-analysis software on quantitation of two-dimensional gel electrophoresis data. *Electrophoresis* 30:325–328
 - 33. Karp NA, McCormick PS, Russell MR et al (2007) Experimental and statistical considerations to avoid false conclusions in proteomics studies using differential in-gel electrophoresis. *Mol Cell Proteomics* 6:1354–1364
 - 34. Thingholm TE, Jensen ON, Larsen MR (2009) Analytical strategies for phosphoproteomics. *Proteomics* 9:1451–1468
 - 35. Robotti A, Natale M, Albo AG et al (2010) Acute-phase proteins investigation based on lectins affinity capture prior to 2-DE separation: application to serum from multiple sclerosis patients. *Electrophoresis* 31:2882–2893
 - 36. Ligat L, Lauber E, Albenne C et al (2011) Analysis of the xylem sap proteome of *Brassica oleracea* reveals a high content in secreted proteins. *Proteomics* 11:1798–1813
 - 37. Catala C, Howe KJ, Hucko S et al (2011) Towards characterization of the glycoproteome of tomato (*Solanum lycopersicum*) fruit using Concanavalin A lectin affinity chromatography and LC-MALDI-MS/MS analysis. *Proteomics* 11:1530–1544
 - 38. Zhang Y, Giboulot A, Zivy M et al (2011) Combining various strategies to increase the coverage of the plant cell wall glycoproteome. *Phytochemistry* 72:1109–1123
 - 39. Komatsu S, Yamada E, Furukawa K (2009) Cold stress changes the concanavalin A-positive glycosylation pattern of proteins expressed in the basal parts of rice leaf sheaths. *Amino Acids* 36:115–123
 - 40. Ishigami A, Ohsawa T, Hiratsuka M et al (2005) Abnormal accumulation of citrullinated proteins catalyzed by peptidylarginine deiminase in hippocampal extracts from patients with Alzheimer's disease. *J Neurosci Res* 80:120–128
 - 41. Job C, Rajjou L, Lovigny Y et al (2005) Patterns of protein oxidation in *Arabidopsis* seeds and during germination. *Plant Physiol* 138:790–802
 - 42. Casteagna A, Thongboonkerd V, Klein JB et al (2003) Proteomic identification of nitrated proteins in Alzheimer's disease brain. *J Neurochem* 85:1394–1401
 - 43. Kanski J, Behring A, Pelling J et al (2005) Proteomic identification of 3-nitrotyrosine-containing rat cardiac proteins: effects of biological aging. *Am J Physiol Heart Circ Physiol* 288:H371–H381
 - 44. Perluigi M, Fai PH, Hensley K et al (2005) Proteomic analysis of 4-hydroxy-2-nonenal-modified proteins in G93A-SOD1 transgenic mice—a model of familial amyotrophic lateral sclerosis. *Free Radic Biol Med* 38:960–968
 - 45. Reed TT, Pierce WM, Markesberry WR et al (2009) Proteomic identification of HNE-bound proteins in early Alzheimer disease: insights into the role of lipid peroxidation in the progression of AD. *Brain Res* 1274:66–76
 - 46. Parker KC, Garrels JI, Hines W et al (1998) Identification of yeast proteins from two-dimensional gels: working out spot cross-contamination. *Electrophoresis* 19:1920–1932
 - 47. Westbrook JA, Yan JX, Wait R et al (2001) Zooming-in on the proteome: very narrow-range immobilised pH gradients reveal more protein species and isoforms. *Electrophoresis* 22:2865–2871
 - 48. Sarma AD, Emerich DW (2006) A comparative proteomic evaluation of culture grown vs nodule isolated *Bradyrhizobium japonicum*. *Proteomics* 6:3008–3028

49. Zhao Z, Zhang W, Stanley BA et al (2008) Functional proteomics of *Arabidopsis thaliana* guard cells uncovers new stomatal signaling pathways. *Plant Cell* 20:3210–3226
50. Righetti PG, Castagna A, Herbert B et al (2005) How to bring the “unseen” proteome to the limelight via electrophoretic pre-fractionation techniques. *Biosci Rep* 25:3–17
51. Rajjou L, Belghazi M, Catusse J et al (2011) Proteomics and posttranslational proteomics of seed dormancy and germination. *Methods Mol Biol* 773:215–236
52. Arc E, Galland M, Cueff G et al (2011) Reboot the system thanks to protein post-translational modifications and proteome diversity: how quiescent seeds restart their metabolism to prepare seedling establishment. *Proteomics* 11:1606–1618
53. Santoni V, Verdoucq L, Sommerer N et al (2006) Methylation of aquaporins in plant plasma membrane. *Biochem J* 400:189–197
54. Fratelli M, Demol H, Puype M et al (2002) Identification by redox proteomics of glutathionylated proteins in oxidatively stressed human T lymphocytes. *Proc Natl Acad Sci USA* 99:3505–3510
55. Lee AY, Park BC, Jang M et al (2004) Identification of caspase-3 degradome by two-dimensional gel electrophoresis and matrix-assisted laser desorption/ionization-time of flight analysis. *Proteomics* 4:3429–3436
56. Jang M, Park BC, Kang S et al (2008) Mining of caspase-7 substrates using a degradome approach. *Mol Cells* 26:152–157
57. Rabilloud T, Heller M, Gasnier F et al (2002) Proteomics analysis of cellular response to oxidative stress—evidence for in vivo overoxidation of peroxiredoxins at their active site. *J Biol Chem* 277:19396–19401
58. Weber H, Engelmann S, Becher D et al (2004) Oxidative stress triggers thiol oxidation in the glyceraldehyde-3-phosphate dehydrogenase of *Staphylococcus aureus*. *Mol Microbiol* 52:133–140
59. Hwang NR, Yim SH, Kim YM et al (2009) Oxidative modifications of glyceraldehyde-3-phosphate dehydrogenase play a key role in its multiple cellular functions. *Biochem J* 423:253–264
60. Hanash SM, Strahler JR, Neel JV et al (1991) Highly resolving 2-dimensional gels for protein sequencing. *Proc Natl Acad Sci USA* 88:5709–5713
61. John JP, Pollak A, Lubec G (2009) Complete sequencing and oxidative modification of manganese superoxide in medulloblastoma cells. *Electrophoresis* 30:3006–3016
62. Seo J, Jeong J, Kim YM et al (2008) Strategy for comprehensive identification of post-cellular proteins, including low abundant modifications: glyceraldehyde-3-phosphate dehydrogenase. *J Proteome Res* 7:587–602
63. Bosisio AB, Rochette J, Wajcman H et al (1985) Electrophoretic and chromatographic techniques for the differential diagnosis of a haemoglobin abnormality: Hb E heterozygosity. *J Chromatogr* 330:299–306
64. Cossu G, Righetti PG (1987) Resolution of G gamma and A gamma foetal haemoglobin tetramers in immobilized pH gradients. *J Chromatogr* 398:211–216

Chapter 5

Standardization of Data Processing and Statistical Analysis in Comparative Plant Proteomics Experiment

Luis Valledor, M. Cristina Romero-Rodríguez, and Jesus V. Jorrin-Novo

Abstract

Two-dimensional gel electrophoresis remains the most widely used technique for protein separation in plant proteomics experiments. Despite the continuous technical advances and improvements in current 2-DE protocols, an adequate and correct experimental design and statistical analysis of the data tend to be ignored or not properly documented in current literature. Both proper experimental design and appropriate statistical analysis are requested in order to confidently discuss our results and to conclude from experimental data.

In this chapter, we describe a model procedure for a correct experimental design and a complete statistical analysis of proteomic dataset. Our model procedure covers all of the steps in data mining and processing, starting with the data preprocessing (transformation, missing value imputation, definition of outliers) and univariate statistics (parametric and nonparametric tests), and finishing with multivariate statistics (clustering, heat-mapping, PCA, ICA, PLS-DA).

Key words 2-DE-based proteomics, Experimental design, Data mining, Univariate statistics, Multivariate statistics, Systems biology

1 Introduction

Comparative or quantitative proteomics is the most utilized subarea of proteomics [1, 2]. It aims at establishing differences in protein profiles between two samples from different individuals or from distinct treatments, defined both from a quantitative and qualitative point of view. In most of the cases the final objective of these methodologies is to differentially identify abundant protein species in order to explain the biological characteristics of each system. Two-DE continues, by far, as the most frequently used tool in plant proteomics among the current proteomic platforms, while gel-free LC-MS methods are still used in a minimum percentage of the papers published. Even so, and despite the continuous technical

advances and improvements in current 2-DE protocols, an adequate and correct experimental design and statistical analysis of the data tend to be ignored or not properly documented in current literature. The design of an experiment and its practical carryover are important for increased robustness, but how the data are analyzed is paramount for maximizing the information obtained [3]. Most of the software packages for image analysis or LC-MS processing embed some statistical tools and they can be considered adequate for most of the researchers' purposes. These packages, however, have several limitations, namely: (1) they do not rationalize decisions when there are gel or chromatographic imperfections; (2) spots/peaks are missing or mismatched, or other artifacts are present, and (3) software-induced variation is not considered most of the times [3, 4].

Traditionally, quantitative proteomics data are assessed by univariate statistics, such as ANOVA. Nevertheless, they are negatively affected by the raw structure of the data, they cannot detect trends or groups, and the possibilities or false positives increase. On the other hand, multivariate analyses, such as PCA, ICA, or PLS, are described to be more effective describing trends and reducing the complexity of the data, since they are less affected by the structure of the data [5]. The combination of both univariate and

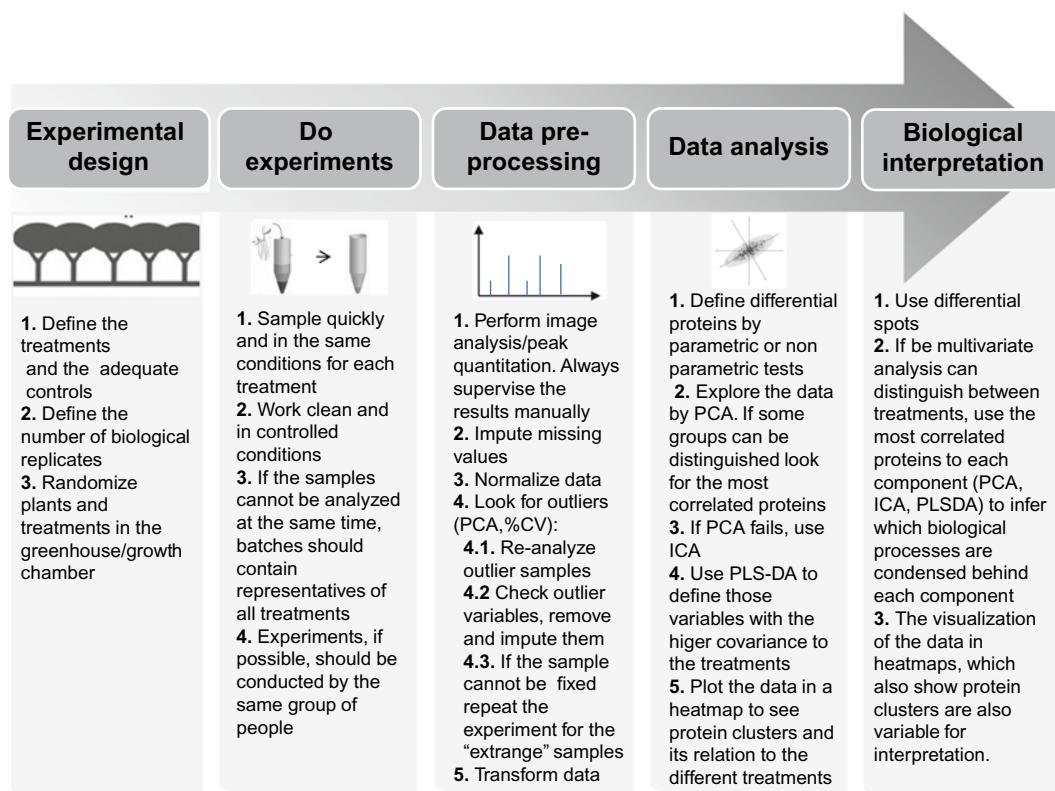


Fig. 1 Model workflow for a complete statistical analysis of proteomic datasets

multivariate approaches provides a comprehensive overview of the data, with single protein studies and multiprotein trends, maximizing the information obtained from most of the datasets.

In this chapter, we provide a model workflow for a complete statistical analysis of proteomic datasets (Fig. 1), starting with some guidelines for image analysis and peptide identification, and following with univariate and multivariate statistics.

2 Materials

Despite the numerous and specific image capture and processing software tools [6], most of them present a limited capability for statistical analyses. The adequate data preprocessing, the presence of missing spots or even the fact that they do not implement error control procedures are some of their limitations. Some authors have used different software packages over the same images, with a difference of up to 50 % in the quantitatively differential spots [3]. To avoid software-induced biases and opaque processing steps, we recommend a dedicated effort in data processing. Here, we present a statistical workflow for maximizing the information obtained by quantitative proteomics. All the analysis described in this chapter can be conducted with any existing statistical package that contains the functions described below. We recommend, however, R, which is a free software environment for statistical computing and graphics [7] (<http://www.r-project.org/>); and/or COVAIN, a user-friendly statistical tool [8].

3 Methods

3.1 Experimental Design

The main limitation of –omics analyses is the reduced number of replicates that are usually performed, which is always limited by the economical and processing capacities of the laboratory. Since only a low number of analyses can be done, biological replication is always better than technical replication. Moreover, it is not recommended to carry out a technical replication if it occurs at the expense of biological replication [9]. To achieve a good statistical power, the recommended number of biological replicates is between 5 and 7, despite statistics can be performed from three biological replicates [10].

3.2 Data Preprocessing

3.2.1 2-DE Electrophoresis Image Acquisition (Spot Abundance)

1. Images should be obtained from a calibrated densitometer/fluorimeter with enough exposure time to capture faint spots without saturating most abundant spots, and with the maximum reasonable resolution. Guidelines for conducting 2-DE and gel staining have been amply described in a number of chapters in this book, as well as in its previous edition.

2. After the use of the image analysis software, the matching and boundaries of all the spots should be manually checked to ensure correctness. The definition of a spot (intensity changes over the background) is flexible, but it should be maintained across all the experimental gels.
3. Normalize spot abundance (SA) values according to the total spot abundance in each gel.

$$(NSA)_k = \frac{(SA)_k}{\sum_{i=1}^N (SA)}$$

The Normalized Spot Abundance (NSA) from spot k of gel i is obtained after the division of the spot abundance by the sum of SA of all N spots present in gel i . This procedure requires a complete match of the gels (*see Note*).

3.3 Missing Value Estimation

Proteomic studies, especially those based on 2-DE, tend to show missing values. For estimating the missing values, one of the better approaches is the use of an imputation algorithm such as k -nearest neighbors. Briefly, this method consists of finding the k nearest neighbors to the variable with missing data, and fills the gaps by taking the mean of the missing variable among the k neighbors. It is not advisable to use this method for the imputation in more than 20 % of missing values (i.e. one missing value in a set of five replicates). It is also recommended to use the whole dataset, including all the treatments, to perform the missing value imputation. In R there are some packages with this algorithm like {impute}, which also provides SVD and SVT (Singular Value Decomposition (SVD) and Singular Value Thresholding (SVT)), two imputation algorithms and tools for determining the optimal k .

3.4 Data Filtering

It is advisable to expend some time in analyzing the data quality, finding and filtering (or reanalyzing) the outlier values. This will allow the removal of artifacts, technical biases, and it will also increase the statistical power of the subsequent analyses.

In most of the cases, the removal of outliers or single hits (proteins/spots that can be only detected in one sample) reduces the noise of the sample, allowing better understandings and more complete analyses. A good practice can be the definition of *consistency* rules for further consideration of spots/proteins (variables from here on). For considering a particular variable, it has to be present in at least 75 % of the replicates of at least one treatment. This will discard mostly noise-artifacts and proteins/peptides near to the detection limit. Methods to determine outlier samples are described below.

In this sense variables with lower values usually show the higher coefficient of variation (CV) due to the presence of outliers. It

is advisable to plot CV after missing value imputation and data transformation in order to check outliers within the replicates and remove them from analysis.

$$CV_k = \left(\frac{\sigma}{\bar{x}} \right) \times 100$$

In the case of 2-DE gels, all of the outliers/non consistent spots should be manually supervised in the image analysis software before removing it, since erroneous or mismatched spots are not unusual.

3.5 Data Transformation

The large proteomic datasets usually shows a linear dependence between mean and standard deviation, as a consequence of the digital acquisition process. Furthermore, in almost all cases, the low number of replicates led to a nonnormal distribution of the values of each variable.

A simple log 10 or cubic root transformations performs adequately for controlling the mean/SD dependence and also for normalizing the variables in most of the cases.

3.6 Univariate Statistics

The use of this group of statistical tests, which compares the variables one by one, represent the classical approach for distinguishing the possible changes in the protein abundance as a consequence of the application of the different treatments. These tests are suitable to define differential proteins, biomarkers, etc. but lack the capacity of distinguishing between groups, multi-variable trends or hidden patterns as multivariate statistics do (see below).

3.6.1 Normality and Homocedasticity

The normality and homogeneity of variance (homocedasticity) of the samples should be tested prior to applying univariate statistics. Shapiro-Wilk test for nonnormality can be applied on R statistical software using the shapiro.test{stats} function. Bartlett test can be applied for testing if samples are from populations with equal variance. In R this test is available in the bartlett.test{stats} function.

3.6.2 Parametric Tests

1. Student's *t*-test

This procedure is classically used to test the null hypothesis that the means of two normally distributed populations are equal (i.e. if Actin abundance is the same in the control and the treated sample). We can compare only two treatments using this test and no block effects can be estimated. Furthermore, and strictly speaking, this test can only be used if the variances of the two populations are assumed to be equal. If this assumption is dropped the form of this test is called Welch's *t*-test. In R we have the function t.test{stats} that can be used for computing both Student's and Welch's *t*-test.

The application of this test over all of the variables of our datasets will determine which ones are differential between treatments. It is important to remark that in the case of those spots/proteins that are qualitatively different (present only in one treatment) this test is not necessary applied. The number of times that this test is applied, the number of analyzed variables, should be considered for controlling the error of multiple comparisons (see below).

2. ANOVA

The analysis of variance (ANOVA) is a collection of statistical models and their associated procedures. The observed variance of a particular variable is divided into components, which are attributed to different sources of variation. The one-way ANOVA, the simplest form of ANOVA, is a generalization of the *t*-test for more than two groups. More complex designs of ANOVA can include more than one source of variation (called factors), repeated measures, consider block effects, interaction between the sources of variation, etc. The null hypothesis of ANOVA, is that the means of the dependent variables are equal to all of groups and factors. As for the *t*-test, the populations from which the samples were obtained must be normally distributed, independent, homoscedastic and the group must have the same sample size. In R `lm{stats}`, `aov{stats}` and `anova{stats}` can be used.

3. Multiple comparisons

There are procedures, like Tukey's HSD (honestly significant difference) or LSD (least significant differences) tests that can be used generally in conjunction with ANOVA, to find which means are significantly different from one another. These test are using for defining which treatments produce significant differences between them, since the ANOVA only tell us that there are differences between groups, but does not point to these particular differences. For example `TukeyHSD{stats}` and `LSD.test{agricolae}` are available in R to perform these tests.

3.6.3 Nonparametric Tests

The samples-datasets of proteomic data not always are homocedastic and normally distributed. This can be partly explained by the nature of these data and the low number of replicates that are considered. In these situations the use of nonparametric tests is mandatory. These tests are based on the median, rather on the mean, so they are less sensitive to outliers than parametric tests. If the number of replicates is high (for -omics studies), with a minimum of 6–7 biological replicates per treatment, nonparametric can perform in a better way than parametric tests for proteomics data [11]. Each parametric test has a nonparametric equivalent, like Mann–Whitney *U* (*t*-test), Wilcoxon signed-rank test (paired

t-test), Kruskal Wallis (one-way ANOVA) or Friedman test (one-way repeated measures ANOVA). These tests can also be easily performed in R using `wilcox.test{stats}`, `kruskal.test{stats}`, or `friedman.test{stats}` respectively.

3.7 Controlling Statistical Error of Multiple Comparisons

When multiple hypotheses are tested as a result of multiple comparisons a method for controlling the error should be implemented in our workflow. For example in a standard proteomic analysis of about 800 variables (spots, identified proteins, etc.) in which univariate statistics have been performed with a significance of 95 % ($p=0.05$), 40 false positives can be expected ($800 \times 0.05 = 40$). For controlling the error different solutions can be applied: the family-wise error rate (FWER), such as the Bonferroni correction, or False Discovery Rate (FDR) procedures [12]. FWER procedures are very conservative because they only control the probability of false rejections of the null hypothesis (false positives), at the cost of increasing the number of false negatives. These methods are adequate when looking for a small number of strong biomarkers. On the other hand the FDR controls the expected proportion of incorrectly rejected null hypotheses (i.e. 1 %, 5 %), being in consequence more powerful than FWER procedures but at the cost of increasing the likelihood of obtaining type I errors between the rejected hypotheses. Classically after a FDR analysis we will have *p*-values and *q*-values (corrected *p*-value after FDR). To interpret the *q*-values it is necessary to look at the ordered list of *q*-values. For example if we now that 52 spots have a *q*-value less than 0.023 we should expect $52 \times 0.0181 = 0.9412$ false positives until the 52th spot. This is another way to use the *q*-values, order them and know the number of expected false positives until different thresholds, to adapt the cut-off to our experiment (it is not the same the establishment of a specific biomarker than descriptive proteomics). Obviously the use of the *q*-values will not always result in a lower number of false positives, but it gives a more accurate indication of the level of false positives for a given cut-off value.

3.8 Multivariate Statistics

These methods are intended to reduce the complexity of the sample to a minimum by “condensing” the original variation of the samples into a reduced number of elements (process which is called reduction of dimensionality). Multivariate data analysis methods are useful for pinpointing the relevant variables for treatment discrimination by focusing not only on single spot differences, but on the covariance structure between proteins [13]. It is, thereby possible to point out which combination of spots could be valuable to identify and characterize them in more detail or also quickly define outlier samples.

3.8.1 Principal Component Analysis

Principal Component Analysis (PCA) generates new variables, called components (PCs), which condense the variability of the samples. PCs are not correlated between each other and are

extracted from the sample one by one, the first component being the one which explains most of the systematic variation in the data. The plotting of principal components quickly visualizes the structure of the data, helping to find sample clusters and identify outliers. If some groups are distinguished, it is always interesting deep into the analysis and study which variables are most correlated to each principal component for defining the biological processes that are hidden in the data. The correlation of each variable to each component is included into the so called loading matrix. These analyses can be performed in R using `prcomp{stats}`, `princomp{stats}`, and `biplot{stats}`. Here we recommend start the processing of the data with a PCA for quickly define outliers, and later on the analysis, when the outliers are removed, perform a complete PCA analysis defining also the biologically interesting variables.

3.8.2 Independent Component Analysis

In contrast to PCA, Independent Component Analysis (ICA) decomposes an input dataset into components so that each component is statistically as independent from the others as possible. ICA can be used to extract mixed signals from the datasets while reducing the effects of noise or artifacts. ICA proved to be more powerful than PCA and faster and more robust than ANOVA dealing with proteomics data [14, 15]. In R, `{FastICA}` package is recommended.

3.8.3 Partial Least Squares, Discriminant Analysis

This is a multivariate projection-based method that, unlike PCA or ICA, maximizes the covariance between two datasets by seeking for linear combinations of the variables from both sets (these linear combinations are called the latent variable). In a classical partial least squares, discriminant analysis (PLS-DA) the response variable is categorical, indicating the different classes (treatments) of the samples, which are used to solve a wide range of classification/discrimination problems in a supervised way determining which variables shows a higher covariance with the different treatments. `{Mixomics}` package contains a set of tools for performing PCA, PLS and other multivariate tests focused on -omics data [16].

3.8.4 Clustering and Heat Maps

Clustering of expression data is usually done to identify proteins with similar behavior, implying that they are correlated. This exploratory technique has clearly proven valuable, and is complementary to multivariate statistics. The representation of the different pathways and visualization of the integrated data across time series or treatment can improve the data interpretation, being also sometimes helpful to select candidate variables. The use of Pearson's correlation coefficient and Ward's aggregation method is the best clustering strategy for proteomics data, being Euclidean distance and UPGMA another valid strategy [17]. R package `{gplots}` can be used for plotting these graphs.

4 Note

Although this chapter is mainly devoted to 2-DE, a similar consideration can be done for label-free approaches, as follows:

- *Label free LC-MS (ion count).*
Consider only the significant peptide and protein hits for quantitation. A general rule for defining a significant hit cannot be given, since it depend both of the quality of the sample, the employed database, and the precision of the mass spectrometer. For example in a Orbitrap instrument, analyzing samples from a fully sequenced organism, using the specific protein database and the 6-frame translation of the genome, and employing SEQUEST algorithm a significant hit can be defined by: (a) at least two peptides with a XCorr > 1.75 or 2.0 (for ions with charge +2 or +3, respectively); (b) only one peptide with XCorr +0.3 greater than charge state are sufficient [18, 19]. The determination of protein scores for orphan species is tricky and dependent of the number and length of the sequences available in the database [20].
- *Label free based quantitation of LC-MS samples a NSAF* [21].

$$(NSAF)_k = \frac{(PSM / L)_k}{\sum_{i=1}^N (PSM / L)_i}$$

in which the total number spectra counts for the matching peptides from protein k (PSM) was divided by the protein's length (L), then divided by the sum of SpC.

Acknowledgements

L.V.'s work was generously supported by a Marie Curie IEF Grant (FP7-PEOPLE-IEF, European Union).

References

1. Abril N, Gion J-M, Kerner R (2011) Proteomics research on forest trees, the most recalcitrant and orphan plant species. *Phytochemistry* 72:1219–1242
2. Jorrín Novo JV, Maldonado AM, Echevarría-Zomeño S (2009) Second generation proteomic techniques, an appropriate experimental design and data analysis to fulfill MIAPE standards, increase plant proteome coverage and biological knowledge. *J Proteomics* 72: 285–314
3. Valledor L, Jorrín JV (2011) Back to the basics: maximizing the information obtained by quantitative two dimensional gel electrophoresis analyses by an appropriate experimental design and statistical analyses. *J Proteomics* 74:1–18
4. Stessl M, Noe CR, Lachmann B (2009) Influence of image-analysis software on quantitation of two-dimensional gel electrophoresis data. *Electrophoresis* 30:325–328
5. Grove H, Jørgensen BM, Jessen F et al (2008) Combination of statistical approaches for analysis of 2-DE data gives complementary results. *J Proteome Res* 7:5119–5124
6. Berth M, Moser FM, Kolbe M et al (2007) The state of the art in the analysis of two dimen-

- sional gel electrophoresis images. *Appl Microbiol Biotechnol* 76:1223–1243
7. Core Team R (2012) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0
 8. Sun X, Weckwerth W (2012) COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics* 8:S81–S93
 9. Horgan GW (2007) Sample size and replication in 2D gel electrophoresis studies. *J Proteome Res* 6:2884–2887
 10. Valledor L, Castillejo MA, Lenz C et al (2008) Proteomic analysis of *Pinus radiata* needles: 2-DE Map and protein identification by LC/MS/MS and substitution-tolerant database searching. *J Proteome Res* 7:2616–2631
 11. Fay MP, Proschan MA (2009) Wilcoxon–Mann–Whitney or t-test? On assumptions for hypothesis tests and multiple interpretation of decision rules. *Stat Surv* 4:1–39
 12. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 57:289–300
 13. Nedenskov Jensen K, Jessen F, Jørgensen BM (2008) Multivariate data analysis of two-dimensional gel electrophoresis protein patterns from few samples. *J Proteome Res* 7:1288–1296
 14. Safavi H, Correa N, Xiong W et al (2008) Independent component analysis of 2-D electrophoresis gels. *Electrophoresis* 29:4017–4026
 15. Morgenthal K, Wienkoop S, Scholz M et al (2005) Correlative GC-TOF-MS-based metabolite profiling and LC-MS-based protein profiling reveal time-related systemic regulation of metabolite–protein networks and improve pattern recognition for multiple biomarker selection. *Metabolomics* 1:109–121
 16. Lê Cao K-A, Boitard S, Besse P (2011) Sparse PLS Discriminant Analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics* 22:253
 17. Meunier B, Dumas E, Piec I et al (2007) Assessment of hierarchical clustering methodologies for proteomic data mining. *J Proteome Res* 6:358–366
 18. Zhou J-Y, Schepmoes AA, Zhang X et al (2010) Improved LC-MS/MS spectral counting statistics by recovering low-scoring spectra matched to confidently identified peptide sequences. *J Proteome Res* 9:5698–5704
 19. Peng J, Elias JE, Thoreen CC et al (2002) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res* 2:43–50
 20. Valledor L, Recuenco L, Egelhofer V et al (2012) The different proteomes of *Chlamydomonas reinhardtii*. *J Proteomics* 75:5883–5887
 21. Paoletti AC, Parmely TJ, Tomomori-Sato C et al (2006) Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc Natl Acad Sci USA* 103:18928–18933

Chapter 6

The Expanding Universe of Mass Analyzer Configurations for Biological Analysis

Juan J. Calvete

Abstract

Mass spectrometry (MS) is an analytical technique that measures the mass-to-charge ratio of electrically charged gas-phase particles. All mass spectrometers combine ion formation, mass analysis, and ion detection. Although mass analyzers can be regarded as sophisticated devices that manipulate ions in space and time, the rich diversity of possible ways to combine ion separation, focusing, and detection in dynamic mass spectrometers accounts for the large number of instrument designs. A historical perspective of the progress in mass spectrometry that since 1965 until today have contributed to position this technique as an indispensable tool for biological research has been recently addressed by a privileged witness of this golden age of MS (Gelpí J. *Mass Spectrom* 43:419–435, 2008; Gelpí J. *Mass Spectrom* 44:1137–1161, 2008). The aim of this chapter is to highlight the view that the operational principles of mass spectrometry can be understood by a simple mathematical language, and that an understanding of the basic concepts of mass spectrometry is necessary to take the most out of this versatile technique.

Key words Mass spectrometry, Mass spectrometers, Mass analyzers, Hybrid mass spectrometers

1 Introduction

The principle of using electric and magnetic fields to accelerate and establish the trajectories of ions inside the spectrometer according to their mass-to-charge ratio is common to all the different designs [1, 2]. Ion motions in a mass analyzers are governed by the field conditions, and thus understanding how the ion's mass-to-charge ratio can be derived from a mass spectrometric experiment requires a knowledge of the parameters involved, which—as described below—can be summarized in simple equations and diagrams. In addition, since there is no single mass analyzer capable of performing all applications required in proteomic researches, knowledge of the analytical capabilities and performances of the different mass analyzer configurations available in the market, proteomic services, or in the laboratories of colleagues is key to take full advantage of this powerful tool for biological research.

2 Pre-proteomics Mass Analyzers: Basic Concepts and Operating Principles

In 1897, experiments by the British physicist Joseph John Thomson at the Cavendish Laboratory at Cambridge, UK, deflecting a beam of cathode rays (discovered by Eugen Goldstein in 1886) by electric and magnetic fields, led to the discovery of the electron [3]. The electric deflection is given by $\Theta = Fqd/mv^2$, where Θ is the angular electric deflection, F is applied electric intensity, q is the charge of the cathode ray particles, d is the length of the electric plates, m is the mass of the cathode ray particles, and v is the velocity of the cathode ray particles. F , d , and θ were measurable, and Thomson had already calculated the ray's velocity. Therefore by measuring the deflection Thomson could calculate the charge-to-mass ratio (q/m) of the cathode ray particles (the electron). Almost at the same time as Thomson, Wilhem Wien, in Berlin (1899) constructed a device with parallel electric and magnetic fields that separated the positive rays according to their charge-to-mass ratio (q/m) and found, as Thomson did, that they are about 2,000 times lighter than the atoms of hydrogen [4]. Thomson later improved on the work of Wien by reducing the pressure in the discharge tube to create the mass spectrograph. In 1912, Thomson and his research assistant, the Canadian-American physicist Arthur Jeffrey Dempster, channeled a stream of ionized neon through a magnetic and an electric field and measured its deflection by placing a photographic plate in its path [5]. The separation of neon isotopes by their mass was the first example of mass spectrometry. Subsequently, A.J. Dempster (1918) [6] and the British chemist and physicist Francis William Aston (1919) [7, 8] established the basic theory and design of mass spectrometers that is still used to this day. The first sector field mass spectrometer was the result of these fundamental physics breakthroughs. Thomson and Aston were honored for their achievements, which laid the foundation of mass spectroscopy, and received Nobel Prizes in Physics and Chemistry in 1906 and 1922, respectively. Wilhem Wien had been awarded the 1911 Nobel Prize for his work on heat radiation.

The fundamental equation of all mass spectrometric techniques that governs the dynamics of charged particles in electric and magnetic fields in vacuum, as operating in a sector instrument, is the Lorentz force law: $F = q(E + vB)$, where E is the electric field strength, B is the magnetic field induction, q is the charge of the particle ($q = ze$, z =charge number and e =elementary charge; -1 for the electron, +1 for the proton), and v is its current velocity (expressed as a vector). Equating the Lorentz force with the centripetal force ($F = mv^2/r$), which is perpendicular to both B and v and keeps the object moving in a circle, gives $qvB = mv^2/r$, where m is the mass of the ion and r the radius of the ion trajectory. Since the speed of the ion is related to its accelerating voltage V by $1/2mv^2 = qV$, the mass-to-charge ratio (m/q) equals $B^2r^2/2V$. Therefore, for a fixed radius of curvature, an ion with a particular m/q can be isolated and

measured by an appropriate combination of magnetic field and accelerating voltage. Figure 1 displays an example of derivation of the m/z ratio of an ion in a magnetic sector mass spectrometer.

The nonmagnetic transmission quadrupole mass filter, which laid the foundation for what we now call the 3D ion trap (“Ionenkäfig”) and the linear quadrupole ion traps, was developed in the 1950s at the University of Bonn by another mass spectrometry pioneer, the German physicist Wolfgang Paul [9, 10]. In 1959 Hans Georg Dehmelt, University of Washington, had built a high vacuum magnetron trap, which he called the Penning Trap [11]. This trap employed a homogeneous magnetic field and a weak electric quadrupole field produced by hyperbolic electrodes. The Penning Trap has advantages over the radio frequency (Paul) trap for precision measurements of properties of ions and stable subatomic particles which have a nonzero electric charge. With this device, Dehmelt was able to trap electrons for about 10 s and to

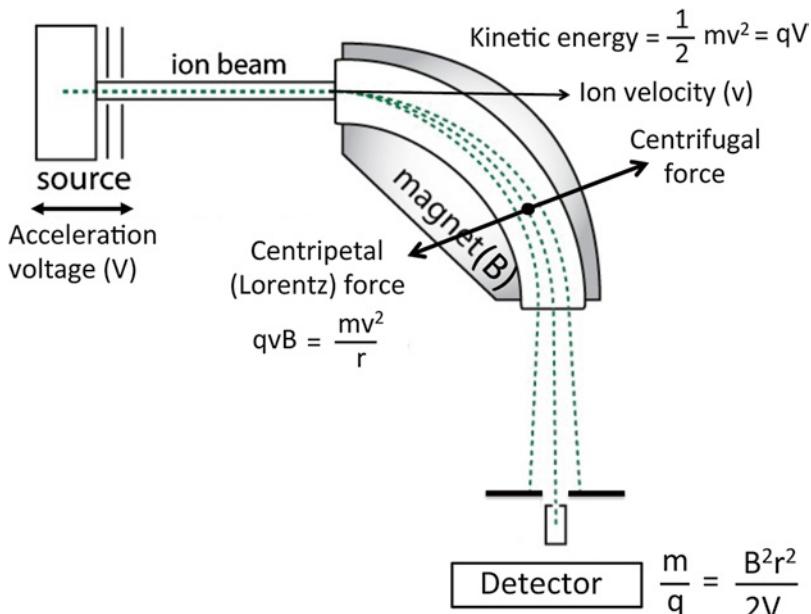


Fig. 1 Scheme of the separation of ions of different m/q by a magnetic sector illustrating how to derive the mass (in unified mass units) of an ion detected at a magnetic field of 1.3 T after acceleration to 9,000 V in the flight tube of a magnetic sector instrument of radius 0.4 m. In a magnetic field of strength $B=1.3$ T, an ion of charge q and velocity v experiences a force given by: $F_B=Bqv$, which is perpendicular to the direction of B and to the direction of the motion of the ion. Hence, the ion travels in a curve path that has a radius ($r=0.4$ m) in a plane perpendicular to the direction of B . Since F_B is counterbalanced by the centrifugal force, $Bqv=mv^2/r$. Substituting the value of v from the equation of the kinetic energy, where m is the mass of the ions in kg ($1\text{u}=1\text{ Da}=1.66 \times 10^{-27}$ kg), v is the velocity of the ion in m/s, q is the charge of the ion ($q=ze$; z is the integral number of charge and e is the fundamental unit of charge $=1.602 \times 10^{-19}$ C), and V is the ion acceleration voltage, m (kg) $=zeB^2r^2/2V=m(\text{u})=[(1)(1.602 \times 10^{-19})\text{ C}(1.3 \text{ T})^2(0.4 \text{ m})^2]/[(1.66 \times 10^{-27})\text{ kg}(2)(9,000 \text{ V})]=1,455.36$ Da. Note that all units will cancel out and the mass of the ion will be calculated in Unified Mass Units ($1\text{u}=1\text{ Da}$) because $T=\text{kg/s C}$; and $CV=J=\text{kg m}^2/\text{s}^2$

detect axial, magnetron, and cyclotron resonances. This technique was also used for high precision measurement of the magnetic moment of the electron. For their work on ion trapping Paul and Dehmelt shared the 1989 Nobel Prize in Physics.

The transmission quadrupole mass filter consists of four parallel metal rods. Opposing rods are connected electrically with a fixed direct current (DC) and an alternating radio frequency (RF) voltage is applied between one pair of rods and the other. Superimposed DC and RF potentials on the quadrupole rods can be set to let only selected mass-to-charge ratio ions travel down the quadrupole. All other ions do not have a stable trajectory through the quadrupole mass analyzer and will collide with the quadrupole rods, never reaching the detector. This permits selection of an ion with a particular m/z or allows scanning for a range of m/z values by continuously varying the applied voltage. Stable trajectories of ions through the quadrupole are achieved by combinations of DC and RF electric fields that define the stability diagram [12, 13]. A derivation of the working equations for a quadrupole mass analyzer is beyond the scope of this discussion, but it is based upon a bounded solution to the Mathieu functions originally derived in 1868 by the French mathematician Émile Léonard Mathieu [12, 13] (<http://www.physics.drexel.edu/~tim/open/mat/mat.html>):

$$\alpha_x = -\alpha_y = \frac{4qU}{m^2 r_0^2}$$

$$q_x = -q_y = \frac{2qV}{m^2 r_0^2}$$

where α and q values are respectively related to the magnitude of the applied DC potential (U) and the applied RF signal (V) that correspond to stable trajectories in the quadrupole mass filter stability diagram (Fig. 1), and r_0 is the distance to the central (z) axis of the quadrupole.

In a transmission quadrupole mass filter ions oscillate in the x - y plane with frequencies which depend on their m/z values. On the other hand, the quadrupole ion trap, is the three dimensional analogue of the linear quadrupole mass filter. In this device too, ions are subjected to forces applied by an RF field but the forces occur in all three (x , y , z), instead of just two dimensions, which can result in ions being trapped in the field. In the 3D ion trap ions experience restoring forces that drive them back toward the center of the trap confining them in the small volume between a ring electrode and two end-cap electrodes by appropriately oscillating electric fields. Stability diagrams can be constructed from the interdependence of the stable trajectories of the ions within the ion trap and the field conditions plotted in Mathieu (α , q) space. The values of α and q from the Mathieu equations depend on the dimensions of the trap according to the following relationships (Fig. 1):

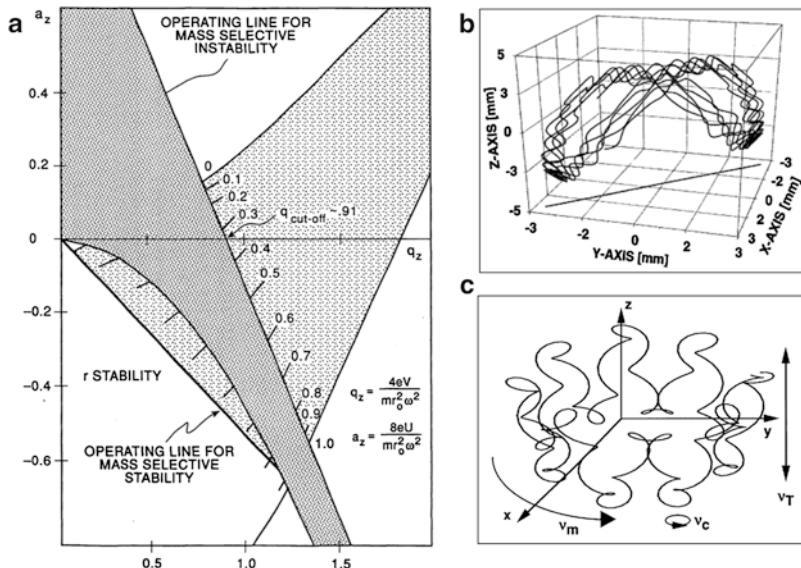


Fig. 2 (a) Diagram of the stability regions of a quadrupole ion trap according to the voltage and frequency applied to the ion trap elements. From Patent 5399857 “Method and apparatus for trapping ions by increasing trapping voltage during ion introduction” US Patent Issued on March 21, 1995. **(b)** Trajectory of an m/z 105 ion trapped in a 3D Ion Trap. The projection onto the x - y plane illustrates planar motion in three-dimensional space. The trajectory develops a shape that resembles a flattened boomerang. Taken from [101]. **(c)** Ion cyclotron resonance motion of an ion of m/z 2,300 in a 2 in. cubic Penning trap in a perfectly homogeneous magnetic field of 3 T, for 10 V trapping voltage [102].

$$a_z = -2a_r = 8 \left(\frac{q}{m} \right) \left(\frac{U}{r_0^2 \omega_{RF}^2} \right)$$

$$q_z = -2q_r = 4 \left(\frac{q}{m} \right) \left(\frac{V}{r_0^2 \omega_{\text{RF}}^2} \right)$$

The subscripts z and r represent, respectively, axial and radial motion perpendicular to and between the end caps, U is the DC potential on the end-cap electrodes, V is the RF potential applied to the ring electrode, r_0 is the radius of the ring electrode, and ω_{RF} is the RF angular frequency. Figure 2a displays a diagram of the stability regions of a quadrupole ion trap according to the voltage and frequency applied to the ion trap elements, and Fig. 2b shows the trajectory of an m/z 105 ion confined in a 3D ion trap.

A unique feature of a Paul ion trap compared with a quadrupole mass analyzer, which will be discussed below, is that the former can perform multiple stage mass spectrometry (MS^n) simply by the use of additional operations which are performed sequentially in time. In a typical multiple-stage mass spectrometry experiment, the ion of interest is isolated, induced to fragment by a resonance signal adjusted to cause collisionally

induced dissociation (CID) of the ion with the helium damping gas, followed by ejection and detection of the fragment ions [14, 15]. Using quadrupole mass analyzers, tandem MS/MS requires combinations of multiple quadrupoles such as the linear series of three quadrupoles (Q1–Q2–Q3) known as triple quadrupole instrument. The first triple-quadrupole mass spectrometer was developed at Michigan State University in the late 1970s [16]. Q1 and Q3 act as mass filters whereas Q2, an RF-only quadrupole, is used as a collision cell where parent ions selected in Q1 are allowed to collide (at ~30 eV) with neutral molecules (often He, Ar, or N at low pressure, $\sim 10^{-3}$ Torr), resulting in bond breakage of the molecular ion thus generating smaller fragments. These daughter ions are then transferred into Q₃ where they may be filtered or fully scanned. The process of collision-induced dissociation allowing the elucidation of the structure of the parent ion provides an added dimension of mass spectral information discussed below.

The principles of time-of-flight (TOF), conceptually the simplest mass analyzer, were introduced in 1946 [17]. An early time-of-flight mass spectrometer, named the Velocitron, was reported in 1948 [18]. Using TOF-MS an ion's mass-to-charge ratio is determined via a time measurement. Ions are accelerated by an electric field of known strength (1–20 kV). The velocity of the ion depends on its mass-to-charge ratio, which can be derived by measuring the time taken for the particle to reach the detector at a known distance in a field-free time-of-flight tube (typically 0.5–2 m) before striking the detector. When the charged particle is accelerated into time-of-flight tube by the voltage U , its potential energy ($E_p = qU$) is converted to kinetic energy: $qU = 1/2mv^2$. Substituting $v = d/t$, $t = [d/(2U)^{1/2}](m/q)^{1/2}$ or, since the TOF tube distance (d) and the accelerating voltage (U) are constants related to the instrument settings and characteristics, $t = k(m/q)^{1/2}$.

In 1974, Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) was developed [19, 20]. The theory of cyclotron resonance was developed in the 1930s by Lawrence (1939 Nobel Prize in Physics) [21]. The operating principles of an ICR cell (a Penning trap) in trapping and detecting ions are based on the observation that a charged particle in a spatially uniform magnetic field experiences a Lorentz force which causes the ion trajectory to bend such that it rotates in a plane perpendicular to the magnetic field axis at a frequency related to its m/z value [20, 22–24]. In a spatially uniform static magnetic field (B), all ions move in circular orbits with characteristic cyclotron frequencies that are inversely related to the m/q : $f = qB/2\pi m$, where f =cyclotron frequency, q =ion charge, B =magnetic field strength and m =ion mass. This is more often represented in angular frequency (ω_c), which is related to frequency by $f = \omega_c/2\pi$, and thus $\omega_c = qB/m$.

Ions in the ICR cell irradiated with a pulse of RF at the same frequency as ω_c oscillate in stable ion cyclotron motions (Fig. 2c). Ions absorbing the energy increase the radius of their orbits inducing thereby an oscillating charge in the walls of the ICR cell as they precess. The frequency of this induced charge oscillation can be detected and amplified. Because a wide variety of energies are transmitted to the ICR cell, FT-ICR MS allows many ions to be detected simultaneously. The advantages of detecting all ions simultaneously derives from the “Fellgett advantage” [25] and is derives from the expression that relates the orbital radius (r) with the excitation electric field (E_0) and the excitation time (t_{excite}), $r = (E_0 t_{\text{excite}})/2B$. The fact that the orbital radius of the excited ions is independent of the m/z ratio, means that ions of different m/z ratios can be excited to the same ICR radius. In addition, because frequency is an easily and accurately measurable parameter, FT-ICR MS has the highest potential for mass accuracy determinations.

Pivotal to any mass spectrometry experiment is the mass resolving power of the mass analyzer, $m/\Delta m$, where Δm is the full peak width at 50 % of maximum peak height. Mass resolving power is a measure of how well two closely spaced peaks can be resolved. In FT-ICR MS $|m/\Delta m| = |\omega/\Delta\omega|$. Because the frequency of ion cyclotron precession is directly proportional to the strength of the magnetic field, increasing the strength of the magnet increases resolution. In addition, Fourier transform MS is unique in that increased measurement time also increases both sensitivity and resolution [26]. Increased sensitivity is derived from the fact that ions are not consumed during the detection process. In a zero-pressure limit, the frequency-domain mass spectral peak width is given by:

$$\Delta\omega = \frac{2\pi(1.2060)}{T}$$

where T is the time-domain data acquisition period [27]. It follows that the mass resolving power is directly proportional to the observation period. In addition, the resolution at low pressure is given by:

$$\frac{m}{\Delta m} = \frac{(1.274 \times 10^7 z B_0 T)}{m}$$

[28, 29] clearly showing that the resolving power, $m/\Delta m$, increases linearly with the strength of the magnetic field (B_0) and the acquisition time. Thus, resolution can be improved either by increasing the strength of the magnet (in teslas) or by increasing the detection duration. Mass resolution (R) exceeding 1,500,000 have been achieved by FT-ICR MS. However, high-performance instruments use superconducting magnets that are large in size and expensive to maintain from the standpoint of cryogenic cooling.

3 The Breakthrough for Proteomics: Molecular Elephants in the Gas Phase

By the 1970s, mass spectrometry had become a standard analytical method in the analysis of organic compounds. However, the application of mass spectrometry to biological studies was hampered due to the inexistence of suitable ionization methods for fragile and nonvolatile biological molecules. The 1980s saw the development of ionization techniques capable of efficiently ionizing biological molecules. This included Californium-252 plasma desorption ($^{252}\text{Cf-PD}$) in 1976 [30], fast atom bombardment (FAB) in 1981 [31, 32], matrix-assisted laser desorption/ionization (MALDI) in 1985–1988 [33], and electrospray ionization (ESI) in 1984–1988 [34, 35]. FAB uses a source of neutral heavy atoms to ionize compounds from the surface of a liquid (i.e. glycerol) matrix. The breakthrough for large molecule ionization came in 1987 when Koichi Tanaka and colleagues from Shimadzu Corporation (Japan) were able to ionize the 34,472 Da protein carboxypeptidase-A using what they called the “ultra fine metal plus liquid matrix method” that combined 30 nm cobalt particles in glycerol with a 337 nm nitrogen laser for ionization [36, 37]. Simultaneously, Michael Karas and Franz Hillenkamp ionized the 67 kDa protein albumin using a nicotinic acid matrix and a 266 nm laser [38]. Meanwhile, John B. Fenn and colleagues at Yale University refined an ion source originally reported by Malcolm Dole of Northwestern University almost two decades earlier to develop the electrospray ionization (ESI) technique [34, 35]. $^{252}\text{Cf-PD}$, FAB, ESI, and MALDI are all relatively soft ionization techniques that produce primarily protonated $[\text{M} + \text{H}]^+$ and deprotonated $[\text{M} - \text{H}]^-$ quasimolecular ions. However, with the development of MALDI and ESI, the upper mass range was extended beyond 100 kDa and this had an enormous impact on the use of mass spectrometry in life sciences. “For the development of soft desorption ionisation methods for mass spectrometric analyses of biological macromolecules” John B. Fenn and Koichi Tanaka were awarded with the Nobel Prize in Chemistry in 2002.

In ESI the liquid containing the analyte of interest is electrostatically dispersed at the ES capillary (internal diameter $<250\text{ }\mu\text{m}$), placed at a potential difference between +500 and +4,500 V, producing a fine aerosol of charged droplets. In 1994, Emmett and Caprioli [39] and Wilm and Mann [40] introduced the microelectrospray, later developed as nanospray [41], demonstrating that a capillary flow of $\sim 25\text{ nL/min}$ can sustain an electrospray at the tip of emitters fabricated by pulling glass capillaries to a few micrometers. To decrease the initial droplet size, compounds that increase the conductivity (e.g. acetic or formic acid) are customarily added to the solution. Regardless of the size of the ESI set up, solvent flowing at the tip of the capillary evaporates from a charged droplet

until it becomes unstable upon reaching its Rayleigh limit. At this point, the droplet deforms and emits charged jets in a process known as Coulomb fission. Repeated charge-induced droplet disintegrations lead ultimately to small highly charged droplets capable of producing gas-phase ions. Multiply charged ions such as $[M + nH]^{n+}$ are often observed. Polypeptides tend to be variably protonated at arginine, lysine, and histidine residues. Multiple charging enables mass analyzers with limited m/z ranges to analyze higher molecular mass molecules. In addition, ESI allows non-covalent protein complexes to be ionized intact, showing its utility for determining the quaternary structure of proteins [42]. A shortcoming of ESI is its susceptibility to ion suppression at relatively low salt concentrations, $> \sim 1$ mM, which means that biological samples need to be desalting before ESI-MS.

Several theories have been proposed to explain the final production of gas-phase ions: the Charge Residue Model (CRM) [43]; the Ion Evaporation Model (IEM) [44, 45]; and a model invoking combined charged residue-field emission [46]. Whatever mechanism operates, and because ion formation involves extensive solvent evaporation, typical solvents for electrospray ionization are mixtures of water with volatile organic solvents (e.g. methanol, acetonitrile). These solvents are typically employed in reverse-phase high-performance liquid chromatography (RP-HPLC). Due to the ease of coupling LC directly to the ion source, the combination of liquid chromatography and mass spectrometry has entered the realm of routine analysis.

The physics behind MALDI is also debated [47–49]. Ionization is triggered by firing a pulsed laser beam at a dried-droplet spot of co-crystallized matrix and analyte molecules usually placed in a vacuum chamber. The matrix (in large excess) absorbs the laser energy and is desorbed and protonated. The ionized matrix transfers the proton to analyte molecules thus generating $M + H^+$ ions. Since the typical MALDI source generates pulse of ions, this ionization method is most compatible with spectrometers which function as pulsed-ion detectors (TOF) and with instruments which trap the ions for later analysis (ion traps and FT-ICR instruments). Mass analyzers which operate on a continuous beam of ions, such as magnetic sector and quadrupole instruments, are generally not suitable for a pulsed-ion source. Although MALDI sources that operate at atmospheric pressure (AP) have been developed, which approach the sensitivity of vacuum MALDI [50], still the most natural form of interfacing a liquid chromatograph to MALDI-MS is off-line. However, it is now well established that time-of-flight mass spectrometry involving instruments with independent (orthogonal) axes for ion generation and mass analysis, generally referred to as orthogonal acceleration time-of-flight (oa-TOF) mass spectrometry, are well suited to continuous ion sources [51]. This approach allows synchronization of ion formation with mass

analysis. Very little sample is wasted and therefore MALDI can achieve high levels of sensitivity, often $<1 \times 10^{-15}$ mol. In addition, MALDI can produce very large singly charged ions and thus the type of mass spectrometer most widely used with MALDI is the TOF due to its large mass range. However, the presence of matrix ions cause large chemical noise at $m/z < 500$, and thus low molecular mass analytes are difficult to analyze by MALDI.

4 Analytical Innovations in the Proteomics Era: The Trend to Hybridization

Concurrent with the development of ionization methods, several innovations in mass analyzer technology expanded the range of applications in biological research. Though innovations involve the design of individual mass analyzers, the trend in contemporary biological mass spectrometry is mass analyzer hybridization because hybrid configurations possess additional functional capabilities. Figure 3 displays hybrid configurations widely used in current biomolecular mass spectrometry. The simplest hybrid configuration, tandem mass spectrometry (MS/MS), involves more than one step of mass selection or analysis, and fragmentation is usually induced between these steps. Several types of tandem mass spectrometers have been designed through combinations of the six classes of major mass analyzers currently used in biomolecular analysis: Time-of-flight (TOF), Quadrupole (Q), 3D Ion Trap (3D-IT),

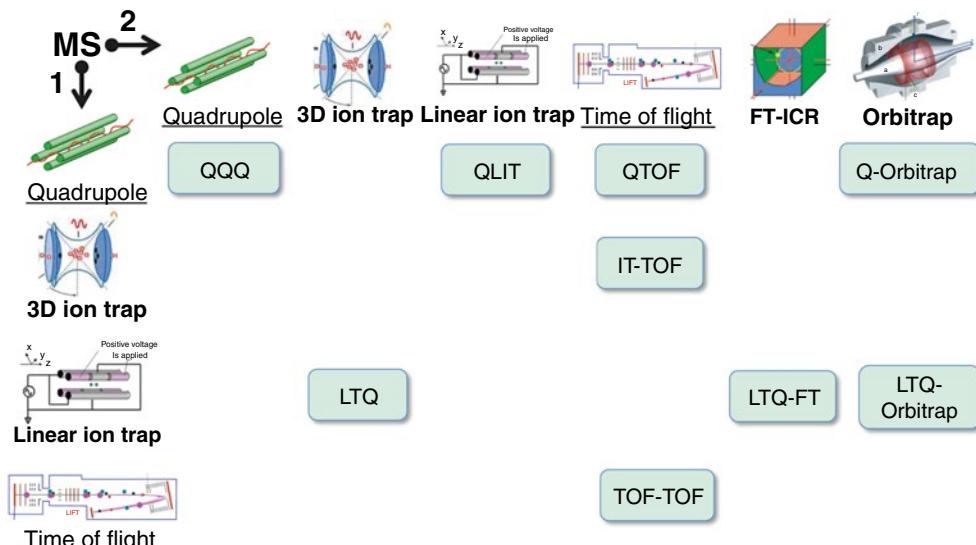


Fig. 3 Hybrid configurations widely used in current biomolecular mass spectrometry. Beam-type instruments' names are *underlined*. Trapping instruments are highlighted in *bold face*. Note that quadrupoles are not used only as mass analyzers; they also are implemented frequently as ion transfer optics, collision cells, and linear ion traps

Linear Ion Trap (LIT), Orbitrap, and Fourier Transform-Ion Cyclotron Resonance (FTICR). Tandem configurations can be classified into two groups: tandem-in-space instruments and tandem-in-time instruments. Tandem-in-space instruments require separate mass analyzers to be utilized for each MS stage and are associated with beam-type technology such as quadrupole mass filters, and TOF mass analyzers. Tandem-in-time instruments perform separate the different MS stages by time, with the various stages of MS/MS being performed in one mass analyzer. Hybrid instruments that combine both in-space with in-time mass analysis have been also built (Fig. 3).

The Q-TOF, which results from replacing the third quadrupole in a triple stage quadrupole by a TOF analyzer [52] is one of the most popular configurations. Targeted absolute quantitation of peptides is a fast developing field with triple quadrupole instruments being traditionally used for this type of work. Recently, AB Sciex has launched its next generation triple quadrupole-TOF (TripleTOF®) family of spectrometers. The new TripleTOF® 5600+ System is a high sensitivity and high-resolution mass spectrometer for simultaneous high-performance qualitative and quantitative analysis. It delivers new innovation for biomolecule research with MS/MS^{ALL} with SWATH™ acquisition [53]. The TOF spectrometer of hybrid systems is usually equipped with a reflectron. The function of the reflectron, an electrostatic mirror, is to compensate for small differences in the velocities of ions with the same m/z , increasing thereby the resolution of TOF MS [54]. The TOF-TOF tandem mass spectrometer was introduced in 1993 [55]. It contained two dual-stage reflectron analyzers and a collision region for producing product ions by collision-induced dissociation. Current TOF-TOF instruments incorporate delayed extraction of ions from the ion source towards the flight tube by ns- μ s in order to allow ion equilibration in the plasma produced during the desorption/ionization, thereby improving mass resolution [56]. A crucial aspect of the MS/MS experiment is the unimolecular ion dissociation that occurs between the two MS stages. High-energy CID of singly charged ($M + H^+$) peptide ions occurring in TOF-TOF instruments generates incomplete series of sequence-specific daughter ions, challenging *de novo* structure elucidation. However, combined peptide mass fingerprinting (in TOF1) and partial MS/MS spectra of selected peptide ions detected (in TOF2) represents a high confident approach for identifying protein sequences present in the searched databases [57]. Automated high-throughput TOF-TOF MS analysis and database search enables a time- and cost-effective protein identification strategy where genome sequence of the particular organism under study is available.

Tandem and multiple-stage MS (MS/MS and MSⁿ, respectively) can also be achieved by the use of additional sequence of operations in the scan function of ion traps. However, a limitation

of the 3D trap is that the trapping volume is small, so that deleterious ion-ion interactions, otherwise known as space-charge effects, allow just a small number of ions to be trapped and analyzed effectively. In addition, in MS/MS mode 3D ion traps typically have a low mass cut-off which usually corresponds to about one-third of the precursor ion mass. The fundamental limitations of the 3D trap prompted researchers to investigate new trapping geometries, ultimately leading to the development of the linear ion trap. The linear ion trap (LIT), sometimes referred to as the two-dimensional (2D) quadrupole ion trap, is such a recent development [58–60]. It makes use of the basic structure of the quadrupole to confine ions radially by a two-dimensional (2D) radio frequency (RF) field, and axially by static electrical potentials applied to on-end electrodes. One of the attractive features of linear ion traps for proteomic experiments is that, in comparison to Paul traps, they have higher injection efficiencies, higher ion storage capacities, and do not have low mass cut-off. Linear traps have been combined with other mass analyzers in hybrid configurations and used to isolate ions of selected mass to charge ratios to perform tandem mass spectrometry experiments (see below). ABSciex's QTRap™ system is a hybrid triple quadrupole ($Q_0Q_1Q_2$) linear ion-trap (LIT) (Q_3) mass spectrometer [58]. It can be operated in a conventional triple quadrupole configuration with the traditional quadrupole scanning modes: precursor ion, neutral loss, and multiple reaction monitoring scans. In LIT mode, a pulse of ions passes through Q_1 operated as a conventional quadrupole mass filter to select the precursor ion of interest. The precursor ion is accelerated into the pressurized Q_2 to promote fragmentation. The fragments and residual precursor ion are then trap in the Q_3 linear ion trap. The Q_3 RF voltage is then ramped to eject ions towards the detector. In addition, enhanced MS survey and high sensitivity product ion spectra are possible because while scanning out the ions from Q_3 , ions can be accumulated in Q_0 and this gives a substantial increase in signal intensity over normal operation mode.

A limitation for interfacing a continuous ion source, as ESI, to a 3D trap is that ESI generates ions while the 3D trap is processing other ions are not used, thereby limiting the duty cycle. The duty cycle (D_c) of a 3D ion trap coupled directly to a continuous ion source is $D_c = T_C / (T_C + T_{A,3D})$, where T_C is the ion trap fill time, and $T_{A,3D}$ is the time for MS analysis in the 3D trap [59]. Injection efficiencies are typically $\leq 5\%$ for externally created ions. Combining a linear trap with a 3D trap, as in Thermo Fisher's linear trap quadrupole (LTQ) instruments, overcomes this limitation [60]. Accumulating ions in the linear trap while the 3D trap is processing other ions increases the duty cycle to nearly 100 %. A TOF mass spectrometer can also have a low-duty cycle when coupled with a continuous ion source. Combining an ion trap with a TOF mass analyzer also improves the duty cycle [61]. Thermo's LTQ FT

instrument is another hybrid instrument, which combines a linear ion trap with a Fourier Transform Ion Cyclotron Resonance (FT-ICR) MS detector [62]. This configuration provides rapid MS and MS/MS analyses, similar to the “data-dependent scanning” found on standard 3-D Paul traps, but with substantially improved internal scan dynamic range, mass measurement accuracy, mass resolution, and detection limits [62].

The new millennium also witnessed the introduction of the Orbitrap™ mass analyzer [63]. The roots of the Orbitrap stem from the principle of orbital trapping in an electrostatic field defined by Kingdon in 1923 [64], and modified in 1981 by Knight, who introduced a modified outer electrode that included an axial quadrupole term that confines the ions on the trap axis [65]. The Orbitrap developed by Alexander Makarov in 1996–1997 is an ion trap that only uses an electrostatic field to trap ions radially around a central spindle electrode [63]. An outer electrode confines ions axially. Mass-to-charge values are measured from the frequency of harmonic ion oscillations along the z -axis of the electric field. The back and forth movement of ions around the central electrode is described by a simple harmonic oscillator with angular frequency of axial oscillations (ω , in rad/s) = $[(q/m)k]^{1/2}$, being k the field constant [63]. The generated image current is amplified and processed by fast Fourier transforms to obtain the mass spectra, resulting in comparable sensitivity and resolution than FT-ICR MS [66, 67]. In the Orbitrap, the mass resolving power, $m/\Delta m = (1/2\Delta\omega)(kq/m)^{1/2}$, increases linearly by increasing the magnetic field. However, since the resolving power also depends on the time-domain data acquisition period, in LC-MS experiments the relevant parameter is the resolution achievable over the limited time of the chromatographic separation ($m/\Delta m$ of around 120,000 (at m/z 400) has been demonstrated at 1 scan/s) [67].

A shortcoming of the Orbitrap analyzer is that when an ion decays under the dynamic trapping, its fragments will have the same velocity as the precursor ion. As their energy is proportional to their individual mass-to-charge ratios, the trajectories of daughter ions with m/q typically below 30–50 % of that of the precursor ion become highly elliptical and eventually these ions are lost. In addition, mass analyzers associated with ion-trapping technology, such as quadrupolar (3D) ion traps, FT-ICR mass analyzer, and Orbitrap, are unable to perform precursor ion or neutral loss scans; for these screenings beam-type instruments have a definite advantage. The challenge of performing MS/MS and data-dependent scans were main reasons behind the concept of using the Orbitrap as an accurate-mass detector for another mass analyzer. The first hybrid system, the LTQ-Orbitrap™ [68], was commercialized in 2005. The ion storage device linking the linear ion trap to the Orbitrap analyzer is called the C-trap [69], which has a high space charge capacity and can also be used for additional fragmentation methods such

as ETD (electron transfer dissociation) [70, 71]. This hybrid system retains the high sensitivity, high charge capacity, accurate control of ions, and short cycle time of the linear trap, and the high-resolution measurements of the Orbitrap. The LTQ-Orbitrap system has also the capability of performing data-dependent scanning for precursor fragmentation studies. The Orbitrap analyzer is used to scan the precursor species and the LTQ mass analyzer is set to pick them, isolate them within the linear C-trap and fragment them. Product ions can then either be detected using the C-trap or the Orbitrap. A variant of the ESI-LTQ-Orbitrap is the MALDI-LTQ-Orbitrap, which couples a vacuum MALDI source to a hybrid LTQ Orbitrap and is particularly suited for Peptide Mass Fingerprint and Tissue Imaging applications [72]. Other Orbitrap configurations include the Q Exactive (hyperbolic quadrupole mass filter-Orbitrap) mass spectrometer (140,000 resolution at m/z 200 and <1 ppm mass accuracy) [73]. This system enables new quantitative methods based on high resolution and accurate mass measurements, including targeted analysis in MS mode (single ion monitoring, SIM) and in MS/MS mode (parallel reaction monitoring, PRM) [74]. The LTQ Velos™ family [75] (Velos™, Velos Pro™, Orbitrap Elite™) was designed to improve sensitivity and scan speed: the first ion trap efficiently captures and fragments ions at relatively high pressure whereas the second ion trap realizes extremely fast scan speeds at reduced pressure.

Drift-time ion mobility mass spectrometry (DT-IMMS) developed in the 1950s and 1960s by Earl W. McDaniel at the Georgia Institute of Technology. With the help of mechanical engineering student, Dan Albritton, McDaniel constructed in 1964 a “drift tube” that revolutionized the field of ion transport. The paper reporting this achievement [76] was chosen as one of the top 100 papers ever published in the reputed journal Physical Review. Drift-time IMMS measures the time that an ion forced to drift in a gas cell takes to migrate through the buffer gas (at a pressure of up to about 0.7 Torr) in the presence of a low electric field [77–79]. Under low-field conditions ion mobility can be thought of as “directed diffusion”, and the velocity of the ion is directly proportional to the electric field. The ion mobility constant (K) is related to the ion’s collision cross-section by:

$$K = \left(\frac{3q}{16N} \right) \left(\frac{2\pi}{kT} \right)^{1/2} \left[\frac{(m+M)}{mM} \right]^{1/2} \left(\frac{1}{\Omega} \right),$$

where q is the charge of the ion, N is the number density of the buffer gas, k is the Boltzmann’s constant, T is the absolute temperature, m is the mass of the buffer gas, M is the mass of the ion, and Ω is the collision cross-section of the ion [80]. Drift-time IMMS can be used to separate isobaric ions differing in

collision cross-section. A novel method of ion mobility, called “traveling-wave” ion mobility spectrometry (TWIMS), has been recently reported [81]. Unlike the traditional drift-time IMMS, a DC pulse is applied sequentially through the ion mobility cell one segment at a time causing the ions to move forth to the next potential groove. When ions differing either in charge state or collision cross sections, or both, are entering the traveling wave ion guides (TWIG), the less mobile ones may occasionally transverse a wave in backward direction. Those ions lagging behind the traveling wave will exit later than those surfing on the wave. This makes the TWIG become an ion mobility separator [82]. A hybrid instrument comprising a mass-selecting quadrupole, a traveling wave ion mobility separator, and an oaTOF analyzer has been marketed by Waters as Synapt™ mass spectrometry systems [81]. An advantage of this Q-IM-oaTOF MS configuration is that the sensitivity of the mass analyzer is not compromised by the duty cycle of the ion mobility cell, as is the case with drift-tube instruments in which the ion gate is only open for about 1 % of the duty cycle [81]. Continuous introduction of mobility-selected ions (referred to as SelexION™ technology) has also been implemented in ABSciex Triple Quad™ and QTRAP® platforms [83–85].

Ion mobility cells have been interfaced to TOF, quadrupole, ion trap, and FTICR mass spectrometers [78], indicating the analytical power of coupling mobility with mass measurements. Ion mobility delivers a new dimension of selectivity and performance for applications requiring the separation of isobaric species. The application of ion mobility in biological research is gaining more and more popularity. Size, mass, and position are important parameters that describe a molecule in a biological system [86]. Ion mobility is capable of separating ions of the same size on the basis of their cross-section area (shape), whereas imaging MS is used to visualize the spatial distribution of molecules. The evolution of ion-mobility-based imaging mass spectrometry, reviewed by Kiss and Heeren [86], provides examples of its application in analytical studies enabling the simultaneous determination of size, mass, and position of molecules in biological surfaces. Advances in the field of proteomics parallel technological advances in mass spectrometry techniques. The unique capabilities of ion mobility offer a useful analytical tool for bottom-up [87] and top-down [88] proteomics. For instance, an increased ability to disperse peptide ion signals in the ion mobility dimension facilitates parallel tandem mass spectrometry experiments on mobility-separated b-type and y-type sequence-specific peptide ions. An approach that combines mass spectrometry, CID, and ion mobility for top-down proteomics has been recently described [88]. Using this approach, CID product ions are dispersed in two dimensions, specifically size-to-charge (IM) and mass-to-charge

(MS), and the resulting 2D data display greatly facilitates mass mapping, amino acid sequence analysis, and determination of site-specific protein modifications.

5 The Big Question: Choice of Mass Analyzer

As a result of its utility in (bio) chemical analysis, MS has become a routinely used standard analytical tool. The emergence of post-genomic disciplines, and efforts in the drug development industry, has fuelled instrumental developments in mass spectrometry to investigate the complexity of living organisms, which has led to a wide variety of configurations. New configurations of mass spectrometry systems, as well as improvements of the current ones, will become available in the near future. This raises the important question of which MS configuration is better suited for addressing one's research needs and expectations. The choice of mass analyzer should be based upon the application, cost, and performance desired. In addition, in assessing the suitability of different mass analyzers, a range of instrument-specific factors must be taken into account. The performance of a mass analyzer can be defined by the following characteristics [89, 90]: mass accuracy (the difference between the true m/z and the measured m/z of a given ion, divided by the true m/z of the ion, usually quoted in terms of parts per million, ppm: $[(m/z(\text{exp}) - m/z(\text{theor})) / m/z(\text{theor})] \times 10^6$) [91]; mass resolving power ($M/\Delta M$, the ratio of peak mass (M) to the peak width at half maximum intensity, ΔM); mass range (the range of m/z over which the mass analyzer can operate to record a mass spectrum); linear dynamic range (the range over which the ion signal is directly proportional to the analyte concentration); tandem (MS/MS) and multi-stage (MS)ⁿ analysis capabilities; abundance sensitivity (the ratio of the maximum ion current recorded at an m/z of M to the signal level arising from the background at an adjacent m/z of $M+1$); and scan speed. Table 1 displays a nonexhaustive comparative overview of these parameters from some major mass analyzer configurations (Fig. 3) currently used for proteomics research.

The most common mass analyzer for quantitative bioanalytical analyses is the triple quadrupole due to its capability to operate under selected-ion monitoring and scanning modes that are used for identifying drug metabolites, biomarkers, and posttranslational modifications such as glycosylation or phosphorylation. The precursor ion scan essentially is the reverse of the product ion scan. The third quadrupole is set to select a specific product ion formed in Q2, and the first quadrupole is scanned for all precursor ions forming the chosen fragment. Neutral loss scanning is also achieved in triple quadrupoles. Neutral loss scans are used routinely to identify common functional groups present in a set of molecules.

Table 1
Conservative figures of merit of mass analyzers typically used in proteomics

| | QQQ | 3D-IT | QQ-qLIT | TOF-TOF | FT-ICR | LTQ-Orbitrap |
|-------------------|---------------------|---------------------|---------------------|---------------------|----------------------|---------------------|
| Mass range | 4,000 | 4,000 | 4,000 | Unlimited | 4,000 | 6,000 |
| | | | | 4,500 (R) | | |
| Mass accuracy | 100 ppm | 100 pm | 100 ppm | 5–10 ppm | <1 ppm | 2–5 ppm |
| Resolution (FWHM) | ≥4,000 | ≥4,000 | ≥4,000 | ~25,000 (R) | ≥1 × 10 ⁶ | ≥150,000 |
| Dynamic range | 1 × 10 ⁷ | 1 × 10 ³ | 1 × 10 ⁴ | 1 × 10 ⁶ | 1 × 10 ³ | 1 × 10 ⁴ |
| (MS) ^a | (MS) ² | (MS) ^a | (MS) ³ | (MS) ² | (MS) ^a | (MS) ^a |
| Sensitivity | Atto-fmol | Picomol | Femtomol | Femtomol | 1 × 10 ⁴ | 1 × 10 ⁴ |
| Scan speed | ~1 s | ~1 s | 4/s | 200/s | ~1 s | ~1 s |

These values may vary with hybrid configurations from different instrument manufacturers; please check the manufacturer's technical specifications

R reflectron, *FWHM* full width at half maximum

Another common application of triple quadrupoles is the quantitative analysis of metabolites or candidate biomarkers by selected or multiple reaction monitoring. SRM and MRM are particularly useful for targeted quantitative proteomics [92, 93].

FT-ICR mass spectrometers provide the highest recorded mass resolution (Table 1). However, stable mass calibration requires superconducting magnets, which are costly to maintain. Orbitrap configurations are characterized by high mass resolution (up to 150,000), large space charge capacity, high mass accuracy (2–5 ppm), a mass/charge range of at least 6,000, and dynamic range greater than 10³ [66]. The high resolving power of the Orbitrap is possible because the ions' *m/z* value is a function of frequency, and frequency is a parameter that can be very accurately measured by existing technology. The ability to obtain accurate mass measurements on precursor and MS/MS fragment ions in proteomic experiments is critical to identify and characterize protein with increased confidence in protein database searches.

Linear ion traps can be used either as ion accumulation devices or as stand-alone mass spectrometers. They have become very popular, particularly as front-end stages prior to trapping instruments, 3D ion trap, FT-ICR, and Orbitrap, or as Q3 in a triple-quadrupole configuration (Fig. 3). In addition to classical triple quadrupole scan functions, such as enhanced product ion scan, time-delayed fragmentation, and enhanced multiply charged scans, the QqLIT design allows an additional stage of MS/MS after the regular triple-quadrupole CID [58, 59].

Unlike the quadrupole, the TOF mass analyzer is not a scanning instrument. The TOF mass analyzer characterizes ions as batches

without scanning for different ions to acquire a complete mass spectrum. TOF-TOF instruments combine the quasi-simultaneous detection of all ions in TOF₁ with high-energy CID MS/MS of selected ions for PMF identification of proteins from genome sequenced organisms [57]. High sequence coverage achieved by the high dynamic range of TOF-MS (Table 1), along with accurate mass measurement coupled with sufficient resolution, makes it possible to greatly restrict the number of false positives in protein database searches. In addition, the high analytical speed and automation makes MALDI-TOF-TOF MS suitable for large number of samples.

Despite the great sophistication of currently available configurations, it is disappointing to note that there is no ideal mass spectrometry system capable of performing all applications required in proteomics [94]. To complicate matters, the increasing trend of employing highly automated mass spectrometry platforms in high-throughput approaches is contributing to create numerous operators and users who are unaware of the fundamentals of MS and for whom mass spectrometers are little more than black boxes able to produce thousands of spectra per day in full computerized workflow. The objective pursued by this chapter was to show that the basic operational principles of mass spectrometry can be understood by a simple mathematical language, and that this understanding is necessary to rationally utilizing this versatile technique in its full potential. However, MS and MS/MS spectra interpretation is outside the scope of this essay. The reader is referred to several excellent monographs [95–100].

References

1. Gelpí E (2008) From large analogical instruments to small digital black boxes: 40 years of progress in mass spectrometry and its role in proteomics. Part I 1965–1984. *J Mass Spectrom* 43:419–435
2. Gelpí E (2008) From large analogical instruments to small digital black boxes: 40 years of progress in mass spectrometry and its role in proteomics. Part II 1985–2000. *J Mass Spectrom* 44:1137–1161
3. Thomson JJ (1897) Cathode rays. *Philos Mag* 44:293 (consult also http://www-outreach.phy.cam.ac.uk/camphy/electron/electron_index.htm)
4. Wien K (1999) 100 years of ion beams: Willy Wien's canal rays. *Braz J Phys* 29:401–414
5. Falconer I (2001) Corpuscles to electrons. In: Buchwald J, Warwick A (eds) *Histories of the electron*. MIT, Cambridge, MA, pp 77–100
6. Dempster AJ (1918) A new method of positive ray analysis. *Phys Rev* 11:316–325
7. Aston FW (1919) A positive ray spectrograph. *Philos Mag Ser* 6(38):707–714
8. Downard KM (2007) Francis William Aston—the man behind the mass spectograph. *Eur J Mass Spectrom* 13:177–190
9. Paul W, Steinwedel H (1953) Ein neues Massenspektrometer ohne Magnetfeld. *Zeitsch Naturforsch A8*:448–450
10. Paul W (1990) Electromagnetic traps for charged and neutral particles. *Rev Mod Phys* 62:531–540
11. Dehmelt HG (1989) Experiments with an isolated subatomic particle at rest. Nobel lecture (http://www.nobelprize.org/nobel_prizes/physics/laureates/1989/dehmelt-lecture.pdf)
12. Miller PE, Denton MB (1986) The quadrupole mass filter: basic operating concepts. *J Chem Ed* 63:617–622
13. Dawson PH (1986) Quadrupole mass analyzers: performance, design and some recent applications. *Mass Spectrom Rev* 5:1–37
14. Cooks RG, Glish GL, McLuckey SA et al (1991) Ion trap mass spectrometry. *Chem Eng News* 69:26–30

15. Jonscher KR, Yates JR (1997) The quadrupole ion trap mass spectrometer—a small solution to a big challenge. *Anal Biochem* 244:1–15
16. Yost RA, Enke CG (1978) Selected ion fragmentation with a tandem quadrupole mass spectrometer. *J Am Chem Soc* 100:2274–2275
17. Stephens WE (1946) A pulsed mass spectrometer with time dispersion. *Phys Rev* 69:691
18. Campana JE (1987) Time-of-flight mass spectrometry: a historical overview. *Instrum Sci Technol* 16:1–14
19. Comisarow MB, Marshall AG (1996) The early development of Fourier transform ion cyclotron resonance (FT-ICR) spectroscopy. *J Mass Spectrom* 31:581–585
20. Marshall AG, Hendrickson CL, Jackson GS (1998) Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom Rev* 17:1–35
21. Lawrence E (1951) The evolution of the cyclotron. Nobel lecture: (http://www.nobelprize.org/nobel_prizes/physics/laureates/1939/lawrence-lecture.html)
22. Amster IJ (1996) Fourier transform mass spectrometry. *J Mass Spectrom* 31:1325–1337
23. Marshall AG (2000) Milestones in Fourier transform ion cyclotron resonance mass spectrometry technique development. *Int J Mass Spectrom* 200:331–356
24. Marshall AG, Hendrickson CL (2002) FTICR detection: principles and experimental configurations. *Int J Mass Spectrom* 215:59–75
25. Fellgett P (1958) A propos de la theorie du spectrometre interferentiel multiplex. *J Phys Radium* 19:187–191
26. Vartanian VH, Laude DA (1995) Optimization of ion cell for FT-ICR MS. *Int J Mass Spectrom Ion Proc* 141:189–200
27. Marshall AG, Hendrickson CL (2001) Charge reduction lowers mass resolving power for isotopically resolved electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Rapid Commun Mass Spectrom* 15:232–235
28. Marshall AG (1979) Theoretical signal-to-noise ratio and mass resolution in Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* 51:1710–1714
29. Marshall AG, Comisarow MB, Parisod G (1979) Relaxation and spectral line shape in Fourier transform ion resonance spectroscopy. *J Chem Phys* 71:4434–4444
30. Macfarlane RD, Torgerson DF (1976) Californium-252 plasma desorption mass spectroscopy. *Science* 191:920–925
31. Barber M, Bordoli RS, Sedgewick RD et al (1981) Fast atom bombardment of solids as an ion source in mass spectrometry. *Nature* 293:270–275
32. Morris HR, Panico M, Barber M et al (1981) Fast atom bombardment: a new mass spectrometric method for peptide sequence analysis. *Biochem Biophys Res Commun* 101:623–631
33. Karas M, Bachmann D, Hillenkamp F (1985) Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Anal Chem* 57:2935–2939
34. Fenn JB, Mann M, Meng CK et al (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246:64–71
35. Fenn JB (2003) Electrospray wing for molecular elephants. *Angew Chem Int Ed Engl* 42:3871–3894
36. Tanaka K, Waki H, Ido Y et al (1988) Protein and polymer analyses up to m/z 100,000 by laser ionization time-of flight mass spectrometry. *Rapid Commun Mass Spectrom* 2:151–153
37. Tanaka K (2002) The origin of macromolecule ionization by laser irradiation. Nobel lecture (http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2002/tanaka-lecture.pdf).
38. Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60:2299–2301
39. Emmett MR, Caprioli RM (1994) Microelectrospray mass spectrometry: ultra-high-sensitivity analysis of peptides and proteins. *J Am Soc Mass Spectrom* 5:605–613
40. Wilm MS, Mann M (1994) Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *Int J Mass Spectrom Ion Proc* 136:167–180
41. Wilm M, Mann M (1996) Analytical properties of the nanoelectrospray ion source. *Anal Chem* 68:1–8
42. Smith RD, Bruce JE, Wu Q et al (1997) New mass spectrometric methods for the study of noncovalent associations of biopolymers. *Chem Soc Rev* 26:191–202
43. Dole M, Mack LL, Hines RL et al (1968) Molecular beams of macroions. *J Chem Phys* 49:2240–2249
44. Iribarne J-V, Thomson B-A (1976) On the evaporation of small ions from charged droplets. *J Chem Phys* 64:2287–2294
45. Nguyen S, Fenn JB (2007) Gas-phase ions of solute species from charged droplets of

- solutions. *Proc Natl Acad Sci USA* 104: 1111–1117
46. Hogan CJ, Carroll JA, Rohrs HW et al (2009) Combined charged residue-field emission model of macromolecular electrospray ionization. *Anal Chem* 81:369–377
47. Knochenmuss R (2006) Ion formation mechanism in UV MALDI. *Analyst* 131:966–986
48. Chang WC, Huang LCL, Wang Y-S et al (2007) Matrix-assisted laser desorption/ionization (MALDI) mechanism revisited. *Anal Chim Acta* 582:1–9
49. Hillenkamp F, Peter-Katalinic J (2007) MALDI MS. A practical guide to instrumentation, methods and applications. Wiley-VCH, Weinheim
50. Laiko VV, Baldwin MA, Burlingame AL (2000) Atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry. *Anal Chem* 72:652–657
51. Guilhaus M, Selby D, Mlynski V (2000) Orthogonal acceleration time-of-flight mass spectrometry. *Mass Spectrom Rev* 19:65–107
52. Chernushevich IV, Loboda AV, Thomson BA (2001) An introduction to quadupole/time-of-flight mass spectrometry. *J Mass Spectrom* 36:849–865
53. Gillet LC, Navarro P, Tate S et al (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* 11:O111.016717. doi:[10.1074/mcp.O111.016717](https://doi.org/10.1074/mcp.O111.016717)
54. Mamyrin BA (2001) Time-of-flight mass spectrometry (concepts, achievements, and prospects). *Int J Mass Spectrom* 3:251–266
55. Cornish TJ, Cotter RJ (1993) Tandem time-of-flight mass spectrometer. *Anal Chem* 15:1043–1047
56. Brown RS, Lennon JJ (1995) Mass resolution improvement by incorporation of pulsed ion extraction in a matrix-assisted laser desorption/ionization linear time-of-flight mass spectrometer. *Anal Chem* 67:1998–2003
57. Clauser KR, Baker P, Burlingame AL (1999) Role of accurate mass measurement (+/–10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal Chem* 71:2871–2882
58. Hager JW (2002) A new linear ion trap MS. *Rapid Commun Mass Spectrom* 16:512–526
59. Le Blanc JC, Hager JW, Ilisiu AM et al (2003) Unique scanning capabilities of a new hybrid linear ion trap mass spectrometer (Q TRAP) used for high sensitivity proteomics applications. *Proteomics* 3:859–869
60. Douglas D, Frank A, Mao D (2005) Linear ion traps in MS. *Mass Spectrom Rev* 24:1–29
61. Douglas DJ (1993) Multipole inlet systems for ion traps. US Patent 5,179,278
62. Syka JE, Marto JA, Bai DL et al (2004) Novel linear quadrupole ion trap/FT mass spectrometer: performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J Proteome Res* 3:621–626
63. Makarov A (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem* 72:1156–1162
64. Kingdon KH (1923) A method for the neutralization of electron space charge by positive ionization at very low gas pressures. *Phys Rev* 21:408–418
65. Knight RD (1981) Storage of ions from laser-produced plasmas. *Appl Phys Lett* 38:221–223
66. Scigelova M, Hornshaw M, Giannakopoulos A et al (2011) Fourier transform mass spectrometry. *Mol Cell Proteomics* 10(7). doi:10.1074/mcp.M11.009431-1. <http://www.mcponline.org>
67. Makarov A, Denisov E, Lange O (2009) Performance evaluation of a high-field Orbitrap mass analyzer. *J Mass Spectrom* 20:1391–1396
68. Makarov A, Denisov E, Kholomeev A et al (2006) Performance evaluation of a hybrid linear ion trap/Orbitrap mass spectrometer. *Anal Chem* 78:2113–2120
69. Makarov A, Horning S, Mann M (2007) Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods* 4:709–712
70. Syka JE, Coon JJ, Schroeder MJ et al (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci USA* 101:9528–9533
71. Olsen JV, Macek B, Lange O et al (2006) The utility of ETD mass spectrometry in proteomic analysis. *Biochim Biophys Acta* 1764:1811–1822
72. Strupat K, Kovtoun V, Bui H et al (2009) MALDI produced ions inspected with a linear ion trap/Orbitrap hybrid mass analyzer. *J Am Soc Mass Spectrom* 20:1451–1463
73. Geiger T, Cox J, Mann M (2010) Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Mol Cell Proteomics* 9:2252–2261
74. Gallien S, Duriez E, Crone C et al (2012) Targeted proteomic quantification on quadrupole–orbitrap mass spectrometer. *Mol Cell Proteomics* 11:1709–1723
75. Olsen JV, Schwartz JC, Griep-Raming J et al (2009) A new instrument for high-speed

- proteomics: Orbitrap mass analyzer interfaced to a dual linear trap. *Mol Cell Proteomics* 8:2759–2769
76. Albritton DL, Miller TM, Martin DW et al (1968) Mobilities of mass-identified H_3^+ and H^+ ions in hydrogen. *Phys Rev* 171:94–102
77. Shvartsburg AA, Smith RD (2008) Fundamentals of traveling wave ion mobility spectrometry. *Anal Chem* 80:9689–9699
78. Kanu AB, Dwivedi P, Tam M et al (2008) Ion mobility-mass spectrometry. *J Mass Spectrom* 43:1–22
79. McDaniel EW, Martin DW, Barnes WS (1962) Drift tube-mass spectrometer for studies of low-energy ion-molecule reactions. *Rev Sci Instrum* 33:2–7
80. Revercomb HE, Mason EA (1975) Theory of plasma chromatography/gaseous electrophoresis—a review. *Anal Chem* 47:970–983
81. Pringle SD, Giles K, Wildgoose JL et al (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int J Mass Spectrom* 261:1–12
82. Giles K, Pringle SD, Worthington KR et al (2004) Applications of a traveling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun Mass Spectrom* 18:2401–2414
83. Schneider BB, Covey TR, Coy SL et al (2010) Chemical effects in the separation process of a differential mobility/mass spectrometer system. *Anal Chem* 82:1867–1880
84. Krylov EV, Coy SL, Vandermey J et al (2010) Selection and generation of waveforms for differential mobility spectrometry. *Rev Sci Instrum* 81:024101
85. <http://www.absciex.com/Documents/Downloads/Literature/SelexION-Forensic-Tech-Note-3460111-01.pdf>
86. Kiss A, Heeren RMA (2011) Size, weight and position: ion mobility spectrometry and imaging MS combined. *Anal Bioanal Chem* 399:2623–2634
87. McLean JA, Ruotolo BT, Gillig KJ et al (2005) Ion mobility mass spectrometry: a new paradigm for proteomics. *Int J Mass Spectrom* 240:301–315
88. Zinnel NF, Pai P-J, Russell DH (2012) Ion mobility-mass spectrometry (IM-MS) for top-down proteomics: increased dynamic range affords increased sequence coverage. *Anal Chem* 84:3390–3397
89. Han X, Aslanian A, Yates JR III (2008) Mass spectrometry for proteomics. *Curr Opin Chem Biol* 12:483–490
90. Hart-Smith G, Blanksby SJ (2012) Mass analysis. In: Barner-Kowollik C, Gruendling T, Falkenhagen J, Weidner S (eds) *Mass spectrometry in polymer chemistry*, 1st edn. Wiley-VCH, Verlag GmbH & Co., Weinheim (Germany), pp 5–32
91. Brenton AG, Godfrey AR (2010) Accurate mass measurement: terminology and treatment of data. *J Am Soc Mass Spectrom* 21:1821–1835
92. Picotti P, Aebersold R (2012) Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat Methods* 9:555–566
93. Carapito C, Aebersold R (Guest editors) (2012) Targeted proteomics. *Proteomics* 8:1073–1300
94. Brunnée C (1987) The ideal mass analyzer: fact or fiction? *Int J Mass Spectrom* 76:125–237
95. McLafferty FW, Turecek F (1993) *Interpretation of mass spectra*, 4th edn. University Science Book, Mill Valley, CA. ISBN 978-0-935702-25-5
96. McCloskey JA (1990) Mass spectrometry. *Methods Enzymol* 193:3–960
97. Smith RM (2004) *Understanding mass spectra: a basic approach*, 2nd edn. Wiley, Hoboken, NJ. ISBN 0-471-42949-X
98. Gross JH (2004) *Mass spectrometry: a textbook*. Springer, Heidelberg. ISBN 3540407391
99. Watson JT, Sparkman OD (2008) *Introduction to mass spectrometry. Instrumentation, applications, and strategies for data interpretation*, 4th edn. Wiley, Chichester. ISBN 978-0-470-51634-8
100. Ma B, Johnson R (2012) *De novo* sequencing and homology searching. *Mol Cell Proteomics* 11:1–16
101. Nappi M, Weil C, Cleven CD et al (1997) Visual representations of simulated three-dimensional ion trajectories in an ion trap mass spectrometer. *Int J Mass Spectrom Ion Proc* 161:77–85
102. Schweikhard L, Alber GM, Marshall AG (1992) FT/ICR of highly-charged atomic ions. *Phys Scripta* 46:598–602

Part II

Experimental

Chapter 7

A Protocol for Protein Extraction from Lipid-Rich Plant Tissues Suitable for Electrophoresis

Agnieszka Zienkiewicz, Juan David Rejón, Juan de Dios Alché, María Isabel Rodríguez-García, and Antonio Jesús Castro

Abstract

Plant tissues contain high levels of nonprotein contaminants such as lipids, phenolic compounds, and polysaccharides among others, which interfere with protein extraction and electrophoretic separation. Preparation of good-quality protein extracts is a critical issue for successful electrophoretic analysis. Here, we describe a three-step method for protein extraction from lipid-rich plant tissues, which is suitable for both 1-D and 2-D electrophoresis and is compatible with downstream applications. The protocol includes prefractionation, filtration, and TCA/acetone precipitation steps prior to protein resolubilization.

Key words Lipid removal, Lipid-rich plant tissues, Protein extraction, TCA/acetone precipitation, Two-dimensional electrophoresis

1 Introduction

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) is a powerful and widely used method to study complex mixtures of hundreds or even thousands of different proteins extracted from cells, tissues, or complete organisms [1–3]. Proteins are separated on a 2-D matrix on the basis of their charge (IEF, first dimension) and size (SDS-PAGE, second dimension) [1]. Sample preparation is a critical factor for achieving good and reproducible 2D-PAGE results. Protein extraction from plant tissues is challenging due to structural (i.e., presence of a cell wall) and chemical (i.e., production of a broad spectrum of metabolites such as pigments, lipids, phenolic compounds, polysaccharides) reasons [4–9].

Oleaginous plants (e.g., sunflower, olive, rapeseed) accumulate high amounts of storage neutral lipids (e.g., TAGs) for energy supply during periods of active metabolism [10]. These compounds are confined to specialized organelles called oil bodies [11], which can be found in different plant tissues like seeds [12], anthers [13],

and pollen [13, 14]. As monomers, lipids can bind proteins, reducing their solubility and changing the properties (pI and molecular mass) used for their electrophoretic separation [15]. As supramolecular assemblies (e.g., oil bodies), an excess of lipids can cause insufficient detergency, and decreased efficiency of the detergent as solubilizing agent [15]. This problem can be overcome by scaling up the separation, and therefore diluting the sample, or by means of chemical delipidation using organic solvents such as chloroform and acetone [16]. However, this method may lead to severe loss of proteins either because some proteins are soluble in organic solvents or because precipitated proteins do not properly resolubilize after pelletizing [15, 17].

Here, we describe a method for cleaning up protein extracts prepared from lipid-rich plant tissues that is suitable for both 1-D and 2-D electrophoresis and is compatible with downstream applications. We routinely use this method in our research with olive pollen and seeds but it is also effective with other tissues and plants. Delipidation of protein extracts is achieved in three steps by (1) centrifugation and lipid pad removal, (2) filtration through a Sephadex G25 matrix-containing column by gravity, and (3) TCA/acetone precipitation prior to protein resolubilization. Lipid removal following this method results in high-resolution and reproducible 1-D and 2-D gels of proteins (Fig. 1).

2 Materials

Prepare all solution using ultrapure water (18 MΩ cm at 25 °C) and analytical grade reagents. Use dust-free gloves and clean glassware and equipment. Proceed as quickly as possible in order to minimize the time of handling.

2.1 Tissue Homogenization and Protein Extraction

1. Liquid nitrogen.
2. Mortar and pestle (*see Note 1*).
3. Extraction buffer: 0.05 M Tris–HCl, pH 7.4 (*see Note 2*), 1 % (v/v) Triton X-100, 4 % (w/v) 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS) (*see Note 3*), 40 mM dithiothreitol (DTT), 5 % (w/v) polyvinylpyrrolidone (PVPP) (*see Note 4*), and 10 μL/mL plant protease inhibitor cocktail (catalog number: P9599, Sigma-Aldrich Corp., St. Louis, MO, USA) (*see Note 5*).
4. Magnetic stirrer.
5. Refrigerated centrifuge 5810R (catalog number: 5810 000.017, Eppendorf, Hamburg, Germany) or similar, equipped with a rotor for 2 mL microcentrifuge tubes.
6. Syringes (3 mL) with very fine needles (purchased at drugstore).

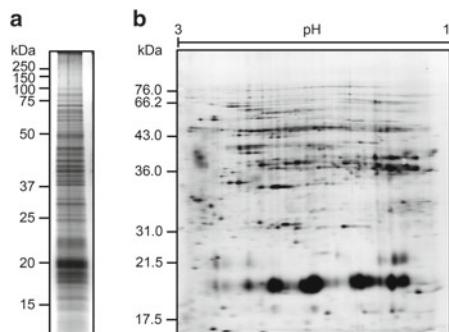


Fig. 1 (a) Representative 1-D gel of mature olive pollen proteins. Approximately 25 μ g of pollen proteins were loaded on 12 % (w/v) polyacrylamide gels with 4 % stacking gels and proteins were electrophoresed using a Mini-Protean 3 apparatus (catalog number: 165-3301, Bio-Rad Corp., Hercules, CA, USA) as previously described [18]. Proteins were stained with BioSafe Coomassie Stain (catalog number: 161-0786, Bio-Rad Corp., Hercules, CA, USA) following the manufacturer's instructions. Protein markers (catalog number: 161-0373, Bio-Rad Corp.) are displayed on the *left*. (b) Representative 2-D gel of the mature olive pollen proteins containing 1,021 spots. The global protein profile was largely reproducible among replicas. Approximately 100 μ g of total protein was loaded on 18 cm polyacrylamide strips, pH 3–10 NL (catalog number: 163-2033, Bio-Rad Corp.) by in-gel rehydration at 30 V for 12 h. IEF was carried out at 20 °C in a Protean IEF Cell (catalog number: 165-4000, Bio-Rad Corp.) as follows: 300 and 1,000 V for 1 h each followed by a linear increase from 1,000 to 10,000 V, and finally 10,000 V to give a total of 40 kWh. Reduction and alkylation steps were performed as previously described [19]. Protein separation in the second dimension was carried out by SDS-PAGE in a Protean II xi Cell (catalog number: 165-1951, Bio-Rad Corp.) as previously described [20]. Proteins were stained with BioSafe Coomassie Stain (Bio-Rad Corp.) following the manufacturer's instructions. 2-D SDS-PAGE standards (catalog number: 161-0320, Bio-Rad Corp.) are displayed on the *left*

2.2 Protein Extract Filtration

1. PD-10 Desalting Columns (catalog number: 17-0851-01, GE Healthcare Biosciences AB, Uppsala, Sweden) or similar (*see Note 6*).
2. Equilibration buffer: 0.05 M Tris–HCl, pH 7.4.
3. Elution buffer: 0.01 M Tris–HCl, pH 7.4.

2.3 Protein Precipitation

1. Precipitation solution: 20 % (v/v) trichloroacetic acid (TCA) and 0.2 % (w/v) DTT in chilled (-20 °C) acetone (*see Note 7*).
2. Washing solution: 0.2 % (w/v) DTT prepared in chilled (-20 °C) acetone.
3. Refrigerated centrifuge Sorvall RC-6 Plus (catalog number: 46910, Thermo Fisher Scientific Inc., Waltham, MA, USA) or similar, equipped with a rotor for 50 mL centrifuge tubes.
4. Oak ridge centrifuge tubes (50 mL), polypropylene copolymer (catalog number: 3119-0050, NALGENE Labware, Thermo Fisher Scientific Inc.).

2.4 Protein Resuspension and Solubilization

1. Solubilization buffer: 7 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 40 mM DTT, 0.5 % (v/v) carrier ampholytes pH 3–10, and 0.002 % (w/v) bromophenol blue.
2. Refrigerated ultracentrifuge Beckman TL-100 (catalog number: 346457, Beckman Instruments, Palo Alto, CA, USA) or similar, equipped with a rotor for 2 mL centrifuge tubes.

3 Methods

3.1 Tissue Homogenization and Protein Extraction

1. Homogenize plant tissues (0.1 g) to a very fine powder in a liquid nitrogen-precooled mortar by using a pestle.
2. Transfer the powder to a 2 mL microcentrifuge tube and add 1.5 mL of extraction buffer. Mix well and stir in a magnetic stirrer for 1–6 h at 4 °C.
3. Centrifuge at $20,000 \times g$ for 30 min at 4 °C.
4. After centrifugation, collect the aqueous supernatant using a syringe with a very fine needle and without disturbing the upper lipid pad formed during centrifugation (Fig. 2).

3.2 Protein Filtration

1. Equilibrate the column with 25 mL of equilibration buffer. Discard the flow-through.
2. Add sample to the column and allow the sample to enter the packed bed completely (see Note 8). If sample volume is less than 2.5 mL, add equilibration buffer to adjust the volume up

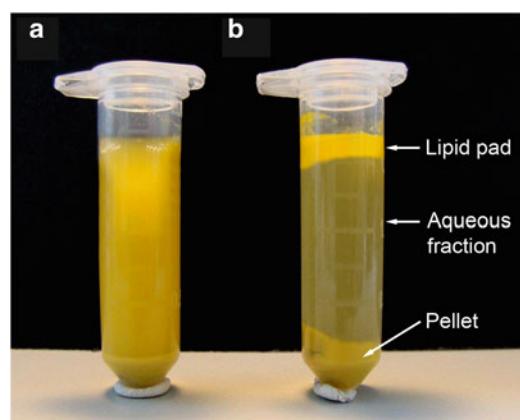


Fig. 2 (a) Protein extracts prepared from olive pollen contain large amounts of lipids. (b) After centrifugation, three fractions are distinguishable: a pellet of cell debris at the *bottom*, an aqueous supernatant in the *middle*, and a lipid pad floating as a *top layer*. The protein-containing aqueous fraction can be carefully collected using a syringe with a very fine needle

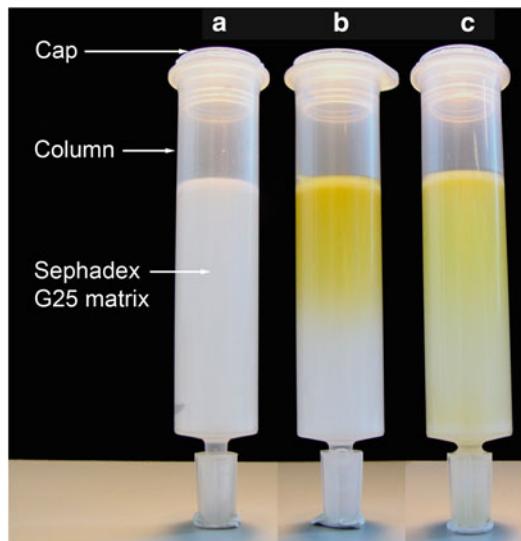


Fig. 3 Filtration of protein extracts prepared from olive pollen through a column containing Sephadex G25 medium (PD-10 column, GE Healthcare Biosciences AB). **(a)** PD-10 column after equilibration and before sample loading. **(b)** and **(c)** The same column as above after sample filtration **(b)** and protein elution **(c)**, respectively

to 2.5 mL after sample enters the SephadexTM matrix completely (Fig. 3). Discard the flow-through.

3. Place a 50 mL centrifuge tube under the column. Elute proteins with 3.5 mL of elution buffer and collect the flow-through.
1. Add 9 volumes of precipitation solution to sample and allow proteins to precipitate at -20°C for 1 h (*see Note 9*).
2. Centrifuge at $20,000\times g$ for 30 min at 4°C .
3. Discard the supernatant and rinse the pellet in washing solution. Spin at $12,000\times g$ for 15 min.
4. Repeat the washing step two times more.
5. After centrifugation, let sample to dry under a fume hood for 5–10 min at room temperature to ensure the total removal of acetone (*see Note 10*).

3.3 Protein Precipitation

3.4 Protein Resuspension and Solubilization

1. Resuspend the resulting pellet in an appropriate volume of solubilization buffer (*see Note 11*).
2. Centrifuge at $100,000\times g$ for 1 h at 4°C to remove insoluble particulates from the sample (*see Note 12*).
3. Quantify extracts and make aliquots. Store aliquots at -80°C until use.

4 Notes

1. Cell lysis can be achieved by different methods. Plant cells possess tough cell walls that are difficult to disrupt. The mortar and pestle procedure described above is usually the method of choice for plant tissues but ultrasonic disintegrators and homogenizers (e.g., blender) can be also used. Proceed as quickly as possible and avoid the homogenate to thaw.
2. Tris buffer can be added when basic conditions are required for full solubilization or to minimize proteolysis.
3. Detergents help to disrupt membranes, solubilize lipids, and delipidate proteins bound to vesicles or membranes. CHAPS is especially useful in purifying **membrane proteins** and can be used in combination with other **nonionic** detergents such as **Triton X-100**.
4. PVPP-bound polyphenols are removed during the centrifugation step.
5. During cell lysis, proteases are released or activated. Protease inhibitors minimize protein degradation during protein extraction and solubilization. The plant protease inhibitor cocktail used contains a mixture of molecules with broad specificity for the inhibition of serine-, cysteine-, aspartic-, and metalloproteases, and aminopeptidases. Protease inhibitors will be displayed in 1-D and 2-D polyacrylamide gels.
6. These columns contain Sephadex™ G-25 medium. They can be used in a wide range of applications such as desalting, buffer exchange, and cleaning up of samples.
7. The mixture TCA/acetone is more effective in precipitating proteins than either TCA or acetone alone.
8. Desalting can be achieved by gravity or centrifugation. There is a slightly higher recovery and desalting capacity using gravity protocol (described above) but sample is diluted 1.4 times.
9. Overnight precipitation results in a higher protein recovery.
10. This step is critical. An insufficient drying will result in an incomplete acetone removal, whereas protein pellets will be hardly resuspended and solubilized after an extended drying.
11. In this step, proteins can be concentrated using low volumes of solubilization buffer.
12. Insoluble particulates cause smearing and block gel pores during 2-D electrophoresis.

Acknowledgements

This work was supported by the Spanish Ministry of Science and Innovation (MICINN) (ERDF-cofinanced projects AGL2008-00517, BFU2011-22779, and PIE-200840I186) and the Junta de Andalucía (ERDF-cofinanced projects P2010-CVI5767 and P2010-AGR6274). A. Zienkiewicz thanks the CSIC for providing JAEdoc grant funding.

References

1. O'Farrel PH (1975) High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 250:4007–4021
2. Görg A, Weiss W, Dunn MJ (2004) Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 4:3665–3685
3. Carrette O, Burkhard PR, Sanchez JC et al (2006) State-of-the-art two-dimensional gel electrophoresis: a key tool for proteomics research. *Nat Protoc* 1:812–823
4. Giavalisco P, Nordhoff P, Lehrach H et al (2003) Extraction of proteins from plant tissues for two-dimensional electrophoresis analysis. *Electrophoresis* 24:207–216
5. Saravanan RS, Rose JKC (2004) A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics* 4:2522–2532
6. Carpentier SC, Witters E, Laukens K et al (2005) Preparation of protein extracts from recalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis. *Proteomics* 5:2497–2507
7. Song J, Braun G, Bevis E et al (2006) A simple protocol for protein extraction of recalcitrant fruit tissues suitable for 2-DE and MS analysis. *Electrophoresis* 27:3144–3151
8. Wang W, Vignani R, Scali M et al (2006) A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. *Electrophoresis* 27:2782–2786
9. Wang W, Tai F, Chen S (2008) Optimizing protein extraction from plant tissues for enhanced proteomic analysis. *J Sep Sci* 31:2032–2039
10. Murphy DJ (2001) Biogenesis and functions of lipid bodies in animals, plants and microorganisms. *Prog Lipid Res* 40:325–438
11. Huang AHC (1996) Oleosin and oil bodies in seeds and other organs. *Plant Physiol* 110:1055–1061
12. Zienkiewicz A, Jiménez-López JC, Zienkiewicz K et al (2011) Development of the cotyledon cells during olive (*Olea europaea* L.) *in vitro* seed germination and seedling growth. *Protoplasma* 248:751–756
13. Zienkiewicz K, Zienkiewicz A, Rodríguez-García MI et al (2011) Characterization of a caleosin expressed during olive (*Olea europaea* L.) pollen ontogeny. *BMC Plant Biol* 11:122
14. Zienkiewicz K, Castro AJ, Alché JD et al (2010) Identification and localization of a caleosin in olive (*Olea europaea* L.) pollen during *in vitro* germination. *J Exp Bot* 61:1537–1546
15. Rabilloud T (1996) Solubilization of proteins for electrophoresis analyses. *Electrophoresis* 17:813–829
16. Wang W, Vignani R, Scali M et al (2004) Removal of lipid contaminants by organic solvents from oilseed protein extract prior to electrophoresis. *Anal Biochem* 329:139–141
17. Radin NS (1981) Extraction of tissue lipids with a solvent of low toxicity. *Methods Enzymol* 72:5–7
18. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685
19. Görg A, Postel W, Günther S (1988) Two-dimensional electrophoresis. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 9:531–546
20. Castro AJ, Bednarczyk A, Schaeffer-Reiss C et al (2010) Screening of Ole e 1 polymorphism among olive cultivars by peptide mapping and N-glycopeptide analysis. *Proteomics* 5:953–962

Chapter 8

Making a Protein Extract from Plant Pathogenic Fungi for Gel- and LC-Based Proteomics

Raquel González Fernández, Inmaculada Redondo, and Jesus V. Jorrín-Novo

Abstract

Proteomic technologies have become a successful tool to provide relevant information on fungal biology. In the case of plant pathogenic fungi, this approach would allow a deeper knowledge of the interaction and the biological cycle of the pathogen, as well as the identification of pathogenicity and virulence factors. These two elements open up new possibilities for crop disease diagnosis and environment-friendly crop protection. Phytopathogenic fungi, due to its particular cellular characteristics, can be considered as a recalcitrant biological material, which makes it difficult to obtain quality protein samples for proteomic analysis. This chapter focuses on protein extraction for gel- and LC-based proteomics with specific protocols of our current research with *Botrytis cinerea*.

Key words Fungal proteomics, Fungal secretome, Gel-based proteomics, LC-based proteomics

1 Introduction

Phytopathogenic fungi are one of the most damaging plant parasitic organisms that cause serious diseases and remarkable yield losses in crops. The biological study of these microorganisms and the interaction with their hosts have experienced great advances in recent years due to the development of modern, holistic, and high-throughput -omic techniques, together with the increasing number of genome sequencing projects and the development of mutants and reverse genetics tools. Within these -omic techniques, proteomics has become a relevant tool in plant–fungus pathosystem research (reviewed in [1–4]). Molecular studies of the fungal biological cycle and their interaction with their hosts are necessary for searching key protein targets, and for developing new more efficient and environment-friendly agrochemicals [5, 6], which may open new ways for crop disease diagnosis and protection. Thus, proteomics aims to identify gene products with a key role in pathogenicity and virulence.

The use of proteomics also allows location-specific analyses (i.e., subproteomes at the level of organelles, cell membranes, cell wall, secretory proteins), the study of posttranslational modifications [7], as well as the study of interactions of host-pathogen [8] or host-pathogen-biocontrol agents [9, 10].

Proteomics involves the combined application of (a) advanced gel-based separation, such as one- and two-dimensional electrophoresis (1-DE and 2-DE), or gel-free based in liquid chromatography (LC) techniques; (b) identification techniques based on mass-spectrometry (MS) analysis; and (c) bioinformatics tools to characterize the proteins in complex biological mixtures [3, 4].

Different fields can be defined in proteomics, including descriptive and differential expression proteomics. In the case of fungi, a new area can also be defined as secretomics (the secretome is defined as the combination of native proteins and cell machinery involved in their secretion), since many fungi secrete an arsenal of proteins to accommodate their saprotrophic lifestyle, such as proteins implicated in the adhesion to the plant surface, host-tissue penetration, and invasion effectors, together with other virulence factors [11]. Over the last years, there has been a great advance in fungal proteomic research due to the availability of powerful proteomics technologies and the increasing number of fungal genome sequencing projects. Currently, more than 20 plant pathogenic fungal genomes have been sequenced (Broad Institute Database, <http://broadinstitute.org/science/project/fungal-genomeinitiative>), and excellent reviews of fungal proteomics methodologies have been recently published [3, 4, 12]. The workflow of a fungal gel-based proteomics experiment includes, among others, the following steps (Fig. 1): experimental design, fungal growth, sampling, sample preparation, protein extraction, separation, MS analysis, protein identification, statistical analysis of data, quantification, and data analysis, management, and storage.

Most plant pathogenic fungi are filamentous and can be considered, just like plants, as recalcitrant biological material due to their robust cell wall accounting for 50–60 % of polysaccharides (glucans), 20–30 % of glycoproteins (mannoproteins), and 10–20 % of chitin [13]. Protein sample preparation is a critical step. The cell breakdown and later protein extraction are difficult because of the presence of a cell wall that makes up the majority of the cell mass [13]. Cell disruption can be performed using mechanical lysis via glass beads [14–16], with a cell mill [17] or by sonication [18–20]. These methods are more efficient than those based on chemical or enzyme extraction [21]. An alternative approach to avoid the difficulty of lysing the fungal cell wall might be the generation of protoplasts (cells whose wall has been completely or partially removed using either mechanical or enzymatic means) [22]. The most widely used method for cell disruption is pulverizing the mycelium in liquid nitrogen using a mortar and pestle [3, 4].

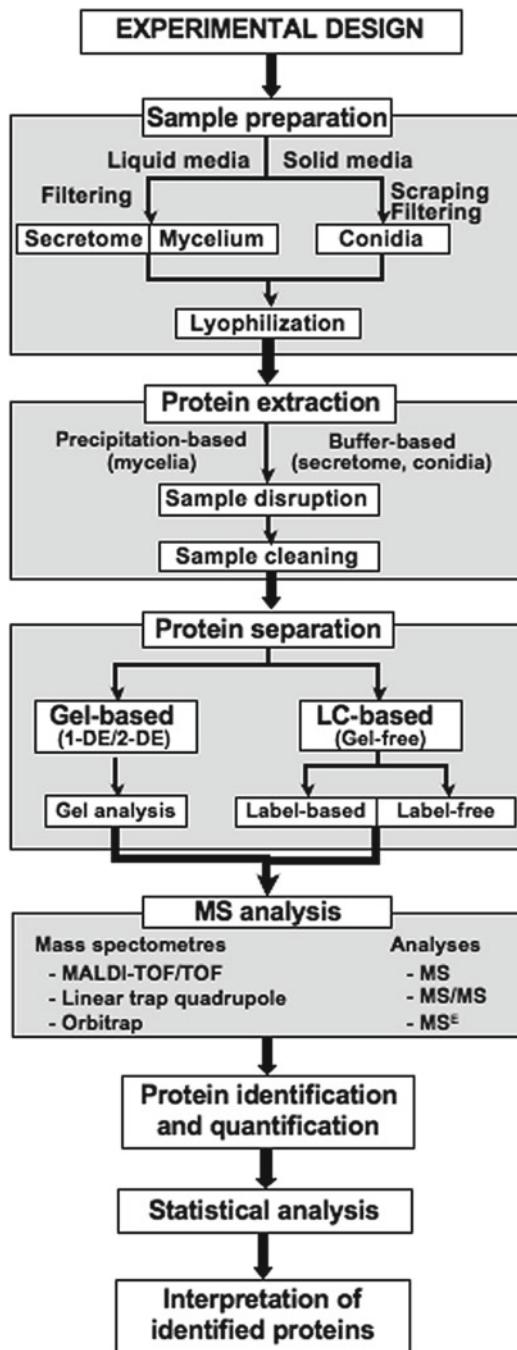


Fig. 1 Overview of proteomics workflow in plant pathogenic fungi

The production of high-quality protein samples is also crucial for proteomic analysis. The most widely employed protocol for fungal proteins uses protein precipitation media containing organic solvents, such as trichloroacetic acid (TCA), followed by solubilization of the

precipitate in an appropriate buffer. This method minimizes protein degradation/modification. Furthermore, it removes interfering compounds such as polysaccharides, polyphenols, pigment, and lipids, which may cause problems during IEF [23], and prevents protease activities [24]. TCA treatment complicates subsequent protein solubilization for IEF, especially with hydrophobic proteins. These problems have been partially overcome by the use of chaotropes (urea and thiourea) [25], new zwitterionic detergents [26–30], and a brief treatment with sodium hydroxide [24] that led to an increase in resolution and capacity of 2-DE gels. Other protein extraction methods have reported an improvement when using an acidic extraction solution to reduce streaking of fungal samples caused by their cell wall [31], as well as with the use of a phosphate buffer solubilization before the precipitation [32, 33]. Finally, the combined use of TCA precipitation and phenol extraction provides a better spot definition due to the fact that it reduces streaking and leads to a higher number of detected spots [34–36]. Alternative protocols for protein extraction from spores of *Aspergillus* ssp. have been optimized, since they use acidic conditions, step organic gradient, and variable sonication treatments (ultrasonic homogenizer and sonic water bath) [20].

Special protocols are required for secreted proteins due to the fact that there may be problems like a very low protein concentration that are sometimes below the detection limit of colorimetric methods (Bradford, Lowry, or BCA), or the presence of polysaccharides, mucilaginous material, salts, and secreted metabolites (low-molecular organic acids, fatty acids, phenols, quinones, and other aromatic compounds). The presence of these extracellular compounds may impair standard methods for protein quantification and may result in a strong overestimation of total protein number [37]. This determination can also be affected by the high concentration of reagents from the solubilization buffer (i.e., urea, thiourea, or DTT) that may interfere in the spectrophotometric measurement, producing an overestimation of the total amount of protein in which, depending on the method, the differences vary in the order of two magnitudes [38]. Comparisons of different standard methods for protein precipitation have demonstrated their limited applicability in analyzing the whole fungal secretome [39–44].

Electrophoresis is the most employed technique for protein separation in fungal research. Despite its simplicity, 1-DE remains as a useful technique that provides relevant information, especially in the case of comparative proteomics with large number of samples to analyze. It is possible by using this technique to distinguish between phenotypes of different wild-type strains of *B. cinerea*, and to identify proteins involved in the pathogenicity mechanisms (Fig. 2) [45, 46]. With appropriate software tools, 1-DE is a simple, reliable technique for fingerprinting of crude extracts, and it is especially useful in the case of hydrophobic and low-molecular-weight proteins [47].

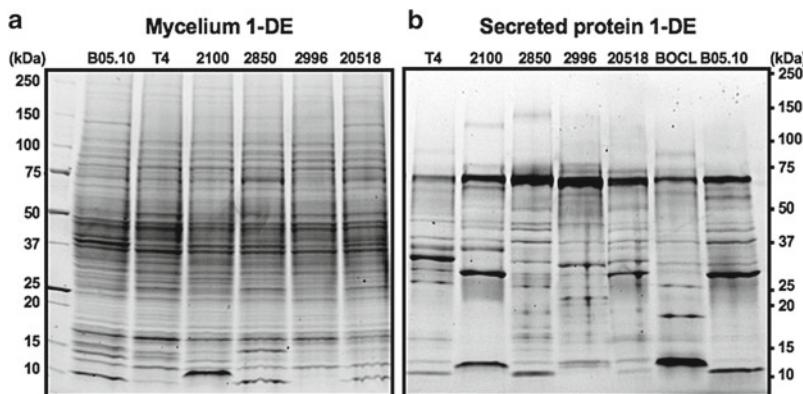


Fig. 2 Protein profiles of seven *B. cinerea* wild-type strains (B05.10, T4, 2100, 2850, 2996, 20518, and BOCL) that differ both in host and virulence from (a) mycelium and (b) secreted proteins. This approach allows the assessment of differences in protein band patterns among strains

Two-DE is the dominant platform in fungal proteomics. The 2-DE consists of a tandem pair of electrophoretic separations: in the first dimension, proteins are resolved according to their isoelectric points (pIs), normally using IEF, while in the second dimension, the proteins are separated according to their approximate molecular weight using SDS-PAGE. Excellent reviews describing and discussing the features and protocols of electrophoretic separations in proteomics strategies have been published [23, 48]. Two main advantages of 2-DE can be emphasized: (a) its high protein separation capacity, and (b) the possibility of making large-scale protein-profiling experiments. Nevertheless, the reproducibility and resolution of this technique are still remaining challenges. This method was reported to under-represent proteins with extreme physicochemical properties (size, isoelectric point, transmembrane domains), as well as those with a low abundance [49].

After separating proteins, they can be detected using different staining techniques [23, 48], namely, (a) organic dyes, like colloidal Coomassie Blue staining, (b) zinc-imidazole staining, (c) silver staining, and (d) fluorescence-based detection, like Sypro Ruby. The criteria used to choose the staining method are the level of sensibility and its compatibility with MS. Gels are digitized, and bands or spots are studied by specific image analysis software (i.e., Quantity-One, PD-Quest, BioRad). Bands or spots are excised from gels and prepared for MS analysis.

The limitations of gel-based analysis have led to the more recent development of techniques based on LC separation of proteins or peptides, including two-dimensional liquid-phase chromatography 2-D LC-MS/MS (based on a high-performance chromatofocusing in the first dimension followed by high-resolution reversed-phase chromatography in the second) [50], and 1-DE-nanoscale capillary LC-MS/MS, like GeLC-MS/MS

(this technique combines a size-based protein separation with an in-gel digestion of the resulting fractions) [51]. This GeLC-MS/MS strategy paves the way towards the analysis on large-scale fungal response environmental cues on the basis of quantitative shotgun protein-profiling experiments. The case of multidimensional protein identification technology (MudPIT), which allows the identification of a much larger number of proteins compared to gel-based methods, has been a drawback due to the lack of quantitative data [52, 53]. MudPIT was used to analyze germling growth mechanisms in *Uromyces appendiculatus* by comparing germinating asexual uredospores to inactive spores [54].

MS is the basic technique for global proteomic analysis due to its accuracy, resolution, and sensitivity (in the femtomole to attomole concentration range), and due to the fact that it has the capacity for a high throughput. Not only does it allow profiling a proteome, but also and more important, it allows the identification of protein species and the characterization of posttranslational modifications and interactions. Proteins are identified from mass spectra of intact proteins (top-down proteomics), or peptide fragments obtained after enzymatic (mostly digested with trypsin) or chemical treatment (bottom-up proteomics). Protein species are identified by comparison of the experimental spectra, while the theoretical ones were obtained *in silico* from protein, genomic, EST sequence, or MS spectra databases. For this purpose, different instrumentation, algorithms, databases, and repositories are available [55, 56].

Although 2-DE remains as a standard tool for fungal proteomic research, current efforts are focussed on alternative gel-free shotgun strategies to identify and quantify proteins. The coupling of nanoscale separations (nanocapillary or nLC) with automated MS/MS has enhanced the development of this methodology. Using LC MS/MS, the complex mixtures of proteins are digested to peptides (normally with trypsin), which are separated according to their hydrophobicity by nLC, and then the eluted peptides are introduced into the mass spectrometer [57]. For example, a gel-free analysis from mycelium and secreted proteins of *B. cinerea* B05.10 and T4 strains has been carried out using a SYNAPT HDMS mass spectrometer (Waters) interfaced with a NanoAcuity UPLC System (Waters) [45, 59, 60]. A total of 197 and 73 proteins were identified from mycelia and secreted proteins, respectively (Fig. 3). Recent reviews study techniques, software, and statistical analyses used in gel-free quantitative proteomics, and discuss about its strengths and limitations [3, 61, 62].

Although several methods for proteomic analysis of limited fungal species have been published [4, 12, 63, 64], procedures for protein extraction as well as gel-based and gel-free analysis conditions are progressively evolving to study individual characteristics of fungal species.

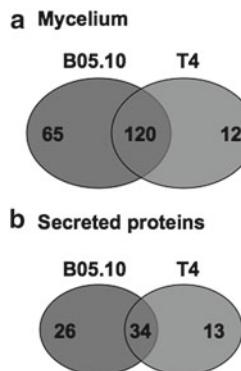


Fig. 3 Venn diagram of B05.10 versus T4 *B. cinerea* strains of proteins identified by using gel-free approach from (a) mycelium extracts and (b) secreted proteins

2 Materials

2.1 Fungal Strains

This protocol has been carried out with different *B. cinerea* strains: B05.10 and T4 (provided by Dr. Julia Schumacher, Prof. Dr. Paul Tudzynski of the Institute of Biology and Biotechnology of Plants, Westfälische Wilhelms-Universität, Münster, Germany), together with CECT2100, CECT2850, CECT2996, and CECT20518 (provided by the Spanish Type Culture Collection).

2.2 Reagents, Solutions, and Buffers

Analytical grade reagents are used, unless other grades are specified. The prepared solutions are kept at 4 °C or -20 °C if indicated. Reagents and solutions must be discarded once used, according to current regulations. It is mandatory to use a fume cupboard when working with volatile or dangerous compounds. Personal protection elements (i.e., lab coats, gloves, glasses) must be used (see Note 1).

1. 10 % (w/v) TCA in 80 % (v/v) acetone.
2. 0.1 M ammonium acetate in 100 or 80 % methanol.
3. 80 % (v/v) acetone.
4. Phenol solution equilibrated with 10 mM Tris-HCl, pH 8 (P4557, Sigma).
5. SDS buffer: 0.1 M Tris-HCl, pH 8, 30 % (w/v) sucrose, 2 % (w/v) SDS, 5 % (v/v) β-mercaptoethanol.
6. Solubilization solution: 9 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 0.5 % (v/v) Triton X-100, 20 mM DTT.
7. Bradford solution (B6916, Sigma).
8. Extraction buffer: 50 mM Tris-HCl, pH 8, 8 M urea, 1 % (w/v) SDS, 1 mM EDTA, 100 mM DTT.
9. TE buffer for secreted proteins: 10 mM Tris-HCl, pH 8, 1 mM EDTA, 2 % (v/v) β-mercaptoethanol, 1 mM PMSF, 10 µL/mL buffer of protease inhibitor cocktail for fungi (P8215, Sigma).

10. Electrophoresis buffer: 50 mM Tris–HCl, pH 8, 192 mM glycine, 1 % (w/v) SDS.
11. Stain-Free Precast Gels (Criterion System, BioRad): 4–20 % Tris–HCl multi-wells for 1-DE and 8–16 % Tris–HCl IPG+1 for 2-DE.
12. IPG strips, 11 cm, pH 5–8 (BioRad).
13. IPG strip rehydration solution: 7 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 2 % (v/v) ampholytes (BioRad), 20 mM DTT.
14. Equilibration buffer: 1.5 M Tris–HCl, pH 8.8, 6 M urea, 20 % (v/v) glycerol, 2 % (w/v) SDS.
15. Distilled water.
16. Liquid nitrogen.

2.3 Equipment

1. Freeze-dryer.
2. Mortar and pestle.
3. Cell strainer, 100 µm nylon.
4. Vortexer.
5. Micropestles.
6. Ultrasonic homogenizer.
7. Microcentrifuge and centrifuge.
8. Disposable microcentrifuge tubes: 1.5 and 2.0 mL.
9. Centrifuge tubes: 50 mL.
10. Microtube mixer.
11. CriterionTM Cell (Biorad).
12. Criterion Stain Free Imager and Image LabTM software (Biorad).
13. Shaker.
14. GS-800TM Calibrated Densitometer and Quantity One[®] 1-D Analysis software (BioRad).

3 Methods

The methods described below have been optimized to mycelium, secreted proteins in liquid media, and conidia from *Botrytis cinerea*, although these procedures can be applied to proteomic analysis of filamentous fungi in general.

3.1 Sample Collection

For in vitro cultures, conidia are produced using rich-media plates at 22 °C under constant black light (UV) during 3–4 weeks. Mycelium and secreted proteins can be obtained from liquid cultures inoculated with conidia or non-sporulating mycelia (see Note 2). Mycelia and media can be separated by centrifugation and filtration, frozen in liquid nitrogen, or lyophilized. At least three biological replicates should be collected [65].

3.2 Protein Extraction by TCA/Acetone–Phenol/Methanol

3.2.1 Mycelium

Protein extraction is carried out by using the TCA/acetone–phenol/methanol method [36, 66] with some modifications [12, 45], and adapted to the initial material (conidia, mycelium, or secreted proteins).

The lyophilized mycelium is ground to a fine powder in liquid nitrogen using a cooled mortar and pestle, and then it is stored at $-80\text{ }^{\circ}\text{C}$ for later analysis (*see Note 3*). The following protocol is used for protein extraction:

1. Transfer 50–100 mg of mycelial powder into a 2-mL tube.
2. Add 1 mL of 10 % (w/v) TCA/acetone and mix well firstly using a micropesle, and secondly by vortexing.
3. Sonicate $3 \times 10\text{ s}$ (50 W, amplitude 60) at $4\text{ }^{\circ}\text{C}$, breaking on ice at 1 min.
4. Fill the tube with 10 % (w/v) TCA/acetone. Mix well by vortexing.
5. Centrifuge at $16,000 \times g$ for 5 min ($4\text{ }^{\circ}\text{C}$) and remove the supernatant by decanting (*see Note 4*).
6. Fill the tube with 0.1 M ammonium acetate in 80 % (v/v) methanol. Mix well by vortexing.
7. Centrifuge at $16,000 \times g$ for 5 min ($4\text{ }^{\circ}\text{C}$) and discard the supernatant.
8. Fill the tube with 80 % (v/v) acetone. Mix well by vortexing.
9. Centrifuge at $16,000 \times g$ for 5 min ($4\text{ }^{\circ}\text{C}$) and discard the supernatant.
10. Air-dry the pellet at room temperature to remove residual acetone.
11. Add 1.2 mL of 1:1 phenol (pH 8, SIGMA)/SDS buffer. Mix well by vortexing and using a pipette. Incubate for 5 min on ice.
12. Centrifuge at $16,000 \times g$ for 5 min. Transfer the upper phenol phase into a new 1.5-mL tube (*see Note 5*).
13. Fill the tube with 0.1 M ammonium acetate in 100 % (v/v) methanol, mix well, and incubate the precipitation overnight at $-20\text{ }^{\circ}\text{C}$.
14. Centrifuge at $16,000 \times g$ for 5 min ($4\text{ }^{\circ}\text{C}$) and discard the supernatant (a white pellet should be visible).
15. Wash the pellet with 100 % methanol, and mix by vortexing.
16. Centrifuge at $16,000 \times g$ for 5 min ($4\text{ }^{\circ}\text{C}$), and discard the supernatant.
17. Wash the pellet with 80 % (v/v) acetone, and mix by vortexing.
18. Centrifuge at $16,000 \times g$ for 5 min ($4\text{ }^{\circ}\text{C}$), and discard the supernatant.

19. Air-dry the pellet at room temperature.
20. Dissolve the proteins with a solubilization solution by shaking for 2 h in a microtube mixer at 4 °C (see **Note 6**).
21. Quantify proteins using the Bradford method [67].
22. Store the protein extracts at -20 °C for further analysis.

3.2.2 Secreted Proteins

Lyophilized media are re-solubilized in 5 mL of TE buffer, and proteins are precipitated using the following protocol:

1. Transfer the re-solubilized medium into a 50-mL tube and add 2/1 (v/v) (10 mL) of 20 % (w/v) TCA/acetone. Mix well by vortexing and allow protein precipitation overnight at 4 °C.
2. Centrifuge at 16,000×*g* for 10 min (4 °C) and remove the supernatant by decanting (see **Note 7**).
3. Add 4/1 (v/v) (20 mL) of 0.1 M ammonium acetate in 80 % (v/v) methanol. Mix well by vortexing.
4. Centrifuge at 16,000×*g* for 10 min (4 °C) and discard the supernatant.
5. Add 4/1 (v/v) (20 mL) of 80 % (v/v) acetone. Mix well by vortexing.
6. Centrifuge at 16,000×*g* for 10 min (4 °C) and discard the supernatant.
7. Air-dry the pellet at room temperature to remove residual acetone.
8. Add 1.5 mL of 1/1 (v/v) phenol (pH 8, Sigma)/SDS buffer. Mix well by vortexing and transfer the results into a new 1.5-mL Eppendorf tube. Incubate for 5 min on ice.
9. Centrifuge at 16,000×*g* for 10 min. Transfer the upper phenol phase into a new 2-mL tube.
10. Fill the tube with 0.1 M ammonium acetate in 100 % (v/v) methanol, mix well, and allow the precipitation overnight at -20 °C.
11. Follow the steps in Subheading 3.2.1 (starting from **step 15**).

3.2.3 Conidia

Conidia can be harvested from H₂O with 0.01 % Tween-80 scraping on the surface of an agar plate. The conidial suspension is filtered through a cell strainer, concentrated in 1.5-mL tubes, centrifuged at 16,000×*g* for 5 min (4 °C), lyophilized, and stored at -80 °C for further analyses. For protein extraction, a buffer-based extraction is used and the proteins are precipitated using the TCA/acetone–phenol/methanol method [36, 66] with some modifications [12, 20]:

1. Add 300 µL of buffer extraction to conidia. Mix well firstly using a micropesle, and secondly by vortexing.
2. Sonicate 3×10 s (50 W, amplitude 60), breaking on ice at 1 min. Mix well firstly using a micropesle, and secondly by vortexing.

3. Centrifuge at $16,000 \times g$ for 5 min (4 °C).
4. Fill the tube with 10 % (w/v) TCA/acetone. Mix well by vortexing.
5. Centrifuge at $16,000 \times g$ for 5 min (4 °C), and discard the supernatant.
6. Fill the tube with 0.1 M ammonium acetate in 80 % (v/v) methanol. Mix well using firstly a micropesle, and secondly by vortexing.
7. Follow the steps in Subheading 3.2.1 (starting from step 5).

3.3 Protein Samples for LC

The fungal protein extracts obtained by the TCA/acetone–phenol/methanol method are re-solubilized in a solubilization solution that contains thiourea, CHAPS, and Triton X-100. These compounds are important for protein solubilization, but they interfere with many downstream analysis methods (as is the case of LC separation). For this reason, it is crucial to remove them. Several methods to solve this problem are described below:

1. Protein solubilization with a solution of 6 M urea and 50 mM $(\text{NH}_4)_2\text{CO}_3$ instead of the solubilization solution (see Note 8).
2. Use of cleaning kits (e.g., 2-D Clean-up kit; GE Healthcare), and protein re-solubilization in detergent-free solutions (i.e., 6 M urea, 50 mM $(\text{NH}_4)_2\text{CO}_3$) (see Note 8).
3. Use of electrophoresis using 5 % acrylamide gels. Proteins (about 50–100 µg) are loaded onto the gel, run at 100 V, stained using the Coomassie method, and finally the protein band is cut out (see Note 9).

3.4 Protein Separation by 1-DE

Proteins can be separated by SDS-PAGE with the Laemmli electrophoresis system [68], for example using the Criterion System (Bio-Rad) with precast Criterion Stain-Free precast Gels, Tris–HCl, and 4–20 % linear gradient (Bio-Rad). The 1-DE is visualized using the Image Lab System (Bio-Rad), and stained by Coomassie Blue Brilliant (CBB) method [69] (see Note 10). After protein staining, bands can be analyzed using the Quantity One software (Bio-Rad).

3.5 Protein Separation by 2-DE

Focusing conditions will vary with sample composition, complexity, and strip pH range. In our conditions, the 11 cm IPG strips, pH 5–8 (Bio-Rad), are rehydrated with 50 µg of protein extract in rehydration solution according to the manufacturer's instructions, and applying 50 V for 16 h (active rehydration) at 20 °C. Before this focusing, a wet wick is placed under each end of the strip (cathode). The conditions for IEF have been adapted to our system from [38]: 150 V for 1 h, 1 h at 200 V, 1 h at 500 V, 1,000 V·h at 1,000 V, followed by 2.5-h gradient from 1,000 to 8,000 V, and finally focused for 20,000 V·h at 8,000 V, with a cell temperature of 20 °C (see Note 11). After IEF, IPG strips are stored at –20 °C.

Before the second dimension, IPG strips are equilibrated following two steps. Firstly, it is carried out with 2 % (w/v) DTT in equilibration buffer for 10 min in agitation at room temperature; secondly, it is done with 2.5 % (w/v) iodoacetamide in equilibration buffer for 10 min in agitation at room temperature.

The second dimension is performed in the same way explained in Subheading 3.3, but using Criterion Stain-Free precast Gels, Tris–HCl, and 8–16 % linear gradient IPG strips (Bio-Rad). After protein staining, spots can be analyzed using the PD-Quest software (Bio-Rad).

3.6 Protein Separation by LC

Protein sample (obtained according to Subheading 3.2) is digested at 37 °C overnight using trypsin. Tryptic peptides are cleaned and concentrated in a C18 Cartridges Octadecyl C18/18 % (Applied Separations), dried in a Speed-vac® Concentrator, and re-suspended in 50 µL of a 5 % ACN and 0.1 % formic acid solution. Finally, peptides are injected to the LC system (i.e., a Finnigan Surveyor an HPLC system). All these processes were made according to the Proteomic Service protocols of SCAI, University of Córdoba.

3.7 Protein Identification

3.7.1 Gel-Based Techniques

The bands or spots are cut out and digested with trypsin. Tryptic peptides are analyzed in a mass-spectrometer (i.e., a 4800 Proteomics Analyzer MALDI–TOF/TOF, Applied Biosystems). In this case, the eight most abundant peptide ions are chosen for an MS/MS analysis. A PMF search and a combined search (+MS/MS) are performed using the nrNCBI database of proteins with the MASCOT algorithm (www.matrixscience.com). All these processes were made according to Proteomic Service protocols of SCAI, University of Córdoba (see Note 12).

3.7.2 Gel-Free Techniques

Peptides are detected for example in an LTQ-Orbitrap equipped with a nanoelectrospray ion source (nESI). The acquired data can be analyzed with Proteome Discoverer v1.3 software (Thermo Fisher Scientific, USA) and MASCOT (<http://www.matrixscience.com>) or SEQUEST (<http://fields.scripps.edu/sequest/>) algorithms, using the public *Botrytis cinerea* Database from *Botrytis cinerea* Sequencing Project of Broad Institute of Harvard and MIT (http://www.broadinstitute.org/annotation/genome/botrytis_cinerea/MultiHome.html), according to Proteomic Service protocols of SCAI, University of Córdoba (see Note 13).

4 Final Remarks

With the use of gel-based techniques, around 50 bands and 500 spots were resolved subjected to 1- and 2-DE, respectively, from *B. cinerea* B05.10 mycelia protein extracts [45]. With the use of label-free LC-based approach, around 240 protein species were

identified. A total of 224 of these identified proteins were also quantified, taking into account only those with at least 3 detected peptides [59].

Summarizing, we have shown in this chapter several protocols to work with different fungal material based on our experience with the phytopathogenic fungus *B. cinerea*. Moreover, the use of complementary proteomic approaches, as gel-based (1-DE and 2-DE) and gel-free (label-free LC-based) techniques, provides higher proteome coverage, relevant information being obtained on fungal biology and their interaction with their hosts [3, 4, 12, 45, 59].

5 Notes

1. Gloves should be used for these procedures, and special care should be taken when handling TCA and phenol (consult safety data sheets) since they are corrosive products. Steps involving phenol and β -mercaptoethanol should be performed in a fume hood.
2. Examples of rich-media are PDAB (potato, dextrose, agar+bean leaves), solid synthetic complete medium (CM) [70], or solid malt extract medium (1.5 % w/v).
3. Be careful when working with liquid nitrogen: due to its cool temperature (-195.8 °C) it may cause severe frostbite. The nitrogen evaporates reducing the concentration of oxygen in the air and may act as an asphyxiating agent, especially in confined spaces. Remember that it may be dangerous because nitrogen is odorless, colorless, and tasteless, and it could cause suffocation without any sensation or warning.
4. Be careful: do not throw out the pellet.
5. Three phases appear, namely, the upper phase (which is the phenolic phase where proteins are), a white interphase, and a lower aqueous phase. Try not to get parts from the white interphase.
6. The volume of solubilization solution added will depend on the quantity of precipitated proteins. It is advisable that samples be well concentrated.
7. In this case, the precipitated pellet may be faint because the proteins secreted to the medium are at a very low concentration.
8. The problem derived from the use of solutions without thiourea, CHAPS, and Triton X-100 is the loss of hydrophobic proteins.
9. Acrylamide gel (5 %) (for 10 mL monomer solution): Mix 5.7 mL of ddH₂O, 1.7 mL of 30 % (v/v) acrylamide-bisacrylamide BioRad mixture, 2.5 mL of 0.5 M Tris-HCl pH 6.8 buffer,

and 100 µL of 10 % (w/v) SDS. Immediately prior to pouring the gel into the hand cast for Mini-PROTEAN two-gel electrophoresis system (BioRad), add 50 µL of 10 % (w/v) APS and 5 µL of TEMED and mix gently. The use of electrophoresis to remove it would lead to a reduced loss of these proteins.

10. More details about 1-DE and 2-DE separation methods are described in two excellent reviews [23, 48].
11. The condition of protein focusing must be optimized for each study method. In our case, we use the PROTEAN IEF cell by Bio-Rad. The conditioning phase involves the application of previous steps at low voltage that allow to remove ions and other contaminants containing the sample, and that interfere on protein focusing. The current should not exceed 50 µA per strips. For more information see the 2-D Electrophoresis for Proteomics Manual by Bio-Rad.
12. More details about the MS analysis are described in [56, 64, 71–73].
13. The *Botrytis cinerea* Database was downloaded from the *Botrytis cinerea* Sequencing Project, Broad Institute of Harvard and MIT (http://www.broadinstitute.org/annotation/genome/botrytis_cinerea/MultiHome.html). More details about gel-free analysis are described in [59, 74, 75].

Acknowledgements

This work was supported by the Spanish Ministry of Science and Innovation (BotBank Project, EUI2008-03686), the Andalusian Regional Government (Junta de Andalucía), and the University of Córdoba (AGR-0164: Agricultural and Plant Biochemistry and Proteomics Research Group).

References

1. Afroz A, Ali GM, Mir A et al (2011) Application of proteomics to investigate stress-induced proteins for improvement in crop protection. *Plant Cell Rep* 30:745–763
2. Quirino BF, Candido ES, Campos PF et al (2010) Proteomic approaches to study plant-pathogen interactions. *Phytochemistry* 71:351–362
3. Gonzalez-Fernandez R, Jorrin-Novo JV (2012) Contribution of proteomics to the study of plant pathogenic fungi. *J Proteome Res* 11:3–16
4. Gonzalez-Fernandez R, Prats P, Jorrin-Novo JV (2010) Proteomics of plant pathogenic fungi. *J Biomed Biotechnol* 2010: Article ID 932527, 36pages.doi:932510.931155/932010/932527
5. Choquer M, Fournier E, Kunz C et al (2007) *Botrytis cinerea* virulence factors: new insights into a necrotrophic and polyphagous pathogen. *FEMS Microbiol Lett* 277:1–10
6. Egan MJ, Talbot NJ (2008) Genomes, free radicals and plant cell invasion: recent developments in plant pathogenic fungi. *Curr Opin Plant Biol* 11:367–372
7. Kim Y, Nandakumar MP, Marten MR (2007) Proteomics of filamentous fungi. *Trends Biotechnol* 25:395–400

8. Vincent D, Tan KC, Cassidy L et al (2012) Proteomic techniques for plant-fungal interactions. *Methods Mol Biol* 835:75–96
9. Marra R, Ambrosino P, Carbone V et al (2006) Study of the three-way interaction between *Trichoderma atroviride*, plant and fungal pathogens by using a proteomic approach. *Curr Genet* 50:307–321
10. Rampitsch C, Bykova NV, McCallum B et al (2006) Analysis of the wheat and *Puccinia triticina* (leaf rust) proteomes during a susceptible host-pathogen interaction. *Proteomics* 6:1897–1907
11. Kamoun S (2009) The secretome of plant-associated fungi and oomycetes. In: Deising HB (ed.) *The mycota* vol. V—plant relationships, 2nd edn. Springer, Berlin, pp. 173–180
12. González-Fernández R, Jorrín-Novo JV (2010) Proteomics of fungal plant pathogens: the case of *Botrytis cinerea*. In: Méndez-Vilas AM (ed) *Current research technology and education topics in applied microbiology and microbial biotechnology*. Formatex Research Center, Badajoz, Spain, pp 205–217
13. Ruiz-Herrera J (2012) Fungal cell wall: structure, synthesis and assembly, 2nd edn. CRC, Boca Raton, FL
14. Ebstrup T, Saalbach G, Egsgaard H (2005) A proteomics study of in vitro cyst germination and appressoria formation in *Phytophthora infestans*. *Proteomics* 5:2839–2848
15. Grinyer J, Kautto L, Traini M et al (2007) Proteome mapping of the *Trichoderma reesei* 20S proteasome. *Curr Genet* 51:79–88
16. Melin P, Schnurer J, Wagner EG (2002) Proteome analysis of *Aspergillus nidulans* reveals proteins associated with the response to the antibiotic concanamycin A, produced by *Streptomyces* species. *Mol Genet Genomics* 267:695–702
17. Bohmer M, Colby T, Bohmer C et al (2007) Proteomic analysis of dimorphic transition in the phytopathogenic fungus *Ustilago maydis*. *Proteomics* 7:675–685
18. Grinyer J, Hunt S, McKay M et al (2005) Proteomic response of the biological control fungus *Trichoderma atroviride* to growth on the cell walls of *Rhizoctonia solani*. *Curr Genet* 47:381–388
19. Grinyer J, McKay M, Nevalainen H et al (2004) Fungal proteomics: initial mapping of biological control strain *Trichoderma harzianum*. *Curr Genet* 45:163–169
20. Sulc M, Peslova K, Zabka M et al (2009) Biomarkers of *Aspergillus* spores: strain typing and protein identification. *Int J Mass Spec* 280:162–168
21. Nandakumar MP, Marten MR (2002) Comparison of lysis methods and preparation protocols for one- and two-dimensional electrophoresis of *Aspergillus oryzae* intracellular proteins. *Electrophoresis* 23:2216–2222
22. Shimizu M, Wariishi H (2005) Development of a sample preparation method for fungal proteomics. *FEMS Microbiol Lett* 247:17–22
23. Gorg A, Weiss W, Dunn MJ (2004) Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 4:3665–3685
24. Nandakumar MP, Shen J, Raman B et al (2003) Solubilization of trichloroacetic acid (TCA) precipitated microbial proteins via NaOH for two-dimensional electrophoresis. *J Proteome Res* 2:89–93
25. Rabilloud T (1998) Use of thiourea to increase the solubility of membrane proteins in two-dimensional electrophoresis. *Electrophoresis* 19:758–760
26. Everberg H, Gustavsson N, Tjerned F (2008) Enrichment of membrane proteins by partitioning in detergent/polymer aqueous two-phase systems. *Methods Mol Biol* 424:403–412
27. Kniemeyer O, Lessing F, Scheibner O et al (2006) Optimisation of a 2-D gel electrophoresis protocol for the human-pathogenic fungus *Aspergillus fumigatus*. *Curr Genet* 49:178–189
28. Luche S, Santoni V, Rabilloud T (2003) Evaluation of nonionic and zwitterionic detergents as membrane protein solubilizers in two-dimensional electrophoresis. *Proteomics* 3:249–253
29. Rabilloud T (1996) Solubilization of proteins for electrophoretic analyses. *Electrophoresis* 17:813–829
30. Rabilloud T, Adessi C, Giraudel A et al (1997) Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 18:307–316
31. Herbert BR, Grinyer J, McCarthy JT et al (2006) Improved 2-DE of microorganisms after acidic extraction. *Electrophoresis* 27:1630–1640
32. Fernandez-Acero FJ, Jorge I, Calvo E et al (2007) Proteomic analysis of phytopathogenic fungus *Botrytis cinerea* as a potential tool for identifying pathogenicity factors, therapeutic targets and for basic research. *Arch Microbiol* 187:207–215
33. Fernandez-Acero FJ, Jorge I, Calvo E et al (2006) Two-dimensional electrophoresis protein profile of the phytopathogenic fungus *Botrytis cinerea*. *Proteomics* 6(Suppl 1):S88–S96

34. Fernandez-Acero FJ, Colby T, Harzen A et al (2009) Proteomic analysis of the phytopathogenic fungus *Botrytis cinerea* during cellulose degradation. *Proteomics* 9:2892–2902
35. Guais O, Borderies G, Pichereaux C et al (2008) Proteomics analysis of "Rovabiot Excel", a secreted protein cocktail from the filamentous fungus *Penicillium funiculosum* grown under industrial process fermentation. *J Ind Microbiol Biotechnol* 35:1659–1668
36. Maldonado AM, Echevarria-Zomeño S, Jean-Baptiste S et al (2008) Evaluation of three different protocols of protein extraction for *Arabidopsis thaliana* leaf proteome analysis by two-dimensional electrophoresis. *J Proteomics* 71:461–472
37. Kao SH, Wong HK, Chiang CY et al (2008) Evaluating the compatibility of three colorimetric protein assays for two-dimensional electrophoresis experiments. *Proteomics* 8:2178–2184
38. Fragner D, Zomorodi M, Kues U et al (2009) Optimized protocol for the 2-DE of extracellular proteins from higher basidiomycetes inhabiting lignocellulose. *Electrophoresis* 30:2431–2441
39. Fernandez-Acero FJ, Colby T, Harzen A et al (2010) 2-DE proteomic approach to the *Botrytis cinerea* secretome induced with different carbon sources and plant-based elicitors. *Proteomics* 10:2270–2280
40. Espino JJ, Gutierrez-Sanchez G, Brito N et al (2010) The *Botrytis cinerea* early secretome. *Proteomics* 10:3020–3034
41. Vincent D, Balesdent MH, Gibon J et al (2009) Hunting down fungal secretomes using liquid-phase IEF prior to high resolution 2-DE. *Electrophoresis* 30:4118–4136
42. Medina, M.L. and Francisco, W.A. (2008) Isolation and enrichment of secreted proteins from filamentous fungi. In: Posch A (ed) 2D PAGE: sample preparation and fractionation. Springer, pp. 275–285
43. Ravalason H, Jan G, Molle D et al (2008) Secretome analysis of *Phanerochaete chrysosporium* strain CIRM-BRFM41 grown on softwood. *Appl Microbiol Biotechnol* 80:719–733
44. Abbas A, Koc H, Liu F et al (2005) Fungal degradation of wood: initial proteomic analysis of extracellular proteins of *Phanerochaete chrysosporium* grown on oak substrate. *Curr Genet* 47:49–56
45. González-Fernández R, Aloria K, Valero-Galván J et al (2013) Proteomic analysis of mycelium and secretome of different *Botrytis cinerea* wild-type strains. *J Prot*, <http://dx.doi.org/10.1016/j.jprot.2013.06.022>
46. González-Fernández R, Redondo I, Jorrín-Novo JV (2010) Gel-based proteomic analysis of *Botrytis cinerea*. The simplest 1-DE reveals differences in virulence-related protein abundance among strains. *Proteómica* 5:128–129
47. Supek F, Peharec P, Krsnik-Rasol M et al (2008) Enhanced analytical power of SDS-PAGE using machine learning algorithms. *Proteomics* 8:28–31
48. Rabillo T, Vaezzadeh AR, Potier N et al (2009) Power and limitations of electrophoretic separations in proteomics strategies. *Mass Spectrom Rev* 28:816–843
49. Haynes PA, Roberts TH (2007) Subcellular shotgun proteomics in plants: looking beyond the usual suspects. *Proteomics* 7:2963–2975
50. Pirondini A, Vissioli G, Malcevschi A et al (2006) A 2-D liquid-phase chromatography for proteomic analysis in plant tissues. *J Chromatogr B Analys Technol Biomed Life Sci* 833:91–100
51. Recorbet G, Rogniaux H, Gianinazzi-Pearson V et al (2009) Fungal proteins in the extra-radical phase of arbuscular mycorrhiza: a shotgun proteomic picture. *New Phytol* 181:248–260
52. Washburn MP, Wolters D, Yates JR 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19:242–247
53. Ye M, Jiang X, Feng S et al (2007) Advances in chromatographic techniques and methods in shotgun proteome analysis. *Trends Anal Chem* 26:80–84
54. Cooper B, Neelam A, Campbell KB et al (2007) Protein accumulation in the germinating *Uromyces appendiculatus* uredospore. *Mol Plant Microbe Interact* 20:857–866
55. Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. *Science* 312:212–217
56. Nesvizhskii AI, Vitek O, Aebersold R (2007) Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat Methods* 4:787–797
57. Roe MR, Griffin TJ (2006) Gel-free mass spectrometry-based high throughput proteomics: tools for studying biological response of proteins and proteomes. *Proteomics* 6:4678–4687
58. González-Fernández R, Aloria K, Redondo I et al (2011) Gel-free/label-free analysis to study the secretome of *Botrytis cinerea*. Tenth annual congress—HUPO, Geneva, Switzerland

59. Gonzalez-Fernandez R, Aloria K, Arizmendi JM et al (2013) Application of label-free shotgun nUPLC-MSE and 2-DE approaches in the study of *Botrytis cinerea* mycelium. *J Proteome Res* 12:3042–3056
60. González-Fernández R, Redondo I, Gomez-Galvez FJ et al (2011) Proteomic approaches to study the phytopathogenic fungus *Botrytis cinerea*. *Botryotis-Sclerotinia* post-genome workshop, Lyon, France
61. de Oliveira JM, de Graaff LH (2011) Proteomics of industrial fungi: trends and insights for biotechnology. *Appl Microbiol Biotechnol* 89:225–237
62. Porteus B, Kocharunchitt C, Nilsson RE et al (2011) Utility of gel-free, label-free shotgun proteomics approaches to investigate microorganisms. *Appl Microbiol Biotechnol* 90:407–416
63. Harder A (2008) Sample preparation procedure for cellular fungi. *Methods Mol Biol* 425:265–273
64. Gusakov A, Semenova M, Sinitsyn A (2010) Mass spectrometry in the study of extracellular enzymes produced by filamentous fungi. *J Anal Chem* 65:1446–1461
65. Valledor L, Jorri J (2011) Back to the basics: maximizing the information obtained by quantitative two dimensional gel electrophoresis analyses by an appropriate experimental design and statistical analyses. *J Proteomics* 74:1–18
66. Wang W, Vignani R, Scali M et al (2006) A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. *Electrophoresis* 27:2782–2786
67. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254
68. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685
69. Neuhoff V, Arold N, Taube D et al (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9:255–262
70. Pontecorvo G, Roper JA, Forbes E (1953) Genetic recombination without sexual reproduction in *Aspergillus niger*. *J Gen Microbiol* 8:198–210
71. Han X, Aslanian A, Yates JR 3rd (2008) Mass spectrometry for proteomics. *Curr Opin Chem Biol* 12:483–490
72. Neilson KA, Ali NA, Muralidharan S et al (2011) Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* 11:535–553
73. Becker CH, Bern M (2010) Quantitation in mass-spectrometry-based proteomics. *Annu Rev Plant Biol* 61:491–516
74. Mueller LN, Brusniak MY, Mani DR et al (2008) An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *J Proteome Res* 7:51–61
75. Schulze WX, Usadel B (2010) Quantitation in mass-spectrometry-based proteomics. *Annu Rev Plant Biol* 61:491–516

Chapter 9

Plant Proteomics Methods to Reach Low-Abundance Proteins

Egisto Boschetti and Pier Giorgio Righetti

Abstract

The question of low-abundance proteins from biological tissues is still a major issue. Technologies have been devised to improve the situation and in the last few years a method based on solid-phase combinatorial peptide ligand libraries has been extensively applied to animal extracts. This method has also been extended to plant extracts taking advantage of findings from previous experience. Detailed methods are described and their pertinence highlighted according to various situations of plant sample origin, size of the sample, and analytical methods intended to be used for protein identifications.

Key words Low-abundance proteins, Combinatorial peptide ligand libraries, Protein depletion, Protein fractionation

Abbreviations

| | |
|------|--------------------------------------|
| CPLL | Combinatorial Peptide Ligand Library |
| TUC | Thiourea–Urea-CHAPS |

1 Introduction

Plant proteomics is experiencing a fast growth in present days and is now well accepted by all dedicated proteomics journals. However, there are still some major shortcomings that hamper proper discoveries in the field [1–3]. One of them is the paucity of data on the various genomes. It is thus clear that for further progress we need the help of geneticists willing to tackle the sequencing of more and more plant genomes (most of the papers published so far deal with the proteomes of *Arabidopsis thaliana* and rice, *Oryza sativa*, and focus on profiling organs, tissues, cells, or subcellular proteomes). The other major drawback is that in some tissues (like leaves) a few proteins dominate the landscape and hamper proper discovery of low-abundance species. This is further aggravated by the presence

of various plant constituents (polyphenols, polysaccharides) that strongly interfere with various sample manipulations, such as protein capture via various chromatographic means and analyses via different electrophoretic methodologies. The present chapter deals with such aspects and in particular offers guidelines in one methodology that is emerging as perhaps the most powerful in detection of low- to very-low-abundance proteins, namely, the combinatorial peptide ligand library (CPLL) technique.

1.1 Proteins: A Minor Component of Plant Extracts

The protein content of plant tissues is much smaller compared to animal tissues [4]. Depending on the organelle, the protein content is about 20 times less since thick cell walls largely contribute to the total biomass. In addition large vacuoles present in the cells also contribute to decrease the average protein mass per biomass unit. Overall when investigating the plant proteomes the amount of tissue to start with must be significantly larger than from animals. The consequence of this situation is that a lot of useless material is to be eliminated during protein extraction with the risk of losing associated proteins. Plant proteins are largely glycosylated and combined to polysaccharides rendering the extraction challenging. For example it is known that some proteins are covalently associated to cell walls [5] such as arabinogalactan proteins. Many of these proteins such as hydroxyproline-rich proteins are also largely glycosylated with a behavior unrelated to animal glycoproteins and thus current technologies cannot directly be applied. For instance their extraction may require calcium-containing buffers or other salts [6]. In some instances, to succeed in protein extraction, a deglycosylation is necessary to release proteins that are strongly involved in cell wall constructs.

In spite of low protein content, the expression of proteins in plant results in a large dynamic concentration range precluding the access to low-abundance species. Well-known massively present proteins in plants, depending on the organ, are RuBisCo [7] in leaves and endosperm storage proteins [8]. To resolve this question several approaches have been proposed: (1) extensive fractionation, (2) removal of dominant species, and (3) compression of concentration difference between low-abundance and high-abundance proteins.

1.2 Necessary Sample Pretreatment to Eliminate Interfering Material

Since interfering substances are present, special treatments are necessary to polish the initial plant crude extract [9]. Extraction and precipitation are the two most popular approaches. TCA associated with acetone with reducing agents allows precipitating most proteins leaving in the supernatant a lot of undesired materials [10]. Pellets containing proteins are then separated by centrifugation and redissolved. Alternatively or in addition to TCA/acetone precipitation other interfering substances are eliminated by using a Tris-HCl solution saturated with phenol followed by a precipitation with ammonium acetate or ammonium sulfate. This option is

especially recommended when the analysis of proteins is based on two-dimensional (2D) electrophoresis, but it depends on the plant organ from where the protein extraction is intended. Also recommended for two-dimensional electrophoresis is the one described by Faurobert et al. [11] involving a combination of phenol extraction followed by a precipitation by ammonium acetate in methanol. Another less popular method is the protein precipitation using a chloroform–methanol mixture with water in 1–4–3 proportions [12]. Several variants to these methods are also described to precipitate proteins [13]; however, the protein solubilization protocol to be followed is not always easy and frequently necessitates the presence of zwitterionic surfactants and chaotropes.

All these preliminary operations do not only remove undesired materials but also contribute to concentrate proteins that are present in very low amounts but in a quite large number.

Indeed while the number of native proteins corresponds to the number of genes, the numerous variants due to extensive post-translational modifications render the reality much different. Nonetheless it is not the number of proteins that constitutes the most important issue in their detection and identification, but rather the level of expression of genes, building up proteomes where the individual component concentration difference spans over several orders of magnitude. Among largely expressed proteins there are species that are present only in few copies. The latter are consequently very difficult to detect. In this context enrichment procedures have been devised to improve the situation. Like in animal proteome investigations, precipitation, fractionation, depletion, and enrichment are the major approaches.

The published papers on the treatments of proteins from plant extract samples to try reaching low-abundance species are not as numerous as described for animal proteins. However, the same concept has here been applied for the same purpose. Special methods of precipitations of *Glycine max* leaf extracts to remove RuBisCO have been used for the detection of rare proteins such as many phosphorylated proteins [14]. Ion-exchange chromatography fractionation has also been used as for instance to enrich protein involved in plant defense from *Arabidopsis thaliana* evidencing more than 30 misregulated proteins [15].

More elegantly immunoaffinity chromatography has been used for the selective removal of RuBisCO [16] in leaf extracts. Besides antibodies, other affinity-based solid-phase techniques have been used for the removal of nucleotide-binding proteins from leaf extracts using PolyA [17] and PolyU [18] as immobilized ligands.

Another affinity-based selective separation method for the separation of phosphoproteome is based on the replacement of the phosphate groups on serine and threonine residues by a biotin tag. Biotinylated species are then separated using avidin-affinity chromatography [19].

Most of these techniques have as essential objective to get rid of RuBisCO except in the case of Widjaja et al. [15] where the authors tried to combine depletion and enrichment. Unfortunately the approach proposed is a combination of various fractionation techniques that are all together labor intensive and add the high risk of losing very low-abundance proteins due to the repeated manipulations. The only method that concomitantly reduces the dynamic range of protein concentration is the one involving CPLL as detailed below. It proved its properties with a number of plant extracts such as maize seeds [20], spinach [21], *Arabidopsis* t. leaves [22], rubber plant latex [23], and wines [24, 25].

1.3 Protocol to Enhance Low-Abundance Proteins in Plant Proteomics

Low-abundance proteins are generally undetectable because of two distinct phenomena: (1) their signal is covered by the high-abundance proteins in either 2D electrophoresis and mass spectrometry and (2) because their concentration is below the sensitivity level of the analytical methods. In the past few years a method has been devised to compress the dynamic concentration range in order to decrease the concentration of most concentrated species reducing thus the signal coverage and also the concentration of rare species. This process is operated by the so-called CPLL. This approach has been extensively described for a number of biological extracts including from plants [26–30].

CPLL is a mixture of small beads (ca. 65 μ m diameters) to which hexapeptides are covalently linked. The number of peptides reaches various millions depending on the number of amino acids used for the synthesis; however, each bead carries a single type of peptide in a large number of copies. This is thus a mixed bed of beads different from each other and individually capable to capture a protein or a group of them. When a plant protein extract is exposed to such a CPLL under large overloading conditions, each bead with affinity to an abundant protein will rapidly become saturated and the vast majority of the same protein will remain unbound. In contrast, trace proteins will not saturate the corresponding partner beads unless the sample volume is large enough to provide for increasing amounts of proteins. Once the excess of unbound proteins is eliminated by filtration or centrifugation, all captured proteins can be collected by elution at a concentration range that is largely lower than in the original biological sample. Trace proteins thus become detectable by current analytical methods. In theory each bead carrying a single peptide ligand should interact with proteins that share the same epitope complementary to the peptide bait. However, because in a number of cases peptide ligands are very similar and due to the mass action law (individual protein concentration, affinity constants, and conditions of exposure) several proteins are found on a single bead (or peptide ligand) and a same protein can be captured by several different beads.

The idea of libraries of millions of peptides for the capture of protein from a crude extract with the aim of reducing the dynamic concentration range is relatively recent [31], but a large number of applications have been described with a variety of protein extracts and biological fluids. Several recent reviews [32, 33] illustrate them and demonstrate the excellent results that can be obtained with this technique and the unique increment in the detection ability.

According to the protein desorption method (single or sequenced) or according to the type of eluents used, certain types of analytical methods can be directly applied. For instance when the desorption is produced by SDS Laemmli buffer at boiling for 5 min, SDS-PAGE is directly applied [34]. Alternatively when thiourea-urea-CHAPS (TUC) solution is used 2D electrophoresis can be directly used with no preliminary treatment [35]. In another instance the captured proteins, still on the beads, can be directly digested by trypsin and the peptides directly used for multidimensional protein fractionation and identification via mass spectrometry [36].

Plant protein extracts are not always dominated by a single or a few largely dominant proteins like human serum or red blood cell extracts; however, there are cases where one protein largely dominates the protein content. This is the case of RuBisCO (ribulose-1,5-biphosphate carboxylase/oxygenase) that accounts for more than 40 % of the total protein amount in leafs [37].

The poor protein content of plant extracts frequently associated to undetectable very-low-abundance proteins and the dominance of few very concentrated proteins justify adapted protocols for protein sample extracts to facilitate the analysis of plant proteomes.

2 Materials

Phosphate-buffered saline, ammonium persulfate, protease inhibitor cocktails, sodium chloride, urea, thiourea, CHAPS, acetic acid, citric acid: All these chemicals are from regular suppliers such as Sigma-Aldrich, MO, USA. Plastic labware is from Pierce Biotechnology, CO, USA. Combinatorial hexapeptide ligand library (ProteoMiner) is from Bio-Rad Laboratories, CA, USA (*see Note 1*). Vortex and benchtop centrifuge are from Thermo Fisher Scientific.

3 Methods

3.1 Tissue Extraction and Pretreatments

Since soluble protein concentration is generally very low in differentiated plant tissues [4] and a number of undesired material is present, before entering the process of proteome studies several points have to be considered. Plant cells are rich in proteases, which require the presence of inactivating agents, and rich in polysaccharides (lots of them polyanionic) interacting directly with CPLL and

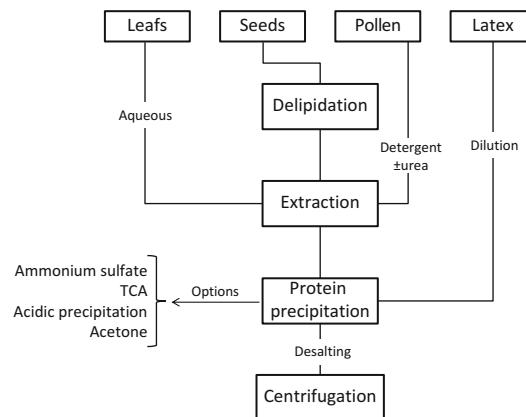


Fig. 1 Scheme of plant sample treatment as a function of extract origin prior to protein recovery for further treatments (fractionation/depletion or enrichment)

thus preventing proper protein capture, as well as in pigments, lipids, polyphenols, and secondary metabolites that may interfere with protein separation and analyses [38].

A large number of plant proteins are heavily glycosylated, thus adding consequent problems of abnormal migrations in 2D electrophoresis and difficult ionization in mass spectrometry. Various cleanup protocols specific to plant extracts are available; however, most protocols are not compatible with CPLLs. Figure 1 illustrates possible options for four typical plant extracts used in conjunction with CPLL. The following general rules should be adopted: (a) the aqueous extraction should be performed in relatively low ionic strength to prevent the solubilization of nucleic acids; (b) with highly viscous material, such as latex and honey, a dilution is recommended; (c) when dealing with proteins that are engaged within the cell wall, such as pollen proteins, some amounts of non-ionic detergent (less than 0.5–1 %) and urea (less than 3 M) should be used at a concentration compatible with CPLL; (d) a treatment with phenol, associated with some amounts of polyvinylpyrrolidone, eliminates in general pigments and polyphenols.

Few examples are given in the literature with detailed technical information [29]. A preliminary lipid removal step is particularly recommended with plant seeds such as soya beans, peanuts, corn, sunflower, and many others. Once undesired material is eliminated plant proteins can be precipitated with ammonium sulfate or trichloroacetic acid or acetone, or polyethylene glycol to collect precipitates that are free or almost free of CPLL interfering substances. Acidic precipitation can also be operated just by acidifying the solution with acetic acid at pH 3–4. It is here to be noticed that the use of trichloroacetic acid as precipitating agent may induce denaturation and therefore the loss of biological properties.

All the above sample pretreatments contribute to obtain significantly better analytical results especially when using 2D electrophoresis and related methods. One of the preferred methods for protein recovery is the precipitation with ammonium sulfate at high concentration (80–90 % saturation). Naturally this salt is to be removed, which is accomplished by simple dialysis at a very low cutoff (e.g., 3,500 Da) or by centrifugation using appropriate filtration-integrated devices.

When dealing with vegetable-derived beverages such as wines no specific pretreatment is really necessary.

3.2 Protein Capture and Dynamic Range Reduction

Once the plant protein extract is obtained (*see Note 2*) the treatment with CPLL is relatively standard. Nevertheless few variations allow capturing more or less proteins. The general protocol is as follows:

1. Solid-phase hexapeptide ligand libraries whatever the presentation, dry or wet, need to be equilibrated with the buffer solution selected for protein capture (e.g., PBS) before use. To this end 100 μ L of CPLL are mixed with 500–1,000 μ L of selected solution and gently agitated for few hours or overnight. Then the beads are washed extensively with the same fresh buffer under mild centrifugation at for example 2,000–3,000 $\times g$. Finally all excess of solution is removed by centrifugation at 12,000 $\times g$ or more.
2. The plant extract in the selected buffer used for protein capture and containing the antiprotease cocktail is centrifuged at 12,000 $\times g$ to remove all solid material in suspension. For an optimal capture the protein concentration of the sample should be at least of 0.1–1 mg/mL. Lower concentrations may render the capture of very-low-abundance proteins difficult when the dissociation constant is too high. The total amount of protein from the sample should be larger than 50 mg when 100 μ L of hexapeptide ligand library is used. The larger the protein load, the greater the probability to enrich and thus detect very-low-abundance proteins. To concentrate a too dilute protein solution current methods adopted in biochemistry techniques can be used. Preferred ways are lyophilization after dialysis or membrane concentration under centrifugation (Centricon, centrifugal filters from Millipore).
3. The protein solution is mixed with the drained CPLL beads and the suspension gently agitated for at least 2 h at room temperature (*see Note 3*). It is critical to maintain an even slurry of the hexapeptide ligand library beads so as to ensure good contact between the beads and the proteins. The agitation should not be too strong so as to avoid formation of foam. The suspension can be left overnight under agitation without negative consequences. These very important protein capture steps can be performed in physiological conditions of ionic strength and pH; however, they can also be made in lower ionic strength to

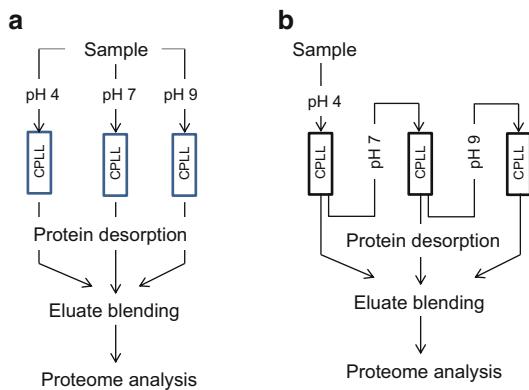


Fig. 2 Schematic representation of protein capture at different pHs. Parallel (a) or sequential (b) protein capture followed by individual desorption and eluate blending prior to proteomics analysis. The second approach is adapted for small amount of proteins, but requires an adjustment of pH after each CPLL contact. For easier understanding the CPLL contact is here represented as chromatographic columns, but the capture phase can be operated in suspension

increase the amount of captured proteins. The pH is also critical to modulate the protein capture since it acts on the affinity constant of proteins for their corresponding peptide baits. To enlarge the captured protein coverage it is advised to operate at three different pHs (e.g., 4, 7, and 9) [39]. This treatment can be performed under different configurations as illustrated in Fig. 2.

4. The excess of proteins is removed from the hexapeptide ligand library by centrifugation at about $2,000 \times g$ followed by at least three washings with 500 μ L of the same buffer used for protein capture. In between centrifugations, the bead slurry is shaken gently for 10 min. All along the washing process CPLLs should not be left completely dry.

3.3 Elution Protocols

While a standard protein elution protocol is generally applied, customized elution protocols are possible as described (global elution, fractionated elution, or direct on-bead protein digestion).

1. Global elution (several options) [35].
 - (a) Beads are added with 200 μ L of urea-CHAPS-acetic acid (8 M-4–0.5 %) and gently shaken for about 1 h at room temperature. Acetic acid could be replaced by citric acid at similar concentration. In both cases the pH should be around 3–3.5.
 - (b) Alternatively CPLLs are added with 6 M guanidine-HCl, pH 6.0. This strong dissociating solution at relatively high ionic strength also allows desorption of most of the proteins.

- (c) Another approach is to add to a TUC desorption solution some amounts of cysteic acid [40].
- (d) When followed by SDS-PAGE analysis, the elution could be performed by boiling the beads with 200 μ L of Laemmli buffer comprising sodium dodecyl sulfate and reducing agents such as 2-mercaptoethanol or dithiothreitol.

3.4 Fractionated Elution

Desorption of proteins is obtained by successive washings with solutions of increased stringency [23].

1. As first eluting agent use 200 μ L of neutral 1 M sodium chloride solution to dissociate all proteins captured by a dominant ion-exchange effect.
2. Then use as second eluting agent 200 μ L of 60 % ethylene glycol in water to dissociate proteins predominantly captured by hydrophobic association.
3. Then incubate the CPLL with a third desorption solution composed of 0.4 M glycine-HCl buffer (pH 2.5).
4. Finally all residual proteins still present on the beads are stripped out by a mixture composed of 6 % acetonitrile, 12 % isopropanol, 10 % of 17 M ammonia, and 72 % distilled water (vol/vol).

In between desorption steps the CPLL beads could be rapidly washed with 200 μ L of distilled water with a small risk of desorbing some proteins. If this washing operation is not performed then some overlap of proteins persists between fractions.

5. Individual fractions are then desalted, cleaned, or used directly for further analysis (see **Notes 4** and **5**).

3.5 Direct On-Bead Protein Digestion

When the analysis of captured proteins is performed by the so-called shotgun approach, the most direct way to proceed is to make a digestion of the captured proteins directly on the beads [36]. The operation requires some excess of trypsin since part of it will be captured by the CPLL beads. Basically the process is as follows:

1. After protein capture on the peptide library beads (whatever the method or the physicochemical conditions), the beads are rapidly washed twice with 200 μ L of 100 mM ammonium bicarbonate containing 0.1 % Rapigest (this is not mandatory, but it facilitates the proteolysis process). This is obtained by adding 1 mL of 100 mM ammonium bicarbonate to the 1 mg Rapigest vial lyophilizate and shake gently for few minutes. The bead suspension is then vortexed for few minutes.
2. Then 300 μ L of 10 mM DTT are added and the bead suspension heated at 65 °C for 1 h under gentle stirring or occasional shaking.
3. 300 other μ L of 55 mM iodoacetamide are added, mixed and stored in the dark for 60 min.

4. Then 60 μ L of 0.2 μ g/ μ L trypsin sequencing grade are added. If the succinylated peptide library beads are used the amount of trypsin should be four times larger.
5. The bead suspension is vortexed and incubated overnight at 37 °C under gentle shaking.
6. 200 μ L of 500 mM formic acid are then added; the suspension is vortexed, for few seconds, and incubated for about 40 min at room temperature.
7. Then the supernatant is recovered by filtration (30,000 MWCO) under centrifugation (microfuge, at least 5,000 $\times g$ for 20 min) in order to separate insoluble material and beads.
8. In order to fully extract the remaining peptides the beads are washed under centrifugation once with 50 μ L of 500 mM formic acid and mixed to previous filtrate.
9. Stripped beads could then be kept at -20 °C for possible further analysis.
10. The solution of peptides is then dried by speedvac and redissolved in 20 μ L HPLC solvent for LC-mass spectrometry analysis.

As a possible alternative to this protocol,

1. 100 μ L are added with 600 μ L of a solution composed of 8 M urea, 100 mM Tris (hydroxyethylamine) pH 8.5, and 5 mM tris (2-carboxyethyl)phosphine. After few seconds of vortex-stirring the suspension is agitated gently for 30 min at room temperature.
2. Cysteine residues are then acetylated by addition of 10 mM iodoacetamide and incubated in the dark for at least 15 min.
3. The bead suspension is then diluted with 100 mM Tris-(hydroxyethylamine) pH 8.5 to reach a final concentration of 2 M urea.
4. Trypsin is added (20 μ g as 0.5 μ g/ μ L, this corresponding roughly to 1:50 protease/protein ratio) along with CaCl₂ to 1 mM concentration and the resulting mixture incubated overnight at 37 °C.
5. The obtained peptides are separated from the beads by centrifugation at 1,000 $\times g$ and stored at -80 °C until the day of analysis.
6. Before analysis the peptide solution is acidified to 5 % formic acid and centrifuged at 18,000 $\times g$ to remove possible solid material.

3.6 Interrelation Between Elution and Analysis

Proteomic analysis may not be compatible with the protein solution after classical treatment with CPLL. In this case two possibilities are available: (1) the collected samples are put in appropriate buffers or (2) elutions are customized to comply with the following analysis. In the first case the protein samples are desalted by extensive dialysis

and then lyophilized. They could also be desalted by gel filtration or through dedicated membranes under centrifugation.

There are also various possibilities to elute proteins captured by CPLL beads and thus render the solution directly compatible with the analysis. This is the case with protein desorption using SDS-Laemmli buffer under boiling conditions. Here the sample can be used directly for SDS-PAGE analysis and then after band slicing and trypsinization the peptides are analyzed by LC-MS/MS [34].

If the proteomic analysis is performed by 2D electrophoresis, the most compatible elution is to use TUC solution. However this does not elute 100 % of proteins from the beads and should be completed by another orthogonal elution. Alternatively the TUC solution could comprise some amounts of cysteic acid that produces an almost exhaustive desorption of proteins. In this case due to the very low pI value of cysteic acid, which collects at the anode, 2D electrophoresis can also be performed [40]. This elution is also compatible with the application of isoelectric focusing separation of components.

When 2D-DIGE is used as 2D electrophoresis analysis, the only elution that is compatible with this technique is the use of 20 mM Tris buffer containing 7 M urea, 2 M thiourea, and 4 % CHAPS, pH 8.5 (sodium carbonate could also be used instead of Tris).

For direct ELISA-based assays of eluted proteins, the denaturing desorption agents that can be selected are 0.2 M glycine-HCl, 2 % NP-40, pH 2.4; 0.1 M acetic acid, 2 % NP-40; 1 M NaCl, 2 % NP-40; or 0.1 M acetic acid containing 40 % ethylene glycol. In case a single eluent does not desorb all proteins from the beads these solutions could be used as a sequence and the eluates blended.

It is here recalled that protein elution from beads may not be necessary. This is the case when proteins are just analyzed after trypsin digestion. This treatment can be operated directly on the beads and the obtained peptides directly analyzed by LC-MS/MS [36]. This approach is recommended especially when dealing with small samples involving small volumes of beads with time saving and largely reduced protein losses.

4 Notes

1. It is recommended not to reuse hexapeptide beads because (a) some level of carryover may appear, with consequent misinterpretation of data, and (b) some peptides may have been modified as a consequence of stringent elution conditions from previous operations.
2. The sample should be clear and not contain lipids in suspension such as in milk or in bile fluid. Large amounts of nucleic acids or viscous polysaccharides, when present, should also be removed using current methods. Samples should not contain a large amount of detergents or denaturing agents. For example, nonionic deter-

gents are tolerated at concentrations not exceeding 0.5 % (wt/vol); urea is also tolerated at a concentration not exceeding 3 M.

The method can be applied to a large variety of plant protein extracts after appropriate elimination of interfering biopolymers. Nevertheless specific aspects of optimization might have to be considered according to the encountered issues. If for instance the protein concentration is below 0.1 mg/mL it may be useful to have a preliminary concentration. This would improve the capture by CPLL in case the affinity is too modest. Among possible concentration methods are dialysis followed by lyophilization or membrane concentration under centrifugation.

The presence of proteases, relatively frequent in plant extracts, is deleterious for the integrity of proteins. Their activity must be stopped with various selected inhibitors or inactivation agents prior to contact with CPLL.

3. It may happen that during the capture stage there is formation of bead aggregates. In this case the supernatant must be separated by high-speed centrifugation and the collected solid material is to be washed extensively with PBS under strong shaking (e.g., vortex) up to the dissociation of beads from each other. Chemical agents are not recommended since they may desorb the captured proteins.
4. The presence of large amounts of high-abundance proteins after CPLL treatment and protein recovery may mean that the amount of proteins in the initial sample was not sufficient to saturate the beads. This could be resolved by either increasing the amount of sample or decreasing the volume of CPLL beads. It is reminded that the enrichment of low-abundance proteins renders the sample to analyze more complex (many more proteins are detectable). The subsequent analytical operations might become of difficult interpretation. To facilitate the analysis it is advised to fractionate the collected proteins or to elute the capture proteins sequentially.
5. Generally CPLL treatments are highly reproducible; however, if it is observed that the results are not exactly the same from an experiment to another, it is advised to check the ionic strength and the pH of the initial sample. Actually even modest modifications of these parameters alter the affinity of proteins for the peptide baits of the CPLL with consequent modification of the molecular interaction process.

5 Practical Examples of CPLL-Treated Plant-Derived Extracts

A number of applications have been described in the recent past with interesting data. To illustrate the power of the described technology, few examples are reported with particular emphasis on the sample treatment and the mode of use of CPLL.

The first example deals with proteins from leaf extracts from *Spinacia oleracea* [41]. Protein extraction (after removal of undesired pigments and polyphenols) was performed in physiological saline solution containing 5 mM EDTA, 1 mM DTT, and a tablet of complete protease inhibitor in order to prevent protease action. The homogenate was clarified by centrifugation at $10,000 \times g$ for 10 min and directly submitted to protein capture with CPLLs, followed by elution in 4 % boiling SDS and 20 mM DTT. The proteins found in the treated sample were compared to untreated extract: 322 unique gene products were found, 114 of which common between native and CPLL-treated extract. While only 18 unique proteins were found in the control, 114 new others have been found thanks to CPLL intervention. They presumably represent low-abundance species since they were undetectable in the crude native extract. A number of proteins related to chloroplast category have been found after CPLL treatment such as inorganic pyrophosphatase 1, phosphoglycerate kinase, glutamate decarboxylase 2, glucose-1-phosphate adenylyltransferase, fructose-bisphosphate aldolase 2, ribulose-phosphate 3-epimerase, and 5-methyltetrahydropteroylglutamate-homocysteine methyltransferase. A second category of enrichment of cellular components was ribosomal proteins such as 30S, 40S, 50S, and 60S.

Figure 3 represents the comparative results between native and treated extracts.

The proteome of *Hevea brasiliensis* latex has been explored after treatment with CPLL [23]. In this case the fresh plant exudate was first centrifuged at high speed to remove latex grains and then proteins were precipitated in the cold with ammonium sulfate at 90 % saturation. The supernatant was discarded, and protein pellets resuspended in distilled water containing protease inhibitor cocktail tablets and then dialyzed (MWCO 3,500 membrane) overnight at 4 °C against PBS. Two successive treatments with CPLL were performed (ProteoMiner and carboxylated ProteoMiner). Protein elution was operated sequentially using 1 M NaCl; 2 M thiourea, 7 M urea, 3 % CHAPS; 9 M urea in 50 mM citric acid, pH 3.3; and hydro-organic mixture composed of acetonitrile, isopropanol, trifluoroacetic acid, and water at final concentrations of, respectively, 16.6, 33.3, 0.5, and 49.5 %. A total of 300 unique gene products have been identified from a proteome that was largely unknown, most of them thanks to the CPLL treatment (Fig. 4). In addition, when searching for allergens by confrontation with the sera of 18 patients, several species have been identified (2D electrophoresis and immunoblots). Among them were major known allergens, and others unknown such as heat-shock protein, proteasome subunit (30 kDa), 8 kDa protease inhibitor, hevamine A, and glyceraldehyde-3-phosphate dehydrogenase.

Proteome investigations have also been operated on cypress pollen in view of finding novel allergens [42]. Pollen is a quite peculiar part of plants with proteins that are part of the tissue wall and

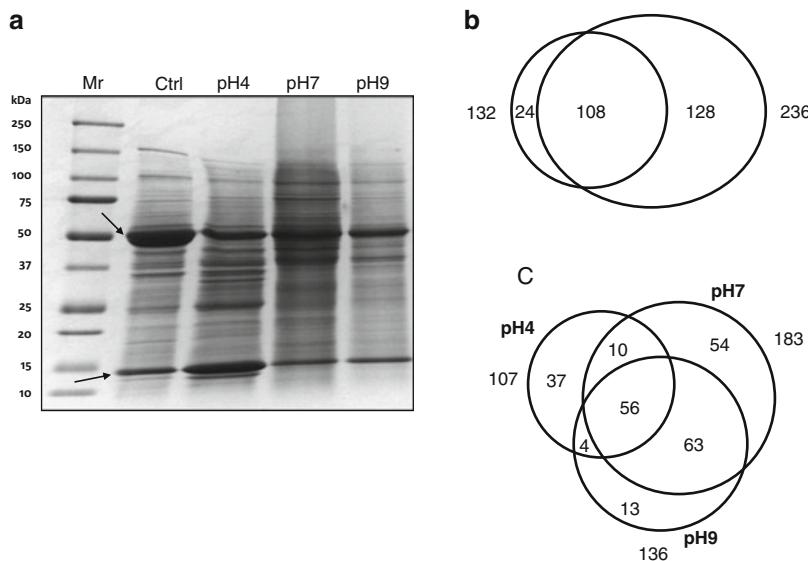


Fig. 3 Spinach proteome investigation from spinach leafs. After removal of undesired vegetable material, the proteins were treated with CPLL at different pHs and analyzed by SDS-PAGE (a). Ctrl means control or untreated protein isolate; the arrows indicate the positioning of RuBisCO light and heavy subunits. Mr: protein mass ladder. (b) Venn diagram of found proteins using LC-MS/MS as analytical method. The circle on the left represents proteins identified from untreated protein isolate, while the circle on the right represents the identified proteins upon complete CPLL treatment. (c) Venn diagram of identified proteins upon CPLL treatment at different pH where the contribution of each pH is shown. Many proteins are common among two or three eluates; others are exclusive of each capture pH (adapted from [41])

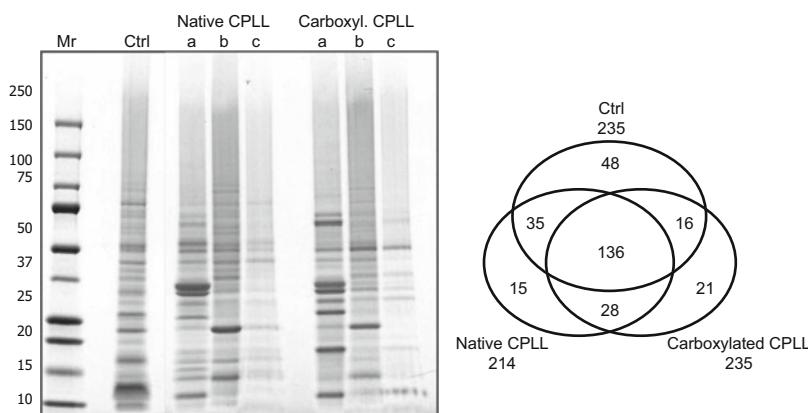


Fig. 4 *Hevea brasiliensis* latex proteome study. After removal of undesired material, proteins were treated with CPLL under physiological conditions; the captured proteins were desorbed sequentially and analyzed by SDS-PAGE (left). Ctrl means control or untreated protein isolate; Mr: protein mass ladder. The protein isolate was then treated sequentially on two distinct CPLLs (native and carboxylated). The captured proteins were in both cases eluted sequentially using 1 M NaCl (a), 2 M thiourea, 7 M urea, 3 % CHAPS (b), 9 M urea in 50 mM citric acid, pH 3.3 (c). Right: Venn diagram of found proteins using LC-MS/MS as analytical method (adapted from [23])

difficult to solubilize. This adds to the large difficulties of eliminating any undesired types of materials such as pigments and polyphenols. Here the authors have first tried to extract the maximum amount and number of proteins involved in insoluble matrices and then to eliminate all that would complicate protein fractionation. The fact that proteins are only a small part of the pollen matrix necessitated to start from very large amounts of dry pollen. The first extraction operation was performed in the presence of antiproteases using a solution of 10 mM phosphate containing 0.2 % of a nonionic detergent—Nonidet NP40—and 3 M urea, followed by a second extraction with physiological buffered saline. After precipitation with 80 % saturation ammonium sulfate, the two collected protein solutions were independently treated with CPLL at three different pHs to maximize the enrichment in low-abundance species. The analysis of treated proteins compared to non-treated extracts was performed by either SDS-PAGE followed by LC-MS/MS or by 2D electrophoresis followed by immunoblots against the serum of patients allergic to cypress pollen, followed by spot excision and identification by mass spectrometry. In a nut, what was found is an extremely large number of peptides of which only few could be assigned to identified proteins; the majority of peptides could not be assigned due to the lack of genome knowledge. As expected however, the number of identifiable proteins was significantly increased by CPLL treatment compared to the control and only few were in common. The number of isoforms was also largely increased after CPLL when comparing 2D map results. Moreover the number of positive spots to IgE from various patients allowed identifying many allergens unknown up to the present investigation and also indicating that allergic profiles are extremely different from an individual to another.

Products derived from plants such as some drinks (e.g., wines, bier, colas, and others) have also been analyzed after treatment with CPLL. Wines were directly loaded on beads for the capture of possible traces of proteins [24, 25]. The volume of these beverages compared to the volume of beads was in all cases huge since around 1,000 mL of each drink was loaded on 0.1 mL of beads. Useless to say that considering the minute amount of protein possibly captured the elution was performed at once using SDS desorption as described. In most investigated wines traces of casein were found (on the average from 20 to 60 µg/L, with a detectability of as low as 1 µg/L), casein being one of refining agents used after fermentation. However, for other younger untreated wine (e.g., white Recioto, from Garganega grapes in the Veneto, Italy) CPLLs captured close to 100 unique gene products derived from the grapes [43]. As far as beer is concerned a large number of proteins were found [41], several from barley, two from maize, and more than four dozens from yeast such as *S. cerevisiae*, *S. pastorianus*, and *S. bayanus*.

6 The Debate Continues

Proteome investigations in plants are a step behind animal proteomes and since very sophisticated technologies developed for animals are readily available one should see in the near future a rapid progress in this domain. The identification of proteins in plants still suffers from the unavailability of genome sequences even if it is possible to work by analogy with other plants. One of the most investigated plants is *Arabidopsis thaliana* for which the genome sequence revealed the presence of about 30,000 genes. In spite of this very large number only a modest number of unique proteins are detected [44] for reasons explained earlier in this chapter. At this point in time it is largely possible to detect many more proteins after treatment with CPLL, as this can easily be evidenced by 2D electrophoresis, but the formal identification is the major technical difficulty. With more plant genomes sequenced this situation is likely to be improved in the next coming years. The presence of very few proteins present in massive amount among others that are very numerous and of low abundance makes plant protein extracts particularly suitable for CPLL treatments. Detecting differences in protein expression may not have a big interest in plant diagnostic, but it can be an elegant way to detect frauds or even genetically modified vegetables through the differential expression of low-abundance proteins.

Difficulties of plant protein extraction may need to devise specific procedure and components that are or are not compatible with CPLL. If so modified, CPLL could be used accordingly. On another level, specialized CPLL could also be designed in order to meet plant protein requirements. These specialized beads could also be designed in order to be specific only for categories of proteins or even modified to avoid nonspecific binding for undesired materials. If such developments will be made they would largely simplify the treatment of crude extracts, which would not need any preliminary treatment prior to fractionation.

On this note, we add here a final paragraph describing recent progress in the field. We felt that extra efforts were sorely needed to study recalcitrant tissues and in general any plant proteome. This extra effort has materialized in more than one extraction protocol, i.e., under native and denaturing conditions, in both cases followed by capture with CPLLs of the extracted proteins. This novel strategy has been applied recently in the case of two recalcitrant fruit tissues, namely, avocado and banana [45, 46]. In both cases about 1 % total protein is embedded either in solid oil (avocado) or in huge amounts of polysaccharides (banana). In order to improve discovery of low-abundance species, in parallel with the standard, native condition extraction, a denaturing solubilization protocol has been implemented, based on 3 % boiling SDS, an anathema in CPLL treatments, since it would completely inhibit the capture. Yet, there are

two ways to go about it: one is to remove the SDS via the classical acetone/methanol precipitation, and the other is to dilute the SDS from 3 to 0.1 % in the presence of another compatible surfactant, like 0.5 % CHAPS, conditions that would be compatible with the CPLL technology. There is a substantial amount of extra work involved, yet the results have been outstanding: in the case of avocado, the total number of unique gene products identified amounted to 1,012 proteins, of which 174 in common with the control, untreated sample, 190 present only in the control, and 648 representing the new species detected via CPLLs of all combined (native and denatured) eluates and likely representing low-abundance proteins. In the case of banana, out of a total number of 1,131 identified proteins, the various captures (under native and denaturing conditions) with the CPLL beads permitted the identification of 849 proteins, whereas the controls allowed identifying 452 proteins, 170 species being in common. Thus it is seen that with this double extraction protocol, one can easily exceed the 1,000-species limit in a single sweep, almost an anathema in plant proteomics. By comparison, in analysis of olive fruit pulp [47], where only native extraction was applied, the booty was rather meagre: only 252 unique gene products could be identified. The readers are referred to these two recent papers [45, 46] for all practical details. For detailed basic information and protocols sample treatment using various approaches including recipes for plant extracts, the reader should refer to a recently published book as referenced [48].

References

1. Jorrín-Novo JV, Maldonado AM, Echevarría-Zomeño S et al (2009) Plant proteomics update (2007-2008): Second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfil MIAPE standards, increase plant proteome coverage and expand biological knowledge. *J Proteomics* 72:285–314
2. Agrawal GK, Rakwal R (2008) Plant proteomics: technologies, strategies, applications. Wiley, Hoboken
3. Agrawal GK, Job D, Zivy M et al (2011) Time to articulate a vision for the future of plant proteomics—a global perspective: an initiative for establishing the international plant proteomics. *Proteomics* 11:1559–1568
4. Rose JK, Bashir S, Giovannoni JJ et al (2008) Tackling the plant proteome: practical approaches, hurdles and experimental tools. *Plant J* 39:715–733
5. Jamet E, Albenne C, Boudart G et al (2008) Recent advances in plant cell wall proteomics. *Proteomics* 8:893–908
6. Millar DJ, Whitelegge JP, Bindschedler LV et al (2009) The cell wall and secretory proteome of a tobacco cell line synthesising a secondary wall. *Proteomics* 9:2355–2372
7. Demirevska-Kepova K, Simova-Stoilova L, Kjurkchiev S et al (1999) Barley leaf rubisco, rubisco binding protein and rubisco activase and their protein/protein interactions. *Bulg J Plant Physiol* 25:31–44
8. Li G, Nallamilli BR, Tan F et al (2008) Removal of high-abundance proteins for nuclear subproteome studies in rice (*Oryza sativa*) endosperm. *Electrophoresis* 29:604–617
9. Carpentier SC, Panis B, Vertommen A et al (2008) Proteome analysis for non-model plants: a challenging but powerful approach. *Mass Spectrom Rev* 27:354–377
10. Méchin V, Damerval C, Zivy M (2007) Total protein extraction with TCA-acetone. *Methods Mol Biol* 355:1–8
11. Faurobert M, Pelpoir E, Chaïb J (2007) Phenol extraction of proteins for proteomic studies of recalcitrant plant tissues. *Methods Mol Biol* 355:9–14
12. Wessel D, Flugge UI (1984) A method for the quantitative recovery of proteins in dilute

- solutions in the presence of detergents and lipids. *Anal Biochem* 138:141–143
13. Isaacson T, Damasceno CM, Saravanan RS et al (2006) Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nat Protoc* 1:769–774
 14. Krishnan HB, Natarajan SS (2009) A rapid method for depletion of Rubisco from soybean (*Glycine max*) leaf for proteomic analysis of lower abundance proteins. *Phytochemistry* 70:1958–1964
 15. Widjaja I, Naumann K, Roth U et al (2009) Combining subproteome enrichment and Rubisco depletion enables identification of low abundance proteins differentially regulated during plant defense. *Proteomics* 9:138–147
 16. Cellar NA, Kuppannan K, Langhorst ML et al (2008) Cross species applicability of abundant protein depletion columns for ribulose-1,5-bisphosphate carboxylase/oxygenase. *J Chromatogr A* 861:29–39
 17. Ni RJ, Shen Z, Yang CP et al (2010) Identification of low abundance polyA-binding proteins in *Arabidopsis* chloroplast using polyA-affinity column. *Mol Biol Rep* 37:637–641
 18. Xu Y, Wang BC, Zhu YX (2007) Identification of proteins expressed at extremely low level in *Arabidopsis* leaves. *Biochem Biophys Res Commun* 358:808–812
 19. Kwon SJ, Choi EY, Seo JB et al (2007) Isolation of the *Arabidopsis* phosphoproteome using a biotin-tagging approach. *Mol Cells* 24:268–275
 20. Fasoli E, Pastorello EA, Farioli L et al (2009) Searching for allergens in maize kernels via proteomic tools. *J Proteomics* 72:501–510
 21. Fasoli E, Aldini A, Regazzoni L et al (2010) Les maîtres de l'orge: the proteome content of your beer mug. *J Proteome Res* 9:5262–5269
 22. Fröhlich A, Lindermayr C (2011) Deep insights into the plant proteome by pretreatment with combinatorial hexapeptide library. *J Proteomics* 74:1182–1189
 23. D'Amato A, Bachi A, Fasoli E et al (2010) In-depth exploration of *Hevea brasiliensis* latex proteome and “hidden allergens” via combinatorial peptide ligand libraries. *J Proteomics* 73:1368–1380
 24. Cereda A, Kravchuk AV, D'Amato A et al (2010) Proteomics of wine additives: mining for the invisible via combinatorial peptide ligand libraries. *J Proteomics* 73:1732–1739
 25. Cereda A, Kravchuk AV, D'Amato A et al (2010) Noah's nectar: the proteome content of a glass of red wine. *J Proteomics* 73:2370–2377
 26. Guerrier L, Thulasiraman V, Castagna A et al (2006) Reducing protein concentration range of biological samples using solid-phase ligand libraries. *J Chromatogr B* 833:33–40
 27. Sennels L, Salek M, Lomas L et al (2007) Proteomic analysis of human blood serum using peptide library beads. *J Proteome Res* 6:4055–4062
 28. Roux-Dalvai F, Gonzalez de Peredo A, Simó C et al (2008) Extensive analysis of the cytoplasmic proteome of human erythrocytes using the peptide ligand library technology and advanced mass spectrometry. *Mol Cell Proteomics* 7:2254–2269
 29. Boschetti E, Bindschedler L, Tang C et al (2009) Combinatorial peptide ligand libraries and plant proteomics: a winning strategy at a price. *J Chromatogr A* 1216:1215–1222
 30. Mouton-Barbosa E, Roux-Dalvai F, Bouyssié D et al (2010) In-depth exploration of cerebrospinal fluid by combining peptide ligand library treatment and label-free protein quantification. *Mol Cell Proteomics* 9:1006–1021
 31. Thulasiraman V, Lin S, Gheorghiu L et al (2005) Reduction of concentration difference of proteins from biological liquids using combinatorial ligands. *Electrophoresis* 26:3561–3571
 32. Righetti PG, Boschetti E, Kravchuk A et al (2010) The proteome buccaneers: how to unearth your treasure chest via combinatorial peptide ligand libraries. *Exp Rev Proteomics* 73:73–385
 33. Righetti PG, Boschetti E (2011) Combinatorial peptide ligand libraries: the conquest of the “hidden proteome” advances at great strides. *Electrophoresis* 32:960–966
 34. Candiano G, Dimuccio V, Bruschi M et al (2009) Combinatorial peptide ligand libraries for urine proteome analysis: investigation of different elution systems. *Electrophoresis* 30:2405–2411
 35. Righetti PG, Boschetti E, Zanella A et al (2010) Plucking, pillaging and plundering proteomes with combinatorial peptide ligand libraries. *J Chromatogr A* 1217:893–900
 36. Fonslow BR, Carvalho PC, Academia K et al (2011) Improvements in proteomic metrics of low abundance proteins through proteome equalization using ProteoMiner prior to MudPIT. *J Proteome Res* 10:3690–3700
 37. McCabe MS, Garratt LC, Schepers F et al (2001) Effects of P(SAG12)-IPT gene expression on development and senescence in transgenic lettuce. *Plant Physiol* 127:505–616
 38. Gengenheimer P (1992) Preparation of extracts from plants. *Methods Enzymol* 182:174–193

39. Fasoli E, Farinazzo A, Sun CJ et al (2010) Interaction among proteins and peptide libraries in proteome analysis: pH involvement for a larger capture of species. *J Proteomics* 73:733–742
40. Farinazzo A, Fasoli E, Kravchuk AV et al (2009) En bloc elution of proteomes from combinatorial peptide ligand libraries. *J Proteomics* 72:725–730
41. Fasoli E, D'Amato A, Kravchuk AV et al (2011) Popeye strikes again: the deep proteome of spinach leaves. *J Proteomics* 74:127–136
42. Shahali Y, Boschetti E, D'Amato A et al (2012) Allergomic study of cypress pollen via combinatorial peptide ligand libraries. *J Proteomics* 77:101–110
43. D'Amato A, Fasoli E, Kravchuk AV et al (2011) *Mehercules, adhuc bacchus!* The debate on wine proteomics continues. *J Proteome Res* 10:3789–3801
44. Peck SC (2005) Update on proteomics in *Arabidopsis*. Where do we go from here? *Plant Physiol* 138:591–599
45. Esteve C, D'Amato A, Marina ML et al (2012) Identification of avocado (*Persea americana*) pulp proteins by nanoLC-MS/MS via combinatorial peptide ligand libraries. *Electrophoresis* 33:2799–2805
46. Esteve C, D'Amato A, Marina ML et al (2013) In-depth proteomic analysis of banana (*Musa* spp.) fruit with combinatorial peptide ligand libraries. *Electrophoresis* 34(2):207–214
47. Esteve C, D'Amato A, Marina ML et al (2012) Identification of olive (*Olea europaea*) seed and pulp proteins by nLC-MS/MS via combinatorial peptide ligand libraries. *J Proteomics* 75:2396–2403
48. Boschetti E, Righetti PG (2013) Low-abundance protein discovery: State of the art and protocols. Elsevier, pp 1–375

Chapter 10

Combination of 2DE and LC for Plant Proteomics Analysis

Sami Irar, Faïçal Brini, Khaled Masmoudi, and Montserrat Pagès

Abstract

The use of analytical biochemical techniques with different separation properties allows us to better understand the proteome. To demonstrate this we have used two different methodologies to analyze embryos from a Tunisian cultivar of durum wheat (*Triticum durum* Desf.), variety Oum Rabiaa. We compared conventional 2D electrophoresis with liquid-phase chromatography. Our results show that a similar number of proteins were detected with both techniques. However, analysis of protein resolution at different pH ranges showed significant differences. By using a large pH gradient we observed that liquid chromatography presents higher resolution at extreme pH, either acidic or basic. Conversely, 2DE is more resolutive at intermediate pH (pH 5–6.5). Taking these results in consideration, we propose that 2DE and liquid chromatography are complementary methods to analyze complex protein extracts and can be used in parallel to acquire a wider perspective and a better understanding of the embryo proteome.

Key words Comparative plant proteomics, Liquid chromatography, ProteomeLab PF-2D, 2D electrophoresis, Wheat embryo proteome

1 Introduction

Plant comparative proteomics is becoming increasingly attractive as the rapidly expanding plant genomic and expressed sequence tag (EST) databases provide new opportunities for protein identification [1]. The evaluation of the proteome, the protein status of a cell type, tissue, organ, and whole organism, is an alternative strategy to address complex biological questions like the link existing between genotype and phenotype [2].

To make the comparison between 2DE and LC techniques we have worked with a sample of wheat embryo, specifically with a salt- and drought-tolerant variety. We believe that by using complementary methodologies in the analysis of the embryo proteome, new proteins required for dehydration tolerance could be discovered. Usually, plant proteins have been analyzed by two-dimensional polyacrylamide gel electrophoresis (2-D PAGE)

followed by mass spectrometry (MS) for their identification [3]. The heterogeneity of polypeptide molecular size, charge, hydrophobicity, complexity, and cellular distribution makes it almost impossible to capture and solubilize the entire complement of proteins in a given sample [4, 5].

More recently, proteome analyses have been performed in a “gel-less” condition by using protein fractionation procedures based entirely on liquid chromatography (LC). Here we use a system of two-dimensional liquid chromatography based on a high-performance chromatofocusing in the first dimension, followed by high-resolution reversed-phase chromatography in the second dimension [6–8].

In the present work we have used the conventional 2-D gel electrophoresis and liquid chromatography. We have isolated wheat embryos from the Tunisian cultivar and analyzed their total protein extracts in a different pH range with two different separation techniques [1].

2 Materials

2.1 Plant Material

Triticum durum Desf., salt- and drought-tolerant Tunisian cultivar (Oum Rabiaa) of wheat.

2.2 Equipment

The materials 1–11 were all purchased from GE Healthcare Life Sciences.

1. Isoelectrofocusing (IEF), IPGphor II Isoelectric Focusing Unit.
2. IPGphor Strip Holders.
3. Second-dimension Ettan DALT^{six} systems.
4. SDS gel-casting box.
5. Cassette rack.
6. Glass plates.
7. Thermostatic circulator (Multitemp III).
8. Power supply (Multidrive XL).
9. Laboratory shaker.
10. ImageScanner.
11. ImageMaster 2D Platinum 5.0 software.
12. LC, ProteomeLab PF-2D (Beckman Coulter).
13. LC software, 32 Karat, and ProteoVUETM (Beckman Coulter).

2.3 Products and Stock Solutions

1. Lysis buffer 1: 7 M urea, 2 M thiourea, 4 % CHAPS, 50 mM Tris–HCl, pH 8.0, supplemented with 53 u/mL DNase I, 4.9 u/mL RNase, a cocktail of protease inhibitors (1 mM PMSF, 0.1 mM pepstatin, 2 mM leupeptine 1 mM E-64, and

- 1 mM aprotinin), 4 % Triton X-100, and dithiothreitol (DTT) 1 M.
2. Strips: Immobiline dry strips 24 cm, with a nonlinear pH gradient 3–10 (GE Healthcare).
 3. Rehydration buffer: 7 M urea, 2 M thiourea, 4 % CHAPS, 50 mM Tris–HCl, pH 8.0, 0.5 % IPG buffer pH 3–10 (GE Healthcare), and 2 % DeStreak (GE Healthcare).
 4. Equilibration buffer 1: 50 mM Tris–HCl, pH 8.8, 6 M urea, 30 % glycerol, 2 % SDS, 10 mg/mL DTT.
 5. Equilibration buffer 2: 50 mM Tris–HCl, pH 8.8, 6 M urea, 30 % glycerol, 2 % SDS, 25 mg/mL iodoacetamide.
 6. Staining solution: 0.1 % CBB R-250 (Bio-Rad), 10 % acetic acid, 40 % methanol as previously described [9].
 7. Lysis buffer 2: 6 M Urea, 2 M thiourea, 10 % glycerol, 50 mM Tris–HCl, pH 8.0, 2 % (w/v) octyl- β -D-glucopyranoside, supplemented with 53 u/mL DNase I, 4.9 u/mL RNase, and a cocktail of protease inhibitors (1 mM PMSF, 0.1 mM pepstatin, 2 mM leupeptine, 1 mM E-64, and 1 mM aprotinin).
 8. ProteomeLab PF-2D kit (Beckman Coulter): Contains start buffer and elution buffer.
 9. ProteomeLab PF-2D buffers for the second dimension: Gradient with *solvent A* (0.1 % trifluoroacetic acid (TFA)–water) and *solvent B* (0.08 % TFA–acetonitrile) and wash buffer (1 M NaCl).
 10. Protein quantification: BCA protein assay reagent Kit (Pierce).

3 Methods

3.1 Protein Extraction and Quantification for 2DE Analysis

1. Mature wheat dry embryos were manually dissected, collected, frozen in liquid nitrogen, and stored at –80 °C until further analysis [10].
2. Proteins from approximately 150 mg of embryos were extracted with 1.2 mL lysis buffer 1 [11].
3. After 20-min incubation at 4 °C, 14 mM DTT was added and samples were centrifuged for 10 min at 35,000 $\times g$ at 4 °C.
4. The supernatant was centrifuged once more and protein content was quantified using the BCA protein assay reagent Kit (Pierce).

3.2 Protein Extraction and Quantification for LC Analysis

1. For the ProteomeLab PF-2D, proteins from approximately 150 mg of embryos were extracted with 1.2 mL of lysis buffer 2 [12] (see Note 1).
2. Samples were centrifuged for 10 min at 35,000 $\times g$ at 4 °C. The supernatant underwent a second centrifugation and

protein content was quantified as mentioned above, using the BCA protein assay reagent Kit (Pierce).

3.3 Two-Dimension Electrophoresis Protein Separation

1. IEF was carried out using an IPGphor electrophoresis system and 24 cm immobiline dry strips, with a nonlinear pH gradient of 3–10 (GE Healthcare). One milligram of protein sample was used for protein separation in the first dimension. Rehydration was performed for 10 h at 50 V. IEF was then performed by gradient 500 V over 1 h 30 min, 1,000 V gradient for 1 h 30 min, 2,000 V gradient for 1 h 30 min, 4,000 V gradient for 1 h 30 min, 8,000 V gradient for 2 h, and 8,000 V holding for 6 h, until a 65,000 Vh was reached [13].
2. After focusing, the strips were equilibrated for 15 min with equilibration buffer 1 and 15 min with equilibration buffer 2.
3. For the second dimension, the 24 cm strips were loaded onto SDS-PAGE 12 % polyacrylamide gels ($25 \times 20 \times 0.1$ cm) and were run at 2.5 W/gel for 30 min followed by 20 W/gel for 4 h.
4. Gels were stained with CBB R-250 (Bio-Rad) as previously described [3].
5. The stained gels were scanned at 300 dpi resolution with an ImageScanner (GE Healthcare). The spot detection and gel comparison were performed using the ImageMaster 2D Platinum 5.0 software (GE Healthcare) (Fig. 1).
6. The 2DE patterns of wheat embryo proteins that are soluble in the standard 2DE rehydration buffer are shown in Fig. 1. With the conventional 2DE system it was possible to detect 899 embryo proteins in a pH gradient 3–10 using Coomassie staining method (see **Notes 2** and **3**).

3.4 Liquid Chromatography: Protein Separation; Chromatofocusing and High-Performance Reverse-Phase Chromatography

1. Total embryo proteins in lysis buffer 2 (2.5 mg) were desalted with a PD-10 desalting columns (GE Healthcare) (see **Note 4**). Samples were further filtered with a 0.45 μ m filter before injection on the chromatography column. The *High-Performance ChromatoFocusing* (HPCF) column was treated according to the manufacturer's instructions (ProteomeLab PF-2D kit, Beckman Coulter).
2. Briefly, the column was washed with 10 volumes of water at a flow rate of 0.2 ml/min for 45 min and then it was equilibrated with 30 volumes of start buffer for 130 min at 0.2 ml/min. Operating conditions or methods are as follows: Column, HPCF 1D column. Mobile phase, start buffer pH 8.0, elution buffer pH 4.0, high ionic strength wash and water distilled, filtered through a 0.45 μ m membrane filter.

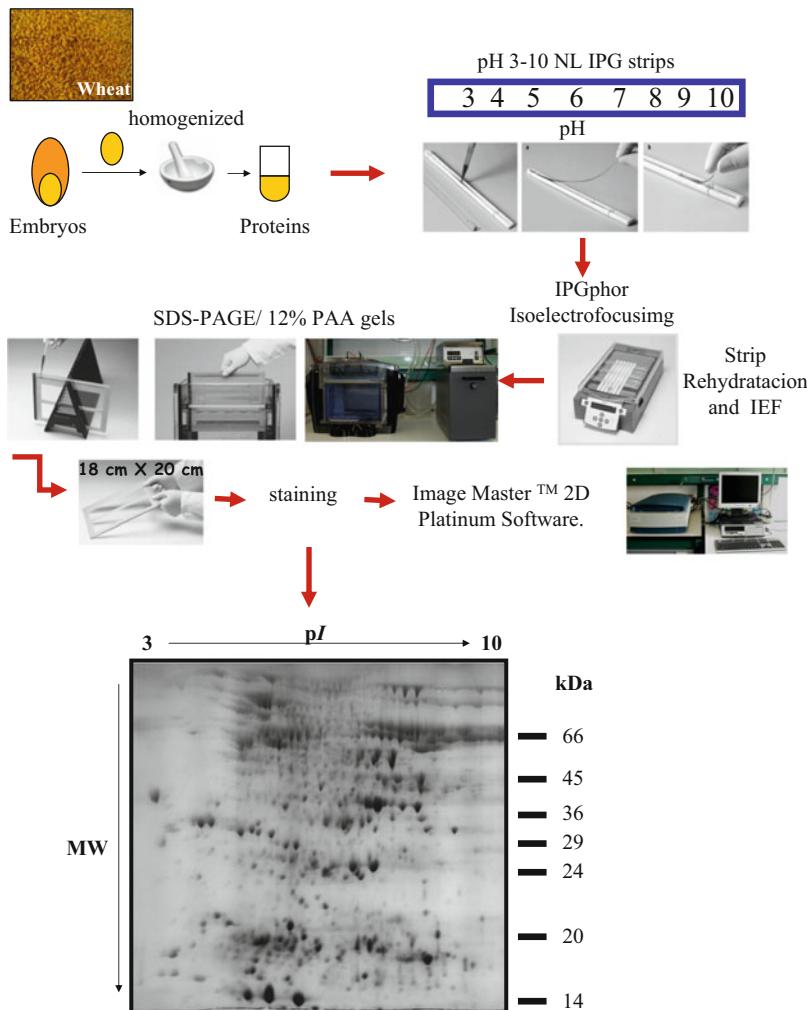


Fig. 1 Summary of extraction and separation in 2DE. First dimension with IPG linear pH 3–10 and second dimension on 12 % SDS-PAGE. Total proteins stained with CBB R250. After gel image analysis using the ImageMaster Platinum software, 768 spots were detected in the pH 4–8 gradient vs. the 899 detected in the pH range 3–10. The relative molecular mass standards ($\times 1,000$ Da) are indicated in the Y-axis and the pI is given at the top. Gels are of 24-cm length

Flow rate, 0.2 ml/min. Temperature, 25–30 °C. Detection, UV 280 nm and back pressure, 600–1,000 psi.

3. After equilibration, the sample was introduced with a manual injector onto the column. In the first dimension the proteins were bound to a strong anion exchanger and eluted with a continuous decreasing pH from 8.0 to 4.0 (elution buffer). The pH began to decrease after 40–45 min.

4. Fractions were collected at a pH interval of 0.3 in a 96 deep-well plate in an FC/I module (fraction collector/injection). At 170 min, the column was washed with 1 M NaCl for 15 min (*see Note 5*).
5. Finally, the first-dimension column was washed with water for 45 min.
6. Proteins were then separated in the second-dimension high-performance reversed-phase (HPRP). When the first-dimension separation was completed (170 min), the pI fractions from the first dimension were then automatically run on the second dimension as indicated in the ProteomeLab PF 2D method: Column, HPRP 2D Column. Mobile phase, solvent A and B. Flow rate, 0.75 ml/min. Temperature, 50 °C. Detection, UV 214 nm and back pressure, 1,500–2,000 psi.
7. Prime lines of the HPRP module with solvent A and solvent B.
8. Injection of 200 µl of sample for each first-dimension fraction.
9. After the sample has been applied to the column from the FC/I module, a gradient of 0–100 % solvent B for 30 min, flow rate 0.75 ml/min, is run.
10. Run 100 % solvent B for 5 min followed by 100 % solvent A for 10 min at a flow rate of 0.75 ml/min before the next sample.
11. The method automatically saves the raw UV absorbance data for each second-dimension analysis of the chromatofocusing fractions for protein mapping and data analysis using 32Karat and ProteoVUE™ in the PF 2D software Suite.
12. In Fig. 2 the chromatofocusing using the ProteomeLab PF 2D system was presented. The fraction of proteins resolved by the ProteomeLab PF 2D gradient (pH 8.0–4.0) corresponded to 0.625 mg. This represented 25 % of the 2.5 mg of total proteins that were initially injected in the column. About 1.1 mg, 44 % of the total proteins were resolved before the pH gradient, thus representing proteins that were not retained by the column. After the pH gradient, the column was washed with a high salinity buffer and different peaks were observed corresponding to 0.775 mg, about 31 % of the total proteins (*see Notes 6 and 7*).

3.5 Combination of 2DE and LC for Plant Proteomics Analysis

In summary we can conclude that both methodologies, 2DE and liquid chromatography, are complementary methods to analyze complex protein extracts and they can be used in parallel to acquire a wider perspective and a better understanding, in this case, of the embryo proteome. Therefore, the choice of the methodology to be applied will depend on the proteins of interest (Fig. 3).

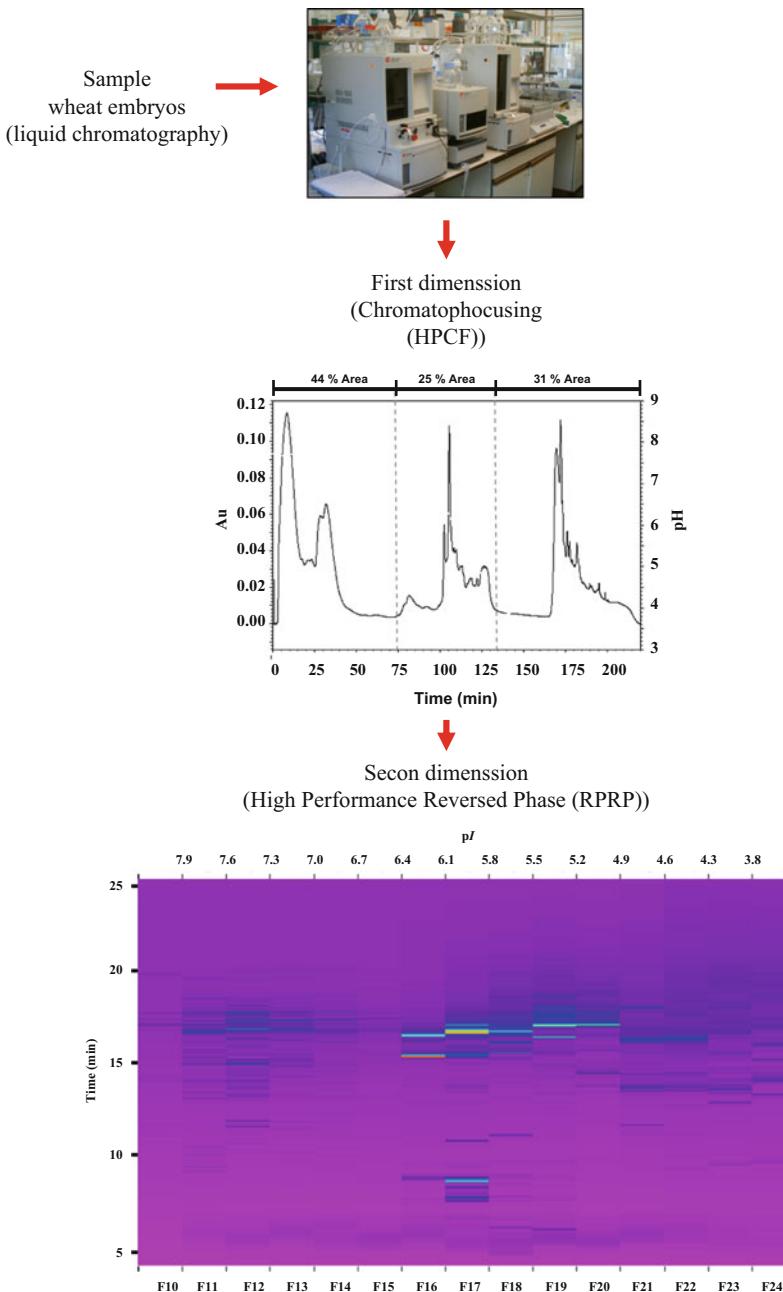


Fig. 2 Liquid chromatography ProteomeLab PF 2D. The first step, chromatofocusing of total wheat embryo proteins in a pH gradient. X-axis is retention time in the column, *left* y-axis pH, and *right* y-axis absorbance and the second step, high-performance reversed phase. The fractions are separated in two dimensions: pH gradient 8.0–4.0 and hydrophobicity. Finally were represented ProteomeLab PF 2D virtual gel. There are a total of 15 fractions in the gradient and were detected 855 bands with a DeltaVue software. In the y-axis the retention time in minutes in the retention column is represented, in the *upper* x-axis the pH gradient, and in the *lower* x-axis the fraction number

| pH Range | BLC | | 2DE | |
|-----------------------|------------|------------|-------|-------|
| | Areas | Spots | Spots | Spots |
| 7.9-3.8 | 855 | 768 | | |
| 3.8-3.0 | 0 | 17 | | |
| 7.9-10.0 | 0 | 114 | | |
| Total Proteins | 855 | 899 | | |

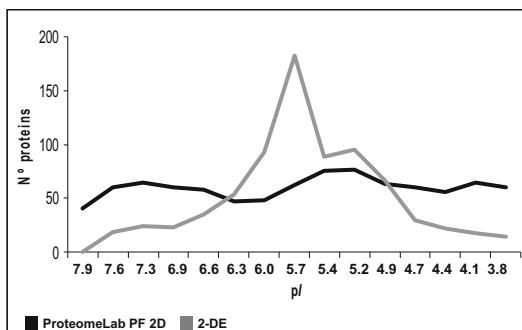


Fig. 3 Summary results. Table and graph of the distribution areas corresponding to proteins in bidimensional liquid chromatography (BLC) and spots corresponding to proteins (2DE) along the pH gradient

4 Notes

1. The protein extraction procedure was identical as for the conventional 2DE, except the removal of CHAPS, since the ProteomeLab PF 2D is not compatible with this detergent.
2. In this study we used two different 2D proteomic methods, with the intent of determining which is better suited to resolve total protein extracts. We compared conventional 2D electrophoresis with ProteomeLab PF 2D liquid chromatography using embryo protein extracts from the Tunisian cultivar of durum wheat (*Triticum durum* Desf.), variety Oum Rabiaa. About 855 bands were detected at a pH range of 8.0–4.0 by ProteomeLab PF 2D. In the same pH range, 2D electrophoresis allowed the detection of 768 Coomassie R-250 stained. The number of proteins stained with Coomassie R-250 is similar to the number of proteins detected with ProteomeLab PF 2D liquid chromatography.
3. For more information of 2DE in plant, please *see* ref. 14.
4. PD-10 desalting columns were equilibrated with 25 ml start buffer. Sample was injected in a desalting column in a volume of 2.5 ml. When the sample has run into the column, elute high-molecular-weight components such as proteins, with 3.5 ml start buffer.

5. Total proteins were separated in the first dimension in different pH fractions of 0.3 range. The total number of fractions obtained was 34 and 15 fractions were selected for the second dimension using the HPRP chromatography. In Fig. 2b the 855 bands obtained in the 15 fractions used for the second dimension are represented. A virtual 2D gel was given by the software, combining the first and the second dimension, thus showing proteins separated by pI and hydrophobicity.
6. We have analyzed the protein resolution at different pH ranges and found that the differences were significant (Fig. 2). Despite the advantage of using the ProteomeLab PF 2D for analyzing protein expression with high loading capacity that enhances the detection of low-abundance proteins with few purification steps, thus achieving a higher level of reproducibility than the chemical 2DE, the resolution of proteins depended on the pH gradient.
7. Liquid chromatography separates proteins uniformly over a pH gradient and 2DE methods follow a normal trend in protein distribution over the same pH gradient. This means that liquid chromatography is better in the resolution of proteins at extreme pH (over 7.0 and below 4.5) while 2DE provides a higher resolution at pH 5–6.5 (Fig. 3).

Acknowledgment

This work was supported by Bio 2009-13044-Co2-01 from MICIN.

References

1. Pirondini A, Vissoli G, Malcevski A, Marmiroli N (2006) A 2-D liquid phase chromatography for proteomic analysis in plant tissues. *J Chromatogr B* 833:91–100
2. Zivy M, de Vienne D (2000) Proteomics: a link between genomics, genetics and physiology. *Plant Mol Biol* 44:575
3. Farinha AP, Irar S, De Oliveira E et al (2011) Novel clues on abiotic stress tolerance emerge from embryo proteome analyses of rice varieties with contrasting stress adaptation. *Proteomics* 11:2389–2405
4. Rose JK, Bashir S, Giovannoni JJ et al (2004) Tackling the plant proteome: practical approaches, hurdles and experimental tools. *Plant J* 39:715
5. Patterson SD (2004) How much of the proteome do we see with discovery-based proteomics methods and how much do we need to see? *Curr Proteomics* 1:3–12
6. Wall DB, Kachman MT, Gong S et al (2000) Isoelectric focusing nonporous RP HPLC: a two-dimensional liquid-phase separation method for mapping of cellular proteins with identification using MALDI-TOF mass spectrometry. *Anal Chem* 72:1099–1111
7. Wang H, Kachman MT, Schwartz DR et al (2002) A protein molecular weight map of ES2 clear cell ovarian carcinoma cells using a two-dimensional liquid separations/mass mapping technique. *Electrophoresis* 23:3168–3181
8. Wall DB, Parus SJ, Lubman DM (2001) Comparison of the capabilities of liquid isoelectric focusing-one-dimensional nonporous silica reversed-phase liquid chromatography-electrospray ionization time-of-flight mass

- spectrometry and liquid isoelectric focusing-one-dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis mass mapping for the analysis of intact protein molecular masses. *J Chromatogr B* 763:139–196
- 9. Yang P, Wnag X, Chen H et al (2007) Proteomic analysis of rice (*Oryza sativa*) seeds during germination. *Proteomics* 7:3358–3368
 - 10. Brini F, Hanin M, Lumbreiras V et al (2007) Functional characterization of DHN-5, a dehydrin showing a differential phosphorylation pattern in two Tunisian durum wheat (*Triticum durum* Desf.) varieties with marked differences in salt and drought tolerance. *Plant Sci* 172:20–28
 - 11. Irar S, Oliveira E, Pagès M, Goday A (2006) Towards the identification of late-embryogenic-abundant phosphoproteome in *Arabidopsis* by 2-DE and MS. *Proteomics* 6:S175–S185
 - 12. Irar S, Brini F, Goday A et al (2010) Proteomic analysis of wheat embryos with 2-DE and liquid-phase chromatography (ProteomeLab PF-2D)-a wider perspective of the proteome. *J Proteomics* 73:1707–1721
 - 13. Sanchez-Pons N, Irar S, Garcia-Muniz N, Vicient CM (2011) Transcriptomic and proteomic profiling of maize embryos exposed to camptothecin. *BMC Plant Biol* 19:91
 - 14. Weiss W, Gorg A (2007) Two-dimensional electrophoresis for plant proteomics. In: Thiellement H, Zivy M, Damerval C, Mechini V (eds) *Plant proteomics methods and protocols*. Humana, Totowa, NJ, pp 121–143

Chapter 11

2DE Analysis of Forest Tree Proteins Using Fluorescent Labels and Multiplexing

Jenny Renaut, Céline Leclercq, and Sébastien Planchon

Abstract

Although proteomists working with gel-free methods are considering the gels as coming from the past, proteomics based on gels has still a lot of opportunities to offer and acquisition of images on which thousands of spots may be resolved is still largely performed. Nowadays, two-dimensional electrophoresis remains a powerful tool to explore the plant proteome and to unravel changes in protein abundance between samples. Some weak points can be pointed out, as for any method, as for example the lack of reproducibility, or the detection of low-abundance proteins. The use of the technique called “difference gel electrophoresis” or “DIGE” can help to overcome or at least to reduce these inconveniences. DIGE requires the labelling of proteins by fluorochromes prior to their separation on 2DE gels. This technique may be applied to a wide array of plant stress studies, among others to trees. Accurate quantitative results can then be obtained and proteins presenting an interest in the studied stress are subsequently subjected to an enzymatic digestion (usually with trypsin) and identified using electrospray ionization, matrix-assisted laser desorption/ionization-time-of-flight-MS, and/or tandem MS.

Key words Two-dimensional electrophoresis (2DE), Quantitative approach, DIGE, Protein abundance, Trees

1 Introduction

Proteomics, i.e., the study of the protein complement of the genome [1], is one of the high-throughput techniques of the post-genome era, such as transcriptomics, metabolomics, and phenomics, taking advantage of the completion of several sequenced genomes (e.g., *Arabidopsis*, rice, poplar) [2–4]. Proteomics allows to study plants at different stages of their life cycle, and to directly relate the proteins and their functions in plants. Two main approaches are used: the gel-based approach (e.g., 2DE) and the gel-free analyses (e.g., LC-MS/MS), each presenting advantages and inconveniences. In plants and especially for trees, a majority of the results are still provided by the use of classical 2DE (see [5, 6] for a review).

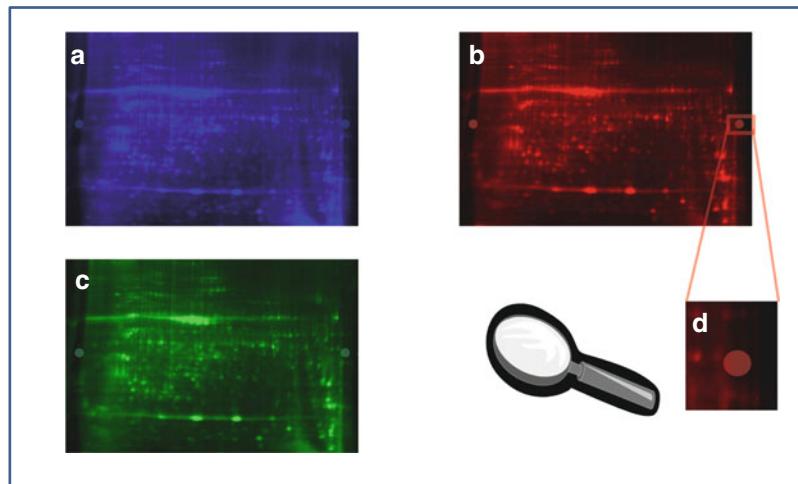


Fig. 1 Visualization of proteins extracted from red cabbage leaves and labelled with (a) Cy2 (in blue), (b) Cy3 (in red), and (c) Cy5 (in green). (d) Represents a magnified area of the gel, carrying the reference markers (see Subheading 2.3, item 7, and Note 14)

The gel-approach includes different steps: extraction of proteins, separation by 2DE, visualization of the gels, software-based analysis, and finally cutting of the spot out of the gel before the identification by mass spectrometry and database search. More or less 10 years ago, a technique called difference gel electrophoresis (DIGE, *see* Fig. 1) allowed an important improvement in the quantification results [7, 8], reducing gel-to-gel variations by the introduction of internal standard and sample multiplexing on the gels.

Trees are of course of economic and ecological importance. In plant science, they also bring their specific features in order to improve our understanding of plant life cycle. Indeed, contrarily to herbaceous, woody plants have to endure the four seasons and to adapt their life cycle to particular conditions. Therefore, they are used to study dormancy, bud burst, seasonal senescence, and so on. For economic aspects, they can be used for phytoremediation studies, for yield study related to stress conditions, for response towards pests, and so on. Proteins give a good picture of what is happening in plants under specific conditions, and therefore combining tree studies with proteomics can improve our understanding of their involvement in plant's metabolism.

2 Materials

In our laboratory, the equipment comes from GE Healthcare, TECAN, and Serva, but many other suppliers with equivalent equipment can also be considered.

2.1 Sample Preparation

1. Reswelling tray (GE Healthcare, Uppsala, Sweden).
2. Extraction solution for TCA/acetone precipitation (extraction 1): 20 % (w/v) trichloroacetic acid (TCA), 0.1 % (w/v) dithiothreitol (DTT) in ice-cold acetone.
3. Rinsing solution for TCA/acetone precipitation (extraction 1): 0.1 % (w/v) DTT in ice-cold acetone.
4. SDS buffer (extraction 2): 4 % (w/v) Sodium dodecyl sulfate (SDS), 5 % (w/v) sucrose, 0.3 % (w/v) DTT, 20 mM Na₂HPO₄, pH 7.0.
5. Precipitation solution I (extraction 2): 10 mM DTT in ice-cold acetone.
6. Polyvinylpolypyrrolidone (PVPP, Sigma, Bornem, Belgium).
7. Phenol (Tris-buffered, pH 8.0).
8. SDS-tris buffer (extraction 3): 2 % (w/v) SDS/30 % (w/v) sucrose/0.1 M Tris-HCl/5 % (v/v) β-mercaptoethanol, pH 8.0.
9. Precipitation solution II (extraction 3): Cold 0.1 M ammonium acetate in methanol.
10. Rehydration buffer: 7 M urea, 2 M thiourea, 0.5/2 % (v/v) pharmalyte or IPG buffer, 0.002 % (w/v) bromophenol blue, 2 % (w/v) 3-((3-cholamidopropyl)dimethylammonium)-1-propanesulfonate (CHAPS).
11. pH test strips 4.5–10.0 pH, resolution: 0.5 pH unit (Sigma Aldrich).
12. CyDyes (Cy2, Cy3, and Cy5) stock solution: Add 25 μL dimethylformamide (DMF, less than 3 months old) to 25 nmol of dye.
13. Labelling solution: 7 M urea; 2 M thiourea; 4 % (w/v) CHAPS; 30 mM Tris-HCl (see Note 1).
14. Stopping solution: 10 mM lysine.

2.2 Isoelectro-focusing

1. Sample buffer (2×): 7 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 2 % (v/v) ampholyte, 2 % (w/v) DTT (see Notes 1 and 2).
2. IPGphor (GE Healthcare, Uppsala, Sweden) or a similar electrophoresis apparatus.
3. Manifold, paper wicks, sample cups.
4. Immobilized pH gradient strips: Gels strips may have different lengths and pH gradients. In the present example, we present the protocol for IPG strips of 24 cm on a pH gradient of 4–7.
5. Mineral oil (e.g., dry strip cover fluid; GE Healthcare).

2.3 Second Dimension—Vertical System

1. Dalt twelve Gel Caster (GE Healthcare, Uppsala, Sweden).
2. Displacing solution: 0.375 M Tris-HCl, pH 8.8, 50 % (v/v) glycerol, trace of bromophenol blue.

3. Acrylamide stock solution: 30.8 % (30 % acrylamide, 0.8 % *N,N*-bisacrylamide). This solution can be adjusted for different applications.
4. Homogenous solution for gels: 12.5 % T acrylamide stock solution, 1.5 M Tris-HCl, pH 8.8, 0.1 % (w/v) SDS, 0.1 % (w/v) ammonium persulfate (APS), 0.01 % (v/v) tetramethylene diamine (TEMED).
5. Low-fluorescence glass plates.
6. Bind-Silane solution: 80 % (v/v) ethanol, 0.2 % (v/v) glacial acetic acid, 0.01 % (v/v) Bind-Silane (3-methacryloxypropyltrimethoxysilane).
7. Fluorescent reference markers (GE Healthcare).
8. Equilibration buffer: 6 M urea, 50 mM Tris-HCl, pH 8.8, 30 % (v/v) glycerol, 2 % (w/v) SDS.
9. Agarose sealing solution: 25 mM Tris base, 192 mM glycine, 0.1 % (w/v) SDS, 0.5 % (w/v) agarose, 0.002 % (w/v) bromophenol blue.
10. Running buffer: 250 mM Tris base, 1.92 M glycine, 2 % (w/v) SDS.
11. Ettan Dalt twelve (GE Healthcare, Uppsala, Sweden) or Protean Plus Dodeca cell (Bio-Rad).
12. Typhoon 9400 (GE Healthcare, Uppsala, Sweden).
13. Fixing solution: 40 % (v/v) ethanol, 1 % (v/v) glacial acetic acid.
14. Ettan Spot Handling Workstation (GE Healthcare, Uppsala, Sweden).

2.4 Second Dimension—Horizontal System

1. HPE-FlatTop Tower (Serva).
2. 2D-Large-Gel flatbed NF 12.5 % kit (Serva) containing electrode buffers (anode and cathode buffers), electrode wicks, equilibration buffer, contact fluid, precast 2D-gel (255 × 200 × 0.65 mm) on a nonfluorescent plastic sheet, DL-DTT, iodoacetamide (IAA), urea.

2.5 Handling of Spots and Peptide Extraction

1. 96-well polypropylene microplates (V-shaped; NUNC).
2. Washing solution: 50 % (v/v) methanol (MS grade), 50 mM ammonium bicarbonate.
3. Acetonitrile (75 %, v/v, MS grade).
4. Trypsin solution: 5 µg/mL Trypsin Gold (mass spectrometry grade; Promega) in 20 mM ammonium bicarbonate.
5. Peptide extraction solution: 50 % (v/v) acetonitrile, 0.1 % (v/v) trifluoroacetic acid (TFA).

6. Sample-spotting solution: 50 % (v/v) acetonitrile, 0.1 % (v/v) TFA.
7. Matrix solution: α -cyano-4-hydroxycinnamic acid (7 mg/mL in sample-spotting solution).

3 Methods

The extraction of the proteins from the plant sample is a crucial step. Many extraction methods exist for tree samples, and none can be considered as “THE” protocol. It is thus very important to optimize the extraction protocol and the sample handling to avoid interfering substances (lipids, phenolics, nucleic acids, etc.). We propose here three examples of protocols successfully tested in our laboratory on different tree species (e.g., poplar, oak, willow, alder, eucalyptus). In the three cases, the extracted proteins are coming from the entire tissue considered (e.g., roots, stems, cambium, leaves). Protein extraction methods using TCA/acetone precipitation, hot SDS, and/or phenol extraction, respectively, are proposed [9–11]. We do not consider here extraction as specific for compartments but rather as a “total” cellular extraction. However, fractionation steps to isolate, e.g., chloroplasts, mitochondria, nuclei, or membranes [12–14], can precede these different methods. However, a protocol should be optimized for individual studies. It is important to note that depending on the physiological status of the plant tissue, protocols may vary. Indeed, we observed that one protocol usually suitable for young or mature leaves was not optimal for senescent leaves. So, optimization remains of uttermost importance.

The labelling reaction between the CyDyes and the proteins is based on the formation of a covalent bond between the lysine residues of the proteins and the *N*-hydroxysuccinimidyl ester reactive group. The CyDyes have been designed to have approximately the same mass (around 500 Da) and to carry a positive charge to replace the intrinsic positive charge of the lysine residue. This will thus ensure that the pI of the protein is not modified and its mass barely changed, although it may be visible in the low molecular mass range on the gel. Once the extraction has been optimized, the proteins from different samples are quantified, labelled, separated, and compared (Fig. 2) (see Note 3).

Fluorescence and multiplexing can also be used to study phospho- or glyco-proteins, namely, with Pro-Q® Diamond Phosphoprotein Stain or Pro-Q® Emerald Glycoprotein Stain from Molecular Probes that can be used in combination with the Cydyes or other fluorescent staining (SyproRuby, LavaPurple).

Some other labels are now available on the market for protein multiplexing, named G100, G200, and G300 developed by NH

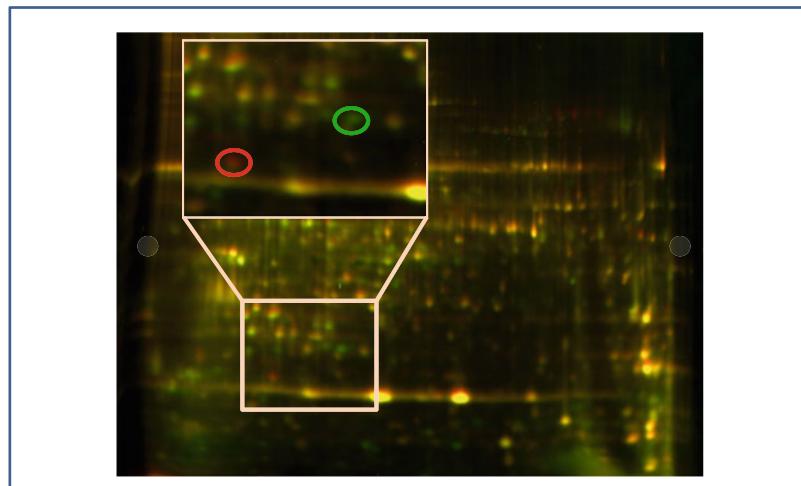


Fig. 2 Overlay of the samples containing proteins extracted from red cabbage, as seen in Fig. 1, with a magnified view showing proteins more abundant in the Cy3-labelled sample (in red) and some more abundant in the Cy5-labelled samples (in green). Proteins showing comparable relative abundance are appearing in yellow

dyeagnostics (Halle, DE), although not considered as equally powerful as Cydyes, in our case. These labels have different molecular weights and this feature may be corrected by a software device.

In conclusion, it is of primary importance to correctly plan the experimental design, the number of samples needed, the extraction, the homemade solutions, and the time needed for the different steps.

3.1 Extraction of the Proteins

3.1.1 TCA/Acetone Precipitation

Approximately 300 mg of plant fresh material is crushed in liquid nitrogen, in cold room, and treated with cold extraction buffer (20 % TCA, 0.1 % DTT, acetone). Proteins are allowed to precipitate overnight at -20 °C. After centrifugation ($30,000 \times g$; 45 min at 4 °C), the pellet is washed with rinsing buffer (acetone/DTT 0.1 %) three times. The supernatant is discarded and the pellet is dried in vacuo.

3.1.2 HOT-SDS Extraction

300–500 mg of plant fresh material is crushed, in cold room. 200 mg of PVPP and 2 mL SDS buffer preheated at 65 °C are added to samples. This mixture is vortexed for 5 min and subsequently put on ice for 15 min. After centrifugation at $15,000 \times g$ for 15 min the supernatant is kept and a re-extraction of the pellet is performed with the same method. The obtained supernatant can be pooled with the first one. Proteins are allowed to precipitate overnight at -20 °C by adding 8 mL of precipitation buffer I. Samples are centrifuged (45 min, $30,000 \times g$, 4 °C). The pellet is washed the same two times. The supernatant is discarded and the pellet is dried in vacuo.

3.1.3 Combination of TCA/Acetone and Phenol-SDS

The first step consists of preparing dry tissue powder (TCA/acetone). Plant tissue is ground in liquid nitrogen in precooled mortar and pestle to a fine powder: 200–400 mg of powder were suspended in 1 mL of cold extraction buffer. Samples were vortexed thoroughly for 30 s, left at –20 °C for 30–60 min, and centrifuged at 10,000 $\times g$ for 5 min, at 4 °C. The supernatant was discarded and the pellet was washed with cold acetone twice (centrifuge at 10,000 $\times g$, for 3 min, at 4 °C). The pellet was dried in vacuo. The second part of this protocol is protein extraction with phenol-SDS. The pellet obtained is resuspended in 800 μ L of buffer phenol (Tris-buffered pH 8.0) added with 800 μ L SDS-Tris buffer. Samples are mixed for at least 30 s and centrifuged for 3 min at 10,000 $\times g$. 300 μ L upper-phase phenol is withdrawn, put in new 2 mL microtube, 5 volumes of cold 0.1 M ammonium acetate in methanol are added and kept for 30 min at –20 °C. After centrifugation (10,000 $\times g$, 5 min), and the pellet is washed twice with precipitation buffer II and twice with 80 % cold acetone. The supernatant is decanted. Each time pellet should be re-suspended completely and centrifuged at 10,000 $\times g$ for 3 min at 4 °C. The final pellet is dried in vacuo.

3.2 Protein Labelling

Before starting the labelling, IPG strips have to be rehydrated. 450 μ L of rehydration solution containing ampholytes of pH 4–7 (24 cm) is placed in each lane of the reswelling tray. The IPG strip can be placed in the lane with or without the light plastic cover protecting the gel itself, with the gel side down. The gel is covered with mineral oil to avoid drying and crystallization of urea. This step has to be carried out overnight or at least for 10 h.

1. The extracted proteins are resolubilized in the labelling solution. The high concentration of urea allows the denaturation of proteins in order to obtain all the proteins present in one conformation before their separation.
2. Reconstitution of CyDye with DMF: An anhydrous solution of DMF (>99.8 % purity) has to be used (*see Note 4*). To prepare a stock solution of dyes, the adequate volume of DMF is added to the tube containing the CyDye to reach the concentration of 1 nmol/ μ L. This solution may be kept for 2 months at –20 °C.
3. The solution containing the proteins has to be equilibrated at a pH between 8.0 and 9.0, the optimum pH for the reaction between the CyDye and the proteins being 8.5. A first test of pH can be done with 0.5 μ L of solution on pH paper and the adjustment is done by using NaOH solutions (0.1 or 1 M depending on the starting pH of the protein solution).
4. Before the labelling and after the pH adjustment, the contents of the protein sample have to be quantified. Several methods

are available from different companies (Bio-Rad, GE Healthcare, Sigma, etc.). Most of these methods are based on a short protocol of precipitation before the quantification by colorimetry. The optimal concentration for the labelling reaction is around 5 µg/µL.

5. 30 µg of protein is labelled for each sample by Cy3 and Cy5, and the internal standard is prepared with an equal part of each sample, before the labelling with Cy2, classically used for the internal standard (*see Note 5*).
6. Each CyDye is added to the Eppendorf tube containing the protein solution. The ratio of 400 pmol/µL of dye to 50 µg of protein has to be respected to optimize the labelling reaction, i.e., in our laboratory, 0.24 µL of dye solution is added to each tube containing 30 µg of protein solution. After vortexing and a short centrifugation to concentrate the solution at the bottom of the tube, the mixture is incubated for 30 min on ice in the dark.
7. After 30 min of incubation, 1 µL of stopping solution is added and the mixture is vortexed and incubated again for 10 min on ice in the dark.
8. The labelled samples can be stored at -80 °C (for up to 3 months) or be used directly (*see Note 6*).

3.3 First-Dimension Separation: Isoelectrofocusing

1. The labelled samples have to be combined in one tube before the loading; each gel should contain 30 µg of protein labelled with Cy3, a second sample of 30 µg labelled with Cy5, and 30 µg of internal standard labelled with Cy2.
2. The volume of each combination has to be set at 120 µL before cup loading by adding 2× sample buffer solution containing DTT and ampholytes.
3. The rehydrated IPG strips are removed from the reswelling tray and transferred, gel side up, into the manifold properly positioned on the IPGphor system.
4. Electrode paper wicks, wetted with deionized water, are placed at the end of the IPG strips. Electrode is set on top of these papers before the positioning of the cup. It is important that the electrode wires are in contact with the paper.
5. IPG strips have to be covered with mineral oil to avoid drying during the first-dimension run. Pour around 100 mL on top of the gels into the manifold (*see Note 7*).
6. The combination of the three samples (labelled with Cy3, Cy5, and internal standard) is loaded in the cup. Before starting the electrophoretic run, the samples have to be covered with mineral oil, still to avoid desiccation and the formation of urea crystals in the cups.
7. The electrophoretic run can be started: in this example, the conditions of the run are the following at 20 °C: 100 V for 2 h,

300 V for 3 h, 1,000 V for 6 h, a gradient step up to 8,000 V during 3 h, and a constant step at 8,000 V to reach 60,000 V h with a maximum current setting of 50 μ A/strip.

8. If the IPG strips are not used directly after the IEF run, they can be stored at -80°C between plastic sheets and maintained horizontally.

3.4 Second-Dimension Separation: SDS-PAGE

3.4.1 Vertical Gels

Gels can be cast in the laboratory or precast gels can be acquired from different companies (e.g., GE Healthcare, Serva).

1. Casting of the gels between low-fluorescence glass plates:

The glass plates are placed in the gel caster. The homogeneous solution for gels is first cooled (4°C) and degassed before addition of APS and TEMED. These last two compounds catalyze the polymerization of the gels and thus can be added directly only before casting. This solution is poured into the gel caster, and 1 mL of water-saturated butanol is added on top of each gel. The polymerization must be performed for at least 4 h.

2. Before the second dimension, the IPG strips are equilibrated in a buffered solution containing SDS (see Note 8), saturating the proteins in the IPG strips with negative charges to allow the mass-dependent separation of the protein in the second-dimension gel. A first step is carried out in equilibration buffer containing 1 % DTT (w/v) for 15 min and then a second step with the equilibration buffer complemented with 2.5 % (w/v) IAA to avoid formation of random S-S (see Note 9).
3. After equilibration, IPG strips are deposited on top of the SDS-PAGE and then sealed with hot agarose solution.
4. The gels are put in the electrophoresis vertical tank “Ettan Dalt twelve” and the running buffer is poured into the tank. The second-dimension run is started either for several hours (first 30 min at 1.5 W per gel, then 17 W per gel for 5 h) or overnight (first 30 min at 1.5 W per gel, then 2.5 W per gel for the night) at 20°C until the dye front reaches the bottom of the gels.

3.4.2 Horizontal Gels

The HPETM (high-performance electrophoresis) FlatTop Tower allows electrophoretic separations of up to four horizontal flatbed gels at the same time.

1. Preparing the HPE gels: The electrode wicks from the kit are soaked for 15 min with electrode buffer, one wick with anode buffer and one with cathode buffer per gel. 4 mL of cooling solution is poured on the cooling plate. The gel is then set, gel side up, and IPG strip-slot into the gel toward the cathode, on the cooling plate. The cooling solution should form a homogeneous layer, without air bubble, between the gel and the plate. Finally, the cathode wick is placed on the gel, with-

out covering the slot, to make the connection between the electrode and the gel. The anode electrode wick is similarly placed on the other side of the gel.

2. As for the vertical system, the IPG strips are incubated in equilibration buffer containing 0.3 % (w/v) urea and 0.8 % DTT (w/v) for 15 min and then a second step with the equilibration buffer complemented with 0.3 % (w/v) urea and 2 % (w/v) IAA.
3. After equilibration, IPG strips are deposited, gel side down, in the strip-slot. The lid is closed on each gel-drawer, allowing the application of the electric current.
4. The second-dimension run is started according to the manufacturer's instructions at 15 °C either for several hours or overnight. The strip is removed from the gel after 70 min of low-voltage steps (4 W for 30 min—12 W for 30 min—20 W for 10 min for 4 gels). The immobilized pH gradient would interfere with the high voltage of the next steps resulting in a poor resolution. The run itself is set at 120 W for 4 h, then 160 W for 50 min for short run and 8 W overnight (10–15 h), and then 120 W for 3 h for overnight runs. The run is stopped when the dye front reaches the end of the gel.

3.5 Image Acquisition and Analysis

1. When the gels are taken out of the SDS-PAGE tank, they are thoroughly rinsed with water (*see Note 10*) and dried with paper before the image acquisition in a scanner with three different lasers—or lasers able to combine three adequate wavelengths of emission and excitation—(here a Typhoon imager). Images are acquired by using three different lasers at 488, 532, and 633 nm (excitation wavelengths) and 520, 590, and 680 nm (emission wavelengths) for Cy2-, Cy3-, and Cy5-labelled proteins, respectively (thus, three images per gel, corresponding to the two samples and the internal standard) (*see Note 11*).
2. After image acquisition, the first step for the image analysis is the detection of the spots. Software dedicated to DIGE image analysis is able to detect the spots simultaneously on the three images obtained from the same gel. One of the main advantages of the DIGE is the use of the internal standard for the gel-to-gel matching (as the same sample is run on all the gels from the experiment) but also for the normalization among the different gels. This feature helps the future analysis of the different gels and the comparison of the protein abundances. After the gel-to-gel matching, powerful statistical analyses can be achieved on thousands of spots present in all the gels [15, 16].
3. Once proteins of interest for the experiment have been selected on the basis of statistical analysis, the next step is the excision of

the protein from a preparative gel (i.e., a gel containing more than 90 µg of proteins to obtain enough material for MS spectra acquisition). Depending on the excision mode (automated or manual), the use of Bind-Silane-coated glass plate has to be considered. A gel containing more or less 300 µg of proteins is run; the IEF step is performed on the same pH range, but with an increased amount of total volt per hour (e.g., 90,000 V h for 24 cm on a pH 4–7). The SDS-PAGE is carried out in the same conditions. This preparative gel is then stained with Coomassie blue or a fluorescent post-migration stain (*see Note 12*).

4. After the scanning of the preparative gel and matching with analytical ones, a pick list is generated and used with the automated picker (e.g., here the Ettan Picker; GE Healthcare). Excised gel plugs are collected in microplates and used for protein identification protocols (*see Note 13*).
5. In this example, the protocol described is automatically performed on the Ettan Spot Handling Workstation (GE Healthcare), composed of an Ettan Picker, a digester, and a spotter, in a closed cabinet to avoid contamination with keratin, for example (*see Note 14*). The digestion is performed in a temperature-controlled closed microplate tower. Washing steps aiming at removing interfering substances on the gel plugs such as Coomassie blue and excess of SDS (providing high background on the MS spectra) are then performed with the washing solution (150 µL, two or three times for 20 min depending on the intensity of the staining if Coomassie blue has been used). A solution of 75 % acetonitrile is then added in the microplate wells for 30 min, then removed, and the gel plugs are left to dry at room temperature.
6. Protein contained in the gel plug is digested by the addition of trypsin or another proteinase; here 6 µL of trypsin solution is added to the dried gel piece, the microplate is covered, and the digestion is performed at 37 °C for 4 h.
7. Extraction of peptides: 35 µL of peptide extraction solution is added in the well at the end of the digestion step and the supernatant is recuperated in a new microplate. This step is repeated twice and the supernatant pooled with the previous one. Peptides are then concentrated by drying.
8. Peptides are resuspended in 2 µL sample-spotting solution and mixed in the microplate wells. Then, 0.7 µL of this mixture is put on a MALDI target, and 0.7 µL of matrix solution is added on top (with three mixing strokes in the spotter needle before final deposition).
9. Once the peptides are on the MALDI target, the instrument is calibrated with a mixture of known peptides and the acquisition of spectra (MS and MS/MS) is started.

10. The data acquired on the mass spectrometer are used for identification searches in a protein databases. Several types of software may be used for automated and manual interpretation of the results. A non-exhaustive list of the Web sites offering this feature is given here:
 - (a) http://fasta.bioch.virginia.edu/fasta_www2/fasta_list2.shtml
 - (b) http://www.matrixscience.com/search_form_select.html
 - (c) <http://prospector.ucsf.edu/>
 - (d) <http://www.expasy.org/tools/>

The MS spectrum provides a list of masses corresponding to peptides generated during the digestion of the protein. The MS/MS spectra allow the sequence of amino acids from a specific peptide selected for fragmentation. The obtained masses and sequences are compared with existing databases. After identification, literature survey is usually requested to link the results obtained to the initial biological question. More information about MS can be found in the following selection of excellent reviews: general mass spectrometry [17] in model plants [18, 19] and in non-model plants [20].

Finally, these steps are given as example of the protocols that are working well in our laboratory, and are used frequently on different matrices: potato, poplar, alfalfa, oak, alder, and other plants, but also on mosses, fungi, etc. The number of spots depends greatly on the tissue used, the explored pI range, whether pre-fractionation steps are used, and so on. Usually, we consider that a good result is between 70 and 95 % of identified proteins, varying according to the species, and the availability of its genome.

4 Notes

1. Urea and thiourea help the solubilization of proteins.
2. DTT preserves the fully reduced state of denatured, unalkylated proteins.
3. The classical colorimetric method of Bradford cannot be used for protein quantification, as the components of the labelling solution interfere with the Bradford solution (e.g., the high concentration of urea, detergent).
4. The DMF solution should not be contaminated with water, which degrades DMF to amine compounds. Select a DMF solution of 0.005 % H₂O, 99.8 % of purity, and this solution should be replaced at least every 3 months.
5. A dye swap is always recommended to avoid preferential labelling.

6. Optionally, the labelling reaction can be checked on a small SDS-PAGE on some of the samples before running the complete series of DIGE.
7. When the mineral oil covers the gels, check that there is no leakage (no oil in the sample cups).
8. SDS denatures proteins and forms negatively charged protein-SDS complexes.
9. IAA alkylates thiol groups on proteins, preventing their reoxidation during electrophoresis. Protein reoxidation during electrophoresis can result in streaking and other artifacts. IAA also alkylates residual DTT to prevent point streaking and other silver-staining artifacts.
10. In the above text, if it is not mentioned otherwise, “water” means double-distilled water.
11. If it is not possible to scan all the gels at the same day, they can be stored at 4 °C for 48 h without variation in the signal and the position on the gel. For longer storage periods, the gels have to be fixed in fixing solution and then kept at 4 °C.
Before scanning, leave the gels on the bench for a while to reach room temperature (as the temperature impacts the signal intensity).
12. Preparative gels, containing high amount of proteins, can be prepared separately for the identification of proteins.
We recommend either to prepare an important amount of proteins coming from the different conditions (e.g., with the same amount as for internal standard) or to run one preparative gel per sampling condition.
In this case, the labelling step is omitted and the gel is stained with Coomassie or another fluorescent solution (Sypro Ruby, Lava Purple, or others). At least 300 µg have to be loaded on the IPG strips, and the time for IEF has to be extended (for a 24-cm gel, pH 4–7, the recommended value to reach is 90,000 V h). It is also recommended to use Bind-Silane-coated glass plates to facilitate the picking (by avoiding shrinking or deformation of the gels).
13. Care is needed with the microplates; digestion carried out in microplates from several suppliers has resulted in the presence of polymers in the mass spectra, rendering the MS/MS spectra unreadable.
14. If the preparative gels are stained with fluorescent dyes and/or used for picking with the Ettan Spot picker or the Spot Handling Workstation, reference markers must be placed on the Bind-Silane-coated plate.

References

- Wilkins MR, Sanchez JC, Gooley AA et al (1996) Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev* 13:19–50
- The Arabidopsis Genome Initiative (2001) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–815
- International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. *Nature* 436:793–799
- Tuskan GA, Difazio S, Jansson S et al (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313:1596–1604
- Jorrin JV, Maldonado AM, Castillejo MA (2007) Plant proteome analysis: a 2006 update. *Proteomics* 7:2947–2962
- Abrial N, Gion J-M, Kerner R et al (2011) Proteomics research on forest trees, the most recalcitrant and orphan plant species. *Phytochemistry* 72:1219–1242
- Skynner H, Rosahl TW, Knowles MR et al (2002) Alterations of stress related proteins in genetically altered mice revealed by two-dimensional differential in-gel electrophoresis analysis. *Proteomics* 2:1018–1025
- Alban A, David SO, Bjorkesten L et al (2003) A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics* 3:36–44
- Saravanan RS, Rose JKC (2004) A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics* 4:2522–2532
- Carpentier SC, Witters E, Laukens K et al (2005) Preparation of protein extracts from recalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis. *Proteomics* 5:2497–2507
- Delaplace P, van der Wal F, Dierick J-F et al (2006) Potato tuber proteomics: comparison of two complementary extraction methods designed for 2-DE of acidic proteins. *Proteomics* 6:6494–6497
- Rolland N, Ferro M, Ephritikhine G et al (2006) A versatile method for deciphering plant membrane proteomes. *J Exp Bot* 57:1579–1589
- Zabrouskov V, Giacomelli L, van Wijk KJ et al (2003) A new approach for plant proteomics: characterization of chloroplast proteins of *Arabidopsis thaliana* by top-down mass spectrometry. *Mol Cell Proteomics* 2:1253–1260
- Keech O, Dizengremel P, Gardestrom P (2005) Preparation of leaf mitochondria from *Arabidopsis thaliana*. *Physiol Plant* 124:403–409
- Fenselau C (2007) A review of quantitative methods for proteomic studies. *J Chromatogr B* 855:14–20
- Gottlieb DM, Schultz J, Bruun SW et al (2004) Multivariate approaches in plant science. *Phytochemistry* 65:1531–1548
- Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. *Nature* 422:198–207
- Baginsky S (2009) Plant proteomics: concepts, applications, and novel strategies for data interpretation. *Mass Spectrom Rev* 28:93–120
- Glinski M, Weckwerth W (2006) The role of mass spectrometry in plant systems biology. *Mass Spectrom Rev* 25:173–214
- Carpentier SC, Panis B, Vertommen A et al (2008) Proteome analysis of non-model plants: a challenging but powerful approach. *Mass Spectrom Rev* 27:354–377

Chapter 12

Differential Plant Proteome Analysis by Isobaric Tags for Relative and Absolute Quantitation (iTRAQ)

María J. Martínez-Esteso, Juan Casado-Vela, Susana Sellés-Marchart, María A. Pedreño, and Roque Bru-Martínez

Abstract

Protein relative quantitation is one of the main targets in many proteomic experiments. Among the range of techniques available for both top-down and bottom-up approaches, *isobaric tags* for *relative* and *absolute quantitation* (iTRAQ) have gained positions within the top-rank techniques used for this purpose in the recent years. Briefly, each iTRAQ reagent consists of three different components: a reporter group (with a variable mass in the range of 114–117 amu), a balance group, and an amino-reactive group. The isobaric nature of iTRAQ-labeled peptides adds a signal to every peptide in the sample which is detectable in both MS and MS/MS spectra, thus enhancing the sensitivity of detection. During MS/MS, the reporter groups are released as singly charged ions with m/z ratios ranking from 114 to 117 amu, visible in the low mass region of MS/MS spectra. The iTRAQ technology can be used to analyze up to four different samples using the 4-plex kit (reporter groups 114–115 amu) or can be scaled up to eight different samples using the 8-plex kit (reporter groups 113–121 amu). In this chapter, we focus on the experimental procedures typically using 4-plex labeling, including tips leading to successful application of iTRAQ technology for the analysis of plant protein mixtures.

Key words iTRAQ, Isobaric tag, Liquid chromatography, Mass spectrometry, Plant proteomics, Quantitation, Strong cation exchange

Abbreviations

| | |
|-------|---|
| ACN | Acetonitrile |
| BCA | Bicinchoninic acid |
| DIGE | Differential in-gel electrophoresis |
| FDR | False-positive rate |
| FWHM | Full-wide half-maximum |
| cICAT | Cleavable isotope-coded affinity tags |
| LIT | Linear ion trap |
| LOPIT | Localization of organelle proteins by isotope tagging |
| MMTS | Methyl methanethiosulfonate |

| | |
|--------|---|
| NCBInr | National center for biotechnology information non-redundant |
| RPB-A | Reverse-phase chromatography buffer A |
| RPB-B | Reverse-phase chromatography buffer B |
| RPB-C | Reverse-phase chromatography buffer C |
| SDS | Sodium dodecyl sulfate |
| SMPW | Spectrum mill proteomics workbench |
| TEAB | Triethylammonium bicarbonate |
| TCEP | Tris(2-carboxyethyl)phosphine |

1 Introduction

Isobaric tags for relative and absolute quantitation (iTRAQ) is a recently developed technique for a mass spectrometric based quantitation of proteins. Since its first description by Ross et al. [1], in 2004, the amine-specific isobaric tag for relative and absolute protein quantification (iTRAQ) has become a consolidated technique in quantitative proteomics because large-fold changes of protein expression within broad dynamic ranges of protein abundance can be measured quite accurately [2]. The iTRAQ method allows for the multiplexed identification and quantification of proteins in four different samples (4-plex) and has been recently scaled to measure protein changes in up to eight different samples (8-plex) [3].

These reagents were designed to bear an isobaric skeleton among the different tags within the experiment. In the case of a 4-plex experiment, these tags are constituted in three parts: a peptide reactive group which is common for all of them, a specific reporter group unique to each tag (114–117 amu), and a neutral balance group (31–28 amu) to maintain a total mass of 145 amu (Fig. 1a). MS/MS fragmentation splits the four unique reporters away from their corresponding tags giving rise to a strong reporter ion signals in the low mass area of any given spectra. As displayed in Fig. 1b, the intensities detected correspond to the addition of the intensities of the four peptide populations, each of them bearing a different tag. Concomitant peptide sequencing and protein quantitation are allowed by combining *–b* and *–y* fragment ions together with reporter ions in a single MS/MS spectrum (Fig. 1c). iTRAQ reagents were designed to react through the peptide reactive group (NHS) with all primary amines found in tryptic peptides, including N-termini, ϵ -amino groups of Lys side chains, and, possibly, Tyr side chains. Labeling all the peptides within the four samples and pooling them prior to the mass spectrometric analyses imply enhancing the peptide coverage for any protein. The general workflow is depicted in Fig. 1 (panels B and C), showing the typical appearance of a 4-plex iTRAQ kit (Life technologies). Each sample is separately reduced, alkylated, digested, and labeled with an iTRAQ tag (114–117 Da) and then pooled together. The iTRAQ technique is currently a popular approach for quantitative

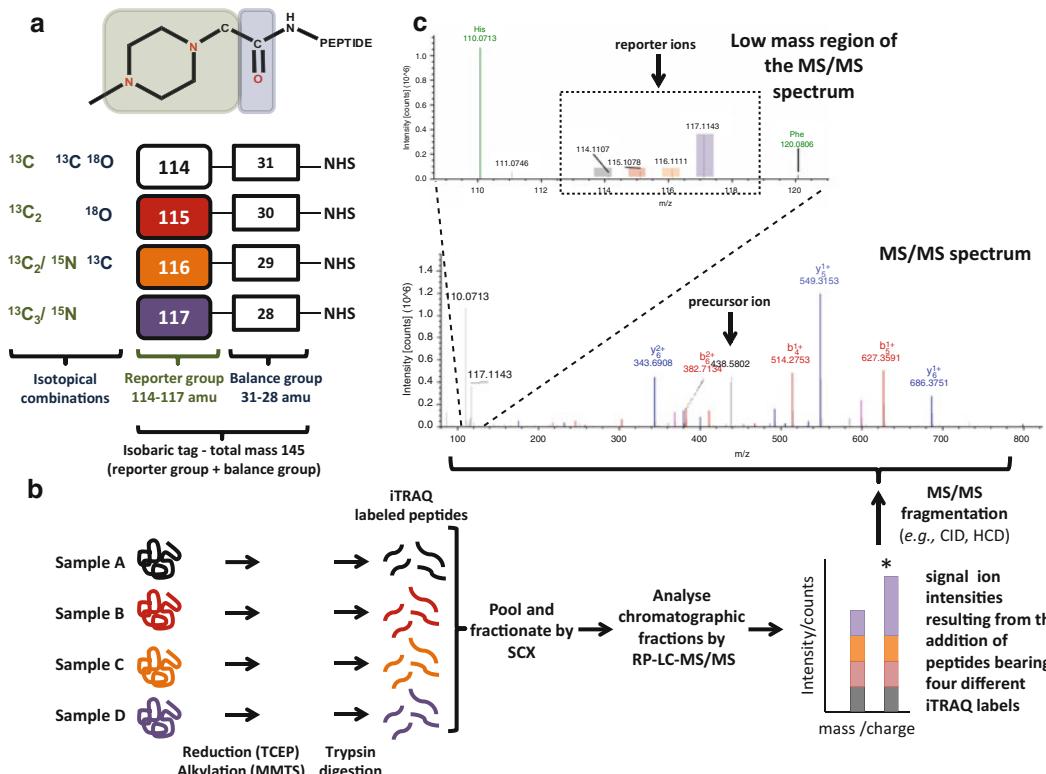


Fig. 1 (a) Diagram showing the components of the iTRAQ tags. The complete molecule consists of a reporter group (based on *N*-methylpiperazine) (shadowed in green), a mass balance group (carbonyl) (shadowed in blue), and an amine-specific peptide-reactive group (NHS ester). The overall mass of reporter and balance components of the molecule display a constant mass of 145 amu, the reporter group ranges in mass from 114 to 117 amu while the balance groups range from 28 to 31 amu. The mass of iTRAQ labels depends on the isotopic combination of N, C, and O atoms within the reporter group and the balance group found in each iTRAQ label. At the left side of each iTRAQ tags is depicted the precise isotopic combinations corresponding to both reporter (green box) and balance (blue box) groups. The tags form an amide linkage to amine groups of peptides (N-terminal or ϵ amino group of lysine) through the NHS group when reacted with a peptide. (b) An overview of a typical iTRAQ workflow. Peptides derived from the trypsin-digested protein samples are labeled with the iTRAQ tags, pooled, and then fractionated by strong cation exchange chromatography. (c) The resulting fractions are analyzed by LC-MS/MS in which a single ion or precursor is generated for each peptide resulting in a summed intensity from the species present on each sample; thus a mixture of four identical peptides each labeled with one iTRAQ tag appears as a single, unresolved precursor ion in MS (identical m/z). As example in the diagram, the precursor ion marked with an asterisk is isolated and subjected to MS/MS fragmentation using CID or HCD. Dashed lines correspond to a zoomed view showing the four reporter ions resolved and detected as fragment ions with m/z values ranging from 114 to 117 amu. All other sequence-informative fragment ions (mainly b - and y -fragment ions, displayed in red and blue color, respectively) remain isobaric after fragmentation, and their individual ion current signals (signal intensities) are also additive. Finally, the relative concentration of the peptides is deduced from the relative intensities of the corresponding reporter ions. Peptides, and their derived fragments ions, labeled with the iTRAQ tags are shaded

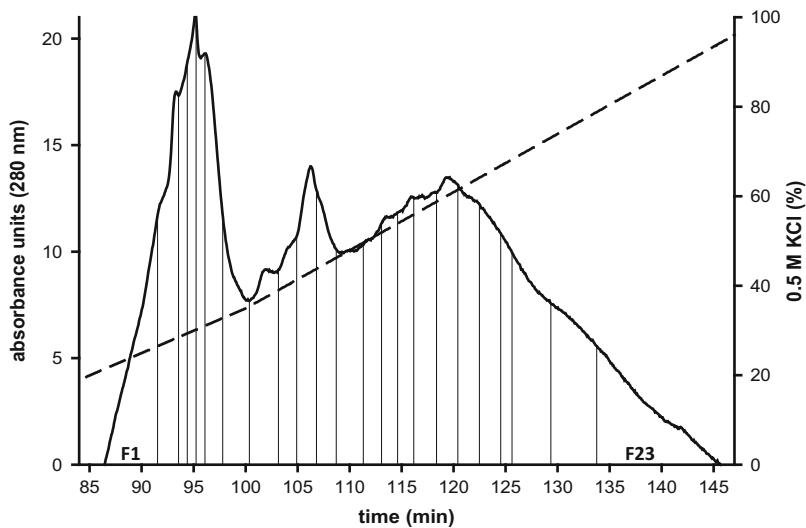


Fig. 2 Representative chromatogram corresponding to the fractionation of an iTRAQ-labeled complex mixture of peptides by strong cation exchange (SCX). Peptide elution was achieved using a gradient of SCX-B buffer containing 0.5 M KCl at 0.1 mL/min constant flow rate. 1-min eluate fractions were collected using an automated sample collector and, then, redistributed in 23 different fractions (F1 to F23) corresponding to approximately equal chromatographic areas ($\text{mAU} \times \text{min}^{-1}$ at 280 nm). The redistribution of peptides according to previously measured UV areas precludes the injection of samples too concentrated or too diluted and improves protein identification and quantitation

proteomics but the analysis of complex pools of peptides typically leads to certain bias such as underestimation of the number of proteins identified and compression of quantitation ratios. A good way to circumvent these limitations is the introduction of peptide fractionation steps prior to mass spectrometric analysis. In this sense, we chose the use of peptide fractionation by strong cation exchange (SCX) chromatography (Fig. 2). The eluted peptide fractions were analyzed by LC-MS/MS.

Currently, stable isotopic labeling followed by analysis through liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) is successfully used to characterize and quantify changes in protein levels in complex biological samples. It is a widely used technique for the differential expression study of proteomes. Since 2005, several publications are reported using iTRAQ technique for the study of plant proteomes [4–7].

The iTRAQ technology is a powerful technique allowing a wide flexibility in the experimental design which means a broad range of applications to address biological questions. iTRAQ has been extensively used to achieve a temporal quantitation of a biological process allowing monitoring time-course experiments [8–12, 29]. Also, its application to study the effect of a treatment [13, 14], a physiological response [15, 16], proteome comparisons [17], or functional studies based on mutant proteome differential

display versus wild type [18] has been reported. iTRAQ technology has expanded as a universal tool with unlimited possibilities only restricted by the experimental design needed to resolve specific biological questions. Among all the possible applications of iTRAQ is its use as a preferential technique to quantify the phosphorylation status of cells [19–21] which is based on the retention and preservation in the labeled peptides of other important structural features such as posttranslational modifications. A brilliant application is the use of iTRAQ for topological localization of organelle proteins which is known as the LOPIT approach [22]. Summarizing, iTRAQ is currently one of the most robust techniques applicable to quantitative proteomics, though some inherent characteristics of this technique may affect the accuracy of the protein quantitative ratios measured, including technical, experimental, and biological variability [23]. Independent analysis of at least two biological replicates is advisable in iTRAQ analyses. In those cases where sample limitation or financial constraints cause a problem, it is possible to use a sample pooling strategy [24]. However, it should be pointed out that the iTRAQ technique offers remarkable advantages compared to other quantitative proteomic techniques such as cleavable isotope-coded affinity tags (cICAT) or differential in-gel electrophoresis (DIGE) in regard to detection limit [25] having a good dynamic range detecting up to 24-fold changes [2] and a great potential to identify and quantify low-abundance proteins [26] which makes this technique a good promise having a great potential in discovery studies.

2 Materials

2.1 Sample Preparation and Labeling (See Note 1)

1. 0.5 M stock solution of triethylammonium bicarbonate (TEAB).
2. 2 % (w/v) sodium dodecyl sulfate (SDS) dissolved in 0.5 M TEAB.
3. iTRAQ reagent labeling kit (Life Technologies). Store at -20 °C. Extremely susceptible to hydrolysis.
4. 50 mM stock solution of Tris-(2-carboxyethyl) phosphine (TCEP).
5. 200 mM Methylmethanethiosulfate (MMTS) dissolved in isopropanol.
6. Trypsin mass spectrometric grade.

2.2 Sample Pre-fractionation

1. SCX chromatography buffer A (SCX-A): 10 mM KH₂PO₄, 20 % acetonitrile (ACN), pH 2.7, made in Milli-Q water or HPLC-grade water.

2. SCX chromatography buffer B (SCX-B): 10 mM KH₂PO₄, 0.5 M KCl, 20 % ACN, pH 2.7 in HPLC-grade water.
3. SCX chromatography buffer C (SCX-C): 10 mM KH₂PO₄, 1 M KCl, 20 % ACN, pH 2.7 in HPLC-grade water.
4. SCX chromatography column: Mono S PC 1.6/5 sulfonate column, 1.6×50 mm, 10 µm beads (GE Healthcare).
5. Desalting activation solution: 50 % ACN.
6. Desalting sample buffer, equilibration, and wash buffer: 0.5 % trifluoroacetic acid (TFA) in 5 % ACN.
7. Desalting elution buffer: 0.1 % TFA in 70 % ACN.

2.3 Sample Analysis by LC-MS/MS

1. Reverse-phase (RP) chromatography buffer A (RPB-A): 0.1 % formic acid (FA) (v/v) and 5 % ACN in water.
2. RP chromatography buffer B (RPB-B): 0.1 % FA in ACN.
3. RP chromatography trap column: Zorbax 300SB-C₁₈ cartridges, 5×0.3 mm, 5 µm particle size (Agilent Technologies).
4. RP chromatography analytical column: Zorbax 300SB RP C₁₈ column, 75 µm×150 mm, 3.5 µm particle size (Agilent Technologies).

3 Methods

The technique and procedure detailed have been used in our group for the proteomic analysis of grapevine (*Vitis vinifera* L) berries [9]. First, the tissue was extensively cleaned up to remove polyphenols and lipids, and then proteins were extracted with phenol and precipitated with ammonium acetate–methanol.

3.1 Sample Preparation and Labeling

1. Prepare a protein sample free of contaminants and impurities avoiding potential interfering substances with iTRAQ protocol. These interfering substances are thiols (i.e., DTT, mercaptoethanol), high amounts of detergents and denaturants, proteases, and primary amines (e.g., Tris buffers). In such case, dilute your sample adding at least six to ten volumes of chilled acetone and incubate overnight at -20 °C. After centrifugation at maximum speed for 10 min using a benchtop centrifuge, decant acetone and centrifuge for an additional 2 min, using a pipette to remove any remaining acetone (see Note 2).
2. The protein pellet without drying is immediately resuspended by addition in sequential order of fresh saturated urea/0.1 M TEAB 1/1 (v/v) and SDS 2 % (w/v) to a final concentration of 0.2 %. Solubilize completely the protein sample if necessary by incubation overnight at 4 °C or alternatively by sonication (see Note 3).

3. Assess protein quantification using a suitable protein assay kit (e.g., BCA assay or Bradford) using at least three serial dilutions of your sample with three replicates each and average quantification values. Check assay compatibility with sample buffer.
4. Aliquot up to 100 µg of protein for labeling. And perform acetone precipitation as described above.
5. Resuspend completely the protein pellet in 30 µL iTRAQ dissolution buffer provided in the kit adding 3 µL of 2 % (w/v) SDS and if needed, followed by incubation of the protein sample overnight at 4 °C (*see Note 4*).
6. Reduce disulfide bonds by adding 2 µL of reducing agent from iTRAQ kit (50 mM TCEP). Vortex, pulse spin, and incubate at 60 °C for 1 h.
7. Bring protein sample to room temperature and alkylate by adding 1 µL of cysteine-blocking reagent from iTRAQ kit (200 mM MMTS). Vortex, pulse spin, and incubate for 10 min at room temperature.
8. Dilute protein sample by adding 250 µL of 100 mM iTRAQ dissolution buffer to ensure that SDS is lower than 0.05 % and perform digestion by addition of trypsin at a 10:1 protein:trypsin ratio (i.e., 10 µg trypsin per 100 µg protein). Vortex, pulse spin, and incubate at 37 °C overnight (*see Note 5*).
9. After overnight incubation, trypsin is added in a 20:1 protein:trypsin ratio (i.e., 5 µg trypsin per 100 µg protein) to ensure complete protein digestion. Vortex, pulse spin, and incubate at 37 °C for 6 h (*see Note 6*).
10. Protein sample digested is evaporated frozen under vacuum or lyophilized (*see Note 7*).
11. Reconstitute peptides in 30 µL iTRAQ dissolution buffer.
12. Thaw iTRAQ reagent vials to room temperature and add 70 µL of absolute ethanol provided in iTRAQ kit. Vortex during 1 min, pulse spin to take all the volume at the bottom of the vials, and add this solution to each sample. Incubate at room temperature for at least 1 h (*see Note 8*).
13. Add to each sample 1 mL of SCX-A buffer. Samples can be kept at 4 °C if the analysis is to be continued immediately or stored at -20 °C but not longer than 1 week (*see Note 9*).

3.2 Sample Pre-fractionation

1. Samples are pooled in a 15 mL clean tube and each sample tube is washed out with an extra 100 µL SCX-A buffer and added to pooled material ensuring recovery of all peptides.
2. Make final volume up to 5 mL with SCX-A buffer and check pH. Ensure that pH is ranged from 2.5 to 3.0 by adding small amounts of phosphoric acid.

3. Centrifuge pooled material either in the same tube or perform stacked centrifugations in a 1.5 mL vial in order to minimize liquid losses. Centrifugation is performed for 2 min at 13,000 rpm to ensure that any insoluble material is not loaded into SCX column. Repeat until the complete sample is processed.
4. Transfer sample to a clean tube for loading onto SCX column.
5. Fractionate labeled peptide mixture by SCX chromatography. Gradient details are sample and LC system dependent. We use ÄKTA Purifier (GE Healthcare) chromatograph. 5 mL of sample are injected into a 50 mL superloop. Sample is then loaded from the loop onto the SCX column and washed for 20 min with 100 % SCX-A at a flow rate of 0.1 mL/min.
6. Next step in the program is triggering of an elution gradient at a flow rate of 0.1 mL/min as follows: 5–35 % SCX-B in 30 min, 35–100 % SCX-B in 50 min, wash at 100 % SCX-B for 20 min, return to 0 % SCX-B (100 % SCX-A), and re-equilibrate column in SCX-A until conductivity signal is flat. 1-min fractions are collected upon triggering of gradient and peptides usually elute between 16 and 70–75 min, thus yielding 45–50 fractions. Peptide recovery is monitored using a UV detector (two different signals at 280 nm and/or 214 nm may be used to monitor peptide chromatographic peaks). This fractionation can be adjusted dependent on the instrument and the program conditions (*see Note 10*).
7. The UV absorption of the different fractions and the summed final chromatographic area ($\text{mAU} \times \text{min}^{-1}$ at 280 nm) are known after SCX. Thus, it is possible to redistribute the fractions in an effort to equalize the amount of peptides in each fraction and the number of MS/MS events after each mass spectrometric analysis. After redistribution of the peptides, around 25 fractions are collected and analyzed by LC-MS/MS (Fig. 2), precluding the injection of samples too concentrated or too diluted and, thus, improving protein identification and quantitation.
8. The combined SCX fractions are dried under vacuum and can be stored at -20 or -80 $^{\circ}\text{C}$ until analysis by reverse-phase LC-MS/MS. Preferably peptide cleanup can be carried out to ensure complete desalting prior to mass spectrometry (*see Note 11*).
9. For desalting of peptides, we use C₁₈ spin-cartridges (Agilent Technologies) following the manufacturer's recommendations. Briefly, resuspend sample in 100–150 μL of sample buffer. Wash C₁₈ beads in activation solution, equilibrate the beads in wash buffer and bind sample, wash resin at least three times in

wash buffer, and elute peptides in two steps by addition of 20 μ L elution buffer.

10. The peptides recovered from each SCX fraction are dried under vacuum for further analyses by MS/MS.

3.3 Sample Analysis by MS/MS

Currently, different mass spectrometers and several types of peptide fragmentation methods may be used to analyze iTRAQ-labeled samples. Q-TOF mass spectrometers are the most widely used instruments in which CID is the common fragmentation used. As a valid alternative, the recently introduced LTQ Orbitrap hybrid mass spectrometer able to perform higher energy collision-induced dissociation (HCD) also allows iTRAQ analysis. Here, we describe an iTRAQ analysis method using a parallel detection method combining CID peptide fragmentation in a linear ion trap (LIT) analyzer and HCD fragmentation in the HCD cell both in the LTQ Orbitrap analyzer. Compared to other methods reported, this method leads to higher protein identification yields together with improved quality of spectra and quantification.

1. The protocol described is for use with an Agilent 1200 HPLC system (chromatographic conditions and columns may vary using alternative instrumentation). Fractions are resuspended in 10 μ L of 0.1 % FA and 8 μ L injections are programmed to ensure reproducible sample injection in this autosampler. Tryptic peptides were pre-concentrated using Zorbax 300SB-C18 cartridges (5 \times 0.3 mm, 5 μ m particle size) at 15 μ L/min for 10 min in RPB-A followed by elution in a Zorbax 300SB RP C₁₈ column (75 μ m i.d. \times 150 mm and 3.5 μ m particle size) (Agilent Technologies) using a 60-min linear gradient from 5 to 40 % RPB-B flowing at 200 nL/min.
2. The eluent is sprayed directly into an LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with a Proxeon electrospray ion source. The electrospray voltage is set to 1,800 V and the capillary voltage to 50 V at 190 °C.
3. The LTQ Orbitrap is operated in the parallel mode, allowing for the accurate measurement of the precursor survey scan (400–2,000 m/z) in the Orbitrap selection, a 60,000 full-width at half-maximum (FWHM) resolution at m/z 400 concurrent with the acquisition of three CID data-dependent MS/MS scans in the LIT, followed by three data-dependent HCD MS/MS scans (100–2,000 m/z) with 7,500 FWHM resolution at m/z 400. The LTQ Orbitrap XL mass spectrometer includes an octopole acting as collision cell able to perform an alternative peptide fragmentation termed HCD [27]. The normalized collision energies used are 40 % for HCD and 35 % for CID. The maximum injection times for MS and MS/MS are set to 50 and 500 ms, respectively. The precursor isolation

width is 2 amu and the exclusion mass width is set to 10 ppm. Monoisotopic precursor selection is allowed and singly charged species are excluded. The minimum intensity threshold for MS/MS is 1,000 counts for the LIT and 8,000 counts for the Orbitrap. Target values for automatic gain control were set as follows: 5×10^5 for the Orbitrap survey scans, 1×10^5 for the Orbitrap MS/MS scans, 3×10^4 for LIT survey scans, and 1×10^4 for the LIT MS/MS scans.

3.4 Protein Identification

To analyze the data, a number of different software platforms may be used. The most extended used is for ProteinPilot software (Life Technologies) which is used to analyze data from instruments such as Q-TOF, all of them from Life Technologies company. The use of the Spectrum Mill Proteomics Workbench (SMPW) (Agilent Technologies) tool with a licensed module allowing the extraction of data or raw files from Thermo instruments is described here (*see Note 12*).

1. The search is jointly performed for all data files and SCX fractions which results in a summed and averaged data in the final results output.
2. Prior to launching the search, the MS/MS spectra dataset is processed with the extraction tool of SMPW to merge the MS/MS spectra with the same precursor ($\Delta m/z < 1.4$ Da and chromatographic $\Delta t < 15$ s) and fragmentation type.
3. A two-step search is performed. The reduced dataset is search first against a protein amino acid database including the following settings: trypsin, up to two missed cleavages, beta-methylthiolation of Cys, incomplete iTRAQ labeling (from no label to only Lys to only N-term to a complete label), Met-oxidation, Asn- and Gln-deamidation, and a mass tolerance for the LIT CID spectra of 2.5 Da for precursor and 0.7 Da for product ions, mass tolerance for the Orbitrap HCD spectra of 25 ppm for precursor and 50 ppm for product ions, and sequence tag length filter ≥ 4 . In a second search, the previous protein hits are researched using the following parameters: phosphorylation at Ser, Thr, and Tyr, and N-term Glu-pyroglutamic acid (*see Notes 13 and 14*).
4. Peptide hits are validated in SMPW in the protein mode and then in the peptide mode according to the manufacturer's recommended score settings. Briefly, identities interpreted for individual spectra are automatically designated as valid by applying the following scoring threshold criteria to all spectra, protein details mode: protein score > 20 , peptide details mode (score, scored percent intensity, delta rank1 – rank2) in the following order: peptide charge +2 (> 6 , $> 60\%$, > 2), peptide charge +1 (> 6 , $> 70\%$, > 2), peptide charge +3 (> 8 , $> 70\%$, > 2),

peptide charge +4 (>8, >70 %, >2), peptide charge +5 (>12, >70 %, >2), and peptide charge +2 (>6, >90 %, >1). Validation is edited to filter out those peptides with equal scoring in the direct and reversed search and Δ mass ($|\text{observed} - \text{expected}|$) >15 ppm. Validated files are summarized in the protein mode to assemble peptides into proteins in order to generate the minimal protein list that best explains the matched peptides. The false-discovery rate (FDR) for automated interpretation of the MS/MS spectra using the above criteria is estimated by searching all spectra passing the quality filter against the same database with all of the protein sequences reversed (*see Note 15*).

3.5 Data Analysis

1. After protein identification searches, a list for all peptides and proteins with the corresponding quantitative data and *p*-value is obtained. The log average ratio (log AR) in each iTRAQ ratio (i.e., in a 4-plex experiment the ratios are 115/114, 116/114, and 117/114) for all the peptides identified in the experiment should be zero. Otherwise, a correction factor of -log AR determined for each couple of labels has to be applied to the ratio data (*see Note 16*).
2. The confident iTRAQ ratios should be selected according to a good *p*-value, at least a *p*-value <0.05 for at least one ratio of the pair tags.
3. It is useful to build a list with all the proteins passing a restricted ratio considered as “significant” change. This list of proteins can be used for global bioinformatics analyses, e.g., pathway analysis (Blas2GO, Cytoscape, MapMan). Another list of proteins can be performed by inclusion of possible quantified proteins such as those having a nonsignificant ratio, a single peptide quantified, or higher quantification ratios but elevated *p*-values. This second list typically helps to interpret the data into a biological context. Those proteins with relevant changes for the biological system under study can be validated by other approaches such as western blot (*see Note 17*).

4 Notes

1. iTRAQ kit provides all the chemicals needed for iTRAQ labeling protocol just in enough amount. When TEAB buffer is needed in larger amounts to perform this protocol, a TEAB buffer from different commercial supplier apart from iTRAQ kit can be used. Extreme caution should be taken to manipulate iTRAQ kit in order to prevent hydrolysis of the tags. For this purpose, remove the tubes needed from the kit and return to the freezer immediately. Also, additional precaution can be

taken storing the kit inside a bag containing silica gel in the freezer.

2. The success for any proteomic experiment including a gel-free approach such as iTRAQ strongly depends on the quality of the protein extracts and on the quality of the solutions. Accordingly, special attention should be paid to optimize a high-quality protein extraction protocol. In this regard, protein extraction from plant material is well suited in a phenol-based protocol [28, 29]. All the chemicals and buffers should be prepared using HPLC-grade water, reducing the potential occurrence of contaminants and favoring peptide detection by mass spectrometry.
3. Protein pellets could be difficult to resuspend in TEAB buffer alone, so alternatively we use this buffer to solubilize protein extracts although having in mind that higher concentration of urea (>1 M) or SDS (>0.05 %) should be avoided during trypsin digestion.
4. The protein pellet should be resuspended in the minimal volume of iTRAQ dissolution buffer. The protocol described is optimized up to 50 μ L although for higher volumes it has to be scaled. Keep a ratio of 10 μ L iTRAQ dissolution buffer per 1 μ L SDS. Incubate sample for solubilization overnight at 4 °C and/or incubate at 60 °C for 30 min. From this step onwards do not use a sonication probe to resuspend the protein sample or any instrument implying sample losses.
5. It is important to use a trypsin of high quality, a mass spectrometric grade trypsin. The reason is that modified trypsin is treated to avoid auto-proteolysis, a process generating fragments that can interfere with protein sequencing. In addition, auto-proteolysis can result in the generation of pseudotrypsin, which has been shown to exhibit chymotrypsin-like specificity. The vials containing trypsin are resuspended in 100 mM of iTRAQ dissolution buffer mixing with a pipette, and the content is pooled in order to add the same volume of trypsin solution to each sample. For 20 μ g vials, resuspend in 250 μ L and add to each sample 120 μ L to ensure that the same volume is added to each sample. Trypsin digestion is a critical step in the iTRAQ protocol and should be performed under equal conditions for all the samples in the same experiment.
6. A critical step for the iTRAQ protocol success is trypsin digestion. So, a double digestion is performed to ensure a complete and exhaustive protein digestion.
7. Digested protein is evaporated frozen under vacuum to avoid sample losses.
8. Add iTRAQ reagents to one sample at a time, and vortex immediately. The iTRAQ reagent tags are extremely unstable

in an aqueous environment. In order to minimize hydrolysis and optimize efficiency of peptide labeling, maintain ratio iTRAQ dissolution buffer:ethanol at least in 4:6 during labeling reaction. The addition of iTRAQ dissolution buffer at high concentration (0.5 M) ensures around 100 mM concentration during labeling reaction (pH 7.0). The addition of ethanol precipitates pectins, which act as contaminants and may interfere with the analyses. Pectin polymers precipitate out of solution, converting samples in a gelatinous state unsuitable for further analysis [30]. Thus, it is strongly advisable to run a previous experiment without labeling exactly performed in the same way as the experiment with samples labeled with iTRAQ in order to guarantee a good condition for the sample under analysis.

9. Addition of SCX-A buffer hydrolyzes free tags and stops labeling reaction.
10. Wash SCX column between runs with SCX-C buffer to avoid cross-contamination among experiments.
11. A desalting step is important for removing salts prior to MS/MS.
12. There are available software solutions not limited to compatibility constraints. A thorough review of different freely available software for analysis of mass spectrometry data can be found in Mueller et al. [31].
13. Database searching can be speeded up by using a restricted database. This database can be created with a subset of National Center for Biotechnology Information non-redundant (NCBInr) protein database (<http://www.ncbi.nlm.nih.gov/>). Sequences are retrieved using the keyword search enquiry “organism,” adding potentially contaminating proteins retrieved using the keyword search enquiry “trypsin” or “keratin.”
14. In spite of the exponential availability of sequences by the recent plant genome sequencing projects and the public release of ESTs for plants within databases, there is still a difficulty for the protein identification for most plant species. Database searching for unsequenced organisms can be helped using different strategies as de novo sequencing [32, 33] or as described by Grossman et al. [34] by a high-throughput analysis based on a spectrum quality scoring, de novo sequencing, and error-tolerant Blast searches.
15. Apply iTRAQ correction factors in the SMPW and the software performs automatically iTRAQ calculations in Protein/Peptide Summary for the application of the correction factors.
16. Radiometric data conversion to \log_2 is recommended. The reason is that in the quantitative data obtained, twofold increase

is 2 while twofold decrease is 0.5. When converted to \log_2 these values become 1 and -1, respectively, and 0 as a value for not change.

17. Gehlenborg et al. [35] published a review for -omics analysis providing a good overview of the range of visualization tools available to deal with the interpretation of large dataset obtained by high-throughput analyses, including iTRAQ.

Acknowledgements

This work has been supported by research grants from MICINN-FEDER (BIO2008-2941 and BIO2011-29856-C02-02). MJME acknowledges a grant from Fundación Caja Murcia-University of Alicante. University of Alicante Proteomics Facility is a member of Proteored (ISCIII). JCV acknowledges financial support from Red Temática de Investigación Cooperativa de Cáncer (RTICC).

References

1. Ross PL, Huang YN, Marchese JN et al (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 3:1154–1169
2. Casado-Vela J, Martínez-Esteso MJ, Rodriguez E et al (2010) iTRAQ-based quantitative analysis of protein mixtures with large fold change and dynamic range. *Proteomics* 10:343–347
3. Pierce A, Unwin RD, Evans CA et al (2008) Eight-channel iTRAQ enables comparison of the activity of six leukemogenic tyrosine kinases. *Mol Cell Proteomics* 7:853–863
4. Jorrín JV, Maldonado AM, Castillejo MA (2007) Plant proteome analysis: a 2006 update. *Proteomics* 7:2947–2962
5. Jorrín-Novo JV, Maldonado AM, Echevarría-Zomeño S et al (2009) Plant proteomics update (2007-2008): second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. *J Proteomics* 72:285–314
6. Oeljeklaus S, Meyer HE, Warscheid B (2009) Advancements in plant proteomics using quantitative mass spectrometry. *J Proteomics* 72(3):545–554
7. Zieske LR (2006) A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies. *J Exp Bot* 57: 1501–1508
8. Owiti J, Grossmann J, Gehrig P et al (2011) iTRAQ-based analysis of changes in the cassava root proteome reveals pathways associated with post-harvest physiological deterioration. *Plant J* 67:145–156
9. Martínez-Esteso MJ, Casado-Vela J, Sellés-Marchart S et al (2011) iTRAQ-based profiling of grape berry exocarp proteins during ripening using a parallel mass spectrometric method. *Mol Biosyst* 7:749–765
10. Yang Y, Qiang K, Owsiany C et al (2011) Evaluation of different multidimensional LC-MS/MS pipelines for isobaric tags for relative and absolute quantitation (iTRAQ)-based proteomic analysis of potato tubers in response to cold storage. *J Proteome Res* 10: 4647–4660
11. Reiland S, Grossmann J, Baerenfaller K et al (2011) Integrated proteome and metabolite analysis of the de- etiolation process in plastids from rice (*Oryza sativa* L.). *Proteomics* 11:1751–1763
12. Lippert DN, Ralph SG, Phillips M et al (2009) Quantitative iTRAQ proteome and comparative transcriptome analysis of elicitor-induced Norway spruce (*Picea abies*) cells reveals elements of calcium signaling in the early conifer defense response. *Proteomics* 9:350–367

13. Alvarez S, Hicks LM, Pandey S (2011) ABA-dependent and -independent G-protein signaling in *Arabidopsis* roots revealed through an iTRAQ proteomics approach. *J Proteome Res* 10:3107–3122
14. Zhao Z, Stanley BA, Zhang W et al (2010) ABA-regulated G protein signaling in *Arabidopsis* guard cells: a proteomic perspective. *J Proteome Res* 9:1637–1647
15. Marsh E, Alvarez S, Hicks LM et al (2010) Changes in protein abundance during powdery mildew infection of leaf tissues of Cabernet Sauvignon grapevine (*Vitis vinifera* L.). *Proteomics* 10:2057–2064
16. Fan J, Chen C, Yu Q, Bransky RH, Li C-G, Gmitter FG Jr (2011) Comparative iTRAQ proteome and transcriptome analyses of sweet orange infected by “*Candidatus Liberibacter asiaticus*”. *Physiol Plant* 143:235–245
17. Pang Q, Chen S, Dai S et al (2010) Comparative proteomics of salt tolerance in *Arabidopsis thaliana* and *Thellungiella halophila*. *J Proteome Res* 9:2584–2599
18. Fukao Y, Ferjani A, Tomioka R et al (2011) iTRAQ analysis reveals mechanisms of growth defects due to excess zinc in *Arabidopsis*. *Plant Physiol* 155:1893–1907
19. Jones AME, Bennett MH, Mansfield JW et al (2006) Analysis of the defence phosphoproteome of *Arabidopsis thaliana* using differential mass tagging. *Proteomics* 6:4155–4165
20. Nühse TS, Bottrill AR, Jones AME et al (2007) Quantitative phosphoproteomic analysis of plasma membrane proteins reveals regulatory mechanisms of plant innate immune responses. *Plant J* 51:931–940
21. Jones AME, Nühse TS (2011) Phosphoproteomics using iTRAQ. In: Dissmeyer N, Schnittger A (eds) *Plant kinases*. Springer, New York, NY, pp 287–302
22. Dunkley TPJ, Hester S, Shadforth IP et al (2006) Mapping the *Arabidopsis* organelle proteome. *Proc Natl Acad Sci USA* 103: 6518–6523
23. Ow SY, Salim M, Noirel J, Evans C et al (2009) iTRAQ underestimation in simple and complex mixtures: “The Good, the Bad and the Ugly”. *J Proteome Res* 8:5347–5355
24. Gan CS, Chong PK, Pham TK et al (2007) Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *J Proteome Res* 6:821–827
25. Wu WW, Wang G, Baek SJ et al (2006) Comparative study of three proteomic quantitative methods, DIGE, cLCAT, and iTRAQ, using 2D gel- or LC-MALDI TOF/TOF. *J Proteome Res* 5:651–658
26. Elliott MH, Smith DS, Parker CE et al (2009) Current trends in quantitative proteomics. *J Mass Spectrom* 44:1637–1660
27. Olsen JV, Macek B, Lange O et al (2007) Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods* 4: 709–712
28. Hurkman WJ, Tanaka CK (1986) Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol* 81:802–806
29. Saravanan RS, Rose JK (2004) A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics* 4:2522–2532
30. Lücker J, Laszczak M, Smith D, Lund ST (2009) Generation of a predicted protein database from EST data and application to iTRAQ analyses in grape (*Vitis vinifera* cv. Cabernet Sauvignon) berries at ripening initiation. *BMC Genomics* 10:50
31. Mueller LN, Brusniak M-Y, Mani DR et al (2008) An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *J Proteome Res* 7: 51–61
32. Sellés-Marchart S, Luque I, Casado-Vela J et al (2008) Proteomics of multigenic families from species underrepresented in databases: the case of loquat (*Eriobotrya japonica* Lindl.) polyphenol oxidases. *J Proteome Res* 7: 4095–4106
33. Shevchenko A, Sunyaev S, Loboda A et al (2001) Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass and BLAST homology searching. *Anal Chem* 73: 1917–1926
34. Grossman J, Fisher B, Baerenfaller K et al (2007) A workflow to increase the detection rate of proteins from unsequenced organisms in high-throughput proteomics experiments. *Proteomics* 7:4245–4254
35. Gehlenborg N, O'Donoghue SI, Baliga NS, Goessmann A et al (2010) Visualization of omics data for systems biology. *Nat Methods* 7(3 Suppl):S56–S68

Chapter 13

Global Quantitative Proteomics Using Spectral Counting: An Inexpensive Experimental and Bioinformatics Workflow for Deep Proteome Coverage

Tiago S. Balbuena, Diogo Ribeiro Demartini, and Jay J. Thelen

Abstract

As the field of proteomics shifts from qualitative identification of protein “subfractions” to quantitative comparison of proteins from complex biological samples, it is apparent that the number of approaches for quantitation can be daunting for the result-oriented biologist. There have been many recent reviews on quantitative proteomic approaches, discussing the strengths and limitations of each. Unfortunately, there are few detailed methodological descriptions of any one of these quantitative approaches. Here we present a detailed bioinformatics workflow for one of the simplest, most pervasive quantitative approach—spectral counting. The informatics and statistical workflow detailed here includes newly available freeware, such as SePro and PatternLab which post-process data according to false discovery rate parameters, and statistically model the data to detect differences and trends.

Key words Computational proteomics, GeLC, Label-free proteomics, MudPIT, Quantitative proteomics

1 Introduction

Mass spectrometry-based quantitative techniques have been extensively used throughout the proteomics community for several years [1, 2]. In discovery-driven experiments, these techniques can be generally distinguished on the basis of the presence or the absence of a labeling procedure. Although significant improvements on key analytical steps and on large-scale data analysis have been achieved, protein identification and quantification using mass spectrometry is still a challenging endeavor, especially for complex biological samples [3].

In the past 5 years, label-free quantitation has gained popularity, being the most published approach for quantitative proteomics [4]. Labeling techniques, such as iTRAQ (isobaric tag [5]), isotope-coded affinity tags (ICAT [6]), and stable isotope labeling

Tiago S. Balbuena and Diogo Ribeiro Demartini equally contributed to this work.

by amino acids in cell culture (SILAC [7]), are often considered to be more accurate at quantifying protein abundance [2]. However, the high expense, limitation of samples that can be analyzed in a single experiment, and need for dedicated software are the main drawbacks of labeling techniques, especially when compared to label-free quantitation. In addition, reports of underestimation of fold changes in iTRAQ experiments are not uncommon [4], as well as sequence-dependent O-acylation modification that may compromise accurate quantification [8].

Label-free strategies can be divided into two distinct groups in bottom-up proteomic experiments: extracted ion current and spectral counting. Spectral counting is a “random sampling” quantitative method based on premise that high-abundance peptides will likely be more selected for fragmentation and thus result in higher number of MS/MS acquired spectra [9, 10]. In other words, the frequency of these MS/MS scans reflects the abundance of this peptide in a sample (Fig. 1). Once the data are searched against a particular database, the protein abundance will be estimated based on the sum of all MS/MS spectra matching to every peptide derived from the protein query (i.e., the number of peptide-spectrum matches (PSMs)). Protein ratios determined by spectral counting correlate well with other label-free strategies [11, 12]. However, this approach has showed to be weak for low-abundance proteins, when few spectral counts are acquired for a particular protein [9, 13] (Table 1).

Although several concerns have been raised, the main limitations of the spectral counting strategy are the “chimera” MS/MS spectra and the protein inference problem (i.e., assembling peptide sequences into proteins). Chimera spectra hamper correct peptide assignment. They are frequently observed when co-eluted peptides present similar m/z , being consequently isolated and fragmented simultaneously generating a merged fragment spectrum. The presence of these MS/MS spectra significantly affects protein identification and may contribute to the usually low number of spectra that are successfully matched to a peptide sequence [14]. Chimera spectra comprise a significant number of MS/MS spectra acquired in modern mass spectrometers. Hoopmann and coworkers estimated that 11 % of all MS/MS spectra acquired in a high-resolution mass spectrometer are chimera spectra [15]. Recently, Houel et al. showed that chimera spectra acquired in an LTQ-OrbitrapTM mass spectrometer may reach a fraction equal to 50 % in complex samples [14].

One way to reduce the peptide co-elution problem is to reduce sample complexity. Multidimensional protein identification technology, MudPIT, was presented by Washburn and coworkers, reporting the identification of 1,484 proteins from *Saccharomyces cerevisiae* [23]. In MudPIT, fractionation is performed with the proteins already digested from a complex sample. Peptides are loaded in a strong cation exchange (SCX) chromatography column at once.

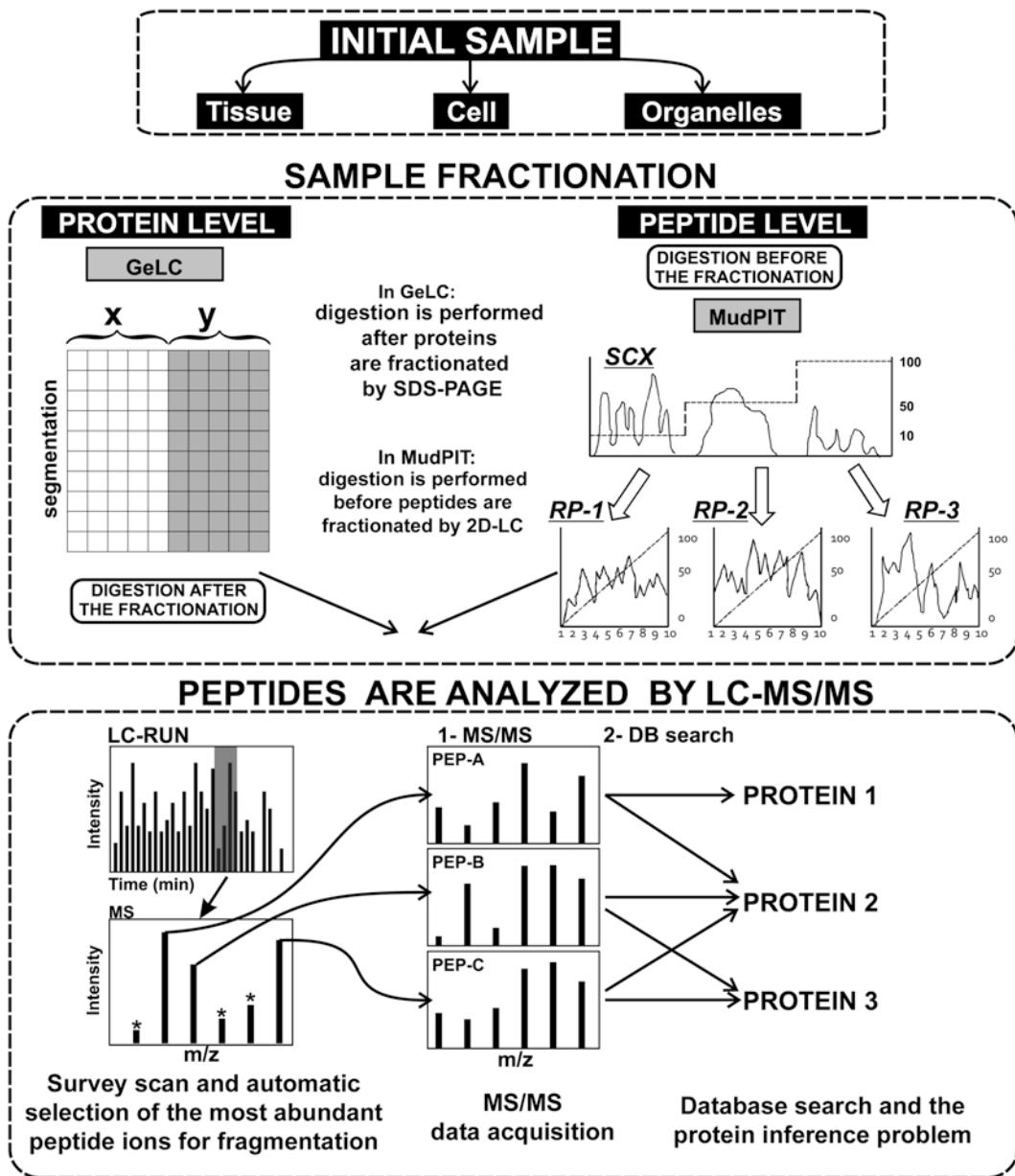


Fig. 1 Summarized workflow for a label-free approach. Initial sample is obtained and fractionated by GeLC (SDS-PAGE represented by the white/gray sections, being *x* and *y* hypothetical samples) or MudPIT (peptides eluted from each salt step on the SCX chromatography are loaded in an RP column prior to injection in the mass spectrometer). In the data-dependent acquisition approach the most intense peptide ions are selected and fragmented and acquired MS/MS spectra are searched against a user-defined database

Elution from SCX column is performed in several steps, increasing the salt concentration. Peptides are then eluted from each step and loaded in an *in line* reverse-phase (RP) column prior to injection in the mass spectrometer (Fig. 1). Another approach largely used in

Table 1
Points to be considered when choosing label-free spectral counting approach

| Label-free spectral counting | | References |
|------------------------------|---|---------------------|
| Strong points | Under controlled conditions, spectral count is highly reproducible between technical replicates | [11] |
| | Faster and cheaper when compared to labeling methods | [1, 16–18] |
| | No limits for replicates; ideal for low amounts of samples | [3] |
| | Straightforward method | [19] |
| Weak points | Inferior accuracy when compared to labeling techniques, especially due to suppression effect | [1] |
| | Data normalization is still challenging | [18] |
| | Low discrimination for peptides shared by multiple proteins | [20] |
| | Low-abundance proteins might not be detected, especially due to co-eluted peptides' effect | [1, 13, 21, 22, 29] |

label-free experiments is prefractionation, prior to tryptic digestion, by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). After electrophoresis, gels are stained and the lanes excised into several smaller segments (the number dependent upon the original amount of protein loaded). Each of these segments is then in-gel digested with trypsin and tryptic peptides contained in each digested segment are analyzed by reversed-phase LC-MS/MS (Fig. 1). The approach is known as GeLC-MS/MS, and it has been extensively used for label-free spectral counting approaches.

Besides prefractionation, an MS-based strategy extensively used to detect low-abundance peptides, or co-eluted peptides, is dynamic exclusion. Dynamic exclusion refers to a particular period of time where the mass spectrometer ignores the n most abundant ions, previously fragmented, and will select the next most abundant ions for fragmentation. So, if the mass spectrometer is set to isolate and fragment the three most abundant ions, it will select the 4th, 5th, and 6th most abundant ions for subsequent fragmentation in the following survey scan [2, 9, 24]. Zhang and coworkers mention that 90 s is an optimal value and that the optimal dynamic exclusion time is proportional to the average chromatographic peak width at the base of the eluting peptides [9]. An alternative method to overcome the co-elution issue is to change the mass spectrometer operational mode from data-dependent (DDA) to data-independent analyses (DIA). This approach was originally suggested by Venable and Dong [25], and does not rely on parent mass scanning but on a continuous MS/MS scan acquisition in narrow m/z ranges until a larger m/z window is covered.

Independent of the quantitative method adopted (labeling or label-free) the protein inference problem is an issue that deserves attention. It is a consequence of the existence of shared, degenerate, or redundant peptides (i.e., when the sequence of an identified peptide is present in more than one protein entry) [26]. When protein identification is performed using “unique peptides only”, it is clear which peptides map to which protein. However, some proteins have high identity at the primary sequence level, turning data analysis into a complex and challenging task [1]. After detection of the shared peptides, they may be ignored and the MS/MS spectra associated to them will be counted multiple times (1), excluded from data analysis which may result in underrepresentation of the quantitative value observed for the proteins sharing peptides (2), or distributed across the protein entries (3).

2 Materials

Noncommercial software and tools with user-friendly interface (Fig. 2):

1. Mspire-sequest (<http://rubydoc.info/gems/ms-sequest/0.2.1/frames>) [27].
2. Search Engine Processor (<http://proteomics.fiocruz.br/Softwares.aspx>) [28].
3. Regrouper (<http://proteomics.fiocruz.br/Softwares.aspx>).
4. PatternLab (<http://proteomics.fiocruz.br/Softwares.aspx>) [29].
5. MUSite (<http://musite.sourceforge.net/>) [30].
6. Blast2GO (<http://www.blast2go.com/b2glaunch>) [31].
7. Permutmatrix (<http://www.lirmm.fr/~caraux/PermutMatrix/>) [32].

3 Methods

3.1 Generation of SQT Files

The executable that reads and supports the conversion from Bioworks Sequest Results File (SRF) to SQT files was developed by Prof. John T. Prince in Ruby programming language. The converter will work on any platform. The original publication [27] and complete documentation of this program can be found at <http://rubydoc.info/gems/mspire-sequest/file/README.rdoc>.

1. Download Ruby Installer for Windows from <http://rubyinstaller.org/downloads/>.

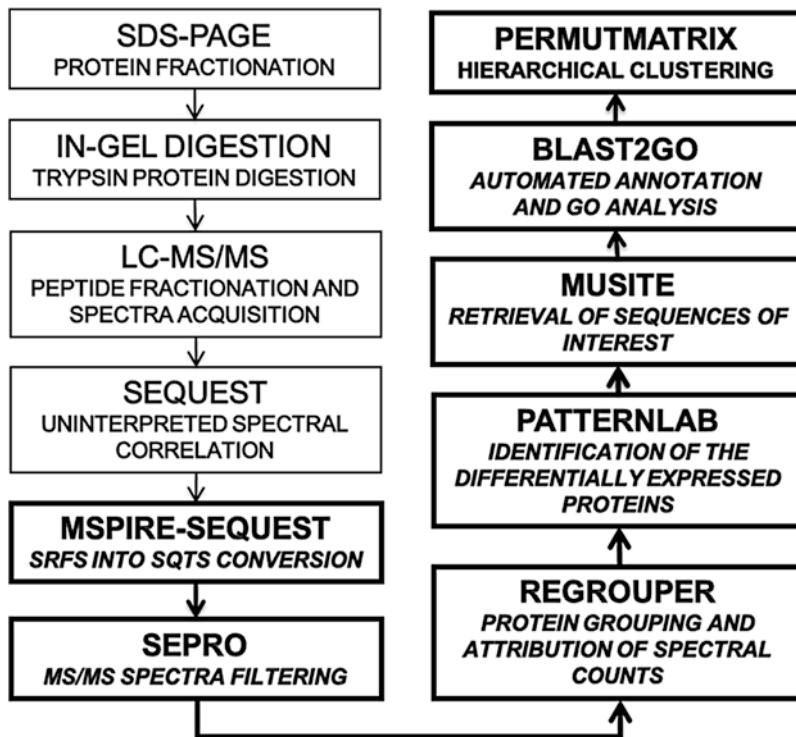


Fig. 2 Pipeline based on freely available tools for spectral counting proteomics. In the present figure we suggest the GeLC approach as a prefractionation strategy. The bioinformatics steps described herein are bold highlighted and span from SEQUEST result file conversion to cluster analysis

2. Install Ruby latest version and be sure to check the options “Add Ruby executables to your PATH” and “Associate .rb and .rbw file with this Ruby installation”.
3. Install mspire-sequest by starting command prompt and running “gem install mspire-sequest”.
4. For conversions, run the executable “srf_to_sqt.rb”. The converter will work on any number of SRF files as input. A simple way to convert a batch of files is to copy-paste all SRF files in a directory and run the executable “srf_to_sqt.rb directoryname*.srf”.

3.2 Peptide Spectrum Match Filtering

Search Processor Engine (SEPro) is a tool developed to filter PSMs proposed by the SEQUEST search engine and improve the number of peptide and protein identifications based on a sophisticated algorithm. For details on the calculation of the Bayesian score and how SEPro uses a three-tier approach to optimize the number of identifications, please refer to Carvalho and coworkers [28].

1. Group all SQTs from each biological replicate into a single directory. Spectrum processing is carried out per replicate (i.e., per biological sample).

2. In SEPro's control panel, indicate the location of the directory containing the SQTs from one biological replicate and the FASTA format database used in SEQUEST searches (*see Note 1*).
3. In SEPro's parameters, select the features that you want to be used for calculation of the Bayesian score. We strongly advise to use all parameters for data processing. If data was acquired in a low-resolution mass spectrometer, uncheck "Delta Mass".
4. Preprocessing filters will be used to speed up the process, leaving fewer PSMs for processing. We recommend removing peptides with a plus one charge state, PSMs with primary score lower than 1, and consider only peptides longer than six amino acids. If MS data was acquired with a high-resolution mass spectrometer, "Delta mass" option equal to 10 ppm is advisable.
5. In the "Acceptable False Discovery Rate" box, type percentage values that obey the logic of the three-tier approach in which a large number of spectra correlates with a lower number of peptides that in turn map to fewer proteins (*see Notes 2 and 3*).
6. Post-processing filters are used after the three-tier filtering is applied. Although being a strong-quality filter, for general spectral counting projects we recommend to consider only proteins having a minimum of two spectral counts and two peptides (*see Note 4*). As mentioned before, "DeltaMass" feature should only be used if high-resolution instrument was used for MS data acquisition.
7. The "Similar Proteins" panel should enable the elimination of internal proteins.
8. In the "General Panel" choose the enzyme used for *in vivo* and *in silico* protein digestion. Check the radio buttons for grouping PSMs by charge state and number of enzymatic termini (*see Note 5*).
9. For classical forward-decoy databases, indicate the tag used to identify decoy protein sequences in the FASTA database. This tag should be typed in the "Labeled Decoy Tag" box (*see Note 6*).
10. After all parameters are adjusted, we recommend saving the parameters as default for future processing.
11. To start the filtering process, go back to "Control" panel and click "GO!". A real-time follow-up can be achieved by clicking on "Followup" panel.
12. Upon completion, go to "ResultBrowser" panel and check the number of retained spectra, peptides, and protein. Also, keep track of the post-processing spectra, peptide, and protein FDR for future comparisons within biological replicates.
13. Save SEPro result.

3.3 Spectral Count Assessment of Protein Groups

ReGrouper is a key tool that connects SEPro and PatternLab software suite. Parameters for the generation of PatternLab’s input (Sparse Matrix and Index File) must be carefully designed in order to facilitate further biological interpretation. Although it allows the generation of quantitative data from multiple formats, we will focus on the analysis of the spectral counts per protein groups in order to adhere to principles of parsimony (i.e., the minimum number of protein sequences that describes a set of peptides).

1. Copy and paste all SEPro files from each treatment (i.e., if three biological replicates, three SEPro files) into different directories. Name the directories with a number (we use 0 for control and 1 for treatment, for instance).
2. In Regrouper’s input panel, click “Add directory” and indicate the location of the directory containing the SEPro files to be processed. If desired, type the descriptions for each directory.
3. In “Processing parameters”, eliminate internal and decoy sequences and click “Load”.
4. In “Spectral Counting Analysis”, select the following “Groups” format: “Spectral Counts” and “Undetermined”. Then, press “GO!”. With this, an Index file and a Sparse Matrix will be created based on the distribution of spectral counts per protein groups that share the same peptide set independently of the protein group type.
5. To verify the content and the quantitative values (i.e., spectral counts) attributed for each protein group, select “generate bird’s eye view” and save the file.

3.4 Detection of Differentially Expressed Proteins

PatternLab for Proteomics is a multi-tool proteomic statistical analysis software [29]. Although other modules are available, we will focus on the detection of the differentially expressed proteins based on the T-Fold test, which integrates Student’s *t*-test and fold change values for the assessment of quantitative differences through pairwise analysis (*see Note 7*).

1. Open PatternLab for proteomics, go to “Analyze” panel, and select “T-Fold”.
2. Load the Sparse Matrix and the Index file generated previously by Regrouper.
3. Select normalization methods (*see Note 8*).
4. Select the minimal number of replicates in which a protein was detected to be considered in the analysis. This value depends on the number of replicates used in the experiment. For conservative approaches type the number of replicates per treatment. For instance, if five biological replicates were used, then one should have five SEPro files and, for a conservative approach, should select five as minimum number of replicas per

class. We prefer to relax this parameter and usually select four as the minimum number of replicates per class (i.e., the protein was identified in four out of the five biological replicates).

5. Select a *q*-value and press the “Optimize F” button. This will make the TFold adjust a variable fold-change cutoff to maximize the identifications for a given *q*-value. The user can also highlight (and separate) proteins that pass the *q*-value and fold-change cutoff by increasing the parameter found in the L-stringency numeric box. We recommend using a value of 0.4.
6. After the parsing step is done, press “Calculate and Plot” to generate a volcano plot on the right side of PatternLab’s display. Proteins that met the statistical and fold-change cutoffs as well as the L-stringency parameter are displayed as blue dots. The ones that satisfy both criteria but do not meet the L-stringency parameter will be highlighted (and separately reported) in orange. These proteins deserve further experimentation to assess their differential expression status. The green dots are those that satisfy the fold change but not the *q*-value statistical filter; and finally, the red dots are proteins that do not pass any criteria.
7. Export results by choosing “Output” panel and selecting “Save Plot” for all proteins (blue, red, orange, and green dots).

3.5 Retrieval of Protein Sequences

In order to gather more information about the proteins suggested as differentially expressed, the protein primary sequence of these identifications is extracted using the Musite application. This software contains several useful features and was originally developed for prediction of phosphorylation sites. For detailed information, please refer to Gao and coworkers [30].

1. Prepare a one-column file containing the accessions suggested by the T-Fold test as differentially expressed (see Note 9).
2. In Musite control panel, convert the FASTA database used by SEQUEST search engine to Musite XML format. For that, select “Tools” and then “File Processing”. For database conversion, select “File Conversion” and “Convert FASTA to Musite XML”. Finally, indicate the file to be converted and the regular expression to parse the FASTA database header.
3. After Musite XML file is created, select “Tools” and then “File Processing”. For filtering the proteins of interest, select “File Filtering” and “Filter Proteins by Accessions”.
4. Indicate the converted Musite XML database and paste the list of proteins of interest in the “Accessions” section.
5. After filtering is done, the retrieved sequences should be converted back to FASTA file by following step 2, but converting Musite XML file to FASTA.

3.6 Annotation and Functional Analysis

BLAST2GO is a Java interface tool with multiple features for annotation and analysis of both nucleotide and protein sequences. A detailed description on the capabilities and uses of the tool can be found in the original publication [31]. Here, we focus on the steps for sequence alignment and functional annotation.

1. Load the FASTA file containing the protein sequences.
2. Run alignment step by selecting “Blast” in the main control panel. Then, select the desired database (nr or SwissProt), number of Blast hits to be reported per sequence, minimum Blast expected value, Blastp as the algorithm, and minimum HSP length. We recommend using a minimum expect value equal to $1e^{-10}$ and minimum HSP cutoff equal to 33 and check the “Low complexity filter” option.
3. After alignment step is done and all sequences received a description, run mapping step to obtain gene ontology information for all the sequences and start the annotation step.
4. To create a pie chart with the distribution of the mapped GO terms for the dataset, select “Analysis” and “Make a combined graph” option (*see Notes 10 and 11*).

3.7 Clustering Analysis

A common strategy applied to analyze large-scale data is to cluster proteins accordingly to their expression profile. We use a freely available program called PermutMatrix [32] to carry out the clustering analysis.

1. In a Microsoft Excel worksheet, input raw spectral count values such that the columns represent samples and rows represent proteins.
2. Replace the missing values (i.e., spectral counts equal to zero) by the mean of the biological replicates (*see Note 12*).
3. Calculate the mean of the spectral count across the biological replicates and normalize the data by dividing each biological replicate mean by the mean of all values for a protein. This ratio-based value should then be submitted to the base 2 logarithmic transformation.
4. Prepare a standard tab-delimited text file containing the transformed data matrix.
5. After loading the file in PermutMatrix, select Pearson distance for the calculation of the dissimilarities. Then, select the average linkage (UPGMA) method for aggregation procedure.
6. After the result is generated, change the display options for a better view of the data. If necessary, try a different clustering method.
7. Export the tree by selecting “Export to Clipboard” and save as a JPEG or Bitmap file.

4 Notes

1. For a complete visualization of acquired spectra, MS2 files must be located in the same directory of the SQT files and the processing option “Include MS2 in Results” must be checked. Processing will take longer, but all filtered spectra and corresponding ion series will be available. To convert RAW into MS2 files, use John Yates’ RawXtract program available at <http://fields.scripps.edu/downloads.php>.
2. For optimum results, we recommend fixing protein FDR to 1 % and select different configurations of spectra and peptide FDRs. It should be noted that user-defined FDRs are not the post-processed FDRs, which are lower than the former due to the three-tier filtering approach.
3. Decreasing spectra and peptide FDR stringency usually results in increasing number of proteins. However, overestimation of the user-defined spectra and peptide FDRs results in the acceptance of a high number of low-scoring PSMs in the dataset. Consequently, for achieving the 1 % protein FDR a more stringent cutoff needs to be adjusted, resulting in a decrease in the number of identified proteins.
4. Alternatively, less stringent quality filters may be applied here in order to maximize the number of identified proteins and a more stringent configuration may be applied later in Regrouper before quantitative analysis is carried out.
5. Grouping PSMs by charge state and number of enzymatic termini allows peptides to be processed separately. If an insufficient number of PSMs is achieved an error message will appear. In this situation, we recommend reducing the number of groups by unchecking one of the grouping procedures.
6. SEPro allows the use of the semi-labeled decoy approach as previously described [33].
7. Briefly, given a user-specified FDR bound, the TFold approach uses a theoretical FDR estimator [34] to maximize the number of identifications that satisfy both a fold-change cutoff, which varies with the *t*-test *p*-value as a power law, and a stringency criterion that aims to detect low-abundance proteins. By following these steps, the TFold capitalizes on two limitations commonly found in competing algorithms. A fixed fold-change cutoff could discard proteins with very low *p*-values but not satisfying the fold-change cutoff. Secondly, low-abundance proteins are more prone to “fooling” common statistical filters because they tend to artificially acquire low *p*-values and ultimately consume the statistical power of the theoretical FDR estimator.

8. From all available methods, we recommend the use of the Row Sigma normalization to deal with samples presenting high spectral counting dynamic range. For a detailed description on the method, please refer to Carvalho and coworkers [29].
9. This can be easily achieved by pasting the saved flexible report in a blank excel file and breaking the first column (use “comma” signal for that) to isolate the first described protein from the others contained in the same protein group. By default, the first described protein in a protein group is the longest one and, thus, has the highest probability of retrieving functional annotations during the sequence alignment step.
10. Selecting the level of depth is an important parameter for the creation of the pie charts. The deeper the analysis, the more detailed information is achieved but less coverage is obtained.
11. As a complement, Fischer’s Exact Test can be applied to the analysis if a background/reference dataset is provided.
12. We recommend replacing the zero values by the mean value if a protein attends to the same parameters used in PatternLab (i.e., protein found in n replicates -1). Otherwise, protein should not be used in the analysis.

Acknowledgements

The authors would like to thank Dr. Paulo C. Carvalho for his critical suggestions to this chapter.

References

1. Elliott MH, Smith DS, Parker CE et al (2009) Current trends in quantitative proteomics. *J Mass Spectrom* 44:1637–1660
2. Neilson KA, Ali NA, Muralidharan S et al (2011) Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* 11:535–553
3. Nanjo Y, Nouri MZ, Komatsu S (2011) Quantitative proteomic analyses of crop seedlings subjected to stress conditions; a commentary. *Phytochemistry* 72:1263–1272
4. Evans C, Noirel J, Ow S et al (2012) An insight into iTRAQ: where do we stand now? *Anal Bioanal Chem* 404:1011–1027
5. Ross PL, Huang YN, Marchese JN et al (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 3:1154–1169
6. Gygi SP, Rist B, Gerber SA et al (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 17:994–999
7. Ong SE, 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
8. Wiktorowicz JE, English RD, Wu Z et al (2012) Model studies on iTRAQ modification of peptides: sequence-dependent reaction specificity. *J Proteome Res* 11:1512–1520
9. Zhang Y, Wen Z, Washburn MP et al (2009) Effect of dynamic exclusion duration on spectral count based quantitative proteomics. *Anal Chem* 81:6317–6326
10. Thelen JJ, Peck SC (2007) Quantitative proteomics in plants: choices in abundance. *Plant Cell* 19:3339–3346
11. Zhang B, VerBerkmoes NC, Langston MA et al (2006) Detecting differential and correlated

- protein expression in label-free shotgun proteomics. *J Proteome Res* 5:2909–2918
12. Old WM, Meyer-Arendt K, Aveline-Wolf L et al (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics* 4:1487–1502
 13. Clough T, Key M, Ott I et al (2009) Protein quantification in label-free LC-MS experiments. *J Proteome Res* 8:5275–5284
 14. Houel S, Abernathy R, Renganathan K et al (2010) Quantifying the impact of chimera MS/MS spectra on peptide identification in large-scale proteomics studies. *J Proteome Res* 9:4152–4160
 15. Hoopmann MR, Finney GL, MacCoss MJ (2007) High-speed data reduction, feature detection, and MS/MS spectrum quality assessment of shotgun proteomics data sets using high-resolution mass spectrometry. *Anal Chem* 79:5620–5632
 16. Podwojski K, Eisenacher M, Kohl M et al (2010) Peek a peak: a glance at statistics for quantitative label-free proteomics. *Expert Rev Proteomics* 7:249–261
 17. Jin S, Daly DS, Springer DL et al (2008) The effects of shared peptides on protein quantitation in label-free proteomics by LC/MS/MS. *J Proteome Res* 7:164–169
 18. Gokce E, Shuford CM, Franck WL et al (2011) Evaluation of normalization methods on GeLC-MS/MS label-free spectral counting data to correct for variation during proteomic workflows. *J Am Soc Mass Spectrom* 22:2199–2208
 19. Friso G, Olinares PDB, Van Wijk KJ (2011) The workflow for quantitative proteome analysis of chloroplast development and differentiation, chloroplast mutants, and protein interactions by spectral counting. *Methods Mol Biol* 775:265–282
 20. Nesvizhskii AI, Aebersold R (2005) Interpretation of shotgun proteomic data: the protein inference problem. *Mol Cell Proteomics* 4:1419–1440
 21. Zhu W, Smith JW, Huang CM (2010) Mass spectrometry-based label-free quantitative proteomics. *J Biomed Biotechnol* 2010:1–6
 22. Zhang Y, Wen Z, Washburn MP et al (2010) Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal Chem* 82:2272–2281
 23. Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19:242–247
 24. Lai X, Wang L, Tang H et al (2011) A novel alignment method and multiple filters for exclusion of unqualified peptides to enhance label-free quantification using peptide intensity in LC-MS/MS. *J Proteome Res* 10:4799–4812
 25. Venable JD, Dong MQ, Wohlschlegel J et al (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat Methods* 1:39–45
 26. Nesvizhskii AI, Keller A, Kolker E et al (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* 75: 4646–4658
 27. Prince JT, Marcotte EM (2008) MsPire: mass spectrometry proteomics in Ruby. *Bioinformatics* 24:2796–2797
 28. Carvalho PC, Fischer JSG, Xu T et al (2012) Search engine processor: filtering and organizing PSMs. *Proteomics* 12:944–949
 29. Carvalho PC, Fischer JSG, Chen EI et al (2008) PatternLab for proteomics: a tool for differential shotgun proteomics. *BMC Bioinformatics* 9:316
 30. Gao J, Thelen JJ, Dunker AK et al (2010) Musite, a tool for global prediction of general and kinase-specific phosphorylation sites. *Mol Cell Proteomics* 9:2586–2600
 31. Conesa A, Götz S, García-Gómez JM et al (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676
 32. Caraux G, Pinloche S (2005) PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order. *Bioinformatics* 21:1280–1281
 33. Barboza R, Cociorva D, Xu T et al (2011) Can the false-discovery rate be misleading? *Proteomics* 11:4105–4108
 34. Benjamini Y, Drai D, Elmer G et al (2001) Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 125: 279–284

Chapter 14

Combining Chymotrypsin/Trypsin Digestion to Identify Hydrophobic Proteins from Oil Bodies

Martina Vermachova, Zita Purkrtova, Jiri Santrucek, Pascale Jolivet, Thierry Chardot, and Milan Kodicek

Abstract

Oil bodies, lipid-storage organelles, are stabilized by a number of specific proteins. These proteins are very hydrophobic, which complicates their identification by “classical” proteomic protocols using trypsin digestion. Due to the lack of trypsin cleavage sites, the achievable protein coverage is limited or even insufficient for reliable protein identification. To identify such proteins and to enhance their coverage, we introduced a modified method comprising standard three-step procedure (SDS-PAGE, in-gel digestion, and LC-MS/MS analysis). In this method, chymotrypsin, single or in combination with trypsin, was used, which enabled to obtain proteolytic peptides from the hydrophobic regions and to identify new oil bodies’ proteins. Our method can be easily applied to identification of other hydrophobic proteins.

Key words Chymotrypsin, Hydrophobic proteins, In-gel digestion, LC-MS/MS, Oil bodies

Abbreviations

| | |
|-------|--|
| AA | Acrylamide |
| ACN | Acetonitrile |
| APS | Ammonium persulfate |
| BIS | N,N'-methylenebisacrylamide |
| CBB | Coomassie Brilliant Blue |
| DTT | Dithiothreitol |
| FA | Formic acid |
| IAA | Iodoacetamide |
| IB | Isolation buffer |
| OBs | Oil bodies |
| PLB | Protein loading buffer |
| SDS | Sodium dodecyl sulfate |
| TEMED | <i>N,N,N',N'</i> -tetramethylethylenediamine |

1 Introduction

In comparison with other plant organelles, oil bodies (OBs) possess an unusual structure: they are composed of a neutral lipid core surrounded by a “half-unit” membrane formed by a phospholipid monolayer with several embedded proteins [1]. The OB proteins stabilize the particles and are expected to participate in OB formation and degradation, but their precise role in these processes is still unclear [2, 3]. Recent reports claim for enzymatic activity of oleosin [4]. The OB proteins are present in a number of homologous isoforms, whose significance has yet to be determined [5]. The proteins are generally very hydrophobic, which complicates matters. This is the main reason why their detailed structure and particular function still have not been described sufficiently. Even some of the protein isoforms have been escaping the identification for decades [6].

Hydrophobic proteins comprise long regions without cleavage sites for trypsin (arginine and lysine), which decreases the number of peptides produced by trypsin digestion and the protein coverage. The extensive protein coverage is a first demand of current challenges in proteomics such as structural analysis and identification of posttranslational modifications.

To overcome the problems connected with hydrophobic protein digestion, we brought chymotrypsin into the play [6]. This protease still has not been appreciated enough, mainly due to its low specificity compared to trypsin. Nevertheless, with the new generation of MS instruments, providing peptide fragmentation, high resolution, and accuracy, the chymotryptic digestion products can be easily “puzzled out” and chymotrypsin becomes a powerful tool for hydrophobic protein identification. We isolated the OBs by the method adapted from Tzen [7], but the purified OBs were dissolved directly in the electrophoresis sample buffer in order to increase the protein concentration in loaded samples. After the SDS-PAGE, the protein bands were excised and in-gel digested. Three different proteolytic systems were used: trypsin, as in classical proteomic protocols, chymotrypsin, and the mixture of both enzymes (1:1, w/w). Total protease concentration was 12.5 mg/L in all cases.

Using chymotrypsin, single or in mixture with trypsin, the coverage of six well-known OB proteins was enhanced from 7.9 to 34.2 [8] up to 80 %; and moreover, four new proteins were identified. Chymotrypsin, digesting in the hydrophobic protein parts, contributed significantly to the protein coverage. The protease mixture digestion led to a number of peptides distributed over the protein sequence and, in some cases, revealed peptides unavailable with single protease digestion.

Our method can be easily applied to extensive identification of other hydrophobic proteins.

2 Materials

Prepare all solutions using ultrapure water and analytical grade reagents. Store all the solutions as indicated. When disposing waste materials, follow waste disposal regulations.

2.1 Plant Grains and Buffers for OB Isolation

Prepare fresh solutions for each isolation. Keep all the buffers at 4 °C.

1. The grains of *Arabidopsis thaliana*, ecotype Columbia, should be stored in a dark, cold (4 °C), and dry place (see Note 1).
2. Phosphate stock solution: 80 mM sodium phosphate, pH 7.5. Weigh 1.25 g of monosodium phosphate dihydrate into a glass beaker and dissolve in 80 mL of water. Mix and adjust pH with NaOH (see Note 2). Make up to 100 mL with water in a graduated cylinder.
3. Isolation buffer 1 (IB1): 0.6 M sucrose, 10 mM sodium phosphate, pH 7.5. Weigh 16.43 g of sucrose, add 10 mL of phosphate stock solution, and make up to 80 mL with water.
4. Isolation buffer 2 (IB2): 0.4 M sucrose, 10 mM sodium phosphate, pH 7.5. Weigh 10.95 g of sucrose, add 10 mL of phosphate stock solution, and make up to 80 mL with water.
5. Isolation buffer 3 (IB3): 0.2 M sucrose, 0.1 % Tween 20 (v/v), 5 mM sodium phosphate, pH 7.5. Weigh 2.74 g of sucrose, add 2.5 mL of phosphate stock solution and 400 µL of 10 % Tween 20 (v/v, see Note 3). Make up to 40 mL with water.
6. Isolation buffer 4 (IB4): 10 mM sodium phosphate, pH 7.5. Dilute 10 mL of phosphate stock solution up to 80 mL with water.
7. Isolation buffer 5 (IB5): 0.6 M sucrose, 2 M sodium chloride, 10 mM sodium phosphate, pH 7.5. Weigh 8.22 g of sucrose and 4.68 g of sodium chloride, add 5 mL of phosphate stock solution, and make up to 40 mL with water.
8. Isolation buffer 6 (IB6): 0.25 M sucrose, 2 M sodium chloride, 10 mM sodium phosphate, pH 7.5. Weigh 3.42 g of sucrose and 4.68 g of sodium chloride, add 5 mL of phosphate stock solution, and make up to 40 mL with water.
9. Urea: 9 M solution in water. Weigh 10.81 g of urea and dissolve in water to reach the final volume of 20 mL (see Note 4).
10. Hexane.
11. Storage buffer: 50 mM phosphate buffer, pH 7.5. Mix 5 mL of phosphate stock solution and 3 mL of water.
1. Resolving gel buffer: 1.5 M Tris–HCl, pH 8.8. Weigh 18.2 g of Tris–HCl into a glass beaker and dissolve in about 80 mL of

2.2 Solutions for SDS-PAGE

water. Adjust pH with HCl (*see Note 2*). Make up to 100 mL with water in a graduated cylinder. Store at 4 °C.

2. Stacking gel buffer: 1 M Tris–HCl, pH 6.8. Weigh 12.1 g of Tris–HCl and dissolve in about 80 mL of water. Adjust pH with HCl (*see Note 2*). Make up to 100 mL with water. Store at 4 °C.
3. 30 % (w/v) acrylamide/bisacrylamide mixture (30 % AA/BIS): 29% (w/v) acrylamide, 1% (w/v) *N,N'*-methylenebisacrylamide. Weigh 29.0 g of acrylamide and 1.0 g of bisacrylamide into a glass beaker (*see Note 5*) and dissolve in about 80 mL of water. Make up to 100 mL with water in a graduated 100 mL cylinder. Store at 4 °C (*see Note 6*).
4. Sodium dodecyl sulfate (SDS) solution: 10 % (w/v) SDS in water. Dissolve 20 mg of SDS in 200 µL of water. Prepare fresh solution each time.
5. Ammonium persulfate (APS): 10 % (w/v) APS in water. Dissolve 20 mg of APS in 200 µL of water. Prepare fresh solution each time.
6. *N,N,N',N'*-tetramethylethylenediamine (TEMED).
7. SDS-PAGE running buffer, 10x concentrated: 0.25 M Tris–HCl, 1.92 M glycine, 1 % (w/v) SDS. Weigh 15.1 g of Tris–HCl, 72.1 g of glycine, and 5 g of SDS and make up to 500 mL with water (*see Note 7*). Do not adjust pH. Store at 4 °C. Before use, dilute ten times.
8. Protein loading buffer (PLB): 0.1 M Tris–HCl, pH 6.8, 1.2 % (w/v) SDS, 30 % (v/v) glycerol, 0.15 M dithiothreitol (DTT), 0.025 % bromophenol blue: Weigh 0.12 g of SDS, 0.23 g of DTT, and 2.5 mg of bromophenol blue. Add 1 mL of stacking gel buffer and 3 mL of glycerol (to measure viscous glycerol proceed similarly as with Tween 20 in **Note 3**). Make up to 10 mL with water. Divide the solution in ten aliquots and store at -20 °C.
9. Protein standards for SDS-PAGE.
10. Coomassie Brilliant Blue (CBB) G-250 solution: 0.4 % CBB G-250 in 96 % ethanol. Weigh 0.4 g of CBB G-250 and dissolve in 100 mL of 96 % ethanol [9]. Store at room temperature.
11. Ammonium sulfate solution: 12.5 % ammonium sulfate in 2.5 % phosphoric acid. Weigh 62.5 g of ammonium sulfate, dissolve in around 400 mL of water, add 15 mL of 85 % phosphoric acid, and make up to 500 mL with water [9]. Store at room temperature.
12. CBB staining solution: 0.08 % CBB G-250, 19 % ethanol, 10 % ammonium sulfate, 2 % phosphoric acid. Mix 20 mL of CBB G-250 solution and 80 mL of ammonium sulfate solution [9]. Mix immediately before use.

2.3 Solutions for In-Gel Digestion

Prepare fresh solutions for each in-gel digestion.

1. Washing buffer: 0.1 M ammonium bicarbonate. Weigh 119 mg of ammonium bicarbonate into a small glass beaker and dissolve in 15 mL of water.
2. Washing buffer/acetonitrile (ACN): 0.1 M ammonium bicarbonate/ACN (1:1) (v/v). Mix equal volumes of washing buffer and ACN in a glass test tube. For in-gel digestion of 20 samples, 10 mL of washing buffer/ACN is needed (*see Note 8*).
3. ACN (*see Note 8*).
4. 50 mM ammonium bicarbonate. Dilute the washing buffer twice. For in-gel digestion of 20 samples, 4 mL of 50 mM ammonium bicarbonate is needed. To be cold enough for further use, keep the solution at 4 °C.
5. DTT solution: 10 mM DTT, 0.1 M ammonium bicarbonate. Weigh 4.8 mg of DTT into a test tube and add 3 mL of washing buffer.
6. Iodoacetamide (IAA) solution: 55 mM IAA, 0.1 M ammonium bicarbonate. Weigh 30 mg of IAA into a test tube and add 3 mL of washing buffer. Keep the solution in a dark place.
7. Trypsin solution: Trypsin 12.5 mg/L in 50 mM ammonium bicarbonate. Dilute the stock solution of trypsin (1 g/L) 80 times with cold (4 °C) 50 mM ammonium bicarbonate to reach the final concentration of trypsin 12.5 mg/L (*see Note 9*). For one sample, according to the amount of gel pieces, 30–60 µL of trypsin solution is needed. Prepare the solution immediately before use. Keep it at 4 °C.
8. Chymotrypsin solution: Chymotrypsin 12.5 mg/L in 50 mM ammonium bicarbonate. Prepare the solution of chymotrypsin as in previous step (*see Note 9*). Keep the solution at 4 °C.
9. Protease mixture solution: Trypsin 6.25 mg/L, chymotrypsin 6.25 mg/L in 50 mM ammonium bicarbonate. Mix equal volumes of trypsin and chymotrypsin solutions in 50 mM ammonium bicarbonate. Keep the solution at 4 °C.
10. Extraction solution 1: 35 % (v/v) ACN, 5 % (v/v) formic acid (FA). In a glass test tube, mix 3.0 mL of water, 1.75 mL of ACN (*see Note 8*), and 250 µL of FA.
11. Extraction solution 2: 70 % (v/v) ACN, 5 % (v/v) FA. In a glass test tube, mix 1.25 µL of water, 3.50 mL of ACN (*see Note 8*), and 250 µL of 5 % FA.
12. ZipTip pipette tips with C18 reversed-phase. We purchase ZipTip pipette tips from Millipore.
13. ZipTip wash solution: 50 % ACN in water. Mix 2 mL of water and 2 mL of ACN in a glass test tube (*see Note 8*). Prepare immediately before use and sonicate in a water bath for 5 min to remove air bubbles.

14. ZipTip equilibration solution: 1 % FA in water. Mix 40 μ L of FA and 3,960 μ L of water in a glass test tube. Prepare immediately before use and sonicate in a water bath for 5 min to remove air bubbles.
15. ZipTip elute solution: 50 % ACN, 1 % FA. Mix 198 μ L of ZipTip wash solution with 2 μ L of FA. Store in a glass vial (see Note 8). Prepare immediately before use and sonicate in a water bath for 5 min to remove air bubbles.

2.4 Solutions for LC-MS/MS

1. Dissolving solution: 5 % ACN, 0.1 % FA. Mix 5 mL of ACN and 0.1 mL of FA and make up to 100 mL with water.
2. Mobile phase A: 0.1 % (v/v) FA in water. Mix 1 mL of FA and 999 mL of water.
3. Mobile phase B: 0.1 % (v/v) FA acid in ACN. Mix 1 mL of FA and 999 mL of ACN (HPLC grade).

3 Methods

3.1 OB Isolation

1. Weigh 1 g of *Arabidopsis thaliana* grains and soak them for 1 h in IB1 (5 mL) at 4 °C.
2. Grind the grains slowly in a mortar with a pestle for 20 min on ice. Successively, during the grinding, add 15 mL of IB1.
3. Rinse the mortar with 5 mL of IB1 and grind the grains 15 times for 30 s with a Potter grinder driven by a Heidolph motor (rate 7). Between each grinding cycle, the suspension should be cooled on ice for 30 s. Rinse the grinder with 5 mL of IB1.
4. Take away 30 μ L of the rude extract into an Eppendorf tube and store at 4 °C for the control of OB isolation.
5. Distribute the suspension into six centrifugation tubes and overlay with 5 mL of IB2.
6. Balance the centrifugation tubes (the difference should be lower than 1 mg) and spin at 100,000 $\times g$ for 1 h in an ultracentrifuge equipped with a swinging bucket rotor.
7. After centrifugation, floating OB fraction should be found on the surface. Collect the OB fraction with a spatula (see Note 10).
8. Resuspend thoroughly the collected OB fraction in 2 mL of IB3 (see Note 11).
9. Take away 30 μ L of resuspended OBs to control the OB isolation. Make the OB suspension up to 30 mL with IB3.
10. Distribute the suspension into six centrifugation tubes and overlay with 5 mL of IB4.

11. Balance the centrifugation tubes (the difference should be lower than 5 mg) and spin at $24,000 \times g$ for 1 h in an ultracentrifuge equipped with a swinging bucket rotor.
12. Collect the OB floating fraction with a spatula (*see Note 12*).
13. Resuspend thoroughly the collected OB fraction in 2 mL of IB5 (*see Note 11*).
14. Take away 30 μ L of resuspended OBs to control the OB isolation. Make the OB suspension up to 30 mL with IB5.
15. Distribute the suspension into six centrifugation tubes and overlay with 5 mL of IB6.
16. Balance the centrifugation tubes (the difference should be lower than 5 mg) and spin at $24,000 \times g$ for 1 h in an ultracentrifuge equipped with a swinging bucket rotor.
17. Collect the OB floating fraction with a spatula.
18. Resuspend thoroughly the collected OB fraction in 2 mL of IB1 (*see Note 11*). Take away 30 μ L of resuspended OBs to control the OB isolation. At this moment, the isolation can be interrupted until the next day and the OB suspension stored overnight at 4 °C.
19. Add 18 mL of 9 M urea to 2 mL of OB suspension, slightly mix, and agitate at room temperature for 10 min.
20. Distribute the suspension into four centrifugation tubes and overlay with 5 mL of IB4.
21. Balance the centrifugation tubes (the difference should be lower than 5 mg) and spin at $24,000 \times g$ for 1 h in an ultracentrifuge equipped with a swinging bucket rotor.
22. Collect the OB floating fraction with a spatula (*see Note 12*).
23. Resuspend thoroughly the collected OB fraction in 2 mL of IB1.
24. Take away 30 μ L of resuspended OBs to control the OB isolation and make the OB suspension up to 15 mL with IB1.
25. Add 15 mL of hexane, slightly mix, and agitate at room temperature for 10 min to test OB integrity (*see Note 13*).
26. Distribute the mixture into two closeable centrifugation tubes.
27. Balance the centrifugation tubes (the difference should be lower than 50 mg) and spin at $7,000 \times g$ for 30 min in a centrifuge equipped with a swinging bucket rotor. Only a centrifuge tolerant to the presence of flammable solvent can be used.
28. Remove the upper organic fraction and resuspend the OBs in the water phase.
29. Remove the rest of hexane under nitrogen flow for 10 min.
30. Distribute the suspension into three centrifugation tubes and overlay with 5 mL of IB2.

31. Balance the centrifugation tubes (the difference should be lower than 5 mg) and spin at $24,000 \times \mathcal{G}$ for 1 h in an ultracentrifuge equipped with a swinging bucket rotor.
32. Collect the OB floating fraction with a spatula.
33. Resuspend thoroughly the collected OB fraction in 2 mL of IB1.
34. Take away 30 μL of resuspended OBs to control the OB isolation and make the OB suspension up to 10 mL with IB1.
35. Remove the rest of hexane under nitrogen flow for 10 min.
36. Distribute the suspension into three centrifugation tubes and overlay with 5 mL of IB2.
37. Balance the centrifugation tubes (the difference should be lower than 5 mg) and spin at $24,000 \times \mathcal{G}$ for 1 h in an ultracentrifuge equipped with a swinging bucket rotor.
38. Collect the OB floating fraction with a spatula.
39. Resuspend the collected purified OBs in 500 μL of storage buffer.
40. To quantify the isolated OB, take away 1 μL of OB suspension, dilute it 1,000 times with the storage buffer, and measure the OD_{600} . The final OD_{600} of 1,000 times diluted sample should be 0.2–0.3. Usually, the yield of the OB isolation is between 700 and 1,200 μL of OB suspension. Store isolated OBs at 4 °C, but no longer than 3 days.

3.2 SDS-PAGE of OBs

1. Prepare the solution for 15 % SDS-PAGE gel mixing 2.3 mL of water, 5.0 mL of 30 % AA/BIS, 2.5 mL of resolving gel buffer, and 100 μL of 10 % SDS in a 50 mL glass beaker. Finally, add 100 μL of 10 % APS and 4 μL of TEMED and slightly mix.
2. Cast quickly two gels within a 7.25 cm \times 10 cm \times 1.0 mm cassette. Leave about 1.5 cm wide spaces for stacking gel and overlay with water.
3. When the gels solidify, remove covering water.
4. Prepare stacking gel mixing 2.7 mL of water, 0.67 mL of 30 % AA/BIS, and 0.5 mL of water in a 25 mL glass beaker. Add 40 μL of 10 % SDS. Finally, add 40 μL of 10 % APS and 4 μL of TEMED and slightly mix.
5. Cast the stacking gels and immediately insert 10-well gel combs without introducing air bubbles.
6. SDS-PAGE can be used either to evaluate the quality of OB isolation, or, mainly, for further LC-MS/MS analysis.
7. To evaluate the quality of OB isolation, mix 15 μL of control samples (one crude extract, five samples from purified fractions, and final OB suspension) with 15 μL of PLB and proceed as follows from **step 10**. The final OB protein profile is in Fig. 1.

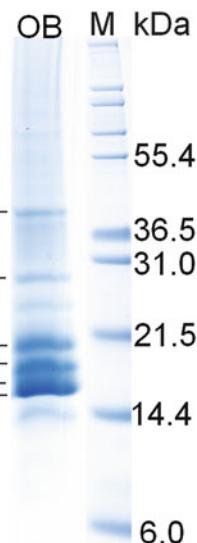


Fig. 1 SDS-PAGE protein profile of isolated OBs. Bands 1–6, corresponding to OB proteins, were excised and analyzed with LC-MS/MS. The OB lane is intentionally overloaded, to increase the protein concentration. This is also the reason why the other, nonspecifically bound proteins are visible at the OB profile (but in lower concentrations). *M*—protein standards (marker)

8. For further LC-MS/MS analysis, distribute 100 μ L of mixed OB suspension into Eppendorf tube. For each proteolytic system, one tube with 100 μ L of OBs is needed. Prepare three samples of OB suspension for all three proteolytic systems.
9. Spin the samples at 20,000 $\times g$ for 20 min at 4 °C. Using an injection needle coupled with a syringe, remove the suspension buffer bellow the OB layer. Leave the OBs in the tube and add 100 μ L of PLB. Resuspend the OBs in the PLB by tapping the Eppendorf tube over the edges of an Eppendorf tube stand. This simple pre-concentration step allows loading higher protein amounts on SDS-PAGE.
10. Heat the samples at 72 °C for 10 min (*see Note 14*).
11. Load the samples on the gel. 100 μ L of the sample should be loaded into four lanes (*see Note 15*).
12. Add protein standards in one lane at each gel.
13. Run the electrophoresis under 160 V till the bromophenol blue dye front has reached the bottom of the gel (usually 1 h).
14. Turn off the electrophoresis and open the gel plates with a spatula. Rinse the gel with water and transfer carefully to a container.

15. Add CBB staining solution to completely cover the gels. Stain overnight, slightly agitating.
16. Remove the CBB staining solution and destain four times for 30 min with water.

3.3 In-Gel Digestion

Work cleanly and cautiously to avoid keratin contamination of the samples. Use an agitator for Eppendorf tubes, which allows rate 300–600 RPM and setting temperature.

1. Excise the bands of interest from the gel using a scalpel. We were able to distinguish six OB proteins in the profile (Fig. 1). Cut the protein bands into small pieces with an edge of 1 mm (*see Note 16*).
2. Collect the gel pieces from corresponding bands in all four lanes into one Eppendorf tube. Usually, 70 μ L of solution is enough to cover gel pieces from four lanes.
3. Wash the gel pieces with 70 μ L of water for 5 min at room temperature.
4. Remove water and add around 70 μ L of washing buffer/ACN mixture to destain the gel. Agitate at room temperature for 20 min. Repeat twice or three times to completely destain the gel pieces.
5. Remove the solution and add 70 μ L of ACN and keep in for 1 min to dehydrate the gel pieces. Repeat, if needed, to completely dry the gel.
6. Add 70 μ L of DTT solution and allow the dry pieces to soak. Make sure that all the pieces are covered with the solution. Agitate at 56 °C for 45 min.
7. Cool the tubes to room temperature, remove the DTT solution, and add 70 μ L of IAA. Agitate at room temperature in the darkness for 30 min.
8. Wash the gel pieces with washing buffer/ACN mixture for 15 min at room temperature.
9. Cool the agitator to 4 °C (*see Note 17*).
10. Remove the solution from tubes, add 70 μ L of ACN, and keep in for 1 min to dehydrate the gel. Repeat, if needed, to completely dry the gel pieces.
11. Prepare protease solution, trypsin, chymotrypsin, or their mixture. Add 30–60 μ L of protease solution to the gel pieces and allow them to soak at 4 °C. Make sure that all the pieces are covered with the solution. Agitate at 4 °C for 45 min.
12. Remove the rest of protease solution and add 30–60 μ L of 50 mM ammonium buffer to completely cover the gel pieces.

13. Digest for 3 h at 37 °C. According to the enzyme efficacy, the time of digestion could be shortened to 1 h, especially for the protease mixture. When the digestion is complete, the experiment can be interrupted until the next day and the samples stored overnight at 4 °C.
14. After digestion, remove the solution around gel pieces from each tube into a new one to gain the peptides produced.
15. Cover the gel pieces with extraction buffer 1 and sonicate in a water bath for 15 min.
16. For each tube, remove the solution and collect it with the first portion.
17. Cover the gel pieces with extraction buffer 2 and sonicate in a water bath for 15 min.
18. Remove the solution and collect with the first two portions.
19. Lyophilize the joined extracts under vacuum.
20. Resolve the content of the tubes in 20 µL of 1 % FA and sonicate for 10 min to complete the dissolution and to remove air bubbles.
21. Clean each solution with ZipTip pipette tips as follows.
22. Put the ZipTip on a 2–20 µL scaled pipette and set the volume to 18 µL. Wet the ZipTip three times with ZipTip wash solution and three times with ZipTip equilibration solution (*see Note 18*).
23. Draw in and out seven times the sample solution. (Do not pour out the sample solution.)
24. Wash the ZipTip twice with ZipTip equilibration solution.
25. Draw in 2 µL of ZipTip elute solution. Draw in and out three times and finally draw out in a clean Eppendorf tube (PCR tube).
26. Repeat once **steps 22–25** for the same samples. Connect both the portions and let dry under vacuum.

3.4 LC-MS/MS Identification

1. Dissolve the sample in 15 µL of dissolving solution (5 % ACN, 0.1 % FA).
2. Load 4–10 µL of sample (according to the instrument sensitivity) onto an HPLC chromatography system equipped with a reversed-phase C18 column (*see Note 19*).
3. Elute the column at 20 °C with 5 % of mobile phase B in A for 2 min and then over 40 min with a linear gradient of B in A from 5 to 45 %. Finally elute the column with 45–95 % gradient of B in A over 5 min before re-equilibration.
4. The eluent from the column is introduced in the electrospray ionization source of an MS/MS instrument operating in positive ion mode (*see Note 20*).

5. Analyze peptide ions using the data-dependent “triple-play” method: (1) full mass spectrum scan, (2) zoom scan of the major ions with higher resolution to determine their charge, and (3) fragmentation of these ions. Acquire mass spectra from mass/charge 200–2,000.
6. Perform protein identification with adequate software and reliable data evaluation (*see Note 21*). Download *A. thaliana* protein sequences from the National Center for Biotechnology Information FTP site (<http://www.ncbi.nlm.nih.gov/>). Set cysteine carbamidomethylation to static modification and methionine oxidation and N-terminal acetylation to possible modifications. For precise analysis, do not specify the protease used (*see Note 22*).

4 Notes

1. We store the grains in a refrigerator, in Eppendorf tubes (1 g of grains per tube) thoroughly closed and sealed with parafilm.
2. At first, concentrated NaOH (10 M) can be used, but when getting closer to the final pH, it is better to use diluted NaOH solutions. Proceed similarly when acidifying solutions with HCl.
3. Dilution of Tween 20 is complicated because of its high viscosity. Use a 10 mL graduated test tube. Fill the tube with 9 mL of ultrapure water and make up to 10 mL with Tween 20. For the manipulation with Tween 20 use a Pasteur pipette. Store the solution in a dark place.
4. Be careful not to exceed the final volume (20 mL) of the solution. Add water in small doses and slightly heat the mixture to support urea dissolution. Before further use, the solution must be slowly returned to room temperature.
5. Work cautiously with acrylamides and their solutions due to their high toxicity.
6. Acrylamide hydrolyzes to ammonia and acrylic acid, if stored for longer time.
7. Be careful when weighing large amounts of SDS. Avoid moving rapidly the open container because SDS irritates mucosa when inhaled.
8. Avoid storing ACN in plastic vessels; ACN, as an organic solvent, may extract polymers hindering LC-MS/MS analysis.
9. We use Sequencing Grade Modified Trypsin purchased from Promega and α -Chymotrypsin from bovine pancreas purchased from Sigma-Aldrich.
10. Usually, the OB layer is solid enough to be easily collected with spatula.

11. Thorough and patient resuspension is an important step of successful OB isolation. Use a blue pipette tip for rough resuspension, and a yellow pipette tip to break the small pieces of OBs.
12. This time, the OB layer is less solid due to lower sucrose concentration, but still solid enough to be collected with spatula. If necessary, the left pieces of the layer could be collected with a 200 μ L pipette, but avoid drawing larger volumes of the solution.
13. If the OBs have good integrity, in further centrifugation we can collect the OB layer as usual. If the OB integrity is broken, after centrifugation, we find wide oil layer between water and hexane phase and it is necessary to repeat all the isolation. The integrity may be broken, only if the temperature during the isolation is excessively increased. For example, hot sunny day in a non-air-conditioned laboratory is not ideal for OB isolation
14. Lower temperature is convenient for hydrophobic proteins.
15. You may use a comb with wider wells and load all the sample into one well, but we have achieved best results with classic 10-well comb and by dividing the sample into four lanes.
16. The gel pieces dry out during the excision and dry gel pieces tend to jump when touched by the scalpel. Avoid “gel jumping” by wetting with a small drop of water.
17. Some agitators cannot reach so low temperature. In this case, incubate the Eppendorf tubes in a container with ice, on a planar agitator.
18. When the ZipTip pipette tip is once wet, avoid drawing in air bubbles.
19. We used BioBasic-18 column (1 \times 150 mm, 300 \AA pore size, 5 μ m film thickness, Thermo Electron). Flow rate was 0.1 mL/min.
20. We used Thermo Electron LCQ Deca ion-trap mass spectrometer. Instrumental parameters were as follows: capillary temperature 280 $^{\circ}$ C; capillary voltage 30 V; spray voltage 4.5 kV; sheath gas flow, 80 a.u.; and auxiliary gas flow, 5 a.u. Normalized collision energy of 35 a.u. was used for ion fragmentation.
21. We used Bioworks 3.1TM software.
22. Chymotrypsin, known for its low specificity, and even trypsin, may cleave apart from their cleavage sites, also behind other amino acids (for example besides histidine). Although the fragmentation spectra were reliable enough, we considered also the peptides semi-tryptic or semi-chymotryptic.

References

1. Tzen JTC, Huang AHC (1992) Surface structure and properties of plant seed oil bodies. *J Cell Biol* 117:327–335
2. Murphy DJ (2001) The biogenesis and functions of lipid bodies in animals, plants and microorganisms. *Prog Lipid Res* 40:325–438
3. Poxleitner M, Rogers SW, Samuels AL et al (2006) A role for caleosin in degradation of oil-body storage lipid during seed germination. *Plant J* 47:917–933
4. Parthibane V, Rajakumari S, Venkateshwari V et al (2012) Oleosin is bifunctional enzyme that has both monoacylglycerol acyltransferase and phospholipase activities. *J Biol Chem* 287: 1946–1954
5. Kim HU, Hsieh K, Ratnayake C et al (2002) A novel group of oleosins is present inside the pollen of *Arabidopsis*. *J Biol Chem* 277: 22677–22684
6. Vermachova M, Purkrbova Z, Santrucek J et al (2011) New protein isoforms identified within *Arabidopsis thaliana* seed oil bodies combining chymotrypsin/trypsin digestion and peptide fragmentation analysis. *Proteomics* 11: 3430–3434
7. Tzen JTC, Peng CC, Cheng DJ et al (1997) A new method for seed oil body purification and examination of oil body integrity following germination. *J Biochem* 121:762–768
8. Jolivet P, Roux E, d'Andréa S et al (2004) Protein composition of oil bodies in *Arabidopsis thaliana* ecotype WS. *Plant Physiol Biochem* 42:501–509
9. Neuhoff V, Arold N, Taube D et al (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9:255–262

Chapter 15

Mass Western for Absolute Quantification of Target Proteins and Considerations About the Instrument of Choice

David Lyon, Wolfram Weckwerth, and Stefanie Wienkoop

Abstract

The Mass Western describes the absolute quantification of proteins based on stable isotope labeled integral standard peptides and liquid chromatography coupled selective reaction monitoring triple quadrupole mass spectrometry (LC-SRM/MS). Here, we present a detailed workflow including tips and we discuss advantages and disadvantages of using different types of MS for absolute quantification.

Key words Mass Western, SRM, Heavy peptide internal standard, Absolute quantification, Triple quadrupole, Orbitrap

Abbreviations

| | |
|-----------|---|
| SID | Stable isotope dilution |
| SRM | Selective reaction monitoring |
| PTM | Posttranslational modification |
| QqQ | Triple quadrupole |
| LC-SRM/MS | Liquid chromatography coupled selective reaction monitoring triple quadrupole mass spectrometry |
| HP | Heavy peptide ($^{15}\text{N}^{13}\text{C}$ labeled synthetic standard) |
| LP | Light peptide (native peptide, no labeling) |
| HCD | High energy collision induced dissociation |
| CE | Collision energy |
| LDR | Linear dynamic range |
| LOD | Limit of detection |
| LOQ | Limit of quantification |

1 Introduction

Absolute protein quantification using stable isotope dilution (SID) in conjunction with liquid chromatography (LC) and selective reaction monitoring (SRM) triple quadrupole (QqQ) mass

spectrometry (MS) has been successfully applied to highly complex crude proteomics samples [1, 2]. In contrast to relative quantification, absolute quantitative data result in absolute concentration levels. Thus besides comparison of different experimental treatments, absolute quantification enables the analysis of protein stoichiometry within a sample, the differentiation of isoforms as well as the comparison of inter-experimental conditions, such as different species. Additionally, it leads to highly verifiable data. In analogy to the well-known Western Blot, Lehmann et al. [3] first coined the suitable term Mass Western. While for Western Blot analysis synthetic peptides can be used for the synthesis of antibodies, they can be directly applied for the sensitive and targeted detection and absolute quantification using the Mass Western. It is the integration of stable isotope labeled synthetic peptides in combination with gel based or gel-free LC-SRM/MS. A theoretical and an experimental approach to set up the Mass Western can be distinguished (Fig. 1). Both approaches start out by defining at least one protein of interest. The theoretical approach continues with *in silico* digestion of proteins and prediction of proteotypic peptides (understood as a unique amino acid sequence of a peptide, unambiguously identifying a specific protein of interest within a given proteome) for the target proteins (in reference to the proteome of the sample). Subsequently, the

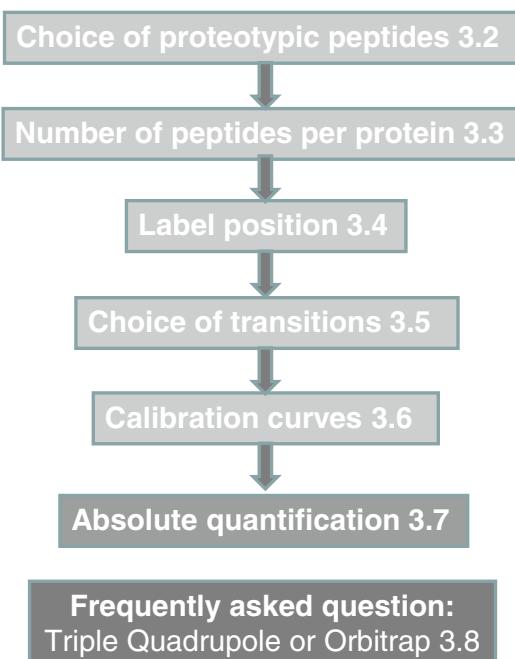


Fig. 1 Workflow diagram of the steps for Mass Western as described in the text

chosen peptides are synthesized, including (at least) one stable isotope labeled amino acid (being highly enriched with ^{15}N and ^{13}C) per peptide. In contrast to the latter, within the experimental approach, proteotypic peptides are chosen through data mining of shotgun-proteomics experiments, mass spectral reference databases such as the plant proteomic spectral library ProMEX [4], as well as protein fractionation, enrichment, and subsequent mass spectral analysis. The ensuing synthesis of stable isotope labeled peptides is identical to the theoretical approach, as well as the rest of the procedure. LC-SRM/MS method development follows and completes the workflow of the Mass Western approach. The detailed workflow is presented in the specific context of quantification capacity of a QqQ, here a TSQ (Thermo Triple Stage Quadrupole Vantage), compared to a Linear Trap Quadrupole-Orbitrap (LTQ-Orbitrap).

2 Materials

2.1 Software

We recommend the use of Skyline ([5], <https://brendanx-uwl.gs.washington.edu/>) for method development and data analysis. Proteolytic cleavage probability, hydrophobicity calculation, and PTM prediction are part of many useful features available at Expasy Tools (<http://ca.expasy.org/>). Skyline in conjunction with SRM collider can be used for the prediction of proteotypic peptides and to find interferences in a given background proteome [6].

2.2 Mass Spectrometer

TSQ (Thermo Triple Stage Quadrupole, Vantage).
LTQ-Orbitrap (Thermo Linear Trap Quadrupole-Orbitrap XL).

3 Methods

3.1 General Overview

The heavy synthetic standard peptide (containing one amino acid labeled with ^{15}N and ^{13}C ; HP) and its native counterpart, the light peptide (with no artificially introduced label; LP), share identical physicochemical properties. Retention time, ionization efficiency, and fragmentation properties of a HP-LP pair are assumed to be identical. The HP is used to tune the mass spectrometer and to create an external calibration curve. The peak area of the LP is set into the calibration curve to calculate the absolute quantity. When measuring samples for quantification, equal amounts of HP are spiked into each sample, as internal standards. The internal standard serves as a quality control. The retention time, peak shape, and relative intensities of individual transitions (see Notes 1 and 2) of the LP are compared to the HP. This gives the experimenter great confidence concerning the accuracy of the data used for

quantification. The incorporation of a HP internal standard is the only approach that enables absolute quantification. The peak area of the HP can be used to normalize data across multiple samples (by dividing each individual peak area of a sample by the peak area of the HP of that particular sample). Since a known amount of HP is spiked into the sample, the peak area of the HP can be compared to the mean and standard deviation of its calibration curve at the given concentration. Sample handling and or technical aberrations can thereby be detected. Subsequently, the measurement can be repeated or a correction factor applied. The discrete steps to set up a Mass Western experiment are described as follows.

3.2 Choice of Proteotypic Peptides

When setting up a Mass Western using the experimental approach, all identified peptides of the target protein are considered as putative candidates. The amino acid sequences of each peptide can be BLASTed against the proteome of the organism, in order to distinguish proteotypic peptides (other tools such as Skyline are available as well). Sequences containing missed cleavage sites should be avoided, due to the potential occurrence of proteolytic peptides consisting of partial sequences of the targeted peptide. Subsequently, the total amount of the targeted peptide will decrease and thus the accuracy and sensitivity of the experiment. Methionine should be avoided, since it is prone to oxidation. Cysteine residues tend to build disulfide bridges and can result in dimers or cyclization, though alkylation can counteract this problem. Generally, reproducibly identifiable proteotypic peptides are very promising candidates for a Mass Western experiment. Peptides with the highest ionization efficiency (and thus the most intense signal), and best chromatographic peak shape should be chosen.

When approaching a Mass Western experiment theoretically, the putative list of proteolytic peptides, resulting from *in silico* digestion of the protein of interest, needs to be reduced to proteotypic candidates. This can be performed by BLAST, as previously mentioned. Further selection criteria are as follows: Peptide length should preferably be between eight and twenty amino acids. Proteolytic efficiency should be as high as possible and can be estimated by tools such as Peptide Cutter (http://web.expasy.org/peptide_cutter/). N-terminal Glutamine has a tendency to form cyclic peptides. Low coupling efficiency due to the hydrophobic and steric characteristics of Tryptophan can pose problems when synthesizing peptides. Generally hydrophobicity and thus solubility of the peptides needs to be considered not only concerning synthesis, but also concerning extraction, digestion and resuspension of proteins/peptides. Sequences containing Proline can produce multiple chromatographic peaks due to enantiomers, but are also known to easily and prominently fragment in MS/MS experiments. Furthermore, potential Post Translational Modifications (PTMs) should be considered. To the best of the author's

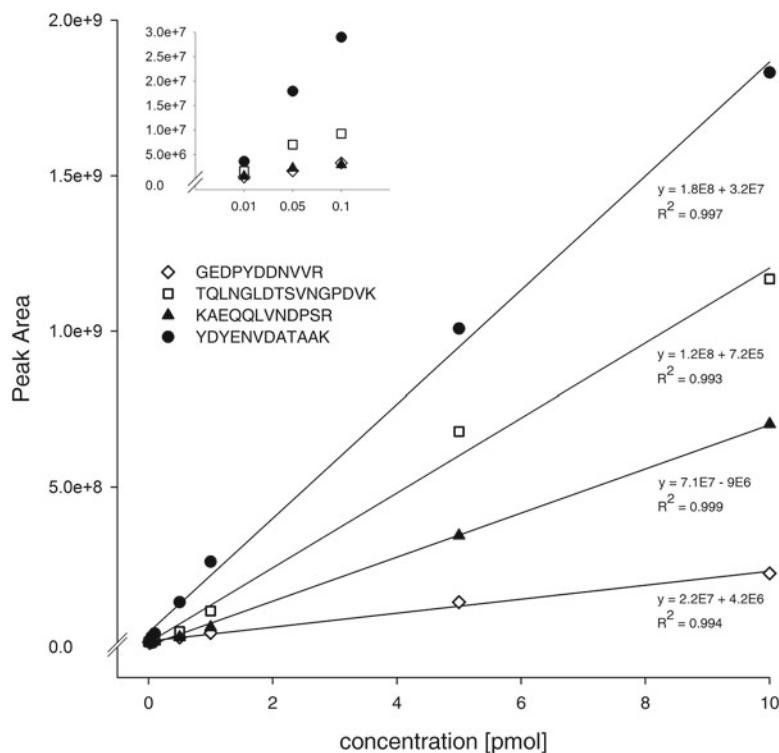


Fig. 2 Concentration versus Peak Area. A dilution series of four HPs was measured on a Thermo TSQ Vantage QqQ-MS. The ionization efficiency, and thus the signal response, is dependent on the amino acid sequence, as can be seen by the differing slopes of the linear regression lines

knowledge, no robust method exists to predict the signal intensity (ionization efficiency) of a given peptide sequence (Fig. 2). This can potentially be a major drawback, since the ionization efficiency can only be determined experimentally. Subsequently, an otherwise suitable peptide can exhibit very low signal intensities, and thus pose analytical difficulties. Digestion efficiency, peptide solubility, ionization efficiency, and matrix effects are accounted for within the experimental approach, in contrast to the theoretical approach.

3.3 Number of Peptides per Protein

Peptides should be chosen from different regions (e.g., the middle of and close to the C-terminus) of the protein. The more peptides per protein measured, leading to identical results, the more certainty about the quality of the results can be assumed. Principally, one peptide per protein should be sufficient for quantification. Nonetheless, at least two peptides per protein should be used for the Mass Western, if possible, due to variability in proteolytic efficiency and differing recovery rates. Necessary validation can thus be performed by comparing individual peptide results for a given protein.

3.4 Label Position

No matter the label position, the precursor m/z values of a HP-LP pair (isotopologues) will differ, dependent on the labeled amino acid. In principal, it doesn't matter which amino acid is chosen, if ^{15}N and ^{13}C labeling is applied, since delta M (the mass shift) will be at least 4 Da (for Alanine) which can easily be separated by modern mass spectrometers. However, it is recommended choosing the C-terminal amino acid of the synthetic peptide to be labeled with ^{15}N and ^{13}C . Y-ions are predominantly used for SRM transitions in QQQ, due to their abundance and intensity but b-ions may be chosen as well. Subsequently, the product m/z values of any Y-ion will also differ for a HP-LP pair, resulting in highly selective transition pairs. In theory, the selection of one transition is enough. However, the more transitions per peptide selected, the higher the certainty. More transitions may also increase signal intensity while improving sensitivity, but increase the duty cycle at the same time.

3.5 Choice of Transitions

It is imperative to use identical types of transitions for a HP and its LP counterpart (The charge state of the precursor of the HP needs to be the same as for the LP. If a singly charged Y₃-type-ion is chosen for the HP, the same needs to be chosen for the LP.) A similar ion fragmentation intensity pattern has been described comparing HCD collision with QQQ fragmentation [7]. However, tuning of the mass spectrometers collision energy (CE) by direct infusion (flow injection) of the synthetic peptides enables maximum sensitivity for specific transitions. Atleast for some mass spectrometers a semiautomatic ramping of CE and selection of the most intense transitions is possible [8]. The occurrence of other parameters such as Declustering Potential, S-lens, Collision Exit Potential, Ion Transfer Capillary Offset Voltage, etc. is vendor specific. Fine tuning of all possible parameters guarantees the highest possible sensitivity of the experiment. Y-ions N-terminal to Proline frequently result in high signal intensities, and are therefore preferably chosen. Selecting the most abundant fragments aims at maximum sensitivity of the assay (see Note 1). High selectivity can usually be achieved by choosing transitions whose product m/z values are higher than their precursor m/z value (possible due to for example precursor charge state 2, product charge state 1). A combination of sensitive and selective transitions often results in the most effective experimental setup. In general two transitions per peptide are sufficient for sensitivity and selectivity since the HP internal standard includes the retention time as additional confidence identification parameter (see Note 2).

3.6 Calibration Curves

Comparing HP and LP peak areas to deduce quantitative results (single point calibration) is not recommended due to the well known fact that peptide ionization efficiency varies significantly (Fig. 2). Except for nearly identical peak areas, quantitation will not be as

accurate as with external calibration curves. An external calibration curve is recorded as follows. A dilution series of the synthetic peptides is measured using optimized settings for the instrument given. The linear dynamic range (LDR), the limit of detection (LOD), and the limit of quantification (LOQ) can thereby be estimated. If the standard peptides are spiked into a complex matrix (reflecting the nature of the sample to be analyzed), the LDR, LOD, and LOQ in matrix can be assessed more precisely concerning the performance of the assay. Accurate quantification can only be achieved within the linear dynamic range. For each HP a linear regression is calculated for data within its linear dynamic range.

3.7 Absolute Quantification

When measuring samples for quantification, the amount of HP added to the individual samples should be kept constant. The heavy standard peptide (HP) is introduced to the sample at the earliest possible point of time. The signal intensity of the HP standard should be as close as possible to that of the native target. The peak area integral of the HP can be used to normalize data across multiple samples. The LP peak area can thus be adjusted and set into the regression equation of the external calibration curve to calculate a quantitative value. Best results are expected if the values are in the middle of the linear regression.

3.8 Frequently Asked Question: Triple Quadrupole or Orbitrap

The SRM approach can generally be executed on different types of MS. At present, for absolute quantification triple quadrupole mass spectrometers (QqQ) are routinely being used due to a wide linear dynamic range, excellent sensitivity and selectivity, as well as acquisition speed. In contrast, due to the very long duty cycles and incomparable sensitivity, quantification based on SRM on the hybrid MS instrument LTQ-Orbitrap-XL is not recommended. The LTQ-Orbitrap-XL MS features high mass accuracy and resolution, and is an excellent tool for discovery experiments (such as shotgun proteomics), and relative quantification [7, 8]. The latter MS instrument has also successfully been used for LC-MS based quantification with the drawback of a reduced linear dynamic range compared to QqQ LC-SRM/MS. Recent articles state that the Orbitrap-Exactive is “equal or better” than QqQ if full scan acquisition based quantification is applied [9, 10] (see Note 3). However, at present the use of QqQ holds advantages compared to other types of instruments such as a Linear Trap Quadrupole-Orbitrap (LTQ-Orbitrap). Figure 3 illustrates the differing linear dynamic range of the aforementioned instruments for three distinct standard HP. The two MS instruments are compared for applicability of absolute quantification (methodological details can be found at <http://www.univie.ac.at/mosys/publications.html>). The figure shows that even though LTQ-Orbitrap MS signal intensities in data dependent mode appears higher, the linear dynamic range is bigger for the QqQ Triple Stage Quadrupole (TSQ) MS (see Note 4).

Comparison of linear dynamic range of 3 standard Heavy Peptides
on the LTQ-Orbitrap-XL (FS) and the TSQ (SRM)

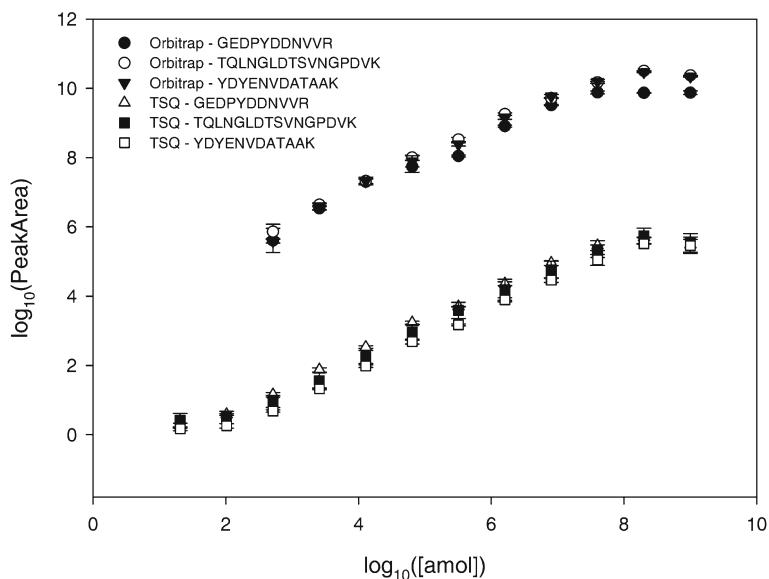


Fig. 3 Concentration versus Peak Area. Three Stable isotope labeled synthetic peptides was subjected to a dilution series, and measured with the LTQ-Orbitrap-XL, in data-dependent acquisition mode, and the TSQ-Vantage, in SRM-mode. Data indistinguishable from noise or simply not present are not depicted, and thus the “missing” data points in the Orbitrap function at lower concentration levels. The linear and nonlinear parts of the two functions are distinguished by the flattening of the respective functions. In order to attain an optimal coefficient of determination (R^2) similar for both instruments, the linear regression was calculated with a subset of the data points. The resulting regression lines are depicted in the graph. The upper but especially the lower limit of the TSQ regression line clearly extends to lower concentration levels, indicating heightened sensitivity over a wider linear dynamic range

This is due to the fact that low concentration signals are detectable in TSQ while the signal-to-noise ratio is already too low in the LTQ-Orbitrap. Thus, when aiming to quantify very low abundant targets in complex matrices, QqQs are the instruments of choice. Recent developments in the field of mass spectrometry include optimized Ion Optics (enabling larger amounts of ions to be focused in a shorter amount of time), data acquisition speed, data acquisition range, sensitivity [10–13] as well as the emergence of novel DIA (Data Independent Acquisition) instruments, methods and software [14]. Thus high resolution and high mass accuracy instruments, such as the latest generation of the Orbitrap, are increasingly competitive to QqQs. Additionally data produced by these instruments is not restricted to quantification as with SRM. These and other studies show that at present the use of an LTQ-Orbitrap MS may be an alternative to QqQ if the sensitivity is not limited [10]. For improved sensitivity of low abundant proteins using FullScan high resolution high mass accuracy MS some instrument adjustments may be useful (see Note 5).

4 Notes

1. When multiple SRM transitions for a given substance are chosen, the individual signal intensities are added to a single signal. This should also be considered when comparing high mass accuracy and resolution FullScan methods with SRM/MRM methods as well as the generation of the mass spectrometer, instrument parameters, and complexity of the underlying sample matrix.
2. False positive signals can be disregarded and do not necessarily distort quantification, since true signals can reliably be identified by comparing the retention time, peak shape, and relative transition order of a HP-LP-pair.
3. Since metabolites most frequently are detected with a single charge, the precursor is larger than its product m/z value, and SRM transitions whose precursor is smaller than its product m/z value can't be selected.
4. The systematic shift of Peak Area values (arbitrary units), between the two instruments does not imply any qualitative or quantitative difference.
5. Reducing the scan range (Full Scan or Selected Ion Monitoring) elevates the signal to noise ratio and improves sensitivity, however, useful information may thus become unavailable [15]. Elevating target Automatic Gain Control (AGC) values can improve the dynamic range, due to elevated sensitivity for low concentrations, but can lead to negative space charge effects at high concentrations [12].

Acknowledgments

We thank Thomas Naegele for useful comments and discussions.

References

1. Desiderio DM, Kai M (1983) Preparation of stable isotope-incorporated peptide internal standards for field desorption mass spectrometry quantification of peptides in biologic tissue. *Biomed Mass Spectrom* 10:471–479
2. Wienkoop S, Weckwerth W (2006) Relative and absolute quantitative shotgun proteomics: targeting low-abundance proteins in *Arabidopsis thaliana*. *J Exp Bot* 57:1529–1535
3. Lehmann U, Wienkoop S, Tschoep H et al (2008) If the antibody fails—a mass western approach. *Plant J* 55:1039–1046
4. Wienkoop S, Staudinger C, Hoehenwarter W et al (2012) ProMEX—a mass spectral reference database for plant proteomics. *Front Plant Sci* 3:125
5. MacLean B, Tomazela DM, Shulman N et al (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26:966–978
6. Röst H, Malmström L, Aebersold R (2012) A Computational Tool to Detect and Avoid Redundancy in Selected Reaction Monitoring. *Mol Cell Proteomics : MCP* 11 (8) (August): 540–549. doi:10.1074/mcp.M111.013045. <http://www.ncbi.nlm.nih.gov/pubmed/22535207>

7. Makarov A, Denisov E, Kholomeev A et al (2006) Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal Chem* 78:2113–2120
8. Olsen JV, Godoy LMFD, Li G et al (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol Cell Proteomics* 4:2010–2021
9. Gallien S, Duriez E, Crone C et al (2012) Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer. *Mol Cell Proteomics mcp.O112.019802*
10. Henry H, Sobhi HR, Scheibner O et al (2012) Comparison between a high-resolution single-stage Orbitrap and a triple quadrupole mass spectrometer for quantitative analyses of drugs. *Rapid Comm Mass Spectrom* 26:499–509
11. Bateman KP, Kellmann M, Muenster H et al (2009) Quantitative-qualitative data acquisition using a benchtop Orbitrap mass spectrometer. *J Am Soc Mass Spectrom* 20:1441–1450
12. Lu W, Clasquin MF, Melamud E et al (2010) Metabolomic analysis via reversed-phase ion-pairing liquid chromatography coupled to a stand alone orbitrap mass spectrometer. *Anal Chem* 82:3212–3221
13. Seppälä U, Dauly C, Robinson S et al (2011) Absolute quantification of allergens from complex mixtures: a new sensitive tool for standardization of allergen extracts for specific immunotherapy. *J Proteome Res* 10:2113–2122
14. Gillet LC, Navarro P, Tate S et al (2012) Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: a New Concept for Consistent and Accurate Proteome Analysis. *Mol Cell Proteomics: MCP* 11(6)(June):O111.016717. doi: [10.1074/mcp.O111.016717](https://doi.org/10.1074/mcp.O111.016717). <http://www.ncbi.nlm.nih.gov/pubmed/22261725>.
15. Venable JD, Dong M, Wohlschlegel J et al (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat Methods* 1:1–7

Chapter 16

Selected Reaction Monitoring Mass Spectrometry: A Methodology Overview

H. Alexander Ebhardt

Abstract

Moving past the discovery phase of proteomics, the term targeted proteomics combines multiple approaches investigating a certain set of proteins in more detail. One such targeted proteomics approach is the combination of liquid chromatography and selected or multiple reaction monitoring mass spectrometry (SRM, MRM). SRM-MS requires prior knowledge of the fragmentation pattern of peptides, as the presence of the analyte in a sample is determined by measuring the m/z values of predefined precursor and fragment ions. Using scheduled SRM-MS, many analytes can robustly be monitored allowing for high-throughput sample analysis of the same set of proteins over many conditions. In this chapter, fundaments of SRM-MS are explained as well as an optimized SRM pipeline from assay generation to data analyzed.

Key words Selected reaction monitoring (SRM), Multiple reaction monitoring (MRM), Targeted mass spectrometry

1 Introduction

SRM-MS is a mass spectrometry method typically performed on triple-quadrupole instruments taking full advantage of the large dynamic range of quadrupoles. The first quadrupole is used to define the precursor ion mass (Q1) with a specified window, e.g., $m/z \pm 0.35$. All precursor ions passing through the first window are fragmented in the second quadrupole. The third quadrupole (Q3) acts again as a filter with a specific window, e.g., $m/z \pm 0.35$, to allow for specific fragment ion to pass through toward the detector. These dual filters provide a higher sensitivity for monitoring specific analytes as supposed to conventional shotgun experiments. The values of Q1, a particular precursor ion, and Q3, a specific fragment ion, are often referred to as *transitions*. The intensity of the transition is recorded by a detector (electron multiplier tube) resulting in an ion chromatograms as a function of time and transition measured. Per precursor ion, typically five transitions are

measured, all of which coelute to positively identify the precursor/peptide of interest [1]. A group of transitions per peptide is termed *assay* and is instrument specific. Due to the nature of SRM-MS, the ion chromatogram is distinct from conventional shotgun MS/MS spectra, which record all possible fragment ions of a particular precursor ion. Hence, SRM-MS does not depend on a single spectrum for positive identification, but on coeluting transitions [1] (see Notes 1, 2 and 6).

Using unscheduled SRM, the number of transitions is defined by the *dwell time* of each transition and the *cycle time*. The cycle time is defined as time it takes to measure all transitions listed in the method, the dwell time per transition and some instrument specific set up time to switch the *m/z* value of the filters. For example if each transition is measured for 25 ms dwell time and instrument set up time is 3 ms, in 2,000 ms approximately 71 transitions can be measured. With 5 transitions per peptide, an unscheduled SRM-MS experiment would allow to monitor 14 peptides. As this number is relatively low, scheduled SRM-MS is employed. For scheduled SRM-MS, prior knowledge of the retention time of a peptide is required (see Note 3). Effectively, a peptide assay is measured only during a 2–4 min window during the LC-SRM-MS run allowing for hundreds of assays to be monitored in a single injection [2] (see Note 5).

Another strength of SRM-MS is the robustness of the methodology allowing for consistent measurements of different sample conditions, but also obtaining reliable results across various laboratories with coefficient of variation of less than 20 % [3] (see Note 4). The robustness of measurements extends to quantifying analytes using stable isotope labeled reference standards. Typically, the endogenous analyte is unlabeled (see Note 8), while the exogenous reference peptide is marked using stable isotope labeling methods (see Note 7). As the synthetic reference peptides fragment identically to the endogenous peptides, the same SRM assay applies, albeit with the appropriate mass shift. If the relative order of transitions is not conserved between endogenous and exogenous peptides, interference might be the cause (see Note 2).

To guide the reader through the SRM-MS workflow, a flowchart of the SRM-MS method development is shown in Fig. 1 and explained in detail in each section.

2 Methods

2.1 Software Tools

There are multiple software tools utilized in a typical SRM-MS workflow. Table 1 gives an overview of the software tools described in this chapter with detailed descriptions included in their respective sections.

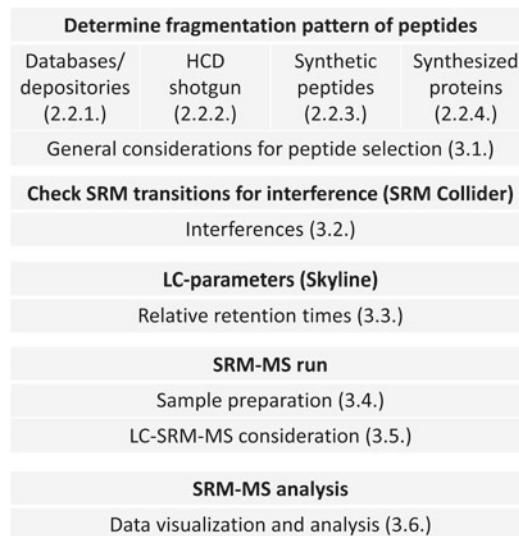


Fig. 1 Flowchart of a SRM-MS project

Table 1
List of software tools mentioned in this chapter, including references, and short descriptions

| | |
|---------------------|---|
| Skyline [4] | Windows based client application supporting quantitative SRM-MS method development and extensive analysis of resulting mass spectrometry data |
| SRM Collider [5] | Web service to detect and avoid shared transitions to increase precision of SRM-MS measurements http://www.srmcollider.org |
| MQuest/Mprophet [6] | Web service to automatically score SRM-MS measurements using decoy transitions and obtain false discovery rates (FDR) http://www.mprophet.org |

2.2 SRM Peptide assays

2.2.1 Prior Knowledge of Transitions: Repositories

2.2.2 HCD shotgun

SRM-MS requires prior knowledge on the properties of peptides to be monitored. This prior knowledge includes retention time as well as fragmentation pattern of peptides obtained with a certain mass spectrometer. Either this prior knowledge is deposited in repositories such as MR Maid [7], MASCP Gator [8], Promex [9], PASSEL [10] or has to be determined prior to SRM-MS measurements take place. Outlined below are common ways to obtain this prior knowledge.

Conventional low energy *collision-induced dissociation* (CID) ion-trap fragmentation patterns of peptides of shotgun mass spectrometers are often distinct from CID fragmentation pattern of the same peptide in a beam type triple quadrupole mass spectrometer. For example, the most intense fragment ions in CID mode are not necessarily the most intense fragment ions in a triple quadrupole due to the differences in energy transfer to the peptide found in resonance (ion-trap) versus beam type (quadrupole) [11]. Recently, changes to instrumentation have lead to the possibility of peptide fragmentation in a quadrupole

with fragments still analyzed in the mass analyzer, e.g., Orbitrap or Time of Flight. This type of fragmentation is also referred to as *higher energy collision dissociation* (HCD) fragmentation and more closely resembles the fragmentation pattern observed in a triple quadrupole instrument and is therefore the method of choice to obtain SRM assays directly from shotgun experiments [12]. Typically, multiple transitions of a HCD spectrum are selected and the same sample analyzed on a triple quadrupole instrument. The five most intense transitions are then used for a high confidence SRM assay.

Obtaining SRM assays from HCD shotgun experiments is very cost efficient. The drawback of this method is that SRM assays are obtained only for peptides that are detectable by shotgun mass spectrometry. If the protein(s) of interest are not in the list of identified peptides, fractionation is necessary. This fractionation can be in form of biological fragmentation of cellular subcomponents, or enrichment of proteins and/or protein complexes using affinity purification mass spectrometry [13]. Another way of fractionation is to separate proteins or peptides on the basis of their physical properties, e.g., first separate peptides by off-gel isoelectric focusing (pI) [14] and then analyze each fraction using shotgun reverse phase LC-MS/MS run.

2.2.3 Synthetic Peptides

As additional fractionation steps are labor intense and require several LC-MS/MS runs, another avenue are chemically synthesized peptides. Very small peptide amounts are already sufficient to generate SRM assays either in HCD shotgun mode or using mass spectrometers with a trapping device in MS1 triggered MS2 mode. The initial peptide spectra are then confirmed in SRM mode in the triple quadrupole instrument.

There are several advantages to this method: chemical peptide synthesis up to 25 amino acids is very cost efficient and many peptides can be analyzed simultaneously. As all chemically synthesized peptides are targeted peptides, resources are used efficiently, compared to off-gel pre-fractionation discussed above. Another advantage: peptide sequences can be synthesized and SRM assays obtained for peptides not detected by shotgun mass spectrometry before [15]. The disadvantage of the method lies in the prediction of peptides to be synthesized. There are many selection criteria: peptides already detected in other shotgun experiments; proteotypic peptides; predicted favorable electrospray mass spectrometry properties; no or little theoretical shared transitions; favorable hydrophobic/hydrophilic properties; fully tryptic peptide [16]. The latter criterion dictates that only a Lys or Arg residue resides at the C-terminus of a peptide. Due to difficulties predicting precise protease cleavage activity, semi-tryptic peptides are typically not considered. Further, once a choice is made to use a certain protease, e.g., trypsin cleaving C-terminal to Lys or Arg, in the future new peptides have to be synthesized if another protease is used, e.g., Asp-N cleaving N-terminal to Asp or Glu. An additional protease digest might be

necessary in cases where closely related proteins do not result in proteotypic peptides using a single protease.

2.2.4 Synthesized Proteins

Many of the disadvantages of non-targeted expeditions tracking down individual peptides in fractionated cell lysates or difficulties in synthetic peptide selection can be overcome by in vitro transcription/translation of the proteins of interest. On the genome scale this approach might be ambitious; however, individual laboratories typically are interested in networks containing hundreds of proteins, making in vitro transcription/translation of proteins for generating SRM assays a viable option [17].

In vitro transcription/translation will generate whole proteins, which can be digested by various proteases; hence, a switch in protease requires only another protein digest. The protease efficiency is considered in the context of the protein, e.g., noncanonical peptides generated are considered for SRM assay development. Peptides of the digested protein are all analyzed in an LC-MS/MS experiment and best performing peptides can be selected for subsequent SRM assay conformation. Best performing peptides are often referred to *high fliers* and frequently results in 10 times more signal than other peptides of the same protein present in equimolar amounts [17]. The proper choice of high flier peptides can make the difference between detecting the peptide in a complex mixture of whole cell lysates, or not.

In vitro transcription/translation systems were described for *Escherichia coli* [18] and are commercially available. Further, cell free transcription/translation systems are available using wheat germ extract [19, 20]. Already the comparison between prokaryotic and eukaryotic in vitro translated proteins might give insights into possible post-translational modifications (PTMs) of proteins [21]. Unexpected PTMs are one of the causes *in silico* search algorithms do not attribute a spectrum to a certain peptide. Also, as outlined below, a less complex sample is very valuable tool to determine signal interferences concerning certain SRM transitions.

In closing, in vitro transcription/translation of proteins might diverge from conventional large scale SRM assay development [16], however, in a more targeted analysis of the dynamic proteome [22] together with the flexibility of protease treatment [23] favors the whole protein synthesis.

3 Notes

1. General Considerations for Peptide Selection

Besides favorably LC and MS properties of the analyte, the unique nature of the peptide should also be considered. Unique sequences are peptide sequences that are found only

once in a given proteome and are termed *proteotypic peptides*. For closely related proteins or splice isoforms, finding proteotypic peptides with favorable LC and MS properties might be challenging. Hence, sometimes statements can be only made for a protein or closely related protein group, but not for splice isoforms of the protein or individual proteins, respectively.

Using trypsin, or any other sequence specific endoprotease, the surrounding protein sequence should be checked for rugged ends, which are sequences of repeating recognition amino acids, e.g., KK, KR, RK, or RR in the case of trypsin. If possible, peptides with rugged ends should be avoided as both the rugged end and canonically cleaved peptide may coexist and are hard to predict a priori. Rugged end peptides are not part of typical peptide libraries (mentioned in Subheading 2.2.3). However, when unavoidable, rugged end proteotypic peptides can be used for SRM-MS and transitions are obtained as described above in Subheadings 2.1.1/2.1.2/2.1.3/2.1.4. It is advisable to monitor the canonical and rugged end peptide.

2. Interferences

SRM transitions are defined by Q1/Q3 pairs, which represent the *m/z* of the precursor ion and the targeted fragment ion, respectively. Unlike shotgun MS/MS where whole spectra are recorded, SRM-MS heavily relies on the specificity of these two filters as the electron multiplier detector only records the amount of ions present following the “filtration.” The window, or pore size, of these filters is typically *m/z* 0.7; tighter windows might increase specificity, but decrease signal intensity. Peptides to be monitored could theoretically share transitions with other peptides. Practically, in whole cell lysates of complex proteomes, shared transitions are observable, especially for short peptides and fragment ions with low sequence information [24]. The most important denominator is the retention time, as these shared transitions only play a role if they coelute during the peak of the targeted peptide. There is a computational tool predicting transition interferences called SRM Collider [25], which bases its calculations on all theoretically possible fragment ions of a given proteome (without any PTM) as well as all peptides ever detected in PeptideAtlas [26] and predicts retention time using SSR-Calc [27]. The output of SRM Collider should be taken into consideration, e.g., if a certain peptide has a long list of shared transitions, its best to choose another peptide. But, if the list of shared transitions is limited or nonexistent, there is no guarantee that the targeted peptide will be free of interferences. After all due diligence, only the actual LC-SRM-MS experiment will determine if the targeted peptide can be measured and quantified.

There are two ways to detect interferences in complex samples. Generally, a low complexity sample is compared to a

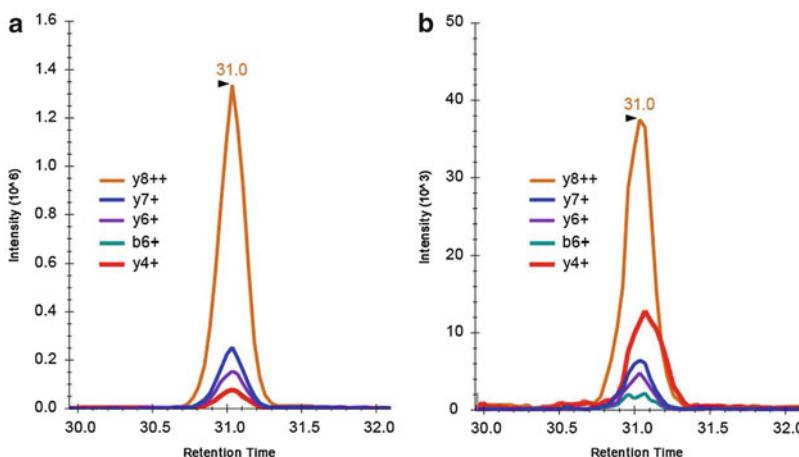


Fig. 2 Ion chromatogram of $\text{NH}_2\text{-IPALDLLIK-COOH}$ in SRM-MS on a triple quadrupole (Thermo TSQ Vantage). (a) Shows the synthetically synthesized peptide with $y8^{++}$ transition as the most intense transition while $y4^+$ is one of the least intense measured transitions. (b) Shows the endogenous measurement of the same peptide, although in a complex cell lysate background. Note that the relative transition order is not conserved between the synthetic less complex peptide and the complex whole cell lysate peptide mixture. Transition $y8^{++}$ is still the most intense however $y4^+$ is now the second most intense transition. This gain in intensity is most likely due to interference, as not only the relative intensity of a single transition changed, but also the peak shape and retention time

high complexity sample. For example, during SRM assay development the sample complexity is relatively low (as described in Subheadings 2.2.3 and 2.2.4) and the relative order of transitions can be compared to the measurements of the same SRM assay in a high complex sample, ideally using the same mass spectrometer. Another way to detect interferences is to use a finite number of stable isotope labeled reference peptides and peptides originating from whole cell lysate and measure both *light* and *heavy* peptide pairs with the same SRM assay. As shown in Fig. 2 there are discrepancies in the relative order of transitions between the low and high complexity measurements due to interference. Transitions with interference should be omitted from quantification and alternative transitions from the peptide's SRM assay be measured.

Further omitted from the SRM transition list should be very small fragment ions, as these carry little sequence information, e.g., a $b2$ fragment ion is comprised of an NH_2 -terminus and two amino acids, frequently resulting in a nondiscriminatory m/z value. As a rule of thumb, the smallest fragment ion for SRM transitions should be $b4$ and $y4$ [24, 28].

3. Relative Retention Time

Crucial to scheduled SRM-MS are standardized retention times, as the window of recording SRM transitions is limited to the time when the peptide is expected to elute from the column. These retention times can be predicted using in silico approaches, such

as SSR-Calc [27]. However, there are discrepancies between in silico predicted retention times of peptides and experimentally observed retention times [29]. Further, the retention time of a peptide was already empirically determined during the initial screening phase (*see* Subheadings 2.2.2/2.2.3/2.2.4), which is a more accurate measure than in silico prediction. In order to universally apply the measured retention times of targeted peptides to various LC-gradients and MS setups, reference peptides are also measured during the same run and relative retention times determined [29].

In detail, the reference retention time peptides span the whole range of the LC-gradient and a linear regression is calculated onto which the targeted peptides are projected. Once these relative values have been established, the same relative retention time peptides can be measured on different LC-MS systems allowing for a robust detection of targeted analytes [29]. Principally there are two ways of establishing relative retention times: using an external standard (multiple vendors sell retention time peptides) or internal standards (using peptides of “housekeeping” proteins). Latter approach has the advantage of not adding any additional peptides to a precious sample, but “housekeeping” proteins might differ depending on tissue type and/or species analyzed.

One such implementation of relative retention times is iRT [29]. There, eleven reference peptides were synthesized which do not share any peptide sequence with currently known proteins and elute off a C18 column during the entire gradient due to their diverse hydrophobic properties. Even measuring only seven reference peptides will decrease retention time precision by only 10 %. These retention time peptides are measured with three transitions each and their retention time is monitored (Table 2).

4. Sample Preparation

The advantage of SRM-MS lies in the fact that targeted proteomics allows for consistent monitoring of analytes from unfractionated lysate. Hence, special care should be taken to obtain peptides well suited for nano-LC systems allowing for multiple subsequent injections of peptide samples. Depending on the species under investigation, e.g., angiosperms or gymnosperms [30], different lysis protocols will be employed [31, 32]. As traditional detergents used for cell lysis are not compatible with LC-MS in general, specialized protocols were established [33, 34] or specialized LC-MS compatible detergents were developed. One such specialized detergent is Sodium3-[(2-methyl-2-undecyl-1,3-dioxolan-4-yl)methoxy]-1-propanesulfonate (World patent number 2005116607), also known under the trade mark RapiGestTM SF (Waters

Table 2
Retention time peptides, their transition and iRT values [29]

| Peptide sequence | Precursor ion | Fragment ions | iRT |
|------------------|----------------|---|--------|
| LGGNEQVTR | 2 ⁺ | y8 ⁺ , y4 ⁺ , y7 ⁺ | -28.31 |
| GAGSSEPVTGLDAK | 2 ⁺ | y8 ⁺ , y6 ⁺ , y10 ⁺ | 0.23 |
| VEATFGVDESNAK | 2 ⁺ | b8 ⁺ , y9 ⁺ , y6 ⁺ | 13.11 |
| YILAGVENSK | 2 ⁺ | y8 ⁺ , y6 ⁺ , y7 ⁺ | 22.38 |
| TPVISGGPYEYR | 2 ⁺ | y8 ⁺ , y9 ⁺ , y7 ⁺ | 28.99 |
| TPVITGAPYEYR | 2 ⁺ | y8 ⁺ , y7 ⁺ , y9 ⁺ | 33.63 |
| DGLDAASYYAPVR | 2 ⁺ | y7 ⁺ , y8 ⁺ , y5 ⁺ | 43.28 |
| ADVTPADFSEWSK | 2 ⁺ | y9 ⁺ , y9 ⁺⁺ , y10 ⁺ | 54.97 |
| GTFIIDPGGVIR | 2 ⁺ | y6 ⁺ , y7 ⁺ , y8 ⁺ | 71.38 |
| GTFIIDPAAVIR | 2 ⁺ | y6 ⁺ , y8 ⁺ , y7 ⁺ | 86.72 |
| LFLQFGAQGSPFLK | 2 ⁺ | y9 ⁺ , y10 ⁺ , y4 ⁺ | 98.09 |

Technologies Ireland Ltd). RapiGest™ SF itself is still not compatible with LC-MS, however, RapiGest™ SF can easily be precipitated by lowering the pH of the solution to pH 2–3. In our hands, protein extraction is more than 10 times more efficient using RapiGest™ SF as compared to 6 M urea/0.1 M ammonium bicarbonate lysis buffer alone.

5. LC-SRM-MS Consideration

Regarding the liquid chromatography gradient, a shorter gradient, e.g., 35 min linear gradient from 5 to 35 % acetonitrile, is typically used for LC-SRM-MS [35]. The rationale behind a shorter gradient is that for a particular peptide the area under the curve is constant, regardless of the length of the gradient. Hence, a shorter gradient will result in sharper and higher peaks versus a longer gradient with wider peaks or elution profile (which is advantageous for data dependent shotgun MS/MS or SWATH-MS [36]). As the peak height is crucial for the signal to rise above the noise, it is advantageous to use a shorter gradient in LC-SRM-MS.

Another consideration regarding short versus long linear LC gradients is the fact that in scheduled SRM a retention time window has to be defined. The size of the window greatly depends on the length of the gradient, for example a 2 min window of a 30 min linear LC gradient is equivalent to a 6 min window of a 90 min linear LC gradient [29]. Hence, the three times longer gradient does not allow for more transitions per injection to be measured and only lowers the peak intensity.

As mentioned in the introduction, cycle time is of essence for (un)scheduled SRM. If a targeted peptide elutes during 20 s of the gradient, a 2 s cycle time will result in approximately 10 positive identifications per peptide. Longer cycle times increase dwell time if the number of transitions are kept constant, but also decrease the number of positive identifications during the elution of the peptide.

An SRM transition is defined as a pair of *m/z* values (Q1/Q3) for precursor ion and fragment ion, respectively. In Q2, the precursor ion is fragmented and the collision energy required for proper fragmentation is a function of the instrument. Typically, the collision energy given for a particular mass spectrometer performs well, but only reflects an average value. One can gain signal increase if the collision energy is optimized [37]. Collision energy optimization is a trade off in terms of how much mass spectrometer time is available, how much time is invested in SRM assay development and how often this optimized collision energy will be used in the future. The resulting optimized collision energy can be unique for each transition. For large scale projects where the targeted list of peptides exceeds 500, we typically do not employ energy optimization and rely on instrument specific collision energy calculations.

6. Data Visualization and Analysis

Data visualization, management and analysis tool of choice currently is Skyline, developed in the MacCoss lab, Department of Genome Sciences, University of Washington [4]. Skyline manages an entire SRM-MS project starting with selecting peptides, using spectral libraries, collision energy optimization options [37], importing all current mass spectrometry vendor instrument files, managing retention time calculators [29] and extensive report function. The program is very well documented and a future development includes integration of MQuest and Mprophet [6]. Besides targeted SRM-MS, Skyline is also suited to import shotgun MS/MS [38] and SWATH-MS data [36].

Skyline has extensive reporting function, e.g., height, area of each transition per peptide which can be used for quantification of peptides (Fig. 3) [28, 39]. For protein significance analysis, SRM Stats is a valuable tool using a family of linear mixed effects models [40] to statistically evaluate a wide range of significant parameters typically used in SRM experiments, e.g., isotopically labeled spike-in reference peptides.

7. Stable Isotope Labeled Reference Peptides

Stable isotope labeling methods are frequently employed to generate reference peptides for relative or absolute quantification of peptides. Dimethyl labeling in this context a cost-efficient, simple, but powerful method for quantification [41].

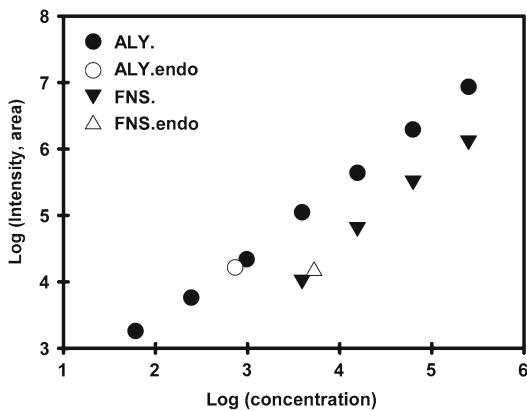


Fig. 3 Dilution series for absolute quantification of endogenous peptide levels. Shown in filled symbols are two dilution series of externally calibrated stable isotope labeled ($^{13}\text{C}_6\text{ }^{15}\text{N}_4$ Arg) reference peptides $\text{NH}_2\text{-ALYDNVAESPDELSPR-COOH}$ (filled circles) and $\text{NH}_2\text{-FNSLNELVDYHR-COOH}$ (filled triangles). The SRM-MS measurements were carried out using a triple quadrupole (Thermo TSQ Vantage) and each measurement, both the exogenous reference peptide and the endogenous peptide were measured. The concentration of endogenous peptides can be inferred from the calibration curve. Together with the number of cells used to obtain the tryptic peptides, absolute values can be determined

Another popular labeling method are stable isotope amino acids $^{13}\text{C}_6\text{ }^{15}\text{N}_2$ Lys or $^{13}\text{C}_6\text{ }^{15}\text{N}_4$ Arg as reference standards for tryptic protein digests. These stable isotope amino acids are favorable, as H to D exchanges affect the C18 reverse phase chromatography of the analytes [42]. On the protein level, $^{13}\text{C}_6\text{ }^{15}\text{N}_2$ Lys or $^{13}\text{C}_6\text{ }^{15}\text{N}_4$ Arg can be used as described in detail by the Brun laboratory [43, 44]. On the peptide level, crude purified $^{13}\text{C}_6\text{ }^{15}\text{N}_2$ Lys or $^{13}\text{C}_6\text{ }^{15}\text{N}_4$ Arg labeled peptides can be added as reference standard to monitor changes of peptides, and therefore proteins, as a function of perturbation. In case of purified externally calibrated (amino acid analysis) stable isotope labeled reference peptides, absolute values can be obtained if the same amount of endogenous lysate is measured in a dilution series of the reference peptide measuring the endogenous/exogenous peptide pair (Fig. 3) [45].

8. Concluding Remarks

Scheduled SRM-MS is a very powerful mass spectrometry method allowing for reliable measurement of analytes [35]. Future developments will undoubtedly increase the dynamic range of mass spectrometers capable of SRM-MS. A limitation of SRM-MS is the limited number of analytes that can be analyzed per LC-SRM-MS run. Using data-independent acquisition of the peptidome, e.g., SWATH-MS [36], has the promise to monitor the peptidome which ionizes well in electrospray ionization mass spectrometry. Also, the development of speedy

shotgun mass spectrometers should not completely be ignored [46], which could record entire spectra of targeted peptides given an extensive inclusion list resulting in a targeted shotgun experiment.

Acknowledgments

HAE is an International Incoming Marie Curie Fellow (FP7) and would like to thank former and current members of the Aebersold lab, ETH Zürich, Switzerland for fruitful discussions.

References

1. Barnidge DR, Dratz EA, Martin T et al (2003) Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. *Anal Chem* 75:445–451
2. Stahl-Zeng J, Lange V, Ossola R et al (2007) High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Mol Cell Proteomics* 6:1809–1817
3. Prakash A, Rezai T, Krastins B et al (2010) Platform for establishing interlaboratory reproducibility of selected reaction monitoring-based mass spectrometry peptide assays. *J Proteome Res* 9:6678–6688
4. MacLean B, Tomazela DM, Shulman N et al (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26:966–968. doi:10.1093/bioinformatics/btq054
5. Rost H, Malmstrom L, Aebersold R (2012) A computational tool to detect and avoid redundancy in selected reaction monitoring. *Mol Cell Proteomics* 11:540–549
6. Reiter L, Rinner O, Picotti P et al (2011) mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat Methods* 8:430–435
7. Cham JA, Bianco L, Barton C et al (2010) MRMaid-DB: a repository of published SRM transitions. *J Proteome Res* 9:620–625
8. Joshi HJ, Hirsch-Hoffmann M, Baerenfaller K et al (2011) MASCP Gator: an aggregation portal for the visualization of *Arabidopsis* proteomics data. *Plant Physiol* 155:259–270
9. Wienkoop S, Staudinger C, Hoehenwarter W et al (2012) ProMEX—a mass spectral reference database for plant proteomics. *Front Plant Sci* 3:125
10. Farrah T, Deutsch EW, Kreisberg R et al (2012) PASSEL: the PeptideAtlas SRMexperiment library. *Proteomics* 12:1170–1175
11. Xia Y, Liang X, McLuckey SA (2006) Ion trap versus low-energy beam-type collision-induced dissociation of protonated ubiquitin ions. *Anal Chem* 78:1218–1227
12. de Graaf EL, Altelaar AF, van Breukelen B et al (2011) Improving SRM assay development: a global comparison between triple quadrupole, ion trap, and higher energy CID peptide fragmentation spectra. *J Proteome Res* 10:4334–4341
13. Glatter T, Wepf A, Aebersold R et al (2009) An integrated workflow for charting the human interaction proteome: insights into the PP2A system. *Mol Syst Biol* 5:237
14. Picotti P, Rinner O, Stallmach R et al (2010) High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. *Nat Methods* 7:43–46
15. Edwards AM, Isserlin R, Bader GD et al (2011) Too many roads not taken. *Nature* 470:163–165
16. Fusaro VA, Mani DR, Mesirov JP et al (2009) Prediction of high-responding peptides for targeted protein assays by mass spectrometry. *Nat Biotechnol* 27:190–198
17. Stergachis AB, MacLean B, Lee K et al (2011) Rapid empirical discovery of optimal peptides for targeted proteomics. *Nat Methods* 8:1041–1043
18. Shimizu Y, Kanamori T, Ueda T (2005) Protein synthesis by pure translation systems. *Methods* 36:299–304
19. Ge X, Xu J (2012) Cell-free protein synthesis as a promising expression system for recombinant proteins. *Methods Mol Biol* 824:565–578

20. Stueber D, Ibrahim I, Cutler D et al (1984) A novel in vitro transcription-translation system: accurate and efficient synthesis of single proteins from cloned DNA sequences. *EMBO J* 3:3143–3148
21. Altelaar AF, Navarro D, Boekhorst J et al (2012) Database independent proteomics analysis of the ostrich and human proteome. *Proc Natl Acad Sci U S A* 109:407–412
22. Lamond AI, Uhlen M, Horning S et al (2012) Advancing cell biology through proteomics in space and time (PROSPECTS). *Mol Cell Proteomics* 11(O112):017731
23. Taouatas N, Mohammed S, Heck AJ (2011) Exploring new proteome space: combining Lys-N proteolytic digestion and strong cation exchange (SCX) separation in peptide-centric MS-driven proteomics. *Methods Mol Biol* 753:157–167
24. Sherman J, McKay MJ, Ashman K et al (2009) How specific is my SRM?: the issue of precursor and product ion redundancy. *Proteomics* 9:1120–1123
25. Rost HL, Malmstrom L, Aebersold R (2012) A computational tool to detect and avoid redundancy in selected reaction monitoring. *Mol Cell Proteomics* 11:540–549. doi:10.1074/mcp.M111.013045
26. Deutsch EW (2010) The PeptideAtlas project. *Methods Mol Biol* 604:285–296
27. Spicer V, Yamchuk A, Cortens J et al (2007) Sequence-specific retention calculator. A family of peptide retention time prediction algorithms in reversed-phase HPLC: applicability to various chromatographic conditions and columns. *Anal Chem* 79:8762–8768
28. Ebhardt HA, Sabido E, Huttenhain R et al (2012) Range of protein detection by selected/multiple reaction monitoring mass spectrometry in an unfractionated human cell culture lysate. *Proteomics* 12:1185–1193
29. Escher C, Reiter L, Maclean B et al (2012) Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* 12:1111–1121
30. Morin RD, Aksay G, Dolgosheina E et al (2008) Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res* 18:571–584
31. Neilson KA, Mariani M, Haynes PA (2011) Quantitative proteomic analysis of cold-responsive proteins in rice. *Proteomics* 11:1696–1706
32. Wang W, Vignani R, Scali M et al (2006) A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. *Electrophoresis* 27:2782–2786
33. Wisniewski JR, Zougman A, Nagaraj N et al (2009) Universal sample preparation method for proteome analysis. *Nat Methods* 6:359–362
34. Bereman MS, Egertson JD, MacCoss MJ (2011) Comparison between procedures using SDS for shotgun proteomic analyses of complex samples. *Proteomics* 11:2931–2935
35. Picotti P, Bodenmiller B, Mueller LN et al (2009) Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. *Cell* 138:795–806
36. Gillet LC, Navarro P, Tate S et al (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* 11, O111016717
37. Maclean B, Tomazela DM, Abbatangelo SE et al (2010) Effect of collision energy optimization on the measurement of peptides by selected reaction monitoring (SRM) mass spectrometry. *Anal Chem* 82:10116–10124
38. Schilling B, Rardin MJ, Maclean BX et al (2012) Platform independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline. Application to protein acetylation and phosphorylation. *Mol Cell Proteomics* 11:202–214
39. Sherrod SD, Myers MV, Li M et al (2012) Label-free quantitation of protein modifications by pseudo-selected reaction monitoring with internal reference peptides. *J Proteome Res* 11:3467–3479
40. Chang CY, Picotti P, Huttenhain R et al (2012) Protein significance analysis in selected reaction monitoring (SRM) measurements. *Mol Cell Proteomics* 11, M111 014662
41. Kovanich D, Cappadona S, Rajmakers R et al (2012) Applications of stable isotope dimethyl labeling in quantitative proteomics. *Anal Bioanal Chem* 404:991–1009
42. Ong SE, Kratchmarova I, 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
43. Dupuis A, Hennekinne JA, Garin J et al (2008) Protein Standard Absolute Quantification (PSAQ) for improved investigation of staphylococcal food poisoning outbreaks. *Proteomics* 8:4633–4636
44. Huillet C, Adrait A, Lebert D et al (2012) Accurate quantification of cardiovascular biomarkers in

- serum using Protein Standard Absolute Quantification (PSAQ) and selected reaction monitoring. *Mol Cell Proteomics* 11, M111-008235
45. Keshishian H, Addona T, Burgess M et al (2007) Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics* 6:2212–2229
46. Michalski A, Damoc E, Lange O et al (2012) Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. *Mol Cell Proteomics* 11, O111.013698

Chapter 17

Mass Spectrometry-Based Imaging of Metabolites and Proteins

Manuela Peukert, Michael Becker, Andrea Matros, and Hans-Peter Mock

Abstract

Imaging techniques based on mass spectrometry (MS) have become powerful approaches to decipher the spatial distribution of metabolites and proteins. MS imaging (MSI) mostly relies on matrix-assisted laser desorption/ionization coupled to MS detection, but desorption electrospray ionization is also frequently used. Here we describe our current protocols for MALDI-MSI of seed sections and for root tissue. Detailed procedures for cryo-sectioning, matrix application, image capture, mass spectrometry measurement and data analysis are given.

Key words Matrix-assisted laser desorption/ionization mass spectrometric imaging (MALDI-MSI), Barley grains, Tobacco roots, Small molecules, Proteins

1 Introduction

MS-based imaging of tissues has found widespread applications in medicinal and pharmaceutical studies [1–4]. The approach provides information about the spatial distributions of small molecules and proteins within tissues. It has been used to study molecular features associated with lung tumors [5], gliomas [6] or with breast cancer [7] and the distribution of drug candidates [8].

The prevalent technique in MS-based imaging involves the use of a MALDI source. Tissue sectioning and the application of a suitable matrix are fundamental steps prior to the sampling of MS data. Both processes are essential to obtain images with high resolution. Tissue sectioning has to be adapted for specific tissues. The selection of a matrix substance will strongly influence the MS profile, and hence a range of matrices has to be tested to obtain sensitivity for different classes of compounds. Image capture of larger tissue sections in high resolution mode generates a vast amount of spectra, and tools for evaluation are central in the processing of the data sets.

Only recently, MS-based imaging techniques have been applied for the analysis of plant samples [9–12].

Here, we present our current protocols for detecting (a) small molecules (b) proteins and peptides in cryo-dissected immature barley grains (a, b) and tobacco roots (a). We will describe the critical steps in detail allowing to reproduce the experiments and to guide the set-up of MALDI imaging for other tissues. The standard workflow for MALDI mass spectrometry-based imaging of molecular compounds in biological tissues is schematized in Fig. 1.

2 Materials for MSI of Small Molecules

2.1 Materials for Sample Preparation, Cryo-sectioning, and Slide Preparation

1. Cryostat (we are using a Leica CM3050, Leica Microsystems GmbH Wetzlar, Germany), brush, forceps, razor blade.
2. Liquid nitrogen, optimal cutting temperature (OCT) medium, water.
3. Indium tin oxide (ITO) glass slides (Bruker Daltonics GmbH, Bremen, Germany).
4. Desiccator and vacuum pump.
5. Whiteout solution (Laco-office products Finke GmbH, Sottrum, Germany).
6. Stereomicroscope (e.g., Leica MZ6,) connected to a digital camera (AxioCam ICc1, Zeiss, Jena, Germany) or a similar microscopic image acquisition system (see Note 1).

2.2 Materials for Matrix Application of Small Molecules MSI

1. ImagePrep device (Bruker Daltonics GmbH, Bremen, Germany; see Note 2).
2. Dilutions of particular matrices were prepared according to the instructions for the ImagePrep. 2,5-Dihydroxybenzoic acid (DHB) was diluted to 30 mg/ml in 50 % (v/v) methanol with 0.2 % (v/v) trifluoroacetic acid (TFA), and α -cyano-4-hydroxycinnamic acid (HCCA) to 7 mg/ml in 60 % (v/v) acetonitrile with 0.2 % (v/v) TFA.

2.3 Software and Materials for MALDI-MS Image Acquisition

1. flexControl (Bruker Daltonics GmbH, Bremen, Germany).
2. flexImaging (Bruker Daltonics GmbH, Bremen, Germany).
3. MTP Slide-Adapter II (Bruker Daltonics GmbH, Bremen, Germany).
4. 1:1 Mixture of PEG200 and PEG600, diluted to 1:300 in TA30 solution (30 %, v/v, acetonitril, 0.1 %, v/v, TFA).

2.4 Software for MALDI-MS Imaging Data Analysis

1. flexImaging (Bruker Daltonics GmbH, Bremen, Germany).
2. ClinProTools (Bruker Daltonics GmbH, Bremen, Germany).

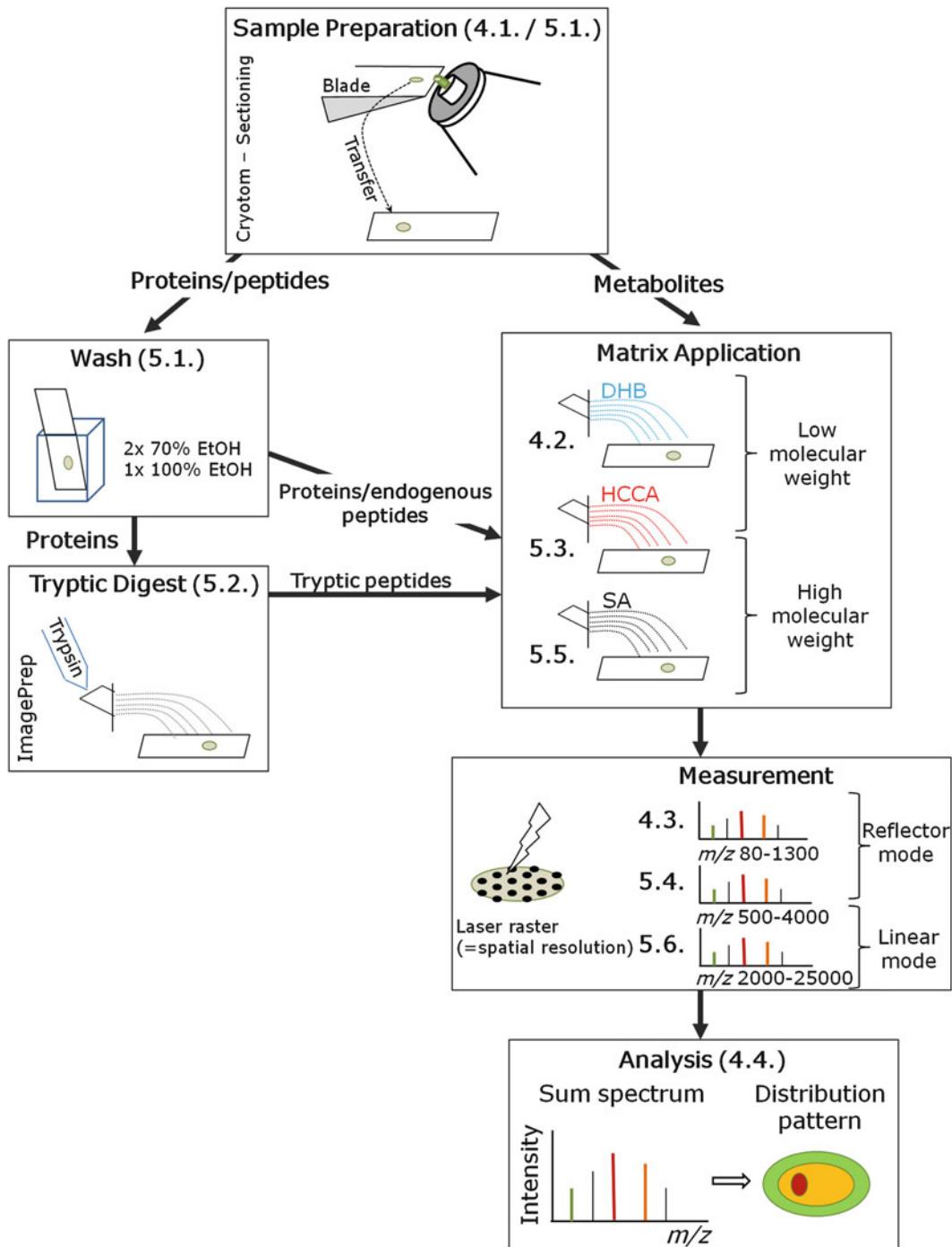


Fig. 1 Workflow for MALDI mass spectrometry-based imaging of molecular compounds in biological tissues. *Bold numbers* represent sections of this chapter

3 Materials for MSI of Proteins/Peptides

3.1 Materials for Sample Preparation, Cryo-sectioning, and Slide Preparation

3.2 Materials for Matrix Application of Protein/Peptide MSI

3.3 Software for MALDI-MS Image Acquisition

3.4 Software for MALDI-MS Imaging Data Analysis

Sample preparation is performed in the same way as stated in Subheading 5.1 using the same material as indicated in Subheading 2.1.

1. ImagePrep device (Bruker Daltonics GmbH; *see Note 2*).
2. Dilutions of particular matrices were prepared according to the instructions for the ImagePrep. Sinapinic acid (SA) was diluted to 10 mg/ml in 60 % (v/v) acetonitrile with 0.2 % (v/v) trifluoroacetic acid (TFA). α -Cyano-4-hydroxycinnamic acid (HCCA) to 7 mg/ml in 60 % (v/v) acetonitrile with 0.2 % (v/v) TFA.

MS image acquisition is performed in the same way as stated in Subheadings 5.4 (peptides) and 5.6 (proteins) using the same software packages as indicated in Subheading 2.3.

Data analysis is performed in the same way as stated in Subheading 4.4 using the same software packages as indicated in Subheading 2.4.

4 Methods for MALDI-MSI of Small Molecules

4.1 Sample Preparation, Cryo-sectioning, and Slide Preparation

1. Clean the sectioning blade or put a new one into the cryostat and adjust the cutting temperature. For most samples, the temperature is set to -20 °C. If the non-embedded part of the sample tends to be still flexible during the cutting process, lower temperatures should be used. Otherwise, for highly fragile samples test higher temperatures (i.e., -10 °C) (*see Note 3*).
2. Fresh plant material has to be frozen in liquid nitrogen. Transfer the frozen sample into the cryostat chamber cooled to the desired temperature (-20 °C) and fix it onto the sample plate via either a droplet of water (ice) or of OCT medium (*see Note 4*). If necessary, pre-cut the plant material with a razor blade to a suitable length before fixation. Select a proper orientation for longitudinal or cross sections.
3. Start the sectioning process by moving the cutting block with the sample plate. To trim the tissue sample a large section thickness can be selected until the desired morphological parts are visible. Then change the setting to the desired thickness. Cut 20–30 μ m sections for seeds and 35–45 μ m for tobacco roots (*see Note 5*). Transfer the fine sections immediately onto indium tin oxide

(ITO) coated glass slides and fix the sections by thaw-mounting. By touching the opposite glass side with a fingertip for ca 20–30 s the ice in the section is melted and evaporated.

4. Transfer the ITO slides with the sections into a desiccator for 0.5–1.0 h until they are completely dried. The time until complete dryness varies and is longest for the high water content tissues (roots). Alternatively, for low water containing samples, the drying step can be skipped.
5. Use whiteout solution to place marks close to the samples for definition of sample position in the flexImaging software before starting the measurement run (*see* Subheading 2.1, item 5f).
6. Capture an optical image with a slide scanner or a stereomicroscope (e.g., Leica MZ6, Wetzlar, Germany) connected to a digital camera (AxioCam ICc1, Zeiss, Jena, Germany) or a similar image acquisition system (*see* Note 1).

4.2 Matrix Deposition for MALDI-MSI of Small Molecules

The following procedures have been optimized for the automated deposition of matrix onto cryo-sections using the ImagePrep device (Bruker Daltonics GmbH, Bremen, Germany; *see* Note 2).

1. Fill the matrix vial with pure methanol and set the global spray power adjustment for the instrument as described in the ImagePrep manual.
2. Insert the slide with the dried sample into the ImagePrep instrument and adjust slide position on the small elevation in the way that the light sensor isn't covered by a section. Fill an appropriate amount of matrix solution (7–8 ml) into the reservoir of the spraying device.
3. Apply matrix solutions via sensor-controlled vibrational vaporization, utilizing the ImagePrep device (Bruker Daltonics GmbH) as per the manufacturer's instructions (*see* Note 6). Different spray protocols specific for the particular matrices are provided.
4. Immediately after matrix application clean the spray device with methanol.

4.3 MALDI-MSI Measurement of Small Molecules

Methods are described for Bruker ultraflexXtreme instrumentation (Bruker Daltonics GmbH), but are generally applicable for Bruker MALDI-TOF instruments.

1. Fit the freshly prepared slides into a Slide Adapter II MALDI target and introduce the adapter into the source of the MS instrument.
2. Select the required flexControl method. For small molecule's MSI measurement select an m/z range of 80–1,300 and a sample

- rate of 0.5 Gs/s (corresponding to 20,000–25,000 data points per spectrum). Make sure there is enough space on the hard disk.
3. Perform mass calibration using a 1:1 mixture of PEG200 and PEG600, diluted to 1:300 in TA30 solution (30 %, v/v, acetonitril, 0.1 %, v/v, TFA) deposited next to your sample on the slide.
 4. Adjust laser power and number of laser shots needed to a relative signal intensity of 25,000 by checking signal intensity from various regions of the tissue section.
 5. Create a file for MSI measurements using flexImaging software (Bruker Daltonics GmbH). Here, the optical image of the sample (Subheading 4.1, step 6) is needed to align sample stage position and laser spot focus of the instrument. Designate the measurement region and select desired spatial resolution. For very small sections (such as young seed cross sections) use 15–20 μm laser raster in order to get sufficient representation of the various tissues (see Note 7). The autoXecute method chosen defines i.e., number of laser shots and the flexControl method. Additionally, mark a small measurement region with just matrix as blank of approximately 25–35 laser raster points.
 6. Perform MSI measurement. Carry out the measurement of the spots in random order to eliminate influence of measurement order (This will result in longer acquisition time compared to FAST modus. Necessity should be tested depending on tissue properties and final number of single spectra acquired in one imaging run.)

4.4 Data Analysis for Small Molecule MALDI-MSI

Methods are described for using flexImaging software.

1. FlexImaging software is used for distribution analysis of selected m/z signals. Not to include matrix derived signals, overlay sum spectra of the tissue measurement region and of the blank region (see Note 8). Avoid selection of m/z values that overlay completely with blank derived signals. It is possible to create a mass filter list with manually selected signals.
2. Using the mass filter list option multiple ion intensity maps can be created. Colors for distinct mass signals to be displayed in the multiple ion intensity maps should be chosen as the distribution patterns of the individual single signals are then clearly observable. In addition, individual ion intensity maps can be visualized for representation of particular molecular ion abundances.
3. With the “Region of interest” (ROI) option different parts of the measured sample can be chosen for comparison of sum spectra according to these specific regions. For statistical analysis the spectra lists of the ROI’s can be exported for further analysis using e.g., ClinProTools software (see Note 9).
4. Targeted analysis for identification of compounds.

The image analysis provides a list of *m/z* signals relevant in the context of each specific analysis. At this point, additional experiments are required to elucidate the chemical structure of the compounds. These include accurate mass measurements, MS/MS analysis, on-tissue digestion, or extraction and analysis of laser-dissected areas containing the unknown compound of interest among other approaches. The choice of an appropriate strategy will depend on the specific experimental context.

5 Methods for MALDI-MSI of Proteins and Peptides

5.1 Sample Preparation, Cryo-sectioning, and Slide Preparation

1. Clean the sectioning blade or put a new one into the cryostat and adjust the cutting temperature. For most samples, the temperature is set to $-20\text{ }^{\circ}\text{C}$. If the non-embedded part of the sample tends to be still flexible touching the blade during sectioning process, lower temperatures should be used. Otherwise, for highly fragile samples test higher temperatures (i.e., $-10\text{ }^{\circ}\text{C}$) (see **Note 3**).
2. Fresh plant material has to be frozen in liquid nitrogen. Transfer the frozen sample into the cryostat chamber cooled to the desired temperature ($-20\text{ }^{\circ}\text{C}$) and fix it onto the sample plate via either a droplet of water (ice) or of OCT medium (see **Note 4**). If necessary, precut the plant material with a razor blade to a suitable length before fixation. Select a proper orientation for longitudinal or cross sections.
3. Start the sectioning process by moving the cutting block with the sample plate. To trim the tissue sample a large section thickness can be selected until the desired morphological parts are visible. Then change the setting to the desired thickness. Cut $20\text{--}30\text{ }\mu\text{m}$ sections for seeds and $35\text{--}45\text{ }\mu\text{m}$ for tobacco roots (see **Note 5**). Transfer the sections immediately onto indium tin oxide (ITO) coated glass slides and fix the sections by thaw-mounting. By touching the opposite glass side with a fingertip for ca $20\text{--}30\text{ s}$ the ice in the section is melted and evaporated.
4. Transfer the ITO slides with the sections into a desiccator for $0.5\text{--}1.0\text{ h}$ until they are completely dried. The time until complete dryness varies and is longest for the high water content tissues (roots). Alternatively, for low water containing samples, the drying step can be skipped.
5. After drying, samples must be subjected to a washing procedure (see **Note 10**). Immerse the slides in 70 % ethanol (two times for 2 min each) and absolute ethanol (once for 2 min). After washing, remove excess ethanol from the slides carefully with a paper towel and return the slides to the desiccator for another $0.5\text{--}1\text{ h}$, until they have completely dried.

6. Use whiteout solution to place marks close to the samples for definition of sample position in the flexImaging software before starting the measurement run (*see* Subheading 5.4).
7. Capture an optical image with a slide scanner or stereomicroscope (e.g., Leica MZ6, Wetzlar, Germany) connected to a digital camera (AxioCam ICc1, Zeiss, Jena, Germany) or a similar image acquisition system (*see* Note 1).

5.2 Spatially Resolved Tryptic Digest (Optional)

1. Reconstitute one vial (20 µg) of trypsin (Sequencing Grade Modified Trypsin, Promega, WI, USA) using 180 µl of ammonium bicarbonate (20 mM) and 20 µl acetonitrile. Avoid foaming.
2. Carefully adjust the global spray power offset of the ImagePrep. Water or ammonium bicarbonate solution (20 mM) can be used for training purposes. Using a spray power of 38 % with no modulation, the offset should be adjusted until a faint but continuous spray for 20–40 s can be achieved using a single 200 µl droplet. To achieve this, the 200 µl droplet must be directly loaded into the flow pipe of the image prep, and not in the matrix reservoir. Detailed information is available from the manufacturer.
3. Load 200 µl of trypsin solution as described under Subheading 2, and place the sample in the ImagePrep, positioning it on the central elevated platform. The scattered light sensor should not be covered by tissue. Execute the trypsin deposition protocol (*see* Note 11). Thoroughly clean the ImagePrep using the “Clean” option of the instrument with pure methanol.
4. After trypsin deposition, carefully remove the slide from the ImagePrep and place it in a sufficiently large container filled with several wet paper towels. Humidity should be high enough to ensure efficient digestion, but there should be no water condensation on the ITO slide. Incubate the container with the sample at 37 °C for 90 min.
5. After incubation, remove the slide from the container, return it into the ImagePrep and follow the standard procedure for depositing matrix on the slide.

5.3 Matrix Deposition for the Analysis of Tryptic Peptides

The following procedures have been optimized for the automated deposition of matrix onto cryo-sections using the ImagePrep device (Bruker Daltonics GmbH, Bremen; *see* Note 2).

1. Fill the matrix vial with pure methanol and set the global spray power adjustment for the ImagePrep as described in the manual.
2. Insert the slide with digested sample into the ImagePrep instrument and adjust slide position on the small elevation in the way that the light sensor isn't covered by a section.

Place a glass coverslip on the ITO slide, directly covering the sensor but not the sample.

3. Fill an appropriate amount of matrix (7–8 ml) solution into the reservoir of the spraying device. Either DHB or HCCA matrix can be used. To compensate for the buffer capacity of the ammonium bicarbonate in the trypsin solution, the TFA content of any matrix used should be increased to 1 % (v/v).
4. Apply matrix solutions via sensor-controlled vibrational vaporization, utilizing the ImagePrep device (Bruker Daltonics GmbH) as per the manufacturer's instructions (*see Note 6*). Different spray protocols specific for the particular matrices are provided.
5. Immediately after matrix application clean the spray device with methanol.

5.4 MALDI-MSI Measurement of Tryptic Peptides

Methods are described for Bruker ultrafleXtreme instrumentation (Bruker Daltonics GmbH), but are generally applicable for Bruker MALDI-TOF instruments.

1. Fit the freshly prepared slides into a Slide Adapter II MALDI target and introduce the adapter into the source of the MS instrument.
2. Select the required flexControl method. For tryptic peptides, use reflector mode in positive polarity and an m/z range of 500–4,000, and a sampling rate of 0.5 Gs/s (corresponding to 20,000–25,000 data points per spectrum). Make sure there is enough space on the hard disk.
3. Perform mass calibration using a peptide standard (e.g., Peptide Calibration Standard II, Bruker Daltonics GmbH), mixed 1:1 with the same matrix used for sample preparation deposited next to your sample on the slide.
4. Adjust laser power and number of laser shots needed to a relative signal intensity of 25,000 by checking signal intensity from various regions of the tissue section (*see Note 12*).
5. Create a file for MSI measurements using flexImaging software (Bruker Daltonics GmbH). Here, the optical image of the sample (Subheading 4.1, step 6) is needed to align sample stage position and laser spot focus of the instrument. Designate the measurement region and select the desired spatial resolution. Typically, the spatial resolution of tryptic digests is limited to raster widths of 50 μm or higher. The autoXecute method chosen defines i.e., number of laser shots and the flexControl method (*see Note 13*).
6. Perform MSI measurement. Carry out the measurement of the spots in random order to eliminate influence of measurement order.

5.5 Matrix Deposition for the Analysis of Intact Proteins

1. Fill the matrix vial with pure methanol and set the global spray power adjustment for the ImagePrep as described in the manual
2. Insert the slide with digested sample into the ImagePrep instrument and adjust slide position on the small elevation in the way that the light sensor isn't covered by a section. Place a glass coverslip on the ITO slide, directly covering the sensor but not the sample
3. Fill an appropriate amount of matrix (7–8 ml) solution into the reservoir of the spraying device. Either SA or HCCA matrix is recommended for the analysis of intact proteins (*see Note 14*).
4. Apply matrix solutions via sensor-controlled vibrational vaporization, utilizing the ImagePrep device (Bruker Daltonics GmbH) as per the manufacturer's instructions (*see Note 6*). Different spray protocols specific for the particular matrices are provided.
5. Immediately after matrix application clean the spray device with methanol.

5.6 MALDI-MSI Measurement of Intact Proteins

Methods are described for Bruker ultrafleXtreme instrumentation (Bruker Daltonics GmbH), but are generally applicable for Bruker MALDI-TOF instruments.

1. Fit the freshly prepared slide into a Slide Adapter II MALDI target and introduce the adapter into the source of the MS instrument.
2. Select the required flexControl method. For intact proteins, use linear mode in positive polarity and a mass range of 2,000–25,000 Da with a sampling rate of 0.1 Gs/s (corresponding to ca. 10,000 data points per spectrum). Make sure there is enough space on the hard disk.
3. Perform mass calibration using a protein standard (e.g., Protein Calibration Standard I, Bruker Daltonics GmbH), mixed 1:1 with the same matrix used for sample preparation deposited next to your sample on the slide.
4. Adjust laser power and number of laser shots needed to a relative signal intensity of 25,000 by checking signal intensity from various regions of the tissue section.
5. Create a file for MSI measurements using flexImaging software (Bruker Daltonics GmbH). Here, the optical image of the sample (Subheading 4.1) is needed to align sample stage position and laser spot focus of the instrument. Designate the measurement region and select desired spatial resolution. The autoXecute method chosen defines, i.e., number of laser shots and the flexControl method. Additionally, mark a small measurement region with just matrix as blank of approximately 25–35 laser raster points.

6. Perform MSI measurement. Carry out the measurement of the spots in random order to eliminate influence of measurement order.

Data analysis is performed in the same way as stated in Subheading 4.4.

6 Distribution Pattern of Small Molecules in Developing Barley Grains

Kernels of developing grains are composed of several different tissue types undergoing massive changes during the first 3 weeks after pollination. The MALDI MSI approach is used to unravel distribution pattern of compounds showing tissue and developmental specificity. Here, the advantages of this MS method are obvious as dissection and micro-extraction procedures for these minute amounts of tissues would require extreme efforts. Kernels of different time points, representing major developmental stages are analyzed. In our analyses several detected m/z values were found to be highly related to distinct tissues. The multiple ion intensity map in Fig. 2a demonstrates accumulation of five individual molecular ions in different seed regions. As mentioned in Subheading 4.4, statistical analysis is performed using ClinProTools software. For this purpose regions of interest (ROIs) were manually defined as indicated in Fig. 2b and subsequently used for statistical analysis. The PCA result for eight ROIs and 131 individual m/z values shows high correlation of molecular ions to distinct ROIs, displayed by defined groups (Fig. 2c). These results clearly illustrate the ability of the presented MALDI MSI approach to significantly assign distribution pattern of molecular ions to defined plant tissues.

7 Notes

1. Regarding optical images of the sample, one has to record low resolution images (<10,000 dpi) which provide a reasonable overview of the sample and can serve as a reference image to define measurement areas. Most scanners provide reasonable images, but in our hands an OpticLab H850 histological slide scanner (Plustek Technology GmbH, Germany) proved very useful. For actual histological evaluation, low resolution images are typically not sufficient. Although high-resolution images from digital cameras and microscopes can be co-registered with the MSI results, their file size often prevents efficient use. Here, microscope slides scanners such as the Pannoramic Desk (3D Histech, Hungary) or ScanScope CS (Aperio, CA, USA) provide an elegant alternative, automatically scanning the slide at several resolution levels in an approach similar to the one used

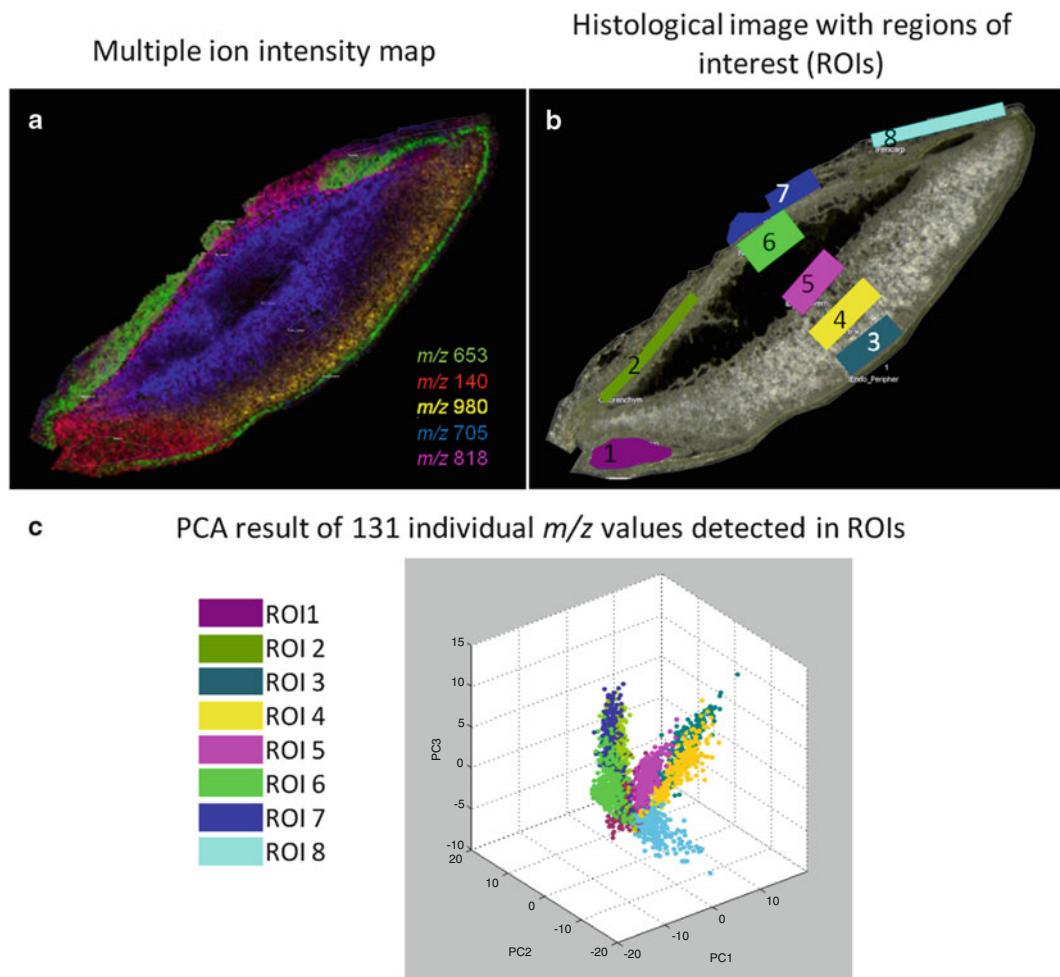


Fig. 2 MALDI MSI analysis of a longitudinal barley grain section, 14 days after pollination. **(a)** Multiple ion intensity map displays distribution of five distinct m/z values. **(b)** Histological image indicating positions of ROIs selected for statistical analysis in CPT. **(c)** Result of PCA analysis of eight ROIs (as given in **b**) including 131 individual m/z values

in Google Maps. This proprietary image formats are compatible with co-registration in the flexImaging software.

2. The ImagePrep apparatus from Bruker Daltonics creates a fine matrix aerosol in a spray chamber filled with nitrogen. The aerosol is applied onto sample sections during several cycles of deposition, incubation and drying. An optical sensor controls the different phases of each cycle. With this technique small droplet sizes and reproducible conditions for sample preparation are provided. Alternative ways for matrix application are the usage of nano-spotting devices or manual pneumatic spray devices. While with nano-spotting devices lateral resolution is limited to at least 200 μ m, the process of manual spraying is

irreproducible. Some efforts have been made for solvent-free approaches, for example via sublimation, but are currently not applied for large scale analysis. Especially, for small molecule MSI experiments carried out on little sample sizes a reproducible application technique with small droplet sizes as well as high lateral resolution is essential.

3. As a simple rule, sample temperature is usually tied to the fat content of the tissue in question. Low fat tissue is usually cut at temperatures around -15°C , whereas highly fatty samples may require temperatures of -25°C or lower. When cutting, observe the behavior of the section. Sticky sections that cling to the blade or the anti-roll bar usually indicate the temperature should be lowered, while brittle and rupturing sections indicate that the temperature is too cold. Plant tissues, with their usually high water content, pose additional challenges when freezing and may require the use of cryoprotectants. For protein analysis of soybean cotyledons, we achieved good results using sucrose solution (20 % w/v). Prior to freezing, the cotyledons were placed in a syringe filled with sucrose solution and negative pressure was applied, soaking the tissue with the cryoprotectant. However, this method is not applicable for the analysis of small molecule distributions as allocation of the molecular compounds may likely be changed and signal intensity of analytes will be suppressed by the very high sucrose level.
4. For the preparation of sections the use of polymeric resins should be avoided, as these compounds will provide a strong background and suppress signals of the tissue's analytes during the subsequent MS imaging procedure. Most commonly used embedding media, such as OCT or paraffin contain polymers, which can be washed from the samples using a combination of water and/or alcohol washes. For example, dipping the sample in warm, de-ionized water is usually sufficient to wash out OCT, and paraffin can be removed by washing in xylenes for 10–15 min. While larger molecular weight peptides and proteins usually do not suffer from extensive delocalization under these washing conditions, more soluble analytes are usually washed out from the tissue, so polymer-embedded tissue is usually not suitable for the analysis of lipids or other small molecules. Alternative embedding media are methyl cellulose, sucrose solution, gelatine, or simply water in case of protein analysis. In contrast for very small, delicate samples, we have had good success using poached egg white as an embedding material for metabolite analysis. A small incision is made in the egg white into which the sample is inserted and then both are frozen in liquid nitrogen. The resulting, styrofoam-like product can be easily cryo-sectioned and creates no noticeable background signal.

5. The optimal thickness has to be determined for each specimen. In general, sections of 8–12 μm thickness are ideal. From an analytical perspective, an increased thickness of the sample represents a deviation from the optimal focus plane of the MALDI laser and has negative effects on the focus of the resulting ion beam, resulting in a loss of sensitivity. On MALDI-TOF instruments in particular, an irregular depth profile of the sample can also induce electrical charging of the instrument source and *m/z*-shifts in the recorded spectra. In addition, thicker sections have an increased risk to detach from the slide during or after vacuum drying. However, a thickness below 15 μm is difficult to obtain with non-embedded plant material, and 20–25 μm sections are quite common. Tobacco root sections have to be thicker than those of seeds to avoid disruption due to the large intercellular spaces in the parenchymal cortex. Essentially, a practical trade-off has to be made between morphological integrity of the sample and the quality of the resulting mass spectra. Thickness has a clear impact on the sensitivity of image analysis. For tobacco roots, a slice thickness above 45 μm led to a notable decrease in sensitivity. Same holds true for a slice thickness above 40 μm in case of barley grains. For young (1–7 DAP) barley grains, 20 μm slices were effective, but as the grain developed, the accumulating starch (associated with a loss of the water content) forced the section thickness to be increased to 30 μm . For some very delicate samples (such as thin leaves or small roots) directly mounting the intact sample onto an ITO slide is also an option. Double sided conductive tape, such as Plano 16081-2 (Plano GmbH, Germany) can be used to maintain conductivity.
6. The optical sensor of the ImagePrep monitors wetness and thickness of matrix by voltage differences of the sensor. As spray power during an application run can decrease, an automatic spray power boost is implemented. Additionally, it is useful to visually control spray application, especially if high resolution images are desired as automatic spray power boost can result in less uniform spray coverage and higher degree of wetness.
7. The size of measurement region and chosen laser raster resolution defines the number of acquired single spectra. A very high spectra number in combination with high number of laser shots per raster spot can result in a significant decrease in spectra intensity due to pollution of the MALDI source. We have generally chosen a maximum of spectra per run of ca. 15,000.
8. Small mass range analysis has the particular disadvantage of the large number of background noise induced by fragmentation and cluster formation of organic matrices during gas-phase ionization. To elucidate this signals a small measurement

region for background acquisition is located next to the tissue section and randomly measured along with the sample. After the MSI run sum spectra of both regions are compared with each other. Signals that totally overlay with matrix background are excluded from further analysis as it is not possible to distinguish if signals are just matrix derived or if other compounds sharing the same molecular mass are additionally present. Furthermore, isolation of such an m/z value for MS/MS measurements carried out directly from the tissue section is not possible without getting matrix fragmentation pattern. Figure 3 shows a mass spectrum from a barley seed methanol

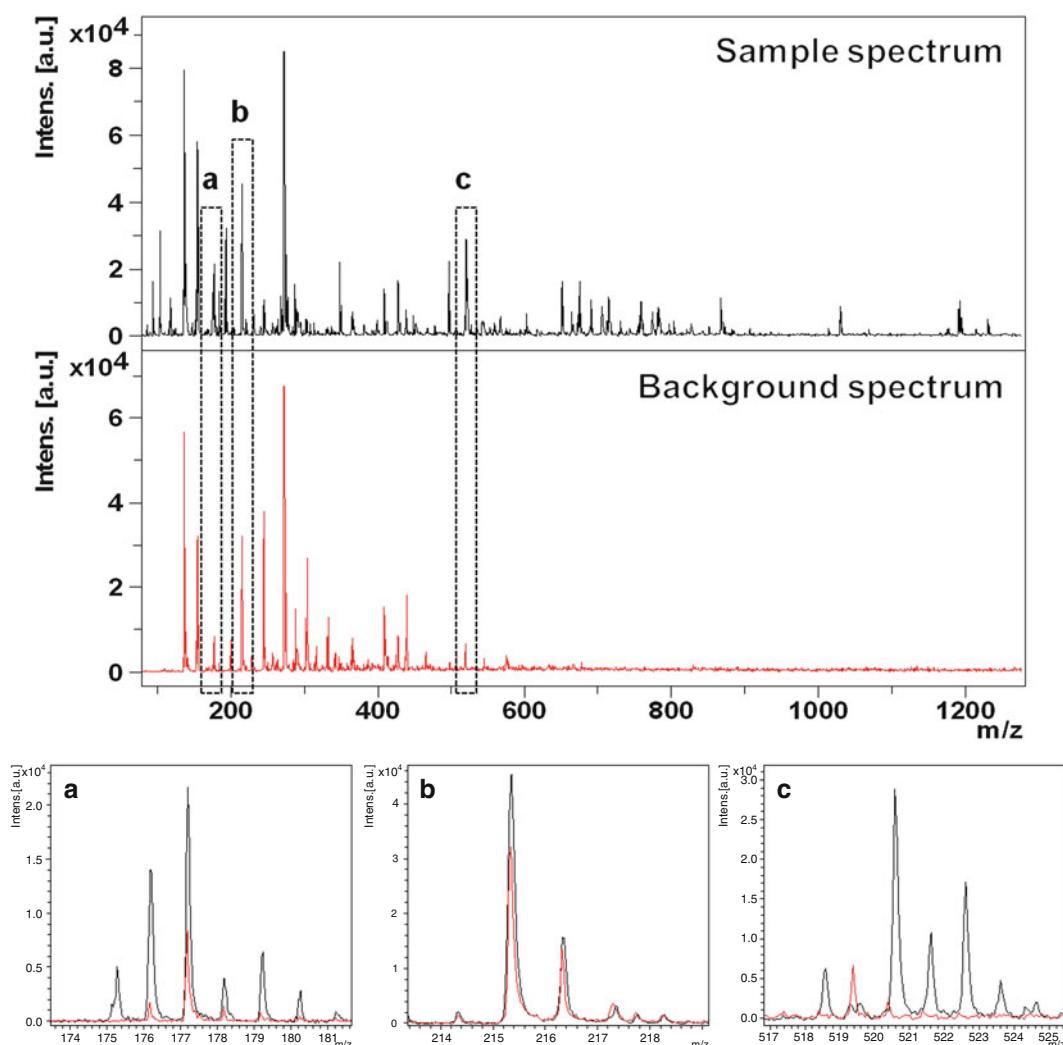


Fig. 3 Sample spectrum of a barley methanol extract (black) compared with DHB derived background signals (red). Measurements were carried out in positive ionization mode using a MALDI TOF/TOF instrumentation (Bruker Daltonics, Bremen, Germany). The lower spectra show overlays in the designated m/z regions (color figure online)

extract (black) with DHB as matrix compound compared to the background spectrum derived from DHB alone (red). This matrix produces very intense background signals in the *m/z* range 100–600 in positive ionization mode (Fig. 3). Therefore, peak selection has to be carried out manually.

9. The number of exported single spectra for analysis by ClinProTools (CPT) has to be reduced to ca. 2,000 if no data reduction in CPT is performed. High rate of data reduction is associated in loss of peak resolution and has therefore to be adapted to the number of spectra and to the kind of analysis. Another possibility for reduction of data amount is to reduce mass range to targeted mass signals.
10. The analysis of peptides and proteins is usually hampered by the presence of other analytes, mainly abundant membrane phospholipids, in the tissue. These lower mass lipids are highly abundant and typically ionize readily. Due to competition for ionization, the presence of lipids and other small molecules usually suppresses the ionization of larger molecules, such as proteins. The ethanol washes suggested here remove most of the lipids and other small molecules, enabling the detection of proteins. Both the solvents and the duration of the washing steps can be modified to optimize for a particular workflow. For example, 2 min immersion as described above is useful before a tryptic digest of the tissue, as it will remove endogenous peptides which could otherwise interfere with the ionization of tryptic peptides. Other, more stringent washing procedures have been suggested, such as the inclusion of isopropyl alcohol or chloroform [13, 14]. Shorter washing steps potentially allow detection of smaller peptides, at the expense of the detection of higher mass proteins. Typically, detection of proteins is limited to masses of 20–30 kDa due to similar ion suppression effects, but specific washing procedures have been suggested to allow detection of proteins up to 70 kDa [15].
11. Similar to the matrix deposition step, the size of the trypsin droplets deposited on the tissue surface has direct impact on the possible spatial resolution of the resulting peptide images. Small trypsin droplets limit the delocalization of the peptides but also negatively affect the efficiency of the enzymatic digestion. The ImagePrep protocol is designed to deposit the total amount of enzyme in a series of 10–20 short (~1 s) spray bursts followed by immediate drying with nitrogen. The setup procedure described will ensure that the instrument is set up correctly to achieve a total of ~20 s spray time with a volume of 200 μ l. Larger volumes of trypsin (at lower concentration) are not recommended, as the ammonium bicarbonate buffer has negative effects on the tissue integrity.

12. The general appearance of the spectra should be checked. Typically, the majority of tryptic peptides have masses of \sim 1,500–2,000 Da and appear in the spectrum as a Gaussian-like distribution of intensities centered around m/z \sim 1,500. Lots of larger peaks may indicate incomplete digestion, whereas many smaller peaks may indicate too long incubation time (leading to nonspecific cleavage), especially if many trypsin autolysis peaks can also be seen. In addition to the incubation time, the concentration of ammonium bicarbonate in the buffer represents another parameter that can be optimized for particular tissue types. Trypsin operates at a slightly alkaline pH optimum, so for samples with a low pH, a higher concentration of ammonium bicarbonate (up to 80 mM) can be helpful.
13. It may be helpful to cover at least a part of the analyzed tissue with a coverslip before the application of trypsin. The coverslip should be removed before matrix application. This area can be measured as an internal control to evaluate the success of the digestion.
14. The primary choice of matrix for the measurement of proteins is usually SA, which usually provides good spectra quality for proteins of up to 25 kDa even from complex mixtures such as tissue samples. The main drawback is the size of matrix crystals, which is commonly in the range of \sim 50 μ m and limits the spatial resolution of the MALDI image. If higher spatial resolution is desired, HCCA matrix can be used instead, which forms smaller crystals. Spatial resolution of up to 20 μ m for MSI of proteins has been achieved using HCCA [16].

References

1. Caprioli RM, Farmer TB, Gile J et al (1997) Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* 69:4751–4760
2. Rohner TC, Staab D, Stoeckli M et al (2005) MALDI mass spectrometric imaging of biological tissue sections. *Mech Ageing Dev* 126: 177–185
3. Burnum KE, Kristin E, Frappier SL et al (2008) Matrix-assisted laser desorption/ionization imaging mass spectrometry for the investigation of proteins and peptides. *Annu Rev Anal Chem* 1:689–705
4. van Hove ERA, Smith DF, Heeren RMA et al (2010) A concise review of mass spectrometry imaging. *J Chromatogr A* 1217:3946–3954
5. Yanagisawa K, Shyr Y, Xu BGJ et al (2003) Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* 362:433–439
6. Schwartz SA, Weil RJ, Johnson MD et al (2004) Protein profiling in brain tumors using mass spectrometry: Feasibility of a new technique for the analysis of protein expression. *Clin Cancer Res* 10:981–987
7. Cornett DS, Mobley JA, Dias EC et al (2006) A novel histology-directed strategy for MALDI-MS tissue profiling that improves throughput and cellular specificity in human breast cancer. *Mol Cell Proteomics* 5: 1975–1983
8. Reyzer ML, Hsieh YS, Ng K et al (2003) Direct analysis of drug candidates in tissue by matrix-assisted laser desorption/ionization mass spectrometry. *J Mass Spectrom* 38:1081–1092
9. Holscher D, Shroff R, Knop K et al (2009) Matrix-free UV-laser desorption/ionization (LDI) mass spectrometric imaging at the single-cell level: distribution of secondary

- metabolites of *Arabidopsis thaliana* and *Hypericum* species. *Plant J* 60:907–918
- 10. Svatos A (2010) Mass spectrometric imaging of small molecules. *Trends Biotechnol* 28: 425–434
 - 11. Kaspar S, Peukert M, Svatos A et al (2011) MALDI-imaging mass spectrometry—an emerging technique in plant biology. *Proteomics* 11:1840–1850
 - 12. Peukert M, Matros A, Lattanzio G et al (2012) Spatially resolved analysis of small molecules by matrix-assisted laser desorption/ionization mass spectrometric imaging (MALDI-MSI). *New Phytol* 193:806–815
 - 13. Deutskens F, Yang J, Caprioli RM et al (2011) High spatial resolution imaging mass spectrometry and classical histology on a single tissue section. *J Mass Spectrom* 46:568–571
 - 14. Seeley EH, Oppenheimer SR, Mi D et al (2008) Enhancement of protein sensitivity for MALDI imaging mass spectrometry after chemical treatment of tissue sections. *J Am Soc Mass Spectrom* 19:1069–1077
 - 15. van Remortere A, van Zeijl RJM, van den Oever N et al (2010) MALDI imaging and profiling MS of higher mass proteins from tissue. *J Am Soc Mass Spectrom* 21:1922–1929
 - 16. Lagarrigue M, Becker M, Lavigne R et al (2011) Revisiting rat spermatogenesis with MALDI imaging at 20- μ m resolution. *Mol Cell Proteomics* 10, DOI: [10.1074/mcp.M110.005991](https://doi.org/10.1074/mcp.M110.005991)

Chapter 18

Using the Yeast Two-Hybrid System to Identify Protein–Protein Interactions

Edgar Rodríguez-Negrete, Eduardo R. Bejarano, and Araceli G. Castillo

Abstract

The yeast two-hybrid system is currently one of the most standardized protein interaction mapping techniques. The rationale of the yeast two-hybrid system relies on the physical separation of the DNA-binding domain from the transcriptional activation domain of several transcription factors. The protein of interest (bait) is fused to a DNA-binding domain, and complementary DNA (cDNA) library-encoded proteins are fused to a transcriptional activation domain. When a protein encoded by the cDNA library binds to the bait, both activities of the transcription factor are rejoined resulting in transcription from a reporter gene. Here, we describe protocols to test interactions between two individual proteins and to look for novel interacting partners by screening a single protein or domain against a library of other proteins using a GAL4 based yeast two-hybrid system.

Key words Two-hybrid system, Protein–protein interaction, Yeast, GAL4, Interactors, Interactome

1 Introduction

The yeast two-hybrid system (Y2H) is an *in vivo* method to determine whether two proteins interact [1–3]. It relies on the modular nature of many site-specific transcriptional activators, which consist of a DNA-binding domain (BD) and a transcriptional activation domain (AD). The system is based on the observation that the two domains need not be covalently linked and can be brought together by the interaction of any two proteins. To determine whether two proteins interact, both must be expressed fused to either BD or AD in a yeast strain that contains one or more reporter genes with binding sites for the DNA-binding domain. If the proteins, bait and prey, interact they create a functional activator by bringing the activation domain into close proximity with the DNA-binding domain detectable by expression of reporter genes (Fig. 1). The technology can also be used to identify novel protein interactions using a bait protein fused to BD and a plasmid library that expresses cDNA-encoded proteins fused to a transcription activation domain.

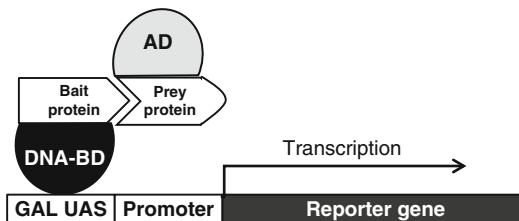


Fig. 1 GAL4 yeast two-hybrid. The DNA-BD of the yeast GAL4 protein binds the GAL UAS upstream of the reported genes. The AD of the GAL4 protein functions as a transcriptional activator

Several varieties of the Y2H exist, commonly involving DNA-binding domains that derive from the yeast Gal4 protein or the *Escherichia coli* LexA protein [4, 5]. Transcriptional activation domains are commonly derived from the Gal4 protein, the herpes simplex virus VP16 protein, or a protein encoded by random *E. coli* sequence, the B42 “acid blob.” Reporter genes include the *E. coli* *lacZ* gene and selectable yeast genes such as HIS3, ADE2 (Gal4 system), or LEU2 (LexA system).

Yeast two-hybrid systems provide a sensitive method for detecting relatively weak and transient protein interactions. In addition, because the two-hybrid assay is performed *in vivo*, the proteins are more likely to be in their native conformations, which may lead to increased sensitivity and accuracy of detection. In spite of its advantages, the yeast two-hybrid system has also some limitations that must be considered prior use.

- Bait proteins may have DNA-binding and/or transcriptional activating properties; hence, deletion of certain portions of bait proteins will be required to eliminate the unwanted activity.
- Some fused proteins may not be able to enter or be stably expressed in the yeast nucleus. The GAL4 BD has its own nuclear localization signal (NLS) within its N-terminal residues while LexA based system does not contain a NLS.
- In some cases, the DNA-BD or AD fusion moiety may hinder the interaction site, or interfere with the folding of the protein, thus altering the ability of the proteins to interact.
- Cellular yeast cell environments may not allow the proper folding or posttranslational modifications required for interaction of some proteins.
- Some protein interactions may not be detectable in a GAL4-based two-hybrid system, but may be detectable using a LexA-based system, and vice versa.

In this chapter we present protocols based on the yeast Gal4 protein where a bait gene is expressed as a fusion to the GAL4

DNA-binding domain (DNA-BD), while another gene or cDNA is expressed as a fusion to the GAL4 activation domain (AD) [1, 6]. Protein interaction is detected by activation of reporter genes under the control of distinct GAL4 upstream activating sequences (UASs). In yeast, galactose metabolism is controlled by two regulatory proteins, GAL4 and GAL80, as well as by the carbon source in the medium. When galactose is present, the GAL4 protein binds to the GAL-responsive elements within the UAS of at least 20 known galactose-responsive genes (including GAL1). In the absence of galactose, Gal80p binds to Gal4p and this interaction blocks transcriptional activation of the galactose-responsive genes. Furthermore, in the presence of glucose, transcription of galactose genes is immediately repressed. To avoid interference by endogenous GAL4 and GAL80 proteins, the yeast host strains used in the GAL4 based two-hybrid system must carry deletions of the GAL4 and GAL80 genes. Due to the deletion of these two genes, the yeast cells grow more slowly as compared to yeast containing the wild type version of these genes. In the GAL4-based two-hybrid system, either the intact GAL1 UAS, which contains four GAL4-binding sites, or an artificially constructed UAS consisting of three copies of the 17-mer consensus binding sequence are used to control de expression of the reported genes.

2 Materials

2.1 Plasmids

A large number of different DNA-binding domain (BD) and transcription activation domain (AD) containing vectors have been successfully used. In our experience, pGKBT7 and pGADT7, designed to obtain a high-level of protein expression, gave the best results. In these vectors, bait and prey/library inserts are expressed as GAL4 fusions with c-Myc and hemagglutinin (HA) epitope tags, respectively. The epitope tags eliminate the need to generate specific antibodies to each new protein and allow convenient identification of the fusion proteins. Otherwise anti-GAL4 BD and AD monoclonal antibodies can be used (Clontech Cat. No 630402 and 630403). Plasmid information is provided in Table 1. A complete list of other plasmids can be found in <http://www.biotech.ubc.ca/Journal/V01I01/8188> yeast.pdf.

2.2 Yeast Strains

Several yeast strains containing different reporter genes are available. Among them, PJ69-4A and AH109 strains use three reporters—*ADE2*, *HIS3*, and *lacZ* (or *MEL1*)—under the control of distinct GAL4 UASs. These promoters yield strong and specific responses to GAL4.

The *ADE2* reporter alone provides strong nutritional selection, however using also the *HIS3* selection reduces the incidence of false positives and allows control of the stringency of selection

Table 1
Common bait and prey vectors used for Y2H

| Plasmid | Fusion | Epitope | Yeast selection | Bacterial selection | Reference and/or map |
|------------|---------|---------|-----------------|---------------------|--------------------------|
| pGBT7 (1) | BD/bait | c-Myc | <i>TRP1</i> | Kanamycin | Clontech Cat. No. 630443 |
| pGBT9 (3) | BD/bait | None | <i>TRP1</i> | Ampicillin | [7] |
| pAS2-1 (1) | BD/bait | None | <i>TRP1</i> | Ampicillin | [8] |
| pGADT7 (1) | AD/prey | HA | <i>LEU2</i> | Ampicillin | Clontech Cat. No. 630442 |
| pACT2 (2) | AD/prey | HA | <i>LEU2</i> | Ampicillin | [9] |

BD or AD fusions are expressed from full length (1) or truncated (2 and 3) ADH promoters to produce high (1), medium (2), or low constitutive expression level

Table 2
Genotypic and phenotypic characteristics of PJ69-4A and AH109

| Strain | Genotype | Reporters | Transformation markers | Reference |
|---------|---|-------------------------|------------------------|-----------|
| PJ69-4A | <i>MATa leu2-3,112 ura3-52 trp1-901 his3-200 gal4Δ gal80Δ GAL2-ADE2 LYS2::GAL1- HIS3 MET2::GAL7- LacZ</i> | <i>HIS3, ADE2, LacZ</i> | <i>trp1, leu2</i> | [10] |
| AH109 | <i>MATa, leu2-3, 112, ura3-52, trp1-901, his3-200, gal4Δ, gal80Δ, GAL2-ADE2, LYS2::GAL1-HIS3, URA3::MEL1-lacZ</i> | <i>HIS3, ADE2, MEL1</i> | <i>trp1, leu2</i> | [10] |

The *GAL1*, *GAL2*, and *GAL7* upstream activating sequences (UASs) are recognized and bound by the Gal4 BD. The *trp1*, *his3*, *gal4*, and *gal80* mutations are all deletions; *leu2-3, 112* is a double mutation. PJ69-4A requires addition of leucine, uracil, tryptophan, histidine, adenine, lysine, and methionine in minimal media to grow

(detailed in Subheading 2.3) [10]. In addition using either *lacZ* or *MEL1*, which encode β-galactosidase and α-galactosidase respectively, allows for quantitative studies of the protein–protein interactions.

For additional information on the growth and maintenance of yeast, see the Guide to Yeast Genetics and Molecular Biology [11]. The complete phenotype and genotype of the PJ69-4A and AH109 strains are shown in Table 2. The protocols presented in this chapter have been optimized for PJ69-4A.

2.3 Media and Supplements

Prepare all media using deionized water. For solutions use ultrapure water and analytical grade reagents. All reagents must be prepared and stored at room temperature (unless otherwise indicated).

1. *Minimal medium (MM)*. It is used for general yeast growth. It contains Yeast Nitrogen Base (YNB, 1.7 g/l), ammonium

Table 3
Supplements to be added to MM, CM and IMs media

| Amino acid | Interaction media (IM) | | | | |
|-------------------------|------------------------|------------------------------|------------------------|---------------------------|-------------------------|
| | Minimal media (MM) | Co-transformation media (CM) | Low stringency (IM-LS) | Medium stringency (IM-MS) | High stringency (IM-HS) |
| Lysine | + | + | + | + | + |
| Methionine | + | + | + | + | + |
| Uracil | + | + | + | + | + |
| Leucine | + | - | - | - | - |
| Tryptophan ^a | + | - | - | - | - |
| Adenine | + | + | + | - | - |
| Histidine | + | + | - | + | - |
| 3-AT | - | - | + | - | + |

^aStock must be protected from light and kept at 4 °C

sulfate (5 g/l), glucose (20 g/l) and a mixture of amino acids to complement the yeast auxotrophies (Table 3). Adjust the pH to 6.5, then autoclave. If preparing solid MM, add agar (20 g/l) and then autoclave it. Amino acid solutions must be sterilized separately and then added before pouring plates. YNB is a base medium for preparation of minimal and synthetic defined yeast media that can purchase from different companies (for example Becton, Dickinson, BD).

2. *Cotransformation medium (CM)*. It is used to select for co-transformants cells that contain both plasmids, the DNA-binding domain (BD) and transcription activation domain (AD) containing vectors. It has the same composition as MM but it is not supplemented with leucine and tryptophan.
3. *Interaction media (IM)*. They are employed to assess protein–protein interactions. Three different IM media are defined to control the stringency of the selection: IM-LS for low stringency, IM-MS for medium stringency and IM-HS for high stringency. The media have the same composition as CM, except that adenine and/or histidine are absent and 3-AT (*3-Amino-1,2,4-triazole*) is added when histidine is not present.
4. *3-Amino-1,2,4-triazole* (3-AT; Sigma-Aldrich) is a competitive inhibitor of the HIS3-gene product and may be used to titrate the minimum level of *HIS3* expression required for growth on histidine-deficient media. A 2 M stock solution should be freshly prepared by dissolving 3-AT in H₂O and filter-sterilized.

Table 4
Recommended final concentration of amino acids

| Amino acid | Final | Stock solution | mL stock for 1 l of media |
|------------|---------|----------------|---------------------------|
| Adenine | 20 mg/l | 2 g/l | 10 |
| Histidine | 20 mg/l | 10 g/l | 2 |
| Leucine | 60 mg/l | 10 g/l | 6 |
| Lysine | 50 mg/l | 10 g/l | 5 |
| Methionine | 20 mg/l | 10 g/l | 2 |
| Tryptophan | 20 mg/l | 10 g/l | 2 |
| Uracil | 20 mg/l | 2 g/l | 10 |
| 3-AT | 2 mM | 2 M | 1 |

Stock solutions may be autoclaved or filter-sterilized using a. 0.2 µm filter and stored at room temperature or at 4 °C for up to 1 year

5. *Yeast Extract Peptone Dextrose Adenine, (YPD)* is a complete medium for *yeast* growth supplemented with adenine. It contains yeast extract (10 g/l), peptone (20 g/l), glucose (20 g/l) and adenine (20 mg/l). Autoclave. If preparing solid YPD, add agar (20 g/l) and then autoclave it.

The media requirements for PJ69-4A strain are summarized in Table 3. The recommended final concentrations of amino acids are shown in Table 4. The amino acids used to supplement the media, will depend on the Y2H strain used.

2.4 Yeast Transformation

Solutions and reagents required for yeast transformation are as follows (use ultrapure water for all solutions and analytical grade reagents):

1. *2 mg/ml of salmon-sperm DNA* or sheared herring testes DNA (Clontech) to be used as DNA-carrier. Immediately prior to use, denature the carrier DNA by boiling for 10 min and cooling it on ice.
2. *50 % PEG 4,000* (Polyethylene glycol, avg. MW = 3,350; Sigma-Aldrich). Filter-sterilized or autoclaved.
3. *1 M Lithium acetate* (Sigma-Aldrich), adjust pH to 7.5 using diluted acetic acid, and filter-sterilized.
4. *TE buffer* 0.1 M Tris-HCl, 10 mM EDTA, adjust pH to 7.5, and autoclave.

2.5 Plasmid DNA and Protein Extraction from Yeast

1. *DNA breakage buffer*. 2 % Triton X-100, 1 % SDS, 100 mM sodium chloride, 10 mM Tris-HCl, pH 8 and 1 mM EDTA.
2. *Phenol-chloroform-isoamyl alcohol* (25:24:1).

3. *3 M sodium acetate, pH 5.2.* Adjust the pH using diluted acetic acid.
4. *RIPA buffer.* 2 mM sodium phosphate buffer, pH 7, 0.2 % Triton X-100, 0.02 % SDS, 0.2 mM EDTA, pH 8, and 10 mM sodium chloride. Store at room temperature.
5. *Protease inhibitor cocktail* (Sigma-Aldrich).
6. *100 mM phenylmethylsulfonyl fluoride, PMSF,* prepared in methanol (Sigma-Aldrich).
7. *2x Laemmli sample buffer.* 10 % SDS, 20 % glycerol, 120 mM Tris-HCl pH 6.8 and 0.02 % bromophenol blue. Store at room temperature.
8. *Acid-washed glass beads (425–600 µm diameter* (Sigma-Aldrich)).

2.6 LacZ Assay

All solutions use ultrapure water and analytical grade reagents.

1. *Z buffer:* 0.1 M sodium phosphate monobasic, pH 7, 10 mM potassium chloride, 1 mM magnesium sulfate. Adjust pH to 7.0 and autoclave. Prepare fresh when needed.
2. 4 mg/l *o-nitrophenyl-beta-D-galactopyranoside.* Dissolve *o-nitrophenyl-beta-D-galactopyranoside* (Sigma-Aldrich) in deionized water. Alternatively, a stock solution may be prepared in phosphate buffer. When used as a substrate for β -galactosidase, a solution of 20.5 mg/ml is prepared in 100 mM sodium phosphate buffer, pH 7.3. The solution is gently warmed to completely dissolve the product. Store in the dark –20 °C.

3 Methods

3.1 Screening a cDNA Library for Interactors

In a *GAL4*-based two-hybrid assay, a bait protein is expressed as a fusion to the Gal4 DNA-binding domain (DNA-BD), while libraries of prey proteins are expressed as fusions to the Gal4 activation domain (AD).

Very few plant cDNA libraries are commercially available, although there are companies that perform screening using pre-made libraries or generate them using *GAL4* or LexA-based systems. Libraries can be found through the following links:

http://www.arabidopsis.org/abrc/catalog/cdna_library1.html
<http://www.hybrigenics-services.com/>

Screening a cDNA library for interactors can be performed by co-transformation or by mating. We include a protocol using co-transformation in a *GAL4*-based two-hybrid system. A protocol to perform Y2H by mating using the *GAL4* system is detailed in

Clontech protocols (Protocol No. PT4084-1, Version No. PR033493). Protocols for the LexA-based system are available in <http://www.protocol-online.org/>.

The screening process consists of the following steps:

1. Clone and test bait for expression, toxicity and autoactivation.
2. Screen the cDNA library.
3. Confirm and interpret results.

3.1.1 Preparing the Bait Vector

The bait clone to express *GAL4* DNA-BD fusioned to your gene of interest can be obtained by cloning it in frame with the *GAL4* DNA binding domain of the bait plasmid pGBKT7 (*see Note 1*).

Before carrying out the library screening it is strongly recommended to determine whether or not your bait is expressed in yeast, does not have toxic effects and does not autonomously activate the reporter genes in the presence of any given prey protein (autoactivation).

Preparation of Yeast Competent Cells

1. Streak an YPDA plate with cells from a frozen yeast stock. Incubate the plate upside down at 30 °C until colonies appear (~3 days). Yeast strains can be stored for up to 2 months at 4 °C on YPDA medium in petri dishes sealed with parafilm. However, fresh colonies will give better results when inoculating a liquid culture.
2. Inoculate one colony into 15 ml YPDA medium in a sterile 50 ml flask or tube. Incubate at 30 °C with shaking at 250 rpm over-night.
3. Transfer an appropriate volume of the over-night culture to 50 ml of YPDA in a 250 ml flask to reach OD₆₀₀ of 0.2. Incubate shaking at 250 rpm until the OD₆₀₀ reaches 0.8–1 (6–8 h).
4. Centrifuge the cells at 3,000×*g* for 4 min at room temperature. Discard the supernatant and resuspend the pellet in 25 ml of sterile deionized water.
5. Centrifuge the cells at 3,000×*g* for 4 min at room temperature. Discard the supernatant and resuspend the pellet in 25 ml of 100 mM lithium acetate.
6. Centrifuge the cells at 3,000×*g* for 4 min at room temperature. Discard the supernatant and resuspend the pellet in 900 µl of 100 mM lithium acetate. Transfer the cell suspensions to a 1.5 ml microcentrifuge tube. The cells are now ready to be transformed with plasmid DNA (*see Note 2*).

Yeast Transformation

1. For each transformation use 50 µl of competent cells and add in the indicated order: 240 µl of 50 % PEG 4,000, 36 µl of 1 M lithium acetate, and the transforming DNA (co-transformation

with the bait construct and a prey empty vector is advisable) dissolved in deionized sterile water (300–600 ng of each plasmid in a final volume of 75 μ l).

2. Mix vigorously and incubate at 30 °C for 20–30 min.
3. Mix by gently inverting the tubes and transfer to a preheated 42 °C circulating water bath or thermo block. Incubate for 30 min.
4. Allow the cells to cool at room temperature for 1–2 min.
5. Centrifuge the cells at $3,000 \times g$ for 3 min at room temperature. Discard the supernatant and resuspend the pellet in 900 μ l of deionized sterile water.
6. Spread different amounts of the resuspended cells onto a plate (10 cm of diameter) containing the appropriate selection medium to obtain isolated colonies (normally 10 and 100 μ l). For pGBK7, use MM without tryptophan. For pGADT7, use MM without leucine. For cotransformations use CM.
7. Incubate the plates at 30 °C for 3–5 days.

Detecting Bait Expression and Toxicity

For detection of the fusion protein, a Western blot of protein extracts obtained from the transformed yeast could be performed using Gal4 DNA-BD Monoclonal Antibody (Clontech) or -Myc Monoclonal Antibody (Clontech).

1. Yeast protein extract can be prepared by growing yeast in MM without tryptophan, to maintain selection for the bait plasmid, to $OD_{600} = 0.8–1.0$.
2. Centrifuge 30–40 ml of the culture at $4,000 \times g$ for 3 min to pellet the cells. Wash the pellet once with cold water, transfer to an ice-cold screw-cap tube containing 0.5 ml of glass beads (425–600 μ m) and resuspend in 200 μ l of RIPA buffer containing protease inhibitor cocktail (50 μ l to 5 ml of buffer) and 1 mM phenylmethylsulfonyl fluoride (PMSF). Grind in FastPrep™ FP120 (BIO 101) at power setting 5.5 for two 15 s intervals separated by 1 min intervals on ice.
3. Add 400 μ l RIPA buffer with protease inhibitors and vigorously vortex. Keep on ice. Heat needle in a flame and quickly pierce bottom of the screw-cap tube containing lysate. Make 2–3 holes in each tube and place this 1 min in cooled centrifuge (do six tubes at a time). Keep on ice.
4. Centrifuge at $4,000 \times g$, 2 min in a cooled centrifuge to pellet the cells. Carefully recover the supernatant containing the protein extract. Protein concentration can be determined using a standard Bradford assays.
5. Load 100 μ g of the protein extract on a polyacrylamide/SDS gel. The proteins can then be transferred to a filter and blotted with standard immunoblotting (Western) methods [12].

A faster but less efficient protocol to extract protein is as follows:

1. Grow and recover the cells as described above.
2. Resuspend the pellet in 50 μ l of 2 \times Laemmli sample buffer. The cells can then be broken by freezing on dry ice followed by boiling for 5 min prior to loading on an SDS polyacrylamide gel (about 15 μ l/lane).

To evaluate the toxicity on yeast growth produced by the expression of the bait, compare the size of the yeast colonies transformed with pGBKT7 (empty vector) or pGBKT7+bait. If the bait is toxic, the colonies containing the bait vector will be significantly smaller than colonies containing the empty pGBKT7 vector (see Note 3).

Testing the Bait for Autoactivation

As a first step for any two-hybrid screen, it is imperative to confirm that your bait does not autonomously activate the reporter genes in the absence of a prey protein.

1. Transform yeast cells with 300 ng of the bait vector and inoculate onto plates of MM without tryptophan to select the transformants (100 μ l of 10 $^{-1}$ and 10 $^{-2}$ dilution).
2. Use sterile toothpicks or inoculating loops to transfer (make small streaks of 2–3 mm in length of each colony) at least five yeast colonies onto the following media:
 - (a) IM-LS, IM-MS and IM-HS (add leucine if cells are transformed only with the bait plasmid) to detect autoactivation.
 - (b) MM without tryptophan to confirm the presence of the plasmid and check yeast growth.
3. Invert the plates and incubate at 30 °C until colonies appear (3–5 days). Growth should be detected only in MM, otherwise the bait clone autoactivates (see Note 4).

3.1.2 Library Screening by Co-transformation

Most cDNA libraries contain over 10 6 individual cDNAs in AD-GAL4 vectors. To isolate the less abundant cDNAs it is required to obtain 2–3 \times 10 6 yeast transformants. Because most common yeast two-hybrid strains transformation efficiency is around 10 5 transformants per μ g of library plasmid DNA, a large-volume transformation or several individual transformations are usually required to obtain the desired number of total transformants. An exploratory transformation should be performed with the selected strain to determine the transformation efficiency.

A yeast transformation protocol readjusted to obtain up to 3 \times 10 7 co-transformants is described (Fig. 2):

1. Streak an YPDA agar plate with cells from a frozen yeast stock. Incubate the plate upside down at 30 °C until colonies appear (~3 days).

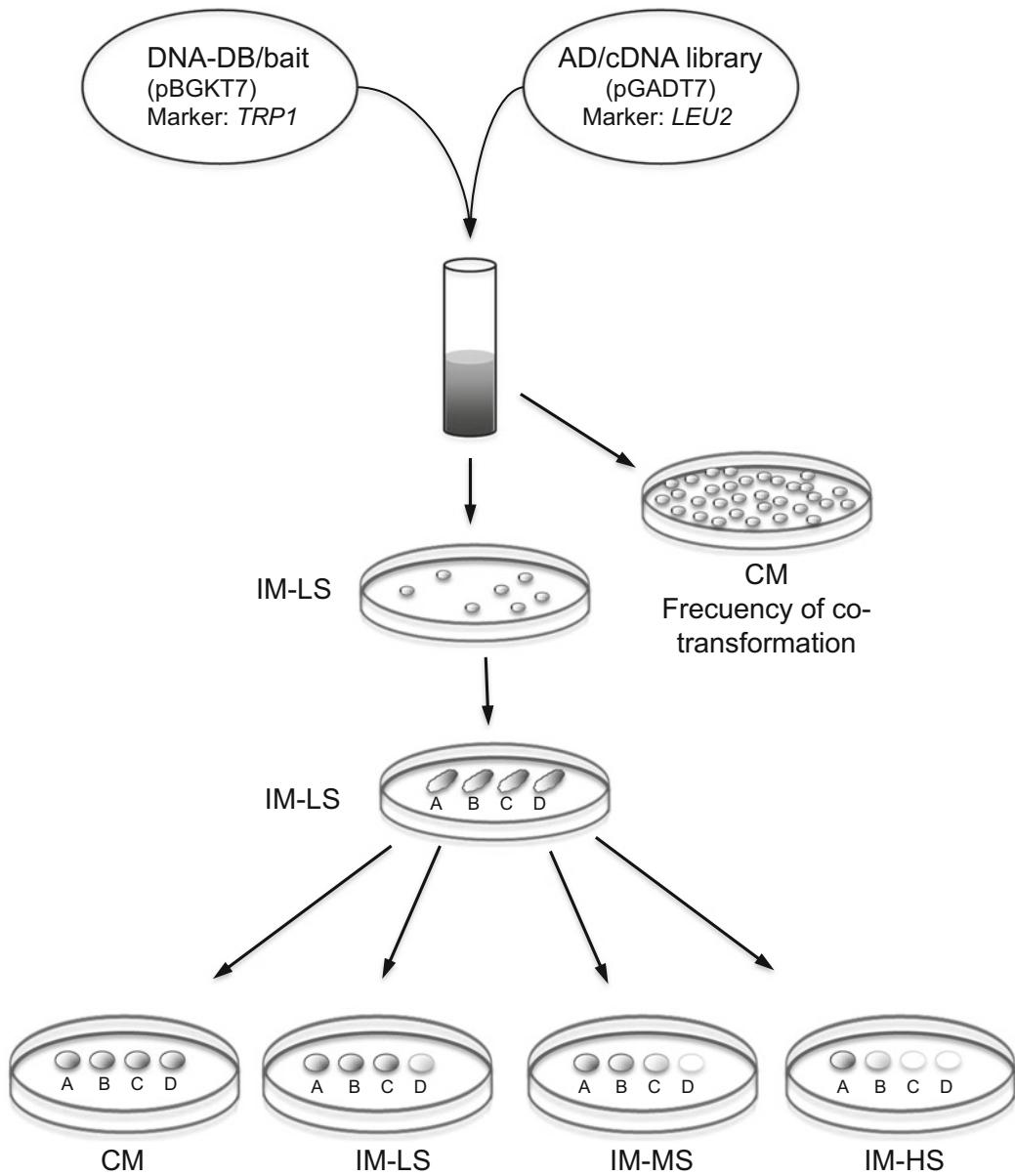


Fig. 2 Screening a cDNA library. A yeast culture is co-transformed with DB-bait and AD-cDNA library. Dilutions of the co-transformed culture are inoculated in CM media to determine the frequency of co-transformation, the rest of the culture is plated in IM-LS media to select for interactors in low stringency conditions. Colonies grown in IM-LS are replicated to a new plate of IM-LS (the diagram represents four positive clones labelled as A, B, C and D). The yeast are collected from one streak using a sterile toothpick and resuspended in 100 μ l of deionized sterile water. Add 5 μ l of each suspension in plates of CM (as a growth control), and IM-LS, IM-MS and IM-HS to determine the strength of the interaction. Examples of growth level expected from strong (clone A), medium (clone B), weak (clone C) and very weak (clone D) interactions are shown

2. Inoculate one colony into 50 ml YPDA medium in a 250 ml sterile flask. Incubate at 30 °C with shaking at 250 rpm over-night.
3. Transfer an appropriate volume of the over-night culture to 300 ml of YPDA in a 500 ml flask to reach OD₆₀₀ of 0.2. Incubate with shaking at 250 rpm until the OD₆₀₀ reaches 0.8–1 (6–8 h).
4. Centrifuge the cells at 3,000×*g* for 4 min at room temperature. Discard the supernatant and resuspend the pellet in 150 ml of sterile deionized water.
5. Repeat **step 4**.
6. Centrifuge the cells at 3,000×*g* for 4 min at room temperature. Discard the supernatant.
7. Add the solutions below to the cell pellet in the following order:

| | |
|--------------------------------------|---------|
| 50 % PEG 4,000 | 14.4 ml |
| 1 M Lithium acetate | 2.1 ml |
| Salmon sperm-DNA (2 mg/ml) | 1.5 ml |
| DNA from the bait-plasmid (10–20 µg) | X µl |
| DNA from the library (5–10 µg) | Y µl |

Complete with sterile deionized water up to a final volume of 3 ml.

The volumes can be scaled down or up to reach the required number of transformants. For best results, competent cells should be used for transformation immediately, although they can be stored on ice for a few hours after the washes, without significant loss in efficiency.

8. Mix vigorously and incubate at 30 °C for 30 min mixing by inversion every 5 min to homogenize the temperature. It is important to resuspend the cells completely before starting the incubation.
9. Transfer the tubes to rack placed in a preheated 42 °C circulating water bath. Store the tubes in the rack for 30 min mixing by inversion every 5 min to homogenize the temperature.
10. Centrifuge the cells at 3,000×*g* for 4 min at room temperature. Gently resuspend the cells in 30–40 ml of sterile deionized water.
11. Remove 100 µl from the transformation mixes and prepare dilutions (10⁻¹, 10⁻², and 10⁻³) in sterile deionized water. Plate 100 µl of each dilution onto 10 cm diameter CM plates and incubate at 30 °C. This will allow the total number of co-transformants to be calculated.

12. Plate the remainder of the transformation mixes onto 24 cm × 24 cm IM-LS plates to select for the induction of HIS3. There is no need to spin the cells or remove the PEG. To achieve an even distribution of cells, pour about 100 sterile glass beads (4 mm diameter) onto the plate with the cells. Gently roll the beads around the plate to distribute the transformation mix. This technique works best when the surface of the plates is not too wet so that the medium absorbs the transformation mix. To achieve this moisture content, put newly solidified plates into a laminar flow hood with the lids ajar for about 1 h before plating.
13. Incubate the plates at 30 °C. Continue incubation until colonies are 1–2 mm in diameter, which should take a total of approximately 3–5 days.
14. Use sterile toothpicks or inoculating loops to transfer all of the grown colonies onto 10 cm diameter IM-LS plates to confirm the result and isolate the colonies. Make small streaks of 2–3 mm in length of each colony. Incubate the plates at 30 °C for 2–4 days. If the colonies are closely spaced it will be necessary to streak to single colonies to separate the different co-transformants.
15. Collect the yeast from one streak using a sterile toothpick and resuspend them in 100 µl of deionized sterile water in 1.5 ml microcentrifuge tubes. Add 5 µl of each suspension to plates with the following media:
 - (a) IM-HS to select for the induction of *ADE2* and *HIS3*.
 - (b) IM-MS to select for the induction of *ADE2*.
 - (c) IM-LS to select for the induction of *HIS3*.
 - (d) CM to check cell growth.

Incubate the plates at 30 °C for 2–7 days. Positive interactions will be determined by comparison with the growth of the positive and negative yeast two-hybrid controls on each media (see Note 4).

Rescue of the Prey Vector

1. To rescue the library plasmid, inoculate 10 ml of CM media with the positive yeast colony and incubate at 30 °C overnight. Centrifuge the culture at 3,000 × *g* for 4 min. Remove all supernatant and resuspend the pellet in 200 µl of DNA breakage buffer. Alternatively use sterile toothpicks or inoculating loops to transfer to make small streaks of 2–3 mm in length in a plate of CM media. Incubate at 30 °C overnight and resuspend a toothpick worth of cells in 200 µl of DNA breakage buffer.
2. Add a scoopful of glass beads (425–600 µm of diameter) and 200 µl of phenol–chloroform.
3. Vortex for 4 min.
4. Centrifuge at 13,000 × *g* during 5 min at room temperature. Transfer the upper aqueous layer to 1 ml of cold ethanol

(do not disturb the interface and re-centrifuge if necessary). Mix by inversion.

5. Centrifuge at $13,000 \times g$ during 5 min at room temperature. Discard the supernatant and air dry the pellet.
6. Resuspend in 50 μ l of TE.
7. Use 2–5 μ l of the extracted yeast DNA to transform the appropriate *Escherichia coli* strain (for example DH5 α) to recover the plasmid from the cDNA library. If using a bait plasmid that contains resistance to kanamycin (pGBKT7) the prey plasmid for each positive clone will be easily isolated by plating the bacteria transformation into ampicillin plates (all GAL4 prey plasmids contain resistance to ampicillin). Select at least six transformed colonies. Extract the plasmids and carry out a restriction analysis, which will determine whether all the recovered clones from each positive yeast colony contain the same cDNA.

Confirm the Interaction

1. Co-transform yeast cells with the bait vector and the recovered cDNA library plasmid following the protocol described in Subheading 3.1.
2. Transfer the yeast colonies one by one onto plates of CM. Make small streaks of 2–3 mm in length of each colony in two plates. Incubate at 30 °C overnight. Store one of the plates at 4 °C and use the other to analyze the interactions.
3. Collect the yeast from one streak using a sterile toothpick and resuspend them in 100 μ l of deionized sterile water placed in 1.5 ml microcentrifuge tubes.
4. Add 5 μ l of each suspension on plates with the following media:
 - (a) IM-HS to select for the induction of ADE2 and HIS3.
 - (b) IM-MS to select for the induction of ADE2.
 - (c) IM-LS to select for the induction of HIS3.
 - (d) CM to check cell growth.
5. Invert the plates and incubate them at 30 °C until colonies appear (3–5 days).

Growth will confirm the result obtained in the screening, then cDNA clones will be them sequenced using the appropriate primers (OPACT-1: ATGATGAAGATACCCC; OPACT-2: TTGAA GTGAACTTGCG, can be used for prey clones in pGADT7).

Determining the Specificity of an Interactor

Many of the clones identified in the yeast two-hybrid screenings may be false positives due to rearrangement of the bait-vector, self-activation of the reporter genes or a nonspecific interaction with the prey-protein. Co-transformation with the rescued prey-vectors and the bait protein should be performed to confirm the interaction. Control for nonspecific interactions must be also included, so plasmids expressing the bait empty vector or expressing non-related

Table 5**Co-transformations required to determine the interaction between two proteins (A and B)**

| Bait clone (BD-) | A | B | A | LamC | B | LamC | P53 | LamC |
|---------------------------|-----|-----|-------|------|------|------|-------|-------|
| Prey clone (AD-) | B | A | T-ant | B | LamC | A | T-ant | T-ant |
| Growth in selective media | Yes | Yes | No | No | No | No | Yes | No |

Positive and negative yeast two-hybrid controls are indicated in bold

LamC lamin C, *T-ant* SV40 large T-antigen, *P53* murine p53

bait proteins should also be co-transformed with the rescued prey-vectors.

This method of testing specificity can be somewhat cumbersome if a large number of different library plasmids were isolated, and if these are to be tested for interaction with several different baits. For this purpose a mating assay to perform the specificity test would be more appropriate (further information could be founded in www.clontech.com, Cat. No. PR033493)

Interestingly, the commonly isolated nonspecific interactors which include heat shock proteins, ribosomal proteins, proteasome subunits, and other proteins, are not isolated in every interactor hunt, and in fact do not appear to interact with every bait. This highlights the importance of using several different bait proteins to test the specificity of an interactor.

3.2 Analysis of a Specific Interaction

To determine if two proteins (A and B) interact, a complete or partial cDNA from each must be cloned into both bait and prey vectors to express them fused either to the BD or the AD of Gal4. Protein expression, toxicity and autoactivation of bait and prey proteins must be tested beforehand to confirm the interaction as described in Subheading 3.1.1.

Interaction is tested by co-transforming yeast cells according protocol Subheading 3.1.1. It is advisable to test the interaction using plasmids expressing protein A fused to AD and B to BD, as well as plasmids expressing protein B fused to AD and A fused to BD. Specificity of the interaction must be also confirmed using bait or prey vectors expressing non-related proteins. Table 5 summarizes the co-transformations required to test the interaction and the expected results for a positive interaction (see Note 4).

3.2.1 Test the Interaction

1. Cells must be plated onto CM plates to select for co-transformants. Incubate the plates at 30 °C for 3–5 days.
2. Transfer the colonies onto fresh CM plates. If the colonies are closely spaced it will be necessary to streak purify to single colonies to separate the different co-transformants. Incubate the plates at 30 °C for 2 days.
3. Collect the yeast from one streak using a sterile toothpick and resuspend them in 100 µl of deionized sterile water placed in 1.5 ml microcentrifuge tubes.

4. Add 5 μ l of each suspension on plates of the following media:
 - (a) IM-HS to select for the induction of ADE2 and HIS3.
 - (b) IM-MS to select for the induction of ADE2.
 - (c) IM-LS to select for the induction of HIS3.
 - (d) CM to check cell growth.
5. Invert the plates and incubate them at 30 °C for 2–7 days. Positive interactions will be determined by comparison with the growth of the positive and negative yeast two-hybrid controls on each media (see Note 4).

3.2.2 Measure the Interaction

Quantitative β -Galactosidase Assays from Liquid Cultures

Interaction strength can be measured by analyzing the induction of *lacZ* activity or by comparison of growth rate in interacting media (IM).

1. Pick three or four co-transformant colonies and inoculate them in 5 ml of CM media (all colonies should have similar growth). Incubate at 30 °C and 250 rpm until OD₆₀₀ reaches 0.5–0.8. The positive and negative yeast two-hybrid interaction controls should always be included (Table 5) (see Note 4).
2. Centrifuge at 3,000 $\times g$ for 4 min at room temperature. Discard the supernatant and wash the pellet with Z buffer (0.1 M NaPO₄ pH 7, 10 mM KCl, 1 mM MgSO₄). Centrifuge at 3,000 $\times g$ for 4 min at room temperature and resuspend the pellet in 300 μ l of Z buffer.
3. Break the cells by subjecting the yeast suspension to three freeze/thaw cycles in liquid nitrogen.
4. Aliquots (100 μ l) are then assayed for β -galactosidase activity by adding 700 μ l of Z buffer containing 40 μ l-mercaptoethanol and 160 μ l of substrate *o*-nitrophenyl- β -D-galactopyranoside (4 mg/ml). Add 400 μ l of 1 M Na₂CO₃ to the mixture when the yellow color develops to stop the reaction.
5. Centrifuge at 12,000 $\times g$ for 10 min. Accumulation of the *o*-nitrophenol product is measured at 420 nm. Protein concentration is measured by the Bradford assay (Bio-Rad). The enzyme-specific activity is calculated as described [13]. The relative activities must be normalized against the maximum value obtained, which was set to 100 %. The different constructs were tested in a minimum of three experiments, each of which assayed three independent co-transformants for each interaction.

Growth Rate Assay

1. Pick two or three co-transformant colonies using sterile toothpicks and resuspend them in 200 μ l of deionized sterile water in a 96 well-plate. Make serial dilutions from 10⁻¹ to 10⁻⁵ of each culture with sterile water (180 μ l of water and 20 μ l of the culture). Add a drop of 5 μ l of each dilution onto the following media:

- (a) IM-HS to select for the induction of *ADE2* and *HIS3*.
 - (b) IM-MS to select for the induction of *ADE2*.
 - (c) IM-LS to select for the induction of *HIS3*.
 - (d) CM to check cell growth.
2. Incubate the plates at 30 °C. Monitor the emergence of colonies during the next several days.

4 Notes

1. In order to confirm that the fusion construct is in-frame, the fusion junction can be sequenced using a standard T7 primer.
2. For best results, competent cells should be used for transformation immediately, although they can be stored on ice for a few hours without significant loss in efficiency.
3. If expression of the bait has toxic effects, clone the bait in a vector (such as pGBT9) with lower level of expression, or use truncated versions of the protein to perform the screening.
4. Positive and negative controls must be included in all tests. For GAL4 system vectors expressing Gal4 DNA-BD fused with murine p53 (pGBT7-53) and Gal4 AD fused with SV40 large T-antigen (pGADT7-T) are commonly used as positive controls [9, 13]. A negative control should also be performed using pGBT7-Lam (which encodes the Gal4 BD fused with lamin C) and pGADT7-T.

Acknowledgments

We are grateful to Sharon White for critical reading the manuscript and making many helpful suggestions. This research was supported by a grant from the Spanish Ministerio de Ciencia y Tecnología (AGL2010-22287-C02-02/AGR) and Fondo Europeo de Desarrollo Regional (FEDER).

References

1. Fields S, Song O (1989) A novel genetic system to detect protein–protein interactions. *Nature* 340:245–246
2. Legrain P, Wojcik J, Gauthier JM (2001) Protein–protein interaction maps: a lead towards cellular functions. *Trends Genet* 17: 346–352
3. Koegl M, Uetz P (2007) Improving yeast two-hybrid screening systems. *Brief Funct Genomic Proteomic* 6:302–312
4. Guo D, Rajamaki ML, Valkonen J (2008) Protein–protein interactions: the yeast two-hybrid system. *Methods Mol Biol* 451:421–439
5. Fashena SJ, Serebriiskii IG, Golemis EA (2000) LexA-based two-hybrid systems. *Methods Enzymol* 328:14–26
6. Chien CT et al (1991) The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc Natl Acad Sci U S A* 88:9578–9582

7. Bartel PL et al (1993) Using the two-hybrid system to detect protein–protein interactions. In: Hartley DA (ed) *Cellular interactions in development: A practical approach*. Oxford University Press, Oxford, pp 153–179
8. Harper J et al (1993) The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1cyclin-dependent kinases. *Cell* 75:805–816
9. Durfee T et al (1993) The retinoblastoma protein associates with the protein phosphatase type I catalytic subunit. *Genes Dev* 7:555–569
10. James P, Halladay J, Craig EA (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* 144:1425–1436
11. Guthrie C, Fink GR (eds) (1991) *Guide to yeast genetics and molecular biology*, vol 169. Academic, San Diego
12. Ausubel FM et al (eds) (1999) *Short protocols in molecular biology: a compendium of methods from current protocols in molecular biology*, 4th edn. Wiley, New York
13. Guarente L (1983) Yeast promoters and lacZ fusions designed to study expression of cloned genes in yeast. *Methods Enzymol* 101: 181–191

Chapter 19

Modifications of Wheat Germ Cell-Free System for Functional Proteomics of Plant Membrane Proteins

Akira Nozawa and Yuzuru Tozawa

Abstract

Functional proteomics of plant membrane proteins is an important approach to understand the comprehensive architecture of each metabolic pathway in plants. One bottleneck in the characterization of membrane proteins is the difficulty in producing sufficient quantities of functional protein for analysis. Here, we describe three methods for membrane protein production utilizing a wheat germ cell-free protein expression system. Owing to the open nature of cell-free synthesis reaction, protein synthesis can be modified with components necessary to produce functional protein. In this way we have developed modifications to a wheat germ cell-free system for the production of functional membrane proteins. Supplementation of liposomes or detergents allows the synthesis of functional integral membrane proteins. Furthermore, supplementation of myristic acid enables synthesis of *N*-myristylated peripheral membrane proteins. These modified cell-free synthesis methods facilitate the preparation and subsequent functional analyses of a wide variety of membrane proteins.

Key words Cell-free protein expression system, Detergents, Liposomes, Membrane proteins, *N*-Myristylation, Wheat germ

Abbreviations

| | |
|----------|--|
| AtDTC | <i>Arabidopsis thaliana</i> dicarboxylate/tricarboxylate carrier |
| bR | Bacteriorhodopsin |
| CBB | Coomassie brilliant blue |
| CF | Cell-free |
| CHAPS | 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate |
| FC | Fos-choline |
| MP | Membrane protein |
| RNase | Ribonuclease |
| SDS-PAGE | Sodium dodecyl sulfate-polyacrylamide gel electrophoresis |
| WG | Wheat germ |
| WGCF | Wheat germ cell-free |

1 Introduction

Membrane proteins (MPs) play critical roles in a variety of biological processes such as substrate transport, signal transduction, energy production, and cellular adhesion. It has been reported that approximately 30 % of genes in the plant genomes encode MPs. Despite their abundance, efforts to study the function and structure of plant MPs have been hampered by their low natural abundance and the toxicity of MP overexpression in host cells. In the last decade, cell-free (CF) protein expression systems have attracted much attention as an alternative strategy for MP production. The CF system has several advantages that better suit it for MP production compared to cell-based methodologies. For instance, MP expression in CF system is decoupled from cell viability allowing for the production of toxic proteins that do not interfere with translation. In addition, CF systems are open for modification, allowing supplementation of cofactors, chaperones, lipids, and detergents in order to produce functional protein.

In this chapter, three methods for the production of MPs using a wheat germ (WG) CF (WGCF) system are described (Fig. 1). The first method is liposome-supplemented WGCF system, which is suitable for functional analysis of MPs (Fig. 1a). We previously reported the development of WGCF-based MP synthesis for *Arabidopsis* phosphoenolpyruvate/phosphate translocators [1]. The WGCF system-based method has been further modified for the synthesis of functional mitochondrial carriers, such as *Arabidopsis thaliana* dicarboxylate/tricarboxylate carrier (AtDTC) [2]. This updated method is described in this chapter.

Another method for WGCF synthesis of integral MPs is through the use of mixed detergents (Fig. 1b). Through the use of detergents, MPs are effectively solubilized, allowing for simplified purification. Recently, we reported mixed-detergent-supplemented

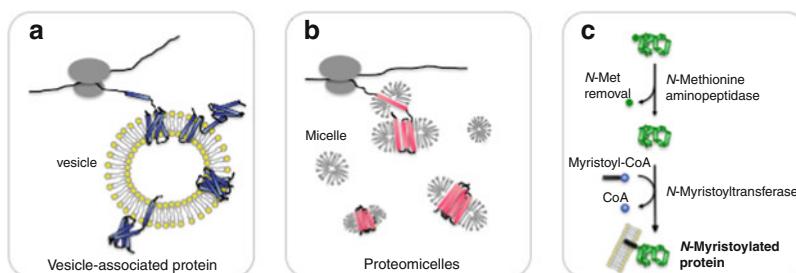


Fig. 1 Modified CF synthesis system for MP production. (a) CF synthesis of MPs in the presence of liposomes. Synthesized MPs are associated with liposomes. (b) CF synthesis of MPs in the presence of mixed detergent. MPs are synthesized in the soluble fraction as proteomicelles. (c) CF synthesis of MPs in the presence of myristic acid. Protein that has a myristylation motif is co-translationally *N*-myristoylated in the presence of myristic acid by the function of intrinsic *N*-methionine aminopeptidase and *N*-myristoyltransferase

WGCF synthesis and purification of properly folded bacteriorhodopsin (bR), which is described below [3].

Lastly, we describe method for WGCF synthesis of peripheral MP modified by *N*-myristoylation (Fig. 1c). In the optimized condition reported here, *N*-myristoylated proteins produced by WGCF system are membrane bound.

2 Materials

2.1 Preparation of mRNA

1. *Ex Taq* DNA polymerase (5 units/μl) and 10× *Ex Taq* buffer (Takara, Shiga, Japan).
2. 2.5 mM dNTP mixture: 2.5 mM each of dATP, dCTP, dGTP, and dTTP.
3. 5× transcription buffer: 400 mM HEPES-KOH, pH 7.8, 80 mM magnesium acetate, 10 mM spermidine, 50 mM DTT.
4. 25 mM NTP mixture: 25 mM each of ATP, CTP, GTP, and UTP.
5. Ribonuclease (RNase) inhibitor (80 units/μl) and SP6 RNA polymerase (80 units/μl) (Promega, Madison, WI).

2.2 Preparation of Liposomes

1. Soybean asolectin (Fluka, Buchs, Switzerland).
2. Chloroform and acetone (Nacalai tesque, Kyoto, Japan).
3. Sonicator (Digital Sonifier model 250 D, 200 W, 20 kHz) (Branson, Danbury, CT).
4. Mini extruder (Avanti Polar Lipids, Alabaster, AL).

2.3 WGCF Protein Synthesis

1. WG extract (WEPRO1240, OD₂₆₀ = 240, and WEPRO7240H, OD₂₆₀ = 240) (CellFree Science Co. Ltd., Matsuyama, Japan).
2. Creatine kinase (Roche Diagnostics K. K., Tokyo, Japan) is dissolved in water at 40 mg/ml and stored in aliquots at -80 °C.
3. Dialysis buffer (4×): 120 mM HEPES-KOH (pH 7.8), 400 mM potassium acetate, 10.8 mM magnesium acetate, 1.6 mM spermidine, 10 mM dithiothreitol, 4.8 mM ATP, 1 mM GTP, 64 mM creatine phosphate, 0.02 % NaN₃, and 1.2 mM of each amino acid. Store at -80 °C (see Note 1).
4. Fos-choline 14 (FC14) (Anatrace, Maumee, OH).
5. 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) (Nacalai tesque).
6. FC14/CHAPS mixture: 25 mM FC14 and 37.5 mM CHAPS.
7. 100 mM retinal: Dissolve 28.4 mg of all-trans retinal (Wako Pure Chemical, Osaka, Japan) in 1 ml dimethyl sulfoxide (sterile-filtered, Nacalai tesque). Store at -20 °C.

8. Dialysis cup (12,000 MWCO) (Cosmo Bio, Tokyo, Japan).
9. Receptacle tube (Maruemu yohki No. 2) (Maruemu, Osaka, Japan).
10. [¹⁴C]Myristic acid (American Radiolabeled Chemicals, St Louis, MO).
11. 3.75 mM myristic acid: Dissolve 8.56 mg of myristic acid (Nacalai tesque) in 10 ml ethanol.
12. [¹⁴C]Leucine (PerkinElmer Japan, Yokohama, Japan).

2.4 Functional Analysis of AtDTC

1. 2× Substrate solution: 60 mM PIPES–NaOH, pH 6.5, 15 mM potassium gluconate, 50 mM 2-oxoglutarate.
2. Dowex AG 1×8 (100–200 mesh) (Bio-Rad, Tokyo, Japan).
3. Pasteur pipette (5 3/4") (Asahi Glass Co., Ltd., Yokohama, Japan).
4. Equilibration solution: 10 mM PIPES–NaOH, pH 6.5, 40 mM potassium gluconate, 100 mM sodium gluconate.
5. [¹⁴C]2-Oxoglutarate (PerkinElmer Japan).
6. 2-Oxoglutarate solution: 4.2 mM 2-oxoglutarate (3 Ci/mol).
7. Stop solution: 360 mM pyridoxal 5'-phosphate, 64 mM mer-salyl acid.

2.5 Purification of bR

1. Ni-NTA-agarose (Qiagen, Tokyo, Japan).
2. *n*-Dodecyl- β -D-maltopyranoside (Anatrace).
3. Buffer A: 50 mM sodium phosphate, pH 7.4, 300 mM NaCl, 20 mM imidazole, and 0.15 % *n*-dodecyl- β -D-maltopyranoside.
4. Buffer B: 50 mM sodium phosphate, pH 7.4, 300 mM NaCl, 50 mM imidazole, and 0.15 % *n*-dodecyl- β -D-maltopyranoside.
5. Buffer C: 50 mM sodium phosphate, pH 7.4, 300 mM NaCl, 250 mM imidazole, and 0.05 % *n*-dodecyl- β -D-maltopyranoside.
6. Poly-Prep Chromatography Column (Bio-Rad, Richmond, CA).

2.6 Radio Isotope Imaging

1. Imaging analyzer (BAS2000 imaging system) (Fuji Photo Film Co., Ltd., Tokyo, Japan).
2. BAS cassette2 2040 (Fuji Photo Film Co., Ltd).
3. Imaging plate (BAS-MS 2040) (Fuji Photo Film Co., Ltd).

2.7 Accudenz Density Gradient Ultracentrifugation

1. Accudenz (Accurate Chemical and Scientific, Westbury, NY).
2. Ultracentrifugation buffer: 30 mM HEPES–KOH, pH 7.8, 100 mM potassium acetate, 2.7 mM magnesium acetate. Store at room temperature.
3. Ultracentrifugation tube (2.2 PA) and rotor (S55S) (Hitachi Koki, Tokyo, Japan).

3 Methods

3.1 *In Vitro Transcription of mRNA from Plasmid DNA*

In the following Subheadings 3.1–3.7, liposome-supplemented WGCF system for integral MP production is described. This method is optimized for the synthesis of functional AtDTC proteins. The typical yield of MPs produced in the bilayer method [4] described here is usually 1–10 µg/150 µl reaction [5]. These amounts are sufficient for functional analysis. The transport assay of AtDTC is also described.

1. Prepare 50 µl of transcription mixture as follows: 10 µl of 5× transcription buffer, 6 µl of 25 mM NTP mix, 5 µl of template plasmid (1 µg/µl), 0.5 µl of RNase inhibitor, 0.625 µl of SP6 RNA polymerase, and 27.875 µl of Milli-Q water.
2. Incubate at 37 °C for 3–6 h.
3. Centrifuge the transcription mixture at 20,000 $\times g$ for 1 min.
4. Transfer the supernatant containing the mRNA for translation to new centrifuge tube (see Note 2).

3.2 *Preparation of Liposomes*

1. Dissolve 10 g of asolectin in 30 ml of chloroform.
2. Add 180 ml of ice-cold acetone to the solution and stir the suspension on a magnetic stirrer for 2 h at room temperature.
3. Turn off the stirrer and allow the solution to stand overnight at 4 °C in order to precipitate phospholipids.
4. Aspirate the supernatant as much as possible using a glass pipette and dry the pellet completely under a flow of nitrogen gas.
5. Store the dried phospholipid mixture at –20 °C until use.
6. Prepare the liposome suspension at 5 mg/ml in Milli-Q water.
7. Sonicate the liposome suspension (10 % amplitude and 30 % duty cycle) on ice until the appearance changes from milky to nearly transparent.
8. Prepare unilamellar liposomes by extrusion through a mini extruder. To ensure proper sizing, pass the liposome suspension 11 times through a 0.4 µm polycarbonate membrane each followed by a 0.1 µm membrane (see Note 3).

3.3 *Bilayer Synthesis of AtDTC Proteins by Liposome-Supplemented WGCF System*

1. Thaw WG extract (WEPRO 1240), creatine kinase, and dialysis buffer on ice. Keep all reagents on ice while in use. Freeze promptly afterwards. Prepare 25 µl of translation mixture on ice as follows: 6.25 µl of WG extract, 0.25 µl of 40 mg/ml creatine kinase, 4.7 µl of 4× dialysis buffer, 2.5 µl of 5 mg/ml liposomes, 7.5 µl of mRNA, and 3.8 µl of Milli-Q water.
2. Prepare 125 µl of substrate mixture on ice as follows: 31.25 µl of 4× dialysis buffer, 93.75 µl of Milli-Q water.

3. Transfer the 125 μ l of substrate mixture to a microtiter plate.
4. Carefully pipette the 25 μ l of translation mixture underneath the substrate mixture to form the bilayer reaction.
5. Seal the plate to avoid evaporation.
6. Incubate the plate at 26 °C for 16 h.

3.4 Preparation of Dowex resin

1. Pour 40 g of Dowex resin into a beaker.
2. Add 400 ml of 1 N HCl and stir for 1 h at room temperature.
3. Pour the resin suspension into Buchner funnel lined with filter paper.
4. After the fluids filter through, transfer the Dowex resin back into the beaker.
5. Repeat steps 2–4 with 400 ml of 1 N NaOH followed by 400 ml of 200 mM sodium acetate.
6. To the beaker of Dowex resin add 400 ml of 200 mM sodium acetate and stir for 1 h at room temperature (see Note 4).
7. Plug the end of Pasteur pipette with absorbent cotton using a stick (see Note 5).
8. Pour the Dowex resin into the Pasteur pipette. The resin bed height should be approximately 5 cm for columns to remove unincorporated substrates from liposomes, or 4.2 cm for transport assay columns.
9. Equilibrate these resins by passing either 6 ml of equilibration solution for columns to remove unincorporated substrates from liposomes, or 5.5 ml of 200 mM sodium acetate for transport assay columns.

3.5 Preparation of Substrate-Preloaded Liposomes

1. Suspend 80 mg of the acetone-washed phospholipid mixture with 1 ml of 2 \times substrate solution.
2. Sonicate the suspension (10 % amplitude and 30 % duty cycle) on ice until the suspension changes in appearance from milky to nearly transparent.

3.6 Preparation of Transport Assay Proteoliposomes

1. After protein synthesis, mix the bilayer reaction mixture (150 μ l) gently, and transfer it to a new microfuge tube.
2. Centrifuge the reaction mixture (20,000 \times g, 20 min, 4 °C) (see Note 6).
3. Resuspend the pellet fraction in 150 μ l of 10 mM PIPES–NaOH, pH 6.5.
4. Sonicate the resuspended pellet at room temperature for 18 s (10 % amplitude and 50 % duty cycle).
5. To verify expression of the target protein, analyze the resuspended pellet by sodium dodecyl sulfate-polyacrylamide gel

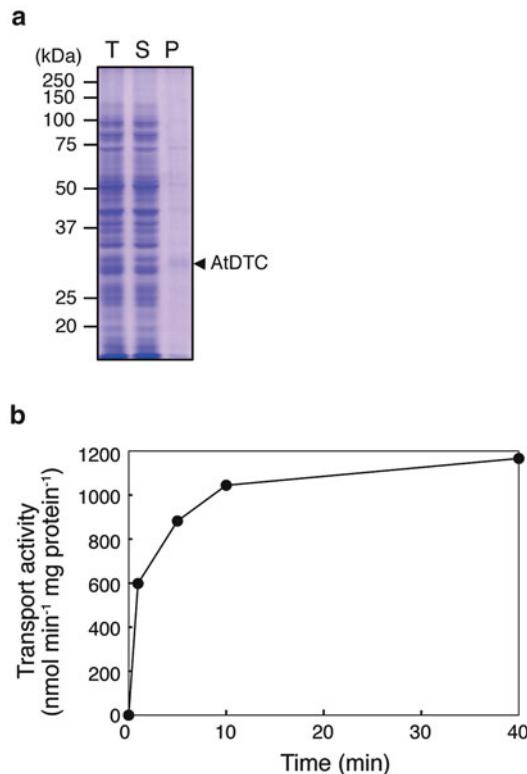


Fig. 2 Synthesis and characterization of AtDTC. (a) Detection of synthesized AtDTC proteins by CBB staining. The proteins were synthesized by liposome-supplemented WGCF system. After protein synthesis, the reaction mixture was centrifuged at $20,000 \times g$ for 20 min at 4°C . The total (T), supernatant (S), and pellet (P) fractions were subjected to SDS-PAGE and analyzed by CBB staining. (b) Time course of the incorporation of external substrate [^{14}C]2-oxoglutarate into proteoliposomes reconstituted with AtDTC. AtDTC was synthesized by liposome-supplemented WGCF system. The liposomes and synthesized protein complexes were sedimented by centrifugation, and reconstituted into proteoliposomes preloaded with 30 mM of 2-oxoglutarate. Intake of [^{14}C]2-oxoglutarate was measured (reproduced from ref. 2)

electrophoresis (SDS-PAGE) and Coomassie brilliant blue (CBB) staining (Fig. 2a).

6. Mix 150 μl of the sonicated resuspended pellet with 150 μl of substrate-preloaded liposomes.
7. Sonicate the mixed liposomes at room temperature for 18 s (10 % amplitude and 50 % duty cycle). This is the proteoliposome solution.
8. Freeze the proteoliposome solution with liquid nitrogen (*see Note 7*).
9. Thaw the proteoliposome solution at room temperature.

10. Sonicate the proteoliposome solution at room temperature for 18 s (10 % amplitude and 50 % duty cycle).
11. In order to remove unincorporated substrates, apply the proteoliposome solution to a 5 cm bed-volume Dowex column. Discard the flow-through.
12. Apply 300 μ l of equilibration solution to the 5 cm bed-volume Dowex column and discard the flow-through.
13. Apply 500 μ l of equilibration solution to the Dowex column and collect the elution. This fraction contains the transport assay proteoliposomes.
14. Divide the transport assay proteoliposome fraction into 100 μ l of aliquots.
15. Store on ice until use (see Note 8).

3.7 Transport Assay of AtDTC

1. Pre-incubate 100 μ l of transport assay proteoliposomes at 25 °C for 2 min.
2. Add 5 μ l of 2-oxoglutarate solution and mix gently.
3. Incubate at 25 °C for 0, 1, 5, 10, or 40 min.
4. Add 15 μ l of stop solution and mix gently.
5. Apply the proteoliposome solution to a 4.2 cm bed-volume Dowex column to eliminate any 2-oxoglutarate remaining outside of the transport assay proteoliposomes. Discard the flow-through.
6. Apply 700 μ l of 200 mM sodium acetate to the column and collect the elution. This fraction contains proteoliposomes.
7. The radioactivity associated with the eluted proteoliposomes is measured with a liquid scintillation spectrometer.
8. The result of a typical AtDTC transport assay is shown in Fig. 2b.

3.8 Preparation of bR mRNA from a PCR Reaction

In the following Subheadings 3.8–3.10, a method for the synthesis of integral MPs using WGCF system in the presence of mixed detergents is described. Recently, we reported that a combination of FC14 and CHAPS is effective for the WGCF synthesis and purification of detergent-solubilized bR in high yield [3]. Furthermore, bR synthesized and purified from the mixed detergent method was confirmed to be functional by measuring the decrease in fluorescence intensity of acridine orange after reconstitution into liposomes (data not shown).

1. Prepare a 50 μ l of PCR reaction mixture as follows: 1 μ l of template plasmid (pEU3SH-bR [3], 1 ng/ μ l), 1.5 μ l of 10 μ M pEU-SPU primer (5'-ACATACGATTAGGTGACACT-3'), 1.5 μ l of 10 μ M Anti3 primer (5'-GGAGAAAGGCGGACAGGTAT-3'), 4 μ l of 2.5 mM dNTP mixture, 5 μ l of 10 \times *Ex Taq* buffer, 0.25 μ l of *Ex Taq* polymerase, and 36.75 μ l of Milli-Q water.

2. The PCR condition is 35 successive cycles of denaturation at 94 °C for 30 s, annealing at 58 °C for 30 s, and elongation at 72 °C for 2.5 min, followed by a final elongation step at 72 °C for 10 min.
3. Verify amplification of DNA by agarose gel electrophoresis. The PCR reaction mixture is directly used as template for in vitro transcription.
4. Prepare 50 µl of transcription mixture as follows: 10 µl of 5× transcription buffer, 6 µl of 25 mM NTP mix, 12.5 µl of PCR reaction mixture (template), 0.5 µl of RNase inhibitor, 0.625 µl of SP6 RNA polymerase, and 20.375 µl of Milli-Q water.
5. Incubate at 37 °C for 3–6 h.
6. Centrifuge the transcription mixture at 20,000×*g* for 1 min.
7. Transfer the supernatant to a new centrifuge tube.
8. To the supernatant, add 6.5 µl of 7.5 M ammonium acetate and 125 µl of ethanol. Mix well and incubate on ice for 15 min.
9. Centrifuge the mixture at 20,000×*g* for 20 min at 4 °C.
10. Discard the supernatant, and rinse the pellet with 500 µl of 70 % ethanol.
11. Dry the pellet and dissolve the dried pellet in 22.1 µl of Milli-Q water.

3.9 Dialysis-Mode Synthesis of bR Proteins by Mixed-Detergent- Supplemented WGCF System

1. Dilute 100 mM retinal with dimethyl sulfoxide to make 10 mM retinal.
2. Thaw WG extract (WEPRO 7240H), creatine kinase, and dialysis buffer on ice. Keep all reagents on ice while in use. Freeze promptly afterwards. Prepare 50 µl of translation mixture on ice as follows: 12.5 µl of WG extract, 2 µl of 40 mg/ml creatine kinase, 9.4 µl of 4× dialysis buffer, 2 µl of FC14/CHAPS mixture, 22.1 µl of mRNA, 1 µl of RNase inhibitor, and 1 µl of 10 mM retinal.
3. Prepare 800 µl of substrate mixture on ice as follows: 200 µl of 4× dialysis buffer, 32 µl of FC14/CHAPS mixture, and 568 µl of Milli-Q water.
4. Transfer the 50 µl of translation mixture into a dialysis cup (*see Note 9*).
5. Transfer 800 µl of the substrate mixture into a receptacle tube.
6. Place the dialysis cup containing the translation mixture into the receptacle tube.
7. Seal the connected portion with parafilm to avoid evaporation.
8. Incubate the reaction mixture at 26 °C for 16 h. The typical yield of dialysis-mode translation is 5–20 µg/50 µl of translation mixture.

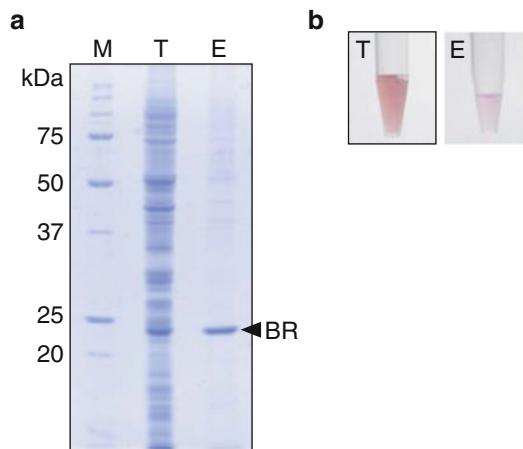


Fig. 3 Purification of bR. **(a)** Purification of bR. His₆-tagged bR synthesized in the presence of 1 mM FC14 and 1.5 mM CHAPS was purified using Ni-NTA-agarose, and visualized by SDS-PAGE and CBB staining. Lane M, molecular size markers; lane T, 1 μ l of translation mixture; lane E, 2 μ l of the imidazole-eluted fraction. **(b)** Photographs of the translation mixture (T) and the imidazole-eluted fraction (E). The *purple color* indicates that the protein was folded correctly (reproduced from ref. 3)

3.10 Purification of bR Protein

1. Equilibrate 200 μ l of Ni-NTA-agarose with buffer A.
2. Incubate 200 μ l of the translation mixture with 200 μ l of the equilibrated Ni-NTA-agarose and 4 ml of buffer A at 4 °C for 1 h with gentle shaking.
3. Load the Ni-NTA-agarose into a gravity-flow column. Save the flow-through for purification analysis.
4. Add 0.5 ml of buffer B to wash out nonspecifically bound proteins. Repeat this step four times.
5. Add 0.2 ml of buffer C to the column.
6. Evaluate the purification by SDS-PAGE and CBB staining (Fig. 3).

3.11 Batch-Mode Synthesis of Myristoylated AtCBL1 in the Presence of [¹⁴C]Myristic Acid

N-myristylation is a co-translational modification of N-terminal glycine residues. This modification is catalyzed by myristoyl-CoA: protein N-myristoyltransferase, and involves the covalent attachment of myristic acid to the N-terminal Gly that is exposed during co-translational elimination of the initiating Met by methionine aminopeptidase [6, 7]. N-myristylation facilitates the interaction of target proteins with membranes by anchoring the proteins through myristoyl group. Recently, we reported the consensus motif for N-myristylation in a WGCF system [8]. In that publication, a method for synthesis of N-myristoylated proteins by the WGCF system was established. In the following Subheadings 3.11–3.14, this method using AtCBL1 as a model protein is described. AtCBL1 is a calcium-binding protein belonging to calcineurin

B-like protein family. Some calcineurin B-like family proteins are reported to regulate transporter functions by connecting calcineurin B-like protein-interacting protein kinases with target transporter proteins [9].

1. Prepare AtCBL1 mRNA by in vitro transcription as described above (Subheading 3.1).
2. Thaw WG extract (WEPRO 1240), creatine kinase, and dialysis buffer on ice. Keep all reagents on ice while in use. Freeze promptly afterwards. Prepare 25 μ l of reaction mixture on ice as follows: 6.25 μ l of WG extract, 0.25 μ l of 40 mg/ml creatine kinase, 4.7 μ l of 4 \times dialysis buffer, 0.5 μ l of [14 C]myristic acid, 7.5 μ l of mRNA, and 5.8 μ l of Milli-Q water.
3. Incubate the reaction mixture at 26 °C for 3 h.
1. Subject the reaction mixture to SDS-PAGE.
2. After electrophoresis, dry the gel with a gel dryer.
3. Put the dried gel into the BAS cassette2 2040 with the imaging plate.
4. Visualize the labeled proteins by the imaging analyzer (Fig. 4a).

3.12 Detection of Myristoylated AtCBL1 with Imaging Analyzer

3.13 Bilayer Synthesis of Myristoylated AtCBL1 in the Presence of [14 C]Leucine

1. Prepare AtCBL1 mRNA by in vitro transcription as described above (Subheading 3.1).
2. Thaw WG extract (WEPRO 1240), creatine kinase, and dialysis buffer on ice. Place and keep all reagents on ice. Prepare 25 μ l of translation mixture on ice as follows: 6.25 μ l of WG extract, 0.25 μ l of 40 mg/ml creatine kinase, 4.7 μ l of 4 \times dialysis buffer, 0.5 μ l of 3.75 mM myristic acid, 0.5 μ l of [14 C]leucine, 5 μ l of 50 mg/ml liposomes, 7.5 μ l of mRNA, and 0.3 μ l of Milli-Q water.
3. Prepare 125 μ l of substrate mixture on ice as follows: 31.25 μ l of 4 \times dialysis buffer, 2.5 μ l of 3.75 mM myristic acid, 2.5 μ l of [14 C]leucine, and 88.75 μ l of Milli-Q water.
4. Transfer 125 μ l of substrate mixture to a microtiter plate.
5. Carefully pipette the 25 μ l of translation mixture underneath the substrate mixture to form the bilayer reaction.
6. Seal the plate to avoid evaporation.
7. Incubate the plate at 15 °C for 24 h.

3.14 Accudenz Density Gradient Ultracentrifugation

Accudenz is a nontoxic reagent used for fractionation of proteins, organelles, and cells [10–12]. Accudenz density ultracentrifugation separates liposomes from WG extract proteins. The liposomes in the CF reaction mixture float to the top of the tube as their contained buffer is less dense than Accudenz. N-myristylation of AtCBL1 is confirmed by co-flootation of protein with liposomes.

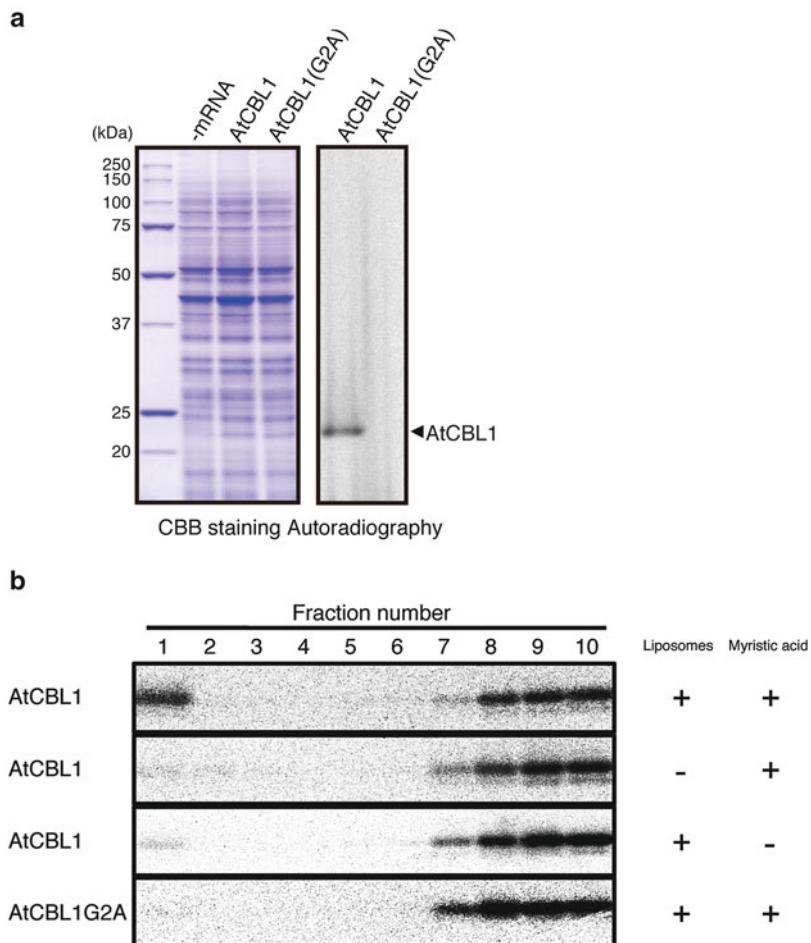


Fig. 4 Synthesis of myristoylated AtCBL1. **(a)** Synthesis of myristoylated AtCBL1. AtCBL1 and AtCBL1(G2A) were synthesized by batch method in the presence of [14 C]myristic acid. After the translation reaction, the reaction mixture was visualized by SDS-PAGE and either CBB staining (*left*) or autoradiography (*right*). **(b)** Accudenz density gradient ultracentrifugation. AtCBL1 and AtCBL1(G2A) were synthesized by the WGCF system in the presence or the absence of liposomes and myristic acid. The reaction mixture was subjected to Accudenz density gradient ultracentrifugation. Fractions were collected from the top of the tube and subjected to SDS-PAGE and autoradiography

As shown in Fig. 4b, AtCBL1 is found at the top of the centrifuge tube when AtCBL1 is synthesized in the presence of myristic acid and liposomes. This technique is useful for rough purification of N-myristoylated proteins.

1. Prepare 30, 35, and 80 % (w/v) Accudenz solution in ultracentrifugation buffer.
2. Mix 150 μ l of the translation mixture with 150 μ l of ultracentrifugation buffer and 300 μ l of 80 % Accudenz solution.
3. Transfer 600 μ l of the resultant 40 % Accudenz solution containing the translation mixture to an ultracentrifugation tube.

4. Overlay 650 μ l of 35 % Accudenz solution, 650 μ l of 30 % Accudenz solution, and 100 μ l of ultracentrifugation solution.
5. Centrifuge the density gradient tube for 1 h at 206,000 $\times g$ and 4 °C.
6. Fractionate the density gradient by carefully removing 200 μ l of aliquots from the top to the bottom of the tube (see Note 10).
7. To determine which fractions contain protein, subject fractions to SDS-PAGE and detect the [¹⁴C]-labeled AtCBL1 with imaging analyzer as described above (Subheading 3.12) (Fig. 4b).

4 Notes

1. All solutions are prepared using Milli-Q water prepared by Autopure WT101UV (Yamato Scientific Co., Ltd., Tokyo, Japan) with BioPak filter (Nihon Millipore K.K., Tokyo, Japan), in order to eliminate deoxyribonucleases, RNases, and pyrogens.
2. The white pellet is magnesium pyrophosphate, a by-product of mRNA synthesis.
3. We freshly prepare liposomes for each experiment and store them on ice until use.
4. The Dowex resin can be stored at 4 °C until use.
5. Micro spatulas are useful for plugging the end of Pasteur pipettes with absorbent cotton.
6. The synthesized MPs co-sediment with liposomes by centrifugation.
7. These proteoliposomes can be stored at -80 °C for at least 1 week.
8. We use these proteoliposomes for transport assay within 1 h.
9. Before use, wash the inside of the dialysis cups and receptacle tubes with Milli-Q water. Ensure that dialysis cup does not leak. Receptacle tubes can be autoclaved and reused.
10. After ultracentrifugation, liposomes are found at the interface between the 30 % Accudenz and ultracentrifugation solution layers.

References

1. Nozawa A, Nanamiya H, Miyata T et al (2007) A cell-free translation and proteoliposome reconstitution system for functional analysis of plant solute transporters. *Plant Cell Physiol* 48:1815–1820
2. Nozawa A, Fujimoto R, Matsuoka H et al (2011) Cell-free synthesis, reconstitution, and characterization of a mitochondrial dicarboxylate-tricarboxylate carrier of *Plasmodium falciparum*. *Biochem Biophys Res Commun* 414:612–617
3. Genji T, Nozawa A, Tozawa Y (2010) Efficient production and purification of functional bacteriorhodopsin with a wheat-germ cell-free

- system and a combination of Fos-choline and CHAPS detergents. *Biochem Biophys Res Commun* 400:638–642
4. Sawasaki T, Hasegawa Y, Tsuchimochi M et al (2002) A bilayer cell-free protein synthesis system for high-throughput screening of gene products. *FEBS Lett* 514:102–105
 5. Nozawa A, Ogasawara T, Matsunaga S et al (2011) Production and partial purification of membrane proteins using a liposome-supplemented wheat cell-free translation system. *BMC Biotechnol* 11:35
 6. Boutin JA (1997) Myristoylation. *Cell Signal* 9:15–35
 7. Bradshaw RA, Brickey WW, Walker KW (1998) N-terminal processing: the methionine aminopeptidase and N α - acetyl transferase families. *Trends Biochem Sci* 23:263–267
 8. Yamauchi S, Fusada N, Hayashi H et al (2010) The consensus motif for N-myristoylation of plant proteins in a wheat germ cell-free translation system. *FEBS J* 277:3596–3607
 9. Weinl S, Kudla J (2009) The CBL-CIPK Ca²⁺-decoding signaling network: function and perspectives. *New Phytol* 184:517–528
 10. Sobrado P, Goren MA, James D et al (2008) A Protein Structure Initiative approach to expression, purification, and in situ delivery of human cytochrome b5 to membrane vesicles. *Protein Expr Purif* 58:229–241
 11. Graham JM, Ford T, Rickwood D (1990) Isolation of the major subcellular organelles from mouse liver using Nycodenz gradients without the use of an ultracentrifuge. *Anal Biochem* 187:318–323
 12. Sbracia M, Sayme N, Grasso J et al (1996) Sperm function and choice of preparation media: comparison of Percoll and Accudenz discontinuous density gradients. *J Androl* 17: 61–67

Part III

Biological Systems

Chapter 20

Arabidopsis Proteomics: A Simple and Standardizable Workflow for Quantitative Proteome Characterization

Anja Rödiger, Birgit Agne, Katja Baerenfaller, and Sacha Baginsky

Abstract

Arabidopsis is the model plant of choice for large-scale proteome analyses, because its genome is well annotated, essentially free of sequencing errors, and relatively small with little redundancy. Furthermore, most Arabidopsis organs are susceptible to standard protein solubilization protocols making protein extraction relatively simple. Many different facets of functional plant proteomics were established with Arabidopsis such as mapping the subcellular proteomes of organelles, proteo-genomic peptide mapping, and numerous studies on the dynamic changes in protein modification and protein abundances. As most standard proteomics technologies are now routinely applied, research interest is increasingly shifting towards the reverse genetic characterization of gene function at the proteome level, i.e., by profiling the quantitative proteome of wild type in comparison with mutant plant tissue. We report here a simple, standardizable protocol for the large-scale comparative quantitative proteome characterization of different Arabidopsis organs based on normalized spectral counting and suggest a statistical framework for data interpretation. Based on existing organellar proteome maps, proteins can be assigned to organelles, thus allowing the identification of organelle-specific responses.

Key words Arabidopsis proteomics, Quantitative plant proteomics, Spectral counting

Abbreviations

| | |
|------|--------------------------------------|
| nSpC | Normalized spectral counting |
| MTP | Measured spectra of tryptic peptides |
| TTP | Theoretical tryptic peptides |

1 Introduction

The systematic identification, quantification, and characterization of proteins is an important strategy to characterize the effect of a gene defect at large scale. It is expected that comparative quantitative proteome analyses between wild-type and mutant plants will provide information about the consequences of a gene defect that reflects the functional footprint of a gene much closer than most

other characterization approaches. Historically, quantitative analyses at the proteome level were performed by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and by the visual, computer-aided comparison between the protein spot pattern of mutant and wild-type plant material [1]. Because 2D-PAGE based on isoelectric focusing (IEF) is technically demanding, restricted to soluble proteins, and, depending on the choice of the staining, quite expensive, alternative approaches that rely on mass spectrometry for peptide or protein quantification were developed [2]. Among the different methods that exist, the linear positive correlation between the number of spectra that identify a protein and protein abundance has attracted considerable attention for quantification purposes [3]. Different flavors of the so-called spectral counting methods were reported and the most advanced methods use sophisticated normalization strategies that take into account physicochemical features of peptides and their probability of being detected in a complex protein mixture [4].

For several reasons, *Arabidopsis thaliana* is excellently suited for large-scale proteome characterization because of its small genome size with well-annotated genes and few sequencing errors. Furthermore, it is prone to genetic manipulations and connects a strong and devoted research community that is willing to share tools and knowledge (see www.arabidopsis.org). Last but not least, the community can utilize information from many organellar proteome studies such that for many proteins their subcellular localization is known. Based on the latter efforts, organellar proteome maps were built that allow performing proteome analyses at the level of the entire cell. Characteristics of organellar proteomes are then inferred by assigning identified proteins to cell organelles *a posteriori*, i.e., after the experiment. Since the study of dynamic proteome changes is incompatible with lengthy organelle isolation procedures, this approach is preferred to study dynamic quantitative changes in protein abundance or status of posttranslational modification in response to a signal. Examples of successful proteome characterizations include studies on the plastid protein import mutant *ppi2* [5], the *ALLENE OXIDE SYNTHASE (aos)* mutant [6], the *clp* protease mutants [7], and the analysis of kinase targets by phosphoproteome profiling [8].

We report here a simple experimental workflow for the characterization of a mutation at the proteome level using protein extracts from *Arabidopsis thaliana* root and leaf (Fig. 1). Our quantification approach relies on normalized spectral counting (nSpC) and is therefore restricted to fast scanning trap instruments that acquire peptide-centric data at high temporal resolution. While higher data quality and deeper proteome coverage can be obtained with Orbitrap- or FTICR-instruments, the workflow reported here is similarly suitable for an LTQ ion trap. The downside of using the

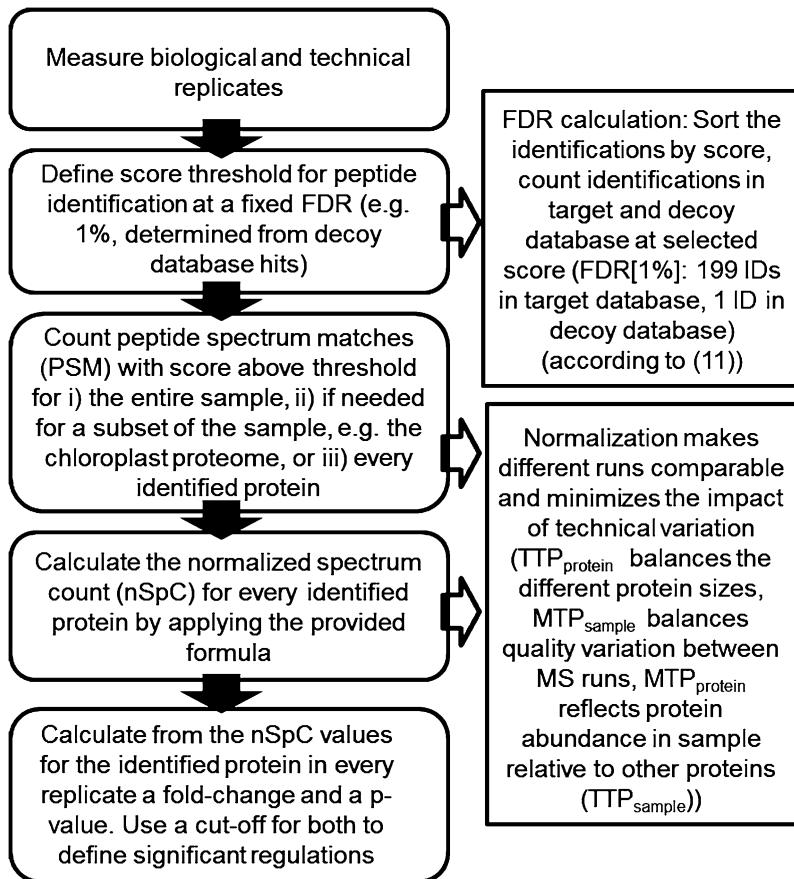


Fig. 1 Flow-scheme of the described methods. *ID* identification, $TTP_{protein}$ theoretical tryptic peptides from the protein of interest, MTP_{sample} measured peptide spectrum matches of the entire sample, $MTP_{protein}$ measured peptide spectrum matches for the protein of interest, retical tryptic peptides from all proteins in sample

inexpensive LTQ ion trap is that mass accuracy is much lower compared to Orbitrap- or FTICR-instruments, so higher identification cutoffs must be applied to decrease the false discovery rate in the dataset because of the higher mass tolerance that needs to be applied in the database-dependent searches. Thus, protein identification numbers decrease because potentially correct identifications must be ignored. When we compare identification rates in comparable experiments between LTQ ion traps and higher accuracy instruments, we find a factor of 2–3 less identifications at the same calculated FDR. With increasing number of reference spectra deposited in databases, LTQ ion trap-derived MS/MS spectra can be compared to those derived from Orbitrap- or FTICR-instruments since MS/MS spectra should be comparable, thus further decreasing the risk of false-positive assignments.

2 Materials

2.1 Protein Extraction

1. Shock-frozen plant tissue of interest (e.g., whole leaves and roots, *see Note 1*). Plan your experiment and the material requirement for at least three to four independent biological replicates [9] (*see Note 2*).
2. Mortar and pestle.
3. Extraction buffer: 40 mM Tris–HCl, pH 6.8, 10 % (v/v) glycerol.
4. Inhibitor cocktail for proteases (e.g., cComplete, EDTA-free Protease Inhibitor Cocktail, Roche, Basel, Switzerland).
5. 20 % (w/v) SDS.
6. 2 M Dithiothreitol (DTT).
7. 1 % (w/v) Bromophenol blue (optional).
8. Heating block.
9. Standard protein assay (e.g., Bradford, bicinchoninic acid (BCA)) (*see Note 3*).
10. Spectrophotometer.
11. Benchtop centrifuge.

2.2 SDS-PAGE

1. Precast Tris–Glycine (Laemmli) SDS-PAGE gel with a uniform, 12 % acrylamide separation gel (*see Note 4*). Recommended dimensions: 18 cm × 16 cm × 1 mm, recommended comb: 10 wells, 1 mm thickness.
2. Vertical slab gel electrophoresis apparatus (e.g., Hoefer SE600).
3. 1–200 µl Gel-saver loading tip or Hamilton syringe.
4. 1× Tris–Glycine (Laemmli) SDS-PAGE running buffer.
5. Molecular weight marker.
6. Power supply.
7. Coomassie brilliant blue staining solution: 1 g Coomassie brilliant blue R250, 0.25 g Coomassie brilliant blue G250, 210 ml ethanol, 25 ml methanol, 50 ml acetic acid, 215 ml H₂O.
8. Coomassie destaining solution: 45 % (v/v) methanol, 10 % (v/v) acetic acid, 45 % (v/v) H₂O.

2.3 Cutting the Gel

1. Glass plate.
2. Scalpel.
3. Eppendorf Safe Lock tubes.

2.4 In-Gel Digest (See Note 5)

1. Ammonium hydrogen carbonate buffer: 100 mM NH₄HCO₃ (adjust pH 8.5 with NH₄OH).
2. Destaining solution: Bring 30 ml acetonitrile (ACN) to a final volume of 100 ml with ammonium hydrogen carbonate buffer.

3. Reducing solution: 10 mM DTT in 100 mM NH₄HCO₃, pH 8.5.
4. Alkylation solution: 54 mM iodacetamide in 100 mM NH₄HCO₃, pH 8.5.
5. Extraction solution: 50 % ACN, 0.1 % trifluoroacetic acid (TFA).
6. Trypsin solution: 3 ng/μl Trypsin (Promega: Sequencing Grade Modified Trypsin V5113), 50 mM NH₄HCO₃, pH 8.5, 5 % ACN.
7. SpeedVac or vacuum concentrator.
8. Spatula or needle for cutting gel pieces.
9. Thermomixer.

2.5 MS Analysis

1. 2 % ACN, 0.1 % formic acid (FA).
2. Ultrasonic bath.
3. HPLC vials.
4. NanoHPLC (Dionex UltiMate® 3000 RSL nano system with autosampler).
5. Trap column: Acclaim® PepMap100 (C18, 3 μm, 100 Å).
6. Analytical column: Acclaim® PepMap RSLC (C18, 2 μm, 100 Å).
7. Solvent A: H₂O, 0.1 % FA; solvent B: ACN, 0.1 % FA.
8. LTQ-XL, LTQ-Orbitrap XL, LTQ Orbitrap Velos, FTICR (Thermo Scientific; NSI).
9. Software for data interpretation (see also Methods for the different options).
10. Arabidopsis TAIR10 database with common contaminants included.

3 Methods

3.1 Protein Extraction

1. Precool mortar and pestle.
2. Chill extraction buffer on ice, and freshly add protease inhibitors (e.g., doubly concentrated complete Protease Inhibitor Cocktail, EDTA-free, Roche).
3. Homogenize plant tissue to a fine powder in liquid nitrogen using mortar and pestle. Store the samples in liquid nitrogen.
4. Resuspend and defrost the powder in extraction buffer (*see Note 6*).
5. Transfer samples to microcentrifuge tubes.
6. Add SDS to obtain a final concentration of 4 % (w/v), and mix gently by inverting the tube.
7. Denature proteins at 70 °C for 10 min in a heating block.

8. Centrifuge at 25 °C for 20 min at 16.000 $\times g$.
9. Transfer supernatant to a new tube.
10. Determine protein concentration using the Bradford or the BCA assay (see Note 3).
11. Prior to loading to the SDS-PAGE gel, add DTT to a final concentration of 20 mM and (optionally) bromophenol blue to a final concentration of 0.01 %.
12. For storage, freeze samples in liquid nitrogen and store at -80 °C.

3.2 SDS-PAGE (See Note 7)

1. Assemble gel apparatus, and fill with 1× Tris-Glycine (Laemmli) SDS-PAGE running buffer.
2. Carefully remove comb and rinse wells with 1× Tris-Glycine (Laemmli) SDS-PAGE running buffer using a loading tip or syringe.
3. Load protein sample yielding 200–400 µg protein per well and molecular weight marker.
4. Run the gel at constant current (25 mA/gel).
5. Stain the gel with Coomassie brilliant blue.

3.3 Cutting the Gel (See Note 8)

1. Place gel on a clean glass plate.
2. Cut each lane of the gel into 15–20 sections.
3. Excise bands with a clean scalpel and transfer to a 1.5 ml Eppendorf Safe Lock tube.

3.4 In-Gel Digest (See Note 9)

3.4.1 Destaining

1. Add 100 µl H₂O to the excised gel slices. Shake for 10 min at room temperature. Take off and discard the supernatant. Repeat this step once.
2. Add 100 µl destaining solution. Shake for 15 min at room temperature. Take off and discard the supernatant. Repeat this step four times.
3. Add 100 µl H₂O. Shake for 15 min at room temperature. Take off and discard the supernatant. Repeat this step once.
4. Add 100 µl ACN. Shake for 15 min at room temperature. Take off and discard the supernatant.
5. Dry the gel slices in a vacuum concentrator.

3.4.2 Reduction and Alkylation of SH-Groups

1. Add 40 µl reducing solution. Shake for 5 min at room temperature.
2. Incubate for 30 min at 50 °C. Take off and discard the supernatant.
3. Add 100 µl ACN. Shake for 15 min at room temperature. Take off and discard the supernatant.

4. Add 40 μ l alkylating solution. Incubate for 15 min in the dark at room temperature. Take off and discard the supernatant.
5. Add 100 μ l destaining solution. Shake for 10 min at room temperature. Take off and discard the supernatant.

3.4.3 Tryptic Digestion

1. Carefully dice gel slices into smaller pieces (approx. 2–3 mm² in size).
2. Add 50 μ l destaining solution. Shake for 10 min at room temperature. Take off and discard the supernatant.
3. Dry gel pieces in a vacuum concentrator.
4. Add 30–100 μ l trypsin solution to a protein-to-trypsin ratio of 20 to 1 (the gel pieces should be covered with the trypsin solution). Incubate overnight at room temperature (*see Note 10*).

3.4.4 Extraction of Peptides

1. Add extraction solution (same volume as trypsin solution). Shake for 40 min at room temperature. Take off the supernatant and transfer to a fresh 1.5 ml Eppendorf tube (do not discard!).
2. Add extraction solution (same volume as trypsin solution). Shake for 10 min at room temperature. Take off the supernatant and combine it with the first supernatant (**step 1**) (do not discard!).
3. Dry the peptides (collected supernatants) in the vacuum concentrator.

3.5 MS Analysis

1. Dissolve dried peptides in 20 μ l 2 % ACN and 0.1 % FA.
2. Incubate for 5 min in an ultrasonic bath.
3. Centrifuge for 3 min at 17,000 $\times g$ and transfer supernatant into an HPLC vial suitable for the autosampler.
4. For the liquid chromatography (LC) inject 1–8 μ l of the sample.
5. Loading: 5 min at 6 μ l/min on the trap column.
6. Gradient (this is a suggestion for a complex protein mixture): Use a column flow rate of 0.240 μ l/min (0–5 min 8 % solvent B; 5–90 min 8–40 % solvent B; 90–95 min 40–85 % solvent B; 95–105 min 85 % solvent B; 105–110 min 85–8 % solvent B; 110–120 min 8 % solvent B).
7. The acquisition: Cycle of survey scan followed by four data-dependent scans of the four most abundant peaks. MS data is acquired over the complete LC run.
8. Cycle time: 30 ms, scan range: 500–2,000 Da.
9. CID; normalized collision energy: 35.0 V.

3.6 Database-Dependent Peptide Spectrum Assignment

After acquisition of the tandem mass spectrometry data, the spectra need to be assigned to peptide sequences. For *Arabidopsis* with its well-annotated genome sequence released by TAIR (www.arabidopsis.org), this is usually done in a protein sequence database search with a search algorithm, which assigns a peptide sequence to the measured pattern of mass-to-charge values of the peptide fragments. Even though the general principle of this assignment is the same, a variety of search algorithms with different scoring schemes exist [10]. To control the quality of the search and to decide on a suitable score cutoff, the database searches are best performed against the *Arabidopsis* protein sequence database expanded by a concatenated decoy database [11]. The decoy database must have the same elemental composition and the same size as the target database. This way, the number of spectrum assignments in the decoy database allows assessing the spectrum false discovery rate in the dataset, or in subsets of the data applying local false discovery rate calculations. Upon defining the database search parameters, the search space should be restricted to parameters necessary for peptide identification because large search spaces will lead to lower scores and a decreased number of identifications. This issue mainly concerns the inclusion of posttranslational modifications as variable modifications in database searches. The general recommendation here is therefore to include low-abundance posttranslational modifications in database searches only if the corresponding modified peptides have been enriched. Alternatively, search algorithms such as PepSplice may be used, which carefully control the search space [12]. However, we recommend that the unexperienced research laboratory applies standard search tools such as Sequest [13] or Mascot (Matrix Science, www.matrixscience.com).

3.7 Working with Database Search Results

When analyzing and interpreting the search results of a mass spectrometry experiment, different questions may be asked from the data depending on the scientific question of the experiment. For complex experimental setups with different samples and biological replicates, multiple measurements for each fraction, one to several spectrum assignments for the same peptide sequence, posttranslational modifications at different peptide positions, and one to several peptide sequences for one protein require integration of the data in a relational database. We have therefore developed the pep2pro database and employed its capacity for the analysis of several large-scale high-throughput proteomics data [14, 15]. To make this analysis pipeline accessible to users, the pep2pro4all system has been developed, which consists of a database schema and a script that will populate the database with mass spectrometry data provided in mzIdentML format [16]. Thus, the database scheme of pep2pro can be individually used for tailored data analysis. The scheme is available from www.pep2pro4all.ethz.ch.

3.8 Quantification of Identified Proteins by Normalized Spectral Counting

As detailed in the introduction, spectral counting is an inexpensive and robust quantification technique for proteins in complex mixtures (see Note 11). This type of quantification works best for complex mixtures and yields reliable results only for very large datasets. Thus, including technical replicas will make the quantification more reliable. Spectral counting should be based on either the original APEX procedure described by Lu and colleagues [4] or the modified APEX-indexing procedure nSpC that was described by Baerenfaller and colleagues [14]. The difference of both procedures is that Lu and colleagues take peptide detection probability into account to define the set of theoretically detectable tryptic peptides (see formula below). This probability estimate is based on a prediction algorithm [17] and as such prone to errors and incapable of dealing with different elemental composition of the sample. We therefore suggest a simplification by taking all theoretically detectable tryptic peptides into account that are in principle accessible to the mass spectrometric analysis. In nSpC, the expected contribution of each individual protein to the sample total peptide and spectra pool is calculated by assuming an equimolar distribution of proteins and then correcting the assumption by a factor that is calculated from the actually measured contribution of the protein. The following formula balances between detected and expected number of spectra assigned to all tryptic peptides in the complete sample:

$$\text{Abundance Protein K} = \text{MTP}_{\text{protein K}} \times \left(\frac{\text{TTP}_{\text{protein K}} \times \text{MTP}_{\text{sample}}}{\text{TTP}_{\text{sample}}} \right)^{-1}$$

$\text{MTP}_{\text{protein K}}$ = Measured spectra of tryptic peptides from protein K (peptide spectrum matches).

$\text{TTP}_{\text{protein K}}$ = Theoretical tryptic peptides (TTP) of protein K.

$\text{MTP}_{\text{sample}}$ = Measured spectra of tryptic peptides (PSM) in sample.

$\text{TTP}_{\text{sample}}$ = TTP of proteins identified in sample.

For the determination of the number of TTP_k , we digested the whole Arabidopsis protein database (TAIR10) with trypsin *in silico*. If Arg or Lys was followed by Pro (KP/RP site), the site was both cut and not cut (resulting in three tryptic peptides). If several of these sequence pairs followed each other, we only considered cutting of one KP/RP site per time. The resulting peptides were labeled as TTP in case they have a mass between 400 and 6,000 Da and comprise at least six amino acids. The list of TTP of the entire Arabidopsis proteome is available at www.pep2pro.ethz.ch.

3.9 Alternative Normalization Schemes

Depending on the biological question to be addressed, different normalization schemes for peptide spectrum matches have been applied. When analyzing the adaptation of the plastid proteome to a mutation, especially when the mutation causes a severe defect in

plastid development, researchers suggested correcting for the actually decreased total mass ratio between chloroplast and other proteins by normalizing exclusively to the TTP of the measured chloroplast proteome. In other words, these approaches consider the chloroplast protein dataset as independent from the other proteins and perform normalization only on the set of chloroplast proteins. The argument that, e.g., chloroplast proteins in albino mutant make up for only a fragment of the total mass compared to wild-type plastids that contain large amounts of photosynthetic proteins is valid, but the differences are small. We have calculated the total mass of plastid proteins in wild type and compared it to the data obtained from three albino mutants *apg1*, *apg2*, and *apg3* [18]. While 73 % of the total spectra identify plastid proteins in wild type, this is the case for 54 % of the spectra in *apg1*, 48 % in *apg2*, and 45 % in *apg3* (data not shown). Because the calculation of statistical significance is usually more reliable when fewer data points are expected to vary, we normally refrain from this type of normalization. However, we recommend checking the PSM distribution between the organelles for all mutants individually.

3.10 Assigning Proteins to Organelles

In order to assign proteins to organelles, we recommend using existing organellar proteome maps for *Arabidopsis thaliana*. These can be generated conveniently using the SUBA database (<http://suba.plantenergy.uwa.edu.au/>) [19] and the interested researcher is provided with a number of different selection criteria for the localization of a protein to a certain organelle. In case of chloroplasts, we have assembled such a reference table based on a number of different criteria. This table is available from our recent publications [5].

3.11 Statistical Data Handling

With the above-described normalization procedure, a numerical value that resembles protein abundance can be assigned to every identified protein (see Note 12). Proteins with detection gaps in mutant or wild type or those that were not identified in every replicate receive a zero as abundance value. This way undetected proteins are treated the same way as low-abundance proteins, i.e., they are represented by a low nSpC value. Because lack of protein detection in complex mixtures is usually a result of low protein abundance, this procedure reflects the experimental reality. The numerical values assigned to proteins can now be used for data evaluation to determine which proteins were up- or down-regulated in the experiment. This mostly entails the calculation of fold changes between the protein abundance values in different samples. If sufficient biological replicates were included in the experiment, a statistical test such as a Student's *t* test (Welch test) or an ANOVA analysis can be performed as well, followed by applying a maximum p-value cutoff. We usually reject the null hypothesis (no change between mutant and wild type) at p-values below 0.05

obtained from Student's *t* test. Because highly variable data can, by chance, result in high or low fold changes, and data with little variance can result in low *p*-values despite very low changes, each test alone can result in false positives (i.e., accepting biologically irrelevant data points as altered between mutant and wild type). Therefore a combination of both criteria—fold change and *p*-value—is recommended as it will minimize the false discovery rate while maximizing the detection of reliable changes [20].

4 Notes

1. The collection of root and leaf material will allow comparing the effect of a mutation on different *Arabidopsis* organs and can be expanded to include other organs of interest. The preparation of different organs is suitable to provide access to different sets of proteins because protein identification reflects functional organ specialization. While leaf material is dominated by photosynthetic proteins root proteins will cover other functional categories. The difference in protein detection that can be obtained by such a distinction is visualized in Fig. 2 (modified from ref. [11]). Root material was collected from

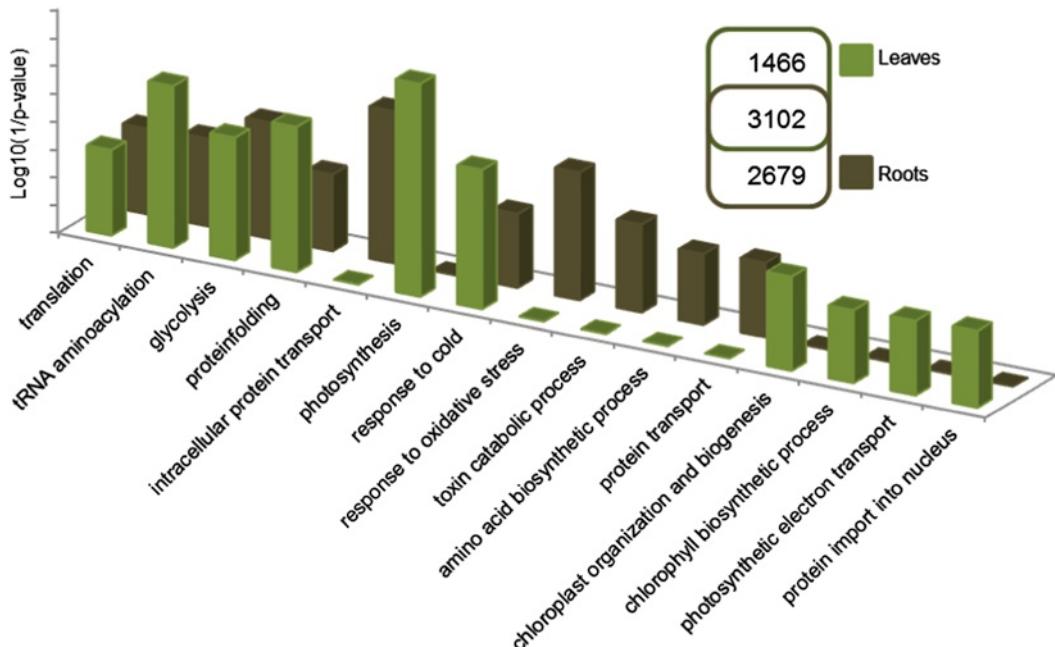


Fig. 2 Functional classification of protein identifications obtained from *Arabidopsis thaliana* leaves and roots (data taken from ref. [14]). Categorization is based on the aspect “biological process” and the numerical values are derived from *p*-values obtained with topGO using the *elim* method [14]. The Venn diagram shows the number of protein identified from roots and leaves and those common to both organs

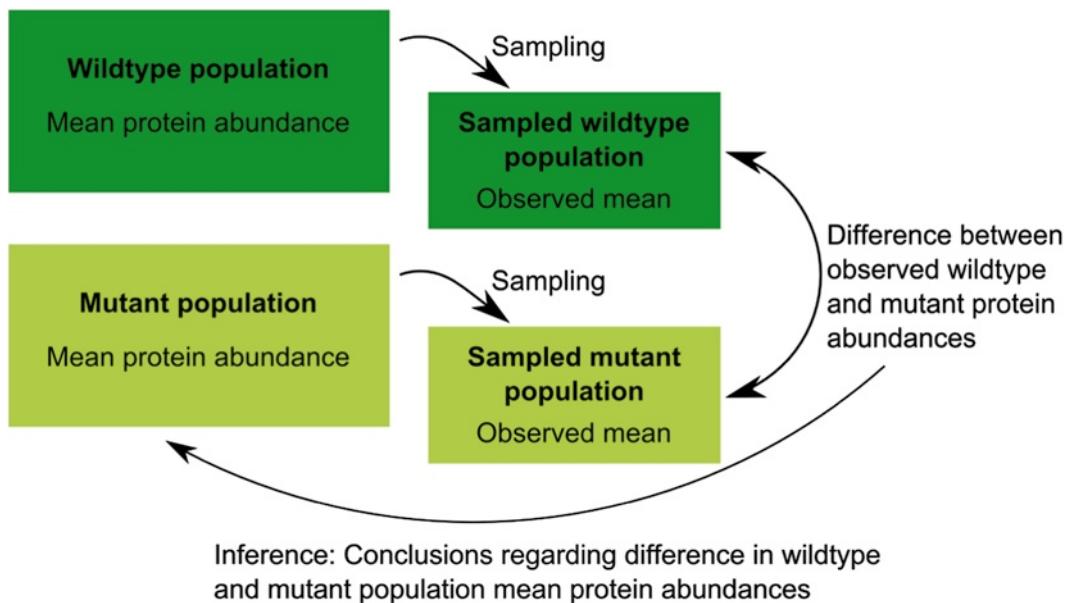


Fig. 3 A schematic representation of the statistical inference procedure adapted from ref. [9]. The sampling should ensure that the samples are representative for the wild-type and mutant populations and that no biases are introduced through different growing conditions (light regime, watering, or position in the growth chamber), seed qualities, or similar. If this is warranted, conclusions regarding differences in population mean protein abundances can be drawn from the observed protein abundance differences in the wild-type and mutant samples

plants grown on Murashige and Skoog medium (1 % sucrose) to avoid soil contaminations in the sample.

2. Only proper experimental design allows statistical inference, meaning that conclusions regarding a whole population can be drawn from the results obtained from a sample [9] (Fig. 3). To ensure that the samples are representative of the population of interest due care should be taken to guard against biases caused by undesirable, and potentially unknown experimental artifacts. In an experimental design with no replicates it cannot be determined whether observed differences are systematic, or due to random chance. Only replication will reveal whether variation is small enough to indicate significant difference. In addition, fold-change differences between wild-type and mutant samples will converge towards the fold-change differences in the population mean protein abundances with more measured independent biological replicates. In situations where the limiting factor is the total number of runs, as it is mostly the case, an experimental design with the maximum number of biological and no technical replicates is most efficient. However, several injections of the same sample (technical replicates) will increase protein detection and make the spectral count quantification more robust [3] and is therefore highly advisable.

3. Pay attention to the compatibility charts of the protein assay used. High concentrations of SDS (>0.1 %) interfere with Bradford, while DTT (>1 mM) interferes with BCA. The optimal choice for protein quantification would be a reducing agent-compatible BCA assay distributed by some suppliers.
4. The SDS-PAGE gel can be prepared in-house with standard laboratory equipment. It is important to allow overnight polymerization of the gels prior to use! Otherwise, unpolymerized acrylamide will be cross-linked to peptides and prevent peptide identification.
5. All solutions should be freshly prepared. Use HPLC-grade H₂O only.
6. To obtain a concentrated protein sample take maximal 10–20 µl extraction buffer per mg fresh weight. Adjust protein concentration to a value of 2 mg/ml.
7. It is advisable to minimize experimental variation and gather the samples from all biological replicates to load them on the same gel. This will minimize differences between the samples that are due to unavoidable technical differences in the protocol application. If not all different biological replicates can be applied on the same gel, alternative statistical tests such as paired t-tests can be considered, which will help in uncovering true differences between wild type and mutants despite the variation introduced by gel-to-gel differences.
8. Avoid keratin contamination! Wear gloves and a lab coat at all times during the experiment! Work in a clean and dust-free environment! Do not lean over gels, tie long hair, and wear a cap or a head cloth; do not wear clothes made from wool!
9. At each step gel pieces should be covered with the appropriate solution. Do not autoclave pipette tips or solutions. Use Eppendorf Safe Lock tubes.
10. Alternatively the tryptic digest can be performed at 37 °C for 4 h.
11. Take care not to confuse *spectral counting* with *peptide counting*. Spectra that should be counted are referred to as peptide spectrum matches (PSMs). PSMs are highly redundant and many hundred spectra often identify a single peptide. It is this redundancy and the repeated sampling of abundant peptides and proteins that make spectral counting so robust for quantification purposes.
12. There is debate on how ambiguous peptides mapping to different protein sequences should be treated. While some colleagues suggest diverting the peptides between the different proteins that they identify, or assign all peptide identifications to the highest scoring gene model, we decided to only accept unambiguous peptide identifications in all our analyses.

While we agree that both procedures have benefits and drawbacks, we decided on our approach, because it was important in our analyses to provide unambiguous protein expression evidence for the identified and quantified proteins.

References

1. Agrawal GK, Bourguignon J, Rolland N et al (2010) Plant organelle proteomics: collaborating for optimal cell function. *Mass Spectrom Rev* 30:772–853
2. Baginsky S (2009) Plant proteomics: concepts, applications, and novel strategies for data interpretation. *Mass Spectrom Rev* 28:93–120
3. Liu H, Sadygov RG, Yates JR 3rd (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 76:4193–4201
4. Lu P, Vogel C, Wang R, Yao X et al (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat Biotechnol* 25:117–124
5. Bischof S, Baerenfaller K, Wildhaber T et al (2011) Plastid proteome assembly without Toc159: photosynthetic protein import and accumulation of N-acetylated plastid precursor proteins. *Plant Cell* 23:3911–3928
6. Gfeller A, Baerenfaller K, Loscos J et al (2011) Jasmonate controls polypeptide patterning in undamaged tissue in wounded *Arabidopsis* leaves. *Plant Physiol* 156:1797–1807
7. Kim J, Rudella A, Ramirez Rodriguez V et al (2009) Subunits of the plastid ClpPR protease complex have differential contributions to embryogenesis, plastid biogenesis, and plant development in *Arabidopsis*. *Plant Cell* 21:1669–1692
8. Reiland S, Finazzi G, Endler A et al (2011) Comparative phosphoproteome profiling reveals a function of the STN8 kinase in fine-tuning of cyclic electron flow (CEF). *Proc Natl Acad Sci USA* 108:12955–12960
9. Oberg AL, Vitek O (2009) Statistical design of quantitative mass spectrometry-based proteomic experiments. *J Proteome Res* 8:2144–2156
10. Sadygov RG, Cociorva D, Yates JR (2004) Large-scale database searching using tandem mass spectra: looking up the answer in the back of the book. *Nat Methods* 1:195–202
11. Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 4:207–214
12. Roos FF, Jacob R, Grossmann J et al (2007) PepSplice: cache-efficient search algorithms for comprehensive identification of tandem mass spectra. *Bioinformatics* 23:3016–3023
13. Eng J, McCormack A, Yates J (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. Baerenfaller K, Grossmann J, Grobeij MA et al (2008) Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. *Science* 320:938–941
15. Baerenfaller K, Hirsch-Hoffmann M, Svozil J et al (2011) pep2pro: a new tool for comprehensive proteome data analysis to reveal information about organ-specific proteomes in *Arabidopsis thaliana*. *Integr Biol (Camb)* 3: 225–237
16. Jones AR, Eisenacher M, Mayer G et al (2012) The mzIdentML data standard for mass spectrometry-based proteomics results. *Mol Cell Proteomics* 11: M111.014381
17. Mallick P, Schirle M, Chen SS et al (2007) Computational prediction of proteotypic peptides for quantitative proteomics. *Nat Biotechnol* 25:125–131
18. Myouga F, Akiyama K, Motohashi R et al (2010) The Chloroplast Function Database: a large-scale collection of *Arabidopsis* Ds/Spm- or T-DNA-tagged homozygous lines for nuclear-encoded chloroplast proteins, and their systematic phenotype analysis. *Plant J* 61: 529–542
19. Heazlewood JL, Verboom RE, Tonti-Filippini J et al (2007) SUBA: the *Arabidopsis* subcellular database. *Nucleic Acids Res* 35:D213–D218
20. Yanofsky CM, Bickel DR (2010) Validation of differential gene expression algorithms: application comparing fold-change estimation to hypothesis testing. *BMC Bioinformatics* 11(63):2

Chapter 21

Analysis of Rice Proteins Using SDS-PAGE Shotgun Proteomics

Karlie A. Neilson, Iniga S. George, Samantha J. Emery, Sridevi Muralidharan, Mehdi Mirzaei, and Paul A. Haynes

Abstract

In this chapter we describe the workflow used in our laboratory to analyze rice leaf samples using label-free shotgun proteomics based on SDS-PAGE fractionation of proteins. Rice proteomics has benefitted substantially from successful execution of shotgun proteomics techniques. We describe steps on how to proceed starting from rice protein extraction, SDS-PAGE, in-gel protein digestion with trypsin, nanoLC-MS/MS, and database searching using the GPM. Data from these experiments can be used for spectral counting, where simultaneous quantitation of several thousand proteins can be obtained.

Key words Shotgun proteomics, Rice, Label-free, Plant proteomics, SDS-PAGE

Abbreviations

| | |
|----------|---|
| 2-DE | Two-dimensional electrophoresis |
| ACN | Acetonitrile |
| BCA | Bicinchoninic acid |
| BSA | Bovine serum albumin |
| DTT | Dithiothreitol |
| FDR | False discovery rate |
| GO | Gene ontology |
| GPM | Global proteome machine |
| IAA | Iodoacetamide |
| MS | Mass spectrometry |
| MS/MS | Tandem mass spectrometry |
| MudPIT | Multidimensional protein identification technology |
| NSAF | Normalized spectral abundance factor |
| RP | Reversed phase |
| SDS-PAGE | Sodium dodecyl sulfate-polyacrylamide gel electrophoresis |
| TCA | Trichloroacetic acid |
| WEGO | Web gene ontology annotation plot |

1 Introduction

To over half the world's population, rice is life. In 2011, rice was harvested from approximately 160 million hectares of land with over 460 million tons of milled rice produced and over 456 million tons consumed, according to the United States Department of Agriculture (as reported by IRRI at <http://ricestat.irri.org>, accessed 27 April 2012). Prices for rice have risen to record highs due to global demand in recent years, with rice constituting 19 % of caloric intake worldwide [1]. In addition to its nutritional and economic importance, rice has become an attractive model system for cereal genomic research because of its relatively small genome (~43 Mb), a high degree of genomic synteny with other cereal crop plants, availability of tens of thousands of searchable insertion lines (e.g., T-DNA, Tos17, Ds), compatibility with genetic transformation, and availability of a sequenced genome that contains approximately 32,000 genes [2]. However, at the time of sequencing, it was estimated that one-third of rice proteins had no known function [3]. One of the major challenges in rice research is to fully annotate the genome with a functional description for each protein to provide a better understanding of rice traits.

Proteomics involves the study of proteins encoded by the genome in a cell, tissue, or organism at a given time or under a particular set of environmental conditions. The study of proteomics is substantially more complex than genomics; the proteome is dynamic and varies with environmental stresses or cellular cues and, furthermore, proteins can be altered by numerous cellular processes such as posttranslational modifications, splicing, degradation, and proteolysis. Traditionally, proteomics has been performed using a two-dimensional electrophoresis (2-DE) gel approach, but with the development of techniques such as multidimensional protein identification technology (MudPIT) [4–7], proteomics has seen rapid growth in shotgun techniques. This shift has been encouraged by changes in mass spectrometry (MS) instrumentation, and driven by the need to analyze many more proteins at a time. Shotgun proteomics is based on the identification of proteins from a complex mixture after separation in at least two dimensions, at either the protein or the peptide level, prior to analysis by tandem MS (MS/MS).

Rice proteomics research has progressed rapidly over the past decade. Increasing proteome coverage and improving genome annotation continue to be a major concern [8, 9], and identification of posttranslational modifications has emerged as an area of intense research [10]. Cataloguing rice proteins is a necessary task; however, the field of proteomics has moved beyond simple protein identification and is now driven to accurately and reliably quantify the differences in protein abundance [11–13]. Plant proteomics research has followed this trend accordingly; quantitation of

differentially expressed proteins between two or more conditions dominates the rice proteomics field [14]. The use of quantitative shotgun proteomics has increased in prevalence and has been used in rice to analyze the effects of abiotic stress in comparative studies [15–18], and for revealing the molecular mechanisms of rice development [19–21].

Despite a recent surge in shotgun techniques observed in the literature, many rice proteomics studies continue to be carried out with the use of 2-DE gels (*see* ref. 14 for a recent review and testament to the technique). This technique is compatible with analysis of plant proteins as the denaturing buffers used for protein separation are well suited to plant protein extraction, proteins can be highly resolved, and protein identification can be performed with simple MS instrumentation and software.

In this chapter we present details of a method, known variously as a “slice and dice” or GeLC-MS/MS experiment, that can be used to simultaneously identify and quantify thousands of proteins at a time without laborious sample handling and visual comparisons of protein spots. This method is executed by separating proteins by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), in-gel digestion of proteins with trypsin, and subsequent extraction of peptides, followed by reversed-phase (RP) separation and direct elution of peptides onto a tandem mass spectrometer [22, 23].

2 Materials

All solutions should be prepared using MilliQ water and highest mass spectrometric grade analytical reagents. Prepared solutions are to be stored at room temperature, unless otherwise indicated. Appropriate waste disposal regulations should be strictly followed. All solutions and samples must be carefully handled to prevent keratin contamination.

1. Extraction buffer was prepared with 8 M urea, 100 mM Tris-HCl, pH 8.5, and 1 % Triton X-100.
2. SDS sample buffer (5× stock solution) was prepared with 6.25 mL 1 M Tris-HCl, pH 6.8, 2 g glycerol, 2.3 g SDS, 617.22 mg dithiothreitol (DTT), 1 mL 5 % (w/v) bromophenol blue, and MilliQ water to 20 mL. Sample buffer was diluted to 2× with MilliQ water.
3. SDS running buffer (10× stock solution) was prepared to 1 L with 144.135 g glycine, 30.285 g Tris, 10 g SDS, and MilliQ water. The stock solution was diluted to 1× with MilliQ water before use.
4. Urea, SDS, Tris, glycine, and Triton X-100 were from Sigma.
5. Bicinchoninic acid (BCA) reagents were from Pierce.

6. DTT, iodoacetamide (IAA), 10 % Tris–HCl precast gels, and Coomassie brilliant Blue G-250 were from Bio-Rad.
7. Trypsin was of Promega sequencing grade.
8. Formic acid was of Fluka 98 % mass spectrometry grade.
9. Acetonitrile was of Merck Lichrosolv liquid chromatography grade.
10. nanoLC-MS/MS was performed using an LTQ-XL linear ion trap mass spectrometer (Thermo, San Jose, CA).
11. Reversed phase C18 columns were packed in-house to approximately 7 cm (100 μ m id) using 100 \AA , 5 μ m Zorbax C18 resin from Agilent Technologies, CA.
12. Spectrum files are converted to .mzXML files prior to database search using Readw.exe, available for free download from <http://sourceforge.net/projects/sashimi/files/ReAdW%20%28Xcalibur%20converter%29/>.
13. XTandem of the Global Proteome Machine software version 2.1.1 is freely available for download from <http://www.thegpm.org/tandem>.
14. The Scrappy program is available as a series of R modules which are available for download from <http://www.proteomecommons.org>.

3 Methods

The workflow detailed below is a simple and effective methodology for a shotgun proteomics experiment that can be routinely used to analyze tissue samples from almost any organism, including recalcitrant tissues (Fig. 1). SDS-PAGE is a fundamental and robust method for protein fractionation and is an amenable technique for sample preparation prior to MS analysis.

This experiment may be performed for both qualitative and quantitative assessments of the proteome of any organism. In-gel digestion can be performed over 2 days once proteins have been separated electrophoretically, and three replicates of 16 fractions can be analyzed over 2½ days of machine analysis time on the mass spectrometer. This means that the workflow can be completed in less than a week for three replicates, and just over a week for a control-versus-treatment experiment of six replicates in total. This approach has been used in several published quantitative proteomic experiments in rice, where the results generated have been analyzed using quantitative tools such as Scrappy [15–17, 24]. The workflow is flexible, in that a variety of protein extraction and MS/MS analysis methods may be applied. The workflow below is written for rice leaves, but can be readily adapted to other tissues including cell cultures, roots, and seeds.

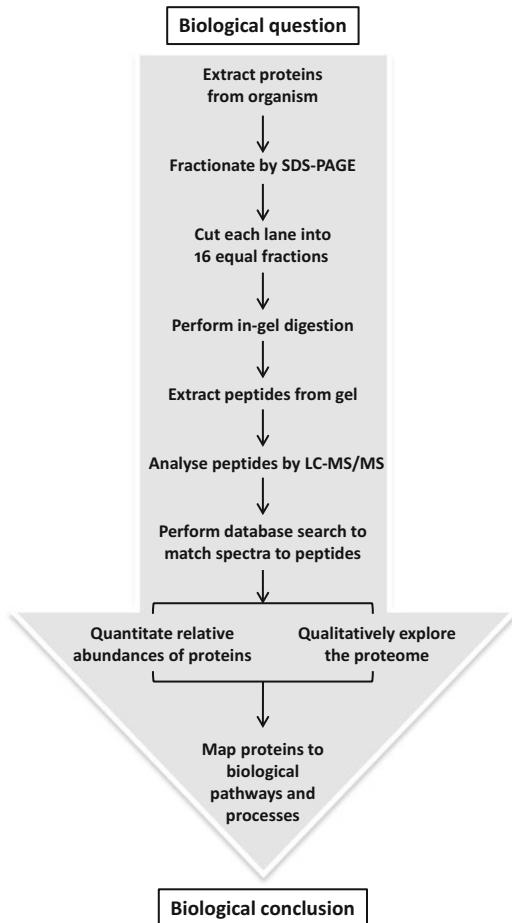


Fig. 1 Schematic diagram of SDS-PAGE shotgun proteomics workflow

3.1 Protein Extraction Protocol

1. The midsection of rice leaves were harvested from rice (*Oryza sativa* cv. Nipponbare) plants grown in a greenhouse.
2. The leaf material was ground into a fine powder under liquid nitrogen with a chilled mortar and pestle (see Note 1). After the nitrogen is evaporated from the powder (it should appear dry), approximately 100 mg aliquots of the fine leaf powder were weighed out into prechilled tubes using a chilled spatula. The leaf material can be stored at -80 °C if not used immediately.
3. The ground leaf material plus 2 mL of extraction buffer (8 M urea in 100 mM Tris-HCl, pH 8.5, with 1 % (v/v) Triton X-100) was subjected to probe sonication on ice for 3 × 20-s pulses.
4. The extract was centrifuged at 17,000×*g* for 10 min at 4 °C and supernatant removed to a clean tube (see Note 2). Repeat if required to remove all debris from the extract.

5. Ice-cold 100 % (v/v) trichloroacetic acid (TCA) was added to the supernatant in a fresh tube to a final concentration of 25 % (v/v). The solution was vortexed briefly, and proteins were precipitated overnight at -20 °C.
6. The protein suspension was centrifuged at $17,000 \times g$ for 10 min at 4 °C to pellet the precipitate. The supernatant was removed and the protein pellet was washed twice with 850 μ L ice-cold acetone.
7. The protein pellet was air dried for 5 min until the acetone evaporated.
8. The protein was solubilized in 100 μ L of 2× SDS sample buffer, without DTT or bromophenol blue, and a BCA assay was performed to determine the concentration of the protein sample using bovine serum albumin (BSA) as a standard (*see Note 3*).
9. 1 M stock solution of DTT was added to 100 μ g of protein extract to give a final concentration of 40 mM DTT. A trace of bromophenol blue was added, and the mixture was heated at 75 °C for 5 min before fractionation by SDS-PAGE using a 10 % precast gel and 1× SDS running buffer. Proteins were electrophoresed at 70 V for 15 min, followed by 160 V for 50 min.
10. The gel was lightly stained with Coomassie blue to provide a visual aid for subsequent handling, although this is not strictly necessary.

3.2 In-Gel Digestion Procedure

This procedure was originally adapted from Shevchenko et al. [25].

1. The stained gel was placed on a clean glass plate, and the lane containing the protein extract was excised from the gel using a clean scalpel (*see Note 4*). The lane was then cut into 16 equal-size fractions from top to bottom of the gel, and each of the 16 gel bands was cut into smaller fractions (*see Note 5*) (Fig. 2).
2. The 16 groups of gel pieces were transferred to 0.65 mL polypropylene tubes or 96-well plate and excess water removed. The gel pieces were washed briefly with 100 mM NH_4HCO_3 to ensure correct pH of gel pieces.
3. The gel pieces were destained with 200 μ L 50 % (v/v) acetonitrile (ACN) and 50 % (v/v) 50 mM NH_4HCO_3 by vortexing, and then incubating for 10 min at room temperature. The liquid was removed, and then this step was repeated. The gel pieces should be clear and free of stain at this stage. If the gel pieces are still stained, then repeat the above washing as necessary. However, it is allowable to proceed with traces of stain, as it will be removed during further washing steps (*see Note 6*).
4. The gel pieces were washed for 5 min with 200 μ L 50 % (v/v) ACN and 50 % (v/v) 50 mM NH_4HCO_3 , and then 5 min with

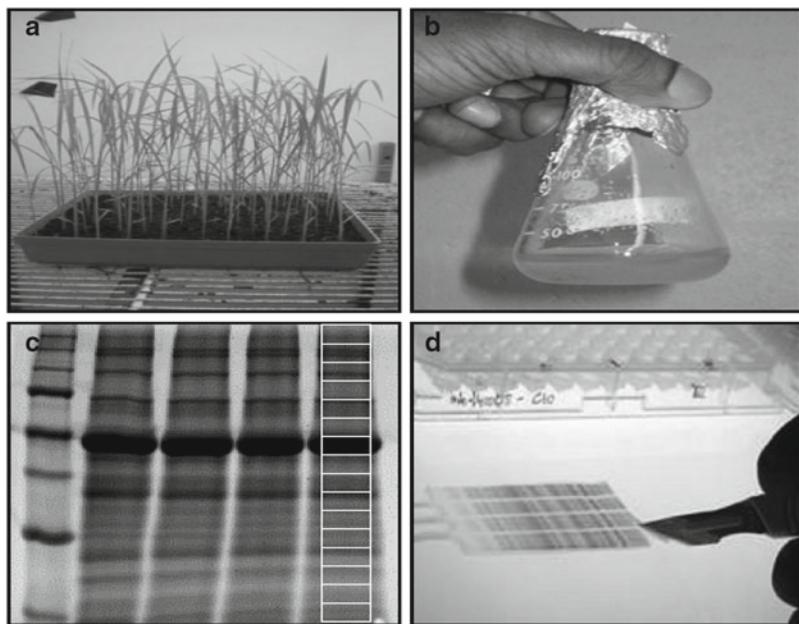


Fig. 2 Proteins are extracted from (a) rice leaves, stems, or roots, or (b) rice cell cultures and separated by SDS-PAGE. (c) After visualizing with a protein stain, (d) lanes are sliced into 16 equal fractions and transferred to a 96-well plate (or microcentrifuge tubes) for in-gel digestion

100 % (v/v) ACN to dehydrate gel pieces. The samples were vortexed during these incubations.

5. The ACN was removed and the gel pieces allowed to either air-dry on bench or in fume hood for 10 min, or evaporated briefly in a vacuum centrifuge without heat. The gel pieces should be noticeably shrunken and white.
6. The gel pieces were rehydrated with 50 μ L 10 mM DTT in 50 mM NH_4HCO_3 and vortexed to mix. The tubes were briefly centrifuged and proteins were reduced for 60 min at 37 °C.
7. The gel pieces were cooled to room temperature, DTT solution was removed, and 50 μ L of 55 mM IAA in 50 mM NH_4HCO_3 was added and vortexed to mix. The tubes were briefly centrifuged and proteins were alkylated for 45 min in the dark at room temperature.
8. The IAA solution was removed and the gel pieces were washed with 100 μ L of 100 mM NH_4HCO_3 for 5 min with vortexing, and then washed twice with 50 % (v/v) ACN and 50 % (v/v) 50 mM NH_4HCO_3 for 5 min with vortexing.
9. The gel pieces were dehydrated with 100 μ L ACN as in **steps 4** and **5**. Again, the gel pieces should be noticeably shrunken and probably white.

10. Trypsin solution was prepared in the buffer provided (Promega) to 1 μ g/ μ L, and then 5 μ L of the solution was added to 395 μ L 50 mM NH₄HCO₃. This should be enough for 20 digestions at a concentration of 12.5 ng/ μ L. An adequate volume of trypsin solution was added to cover the gel bands. This volume will vary but is usually around 20 μ L and up to 50 μ L for large gel bands.
11. The gel pieces were rehydrated at 4 °C for at least 30 min in trypsin solution. Incubation must be kept cool so as to allow as much active trypsin to be absorbed into the gel as possible before auto-digestion occurs (*see Note 7*).
12. The tubes were centrifuged briefly, and an additional 25 μ L 50 mM NH₄HCO₃ was added to cover gel pieces.
13. Proteins were digested overnight at 37 °C, or for a minimum of 4 h.
14. The digest solution supernatant (if any) was transferred into a clean 0.65 mL tube.
15. Peptides were extracted from the gel pieces with 30 μ L (or enough to cover) of 50 % ACN (v/v) and 2 % (v/v) formic acid and incubated for 20 min with vortexing at room temperature. The tubes were centrifuged briefly, and the supernatant was removed and combined with initial digest solution supernatant. For large pieces of gel, use more liquid where required.
16. **Step 15** was repeated to give a combined peptide extract volume of around 60–100 μ L (*see Note 8*). For large gel slices, use slightly larger volumes, repeat extraction a third time, and combine all extracts.
17. The extracted digests were vortexed, centrifuged briefly, and then concentrated in a vacuum centrifuge to approximately 5 μ L each. Do not dry completely if possible. The peptides were reconstituted to a volume of 10 μ L with 1 % (v/v) formic acid. If peptides are reduced to dryness, reconstitute peptides in 10 μ L of 1 % (v/v) formic acid (*see Note 9*).
18. The 10 μ L extracts were centrifuged at 17,000 $\times g$ for at least 15 min to pellet any microparticulates. The supernatant was very carefully transferred to a fresh 0.65 mL polypropylene tube, or into a 96-well plate. The sample is now ready for analysis by LC-MS/MS.

3.3 nanoLC-MS/MS of Gel Fractions

1. Each of the 16 fractions prepared was analyzed sequentially using a nanoLC-MS/MS system, employing an LTQ-XL ion-trap mass spectrometer (Thermo, San Jose, CA).
2. Reversed phase columns were packed in-house to approximately 7 cm (100 μ m i.d.) using 100 Å, 5 mM Zorbax C18 resin (Agilent Technologies, CA, USA) in a fused silica capillary with an integrated electrospray tip (*see Note 10*).

3. A 1.8 kV electrospray voltage was applied via a gold-electrode liquid junction upstream of the C18 column.
4. Samples were injected onto the C18 column using a Surveyor autosampler (Thermo, San Jose, CA).
5. Each sample was loaded onto the C18 column followed by an initial wash step with buffer A (5 % (v/v) ACN, 0.1 % (v/v) formic acid) for 10 min at 1 μ L/min.
6. Peptides were subsequently eluted from the C18 column with 0–50 % buffer B (95 % (v/v) ACN, 0.1 % (v/v) formic acid) over a 30-min linear gradient at 500 nL/min followed by 50–95 % buffer B over 5 min at 500 nL/min and 5 min was with 95 %B prior to column re-equilibration (*see Note 11*).
7. The column eluate was directed into the nanospray ionization source of the mass spectrometer.
8. Spectra were scanned over the range 400–1,500 amu. Automated peak recognition, dynamic exclusion (90 s), and tandem MS of the top six most intense precursor ions at 40 % normalization collision energy were performed using Xcalibur software (Thermo).

3.4 Protein and Peptide Identification

1. The set of 16 data files from one experiment were acquired in the proprietary .RAW format. These were first converted to .mzxml format using the freeware Readw.exe program (*see Note 12*).
2. The set of 16 .mzxml data files from a given sample were then placed into one directory and peptide-to-spectrum matching was performed using the XTandem algorithm. We use the Global Proteome Machine (GPM) software [26], which is freely available and runs the XTandem Tornado version. Searching the set of .mzxml files stored in a directory enables the user to choose for a single combined summary output file to be created in addition to all 16 individual result files (*see Note 13*).
3. The combined protein and peptide identification output file for all 16 gel slices was then exported to an Excel spreadsheet. This contains six columns of data, with the headers Identifier, $\log(I)$, rI, $\log(e)$, pI, and Mr (kDa). Two other headers are included, Description and Annotated Domains, but will only have entries automatically provided if this information is available within the searched database. The Excel file is then exported to comma-separated value format, which is then compatible with input into the Scrappy software described elsewhere [27].

An example of the output produced by this approach is provided in Table 1. Included in this abridged table are the first 20 proteins identified from a rice leaf extract, sorted by the confidence

Table 1

Table of results from a GPM search showing the top 20 proteins identified in a previous study, sorted by $\log(e)$ value

| Identifier | $\log(I)$ | rI | $\log(e)$ | pI | Mr (kDa) | Description |
|--------------|-----------|-------|-----------|------|----------|--|
| gi 11466795 | 9.36 | 3,512 | -1,288 | 6.22 | 52.8 | RuBisCO large subunit |
| gi 115468792 | 9.17 | 3,085 | -1,057.5 | 6.95 | 48 | RuBisCO large subunit |
| gi 115440423 | 8.56 | 1,324 | -838.3 | 8.14 | 172.7 | ATPase delta/epsilon subunit |
| gi 115481654 | 8.65 | 914 | -802.8 | 5.95 | 84 | ATP synthase beta chain |
| gi 11466794 | 8.65 | 932 | -795.2 | 5.47 | 54 | ATP synthase CF1 beta subunit |
| gi 115439533 | 7.82 | 224 | -671.8 | 6.51 | 111.4 | Glycine dehydrogenase |
| gi 115487804 | 9.13 | 1,678 | -634.5 | 9.01 | 59 | RuBisCO large subunit |
| gi 115473843 | 7.68 | 182 | -590.1 | 6.04 | 125.6 | Glutamate synthase |
| gi 115478691 | 7.75 | 727 | -560.6 | 5.14 | 28 | Chlorophyll A-B-binding protein |
| gi 115438250 | 7.72 | 712 | -515.7 | 5.29 | 27.5 | Chlorophyll A-B-binding protein |
| gi 11466784 | 8.45 | 844 | -508.3 | 5.95 | 55.6 | ATP synthase CF1 alpha subunit |
| gi 115439621 | 7.59 | 643 | -505.4 | 5.29 | 27.9 | Chlorophyll A-B-binding protein |
| gi 115468926 | 7.73 | 191 | -501.9 | 5.98 | 96.9 | Glycine cleavage system P-protein |
| gi 115466224 | 7.98 | 377 | -495.3 | 5.44 | 73.4 | Thiamine pyrophosphate family, transketolase |
| gi 115486823 | 8.43 | 562 | -477.3 | 5.86 | 47.9 | Ribulose-bisphosphate carboxylase activase |
| gi 115487910 | 7.24 | 186 | -450.5 | 6.62 | 102 | Clp protease ATP-binding subunit |
| gi 115436780 | 8.1 | 300 | -426.3 | 6.08 | 34.8 | Manganese-stabilizing protein |
| gi 115488160 | 7.49 | 153 | -423.7 | 5.12 | 61.1 | GroEL-like type I chaperonin |
| gi 115438793 | 8.41 | 564 | -371.3 | 8.78 | 30 | Carbonic anhydrase |
| gi 11466786 | 7.96 | 339 | -363.4 | 6.63 | 82.5 | Photosystem I P700 chlorophyll a apoprotein A2 |

The Identifier indicates the gene identifier number associated with the protein, and will depend on the origin of the FASTA database, $\log(I)$ represents the log sum intensity of the spectra acquired, rI is the number of redundant peptides identified for the protein (used in spectral counting quantitation), $\log(e)$ is the log expectation value that the protein match was random, pI is the isoelectric point of the protein, and Mr (kDa) indicates the mass of the protein

score, $\log(e)$. From this single experiment, which was first published as part of a larger data set in Gammulla et al. [16], 760 proteins were identified to be common to three replicates at 0.79 % false discovery rate (FDR) with an average of 37,469 peptides at 0.21 % FDR.

3.5 Beyond Protein Identification: Quantitation and Extraction of Biological Relevance

Quantitative proteomics can be divided into two categories: label-free methods and the use of isotope labelling. Label-free proteomics has several advantages over labelling techniques: all organisms and sample types may be analyzed, the number of samples that can be compared is not limited, and data analysis is comparatively straightforward and does not require the use of specific software. However, the volume of data accumulated in a global, label-free proteomics study, such as in a GeLC-MS/MS experiment, can be overwhelming. An important goal in functional proteomics is to globally profile changes in protein abundances in biological systems and also provide a snapshot of the protein expression state in response to biological perturbations. The eventual outcome of any quantitative proteomics study is to draw a biological conclusion from the large volume of data acquired. This can be systematically achieved by measuring the protein abundance differences between proteins from two or more conditions, applying statistical tests for significance, and visualizing results in a biological context (*see* ref. 28 for a recent review).

We analyze our data by spectral counting, specifically by calculating normalized spectral abundance factors (NSAFs) [29] for each protein in a data set. This calculation takes into consideration that the number of spectra identified for a protein will be dependent on the length of that protein. This calculation is incorporated in a series of freely available R modules assembled in the form of the Scrappy program [27]. Spectral counting using NSAFs has been demonstrated to be an accurate method for calculating relative abundances of proteins between two or more samples [30–32]. In a quantitative study comparing spectral counting using NSAFs with isobaric labelled peptides, we found that a much greater volume of differentially expressed proteins were identified with the label-free method, although the two methods yielded similar biological conclusions [33].

Extracting biological relevance from large-scale proteomics data sets is a challenge in both label-free and labelled proteomic studies. Numerous tools exist to functionally categorize protein identification data and map proteins to biological processes. In our experience, we have found it useful to begin functional analysis by accumulating Gene Ontology (GO) annotations [34] for proteins identified in the data set, followed by functional categorization using Web Gene Ontology Annotation Plot (WEGO) [35]. WEGO maps GO annotations to either functional pathways or processes and calculates a p-value for category enrichment. Quantitative analysis of functional categories can be performed using PloGO, a freely available, open-source tool [24]; rather than simply summing the number of proteins in a category, the sum of NSAF values can be used to estimate the relative abundance of proteins in a pathway or a process. This information can be used to highlight areas of the proteome that are most responsive to a stress or a specific treatment and hence, that particular area may be studied in further detail.

4 Conclusion

Label-free shotgun proteomics based on SDS-PAGE fractionation of proteins from rice is a simple, robust, and versatile technique. The method can be used with proteins extracted from any tissue type as varying extraction methods can be applied and data can be collected for qualitative or quantitative purposes. In our experience, the technique is highly reproducible for the identification and quantitation of large protein data sets using spectral counting, although other methods of label-free quantitation can be applied. The range of achievable outcomes of label-free quantitative proteomics in the field of systems biology is bound to expand with improvements in instrumentation and computer software.

5 Notes

1. During protein extraction, keep falcon tubes, scissors, and spatula on dry ice. Wipe mortar, pestle, and spatula clean with a lint-free tissue between samples.
2. Chill the centrifuge and rotor to 4 °C before starting protein extraction.
3. DTT and bromophenol blue are not compatible with a BCA assay and so were added after measuring protein concentration. SDS needs to be diluted to <5 % to be compatible with BCA reagents.
4. Spray enough water on the plate to stop the gel drying out. Wipe down all surfaces before starting with a lint-free cloth in methanol, work in a hood if possible, and wear gloves and sleeve protectors. Cover hair and beards, and take all available steps to avoid keratin contamination.
5. It is not necessary to contain noticeable bands within a single fraction; it is more important to make the gel fractions equal sized and reproducible across replicates. This may be done by dividing the gel into two fractions, and then each two fractions into four fractions until 16 fractions are obtained. This may be achieved by simply cutting each fraction down the middle for a total of four rounds.
6. Treat gels stained with Sypro ruby, Deep Purple, or other fluorescent stains the same as a Coomassie-stained gel.
7. Allow trypsin to incubate with the gel pieces on ice for 20–30 min.
8. For the most complete extraction of peptides, peptides may be extracted in sequential steps with the first in 1 % (v/v) formic acid, the second in 50 % (v/v) ACN, and the third in 90 % (v/v) ACN (each diluted in water).

9. Dried or reconstituted peptides can be stored at -80 °C for as long as required.
10. nanoLC columns can also be purchased from suppliers such as Michrom or New Objective.
11. It is not required to completely eliminate carryover between injections as all the samples in a given set will be combined for subsequent analysis.
12. Our experimental design consists of 16 SDS-PAGE gel slices from each of the three biological replicates of a given sample. This can be completed in 2½ days of mass spectrometric analysis time.
13. We use the “mudpit combine” option in GPM searches to produce a unified output file from the individual search result files.

Acknowledgements

KAN, ISG, and SM acknowledge support in the form of iMQRES awards. SJE acknowledges support in the form of an APA scholarship. PAH acknowledges funding support from the Australian Research Council and wishes to thank Gayani Gammulla for providing images.

References

1. Weinberger KM, Easdown WJ, Yang R et al (2009) Food crisis in the Asia-Pacific region. *Asia Pac J Clin Nutr* 18:507–515
2. Goff SA, Ricke D, Lan TH et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296:92–100
3. Rabbani MA, Maruyama K, Abe H et al (2003) Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. *Plant Physiol* 133:1755–1767
4. Koller A, Washburn MP, Lange B et al (2002) Proteomic survey of metabolic pathways in rice. *Proc Natl Acad Sci U S A* 99: 11969–11974
5. Link AJ, Eng J, Schieltz DM et al (1999) Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* 17:676–682
6. Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19:242–247
7. Wolters DA, Washburn MP, Yates JR 3rd (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* 73:5683–5690
8. Helmy M, Tomita M, Ishihama Y (2011) OryzaPG-DB: rice proteome database based on shotgun proteogenomics. *BMC Plant Biol* 11:63
9. Liu L, Bai L, Luo C et al (2011) Systematic annotation and bioinformatics analyses of large-scale *Oryza sativa* proteome. *Curr Protein Pept Sci* 12:621–630
10. Nakagami H, Sugiyama N, Mochida K et al (2010) Large-scale comparative phosphoproteomics identifies conserved phosphorylation sites in plants. *Plant Physiol* 153:1161–1174
11. Ong SE, Foster LJ, Mann M (2003) Mass spectrometric-based approaches in quantitative proteomics. *Methods* 29:124–130
12. Ong SE, Mann M (2005) Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* 1:252–262
13. Steen H, Pandey A (2002) Proteomics goes quantitative: measuring protein abundance. *Trends Biotechnol* 20:361–364
14. Agrawal GK, Rakwal R (2011) Rice proteomics: a move toward expanded proteome

- coverage to comparative and functional proteomics uncovers the mysteries of rice and plant biology. *Proteomics* 11:1630–1649
15. Gammulla CG, Pascovici D, Atwell BJ et al (2010) Differential metabolic response of cultured rice (*Oryza sativa*) cells exposed to high- and low-temperature stress. *Proteomics* 10:3001–3019
 16. Gammulla CG, Pascovici D, Atwell BJ et al (2011) Differential proteomic response of rice (*Oryza sativa*) leaves exposed to high- and low-temperature stress. *Proteomics* 11:2839–2850
 17. Mirzaei M, Pascovici D, Atwell BJ et al (2012) Differential regulation of aquaporins, small GTPases and V-ATPases proteins in rice leaves subjected to drought stress and recovery. *Proteomics* 12:864–877
 18. Mirzaei M, Soltani N, Sarhadi E et al (2012) Shotgun proteomic analysis of long-distance drought signaling in rice roots. *J Proteome Res* 11:348–358
 19. Hamamoto K, Aki T, Shigyo M et al (2012) Proteomic characterization of the greening process in rice seedlings using the MS spectral intensity-based label free method. *J Proteome Res* 11:331–347
 20. He D, Han C, Yao J et al (2011) Constructing the metabolic and regulatory pathways in germinating rice seeds through proteomic approach. *Proteomics* 11:2693–2713
 21. Lee J, Jiang W, Qiao Y et al (2011) Shotgun proteomic analysis for detecting differentially expressed proteins in the reduced culm number rice. *Proteomics* 11:455–468
 22. Schirle M, Heurtier MA, Kuster B (2003) Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry. *Mol Cell Proteomics* 2:1297–1305
 23. Simpson RJ, Connolly LM, Eddes JS et al (2000) Proteomic analysis of the human colon carcinoma cell line (LIM 1215): development of a membrane protein database. *Electrophoresis* 21:1707–1732
 24. Pascovici D, Keighley T, Mirzaei M et al (2012) PloGO: plotting gene ontology annotation and abundance in multi-condition proteomics experiments. *Proteomics* 12:406–410
 25. Shevchenko A, Wilm M, Vorm O et al (1996) A strategy for identifying gel-separated proteins in sequence databases by MS alone. *Biochem Soc Trans* 24:893–896
 26. Craig R, Beavis RC (2004) TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 20:1466–1467
 27. Neilson KA, Keighley T, Pascovici D et al (Submitted May 2012) Label-free quantitative shotgun proteomics using normalized spectral abundance factors. *Meth Mol Biol*
 28. Neilson KA, Ali NA, Muralidharan S et al (2011) Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* 11:535–553
 29. Zyballov B, Mosley AL, Sardiu ME et al (2006) Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *J Proteome Res* 5:2339–2347
 30. Colaert N, Vandekerckhove J, Gevaert K et al (2011) A comparison of MS2-based label-free quantitative proteomic techniques with regards to accuracy and precision. *Proteomics* 11:1110–1113
 31. Gokce E, Shuford CM, Franck WL et al (2011) Evaluation of normalization methods on GeLC-MS/MS label-free spectral counting data to correct for variation during proteomic workflows. *J Am Soc Mass Spectrom* 22:2199–2208
 32. Pavelka N, Fournier ML, Swanson SK et al (2008) Statistical similarities between transcriptomics and quantitative shotgun proteomics data. *Mol Cell Proteomics* 7:631–644
 33. Neilson KA, Mariani M, Haynes PA (2011) Quantitative proteomic analysis of cold-responsive proteins in rice. *Proteomics* 11: 1696–1706
 34. Ashburner M, Ball CA, Blake JA et al (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25:25–29
 35. Ye J, Fang L, Zheng H et al (2006) WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res* 34:W293–W297

Chapter 22

***Medicago truncatula* Proteomics for Systems Biology: Novel Rapid Shotgun LC-MS Approach for Relative Quantification Based on Full-Scan Selective Peptide Extraction (Selpex)**

**M. Angeles Castillejo, Christiana Staudinger, Volker Egelhofer,
and Stefanie Wienkoop**

Abstract

Medicago truncatula has become the focus of systems biology research for improved legume crop breeding. In plant systems biology, several comparative studies have been carried out using liquid chromatography shotgun mass spectrometry (LC-MS/MS) and database-dependent protein identification analyses in combination with the spectral count for relative quantification. In order to receive optimal protein identification rates and spectral count quantification, data-dependent tandem mass spectrometry with LC separation of more than 1 h is required. Thus LC-MS/MS analyses time is the bottleneck for high-throughput research of experiments with high sample number.

We describe a novel method, called full-scan (FS) selective peptide extraction, that allows for comparative quantification of target peptides combined with a significant reduction in LC-MS analysis time. In future, it will be a useful tool to detect ¹⁵N-labeled selected peptide patterns for the targeted analysis of protein turnover and synthesis. We provide a first reference library of selected target peptides generated for *M. truncatula* leaf tissue. These peptides are also suitable candidates for selective reaction monitoring approaches.

Key words *Medicago*, Selpex, Target peptides, Peptide quantification

Abbreviations

| | |
|----------|---|
| ACN | Acetonitrile |
| DDP | Data-dependent |
| FA | Formic acid |
| FS | Full-scan |
| LC-MS/MS | Liquid chromatography mass spectrometry |
| MAPA | Mass accuracy precursor alignment |
| SC | Spectral count |
| Selpex | Selective peptide extraction |
| SRM | Selective reaction monitoring |

1 Introduction

The model legume *Medicago truncatula* has recently been subject to proteomic profiling studies such as of root plastids [1], mitochondria [2], and comparative proteomics of root nodules exposed to drought stress [3, 4]. For the latter systems biology approach, gel-free high-throughput shotgun mass spectrometry was used to analyze 35 samples.

The analysis was based on database-dependent protein identification and the corresponding spectral count (SC) for the detection of relative changes on protein levels. To maximize peptide identification rates, data-dependent (DDP) LC-MS/MS analysis takes several hours [5]. To enhance the explanatory power of the quantitative information extracted, replicate MS measurements are also required. Thus, MS runtime is limiting to experiments on large datasets. The DDP triggers a series of tandem MS (MS/MS) for peptide identification. It is typically a “top 5-” or “top10”-type MS experiment in which the 5–10 most intense ions of a survey scan (also called full scan) were selected for MS/MS. This common type of DDP approach usually takes less than 1 s per cycle time, however, slowing down data acquisition on full-scan level which is in the range of a few milliseconds. Thus a data-independent analysis (DIA) using only the full scan (FS) results in many more scans per ion (peptide peak) allowing for a strongly reduced gradient time. In systems biology the sample number can easily increase up to one hundred and more including many biological replicates [6]. Runtime on a mass spectrometer may therefore exceed weeks which needs reduction in future.

We demonstrate a method that allows for a significant reduction in LC-MS analysis time without loss in the number of protein identification for relative quantification. This approach is based on prior peptide identification and retention time extraction and the preparation of selected target peptide lists in combination with full-scan analysis without further MS/MS data acquisition. Instead of the common spectral counting for label-free relative quantification (see also Chapter 13), the ion intensity count (the sum of all ion intensities of the target peptide) is being used. The target peptides have been picked from *M. truncatula* leaf protein extracts.

2 Materials

2.1 Plant Material

Leaves of 7-week-old *M. truncatula* (Jemalong) plants. Growth conditions as described previously [3] with some changes. Plants were fertilized with 2.5 mM NH₄NO₃ instead of rhizobia inoculation.

2.2 Buffers and Solutions

1. Extraction buffer: 50 mM HEPES, pH 7.8, 1 mM EDTA, 1 mM KCl, and 2 mM MgCl₂.
2. Solubilization buffer: 50 mM HEPES, pH 7.8, 8 M urea.
3. Trypsin buffer: 10 % ACN, 50 mM ammonium bicarbonate, 2 mM CaCl₂.
4. Solvent A: 0.1 % FA.
5. Solvent B: 0.1 % FA and 99.9 % ACN.

2.3 Equipment and Software

1. Thermo Scientific LTQ Orbitrap XL mass spectrometer and Xcalibur software.
2. ProtMAX (<http://www.univie.ac.at/mosys/software.html>). A detailed protocol for ProtMAX can be found elsewhere [7].
3. RAW to mzXML file converter (MassMatrix MS Data File Conversion, v3.9).

3 Methods

3.1 Protein Extraction of a *Medicago* Leaf Master Sample

Medicago leaves (100 mg fresh weight) were homogenized at 4 °C in the extraction buffer. After centrifugation (10,000 $\times g$, 20 min, 4 °C), proteins in the supernatant were precipitated overnight using ice-cold acetone and 0.5 % β -mercaptoethanol. After centrifugation for 15 min at 4,000 $\times g$ at 4 °C, the supernatant was removed and the pellet air dried for 10 min.

3.2 Sample Preparation

The precipitated proteins were resuspended in solubilization buffer. For digestion, 50 μ g of protein were incubated with 0.5 μ g endoproteinase LysC (Roche Diagnostics Corp.) during 5 h at 30 °C. Samples were then diluted to 2 M urea using trypsin buffer and 10 μ l Porosozyme® immobilized trypsin (Applied Biosystems, CA, USA) were added for further digestion overnight at 37 °C. The digests were then desalted on C-18 Selpex cartridges (Varian, CA, USA), eluted with methanol and dried in a vacuum concentrator.

3.3 Initial Identification and Selective Peptide Extraction

For protein identification, a common DDP analysis was performed (top5 MS/MS, $n=3$) using a 120-min gradient from 2 % solvent A and 98 % solvent B to 60 % B. 500 ng of protein digest were loaded to a C18 reversed monolithic column (Chromolith® CapRod 150 \times 0.1 mm i.d., Merk KGsA, Germany). After MS analysis the Orbitrap raw files were searched against a *M. truncatula* database (DFCI release 10) using the SEQUEST algorithm of the Thermo Proteome Discoverer 1.3 software. In order to identify with high statistical confidence, only precursor ions within a mass tolerance limit of 3 ppm were allowed. Additionally, the occurrence of at least two distinct peptides per protein is required. The output file was a multi-consensus Excel sheet with the identified proteins and their

corresponding peptide sequences, the number of spectral counts, as well as the retention time and the mass-to-charge ratio (m/z) (see Note 1). The Qual Browser (Xcalibur v2.1, Thermo) software was used for peptide peak extraction, visualization, and evaluation. Doubly charged peptides with robust and reproducible high-quality peak shape, reproducible found in all replicates, were selected for the target peptide list for further selective peptide extraction (Selpex) analysis (Table 1) (see Note 2). All peptide sequences have been tested and marked for protein specificity (proteotypic), as far as the present *Medicago* protein annotation information allows for (see Note 3). Retention times for these peptides were adapted to a very short 30-min LC-MS gradient analysis.

3.4 Full-Scan Selpex for Rapid Relative Quantification Based on Ion Intensity Count Using ProtMAX

For a Selpex check, DDP LC-MS/MS analyses were compared with the FS-based LC-MS detection (Fig. 1). A file (tab-delimited text) with the target list of the selected m/z peptides and corresponding retention times was generated (Table 1) and imported into the ProtMAX tool (<http://www.univie.ac.at/mosys/software.html>). The converted RAW chromatogram files in mzXML format (MassMatrix MS Data File Conversion, v3.9) were imported into ProtMAX. The selected preference settings used to perform the quantitation were as follows: (1) target list was chosen as method; (2) intensity for quantification; (3) cut to two decimals; (4) +2 charge state and retention time (RT) [min] environment was set to 3. A data matrix was retrieved from the software including the target m/z values with corresponding ion intensity counts (cumulative ion intensity), scan number, and retention time for each sample analyzed.

4 Anticipated Results

The Selpex approach is useful in order to minimize MS analysis time enabling high-throughput measurements of huge sample sets that often occur in systems biology experiments. Selpex subsequently allows for the database-independent protein (peptide) identification and relative quantification. Due to peptide-level resolution, Selpex also enables the recognition of modification-induced changes as previously shown for the MAPA approach [8].

Here, we demonstrate the possibility to quantify around 40 identified proteins using a selected peptide list even from an extremely short 30-min LC gradient using FS MS only (Fig. 1) (see Note 4). Target 738.40 m/z shows most abundant ion intensity count (peak intensity). The lower ratio of that target in the 30-min gradient compared to the 120-min gradient might be due to ion suppression effects (see Note 5). Thus, the dynamic range of the 30-min gradient seems to be the limitation for high-abundance peptides (see Note 6). As a consequence, quantitative changes of

Table 1
Peptide target list for the Sepex approach

| <i>m/z</i> (Da) | TC | UniRef100 unreviewed | Annotation | Sequence | RT (min) | | |
|-----------------|----------|--------------------------|------------------|-------------------------|-------------|-----|----|
| | | | | | Proteotypic | 120 | 30 |
| 1,085.49 | TC112983 | A2Q5A2_MEDTR | RLS | GGLDDFTKDDENVNSQPFMR | No | 29 | 11 |
| 871.37 | TC112983 | A2Q5A2_MEDTR | RLS | GHYNATAGTCEDMMK | Yes | 18 | 9 |
| 745.36 | TC112983 | A2Q5A2_MEDTR | RLS | WSPELAAACEVWK | No | 49 | 16 |
| 488.28 | TC113878 | A2Q5A3_MEDTR | ATPasc | IGLFGGGAGVGK | No | 12 | 8 |
| 801.40 | TC113878 | A2Q5A3_MEDTR | ATPasc | VALVYQQMNEPPGAR | No | 16 | 9 |
| 983.01 | TC171155 | G7KMR3_MEDTR | RSS | FETLSYLPPLTEDQLAK | Yes | 50 | 16 |
| 1,047.06 | TC171155 | G7KMR3_MEDTR | RSS | KFETLSYLPPLTEDQLAK | Yes | 42 | 14 |
| 738.40 | TC171155 | G7KMR3_MEDTR | RSS | LPLFGATTDSSQVLK | No | 32 | 12 |
| 718.89 | TC167788 | G8A0J3_MEDTR | RSS | FETLSYLPPLTR | Yes | 40 | 14 |
| 467.26 | TC167788 | G8A0J3_MEDTR | RSS | IIGFDNNVR | No | 13 | 8 |
| 730.41 | TC151243 | G7J252_MEDTR | RSS | LPLFGATIDASQVLK | Yes | 39 | 14 |
| 782.94 | TC151243 | G7J252_MEDTR | RSS | KFETLSYLPPLTR | Yes | 31 | 12 |
| 788.88 | TC120394 | G7K9H5_MEDTR | OEE1 | GASTGYDNAVALPAGGR | No | 12 | 8 |
| 618.83 | TC120394 | G7K9H5_MEDTR | OEE1 | RLTFDEIQSK | No | 8 | 7 |
| 974.98 | TC135344 | Similar to Q7Y1T5 (89 %) | OEE3 | VGGPPAPSGGGLPGTLNSDEAR | Yes | 16 | 9 |
| 855.40 | TC137623 | G7KGT1_MEDTR | Rubisco activase | GLAYDISDDQQDITR | No | 22 | 10 |
| 1,152.53 | TC137623 | G7KGT1_MEDTR | Rubisco activase | MGINPIMMSAGELESGNAGEPAK | No | 45 | 15 |
| 824.91 | TC137623 | G7KGT1_MEDTR | Rubisco activase | YLEGAALGDANQDAIK | No | 17 | 9 |

(continued)

Table 1
(continued)

| <i>m/z</i> (Da) | TC | UniRef100 unreviewed | Annotation | RT (min) | | | |
|------------------------|-----------|-----------------------------|-------------------|--------------------------|------------|-----------|----|
| | | | | Proteotypic | 120 | 30 | |
| 807.42 | TC149800 | G7KYC0_MEDTR | GAPDH | DSPLDVIAINDTGGVK | Yes | 38 | 14 |
| 890.51 | TC149800 | G7KYC0_MEDTR | GAPDH | VPTPNWSVVDLWVQVSK | Yes | 56 | 17 |
| 728.89 | TC149800 | G7KYC0_MEDTR | GAPDH | VVDLADIVANNWK | Yes | 53 | 17 |
| 1,285.14 | TC149800 | G7KYC0_MEDTR | GAPDH | YDSTLGFADAVKPVGTDGIVSDGK | Yes | 46 | 15 |
| 631.82 | TC149435 | G7IAW2_MEDTR | ATPase | TIAMDATEGVVR | No | 28 | 12 |
| 587.33 | TC149435 | G7IAW2_MEDTR | ATPase | VVDLLAPYQR | No | 11 | 7 |
| 872.98 | TC163858 | G7JUY1_MEDTR | ATPase CF1 | ATGIIAQIPVSEGVLGR | Yes | 40 | 14 |
| 806.95 | TC163858 | G7JUY1_MEDTR | ATPase CF1 | IVNTGTVLQVGDGIAR | Yes | 28 | 11 |
| 626.87 | TC163858 | G7JUY1_MEDTR | ATPase CF1 | VVNALAKPIDGR | Yes | 16 | 9 |
| 1,541.76 | TC149938 | B7FL07_MEDTR | PGK | FAPDANSQIVPASAIPDG | Yes | 70 | 30 |
| | | | | WMGLDIGPDSIK | | | |
| 1,173.06 | TC150861 | G7JJV9_MEDTR | FBA | GLVPLAGSNDSEWCQGLDGLASR | No | 58 | 17 |
| 825.49 | TC150861 | G7JJV9_MEDTR | FBA | IVDVLLIEQNIIVPGIK | No | 53 | 16 |
| 772.41 | TC150861 | G7JJV9_MEDTR | FBA | RLASIGLENTEANR | No | 7 | 6 |
| 672.33 | TC167914 | B7FIZ8_MEDTR | PRK | FYGEVTQQMLK | Yes | 15 | 9 |
| 637.34 | TC167914 | B7FIZ8_MEDTR | PRK | LTSVFGGAAEPPK | Yes | 12 | 8 |
| 1,214.61 | TC167914 | B7FIZ8_MEDTR | PRK | QYADAVIEVLPTQLIPDDNEGK | Yes | 59 | 18 |
| 939.45 | TC157261 | G7ILC6_MEDTR | ATPase chl | VFDYFSPTVEDSTKR | No | 22 | 11 |
| 732.39 | TC142539 | G7IW95_MEDTR | FNR chl | RLVYTNDAGEVVK | Yes | 23 | 9 |

| | | | | | | | |
|----------|----------|----------------------------|-------------------------|--------------------------|-----|----|----|
| 802.40 | TC154744 | G7IF28_MEDTR | Transketolase | ALPTYTPETPADATR | Yes | 15 | 8 |
| 529.80 | TC143176 | Homologue to Q39640 (96 %) | Glyoxylate oxidase | VPVFELDGGR | Yes | 19 | 10 |
| 756.46 | TC143176 | Homologue to Q39640 (96 %) | Glyoxylate oxidase | WLQTITSLPLVVK | Yes | 57 | 17 |
| 1,278.60 | TC164554 | G7J5X6_MEDTR | Plastocyanin | NINAGFPHNVTDFDEIPSGVDAAK | No | 40 | 13 |
| 721.39 | TC156968 | G7KQT5_MEDTR | FGS | FAQVTNPAIDPLR | Yes | 24 | 11 |
| 749.93 | TC161515 | G7I9Z0_MEDTR | Glycine dehydrogenase | IGVSVDSSGKPALR | Yes | 15 | 8 |
| 748.41 | TC147215 | G7L1U4_MEDTR | Ribose-5-P isomerase | SLGIPLSVLDDDNPR | Yes | 41 | 14 |
| 790.40 | TC156165 | Similar to A7PCN3 (85 %) | Unknown | IFVGNLPLFDVDSK | Yes | 43 | 15 |
| 842.46 | TC158157 | G7JZW9_MEDTR | GCS | IKPSDPSELESSLGAK | No | 35 | 13 |
| 944.01 | TC144015 | G7J013_MEDTR | AGT | AALDLIFEEGLENIAR | No | 87 | 24 |
| 751.42 | TC149279 | G7IE32_MEDTR | 2-Cys peroxiredoxin | SYGVLPDQGIALR | No | 38 | 13 |
| 744.9 | TC146112 | G7K4T4_MEDTR | FBA | GILAADESTGTIGKR | No | 27 | 11 |
| 1,211.08 | TC152332 | G7KIR1_MEDTR | Carbonic anhydrase | GLLSFPFDGAYSTDFFEEWVK | No | 93 | 24 |
| 843.45 | TC148305 | G7JA12_MEDTR | ATPase | ALDSQIAALSQDIVNK | Yes | 47 | 15 |
| 794.37 | TC149890 | G7K4R2_MEDTR | Hsp90 | SGDDDMTSLKDYVTR | Yes | 27 | 9 |
| 1,127.53 | TC143554 | Q84UC1_MEDTR | Uncharacterized protein | APITQQLPGESDTDFADFSSK | Yes | 38 | 13 |
| 781.43 | TC151139 | E1ANG4_MEDTR | GS2b, chl | TISKPVVEHPSELPK | No | 5 | 5 |
| 881.47 | TC144029 | G7JZ05_MEDTR | Polygalacturonase-1 | LLEATGISTVPGSQFGQK | Yes | 28 | 11 |
| 513.31 | TC144429 | G7I9L3_MEDTR | EF1 α | IGGIGTVPVGR | No | 10 | 7 |
| 767.37 | TC142554 | G7J7Y8_MEDTR | UGPase | YLSGEAQHVEWSK | No | 9 | 6 |
| 960.97 | TC144622 | G7ZZM3_MEDTR | RLS | DSTTHIADAASKDELQSR | No | 33 | 30 |

(continued)

Table 1
(continued)

| <i>m/z</i> (Da) | TC | UniRef100 unreviewed | Annotation | Sequence | RT (min) | | |
|------------------------|-----------|-----------------------------|-------------------------|--------------------------|--------------------|------------|-----------|
| | | | | | Proteotypic | 120 | 30 |
| 849.99 | TC144593 | G7KZM0_MEDTR | Hsp70 | IISGPAEKPLIGVNYK | No | 17 | 9 |
| 605.84 | TC142931 | G7J0Q8_MEDTR | Cinnamoyl-CoA reductase | IPTDTQPGLLR | Yes | 11 | 7 |
| 737.89 | TC152702 | Similar to P36212 (84 %) | 50S ribosomal protein | LGNDISDLTLSQAK | Yes | 27 | 11 |
| 659.34 | TC145241 | Similar to A7P862 (94 %) | TPI | IEISGQNSWVGK | Yes | 15 | 9 |
| 693.86 | TC154443 | Homologue to P35100 (71 %) | ATP-dependent protease | VIGQDEAVEAISR | Yes | 11 | 9 |
| 766.88 | TC144388 | Homologue to Q40977 (95 %) | MDAR | VVGAFLEGGTDPDENK | Yes | 15 | 9 |
| 1,129.11 | TC165606 | G7IH13_MEDTR | Elongation factor 2 | STLTDSLVAAGIIAQEVAGDVR | Yes | 52 | 28 |
| 712.84 | TC165606 | G7IH13_MEDTR | Elongation factor 2 | LWGENNFFDPATK | Yes | 19 | 10 |
| 873.96 | TC144042 | G7J567_MEDTR | 14-3-3 like | AAQDIAAADLPSITHPIR | Yes | 15 | 8 |
| 557.28 | TC153480 | Similar to A7PDS7 (91 %) | Unknown | DGFETYTLR | Yes | 8 | 7 |
| 1,026.97 | TC154064 | G7JNN9_MEDTR | Cysteine synthase | STPDAYMLQQFDNPSNPK | No | 33 | 29 |
| 1,275.62 | TC150490 | G8A0S6_MEDTR | FBPase | VLYEVFPMSFLMEQAGGQAAFTGK | Yes | 53 | 16 |

TC numbers of the in-house amino acid six-frame-translation of the *Medicago truncatula* database (DFCI release 10). UniRef100 unreviewed accessions specific for *M. truncatula* if available. RT (min) 120: Retention times of the selected target peptides for the 120-min shotgun LC-MS analysis. RT (min) 30: Retention times of the selected target peptides for the 30-min shotgun LC-MS analysis. Functional annotation abbreviations for *RJS* RubisCo large subunit, *RSS* Rubisco small subunit, *OEE* oxygen-evolving enhancer protein, *GAPDH* glyceraldehyde-3-phosphate-dehydrogenase, *PGK* phosphoglycerate kinase, *FNR* ferredoxin-NADP reductase, *FGS* ferredoxin-dependent glutamate synthase, *GCS* glycine cleavage system H protein, *AGT* alanine glyoxylate aminotransferase, *FBA* fructose-bisphosphate aldolase, *EF1 α* elongation factor 1-alpha, *UGPase* UDP-glucose pyrophosphorylase, *MDAR* monodehydroascorbate reductase, *FBPase* fructose-1,6-bisphosphatase

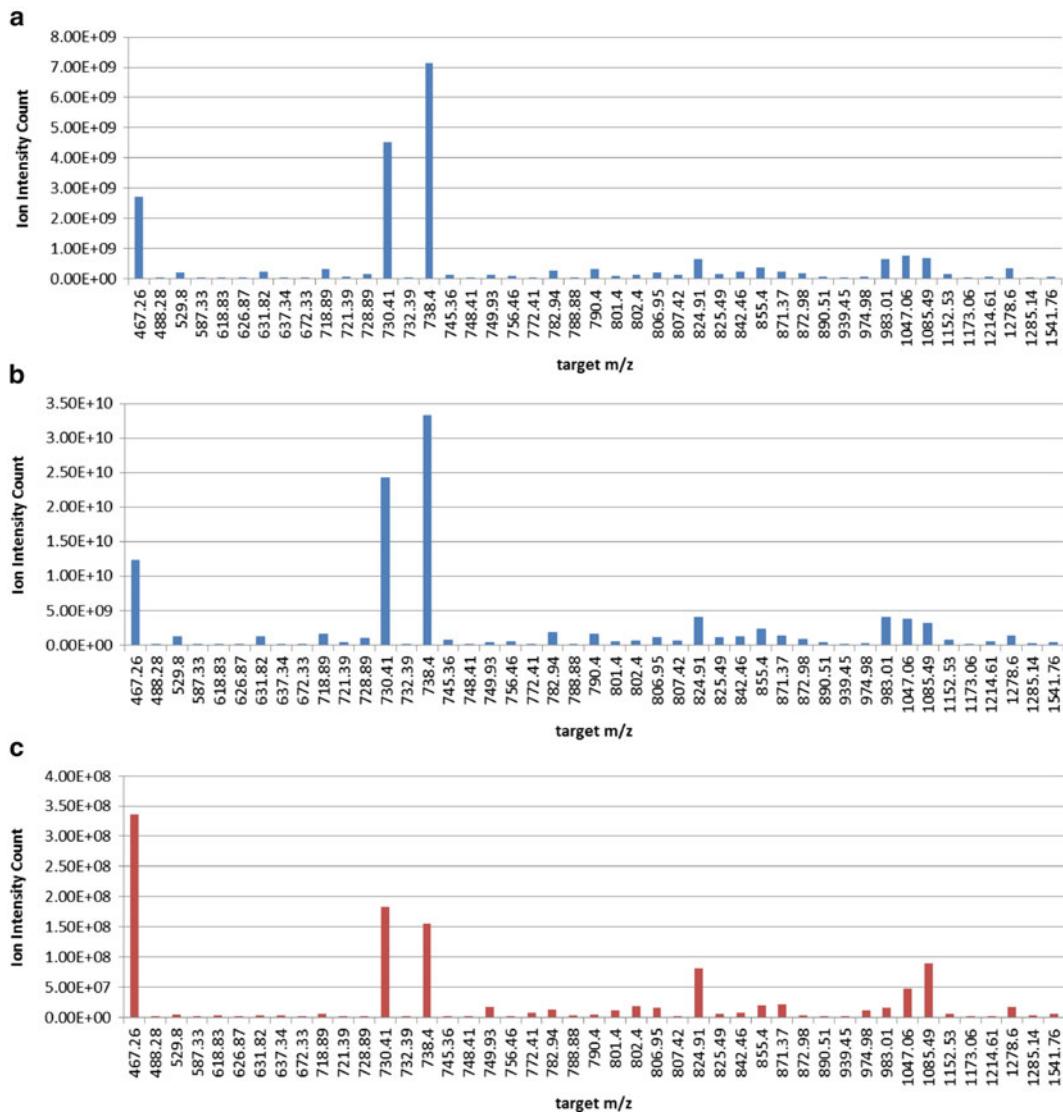


Fig. 1 Application of the Selpex approach to a 120- and 30-min LC-MS analysis. Mean ion intensity of three biological replicates for (a) target m/z during 120-min shotgun LC-MS DDP scan analysis, (b) 120-min shotgun LC-MS FS analysis, (c) 30-min shotgun LC-MS FS analysis

high-abundance peptides might be dampened. However, the FS analysis enhances ion count on MS1 level that is reduced during a DDP MS analysis and therefore enables significant reduction of MS runtime (Fig. 1). Runtime for the Selpex approach is linear to the number of samples. The analysis of three replicates including ProtMAX processing and identification time took about 5 min (see Note 7). The list of selected peptide targets is very flexible, is easy to extend, and can be adapted according to specific goals.

5 Notes

1. Although after database search several different peptides per proteins were identified, for Selpex here only those peptide targets with robust peak shape and reproducibility have been selected out of 45 different proteins. Thus, stringent filter criteria for peptide selection often result in one peptide per protein.
2. The selected peptides are not only useful for the described Selpex application but also excellent candidates for the selective reaction monitoring (SRM) absolute quantification.
3. This procedure needs to be done once for each organ or organism of interest. For *Medicago* leaves however, the present list can be directly used as a reference library of selected peptides for relative quantification using Selpex.
4. The FS-based MS analysis reduces/eliminates missing values.
5. The use of very short gradients may lead to ion suppression. The optimal gradient time required for the highest target coverage should therefore be determined prior to the final experiment.
6. It is recommended to use target-specific RT windows and adapt those depending on the gradient.
7. ProtMAX output can only be as good as MS analyses. For example, if two or more results are shown for the same target at a specific RT, reproducibility of chromatograms between samples might not be good or chosen environmental RT window was set too small. If targets are missing in the result file, it may not be in the sample, mass accuracy does not match between MS analysis and preference settings, or RT window does not match.

Acknowledgements

The work was funded by the Austrian Science Fund (FWF): P23441-B20. M.A.C. is funded by the Spanish Ministry of Science and Innovation.

References

1. Daher Z, Recorbet G, Valot B et al (2010) Proteomic analysis of *Medicago truncatula* root plastids. *Proteomics* 10:2123–2137
2. Dubinin J, Braun HP, Schmitz U, Colditz F (2011) The mitochondrial proteome of the model legume *Medicago truncatula*. *Biochim Biophys Acta* 1814:1658–1668
3. Larrainzar E, Wienkoop S, Weckwerth W et al (2007) *Medicago truncatula* root nodule proteome analysis reveals differential plant and bacteroid responses to drought stress. *Plant Physiol* 144:1495–1507
4. Larrainzar E, Wienkoop S, Scherling C et al (2009) Carbon metabolism and bacteroid

- functioning are involved in the regulation of nitrogen fixation in *Medicago truncatula* under drought and recovery. *Mol Plant Microbe Interact* 22:1565–1576
- 5. Kocher T, Swart R, Mechtr K (2011) Ultra-high-pressure RPLC hyphenated to an LTQ-Orbitrap Velos reveals a linear relation between peak capacity and number of identified peptides. *Anal Chem* 83:2699–2704
 - 6. Hoehnwarter W, Larhlimi A, Hummel J et al (2011) MAPA distinguishes genotype-specific variability of highly similar regulatory protein isoforms in potato tuber. *J Proteome Res* 10:2979–2991
 - 7. Egelhofer V, Hoehnwarter W, Lyon D, Weckwerth W, Wienkoop S (2013) Using ProtMAX to create high-mass-accuracy precursor alignments from label-free quantitative mass spectrometry data generated in shotgun proteomics experiments. *Nat Protoc* 3:595–601. doi:[10.1038/nprot.2013.013](https://doi.org/10.1038/nprot.2013.013)
 - 8. Hoehnwarter W, van Dongen JT, Wienkoop S et al (2008) A rapid approach for phenotype-screening and database independent detection of cSNP/protein polymorphism using mass accuracy precursor alignment. *Proteomics* 8: 4214–4225

Chapter 23

Soybean Proteomics

Zahed Hossain and Setsuko Komatsu

Abstract

Soybean, the world's most widely grown seed legume, is an important global source of vegetable oil and protein. Though, complete draft genome sequence of soybean is now available, but functional genomics studies remain in their infancy, as this agricultural legume species exhibits genetic constraints like genome duplications and self-incompatibilities. The techniques of proteomics provide much powerful tool for functional analysis of soybean. In the present review, an attempt has been made to summarize all significant contributions in the field of soybean proteomics. Special emphasis is given to subcellular proteomics in response to abiotic stresses for better understanding molecular basis of acquisition of stress tolerance mechanism. Detailed protocols of protein extraction, solubilization, fractionation of subcellular organelle, and proteins identification are explained for soybean proteomics. All this information would not only enrich us in understanding the plants response to environmental stressors but would also enable us to design genetically engineered stress tolerant soybean.

Key words Soybean, Proteomics, Methodology, Stress

1 Introduction

In the present “omics” era, proteomics has gained ample popularity over the genome-based technologies as it directly deals with the functional translated part of the genome. The journey of modern day proteomics started in 1975 with the establishment of high-resolution two-dimensional gel electrophoresis (2-DE) technique by O’Farrell [1]. Introduction of immobilized pH gradient strips [2, 3] followed by advancement in mass spectrometry (MS) has made proteomics a fast, sensitive and reliable technique for separation, identification and characterization of complex protein mixtures.

Soybean, one of the most important global sources of vegetable oil and protein, is the first major crop legume species with a published complete draft genome sequence [4]. However, improvements in the functional genomics studies on soybean are still in their infancy, as this agricultural legume species exhibits genome duplications and self-incompatibilities with a long generation time.

Identification and understanding of the biological function of any novel gene is a more ambitious goal than merely determining its sequence. Proteomics approach supported with genome-sequence data is thus a powerful tool to identify novel proteins and to follow temporal changes in protein expressions. In spite of the agricultural importance, soybean yield increment through conventional breeding over the past few decades has lagged behind those of cereals. Different abiotic and biotic impediments including flooding, drought, salinity, metal toxicity, nutrient limitations, and pathogen attacks curtail yield potential in soybean. Stress-induced changes in gene expression modulate metabolic processes through alteration of cellular protein abundance and function. Therefore, understanding how the function of proteins changes under stressed conditions is crucial for clarifying the molecular mechanisms underlying stress tolerance and crop injury.

The proteomic research done so far in soybean has mainly focused on the comparative analyses of protein abundance between control and stressed or tolerant/susceptible cultivars (Table 1). In addition, changes in protein expression profiling at various organ developmental stages including seed filling period of soybean have been well exploited. In the present review, an attempt has been made to summarize all significant contributions in the field of soybean proteomics. Special emphasis is given to subcellular proteomics in response to abiotic stresses for better understanding of changes in the protein abundance and plant stress tolerance.

2 Protein Extraction of Soybean

Protein extraction is the most critical step in a two-dimensional gel electrophoresis (2-DE) approach, as the amount and quality of the extracted proteins ultimately determine the protein spot number, resolution and intensity. Phenolic compounds, proteolytic and oxidative enzymes, terpenes, pigments, organic acids, inhibitory ions, and carbohydrates are the most common interfering substances present in recalcitrant plant tissues, resulting inferior results such as proteolytic breakdown, streaking and charge heterogeneity. Efforts have been made to standardize sample preparation protocols to get optimized yield. Instead of having physicochemical limitations of each and every protocol, the trichloroacetic acid (TCA)-acetone precipitation and phenol extraction methanol-ammonium acetate precipitation methods are most used as standard methods for removing interfering substances to obtain high quality gels [5].

Soybean seeds contain approximately 36 % protein, 30 % carbohydrate, 20 % oil, 9 % crude fiber and 5 % ash [6]. In addition, large amounts of secondary metabolites like kaempferol and quercetin are present in both seeds and leaves of soybean [7]. Increased activities of enzymes involved in phenolic biosynthetic

Table 1
Summary of functional proteomic analyses of soybean

| Conditions | Cultivar | Organ/organelle | Proteomic methodologies | Reference |
|----------------|---------------------------------|----------------------------------|--|-----------|
| Flooding | Enrei | Hypocotyl/root | IEF/IPG tube gel, SDS-PAGE, MALDI-TOF MS, nanoLC-MS/MS, protein sequencing | [28] |
| | Enrei | Hypocotyl/root (mitochondria) | IPG, SDS-PAGE, BN-PAGE, nanoLC-MS/MS | [42] |
| | Enrei | Hypocotyl/root (plasma membrane) | IEF tube gel, SDS-PAGE, MALDI-TOF MS, nanoLC-MS/MS, protein sequencing | [40] |
| | Enrei | Hypocotyl/root (cell wall) | IEF tube gel, SDS-PAGE, MALDI-TOF MS, nanoLC-MS/MS, protein sequencing | [39] |
| Waterlogging | Enrei | Hypocotyl/root | IEF/IPG tube gel, SDS-PAGE, MALDI-TOF MS, protein sequencing | [44] |
| | Asoagari | Root | IPG, SDS-PAGE, MALDI-TOF MS, ESI-MS/MS | [27] |
| Salinity | Enrei | Hypocotyl/root | IEF tube gel, SDS-PAGE, ESI-Q/TOF-MS/MS, protein sequencing | [32] |
| | Enrei | Leaf/hypocotyl/root | IEF tube gel, SDS-PAGE, MALDI-TOF MS, protein sequencing | [33] |
| Drought | Enrei | Leaf/hypocotyl/root | IPG, SDS-PAGE, nanoLC-MS/MS | [31] |
| Osmotic stress | Enrei | Hypocotyl/root (plasma membrane) | IEF tube gel, SDS-PAGE, LC MS/MS, nanoLC-MS/MS | [41] |
| Cadmium | Naviko | Suspension cell | SDS-PAGE/Q-TOF MS | [36] |
| | Enrei Harosoy | Root (microsome) | IPG, SDS-PAGE, nanoLC-MS/MS | [34] |
| | Harosoy Fukuyutaka CDH-80 | Leaf, root | IPG, SDS-PAGE, nanoLC-MS/MS, MALDI-TOF MS | [35] |
| Aluminum | BaXi 10 | Root | IPG, SDS-PAGE, MALDI-TOF MS | [17] |
| UV-B | Clark | Leaf | IPG, SDS-PAGE, MALDI-TOF MS | [45] |

(continued)

Table 1
(continued)

| Conditions | Cultivar | Organ/organelle | Proteomic methodologies | Reference |
|----------------------|----------------------------|--------------------|--|-----------|
| Ozone | Enrei | Leaf (chloroplast) | IEF tube gel, SDS-PAGE, MALDI-TOF MS, protein sequencing | [24] |
| Infection by | Williams 82 | Root hairs | IPG, gradient acrylamide gel, MALDI-TOF MS, Q-TOF MS | [37] |
| <i>B. japonicum</i> | Eb-b0-1 En1282 Enrei | Root | IPG, SDS-PAGE, nanoLC-MS/MS | [16] |
| Seed filling | Maverick | Developing seed | IPG, SDS-PAGE, MALDI-TOF MS | [23] |
| | Maverick | Developing seed | IPG, SDS-PAGE, nESI-LC-MS/MS | [26] |
| Storage protein | Jefferson | Seed | IPG, SDS-PAGE, MALDI-TOF MS | [22] |
| Developmental stages | Enrei | Leaf/flower | IEF tube gel, SDS-PAGE, MALDI-TOF MS, protein sequencing | [12] |

pathways in mesophyll cells indicate the presence of secondary phenolic metabolism in soybean leaves [8]. Presence of these interfering substances not only hampers high-quality protein extraction [9], but also impedes protein spot separation in high resolution 2-DE gels, resulting in streaking, smearing, and a significant reduction in the number of distinctly resolved protein spots [10]. Instead of these limitations, much progress has been made in standardization of the protein extraction methodologies from different tissues of soybean. Natarajan et al. [11] compared four different protein extraction/solubilization methods—urea, thiourea/urea, phenol, and a modified TCA /acetone to determine their effectiveness in separating soybean seed proteins by 2-DE. The thiourea/urea and TCA methods were found to be more suitable in resolving less abundant and high molecular weight proteins. In addition, these two methods exhibited higher protein resolution and spot intensity as compared to rest of the methods.

To compare proteomic changes of soybean leaves and flowers at various developmental stages, Ahsan and Komatsu [12] evaluated three different protein extraction protocols—TCA precipitation [13], phenol extraction method [14] with modifications and direct tissue homogenizing in suitable protein solubilization buffers. To optimize protein pellet solubilization buffer, A-buffer [1]

containing 8 M urea, 2 % Nonidet P-40, 2 % ampholine (pH 3.5–10), 5 % 2-mercaptoethanol, and 5 % Polyvinylpyrrolidone (PVP)-40; B-buffer [13] containing 7 M urea, 0.2 M thiourea, 0.2 mM tributylphosphine (TBP), 0.4 % CHAPS, 5 % PVP-40, and 2 % ampholine (pH 3–10); and C-buffer containing 8.5 M urea, 2.5 M thiourea, 5 % CHAPS, 1 % dithiothreitol (DTT), 1 % Triton X-100, and 0.5 % ampholines (pH 3–10 and 5–8) were tested. Combination of the phenol-based method with C-solubilization buffer generated high quality proteome maps in terms of well-separated resolved spots, spot intensity, and the number of proteins in the 2-DE gels with no horizontal streaking and high background noise levels. Authors claim that this optimized protein extraction protocol is equally efficient for all other tissues and organs of soybean, including the hypocotyl, leaf, petiole, stem, flower bud, and flower.

Soybean root proteins extracted in 10 % TCA and 0.07 % 2-mercaptoethanol in acetone followed by subsequent solubilization in the lysis buffer containing 8 M urea, 2 M thiourea, 5 % CHAPS, and 2 mM TBP results high quality gel with a good number of resolved protein spots [15, 16]. Addition of DTT and PVP in the soybean protein extraction buffer was found to be effective in enhancing the number of resolved spots in gels [17, 18]. Xu et al. [19] and Natarajan et al. [20] used a modified lysis buffer devoid of thiourea and TBP, that comprises of 9 M urea, 1 % CHAPS, 1 % DTT and 1 % ampholytes (pH 3–10), to solubilize the TCA/acetone precipitated proteins of soybean leaf and mature seed, respectively (Table 2).

Recently, Barbosa et al. [21] have analyzed mature seed proteome by extracting proteins in 50 mM Tris–HCl (pH 8.8), 1.5 mM KCl, 10 mM DTT, 1.0 mM phenylmethylsulfonyl fluoride (PMSF), and 0.1 % SDS followed by precipitation in 0.1 M ammonium acetate in methanol. However, phenol based protein extraction was shown to be more effective in extracting seed proteins both at mature [22] and seed filling stages [23]. In contrast to TCA/acetone or Tris–HCl buffer, protein extracted in buffer comprises of 50%phenol,0.45Msucrose,5mMEDTA,0.2%2-mercaptoethanol, 50 mM Tris–HCl (pH 8.8) when solubilized in 8 M urea, 2 M thiourea, 4 % CHAPS, 2 % Triton X-100, 50 mM DTT, ampholytes produced large number of reproducible protein spots (Table 2). Mooney and Thelen [22] compared the quality and purity of the acetone-precipitated and phenol extracted proteins. Occurrence of minimal horizontal streaking in 2-DE gels for phenol-extracted proteins indicates that acetone-precipitated sample contains more non-protein molecules. Ahsan et al. [24] evaluated a modified phenol method for extraction and analysis of soybean leaf proteome in response to ozone stress (Table 2). The leaf tissue was first homogenized in extraction buffer containing 0.5 M Tris–HCl (pH 8.3), 2 % Nonidet P-40, 20 mM MgCl₂, 2 % 2-mercaptoethanol, 1 mM PMSF and 0.7 M sucrose. An equal

Table 2
Summary of different proteomic methodologies used for soybean proteome analyses

| Plant tissue | Protein extraction buffer | Protein solubilization/lysis buffer | Proteomic methodologies | Spot resolved | Reference |
|---|--|---|--|-------------------------------|-----------|
| Roots | 10 % TCA in acetone containing 0.07 % DTT, 1 % PVP | 7 M urea, 2 M thiourea, 2 % CHAPS, 1 % DTT, 2 % ampholytes (pH 3–10) | IPG, 2-DE, MALDI-TOF MS | >1,200 | [17] |
| | 10 % TCA, 0.07 % 2-ME | 8 M urea, 2 M thiourea, 5 % CHAPS, 2 mM tributylphosphine, ampholytes (pH 3–10) | IPG, 2-DE, nanoLC-MS/MS | 439 | [16] |
| | 0.5 M Tris-HCl (pH 8.3), 2 % NP-40, 20 mM MgCl ₂ , 2 % 2-ME, 1 mM PMSF, 1 % PVP/fractionation with PEG 4000 | 8 M of urea, 1 % CHAPS, 0.5 % IPG buffer pH 4–7, 20 mM DTT, BPP | IPG, 2-DE, MALDI-TOF MS, ESI-MS/MS | ~900 | [27] |
| Root apex/ differentiated root zone | Acetone containing 10 % TCA, 0.07 % DTT | 9 M urea, 4 % CHAPS, 1 % DTT, 0.8 % ampholytes (pH 3–10), 35 mM Tris-HCl, 1 mM PMSF, 5 mM EDTA | IPG, 2-DE, DIGE Labeling, CBB staining, MALDI-TOF MS/MS | 1630 Cy-dye-/ 550 CBB | [18] |
| Root tips | 10 % TCA, 0.07 % 2-ME in acetone | 8 M urea, 2 M thiourea, 5 % CHAPS, 2 mM tributylphosphine | SDS-PAGE, ProQ Diamond phosphoprotein gel stain, nanoLC-MS/MS | – | [46] |
| Root hairs | 50 % phenol, 0.45 M sucrose, 5 mM EDTA, 0.2 % 2-ME, 50 mM Tris-HCl pH 8.8 | 8 M urea, 2 M thiourea, 2 % CHAPS, 2 % Triton X-100, 50 mM DTT, 2 mM TBP, 0.5 % ampholytes | IPG, 2-DE, MALDI-TOF MS, Q-TOF-MS | – | [37] |
| Roots and hypocotyls | Phosphate saline buffer (pH 7.6) containing 65 mM K ₂ HPO ₄ , 2.6 mM KH ₂ PO ₄ , 400 mM NaCl and 3 mM NaN ₃ followed by 10 % trichloro acetic acid | 8.5 M urea, 2.5 M thiourea, 5 % CHAPS, 100 mM dithiothreitol and 0.5 % ampholytes (pH 3.0–10.0/5.0–8.0) | IEF tube gel, 2-DE, MALDI-TOF MS | 235 (R) 330 (H) 340 (L) | [33] |
| Leaves | – | 7 M urea, 0.2 M thiourea, 0.2 mM TBP, 0.4 % CHAPS, 0.2 % ampholytes (pH 3.0–10.0), 5 % PVP-40 | | | |

| | | | | | |
|--------------------|---|--|--|--------|------|
| Leaves | 10 % TCA in acetone containing 0.07 % 2-ME | 9 M urea, 1 % CHAPS, 1 % DTT, 1 % ampholytes pH 3–10 | IPG, 2-DE, MALDI-TOF MS, LC-MS/MS | >300 | [19] |
| | 100 mM Tris-HCl pH 8.8, 50 mM DTT, 10 mM EDTA, 1 mM PMSF, 1 μ M peptstain A, 10 mM leupeptine/Tris-saturated phenol/100 % methanol containing 100 mM ammonium acetate, 10 mM DTT | 8.5 M urea, 2.5 M thiourea, 5 % CHAPS, 2 % Triton X-100, 100 mM DTT, 1 % ampholytes pH 3–10 | IPG, 2-DE | >1,400 | [47] |
| | 0.5 M Tris-HCl (pH 8.3), 2 % NP-40, 20 mM MgCl ₂ , 2 % 2-ME, 1 mM PMSF, 0.7 M sucrose/Tris-HCl saturated phenol (pH 8.0)/cold methanol containing 0.1 M ammonium acetate for precipitation | 8.5 M urea, 2.5 M thiourea, 5 % CHAPS hydrate, 1 % DTT, 1 % Triton X-100, and 0.5 % ampholytes pH 3–10/5–8 | IEF tube gel, 2-DE, MALDI-TOF MS | >500 | [24] |
| Mature seeds | 50 % phenol, 0.45 M sucrose, 5 mM EDTA, 0.2 % 2-ME, 50 mM Tris-HCl pH 8.8 | 8 M urea, 2 M thiourea, 2 % CHAPS, 2 % Triton X-100, 50 mM DTT, 2 mM TBP, 0.5 % ampholytes pH 3–10 | IPG, 2-DE, MALDI-TOF MS, LC-MS/MS | ≥128 | [22] |
| | 10 % TCA in acetone containing 0.07 % 2-ME | 9 M urea, 1 % CHAPS, 1 % DTT, 1 % ampholytes pH 3–10 | IPG, 2-DE, MALDI-TOF MS, LC-MS/MS | <50 | [20] |
| | 50 mM Tris-HCl pH 8.8, 1.5 mM KCl, 10 mM DTT, 1.0 mM PMSF, 0.1 % SDS/0.1 M ammonium acetate in methanol for precipitation | 7 M urea, 2 M thiourea, 2 % CHAPS, 0.002 % BPP, 0.5 % ampholytes | IPG, 2-DE, DIGE labeling, – CBB staining MALDI-QTOF MS/MS, ESI-QTOF MS/MS. | | [21] |
| Seed filling stage | 50 % phenol, 0.45 M sucrose, 5 mM EDTA, 0.2 % 2-ME, 50 mM Tris-HCl pH 8.8 | 8 M urea, 2 M thiourea, 2 % CHAPS, 2 % Triton X-100, 50 mM DTT, 2 % ampholytes | IPG, 2-DE, MALDI-TOF MS | 679 | [23] |
| | 50 % phenol, 0.45 M sucrose, 5 mM EDTA, 0.2 % 2-ME, 50 mM Tris-HCl pH 8.8 | 8 M urea, 2 M thiourea, 4 % CHAPS, 2 % Triton X-100, 50 mM DTT, 2 % ampholytes | IPG, 2-DE, Sec-MudPIT, nESI-LC-MS/MS | >900 | [26] |

TCA trichloroacetic acid, BPB bromophenol blue, CBB Coomassie brilliant blue, Sec-MudPIT semicontinuous multidimensional protein identification technology, PMSF phenyl methyl sulfonyl fluoride, 2-ME 2-mercaptoethanol

volume of Tris–HCl (pH 8.0) saturated phenol was added and proteins were precipitated by mixing with cold methanol containing 0.1 M ammonium acetate. Using this optimized protocol, authors successfully constructed soybean leaf proteome map of high quality with more than 500 reproducibly detected protein spots in CBB-stained gel. All these observations indicate that both phenol and TCA/acetone based extraction buffers are suitable for protein extraction from different tissues of soybean. However, as compared to TCA/acetone method, the protein yield is comparatively higher in phenol-based method in particular for soybean seed proteins.

3 Comprehensive Analysis of Soybean Proteins

In spite of the difficulties in extracting proteins from soybean, substantial soybean proteomic research has been carried out at different developmental stages in particular seed maturation and young seedling stages.

3.1 Seed Maturation Stage

Seed filling is the developmental period when rapid metabolic and morphological (seed size, weight and color) changes take place [25]. To better understand the metabolic processes associated with seed filling in soybean, Agrawal et al. [26] investigated the seed proteome at five developmental stages by two complementary proteomic approaches—2-DE and semicontinuous multidimensional protein identification technology (Sec-MudPIT) coupled with liquid chromatography-MS. In total, 478 nonredundant proteins were collectively identified through 2-DE and Sec-MudPIT analyses, of which only 70 proteins were common to both datasets. Major functional classes of proteins revealed from both 2-DE and Sec-MudPIT analyses were involved in primary metabolism, protein destination and storage, and energy. Sec-MudPIT approach identified threefold higher number of membrane proteins than 2-DE. Comparisons of quantitative seed-filling proteome of soybean and rapeseed were done to further understand the regulation of intermediary metabolism in protein-rich versus oil-rich seeds. Authors conclude that 2-DE and Sec-MudPIT could be used as complementary proteomic approaches for characterizing the soybean seed-filling proteome.

Similar proteomic study was previously performed by Hajduch et al. [23] to determine the expression profile of soybean seed proteins at 2, 3, 4, 5, and 6 weeks after flowering. Using 2-DE and matrix-assisted laser desorption ionization time-of-flight (MALDI) MS, high-resolution proteome reference maps, expression profiles of 679 spots, and corresponding MALDI MS spectra for each spot were successfully established. Searching NCBI soybean UniGene database using the MS-Fit program of Protein Prospector, led to

identify 422 proteins primarily involved in plant metabolism, protein destination and storage, metabolite transport, and disease/defense. An overall decrease in metabolism-related proteins versus an increase in proteins associated with destination and storage was observed during seed filling. A user-intuitive database (<http://oilseedproteomics.missouri.edu>) was constructed to access all expression profile data for proteomics research on soybean and other oilseeds plants.

Barbosa et al. [21] have recently compared the expression patterns of seed proteins of non-transgenic soybean variety with that of transgenic soybean which confers resistance to the herbicide glyphosate-*N*-(phosphonomethyl)glycine. After separation of seed proteins by 2-DE, 192 proteins were identified using MALDI quadrupole time-of-flight (QTOF) MS and electrospray ionization (ESI) QTOF MS. Majority of the identified proteins (50 %) were related to storage function. The second most abundant class of proteins (18 %) was found to be involved in cell growth/cell division. Authors were successful in identifying the enzyme CP4 EPSPS, involved in the genetic modification by ESI-QTOF MS/MS and using trypsin as a cleavage enzyme. Using two-dimensional difference gel electrophoresis (2D-DIGE) technique actin fragment, cytosolic glutamine synthetase, gycinin subunit G1, and glycine-rich RNA-binding protein were shown to be differentially expressed in transgenic seeds. In addition, transgenic seeds exhibited higher level of lipid peroxidation and increased activity of reactive oxygen species (ROS) scavenging enzymes. These findings led them conclude that the transgenic seeds experienced higher level of stress even when no herbicides were used.

3.2 Young Seedling Stage

Stress induced changes in protein profiling of soybean at young seedling stage have been well explored. Alam et al. [27] carried out proteomic analysis of 2 weeks old soybean roots exposed to waterlogging stress. 2-DE technique was exploited to separate the protein spots. Five novel proteins viz. translation initiation factor, apyrase, auxin-amidohydrolase and coproporphyrinogen oxidase induced under waterlogging condition were identified by MALDI-TOF MS or ESI-MS/MS analysis. Functional analysis revealed that these proteins were involved in signal transduction, programmed cell death, RNA processing, redox homeostasis and metabolisms of energy. Waterlogging induces progressive reduction in soil O₂ concentration, creating an anaerobic condition around the root zone. Higher abundance of proteins involved in glycolysis and fermentation pathway suggests that stressed plants meet their energy requirement through fermentation process. Authors propose that the soybean plant cope with the waterlogged condition through the management of carbohydrate consumption and by regulating programmed cell death.

Komatsu et al. [28] exploited proteomic technique in combination with transcriptomic to unravel the underlying molecular mechanism of soybean conferring flooding tolerance. High-coverage gene expression profiling analysis was carried out after 12 h of flooding of 2 days old soybean seedlings. Transcriptome technique revealed that genes associated with alcohol fermentation, ethylene biosynthesis, pathogen defense, and cell wall loosening were significantly upregulated following flood-induced stress. In total, 97 genes and 34 proteins were found to be altered after flooding. Interestingly, disease/defense-related proteins were increased at a transcriptional level but were decreased at a translational level following flooding. Altered expressions of hemoglobin, acid phosphatase, and Kunitz trypsin protease inhibitor were evident at both transcriptional and translational levels. However, abundance of molecular chaperons and ROS scavengers were changed only at the translational level. Findings of this study indicate that early responses might be important stress adaptation factors to ensure survival of the soybean plants against flood-induced hypoxia. Separate study revealed that proteins related to glycolysis and ROS scavenging were increased in roots of 2 days old soybeans exposed to 1 day of flooding stress [29].

In contrast to other increased ROS scavenging enzymes, two ascorbate peroxidases (APXs) were found to be decreased in response to flooding. This finding is consistent with the previous proteomic study by Shi et al. [30]. Authors identified ten flood-responsive proteins, among which cytosolic APX 2 was predominantly decreased under flooding condition. Northern-hybridization also confirmed that the abundance of cytosolic APX 2 transcript decreased significantly after flooding, as did the enzymatic activity of APX. Results suggest that cytosolic APX 2 is involved in flood-induced stress response of young soybean seedlings. A recent analysis of proteomic changes in roots of 2 days old soybean seedlings at post-flooding recovery stage identifies seven proteins involved in cell wall modification and S-adenosylmethionine synthesis [16]. Comparative root proteome analysis of control and 3-day flooding-experienced soybean reveals 70 differentially expressed protein spots, primarily involved in protein destination/storage and metabolic processes. Clustering analysis of the differentially expressed protein spots strongly indicates that 3 days of flooding causes significant changes in protein expression, even during post-flooding recovery period. Three days of flooding results in decrease of ion transport-related proteins and increase of proteins involved in cytoskeletal reorganization, cell expansion, and programmed cell death. Results indicate that alteration of cell structure through changes in cell wall metabolism and cytoskeletal organization may be involved in post-flooding recovery processes in soybean seedlings.

Drought constitutes another form of water stress that results from shortage of water. Unlike flooding, drought induces osmotic stress, which affects plants metabolism and yield. Mohammadi et al. [31] have recently investigated the response of 3-day-old soybean seedlings to drought and polyethylene glycol (PEG)-induced osmotic stress. Proteomic study revealed that root was the most drought-responsive organ as evident from 32, 13, and 12 proteins with altered expressions in response to drought stress, PEG treatment, and both, respectively. Expressions of metabolism-related proteins were shown to be increased in leaves of both PEG-treated and drought-stressed seedlings, while proteins related to energy production- and protein synthesis were decreased. In contrast to waterlogging stress, increased APX abundance was evident in drought stressed soybeans. Down regulation of methionine synthase both at mRNA and protein levels of drought-stressed plants irrespective of organs indicates its possible role in impairment of seedling growth under drought condition.

Soil salinity is considered as one of the environmental constraints that limits productivity of crop plants including legumes. To elucidate the response of soybean to salt stress, the related changes in protein expressions were investigated using proteomic approach [32]. Proteins from the hypocotyls and roots of 100 mM NaCl treated soybean were extracted and separated by 2-DE. MS analysis revealed increase of late embryogenesis-abundant protein, β -conglycinin, elicitor peptide three precursors, and basic/helix-loop-helix protein, while protease inhibitor, lectin, and stem 31-kDa glycoprotein precursor were decreased. Expressions of metabolism related proteins are mostly affected by salt stress. Sobhanian et al. [33] reported significant decrease of glyceraldehyde-3-phosphate dehydrogenase and fructokinase 2 in hypocotyls of 1-week-old soybean seedlings exposed to 40 mM NaCl treatment. Low expression of glyceraldehyde-3-phosphate dehydrogenase at both protein and mRNA levels coupled with decreased plant growth indicates that metabolism of glucose through glycolysis is important to meet the required energy to overcome the salinity stress.

Contamination of agricultural soil by heavy metals has become a global concern. High accumulation of toxic metals significantly affects soybean growth and development. Soybean cultivars although differ in their ability to uptake, accumulation and translocation of cadmium to aerial part, little attention has been paid so far to unravel the underlying molecular mechanism of cadmium tolerance. Ahsan et al. [34] investigated differential responses of root microsomal proteins in contrasting cadmium accumulating soybean cultivars exposed to 100 μ M of CdCl₂. Combined proteomic and metabolomics analyses reveal that proteins and amino acids associated with cadmium chelating pathways are highly active in low root-to-shoot cadmium translocating cultivar. In addition, proteins involved in lignin biosynthesis were shown to be increased

under stress. Proteomic findings suggest that translocation of cadmium from the root to the aerial parts might be prevented by the increased xylem lignification. To further dissect the protein networks involved in cadmium stress response in soybean, we performed a comparative proteomic study among contrasting cadmium accumulating soybean cultivars [35]. Study revealed that both high (Harosoy) and low (Fukuyutaka) cadmium accumulating cultivars and their recombinant inbred line CDH-80 share some common defense strategies to cope with the cadmium stress. High abundance of enzymes involved in glycolysis and TCA cycle was evident, that might help cadmium challenged cells to produce more energy necessary to meet the high energy demand. Moreover, enhanced expression of glutamine synthetase might be involved in phytochelatin-mediated detoxification of cadmium ions. Increased abundance of antioxidant enzymes (SOD, APX, CAT) further ensures cellular protection from ROS mediated damages under cadmium stress.

Role of stress-responsive proteins in conferring aluminum resistance in soybean has been elucidated by Zhen et al. [17]. To induce stress, 1-week-old soybean seedlings were exposed to 50 μ M AlCl_3 for 24, 48 and 72 h. 2-DE analysis of root proteins revealed 39 differentially expressed protein spots, with 21 increased, 13 newly induced and 5 decreased. The heat shock protein, glutathione S-transferase, chalcone-related synthetase, GTP-binding protein, and ABC transporter ATP-binding protein were identified as aluminum-responsive proteins. Results indicate that plants stress/defense, signal transduction, transport, protein folding, gene regulation, and primary metabolisms are primarily affected under aluminum stress.

3.3 Others

To reveal the stress-induced changes in protein expression, in vitro grown cells offer advantages over the young seedlings as all the cultured cells grow in uniform environment and uniform stress pressure can be applied. Sobkowiak and Deckert [36] studied the cadmium-induced changes in protein pattern in soybean cell suspension culture. Synthesis of [^{35}S]-labeled proteins and their accumulation were analyzed by SDS-PAGE and identification of selected protein bands was performed by mass spectrometry. Cadmium was found to induce superoxide dismutase, histone H2B, chalcone synthase and glutathione transferase in soybean cells.

As compared to root, root hair offers an excellent single-cell plant tissue for the proteomic study. Root hair proteome analysis following the establishment of soybean-*Rhizobium* symbiotic interaction led to identify novel proteins—phospholipase D and phosphoglucomutase that appeared to respond to *Bradyrhizobium japonicum* inoculation [37]. In addition, differential protein expressions in roots and root hairs of soybean were evident. Chitinase class I and a stress-induced gene H4 were found to be

present only in root hairs, while CDPK:calmodulin-like domain protein kinase isoenzyme β , phosphoenolpyruvate carboxylase, and ascorbate peroxidase 2 appeared more abundant in root hairs than in stripped roots. This work opened new avenues for further proteomic study of the legume–*Rhizobium* symbiotic interactions.

Recently, Salavati et al. [16] have analyzed the early symbiotic response in the roots of super-nodulating (En-b0-1) and non-nodulating (En1282) varieties, and their parental normal-nodulating variety (Enrei). The nanoLC MS/MS analysis identified 56 proteins from 48 differentially expressed protein spots in normal-nodulating soybean variety after bacterial inoculation. Proteins involved in metabolism and energy production were shown to be increased in super-nodulating and decreased in non-nodulating varieties compared to normal-nodulating variety. The super-nodulating and non-nodulating varieties responded oppositely to bacterial inoculation. Seven proteins out of 11 were decreased in super-nodulating varieties, while expression of protease some subunit alpha type 6, gamma glutamyl hydrolase, glucan endo-1,3-beta glucosidase, and nodulin 35 were increased in En-b0-1 compared to non-nodulating En1282. Results suggest that suppression of the autoregulatory mechanism in the super-nodulating soybean variety might be due to negative regulation of defense and signal transduction related processes.

4 Subcellular Proteomics of Soybean

Proteomic analysis of subcellular organelle provides fundamental information about the plants response to a given stress at the functional level and thus refines our knowledge about plant stress related signaling pathways. We have recently reviewed plant cell organelle proteomics in response to abiotic stress [38]. In this section, significant contributions of only soybean proteomic works related to abiotic stress responses are summarized.

4.1 Cell Wall Proteomics

In plants, cell wall is the first compartment that responds to environmental stress. Signals are then transmitted to the cell interior to trigger the plants defense cascade. Change in cell wall proteome in response to flooding stress was investigated in detail in soybean cultivar Enrei [39]. The CaCl_2 -extracted cell wall proteins were isolated via sucrose gradient centrifugation and analyzed using gel-based proteomic techniques. All together 16 protein spots out of 204 showed changes in abundance under flooding stress. Of the identified proteins, methionine synthases and copper amine oxidase were found to be increased. Interestingly, proteins related to lignification, such as lipoxygenases, germin-like protein precursors, stem 31 kDa glycoprotein precursors, Cu–Zn-superoxide dismutase, and copper amine oxidase were decreased by flooding. Furthermore,

lignin staining confirmed suppression of lignification in the roots of flooding stressed soybeans. Results suggested that the suppression of cell wall lignification was caused by down regulation of jasmonate biosynthesis following flooding stress.

4.2 Plasma Membrane Proteomics

Plasma membrane acts as the communication interface between the cellular cytoplasm and the extracellular environment. It has been a subject of interest of many proteomic researchers as processing of biotic and abiotic stress signals primarily occur in the plasma membrane. Komatsu et al. [40] explored the effects of flooding stress on soybean plasma membrane proteins using an aqueous two-phase partitioning method. 2-DE MS/protein sequencer-based proteomics and nanoLC-MS/MS-based proteomic techniques were compared to identify the flooding stress responsive proteins. Higher abundance of cell wall proteins in the plasma membrane of flooded plants suggests the contribution of plasma membrane in constructing cell wall. In addition, increased expression of superoxide dismutase indicates that the antioxidative system play a crucial role in protecting cells from oxidative stress damage following exposure to flooding stress. Moreover, heat shock cognate 70 kDa protein likely play a significant role in protecting other proteins from denaturation and degradation during flooding stress. Authors suggest that signaling proteins, such as 14-3-3 and serine/threonine protein kinase and band 7 family proteins, might work cooperatively to regulate plasma membrane H⁺-ATPase and thus maintain ion homeostasis.

Nouri and Komatsu [41] compared both gel-based and LC MS/MS-based proteomic techniques to study the PEG mediated osmotic stress induced changes in soybean plasma membrane proteome. Interestingly, LC MS/MS-based method was found to be more effective in identifying most of the transporter proteins and proteins with high number of transmembrane helices as well as low-abundance proteins. Using the gel-based proteomics, four and eight protein spots were identified as increased and decreased, respectively, while in the nanoLC MS/MS approach, 11 and 75 proteins were found as increased and decreased, respectively, under PEG treatment. Among the identified proteins, only seven were found to be mutual in two proteomics techniques. Increased abundance of calnexin protein was evident under osmotic stress by both proteomic approaches. Activation of plasma membrane H⁺-ATPase protein implies acceleration of ion efflux under hyperosmotic condition.

4.3 Mitochondrial Proteomics

Mitochondria have been a target for subcellular proteomic as most of the abiotic stresses primarily impair mitochondrial electron transport chain resulting excess ROS generation. Proteomics technique coupled with metabolomics was successfully used to elucidate the flooding stress impacts on mitochondrial function of flooded soybeans [42]. 2-DE and blue-native PAGE followed by nanoLC MS/MS were used to separate and identify the differentially

expressed mitochondrial matrix and membrane proteins respectively. To obtain a better insight of the modulation of TCA cycle in response to flooding condition, metabolites were identified using capillary electrophoresis (CE)-MS. Flooding appears to cause considerable impairment to the electron transport chain in the roots and hypocotyls of soybean seedlings. Abundance of inner membrane carrier proteins and proteins related to complexes III, IV, and V of the electron transport chain were found to be reduced, while proteins and metabolites related to TCA and gamma-amino butyrate shunt were increased by flooding stress resulting high NADH production. These responses of mitochondria to flooding stress appear to be very similar with that of low oxygen stress, however considerably different from their response to osmotic stress.

4.4 Peroxisomal Proteomics

Peroxisomes, the ubiquitous cell organelles of eukaryotic cells, play a major role in the breakdown of long chain fatty acids through beta-oxidation. Although, peroxisomal proteins are encoded by nuclear genes but they are finally targeted to peroxisomes due to the presence of Peroxisome Targeting Signals (PTS). A majority of the peroxisomal proteins are involved in a variety of oxidative metabolic reactions, however functions of a large number of proteins remain unknown. Arai et al. [43] successfully characterized peroxisomal proteins by a combination of 2-DE and peptide mass fingerprinting using MALDI-TOF MS. Percoll density gradient centrifugation followed by iodixanol density gradient centrifugation were used to isolate purified peroxisomes from etiolated soybean cotyledons. Identified proteins are mainly enzymes involved in fatty acid β -oxidation, glyoxylate cycle, photorespiratory glycolate metabolism, stress response and metabolite transport. Authors demonstrated that the plant peroxisomes contain a short-chain dehydrogenase/reductase family protein, enoyl-CoA hydratase/isomerase family protein, 3-hydroxyacyl-CoA dehydrogenase-like protein and a voltage-dependent anion-selective channel protein.

5 Conclusion

The scenario of soybean proteomics has started changing since completion of soybean draft genome sequence. In spite of being a recalcitrant plant species, protein extraction protocols have been standardized to get optimized 2-DE results in terms of protein spot number and spot intensity. Different tissue specific proteomic studies reveal that phenol and TCA/acetone based extraction protocol are most suitable for soybean protein extraction. Construction of detail quantitative soybean proteome reference maps not only help functional genomic studies but also provide essential tool for rapid identification of soybean mutants/transgenic lines. Identifications of low-abundance proteins become possible with the development of sensitive stains and rapid technical advancement in

the MS technology. Proteomic research on soybean response to abiotic stresses both at whole plant and organelle levels provides new insights of stress adaptation. More initiatives need to be taken to delineate the molecular basis of acquisition of stress tolerance mechanism at the organelle level. In depth information about the expression of stress induced novel marker proteins would further enable us to design genetically engineered stress tolerant soybean.

References

1. O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 250:4007–4021
2. Bjellqvist B, Ek K, Righetti PG et al (1982) Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys Methods* 6:317–339
3. Görg A, Postel W, Günther S (1988) The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 9:531–546
4. Schmutz J, Cannon SB, Schlueter J et al (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178–183
5. Carpentier SC, Witters E, Laukens K et al (2005) Preparation of protein extracts from recalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis. *Proteomics* 5:2497–2507
6. Padgett SR, Taylor NB, Nida DL et al (1996) The composition of glyphosate-tolerant soybean seeds is equivalent to that of conventional soybeans. *J Nutr* 126:702–716
7. Buttery BR, Buzzell RI (1973) Soybean flavonol glycosides: identification and biochemical genetics. *Can J Bot* 53:309–313
8. Cosio EG, McClure JW (1984) Kaempferol glycosides and enzymes of flavonol biosynthesis in leaves of a soybean strain with low photosynthetic rates. *Plant Physiol* 74:877–881
9. Tsugita A, Kamo M (1999) 2-D electrophoresis of plant proteins. *Methods Mol Biol* 112: 95–97
10. Saravanan RS, Rose JK (2004) A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics* 4:2522–2532
11. Natarajan S, Xu C, Caperna TJ et al (2005) Comparison of protein solubilization methods suitable for proteomic analysis of soybean seed proteins. *Anal Biochem* 342:214–220
12. Ahsan N, Komatsu S (2009) Comparative analyses of the proteomes of leaves and flowers at various stages of development reveal organ-specific functional differentiation of proteins in soybean. *Proteomics* 9:4889–4907
13. Toorchi M, Nouri MZ, Tsumura M et al (2008) Acoustic technology for high-performance disruption and extraction of plant proteins. *J Proteome Res* 7:3035–3041
14. Hurkman WJ, Tanaka CK (1986) Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol* 81:802–806
15. Nanjo Y, Skultety L, Ashraf Y et al (2010) Comparative proteomic analysis of early-stage soybean seedlings responses to flooding by using gel and gel-free techniques. *J Proteome Res* 9:3989–4002
16. Salavati A, Khatoon A, Nanjo Y et al (2012) Analysis of proteomic changes in roots of soybean seedlings during recovery after flooding. *J Proteomics* 75:878–893
17. Zhen Y, Qi JL, Wang SS et al (2007) Comparative proteome analysis of differentially expressed proteins induced by Al toxicity in soybean. *Physiol Plant* 131:542–554
18. Mathesius U, Djordjevic MA, Oakes M et al (2011) Comparative proteomic profiles of the soybean (*Glycine max*) root apex and differentiated root zone. *Proteomics* 11:1707–1719
19. Xu C, Garrett WM, Sullivan JH et al (2006) Separation and identification of soybean leaf proteins by two-dimensional gel electrophoresis and mass spectrometry. *Phytochemistry* 67:2431–2440
20. Natarajan S, Xu C, Bae H et al (2006) Characterization of storage proteins in wild (*Glycine soja*) and cultivated (*Glycine max*) soybean seeds using proteomic analysis. *J Agric Food Chem* 54:3114–3120
21. Barbosa HS, Arruda SC, Azevedo RA et al (2012) New insights on proteomics of transgenic soybean seeds: evaluation of differential expressions of enzymes and proteins. *Anal Bioanal Chem* 402:299–314
22. Mooney BP, Thelen JJ (2004) High-throughput peptide mass fingerprinting of soybean seed proteins: automated workflow and utility of UniGene expressed sequence tag databases for protein identification. *Phytochemistry* 65:1733–1744

23. Hajduch M, Ganapathy A, Stein JW et al (2005) A systematic proteomic study of seed filling in soybean. Establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. *Plant Physiol* 137:1397–1419
24. Ahsan N, Nanjo Y, Sawada H et al (2010) Ozone stress-induced proteomic changes in leaf total soluble and chloroplast proteins of soybean reveal that carbon allocation is involved in adaptation in the early developmental stage. *Proteomics* 10:2605–2619
25. Agrawal GK, Thelen JJ (2006) Large scale identification and quantification profiling of phosphoproteins expressed during seed filling in oilseed rape. *Mol Cell Proteomics* 5:2044–2059
26. Agrawal GK, Hajduch M, Graham K et al (2008) In-depth investigation of the soybean seed-filling proteome and comparison with a parallel study of rapeseed. *Plant Physiol* 148: 504–518
27. Alam I, Lee DG, Kim KH et al (2010) Proteome analysis of soybean roots under waterlogging stress at an early vegetative stage. *J Biosci* 35:49–62
28. Komatsu S, Yamamoto R, Nanjo Y et al (2009) A comprehensive analysis of the soybean genes and proteins expressed under flooding stress using transcriptome and proteome techniques. *J Proteome Res* 8:4766–4778
29. Hashiguchi A, Sakata K, Komatsu S (2009) Proteome analysis of early-stage soybean seedlings under flooding stress. *J Proteome Res* 8:2058–2069
30. Shi F, Yamamoto R, Shimamura S et al (2008) Cytosolic ascorbate peroxidase 2 (cAPX 2) is involved in the soybean response to flooding. *Phytochemistry* 69:1295–1303
31. Mohammadi PP, Moieni A, Hiraga S et al (2012) Organ-specific proteomic analysis of drought-stressed soybean seedlings. *J Proteomics* 75:1906–1923
32. Aghaei K, Ehsanpour AA, Shah AH et al (2009) Proteome analysis of soybean hypocotyl and root under salt stress. *Amino Acids* 36:91–98
33. Sobhanian H, Razavizadeh R, Nanjo Y et al (2010) Proteome analysis of soybean leaves, hypocotyls and roots under salt stress. *Proteome Sci* 29:8–19
34. Ahsan N, Nakamura T, Komatsu S (2012) Differential responses of microsomal proteins and metabolites in two contrasting cadmium-accumulating soybean cultivars. *Amino Acids* 42:317–327
35. Hossain Z, Hajika M, Komatsu S (2012) Comparative proteome analysis of high and low cadmium accumulating soybeans under cadmium stress. *Amino Acids*. doi:[10.1007/s00726-012-1319-6](https://doi.org/10.1007/s00726-012-1319-6)
36. Sobkowiak R, Deckert J (2006) Proteins induced by cadmium in soybean cells. *J Plant Physiol* 163:1203–1206
37. Wan J, Torres M, Ganapathy A et al (2005) Proteomic analysis of soybean root hairs after infection by *Bradyrhizobium japonicum*. *Mol Plant Microbe Interact* 18:458–467
38. Hossain Z, Nouri MZ, Komatsu S (2012) Plant cell organelle proteomics in response to abiotic stress. *J Proteome Res* 11:37–48
39. Komatsu S, Kobayashi Y, Nishizawa K et al (2010) Comparative proteomics analysis of differentially expressed proteins in soybean cell wall during flooding stress. *Amino Acids* 39: 1435–1449
40. Komatsu S, Wada T, Abaléa Y et al (2009) Analysis of plasma membrane proteome in soybean and application to flooding stress response. *J Proteome Res* 8:4487–4499
41. Nouri MZ, Komatsu S (2010) Comparative analysis of soybean plasma membrane proteins under osmotic stress using gel-based and LC MS/MS-based proteomics approaches. *Proteomics* 10:1930–1945
42. Komatsu S, Yamamoto A, Nakamura T et al (2011) Comprehensive analysis of mitochondria in roots and hypocotyls of soybean under flooding stress using proteomics and metabolomics techniques. *J Proteome Res* 10: 3993–4004
43. Arai Y, Hayashi M, Nishimura M (2008) Proteomic analysis of highly purified peroxisomes from etiolated soybean cotyledons. *Plant Cell Physiol* 49:526–539
44. Komatsu S, Sugimoto T, Hoshino T et al (2010) Identification of flooding stress responsible cascades in root and hypocotyl of soybean using proteome analysis. *Amino Acids* 38:729–738
45. Xu C, Sullivan JH, Garrett WM et al (2008) Impact of solar ultraviolet-B on the proteome in soybean lines differing in flavonoid contents. *Phytochemistry* 69:38–48
46. Nanjo Y, Skultety L, Uváčková L, Klubicová K, Hajduch M, Komatsu S (2012) Mass spectrometry-based analysis of proteomic changes in the root tips of flooded soybean seedlings. *J Proteome Res* 11:372–385
47. Sarma AD, Oehrle NW, Emerich DW (2008) Plant protein isolation and stabilization for enhanced resolution of two-dimensional polyacrylamide gel electrophoresis. *Anal Biochem* 379:192–195

Chapter 24

Proteome Analysis of Orphan Plant Species, Fact or Fiction?

Sebastien C. Carpentier and Twan America

Abstract

Biological research has focused in the past on model organisms and most of the functional genomics studies in the field of plant sciences are still performed on model species or reference species that are characterized to a great extent. However, numerous non-model plants are essential as food, feed, or energy resource. Some features and processes are unique to these plant species or families and cannot be approached via a model plant. The power of all proteomic and transcriptomic methods, i.e., high throughput identification of candidate gene products, tends to be lost in orphan species due to the lack of genomic information, the complexity of the genome (protein inference problem, polyploidy) or due to the sequence divergence to a related sequenced reference variety or to a related model organism. Nevertheless, a proteomics approach has a great potential to study orphan species. This chapter reviews concisely orphan plants from a proteomic angle and provides an outline of the problems encountered when initiating the proteome analysis of a non-model organism. We discuss briefly the problems and solutions for orphan plants associated with sample preparation and focus further on the difficulties associated with protein redundancy in polyploid species and the protein inference issue which is particularly associated with a peptide based proteomics approach.

Key words Polyploidy, Uncharacterized complex genome, Protein inference, Databases, Protein extraction, Peptide based analysis, Protein based analysis

1 Introduction

What is a model organism? In past decades, the term “model organism” has been applied to species that facilitate experimental laboratory research. Research communities have focused on model organisms to gain insight into some general principles that underlie various disciplines. The first and only classical plant model, *Arabidopsis thaliana*, is ideal for laboratory studies. It has a short life cycle, a small size, a good production of seeds, and a relatively small genome that is completely sequenced. Although *Arabidopsis* is the preeminent model plant, to date still 48 % of the *Arabidopsis* genome has an unknown molecular function, 51 % belongs to an unknown biological process and of 48 % the cellular localization is unknown (TAIR10 genome statistics). This illustrates the challenge to work with orphan species.

What is an orphan species? Orphan plant species are crops/plants that are poorly characterized at the genetic and/or physiological level. With the recent increase in the number of genome-sequencing projects, the definition of a model organism has broadened [1] and many species that could be considered as an orphan have become a reference. Currently, 15,902 genomes are considered to be completely sequenced and publicly available and the number is still growing (www.genomesonline.org last update: 2012-03-12). Those species fall under the new and broad definition of “model organism” or reference species. The green plants or Viridiplantae are largely under-represented with only eight plant genomes completed, publicly available and reasonably well annotated: *A. thaliana* (thale cress), *Oryza sativa* (japonica cultivar-group) (rice), *Glycine max* (soybean), *Medicago truncatula* (barrel medic), *Populus trichocarpa* (black cottonwood), *Sorghum bicolor*, *Vitis vinifera* (wine grape), and *Zea mays* (corn) (<http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html>). All those new model plants have a relatively small genome size but are not necessary ideal as a laboratory model. The NCBI Plant Genomes Central expects that the genome sequencing of *Lotus japonicus* (lotus), *Manihot esculenta* (cassava), *Solanum lycopersicum* (tomato), *Solanum tuberosum* (potato), *Brachypodium distachyon*, and *Carica papaya* (papaya) will be completed and the data will be publicly available in the near future. However, there are approximately 300,000 known species of land plants and the new reference species represent only a handful of species and families. The genome size of many crops is large, complex and uncharacterized making them orphans. Polyploidy and allopolyploidy have played an important role in the evolution of plants including important crops [2]. While the presumed advantage of allopolyploidy is an increase in genetic diversity, its occurrence considerably complicates the proteome analysis of crops [3]. Numerous crops and plant species have different levels of ploidy. Cultivated sugar cane (*S. officinatum*) is a good example. It is a hybrid of different species and it has a complex octoploid genome with chromosome number ranging from $2n=70-140$ [4]. Though advances in high-throughput and computational technologies have resulted nowadays in the sequencing of the genome of many plant species, for simplicity reasons, most genome sequencing projects of crops focus on double haploid or diploid varieties to set up a reference genome. It is indeed extremely useful and necessary to have reference species/varieties. Probably in the future the price of sequencing will continue to drop so dramatically that each variety can be sequenced, but currently we still need a technique that is able to analyze existing orphan species and new hybrids. As a matter of fact, also most of the proteome studies are still performed on *Arabidopsis* (Fig. 1). Insights from the model *Arabidopsis* will undoubtedly boost crop science but *Arabidopsis* is not a crop and will never feed the world. Consequently, there is a great need for proteomics of orphans. Some features and processes are unique and cannot be approached via a model or a related reference species. However, to be able to analyze orphans, one must take several hurdles.

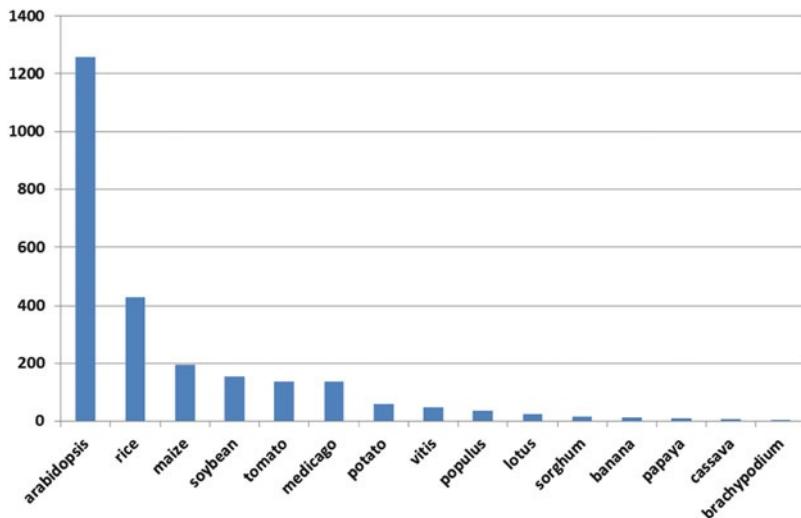


Fig. 1 Number of plant proteomics publications till March 2012 reported in ISI Topic=(proteomics and “species”)

2 Protein Extraction

Most plant tissues are not a ready source for protein extraction and need specific precautions. The cell wall and the vacuole make up the majority of the cell mass. Consequently, plant tissues have a relatively low protein content. Moreover, the cell wall and the vacuole are associated with numerous substances responsible for irreproducible results such as proteolytic breakdown and charge heterogeneity. Compared to *Arabidopsis*, several orphan species contain extremely high levels of interfering compounds. The majority of the plant protocols introduce a precipitation step to concentrate the proteins and to separate them from the interfering compounds. The protein precipitation step can be preceded by a denaturing or non-denaturing extraction step and is combined with one or two washing steps to remove introduced salt ions and other remaining interfering substances. The most commonly used method for extraction of plant proteins is the trichloroacetic acid (TCA)/acetone precipitation method [5]. Proteins are very sensitive to denaturation at low pH. TCA is a strong acid (pK_a 0.7) that is soluble in organic solvents. We have tested in the past different protocols to extract proteins from plant samples [6]. In our hands and also according to others [7, 8], the phenol extraction protocol proved to be the most powerful for recalcitrant plant tissues. The protein precipitation step is in this method achieved by ammonium acetate and methanol and is preceded by a denaturing phenol extraction. Moreover, the low pH of TCA might create problems with basic chemical labeling methods. For a more detailed discussion of protein extraction the reader is referred to [6, 9, 10].

3 Protein/Peptide Separation and Identification

Two approaches are generally distinguished in the field of proteome analysis a protein based approach (in general referred to as gel based approach) and a peptide based approach (in general referred to gel free or as shot gun approach). In the first approach proteins are separated and quantified. The proteins of interest are then digested and the resulting peptides identified via mass spectrometry. In the second approach, protein digestion precedes the separation and quantification of peptides. Yates and coworkers were one of the pioneers to explore the use of two-dimensional liquid chromatography coupled to electrospray tandem mass spectrometry (the Mudpit approach) to realize automated high-throughput shot gun proteomics [11]. A hybrid form where separation of proteins by gel electrophoresis is followed by the separation of peptides via liquid chromatography is called geLC. Almost all proteome studies on orphan species preferentially use a protein-based approach since all peptide-based strategies have the disadvantage to lose connectivity between peptides derived from the same protein [9]. The high resolving power of the two-dimensional protein-based separation technique is of great importance aiming for an individual protein separation. This assures that all derived peptides come from the same protein. In the case of protein separation techniques with a lower resolution (e.g., broad IPG strips, 1D SDS-PAGE, or geLC-MS) multiple proteins are digested simultaneously resulting in a more complex peptide pool. When a combination of masses of possibly non-related peptides is submitted to a database search, the risk of generating false positive identifications increases. It is therefore crucial that the parent ions are analyzed further by Tandem mass spectrometry (MS/MS). MS/MS has been used for decades to obtain structural information of (bio)molecules. The main advantage of a protein based approach is that it can be successfully used for the identification of protein orthologs based on structural information of the different peptides. For peptides derived of the same protein, MS/MS generates sequence specific information and the information content of such spectra in combination with the parent masses (Peptide Mass Fingerprinting (PMF)) can lead to a successful identification. In the MS mode, peptides with a high signal-to-noise (S/N) ratio are selected for further fragmentation. In the case of orphan species, it is possible that such intense peptides are less or not informative (i.e., not present in a database and not conserved in a sequenced species) and classical identification will fail since more homologous peptides with a lower S/N ratio will not be selected for MS/MS. However, when the peptide mixture consists of a limited amount of peptides which are all derived from a single protein, the informative peptides with a lower S/N ratio have a bigger chance of being selected for MS/MS, making protein identification more likely. Unfortunately, almost all software tools are developed to

search in a non-error tolerant way against a database of known proteins, which produces low protein scores when several peptide sequences of the orphans are not identical. EST sequences are nowadays quite efficiently and cheaply achieved with the next generation sequencing techniques and are indispensable.

Depending on the availability of EST sequences, the genome status of the species under investigation and on the degree of homology to a model organism, de novo sequencing is essential to obtain sequence information. Sequence reconstruction of an unknown peptide based solely on the acquired mass data is referred to as de novo peptide sequencing. Early de novo sequencing involved the use of chemical microsequencing using Edman chemistry. This method, however, is quite expensive in terms of reagent cost and suffers from a low throughput and sensitivity. MS based de novo sequencing is much more sensitive (needs less starting material). ESI-MS/MS routinely provides more informative MS/MS spectra than MALDI-TOF MS/MS. Multiple charged peptide ions tend to fragment more equally across a given sequence than do singly charged ion species [12, 13]. High precision mass spectrometry dramatically improves the performance for de novo sequencing. The high precision eliminates drastically the number of candidate peptide sequences that could fit a tandem mass spectrum. Unfortunately, separation of peptides prior to MS/MS is relatively expensive and time consuming. For these reasons, MALDI is often preferred in a protein based approach because of ease of use, speed and the ability to include MALDI spotting in automated digestion protocols on liquid handling systems. In addition, a MALDI approach has the advantage that it has the potential to store temporarily the targets for reanalysis when certain data are not yet fully explored. We have recently optimized an automated approach for the derivatization of peptides enabling a facilitated MALDI based MS/MS de novo identification [3].

The protein based 2-DE technique is an excellent choice for proteomics in orphan plants [9]. Unfortunately the technique has some major drawbacks, i.e., it has a very poor performance regarding the analysis of hydrophobic and basic proteins and is limited in throughput. A peptide based proteome analysis is the perfect answer to these limitations. However, a peptide based approach is this feasible for an orphan plant species? For a review on the analysis of hydrophobic membrane proteins of orphan species the reader is referred to [14]. We will focus here further on a peptide-based workflow in general.

As indicated above, proteomics on orphans relies the presence of EST libraries, on de novo sequencing and on cross-species comparison. To increase the chance of identifying conserved peptides, it is important to ensure a good peptide separation. In this proposed workflow this is achieved by using a RP-RP UPLC system coupled to a high accuracy mass spectrometer as described by Gilar and colleagues [15]. It is important to build a species specific in-house database and to search this database in a non-error tolerant manner. Subsequently

an error tolerant search is performed automatically to the unassigned spectra. When this results in a negative result, *de novo* sequencing is automatically applied to the unassigned spectra. This *de novo* approach is a crucial step since it allows a homology search of sequences instead of a search based on *m/z* values. An *in silico* innovation in the proposed workflow is the visualization of the peptides that were assigned to specific proteins using Cytoscape. By a simultaneous visualization of the significant hits of the orphan specific search and the hits of the cross-species search, the confidence of the orphan specific hits increases (Fig. 2). Moreover, some peptides wrongly assigned to cross-species hits are visualized and can be removed. Moreover, Cytoscape also allows the visualization of the protein inference, a common problem in a cross-species approach (Figs. 3 and 4).

4 Materials (Peptide Based Approach)

4.1 Protein Extraction

1. Liquid nitrogen.
2. Pestle and mortar (washed and baked in an oven at 180 °C overnight prior to use).
3. Phenol, acetone and methanol compatible tubes and pipettes.
4. 10 % (w/v) DTT. Make freshly in MQ H₂O.
5. Extraction buffer: 50 mM Tris–HCl, pH 8.5, 5 mM EDTA, 100 mM KCl, 1 % (w/v) DTT, 30 % (w/v) sucrose, and complete protease inhibitor cocktail (according to the manufacturer, Roche Applied Science) in MQ H₂O. Prepare freshly.
6. Tris buffered phenol, pH 8 (see Note 1).
7. Precipitation solution: 100 mM ammonium acetate in methanol. Store at -20 °C.
8. Cold acetone. Store at -20 °C.
9. Lysis solution: 8 M urea, 5 mM DTT.
10. Protein Quantification kit 2D Quant (GEhealthcare, Diegem, Belgium).

4.2 Protein Digestion and Cleanup

1. 0.5 M Iodoacetamide (IAA).
2. 100 mM Ammonium bicarbonate (ABC).
3. 0.2 µg/µl Trypsin (Promega).
4. 10% trifluoro acetic acid (TFA).
5. Solid phase extraction (SPE) column (Supelco Inc, Bellefonte, PA, USA).
6. Acetonitrile.
7. 0.1 M ammonium formate (AF).
8. SpeedVac.
9. NanoDrop (Thermo Scientific, Waltham, MA, USA).

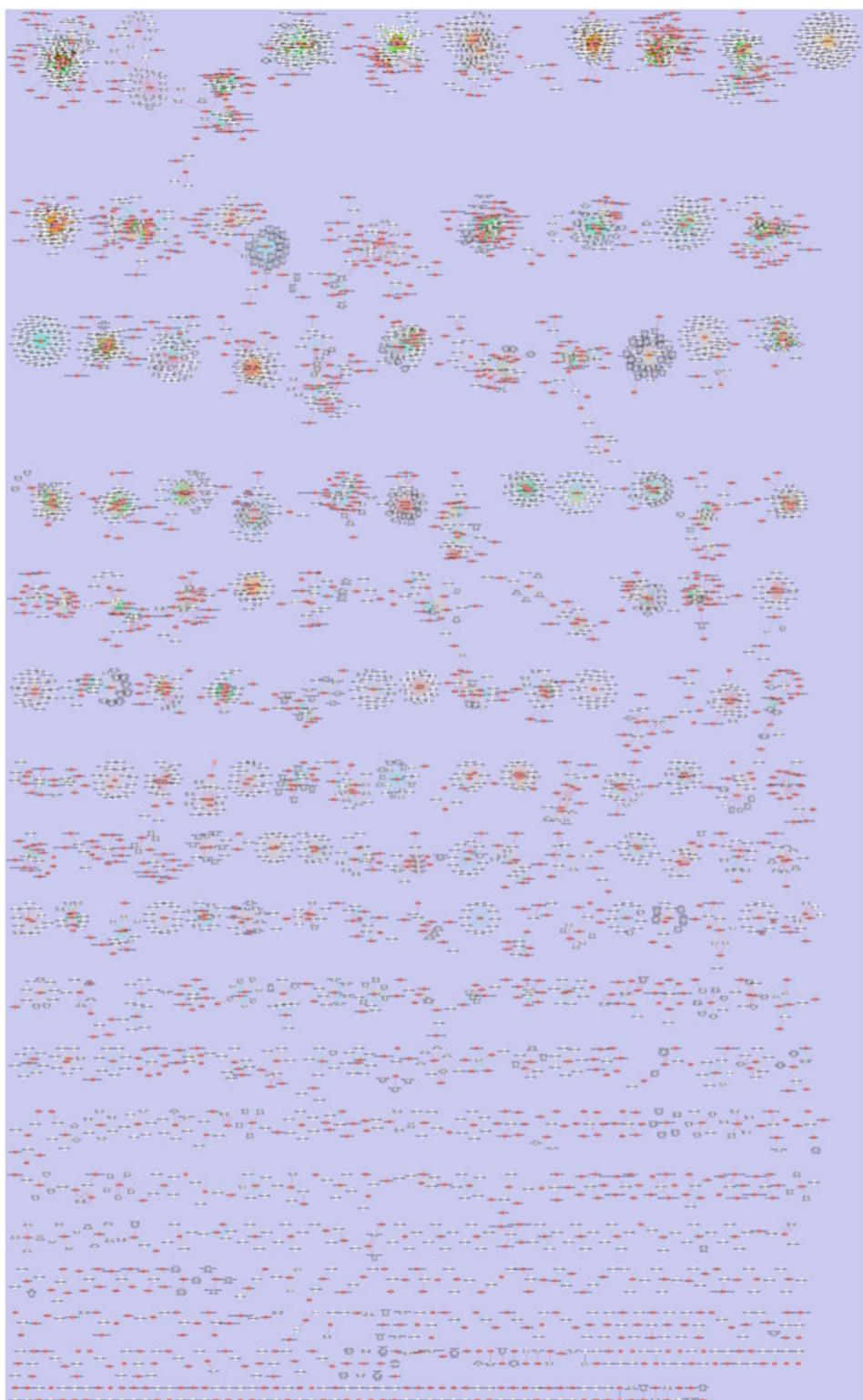


Fig. 2 Cytoscape view of the identified proteins in *Musa*. Nodes that represent a protein are visualized in *white*, peptides in *red*. Interactions with a ladder score between 0 and 50 are visualized in *red*, a score between 50 and 60 between in *orange* and >60 in *green* (color figure online)

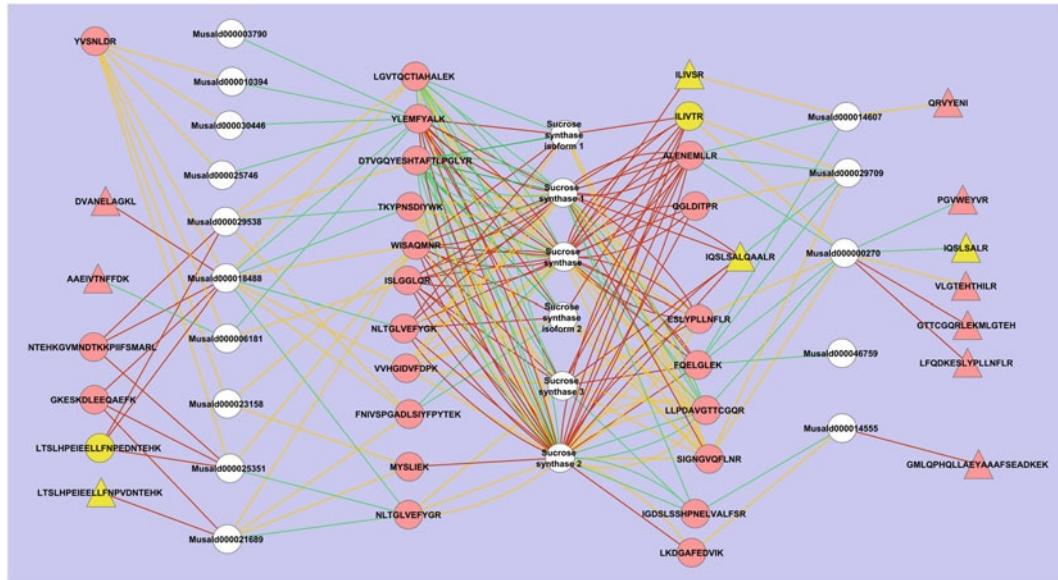


Fig. 3 Example from Fig. 2, of *Musa* sucrose synthase. Nodes that represent a protein are visualized in white, peptides in red. Interactions with a ladder score between 0 and 50 are visualized in red, a score between 50 and 60 between in orange and >60 in green. Fifteen *Musa* specific EST sequences have at least one peptide with a ladder score above 50. Those 15 EST sequences are all unique and have at least one amino acid difference. The clustering compared to known full sequences of sucrose synthase shows that none of the *Musa* EST sequences present in the database cover the whole protein range. The EST database is from a mixed origin (5'-end and 3'-end; different varieties), which is illustrated in the clustering. Peptide nodes unique for one sequence have a triangular shape. Homologous peptide nodes are colored in yellow. The Cytoscape view gives an impression about the protein inference and shows that we have experimental evidence for at least two different isoforms. Thirty percent of the peptides are not conserved and are unique for *Musa* (color figure online)

4.3 Peptide Separation, MS, and Identification

1. 2-D nanoAcuity UPLC system online coupled to a Synapt HDMS Q-TOF MS instrument (Waters, Milford, MA, USA).
2. First RP (bridge C18, 5 μ m) column (300 μ m \times 50 mm, Waters).
3. Second C18 (Symmetry 5 μ m) trap column (180 μ m \times 20 mm, Waters, Milford, MA, USA).
4. Syringe pump for lock spray.
5. Lock mass, [Glu1]fibrinopeptide B (1 pmol/L) (Waters, Milford, MA, USA).
6. Proteinlynx global server (Waters, Milford, MA, USA).
7. Cytoscape <http://www.cytoscape.org/> (free ware).
8. MS homology <http://prospector.ucsf.edu/prospector/mshome.htm>
9. Microsoft Excel 2010.

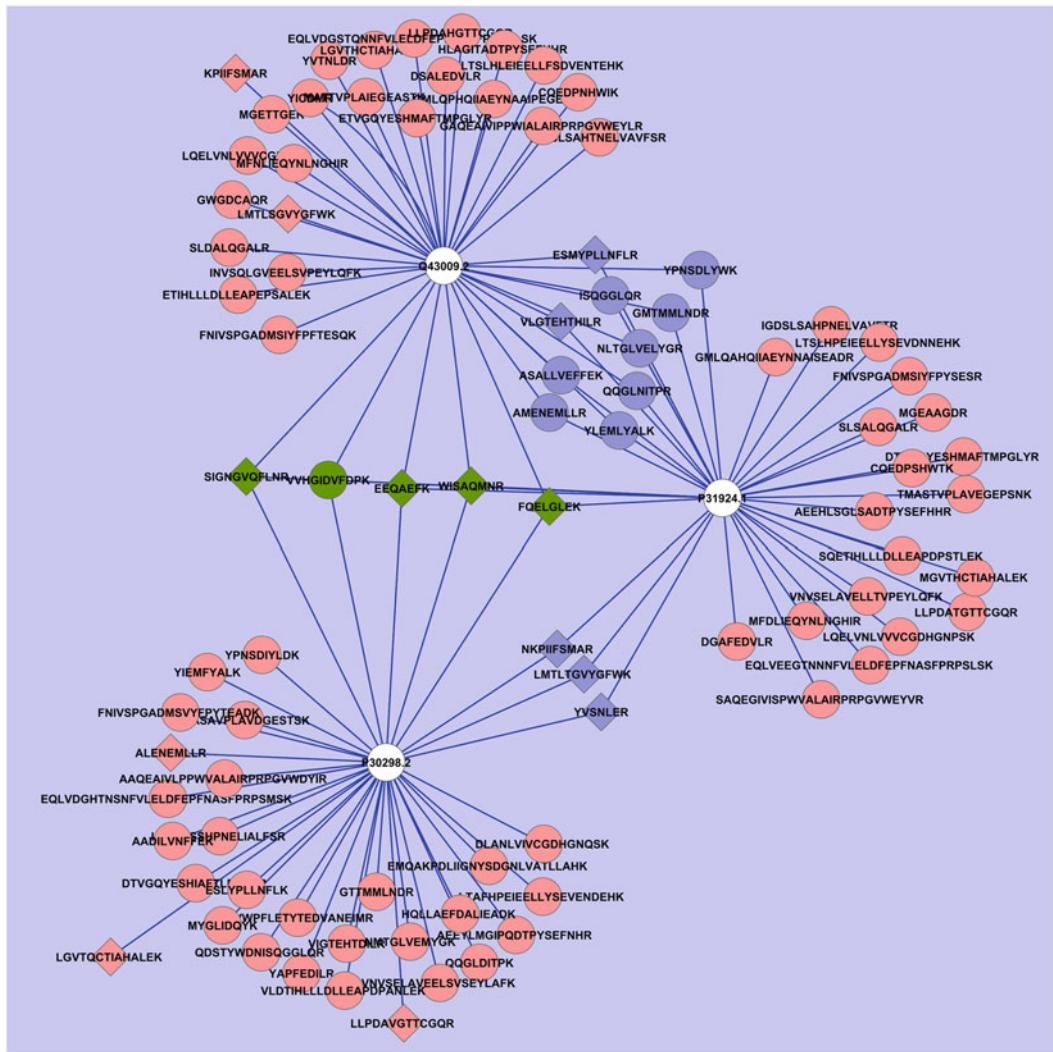


Fig. 4 Protein inference of sucrose synthase in rice. The theoretical tryptic peptides are visualized for the three known isoforms. Nodes that represent a protein are visualized in white, peptides that are tryptic specific in red, peptides that are common for two isoforms in blue and peptides common to three isoforms in green. In order to identify a specific isoform at least one tryptic specific peptides should be identified with sufficient confidence. Nodes that represent peptides that are conserved in *Musa* are represented as diamonds (color figure online)

5 Methods

5.1 Protein Extraction

1. Transfer fresh plant tissue to a liquid nitrogen precooled mortar and grind in liquid nitrogen (see Note 2).
2. Transfer 50–100 mg of frozen tissue powder to an extraction tube (2 ml), add 500 μ l extraction buffer and vortex 30 s (see Note 3).
3. Add 500 μ l of buffered phenol and vortex 10 min at 4 °C.
4. Centrifuge for 3 min, 6,000 $\times g$ at 4 °C, collect the upper phase, i.e., the phenolic phase, and transfer to a new tube and discard the lower phase.

5. Re-extract by adding 500 μ l of new extraction buffer.
6. Centrifuge for 3 min, 6,000 $\times g$ at 4 °C.
7. Transfer the phenolic phase into a new 2 ml tube and precipitate the proteins overnight (or at least 150 min) with 5 volumes 100 mM ammonium acetate in methanol at -20 °C.
8. Centrifuge 60 min, 16,000 $\times g$, at 4 °C.
9. Remove the supernatant and rinse the pellet twice (do not resuspend) in 2 ml rinsing solution (cold acetone/0.2 % DTT). After the first rinse leave in rinsing solution for 1 h at -20 °C. After rinsing, centrifuge 30 min, 16,000 $\times g$, at 4 °C.
10. Dry the pellet at RT.
11. Suspend the pellet in 100 μ l lysis solution (optimum concentration is 1–5 μ g/ μ l).
12. Clear samples by centrifugation (16,000 $\times g$, 2 $\times g$ 30 min at 18 °C).
13. Quantify the samples (*see Note 4*).

5.2 Protein Digestion and Cleanup

1. From the quantification analysis calculate the amount of μ l necessary to digest 50–60 μ g protein.
2. Add $x\mu$ l 0.5 M IAA to reach a final concentration of 0.015 M and incubate for 0.5 h in the dark at RT (shaker) (*see Note 5*).
3. Add $x\mu$ l 100 mM ABC to reach a final concentration of 75 mM (*see Note 5*).
4. Add 3 μ l trypsin (0.2 μ g/ μ l) and incubate at 37 °C (overnight).
5. Add $x\mu$ l of 10 % TFA to reach a final concentration of 0.5 % (*see Notes 5 and 6*).
6. Wash SPE columns with 1 ml 95 % ACN.
7. Equilibrate SPE columns with 1 ml 2 % ACN 0.1 % TFA.
8. Apply the sample to the column.
9. Wash SPE columns with 1 ml 2 % ACN 0.1 % TFA.
10. Elute the sample with 84 % ACN 0.1 % TFA.
11. Dry the sample in the speed vac.
12. Dissolve in 40 μ l 0.1 M AF.
13. Vortex 60 min.
14. Centrifuge 5 min 16,000 $\times g$.
15. Determine the peptide concentration through measurement at A280 nm (NanoDrop, Thermo Scientific, Waltham, MA, USA) (*see Note 4*).

5.3 Peptide Separation, MS and Identification

1. Inject 2.5 μ g of digested proteins on the first RP column.
2. Elute the sample from this column stepwise under high pH and ultrahigh pressure at 2 μ l/min. Samples are subsequently eluted in fractions of 12, 15, 18, 20, 25, 35, and 65 % ACN in

20 mM ammonium formate (pH 10). The Elution of the peptides takes place in an excess of 0.1 % FA in water at a flow rate of 20 μ l/min to reach a tenfold dilution before loading on the analytical column.

3. Elute the peptide from the analytical column at 0.2 μ l/min using 0.1 % formic acid in water as eluent A and 0.1 % formic acid in ACN as eluent B. The elution pattern is dependent on the nature of the eluted sample from the first column and is controlled by a curve setting. The curve determines the rate at which the solvent is to change to the new proportions and/or flow rates. Curves are specified by number with available choices from 1 to 11. Curve 1 immediately goes to specified conditions; curves 2–5 are convex; curve 6 is linear; curves 7–10 are concave and curve 11 maintains start condition until next step. All gradient slopes are linear (curve 6), except for the 10–40 % B step of the 12 % ACN fractions (concave curve 7) and for the fractions 35 and 65 % ACN (convex curve 5). Separation should be carried out using 5 % B for 1 min, 10 % B for 2 min, 10–40 % B over 62 min, and 40–85 % B over 9 min. After 6 min of rinsing with 85 % B and a linear gradient back to 5 % B over 2 min the column is re-equilibrated at initial conditions. The analytical column temperature should be maintained at 35 °C.
4. Perform the mass spectrometric analyses in positive mode using ESI with a NanoLockSpray source. Spray the eluates immediately into the Q-TOF device with lock mass at a flow rate of 0.2 μ l/min. Set the lock mass channel at 30 s. Select for MS/MS, the three most intensive multiply charged ions eluting from the column for fragmentation. Detect the eluting peptide ions in the MS survey scan (0.6 s) from *m/z* of 300 to 1,400. Set a dynamic exclusion window was at 60 s (*see Note 7*).
5. Search the obtained peak lists against the species specific database using Proteinlynx Global Server (PLGS 2.4, Waters) with the following settings peptide tolerance 20 ppm, fragment tolerance 0.05 Da and one tryptic miscleavage, Carbamidomethylation of C fixed modification, and oxidation of M as variable modification (*see Note 7*).
6. Perform an error tolerant search (“automod” query) on the spectra that did not result in an identification in the non-error tolerant query. Allow one nonspecific cleavage and one amino acid substitution or modification per peptide. Proteins identified with at least one peptide with a ladder score above 50 are considered significant.
7. Perform a de novo determination on the spectra that did not result in an identification “automod” query (*see Notes 8 and 9*).
8. Export all the peptide tables of the different 1D fractions and merge them into one Excel file 2010, i.e., original file.

9. Create a pivot table and filter the proteins with minimum one peptide with ladder score >50. Discard the rest of the proteins and create a new file, i.e., database file.
10. Import the new Excel file in Cytoscape, i.e., database network (Fig. 2).
11. Sort from the original file the data according to *m/z* (ascending) and ladder score (descending) and keep the best hit per *m/z*. Remove the redundancy in the peptides by filtering for unique *m/z*.
12. Filter all the peptides with a ladder score greater than 50 and discard peptides assigned to keratin or trypsin.
13. Submit the high quality peptide characterizations of mixed origin (database, automod, de novo) to MS homology taxonomy Green plants (freely available at the Web site of the University of California (<http://www.ucsf.edu/>)). Set the number of allowed amino acid substitutions for the MS Homology search as “the length of the peptide divided by 5” (see Note 10).
14. Save the results as a new Excel file, i.e., MShomology file.
15. Sort from the MShomology file the data according to sequence (ascending) and peptide score (descending) and protein score (descending), keep the best hit per sequence. Remove the redundancy in the peptides by filtering for unique sequences.
16. Import the new nonredundant Excel file in Cytoscape, i.e., MS homology network.
17. Merge both networks in Cytoscape to create a final network, i.e., union network.
18. Give a different layout to the nodes of peptides and proteins.
19. Give a different color to the different interactions between the peptides and proteins correlated to the confidence level of identification (ladder score, MS homology score) (see Figs. 2, 3 and 4).

6 Notes

1. Phenol is toxic and needs to be handled with great care. Wear gloves, work in a dedicated room. Use only under the hood and make sure that the waste (tips, pipettes, tubes) is disposed safely.
2. The grinding process is very important. Do not stop grinding until the power is very fine and homogeneous.
3. Depending on the protein concentration of the particular tissue, more or less starting material is needed.
4. For a quantitative peptide based approach, the quantification of the proteins and peptides is very important but challenging. Each quantification method has its limitations since the concentration

of a complicated protein mixture is estimated based on the quantification of one reference protein or peptide mixture.

5. The amount x is dependent on the volume that has been taken to have 50 μg of proteins in **step 1**.
6. Adding TFA will acidify the sample and will stop the enzymatic digestion.
7. Depending on the type of mass spectrometer (speed) those settings can be adjusted to select more or less peaks.
8. The sequences of the different isoforms of keratin and trypsin are present in our database to avoid that those peptides are selected to perform de novo analysis, where they could contribute to false positive identifications.
9. The sequence specific database should not contain duplicate sequences in order to avoid additional redundancy. The database can be made nonredundant via the program CD hit. <http://cd-hit.org> [16].
10. In this step the peptides belonging to different partial protein sequences of the EST database are compared with known protein sequences. EST's belonging to the same protein are clustered to a full protein sequence (Fig. 3). This step is for orphan species the only way to reconstruct the proteins based on the measured peptides.

Acknowledgment

Dr. Jan Cordewener is gratefully acknowledged for technical assistance.

References

1. Hedges SB (2002) The origin and evolution of model organisms. *Nat Rev Genet* 3:838–849
2. Soltis DE, Soltis PS (1999) Polyploidy: recurrent formation and genome evolution. *Trends Ecol Evol* 14:348–352
3. Carpentier SC, Panis B, Renaut J et al (2011) The use of 2D-electrophoresis and de novo sequencing to characterize inter- and intra-cultivar protein polymorphisms in an allopolyploid crop. *Phytochemistry* 72:1243–1250
4. Asano T, Tsudzuki T, Takahashi S et al (2004) Complete nucleotide sequence of the sugarcane (*Saccharum officinarum*) chloroplast genome: a comparative analysis of four monocot chloroplast genomes. *DNA Res* 11:93–99
5. Damerval C, Devienne D, Zivy M et al (1986) Technical improvements in two-dimensional electrophoresis increase the level of genetic-variation detected in wheat-seedling proteins. *Electrophoresis* 7:52–54
6. Carpentier SC, Witters E, Laukens K et al (2005) Preparation of protein extracts from recalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis. *Proteomics* 5:2497–2507
7. Saravanan RS, Rose JKC (2004) A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics* 4:2522–2532
8. Vincent D, Wheatley MD, Cramer GR (2006) Optimization of protein extraction and solubilization for mature grape berry clusters. *Electrophoresis* 27:1853–1865
9. Carpentier SC, Panis B, Vertommen A et al (2008) Proteome analysis of non-model plants:

- a challenging but powerful approach. *Mass Spectrom Rev* 27:354–377
- 10. Carpentier SC, Swennen R, Panis B (2009) Plant protein sample preparation for 2DE. In: Walker JM (ed) *The protein protocols handbook*. Humana Press, Totowa, pp 107–117
 - 11. Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19:242–247
 - 12. Tabb DL, Smith LL, Breci LA et al (2003) Statistical characterization of ion trap tandem mass spectra from doubly charged tryptic peptides. *Anal Chem* 75:1155–1163
 - 13. Cramer R, Corless S (2001) The nature of collision-induced dissociation processes of doubly protonated peptides: comparative study for the future use of matrix-assisted laser desorption/ionization on a hybrid quadrupole time-of-flight mass spectrometer in proteomics. *Rapid Commun Mass Spectrom* 15:2058–2066
 - 14. Vertommen A, Panis B, Swennen R et al (2011) Challenges and solutions for the identification of membrane proteins in non-model plants. *J Proteomics* 74:1165–1181
 - 15. Gilar M, Olivova P, Daly AE et al (2005) Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. *J Sep Sci* 28:1694–1703
 - 16. Li WZ, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22:658–659

Chapter 25

An Improved Detergent-Compatible Gel-Fractionation LC-LTQ-Orbitrap-MS Workflow for Plant and Microbial Proteomics

Luis Valledor and Wolfram Weckwerth

Abstract

In proteomics, liquid chromatography coupled to mass spectrometry (LC-MS/MS) is an invaluable technique to accurately identify and quantify large sets of proteins. In this chapter we show a time-effective, and detergent compatible, Gel-LC-LTQ-Orbitrap/MS proteomics workflow. The compatibility of this protocol with high concentrations of detergents significantly increases the extraction yield and the abundance of membrane proteins while gel fractionation increases the number of protein identifications. In our hands this workflow allows the identification of more than 1,500 proteins per sample, harvesting less than 20 mg of fresh weight material, in many different organisms such as *Chlamydomonas*, *Cyanothecce*, *Arabidopsis*, or *Nicotiana*, various microbes and enriched microbial samples.

Key words Green proteomics, LC-MS, Orbitrap, Detergent-compatible, Membrane proteins, Plant Systems Biology

1 Introduction

Nowadays LC-MS/MS provides the highest resolution for the identification of peptides and the mapping of the proteomes [1]. Proteins are extracted and fractionated using standard protocols, which have been adapted to be compatible with MS (reduced concentration of detergents and buffers), and then digested to tryptic peptides by the use of endopeptidases. These tryptic peptides are first separated in a nanoHPLC system, generally coupled to a C18 column, and then identified using tandem mass spectrometry [2]. This methodology give us the possibility of identify and accurately quantify a huge number of proteins per run.

Even with the use of the next generation of mass spectrometers, with an increased resolution, sensitivity, dynamic range, and speed, the identification of all the ions co-eluting from the column remains one of major bottlenecks. The pre-fractionation of

the proteins before digestion is a commonly used strategy for reducing the complexity of the sample, increasing the total number of analyzed peptides after the combination of the different fractions which usually lead to the identification of large number of proteins [3].

In laboratory practice, the use of mass spectrometry means that all of the work should be done in a very clean environment, and mass-spectrometry compatible buffers and reagents should be used during protein extraction, fractionation and digestion. One of the classical drawbacks of using standard protocols is the poor identification rate of membrane and cell-wall proteins. The solubilization of these proteins requires a high concentration of detergents, not only during the initial steps of the extraction but also for the solubilization of the purified protein pellet. There are systems for removing detergents such as SDS from the sample; they are expensive and not effective in all of the situations. Another problematic step is the protein fractionation. Despite the fact that fractionation protocols are not very complex (classically done by column chromatography, i.e., FPLC), they are time consuming and require large amounts of proteins.

In this chapter we show a complete LC-MS/MS proteomics workflow, fully compatible with high concentrations of SDS. We implemented a rapid standard SDS-PAGE step that is used for protein pre-fractionation, washing (removal of detergents and other contaminants), and a high performance in-gel digestion of proteins. This workflow is time effective and can be completed in less than 3 days (day 1, protein extraction and quantitation; day 2, protein fractionation and digestion; day 3, peptide desalting and LC-MS/MS analysis). Peptides are resolved and identified using an nano HPLC system coupled to a LTQ-Orbitrap mass spectrometer. In our hand this protocol leads to a reproducible identification of more than 1,500 proteins starting from 50 µg of total protein in many different biological systems from microbes to plants.

2 Materials

Prepare all solutions using ultrapure water (bidistilled, deionized) and analytical/ultra HPLC grade reagents.

2.1 *Buffers for Protein Extraction and Quantitation*

1. Extraction buffer: 100 mM Tris-HCl, pH 8.0, 10 % (v/v) Glycerol, 2 mM PMSF, 10 mM DTT, 1.2 % (v/v) Plant protease inhibitor cocktail (Sigma P9599). Prepare fresh.
2. Buffer RS: 8 M Urea, 4 (w/v) % SDS. Buffer RS can be stored for 1 year at room temperature.
3. BCA Reagent A: sodium bicinchoninate (0.1 g), $\text{Na}_2\text{CO}_3 \cdot \text{H}_2\text{O}$ (2.0 g), sodium tartrate (dihydrate) (0.16 g), NaOH (0.4 g),

NaHCO₃ (0.95 g), made up to 100 mL. If necessary, adjust the pH to 11.25 with NaHCO₃ or NaOH.

4. BCA Reagent B: CuSO₄·5H₂O (0.4 g) in 10 mL of water. Reagents A and B are stable at room temperature.

2.2 SDS-PAGE

1. Running Buffer (TGS): For a 10x stock mix 30.2 g Tris-Base, 144 g of glycine, and 1 g of SDS; add _{dd}H₂O up to 1 L.
2. Resolving gel. Amount required for 2 Mini-Protean (Bio-Rad) gels: mix 3 mL of 40 % (w/v) acrylamide/bisacrylamide (37.5:1), 2.5 mL 1.5 M Tris-HCl, pH 8.8, 50 μ L of 20 % (w/v) SDS, and 4.5 mL of H₂O. Mix well and degasify in vacuo. Add 50 μ L of 10 % (w/v) APS and 5 μ L of TEMED for starting the polymerization. Mix slowly by inversion before casting the gel. Add 5 mL of acrylamide mix to each 7 cm gel cassette. Avoid bubbles during casting and quickly cover the acrylamide with 2-propanol. Let the gels polymerize for 1 h. If larger gels are used, increase the volumes accordingly.
3. Stacking gel. Amount required for 2 Mini-Protean (Bio-Rad) gels: mix 0.7 mL of 40 % (w/v) acrylamide/bisacrylamide (37.5:1), 2.5 mL 0.5 M Tris-HCl, pH 6.8, 50 μ L of 20 % (w/v) SDS and 2.25 mL of H₂O. Mix well and degasify in vacuo. Add 50 μ L of 10 % (w/v) APS and 5 μ L of TEMED for starting the polymerization. Before adding the APS and TEMED discard the 2-propanol layer covering the resolving gels and briefly rinse with water. Then pour the stacking solution containing APS and TEMED, and carefully place the comb avoiding bubbles.
4. Coomassie staining solution: 40 % (v/v) methanol, 10 % (v/v) acetic acid, 0.1 % (w/v) Coomassie R-250, in distilled water. Dissolve the Coomassie in methanol and then add the other components.
5. Coomassie destaining solution: 40 % (v/v) methanol, 2 % (v/v) acetic acid in distilled water.
6. 5 \times Laemmli buffer: 20 % (w/v) SDS, 80 % (v/v) Glycerol, 20 % (v/v) β -Mercaptoethanol, 0.01 % (w/v) bromophenol blue. Store at -20 °C. Before using make sure that SDS is completely dissolved. Heating may be necessary [4].

2.3 Protein Digestion and Desalting

1. Trypsin Buffer: 25 mM NH₄HCO₃ (AmBic), 10 % (v/v) acetonitrile, 5 mM CaCl₂. For preparing 1 mL mix 250 μ L of 0.1 M AmBic, 100 μ L of pure acetonitrile and 645 μ L of H₂O. Add 5 μ L of 1 M CaCl₂, Calcium must be added in last place to avoid precipitation. Prepare fresh.
2. Trypsin Solution: dilute Trypsin Sequencing Grade (Roche 11 418 475 001) to 10–12.5 ng/ μ L. Add 1 mL of trypsin buffer to the vial, incubate on ice for 10 min and then pipette up and

down to completely resuspend the pellet. Transfer to a 2 mL low binding tube. Add 1 mL of trypsin buffer to the vial, washing all the walls. Then transfer to the previous tube. Trypsin can be stored at $-20\text{ }^{\circ}\text{C}$ for 1 month. It is recommended to make aliquots to avoid melt/freezing cycles that can damage the enzyme. Aliquots should be stored in low binding tubes.

3. Peptide Solubilization Buffer (PS): 4 % (v/v) acetonitrile, 0.25 % (v/v) formic acid. If properly closed this buffer is stable for 3 months at room temperature.

3 Methods

3.1 Protein Extraction

1. Plant and cell samples (10–50 mg of fresh weight) can be processed using the volumes indicated in this protocol. For bigger amount of starting material, the proportions and volumes must be adjusted properly. If not processed immediately after sampling, store at $-80\text{ }^{\circ}\text{C}$ until analysis.
2. Add 300 μL of extraction buffer, and pipette up and down until the sample pellet is completely resuspended. Transfer to a screw-cap 2 mL tube previously filled with 25 mg of quartz sand and homogenize in a regimill/fastprep for 1 min at maximum speed. Non pelleted samples can be first transferred to the screw cap tube, and then add the extraction buffer. Add 100 μL of 20 % SDS to the sample tube, mix gently by inversion. Incubate 4–5 min in a shaker at $95\text{ }^{\circ}\text{C}$. Cool down at room temperature (*see Notes 1 and 2*).
3. Centrifuge 5 min at $4\text{ }^{\circ}\text{C}$ in a bench top centrifuge at full speed to pellet all insoluble materials. Transfer supernatant to a new 2 mL tube (for better yields it is recommended the use low binding tubes in all of the steps of this protocol).
4. Add at least 4–5 volumes of cold ($-20\text{ }^{\circ}\text{C}$) acetone with 0.5 % of β -Mercaptoethanol. Mix it gently by inversion. Keep a minimum of 2 h at $-20\text{ }^{\circ}\text{C}$.
5. Centrifuge 15 min at $5,000\times g$ and $4\text{ }^{\circ}\text{C}$. Carefully discard the supernatant by pipetting it out.
6. Wash the pellet with 1 mL of acetone (0.5 % β -Mercaptoethanol). Pipette up and down or use a soft ultrasound bath to disaggregate the pellet. Centrifuge 5 min at $5,000\times g$ and $4\text{ }^{\circ}\text{C}$ to pellet down the proteins. Use the micropipette for removing the supernatant, be careful not disturbing the pellet.
7. Wash the pellet with 1.2 mL of acetone. Pipette up and down or use a soft ultrasound bath to disaggregate the pellet. Centrifuge 5 min at $5,000\times g$ and $4\text{ }^{\circ}\text{C}$ to pellet down the proteins. Use the micropipette for removing the supernatant, be careful not disturbing the pellet.

8. Wash the pellet in 1.2 mL of 90 % acetone (diluted with distilled water). Pipette up and down to disaggregate the pellet. Centrifuge 5 min at $5,000 \times g$ and 4 °C to pellet down the proteins. Discard supernatant.
9. Air-dry the pellets until the acetone is completely eliminated. Pellets may retain some water. This will help the pellet solubilisation (*step 9*).
10. Resuspend pellet in RS buffer. A complete pellet solubilisation may take up to 3 h in a thermal shaker (35 °C and 750 rpm), being also some occasional pipetting needed (*see Note 3*).
11. Centrifuge a maximum speed and room temperature for 5 min to remove insoluble particles. Transfer the supernatant to a new tube (*see Note 4*).

3.2 Protein Quantitation

1. Prepare Standard Working Reagent (SWR): Mix 100 vol. of reagent A with 2 vol. of reagent B. The solution is stable at room temperature for 1 week.
2. Add BSA standards to an ELISA plate. 0, 1, 2, 3, 4, 6, 8, 10, 15 and 20 µg of BSA are using for the calibration curve (three technical triplicates are advised).
3. Add 2 µL sample to each well.
4. Add 200 µL of SWR to each well. Shake the plate gently.
5. Close the plate with a cap or aluminum foil to minimize evaporation. Incubate the plate 20 min at 60 °C (in an oven) or 40 min at 45 °C in the plate reader. Read absorbance at 562 nm. Use the spectrophotometer bundled software or spreadsheet for calculating concentrations (*see Note 5*).

3.3 Protein Fractionation by SDS-PAGE

1. Mix 40–80 µg of protein (minimum recommended 50 µg), 5 µL of 5× Laemli buffer, and water to a final volume of 25 µL (this may vary in function of the well size/electrophoresis system).
2. Perform a standard SDS page fractionation in a 12.5 % acrylamide gel (maximum thickness 1 mm). Before loading the protein, mark a line 1 cm below the stacking-resolving interface. Load the samples leaving a blank well between them to avoid contaminations (*see Note 6*) (Fig. 1).
3. Run the gel at 80 V, constantly, until the bromophenol blue reach the marked line.
4. Stop the electrophoresis and immediately transfer the gel to a plate with Coomassie staining solution. Incubate in an orbital shaker for 30 min (recovery of the tryptic peptides is dramatically reduced in overfixed gels).
5. Destain the gels in destaining solution for 80 min, replace destaining solution every 20 min.

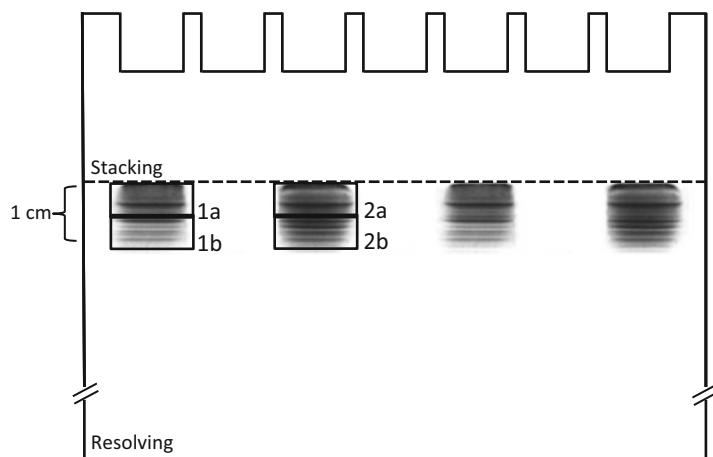


Fig. 1 Schematic representation of a gel, in which loading, staining, and band cutting for a high-performance proteomics analysis are represented. The band pattern corresponds to *Chlamydomonas reinhardtii* cells grown in mixotrophic media

6. Wash the gel in distilled water for 30 min.
7. Cut the lanes into two gel pieces, aiming to divide the total intensity of the band equally between them. Cut all of the lanes in the same way. We recommend choosing one clear band as a cutting reference to make the divisions of all of the lanes of the experiment. Transfer gel pieces to individual 1.5 mL tubes. At this point the gel pieces can be long term stored at -20 °C. Before freezing cover the pieces with distilled water.

3.4 Protein Digestion, Peptide Desalting and Concentration

3.4.1 Protein Digestion

This procedure is an adaption of [5] with increased effort in gel washing and digestion. Reduction and alkylation of proteins is not needed using the protocol described above. Double digestion (LysC-Trypsin) do not improve the final results.

1. Chop the gel pieces with a scalpel over a glass plate. The final size of acrylamide pieces should be around 1 mm³. Glass plate should be carefully cleaned after each gel piece. Avoid taking too much gel without proteins.
2. Transfer the gel pieces to a 1.5 mL low binding tube and keep them on water until all samples are processed.
3. Remove water and add 1 mL of 25 mM AmBic, incubate 15 min at 37 °C in a thermal shaker. Discard supernatant.
4. Remove water and add 1 mL of 25 mM AmBic/50 % acetonitrile, incubate 15 min at 37 °C in a thermal shaker. Discard supernatant.

5. Repeat **step 4**. At this step the supernatant must be almost transparent or with a very light blue color, if the blue color is intense, repeat this step (*see Note 7*).
6. Dehydrate gel pieces with 300 μ L 100 % Acetonitrile (AcN) incubate 5 min at room temp in a shaker. Discard supernatant. At this point gel pieces should be very reduced in volume, hard, and completely white, having a “plastic” look. If they are still translucent or soft (you can use the pipette tip to test it) repeat this step.
7. Dry out in a SpeedVac for 5 min.
8. Add 50 μ L of Trypsin Solution to each tube and incubate 15 min at 37 °C. If needed, add more trypsin solution until the gel pieces are completely rehydrated. Wait another 15 min. Completely cover the gel pieces with trypsin buffer.
9. Incubate 14–16 h at 37 °C. Incubation in an oven is recommended, if a thermal block is used, cover it with aluminum foil to maintain the temperature also in the caps of the tubes. Shaking during the digestion is not advised (*see Note 8*).

3.4.2 Peptide Extraction

1. Add 150 μ L of 50 % ACN/1 % formic acid (FA) to each tube with gel pieces and incubate 5 min at room temperature. After incubation sonicate 3 min in a low intensity ultrasound bath.
2. Spin the tubes and transfer the supernatant to a new tube (“A”).
3. Repeat 1. Transfer the supernatant to tube A.
4. Add 100 μ L of 90 % ACN/1 % formic acid (FA) and incubate for 5 min at room temperature. Transfer the supernatant to the A tube. The gel pieces should look in the same way than in **step 6** of protein digestion. If not, repeat this step once more.
5. Dryout in SpeedVac. Keep at –20 °C or proceed with desalting. For long term storage –80 °C is recommended.

3.4.3 Peptide Desalting

1. Add 75 μ L of PS Buffer to peptide pellets. Incubate for 10 min on ice.
2. Resuspend the pellet by pipetting and sonicate 3 min in a low intensity ultrasound bath. Incubate 5 min at room temperature and then keep on ice.
3. Use C18 stage tips or 96-well plates for peptide desalting. Independently of the chosen support the protocol remains unchanged (only volumes should be adjusted). All of the steps are performed at room temperature. We routinely use 96-well plates (Spec 96-Well C18, Agilent) as follows:
 - (a) Plate activation and washing: Add 700 μ L MeOH to each well (repeat once) (*see Note 9*).
 - (b) Plate washing and equilibration: Add 700 μ L $_{dd}$ H₂O to each well (repeat once).

- (c) Sample binding: Pipette solution in the center of the well all of the volume in the tubes, including all of the insoluble parts that may remain in the solution. Confirm visually that the membrane absorbs the peptides. Incubate for 3 min at room temperature.
- (d) Sample desalting: Add 400 μ L of $_{\text{dd}}$ H₂O to each well (repeat five times)
- (e) Sample recovery: Replace the under well tray with a clean one and elute peptides by adding 200 μ L of MeOH (repeat once). Transfer the MeOH-peptides to low binding tubes and evaporate completely in speedvac.
- (f) Store the tubes at -20 °C until LC-MS/MS analysis

3.5 Protein Identification and Quantitation

3.5.1 Peptide Preparation

1. Resuspend peptide pellets in 15 μ L of PS buffer. The amount of buffer should be proportional to the initial amount of peptides (*see Note 10*).
2. Incubate for 10 min on ice. Resuspend the pellet by pipetting and sonicate for 3 min in a low intensity ultrasound bath. Incubate 5 min at room temperature.
3. Spin down at $>20,000 \times g$ for 10 min and at 4 °C to remove insolubles.
4. Transfer supernatant to a LC microvial, take special care in not disturbing the pellet or taking any clump.

3.5.2 LC-MS/MS

1. LC settings: 10 μ g of digested peptides in 5 μ L were loaded per injection onto a one-dimensional (1D) nano-flow LC-MS/MS system (Eksigent, Germany) equipped with an in line pre-microfilter (Scivex, USA). Peptides were eluted using a monolithic C18 column Chromolith RP-18r (Merck, Germany) of 15 cm length and 0.1 mm internal diameter during a 90 min gradient from 5 to 40 % B with a controlled flow rate of 400 nL per minute. LC was coupled to MS using a ESI source. Mobile phase A: 0.1 % Formic Acid; Mobile phase B: 90 % Acetonitrile, 0.1 % Formic Acid (*see Note 11*).
2. MS settings: MS analysis was performed on an Orbitrap LTQ XL mass spectrometer (Thermo, Germany). Specific tune settings for the MS were set as follows: spray voltage was set to 1.8 kV using a 30 μ m inner diameter needle (PicoTip Emitter; NewObjective, USA); temperature of the heated transfer capillary was set to 180 °C. FTMS was operated as follows: fullscan mode, centroid, resolution of 30,000, covering the range 300–1,800 m/z , and Cyclomethicone was used as lock mass. Each full MS scan was followed by ten dependent MS/MS scans, in which the ten most abundant peptide molecular ions were dynamically selected, with a dynamic exclusion window set to 90 s and exclusion list set to 500. Dependent fragmentations were performed in CID mode, with a normalized collision

energy of 35, iso width of 1.0, activation Q of 0.250 and activation time of 30 ms. Ions with unassigned charge or +1 were excluded for fragmentation. The minimum signal threshold was set to 750.

3. Protein identification: Raw-files are processed in Proteome Discoverer software (Thermo, Germany) using in house protein databases containing the latest available protein sequences and six-frame translations using the SEQUEST algorithm (as is available in Proteome Discoverer 1.3, Thermo, USA). The following settings were used: precursor mass tolerance was set to 10 ppm and fragment ion mass tolerance to 0.8 Da. Only charge states +2 or greater were used. Identification confidence was set to a 5 % FDR and the variable modifications were set to: acetylation of N terminus, oxidation of methionine, and carbamidomethyl cysteine formation. No fixed modifications were set. A maximum of two missed cleavages were set for all searches [6]. For fully sequenced organisms, like Chlamydomonas or Arabidopsis we establish as a threshold for protein identification one unique peptide (peptide that only appears once in the entire database) with a X-Correlation value 0.5 greater than the charge state (i.e., 2.5 for peptides with charge +2). For non-sequenced organisms the use of two peptides are advised.
4. Quantification: identified proteins are quantified by a standard peptide count measurement using a NSAF approach [7]. This measurement is limited to peptides that have been assigned to proteins, and not those which are not present in the database or those with posttranslational modifications (PTM's) not defined in the SEQUEST search step. A complete unbiased approach for quantification independent from database search is the MAPA (mass accuracy precursor alignment) approach [8, 9]. Here the identification relies on feature selection and abundance. Using this method posttranslational modifications (PTM's) and non-sequenced organisms can be analyzed as well [8, 10]. Independently of the employed approach, a power multivariate statistical analysis is paramount [11] (see Chapter 5). We routinely performed these analyses using R statistical environment, but recently we have developed COVAIN, a tool that allows a deep statistical analysis and data pre-processing in a user-friendly environment [12].

4 Notes

1. The indicated protocol is intended for microalgae and bacteria, for strains or plant organs with thicker cell wall the homogenization step probably should be stronger, or alternative homogenization methods like French press or mortar and pestle in liquid nitrogen used.

2. If required a phenol purification can be performed after this step.
 - (a) Add 3 volumes of sucrose buffer (0.9 M Sucrose, 0.1 M Tris-HCl pH 8, 10 mM DTT, 1 % Protease inhibitor cocktail) and mix vigorously by vortex.
 - (b) Add 300 μ L of phenol (equilibrated at pH 8) and vortex. Incubate 10 min at room temperature in a shaker.
 - (c) Centrifuge 5 min $>10,000 \times g$ at room temperature. After centrifugation two clear layers with a sharp interphase should be seen. Transfer supernatant (phenol, should be green) to a new 2 mL tube. If there is some turbidity add 200 μ L of 2:1 of sucrose buffer:phenol, vortex and centrifuge again.
 - (d) Re-extract the aqueous phase by adding 300 μ L of phenol to the original. Vortex thoroughly and centrifuge 5 min $>10,000 \times g$ at room temperature. Transfer supernatant to the tube with the previous phenolic phases.
 - (e) Optional clean-up of the phenol phase: adding 1 volume sucrose buffer, vortex, and centrifuge. Transfer upper (phenolic) phase to a new tube. If a thick interphase is visible, repeat this step once.
 - (f) Precipitate the proteins by adding 2 volumes of 0.1 M ammonium acetate in methanol. Incubate overnight at -20°C .
 - (g) Continue the protocol in **step 5**.
3. The resolubilization of the pellet is a critical step since the proteins should be in an adequate concentration. Here we recommend starting with the addition of a low volume of buffer, 20–30 μ L. If the buffer become very viscous (you can invert the tube and nothing is coming down), or there is some undissolved proteins more buffer can be added. Protein solubilisation takes time so we recommend wait at least 30 min before starting pipetting or adding more buffer. Always try to dissolve the protein clumps by pipetting before adding more buffer. Use adequate pipette tips to avoid losses.
4. The pellet may vary in size because it contains some carryover of contaminants that are not dissolved in the protein resolubilization step so a big pellet is not a sign of poor resolubilization or yield. Samples washed with phenol generally do not generate a noticeable pellet in this step.
5. A good protein quantitation is needed because this protocol is very sensitive to protein amount. We regularly use BCA method because it is robust and also compatible with detergents.
6. We routinely fractionate the proteins into two fractions, but for a higher resolution three fractions can be picked. In this

case load 90 µg of protein per well and run the gels 1.5 cm. It is important using gels of 1 mm thick (or less). Thicker gels present problems for destaining and digestion. Leave one blank well between samples to avoid contaminations. Fill these wells with 1× Laemmli buffer.

7. Band destaining and cleaning is paramount for a good LC-MS analysis. In these steps not only Coomassie, but also SDS and other contaminants are removed. If there is still blue color inside the gel pieces after three washings, continue with the protocol, it will not interfere with protein digestion. This usually happens when the gel pieces are not small enough.
8. Incubating the samples below 4 °C can stop trypsin digestion, but better results are obtained when peptide extraction is done just afterwards. We do not recommend spending more than a couple of hours at this temperature or freezing the samples for a later processing.
9. When using the 96-well plates in a vacuum manifold system a special care should be taking regarding to the vacuum. The vacuum should be the minimum to dry the membranes. The vacuum should not be released sharply, because it can damage the membranes and the plate. Bottom part of the wells should not contact directly the waste tray at any moment. Collection trays can hold up to 800 µL before the flow-through reaches the bottom of the wells, so waste tray should be emptied after each equilibration/washing step.
10. In our experience the best results are obtained when the desalted peptides are resuspended in a volume that is half of the loaded micrograms. In our case we load 60 µg, which are divided into two gel pieces containing approximately 30 µg each. Our setup injects 5 µL per run. So our recommended resuspension volume is 15 µL. If the amount of loaded proteins or the HPLC injection volume varies, this resuspension volume should change accordingly.
11. The specific gradient we use is the following:

| Time (min) | % A (0.1 % FA) | % B (90 % ACN, 0.1 % FA) |
|------------|----------------|--------------------------|
| 0 | 95 | 5 |
| 5 | 90 | 10 |
| 90 | 60 | 40 |
| 100 | 15 | 85 |
| 105 | 15 | 85 |
| 107 | 95 | 5 |
| 120 | 95 | 5 |

Acknowledgment

L.V.'s work was generously supported by a Marie Curie IEF Grant (European Union).

References

1. Weckwerth W (2011) Green systems biology, from single genomes, proteomes and metabolomes to ecosystems research and biotechnology. *J Proteomics* 75:284–305
2. Cottrell JS (2011) Protein identification using MS/MS data. *J Proteomics* 74:1842–1851
3. Ly L, Wasinger VC (2011) Protein and peptide fractionation, enrichment and depletion: tools for the complex proteome. *Proteomics* 11:513–534
4. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685
5. Shevchenko A, Tomas H, Havlis J et al (2007) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 1:2856–2860
6. Valledor L, Recuenco-Munoz L, Wienkoop S et al (2012) The different proteomes of *Chlamydomonas reinhardtii*. *J Proteomics* 75(18):5883–5887
7. Paoletti AC, Parmely TJ, Tomomori-Sato C et al (2006) Quantitative proteomic analysis of distinct mammalian mediator complexes using normalized spectral abundance factors. *Proc Natl Acad Sci USA* 103:18928–18933
8. Hohenwarter W, van Dongen JT, Wienkoop S, Steinfath M, Hummel J, Erban A et al (2008) A rapid approach for phenotype-screening and database independent detection of cSNP/protein polymorphism using mass accuracy precursor alignment. *Proteomics* 8:4214–4225
9. Egelhofer V, Wolfgang H, Lyon D, Weckwerth W, Wienkoop S (2013) Using ProtMAX to create high mass accuracy precursor alignments from label-free quantitative mass spectrometry data generated in large shotgun proteomics experiments. *Nat Protoc* 8(3):595–601
10. Hoehenwarter W, Larhlimi A, Hummel J, Egelhofer V, Selbig J, van Dongen JT et al (2011) MAPA distinguishes genotype-specific variability of highly similar regulatory protein isoforms in potato tuber. *J Proteome Res* 10:2979–2991
11. Valledor L, Jorrin J (2011) Back to the basics: maximizing the information obtained by quantitative two dimensional gel electrophoresis analyses by an appropriate experimental design and statistical analyses. *J Proteomics* 74:1–18
12. Sun X, Weckwerth W (2012) COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics* 8:81–93

Part IV

Descriptive Proteomics

Chapter 26

Seed Proteomics

Ján A. Mierny

Abstract

Rather than providing a single specific protocol, the inclusive area of seed proteomics is reviewed; methods are described and compared and primary literature citations are provided. The limitations and challenges of proteomics as an approach to study seed biology are emphasized. The proteomic analysis of seeds encounters some specific problems that do not impinge on analyses of other plant cells, tissues, or organs. There are anatomic considerations. Seeds comprise the seed coat, the storage organ(s), and the embryonic axis. Are these to be studied individually or as a composite? The physiological status of the seeds must be considered; developing, mature, or germinating? If mature, are they quiescent or dormant? If mature and quiescent, then orthodox or recalcitrant? The genetic uniformity of the population of seeds being compared must be considered. Finally, seeds are protein-rich and the extreme abundance of the storage proteins results in a study-subject with a dynamic range that spans several orders of magnitude. This represents a problem that must be dealt with if the study involves analysis of proteins that are of “normal” to low abundance. Several different methods of prefractionation are described and the results compared.

Key words Seed proteomic, Seeds, Seed storage proteins

Abbreviations

| | |
|------|------------------------------------|
| DIGE | Difference in gel electrophoresis |
| LC | Liquid chromatography |
| MS | Mass spectrometry |
| PAGE | Polyacrylamide gel electrophoresis |
| SSP | Seed storage proteins |

1 Seeds

The formation of an embryonic plant, the product of the ripened ovule, completes the reproductive cycle of seed plants. In angiosperm plants the seeds are fruit-enclosed while in gymnosperms

“*Seeds of every generation between our hands. And the promise to teach you the little I have learned so far*”—
Brooke Fraser.

the seeds are naked. Seeds comprise an outer covering (the seed coat, which developed from the integuments of the ovule), the endosperm, cotyledons, and embryonic axes [1, 2]. The reserve polymers necessary for the support of postgerminative seedling growth are contained either within the endosperm (endosperm-dominant) or cotyledons (cotyledon-dominant) of the seeds [3, 4]. In cotyledon-dominant angiosperm seeds, such as those from legumes, the endosperm at maturity can be as thin as 2–3 cells, or be absent entirely. Conversely, the endosperm is the more prominent organ in grains, and the cotyledons are relatively small. Embryos of cotyledon-dominant seeds generally have two cotyledons (dicot) while those of endosperm-dominant seeds have a single cotyledon (monocot) or cotyledon-like (e.g., scutellum) storage organ.

There is an enormous range in the size of seeds, from the dust-like seeds of orchids (e.g., *Gomesa crispula*, 810 ng dry weight/seed) to the formidable seeds of Coco de Mer (*Lodoicea maldivica*), samples of which have been reported to exceed 18 kg dry weight! The more familiar seeds of soybean (*Glycine max* (L.) Merr, cv Jack) and “mouse-eared cress” (*Arabidopsis thaliana* L. ecotype Columbia) weigh in at 225 mg dry weight and 30 µg dry weight, respectively [5]. Size is one of many criteria that can impinge on the choice of a research subject by the seed biologist. While I am unaware of specific supporting data, it seems likely that the smaller seeds have the greater biological variation.

One seldom sees reference to, e.g., “leaf proteomics,” or “stem proteomics.” Are seeds then actually deserving of their own unique sub-discipline? Well ... yes! Any one of multiple seed-specific characteristics is adequate to justify unique treatment, and together they virtually require it. For any sort of meaningful comparisons, sample uniformity is an important issue. In addition to anatomical considerations, any meaningful analysis of seed composition must accommodate their physiological state; quiescent or dormant [6], dry or hydrated [7, 8], orthodox or recalcitrant [9]. The occurrence of seed storage proteins (SSP) is a major technical problem because of the extended protein dynamic range [10].

As a discipline, proteomics is a broad, instrument-intensive research area that has progressed rapidly since its genesis 20 years ago. While methods of protein isolation and separation have improved during this period, it has been the improvements in instrumentation that have driven expansion of the field [11–14]. Significant improvements in sensitivity, mass accuracy, and fragmentation in recent years have led to widespread adoption of proteomic strategies. As proteomics has matured as a discipline, there have been an increasing number of specialized studies; organism-specific, cell or organelle-specific, and specific for increasingly narrow physiological or developmental targets. Herein I address the burgeoning field of seed proteomics.

2 Samples

All meaningful proteomic comparisons, whether gel-based [15] or gel-free [16] are complicated by both technical and biological variations. It is paramount that this not be further exacerbated by sample variation. Otherwise, seemly important quantitative or qualitative differences might simply be artifacts because of nonuniform starting material. This can be particularly important with regard to studies of seeds where it is well-known that there can be substantial developmental and physiological changes [3, 17–19].

2.1 Anatomy: *The Whole, the Sum, or Just Some of the Parts?*

Formation of the seed completes a reproductive cycle that begins with development of flowers, and pollination [4]. The embryo develops from the zygote formed by the fusion of an egg and a sperm cell, and the seed coat or testa develops from the integuments of the ovule [20]. Angiosperm seeds typically have the nutrients, including the SSP, stored in either the cotyledons of dicots [3] or endosperm of monocots [2]. The SSP in Gymnosperm seeds are stored within the endosperm-related megagametophyte [21]. During development or maturation it is typical to analyze whole seeds, as the seed coats and embryonic axes make a relatively small contribution to the mass of the whole seed. Thus, “seed proteomics” often means proteomic analysis of the storage tissues/organs. Post-germination the seed coat remnants generally slough off and are discarded, and the storage organs are often separated from the growing seedling.

Is whole-seed proteomic analysis justified? It depends upon the nature of the seeds and the intended goals of the study. For example if focus is on the post-germination, heterotrophic to autotrophic transition, or mobilization of SSP, then one can reasonably assume that the relatively small contribution of the seed coat and axis proteins will not confound the analysis. There is also the matter of practicality. Removal of the seed coat and embryonic axis from a tiny *A. thaliana* seed is a near impossible task versus removal of the relatively large castor oil (*Ricinus communis*) seed coat and axis. During seed development and maturation, however, the different organs and tissues have their own distinctive protein composition (c.f., Fig. 1), which reflects their specialized origins and roles.

In these cases it would seem that separation of the seeds into component parts is justified, if not essential. What certainly is essential is for the Methods sections to include a description of what was done during sample preparation!

All seeds can be separated into two categories depending upon the nature of the storage organ; endosperm-dominant or cotyledon-dominant [4]. Endosperm-dominant seeds typically store starch and proteins, while cotyledon-dominant seeds seem approximately evenly distributed among starch plus proteins and starch plus lipids groups.

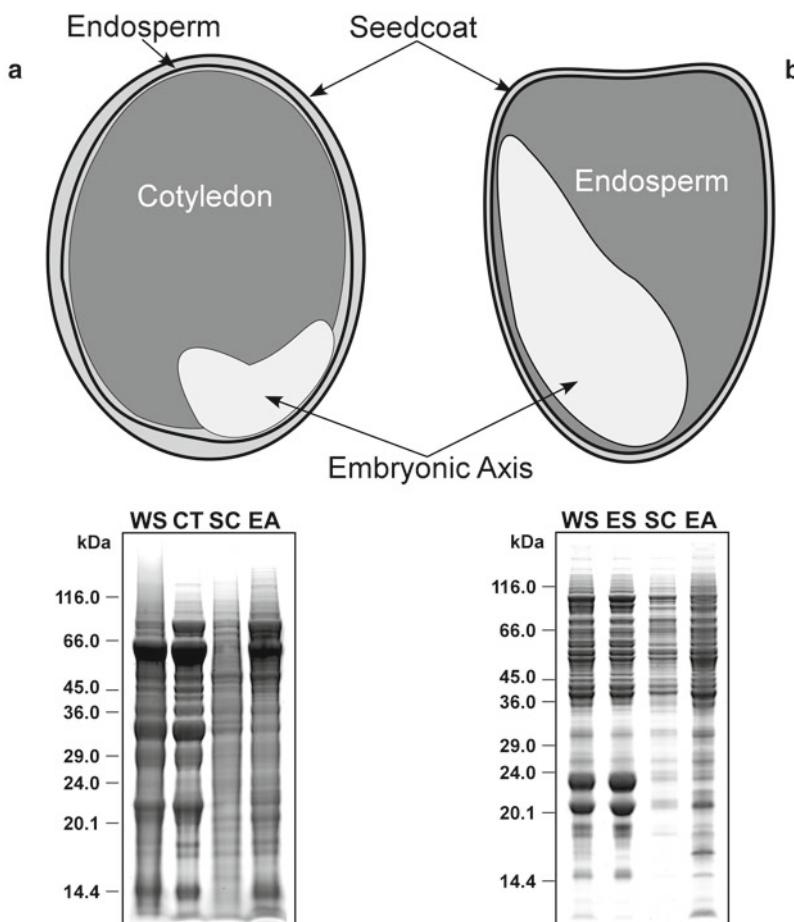


Fig. 1 Seed anatomy: The whole, the sum, or just some of the parts? (a) Is a diagrammatic presentation of a S4 developing soybean (*G. max* (L.) Merr.) seed; a representative cotyledon-dominant seed. Below is a Coomassie-stained 1D SDS-gel of a total protein fraction isolated from whole seeds (WS), cotyledons (CT), seed coats (SC), and embryonic axes (EA). The soybean seed endosperm is only one to three cells thick and is not shown in these analyses. (b) Is a diagrammatic presentation of a developing maize (*Z. mays* L.) seed

There are of course exceptions. Castor (*R. communis* L.) is a large endosperm-dominant seed that stores proteins plus oil [21, 22].

2.2 Physiological Uniformity

Seeds develop [3, 4], reach maturity, and then when the conditions are appropriate germinate [19, 23]. The first 7–10 days after pollination are termed embryogenesis, and are dominated by cell division and organ formation. Post-embryogenesis is an extended period of cellular specialization which includes the synthesis and accumulation of reserve polymers. The final aspect of seed development is that of maturation/dehydration.

The rate of seed development is variable and is controlled by multiple environmental factors. It is common to see reference to

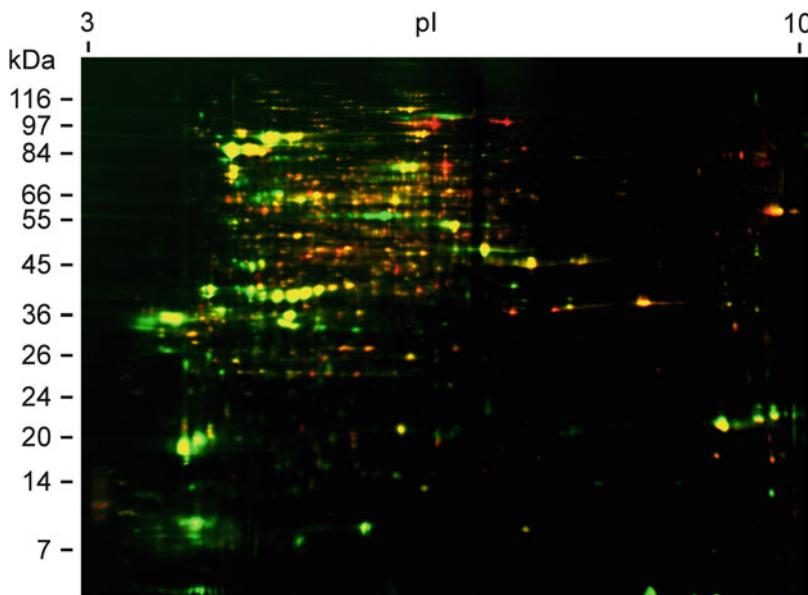


Fig. 2 Protein dynamics during *G. max* (L.) Merr. seed development. Total protein fractions were isolated from S2 (27–42 mg) and S4 (115–150 mg) soybean seeds, and compared by Difference InGel Electrophoresis (DIGE). Proteins from S2 seeds were labeled with Cye 3, while proteins from S4 seeds were labeled with Cye 5. The combined proteins were separated in the first direction by isoelectric focusing using an immobilized 3–10 pH gradient, and by SDS-PAGE in the second dimension using a 10–18 % acrylamide gradient gel. Proteins that are equally abundant in the two stages appear as *yellow dots*, while proteins more abundant in the S4 seeds appear *green* and those more abundant in the S2 seeds appear *red*

seed age expressed as a chronological age (Days After Flowering, DAF, or Days After Anthesis, DAA). This is unfortunate because there is essentially no correlation between the chronological ages of a seed from a growth-chamber grown plant in January and a field-grown plant in June. Instead of chronological age, seed biologists have developed various staging platforms that are indicative of physiological age. This staging is typically based upon easily quantifiable characteristics such as fresh weight, size (length \times width), color, and a series of morphological and anatomical features [24–26]. These staging systems can be very robust, and are far more reliable than chronological age. Especially during the period of cellular specialization, there are substantial differences in the protein profiles of seeds at different physiological ages. The results of two-dimensional DIGE analysis of developing soybean cotyledons highlight both protein complexity and dynamics as seed development proceeds (Fig. 2). Similar dynamic complexity has also been reported for developing castor [21] and rape (*Brassica napus*) [26] seeds.

Mature seeds can be quiescent or dormant. Quiescent seeds are “dry” (5–15 % moisture) and metabolically inert. They are capable of shifting to metabolic activity, requiring only hydration

and a permissive temperature in order to proceed to germination. However, quiescent seeds can be orthodox or recalcitrant [19, 27]. Orthodox seeds present higher dehydration tolerance, whereas recalcitrant seeds are dehydration-sensitive both throughout development and after shedding from the parent plant [28]. Information on the biochemical and cellular alterations produced by dehydration in recalcitrant seeds is rare, and the bases of this sensitivity remain obscure. Recalcitrant seeds lose their viability if stored for any length of time, even under conditions that are normally conducive to seed longevity, i.e., low moisture content and low temperatures. Many tropical plants, e.g., coconut (*Cocos nucifera* L.), rubber (*Hevea brasiliensis* Müll. Arg), and tea (*Camellia sinensis* (L.) Kuntze), have recalcitrant seeds [28].

In contrast, after reaching physiological maturity, seeds of many plant species enter a state of dormancy during which they will not germinate (or germinate very slowly). Seed dormancy is believed to be an adaptive trait that prevents premature germination during adverse environmental conditions. Under natural conditions, seed survival in the soil and cycling through states of dormancy are major ecological characters determining entry and persistence in ecosystems [29, 30], and seed dormancy is a major trait altered during domestication of wild species [31]. Embryo dormancy is a genetically and environmentally determined developmental state imposed following imbibition of mature seeds, in which cells are metabolically active, but growth processes are repressed [23].

Breaking of dormancy can either occur gradually in the dry state (after-ripening) or be initiated by imbibition under defined conditions (e.g., stratification). Changes in the proteome (including changes due to PTM) accompany both after-ripening [7, 8], and stratification (Fig. 3).

Germination sensu stricto encompasses the events beginning with hydration of the mature dry seed and elongation of the embryonic radicle such that it penetrates the seed coat [19, 23, 32, 33]. Following germination there is a period of postgerminative growth and development that precedes autotrophy [19]. Many studies, including most that have used a proteomics-based strategy, use the term “germination” when they are actually studying postgerminative growth [34]. In most instances the time between seed hydration and true germination is relatively short, and while there are a few instances of bona fide changes in protein abundance associated with germination [35], major changes in protein profiles are more typically associated with de novo synthesis of enzymes involved with mobilization of reserve polymers during postgerminative growth [3, 36, 37]. Considering the breadth of seed physiology, except in instances where the changes that occur within a physiological stage or that accompany transition from one stage to another are specifically targeted, it is critical that comparisons are made between samples that are truly comparable!

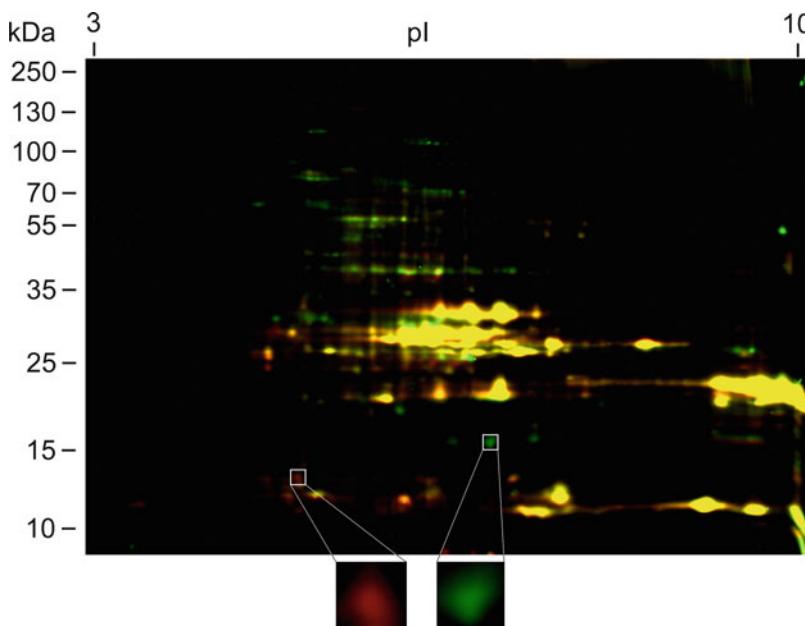


Fig. 3 Protein dynamics during cold treatment of *A. thaliana* seeds. Proteins were isolated from mature, dry, dormant seeds of *A. thaliana* and compared with proteins from seeds incubated for 2 days at 4 °C, using DIGE. A total protein fraction was isolated as described by Schober et al. [90] with minor modifications. Proteins from dormant seeds were labeled with Cye 3, and proteins from cold treated seeds were labeled with Cye 5. The combined proteins were separated as described in the legend for Fig. 3. Proteins that are equally abundant in the two treatments appear as *yellow dots*. Proteins more abundant in the dormant seeds appear *green* and those more abundant after cold treatment appear *red*

2.3 Genetic Uniformity

Not surprisingly, most seed proteomics research has addressed agricultural crop plants. Typically these seeds are obtained from a commercial source, and it is reasonable to assume that they are genetically uniform so results will not be biased by the contributions of a “contaminating” proteome. This important point is often overlooked, however there are reports of substantial differences between samples of a common species that are distinct cultivars/genotypes [38, 39].

The issue of genetic uniformity is much more serious during analyses of non-cultivated plants. In a few exceptional instances there are repositories for weed/wild seeds. *A. thaliana*, for example, has become such a common experimental system that genetically uniform cultivar/ecotype seeds are readily available (e.g., the Arabidopsis Biological Resource Center; <http://abrc.osu.edu/home>). With soybeans or maize, the wild progenitors of extant crop plants, *G. soja* and *Z. diploperennis*, respectively, have been used in genetics and genomics programs, and there are sources for uniform wild seeds (e.g., USDA Soybean Germplasm Collection, <http://www.ars.usda.gov/is/np/SoybeanGermplasm/SoybeanGermplasmIntro.htm>; USDA/ARS Maize Germplasm Collection, <http://www.ars-grin.gov/cgi-bin/npgs/html/crop.pl?89>).

Sample uniformity with non-crop related, long-lived seeds, such as those of trees, can be especially problematic. While seeds collected from mature forests are potentially of clonal origin, it is probably best to assume that they are not. In such cases (e.g., 40, 41), the only real option is to describe as thoroughly as practical where/how the seeds were collected in the wild, and either maintain or deposit a sample with some sort of botanical archiving service. Finally, it is best to remain conservative in interpretation of the results of proteomic analyses especially in regard to low abundance proteins.

3 Storage Proteins

3.1 Terminology

The systematic terminology used to define classes of SSP is solubility-based [42]. Thus, the SSP soluble in H₂O are called albumins, while those soluble in dilute saline, aqueous alcohol, or dilute alkali or acid are globulins, prolamins, or glutelins, respectively. While albumins are found in all seeds, prolamins and glutelins are abundant in monocot seeds and globulins are prevalent in dicot seeds [43, 44]. The globulin SSP has additionally been grouped based upon sedimentation into 7S or 11S [45]. Furthermore, there are a plethora of trivial names assigned to SSP based on their Latin generic names: the zeins from maize (*Zea mays*), hordeins from barley (*Hordeum vulgare*), secalins from rye (*Secale cereale*), etc.

3.2 Synthesis and Processing

The SSP are synthesized as precursors on the rough-endoplasmic reticulum (ER), targeted to the ER lumen by an N-terminal signal sequence which is co-translocationally removed by signal peptidase [3]. The SSP are then sorted from bulk protein traffic through the secretory pathway to their site of deposition, the protein storage vacuole (PSV). There are two targeting pathways for SSP, one for prolamins and a second for the non-prolamin SSP. The prolamin-containing PSV arise by direct vesiculation from the ER; prolamins do not traverse the classical secretory pathway. Non-prolamin SSP passes through the classical secretory pathway, and is sorted from a post-Golgi compartment to the PSV [46].

The prolamin PSV can have a complex internal architecture, which is thought to be the result of different rates of synthesis of the various subunits [47]. In contrast, non-prolamin PSV has a uniformly granular appearance [48]. An additional morphological characteristic of prolamin-PSV is, since they are directly derived from the rough-ER, that they are studded with ribosomes [49]. Rice (*Oryza sativa*) endosperm is unusual in containing both prolamin- and non-prolamin PSV [50].

The 2S, 7S, and 11S SSP are synthesized as large precursors (now known to be prepro-polyproteins) [3] which undergo

initial processing and assembly while still within the ER [51, 52]. In addition to the N-terminal signal sequence, the canonical SSP-precursor primary-sequence includes a pro-sequence containing the PSV-targeting information, plus at least two linker/protease cleavage sites. It is not uncommon for the SSP pro-proteins to display the canonical Asn-X-Ser/Thr glycosylation sequon, and become N-glycosylated during passage through the ER. In some instances the N-glycosylation sequons are within the pro-sequence or linker regions, and are subsequently removed during proteolytic processing. In other instances, however, they persist and are present in the final fully processed SSP. Proteolytic processing and assembly continue during transit through the secretory pathway, and are completed within the PSV [53].

After proteolytic processing has been completed, the final polypeptide associations are stabilized by formation of at least one disulfide-bond. The final stable SSP structures can be as simple as a “ $\alpha\beta$ -heterodimer” or as complex as a $\alpha_3\beta_3$ -heterohexamer. It is noteworthy that the incisive SSP-based research of a generation of plant cell biologists has been mechanistically verified and extended by recent results based upon availability of genome sequence information and application of tandem mass spectrometry (MS/MS) [4]. It is equally important to note that the understanding of SSP processing can greatly simplify interpretation of MS results. This is especially true of gel-based results, where it could be difficult to interpret the position of a protein that is very different from the MW and pI of the primary translation product unless one has prior knowledge of both proteolytic and glycolytic processing events [3].

3.3 Addressing the Dynamic Range Problem

The abundance of the SSP can be a great benefit ... if you are studying SSP. If not, especially if a gel-based strategy is employed, then the SSP can substantially interfere with analysis of total proteins. Even if the seeds of a hypothetical plant have only a 2S albumin and an 11S globulin as SSP, the complexity in 2D gel spot-patterns can be nearly overwhelming because of the contributions of extended multigene families plus heterogeneity in both proteolytic and glycolytic processing. Advances in high performance ion trap mass spectrometry (MS) employing electrospray ionization (ESI) and nanoflow liquid chromatography (nLC) have significantly improved the ability to analyze proteins [14]. Despite these advances there remain serious inherent limitations. Simply put, the range of protein concentrations in biological samples is very large (as much as 10^{12}) [10] and the dynamic range of the analytical methods used is small (less than 10^3 in most cases) [54]. The only practical way to circumvent this problem is by including a prefractionation/depletion step to the work-flow (Table 1).

There are several different strategies by which this can be achieved [55]. With the exception of the prolamins (Fig. 4a), fractional solubility has not proved to be generally useful in SSP-depletion.

Table 1
Sample pretreatment strategies for reduction of the seed storage proteins, and shortening of the sample dynamic range

| Method | Ease of application | Efficiency | Capacity | Cost |
|---------------------------|---------------------|------------|----------|-------|
| Fractional solubility | +++ | + | +++++ | + |
| “Salting out” | +++++ | +++++ | +++++ | + |
| Lectin-chromatography | ++ | ++ | ++ | +++ |
| Immunoremoval | ++ | ++ | ++ | +++++ |
| Isoelectric precipitation | + | + | +++ | +++++ |
| Random peptide-beads | + | +++ | ++ | +++++ |

Scale: Five is the easiest, has the greatest efficiency, has the highest capacity, and is the most expensive

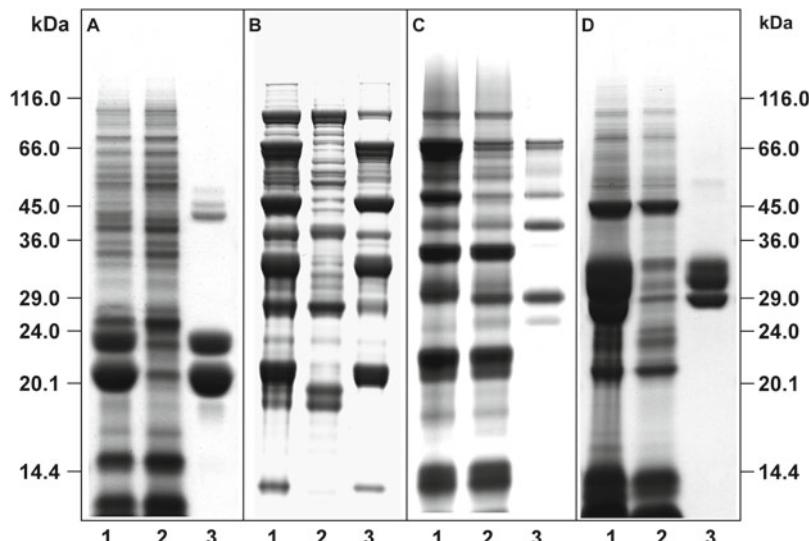


Fig. 4 Depletion of the supra-abundant seed storage proteins as a prelude to proteomic analyses. *Lane A*, fractional solubility of the maize endosperm prolamins. A1, total proteins; A2, dilute-buffer soluble proteins; A3, extraction of the zein SSP with 35 % (v/v) ethanol. *Lane B*, “salting out” of soybean globulin SSP [3]. B1, total proteins from mature seeds; B2, globulin-depleted proteins; B3, proteins precipitated by incubation with 10 mM CaCl₂. *Lane C*, reduction in the beta-conglycinin SSP from a total protein fraction isolated from developing soybean seeds by immobilized-lectin chromatography [58]. *Lane D*, immunoremoval of the agglutinin proteins from mature castor seed endosperm preparations. D1, total endosperm proteins; D2, ricin-depleted endosperm proteins; D3, removed castor SSP [3]. The positions of size marker proteins are indicated to the *left* and *right* of the Y-axes

However, if a “total protein” fraction has been initially isolated, then fractional solubility of the globulins can be exploited by selective precipitation [56, 57] (Fig. 4b). If the interfering SSP are N-glycosylated, it is possible to use immobilized-lectin affinity chromatography for their removal. This strategy was useful in

removing the acidic subunit of beta-conglycinin from total soybean (*G. max* (L.) Merr.) seed proteins [55] (Fig. 2c). If suitable antibodies are available [55], then immunoremoval can be a useful strategy for depletion of SSP (Fig. 2d). Application of any of these methods will need to be individually refined. For example, antibodies and lectins might need to be chemically cross-linked in order to be useful if SSP are prepared under denaturing conditions [58].

In principle, it should be possible to deplete a sample of SSP by either differential solubility or isoelectric precipitation (Table 1). In fact, these simple strategies seldom work satisfactorily. For example, when a total protein fraction prepared from developing soybean seeds was dialyzed exhaustively against deionized water it should have yielded a globulin-depleted albumin fraction and an insoluble globulin fraction. Both fractions, however, were heavily cross-contaminated (JAM, unpublished). Results were better with rice seed glutelins, and best with maize prolams (Fig. 4a). Extraction of dry, mature maize endosperm did yield a prolamin (zein) enriched fraction, but some of the zeins remained associated with the ethanol-insoluble material. This might be because of the oxidative cross-linking of the zein subunits [59]. If this were the case then results might have been improved by including 2 % (V/V) 2-mercaptoethanol in the 70 % ethanol solvent.

The first practical, usable, and commercially available device for Liquid Phase-IEF/isolectric precipitation was dubbed the Bio-Rad Rotofor. This method has been used extensively in both shotgun proteomic analyses [60] and in individual, specific applications [61]. More recently, Agilent Technologies has marketed a LP-IEF system called the 3100 OFFGEL Fractionator. In this system, the pH gradient results from use of an immobilized pH gradient strip, rather than pH limit ampholytes. The Rotofor and OFFGEL, and similar devices such as the ZOOM IEF Fractionator [62] are capable of suitable SSP depletion. The disadvantages of include being technically difficult and expensive.

Krishnan and associates [57, 63] adopted a somewhat different strategy for depleting total soybean seed proteins of the very abundant 7S and 11S globulins. They found that incubation of a total protein fraction with 10 mM Ca²⁺ led to precipitation of the seed globulins glycinin and β -conglycinin (Fig. 4b). This simple and inexpensive method removed 87 ± 4 % of the SSP from the sample and allowed identification of 541 previously inconspicuous proteins. In preliminary experiments, this method appeared to work equally well with seeds from other legumes (e.g., peanut, bean, pea, and alfalfa). While it would be difficult to improve the ease or cost of the calcium-precipitation strategy, it does not always work perfectly and might be more productively combined with one of the other methods described herein.

Many of the abundant SSP are N-glycosylated, making lectin-affinity chromatography a potentially useful depletion method in this context [64, 65]. Concanavalin A (con A) recognizes and binds

high-mannose type glycans (those with terminal α -D-mannopyranoside or α -D-glucopyranoside residues). Many additional lectins are available, with glycan-specificities that encompass most of the typical complex-type glycans. The soybean β -conglycinin SSP are known to be N-glycosylated [66], and at least one glycan is of the high-mannose type which should allow use of con A in a lectin-based affinity-depletion step [55, 58]. Using con A-affinity chromatography removes all high-mannose type glycoproteins from a complex mixture (Fig. 4c), including low-abundance non-SSP proteins. These can be eluted from the lectin with the hapten α -methyl-mannoside, separated from the β -conglycinin by electrophoresis, digested, and analyzed by LC-MS. Immobilized lectins can be reused multiple times, reducing the disadvantages of initial expense and capacity (Table 1).

Removal of supra-abundant proteins with antibodies is the most specific of all of the depletion methods. The efficacy of immunoremoval depends upon the avidity of the antibodies used. In general, antibodies prepared against a native antigen are more useful in immunoremoval than antibodies prepared against a denatured antigen. This can be problematic with SSP, some of which are only slightly soluble in aqueous solutions. However, at least some of the difficulties in antigen solubility can be overcome through the use of synthetic peptide-antibodies [21] (Fig. 4d). There are several potential difficulties that must be considered before committing to an immunoremoval strategy (Table 1) [55]. Antibody preparation can be expensive, and is not always efficient. Nonetheless, the selectivity and efficiency of immunoremoval make it a good first choice when designing an experimental strategy.

Combinatorial-ligand random-peptide beads allow a protein to be removed from a complex mixture, up to the point of saturation of the ligand [67]. With sufficient diversity, it is theoretically possible to have an immobilized ligand complementary to each protein in a complex mixture, ensuring that they would all be adsorbed. When a biological sample is incubated with such a ligand-library under capacity-restrained conditions, abundant proteins will saturate all available high-affinity ligands and the remaining non-binding majority of the protein will remain in solution [68]. In contrast, a low abundance protein will not saturate the corresponding high-affinity ligand, and most of this protein will be removed from solution [55]. Based on this saturation-overload principle, use of a combinatorial library should enrich for low-abundance proteins relative to those of high abundance [69]. Elution of the entire population of proteins adsorbed to the beads should result in a solution with a smaller dynamic range than the starting material but still including representatives of all of the original proteins.

The combinatorial-peptide beads, in the format of the ProteoMiner kit [70], have been productively used in analyses of

Hevea brasiliensis latex [69] and the cytoplasmic proteome from maize [69], and more recently spinach leaves [71]. At the time of writing, there were no publications describing the use of ProteoMiner for analysis of seed proteins. Our attempts to use ProteoMiner kit, under the conditions developed for spinach leaves, to either deplete samples from developing soybeans of the abundant SSP or to prepare samples that yielded an increased number of protein ID's have been without success to date (JAM, unpublished). The method is known to be sensitive to pH and solvent characteristics, and will likely have to be customized for individual applications. Furthermore, we speculate that the often unusual solubility of SSP might be problematic in terms of nonspecific adsorption.

4 The Future of Seed Proteomics

4.1 More, Better, and Quantitative

The short-term future of seed proteomics will feature more and better protein IDs. This goal will be achieved through combined advances in instrumentation [72], bioinformatic analyses and methods for protein identification [73], improved databases (e.g., phytozome.net/), and, of course, improvements in the methods used to deplete the supra-abundant SSP from input samples. In the latter case, the strategy of using combinatorial-ligand random-peptide beads appears to have substantial potential. One obvious improvement would be to remove the “random” component. In part because of their abundance, genes and cDNAs for SSP were among the first sequenced at the beginning of the genomics era [3]. Because the sequences of many SSP are extant [74–76], it should be relatively simple to prepare “designer sequence”—beads that would have both the capacity and specificity to efficiently remove these supra-abundant components of the seed proteome. This strategy should additionally be amenable to incorporation into high-throughput experimental designs.

Another obvious need is for more comparative analyses, especially in terms of gymnosperm seeds, although in many cases this will require parallel development of better genomic/EST resources to facilitate protein identification.

There can be no debate about the need to move from qualitative to quantitative proteomic analyses, and there are proponents of both protein labelling and of label-free methods to achieve this end [14]. While the debate will likely continue into the future, an increasing number of researchers are using a simple, inexpensive label-free method termed spectral counting [77–79]. Regardless of the method ultimately adopted, it is clear that future proteomics-based studies of seeds can only benefit from employing quantitative approaches.

4.2 Posttranslational Modifications

One aspect of biology that cannot be addressed in any other large-scale survey (e.g., transcript profiling) is the posttranslational modification (PTM) of proteins. Much of the elaboration of the proteome is the result of PTM [80]. In addition to the nature of the PTM, it is also important to determine which residue is modified and the result of the modification. Do the PTM affect protein turnover, activity, interactions? In addition to using bottom up analysis strategies, it is important that top down proteomic analyses are included in analysis of PTM [81]. Thus far only a handful of studies of seed proteins have included systematic analysis of PTM, and these have addressed only phosphorylation [82]. In addition to identification of the nature of PTM, it is important to determine their stoichiometry in the proteome. Changes in PTM stoichiometry likely indicate a specific functional change, while increase in protein amount with a parallel increase in the extent of PTM means that the stoichiometry has not changed and there is unlikely to be any functional consequence [83].

4.3 Protein Interactions

The other major aspect of biology that cannot be addressed in any sort of high-throughput or computational context is that of protein interactions (e.g., the interactome). It is increasingly appreciated that proteins do not function in isolation [84]. Unfortunately, however, there are no extant publications addressing MS-based analysis of protein interactions as a component of seed biology. But there should be! The next wave of understanding of how proteins function will involve analysis of protein interactions in a cellular or subcellular context [85, 86].

4.4 Imaging Mass Spectrometry

Finally, an exciting but relatively new application of MS in proteomics studies is at the tissue- and sometimes even cellular level of analysis; imaging MS or MALDI-imaging [14, 87, 88]. By gating the detector to a specified mass (or range), it is possible to detect the location of a protein in whole tissue/organ mounts. Especially in conjunction with high-sensitivity instruments, this method could provide significant insight into the cellular and even subcellular molecular organization of seeds and their component organs and tissues. To date the greatest success of MALDI imaging has been in studies of small molecules [89] and membrane lipids [90]. There is, however, a constant evolution and development of methods for improved protein detection and spatial resolution [91].

Acknowledgments

The graphics were created by M.L. Johnston. Professor J.J. Thelen provided the antibodies used in Fig. 2d.

References

- Kanei M, Horiguchi G, Tsukaya H (2012) Stable establishment of cotyledon identity during embryogenesis in *Arabidopsis* by ANGUSTIFOLIA3 and HANABA TARANU. *Development* 139:2436–2446
- Li J, Berger F (2012) Endosperm: food for humankind and fodder for scientific discoveries. *New Phytol* 195:290–305
- Miernyk JA, Hajduch M (2011) Seed proteomics. *J Proteomics* 74:389–400
- Miernyk JA, Petrova A, Olmedilla A et al (2011) Using proteomics to study sexual reproduction in angiosperms. *Sex Plant Reprod* 24:9–22
- Herridge RP, Herridge RP, Day RC et al (2011) Rapid analysis of seed size in *Arabidopsis* for mutant and QTL discovery. *Plant Methods* 7:3
- Chibani K, Ali-Rachedi S, Job C et al (2006) Proteomic analysis of seed dormancy in *Arabidopsis*. *Plant Physiol* 142:1493–1510
- Finch-Savage WE, Leubner-Metzger G (2006) Seed dormancy and the control of germination. *New Phytol* 171:501–523
- Finkelstein R, Reeves W, Ariizumi T et al (2008) Molecular aspects of seed dormancy. *Annu Rev Plant Biol* 59:387–415
- Balbuena TS, Jo L, Pieruzzi FP et al (2011) Differential proteome analysis of mature and germinated embryos of *Araucaria angustifolia*. *Phytochemistry* 72:302–311
- Hortin GL, Sviridov D (2010) The dynamic range problem in the analysis of the plasma proteome. *J Proteomics* 73:629–636
- Armiotti A, Damonte G (2010) Achievements and perspectives of top-down proteomics. *Proteomics* 10:3566–3576
- Jacob RJ (2010) Bioinformatics for LC-MS/MS-based proteomics. *Methods Mol Biol* 658:61–91
- May C, Brosseron F, Chartowski P et al (2011) Instruments and methods in proteomics. *Methods Mol Biol* 696:3–26
- Thelen JJ, Miernyk JA (2012) The proteomic future: where mass spectrometry should be taking us. *Biochem J* 444:169–181
- Rabilloud T (2012) The whereabouts of 2D gels in quantitative proteomics. *Meth Mol Biol* 893:25–35
- Nikolov M, Schmidt C, Urlaub H (2012) Quantitative mass spectrometry-based proteomics: an overview. *Meth Mol Biol* 893:85–100
- Weber H, Borisjuk L, Wobus U (2005) Molecular physiology of legume seed development. *Annu Rev Plant Biol* 56:253–279
- Hajduch M, Matusova R, Houston N et al (2011) Comparative proteomics of seed maturation in oilseeds reveals differences in intermediary metabolism. *Proteomics* 11:1619–1629
- Rajjou L, Duval M, Gallardo K et al (2012) Seed germination and vigor. *Annu Rev Plant Biol* 63:507–533
- Moise JA, Han S, Gudynaitė-Savitch L et al (2005) Seed coats: structure, development, composition, and biotechnology. *In Vitro Cell Dev Biol Plant* 41:620–644
- Houston NL, Hajduch M, Thelen JJ (2009) Quantitative proteomics of seed filling in castor: comparison with soybean and rapeseed reveals differences between photosynthetic and nonphotosynthetic seed metabolism. *Plant Physiol* 151:857–868
- Campos FAP, Nogueira FCS, Cardoso K et al (2010) Proteome analysis of castor bean seeds. *Pure Appl Chem* 82:259–267
- Bewley JD (1997) Seed germination and dormancy. *Plant Cell* 9:1055–1066
- Miernyk JA, Trelease RN (1981) Role of malate synthase in citric acid synthesis by maturing cotton embryos: a proposal. *Plant Physiol* 67:875–881
- Miernyk JA, Dennis DT (1992) A developmental analysis of the enolase isozymes of *Ricinus communis*. *Plant Physiol* 99:748–750
- Hajduch M, Casteel JE, Hurrelmeyer KE et al (2006) Proteomic analysis of seed filling in *Brassica napus*. Developmental characterization of metabolic isozymes using high-resolution two-dimensional gel electrophoresis. *Plant Physiol* 141:32–46
- Linkies A, Graeber K, Knight C et al (2010) The evolution of seeds. *New Phytol* 186:817–831
- Berjak P, Pammerer NW (2008) From Avicennia to Zizania: seed recalcitrance in perspective. *Ann Bot* 101:213–228
- Baskin CC, Baskin JM (1998) Seeds. Ecology, Biogeography, and Evolution of Dormancy and Germination. Academic, San Diego, p 666
- Bewley JD, Black M (1994) Seeds. NY, Plenum Press, Physiology of Development and Germination. Second Ed., p 445
- Harlan JR (1992) Crops and Man. *Am Soc Agron. Am Soc Crop Sci*, Madison, WI, p 294
- Arc E, Galland M, Cueff G et al (2011) Reboot the system thanks to protein post-translational modifications and proteome diversity: how quiescent seeds restart their metabolism to prepare seedling establishment. *Proteomics* 11:1606–1618

33. Weitbrecht K, Müller K, Leubner-Metzger G (2011) First off the mark: early seed germination. *J Exp Bot* 62:3289–3309
34. Kim HT, Tae C, Ung-Kyu R et al (2011) Mobilization of storage proteins in soybean seed (*Glycine max* L.) during germination and seedling growth. *Biochim Biophys Acta* 1814:1178–1187
35. Finnie C, Andersen B, Shahpuri A et al (2011) Proteomes of the barley aleurone layer: a model system for plant signalling and protein secretion. *Proteomics* 11:1595–1605
36. Müntz K, Belozersky MA, Dunaevsky YE et al (2001) Stored proteinases and the initiation of storage protein mobilization in seeds during germination and seedling growth. *J Exp Bot* 52:1741–1752
37. Graham IA (2008) Seed storage oil mobilization. *Annu Rev Plant Biol* 59:115–142
38. Natarajan S, Xu C, Bae H et al (2007) Proteomic and genomic characterization of Kunitz trypsin inhibitors in wild and cultivated soybean genotypes. *J Plant Physiol* 164: 756–763
39. Natarajan S, Xu C, Bae H et al (2007) Proteomic and genetic analysis of glycinin subunits of sixteen soybean genotypes. *Plant Physiol Biochem* 45:436–444
40. Dias LLC, Balbuena TS, Silveira V et al (2010) Two-dimensional gel electrophoretic protein profile analysis during seed development of *Ocotea catharinensis*: a recalcitrant seed species. *Braz J Plant Physiol* 22:23–33
41. Pawłowski TA (2010) Proteomic approach to analyze dormancy breaking of tree seeds. *Plant Mol Biol* 73:15–25
42. Osborne TB (1908) Our present knowledge of plant proteins. *Science* 28:417–427
43. Shewry PR, Tatham AS (1990) The prolamin storage proteins of cereal seeds: structure and evolution. *Biochem J* 267:1–12
44. Xu JH, Messing J (2009) Amplification of prolamin storage protein genes in different subfamilies of the Poaceae. *Theor Appl Genet* 119: 1397–1412
45. Shewry PR, Napier JA, Tatham AS (1995) Seed storage proteins: structures and biosynthesis. *Plant Cell* 7:945–956
46. Herman EM, Larkins BA (1999) Protein storage bodies and vacuoles. *Plant Cell* 11:601–614
47. Lending CR, Larkins BA (1989) Changes in the zein composition of protein bodies during maize endosperm development. *Plant Cell* 1:1011–1023
48. zur Nieden U, Neumann D, Manteuffel R et al (1982) Electron microscopic immunocytochemical localization of storage proteins in *Vicia faba* seeds. *Eur J Cell Biol* 26:228–233
49. Burr B, Burr FA (1976) Zein synthesis in maize endosperm by polyribosomes attached to protein bodies. *Proc Natl Acad Sci USA* 73:515–519
50. Krishnan HB, White JA (1995) Morphometric analysis of rice seed protein bodies (implication for a significant contribution of prolamin to the total protein content of rice endosperm). *Plant Physiol* 109:1491–1495
51. Chrispeels MJ, Higgins TJ, Spencer D (1982) Assembly of storage protein oligomers in the endoplasmic reticulum and processing of the polypeptides in the protein bodies of developing pea cotyledons. *J Cell Biol* 93:306–313
52. Bollini R, Vitale A, Chrispeels MJ (1983) In vivo and in vitro processing of seed reserve protein in the endoplasmic reticulum: evidence for two glycosylation steps. *J Cell Biol* 96:999–1007
53. Kumamaru T, Uemura Y, Inoue Y et al (2010) Vacuolar processing enzyme plays an essential role in the crystalline structure of gluten in rice seed. *Plant Cell Physiol* 51:38–46
54. Vestal ML (2011) The future of biological mass spectrometry. *J Am Soc Mass Spectrom* 22:953–959
55. Miernyk JA, Johnston ML (2012) Digging deeper into the seed proteome: prefractionation of total proteins. In: Agrawal GK, Rakwal R (eds) *Seed development: Omics technologies toward improvement of seed quality and crop yield*. Springer, Berlin
56. Ferreira RB, Franco E, Teixeira AR (1999) Calcium- and magnesium-dependent aggregation of legume seed storage proteins. *J Agric Food Chem* 47:3009–3015
57. Krishnan HB, Oehrle NW, Natarajan SS (2009) A rapid and simple procedure for the depletion of abundant storage proteins from legume seeds to advance proteome analysis: a case study using *Glycine max*. *Proteomics* 9:3174–3188
58. Miernyk JA, Johnston ML (2006) Chemical cross-linking immobilized-Concanavalin A for use in proteomic analyses. *Prep Biochem Biotechnol* 36:203–213
59. Field JM, Shewry PR, Miflin BJ (1983) Aggregation states of alcohol-soluble storage proteins of barley, rye, wheat and maize. *J Sci Food Agric* 34:362–369
60. Righetti PG, Castagna A, Herbert B et al (2003) Prefractionation techniques in proteome analysis. *Proteomics* 3:1397–1407
61. Vincent D, Balesdent M-H, Gibon J et al (2009) Hunting down fungal secretomes using liquid-phase IEF prior to high resolution 2-DE. *Electrophoresis* 32:4118–4136

62. Richardson MR, Liu S, Ringham HN et al (2008) Sample complexity reduction for two-dimensional electrophoresis using solution isoelectric focusing prefractionation. *Electrophoresis* 29:2637–2644
63. Natarajan SS, Savithiry S, Krishnan HB et al (2009) An efficient extraction method to enhance analysis of low abundant proteins from soybean seed. *Anal Biochem* 394: 259–268
64. Hirabayashi J (2004) Lectin-based structural glycomics: glycoproteomics and glycan profiling. *Glycoconj J* 21:35–40
65. Van Damme EJ (2011) Lectins as tools to select for glycosylated proteins. *Methods Mol Biol* 753:289–297
66. Maruyama N, Katsume T, Wada Y et al (1998) The roles of the N-linked glycans and extension regions of soybean beta-conglycinin in folding, assembly and structural features. *Eur J Biochem* 258:854–862
67. Boschetti E, Monsarrat B, Righetti PG (2007) The “Invisible Proteome”: how to capture the low abundance proteins via combinatorial ligand libraries. *Curr Proteomics* 4:198–208
68. Boschetti E, Righetti PG (2009) The art of observing rare protein species in proteomes with peptide ligand libraries. *Proteomics* 9:1492–1510
69. Boschetti E, Bindschedler LV, Tang C et al (2009) Combinatorial peptide ligand libraries and plant proteomics: a winning strategy at a price. *J Chromatogr A* 1216:1215–1222
70. Boschetti E, Righetti PG (2008) The ProteoMiner in the proteomic arena: a non-depleting tool for discovering low-abundance species. *J Proteomics* 71:255–264
71. Fasoli E, D'Amato A, Kravchuk AV et al (2011) Popeye strikes again: the deep proteome of spinach leaves. *J Proteomics* 74:127–136
72. Chait BT (2011) Mass spectrometry in the post-genomic era. *Annu Rev Biochem* 80:239–246
73. Bandeira N, Nesvizhskii A, McIntosh M (2011) Advancing next-generation proteomics through computational research. *J Proteome Res* 10:2895
74. Gibbs PE, Strongin KB, McPherson A (1989) Evolution of legume seed storage proteins—a domain common to legumins and vicilins is duplicated in vicins. *Mol Biol Evol* 6:614–623
75. Esen A (1990) An immunodominant site of gamma-zein1 is in the region of tandem hexapeptide repeats. *J Protein Chem* 9:453–460
76. Feeney KA, Wellner N, Gilbert SM et al (2003) Molecular structures and interactions of repetitive peptides based on wheat glutenin subunits depend on chain length. *Biopolymers* 72: 123–131
77. Stevenson SE, Chu Y, Ozias-Akins P et al (2009) Validation of gel-free, label-free quantitative proteomics approaches: applications for seed allergen profiling. *J Proteomics* 72:555–566
78. Huang Y, Houston NL, Tovar-Mendez A et al (2010) A quantitative mass spectrometry-based approach for identifying protein kinase-clients and quantifying kinase activity. *Anal Biochem* 402:69–76
79. Lundgren DH, Hwang S-I, Wu L et al (2010) Role of spectral counting in quantitative proteomics. *Expert Rev Proteomics* 7:39–53
80. Farley AR, Link AJ (2009) Identification and quantification of protein posttranslational modifications. *Methods Enzymol* 463:725–763
81. Siuti N, Kelleher NL (2007) Decoding protein modifications using top-down mass spectrometry. *Nat Methods* 4:817–821
82. Agrawal GK, Thelen JJ (2009) A high-resolution two dimensional gel- and Pro-Q DPS-based proteomics workflow for phosphoprotein identification and quantitative profiling. *Meth Mol Biol* 527:3–19
83. Lewandowska-Gnatowska E, Johnston M, Antoine W et al (2011) Using multiplex-staining to study changes in the maize leaf phosphoproteome in response to mechanical wounding. *Phytochemistry* 72:1285–1292
84. Nesvizhskii AI (2012) Computational and informatics strategies for identification of specific protein interaction partners in affinity purification mass spectrometry experiments. *Proteomics* 12:1639–1655
85. Gache V, Waridel P, Winter C et al (2010) Xenopus meiotic microtubule-associated interactome. *PLoS One* 5:e9248
86. Sharon M (2010) How far can we go with structural mass spectrometry of protein complexes? *J Am Soc Mass Spectrom* 21:487–500
87. Heeren RM, Ron MA, Smith DF et al (2009) Imaging mass spectrometry: hype or hope? *J Am Soc Mass Spectrom* 20:1006–1014
88. Stauber J, MacAleece L, Franck J et al (2010) On-tissue protein identification and imaging by MALDI-ion mobility mass spectrometry. *J Am Soc Mass Spectrom* 21:338–347
89. Lee YJ, Perdian DC, Song Z et al (2012) Use of mass spectrometry for imaging metabolites in plants. *Plant J* 70:81–95
90. Sparvero LJ, Amoscato AA, Dixon CE et al (2012) Mapping of phospholipids by MALDI imaging (MALDI-MSI): realities and expectations. *Chem Phys Lipids* 165:545–562
91. Schober Y, Guenther S, Spengler B et al (2012) High-resolution matrix-assisted laser desorption/ionization imaging of tryptic peptides from tissue. *Rapid Commun Mass Spectrom* 26:1141–1146

Chapter 27

Back to Osborne. Sequential Protein Extraction and LC-MS Analysis for the Characterization of the Holm Oak Seed Proteome

M. Cristina Romero-Rodríguez, Ana M. Maldonado-Alconada, Luis Valledor, and Jesus V. Jorrin-Novo

Abstract

It is impossible to capture in just one experiment all or most of the total set of protein species that constitute the cell's proteome. Thus, according to our results, and even considering that they depend on the experimental system carried out (plant, yeast, fungi, or bacteria), the best protein extraction protocol yielded less than 20 % of the total amount of proteins, as determined by the Kjeldahl method. For this reason, protein cataloguing and the whole proteome characterization require the use of firstly, fractionation techniques at the cellular, subcellular, protein, or peptide level, and secondly, the use of complementary approaches.

Within our current research on Holm oak (*Quercus ilex* subsp. *ballota*), we aim to characterize its seed proteome. For that we have optimized an experimental workflow in which the Osborne sequential protein extraction (Osborne, *Science* 28:417–427, 1908) is combined with downstream electrophoretic protein separation or shotgun MS analysis. In general, it can be used to study any plant seed, as well as to investigate on seed maturation and germination, genotype characterization, allergens identification, food traceability, and substantial equivalence, among others.

Key words Holm oak proteomics, Seed proteomics, Osborne

1 Introduction

Osborne (1924) classified seed proteins into four categories according to their solubility in various solutions: (1) albumins, soluble in water, which primarily correspond to enzymes; (2) globulins, insoluble in water but soluble in salt solutions (i.e., 1 M NaCl); (3) prolamins, soluble in ethanol; and (4) glutelins, which can be extracted with either very acidic or very basic solutions [1, 2].

The major seed storage proteins are albumins, globulins, and prolamins. All storage protein fractions are mixtures of components that exhibit polymorphism both within single genotypes and among genotypes of the same species according to species [3]. These protein classes are not present systematically in the seeds of all

plant species. For example, prolamins exist only in the herbaceous plant species and cereals as the major storage proteins, whereas only fractions of albumin, globulin, and glutelin are found in sugarbeet [4].

Eighteen years after the Osborne work, in 1994, Marc R. Wilkins coined the term proteome [5, 6]; this being the beginning of a paradigm shift in the protein studies from protein chemistry/biochemistry to proteomics. This change was possible thanks to the progress in genome sequencing projects, the applicability of mass spectrometry to the analysis of peptide and proteins, and the development of bioinformatic tools for protein identification and data analysis. The potential of proteomics in plant biology research is extensively validated with the number of original papers and revisions, some of them dealing with plant seeds, devoted to different purposes: descriptive studies, cataloguing of genotypes, phylogenetic studies, seed germination, development, and responses against biotic and abiotic stresses [7–12]. The full potential of proteomics, however, is far from being fully exploited in biological research, due to the inherent difficulty of research with proteins, together with techniques and equipment limitations. We should highlight that the number of protein species is much higher than the number of genes, they are quite different physicochemically from each other, its dynamic range is of up nine orders of magnitude, and there is no possibility of being amplified, as it occurs with DNA. For all the above reasons, it is impossible to capture in just one experiment all or most of the total set of protein species that constitute the cell's proteome. Thus, according to our results, and considering that they depend on the experimental system carried out (plant, yeast, fungi, or bacteria), the best protein extraction protocol yielded less than 20 % of the total amount of proteins as determined by the Kjeldahl method [13]. For this reason, the proteome characterization and the protein species cataloguing require the use of firstly fractionation techniques at the cellular, subcellular, protein or peptide level; and secondly, the use of complementary approaches.

We have optimized for Holm oak seeds an experimental workflow in which the Osborne sequential protein extraction is combined with downstream electrophoretic separation or shotgun MS analysis. Up to now, most analyses of plant seed proteomic have been based on a total protein extract, and also with some examples of a similar strategy [4, 14–17].

In our current project on *Quercus ilex* subsp. *ballota*, we aim to obtain a full description of the seed proteome, in order to: (1) search for proteins of practical interest (i.e., enzymes); (2) identify possible allergens; (3) characterize the acorn nutritional quality; (4) catalog populations; and (5) carry out studies of seed germination and responses to stresses. The optimized workflow is schematized in Fig. 1.

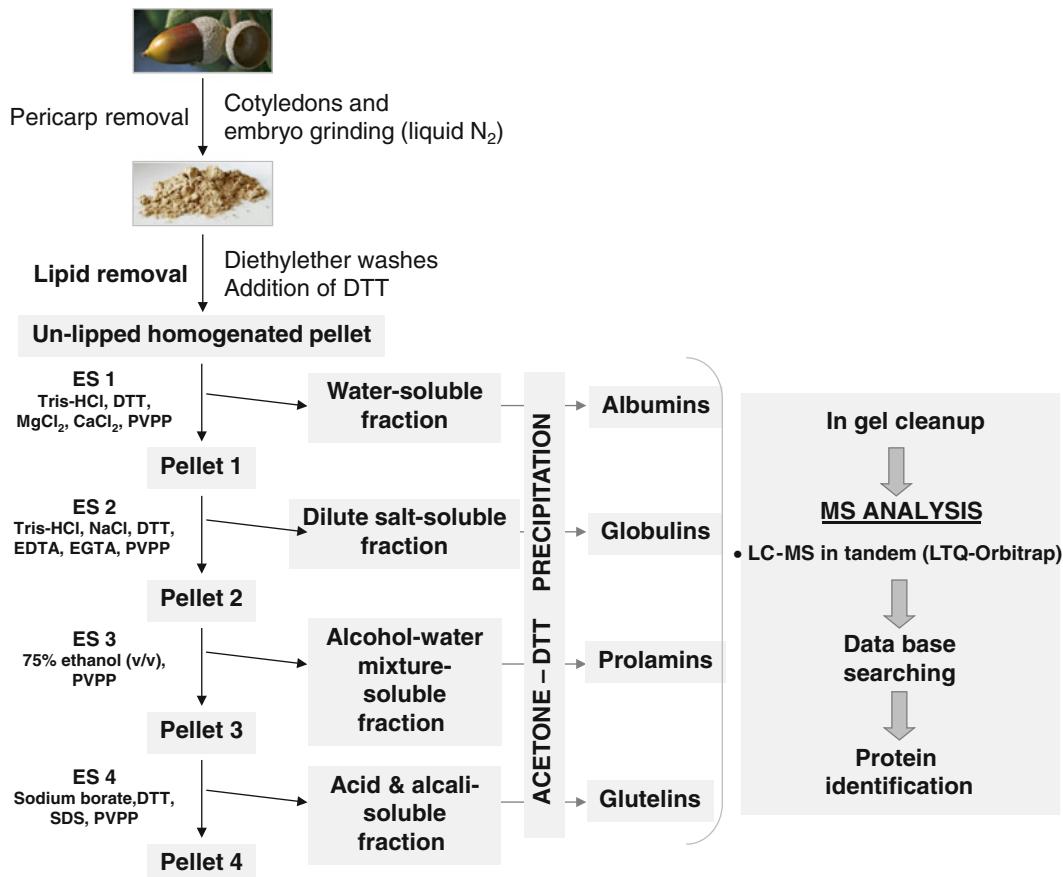


Fig. 1 Flowchart of the protein extraction procedure and MS analysis

2 Materials

2.1 Equipment (See Note 1)

1. Analytical Mill, Model A-10 (115 V), Janke & Kunkel IKA Labortechnik.
2. Mortar, pestle, and knife.
3. *Ultrasonic Homogenizer 4710 series*, Cole-Parmer Instrument Co.
4. Refrigerated centrifuge Model J2-21; Beckman.
5. Microcentrifuge Model 5415 D, Eppendorf.
6. BioPhotometer, Model 22331, Eppendorf.
7. Vertical Electrophoresis equipment. Mini PROTEAN II, Bio Rad.
8. Power supply, PowerPac HC, Bio-Rad

2.2 Reagents and Solutions

Prepare all solutions in distilled water (conductivity lower than $3 \mu\text{S}\cdot\text{cm}^{-1}$). Store all solutions at 4°C (unless otherwise indicated).

1. Extraction solution for albumins (ES1): 10 mM Tris–HCl, pH 7.5, 0.1 % (w/v) DTT, 10 mM MgCl₂, 10 mM CaCl₂.
2. Extraction solution for globulins (ES2): 10 mM Tris–HCl, pH 7.5, 1 M NaCl, 0.1 % (w/v) DTT, 10 mM EDTA, 10 mM EGTA.
3. Extraction solution for prolamins (ES3): 75 % (v/v) ethanol.
4. Extraction solution for glutelins (ES4): 50 mM sodium borate buffer, pH 10, 2 % (w/v) DTT, 1 % (w/v) SDS. Store this solution at room temperature.
5. Solubilization solution: 7 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 2 % (v/v) Tritón-X-100, 100 mM DTT. Add DTT to 10 mM just before use.
6. Bradford reagent (Sigma-Aldrich).
7. SDS-PAGE running buffer: 25 mM Tris–HCl, pH 8, 192 mM glycine, 0.1 % (w/v) SDS.
8. Acrylamide/Bis-acrylamide solution (30 %) commercial solution (Bio-Rad).
9. Ammonium persulfate: 10 % (w/v).
10. SDS: 10 % (w/v).
11. Resolving gel buffer: 1.5 M Tris–HCl, pH 8.8.
12. Stacking gel buffer: 0.5 M Tris–HCl pH 6.8.
13. Colloidal Coomassie staining gel: 10 % (w/v) ammonium sulfate, 3 % (w/v) phosphoric acid, 0.1 % (w/v) Coomassie Brilliant Blue (CBB) G250, 20 % (v/v) methanol.
14. Tris-phosphoric solution: 0.1 M Tris-phosphoric acid, pH 6.5.
15. Ammonium sulfate: 20 % (w/v).
16. Methanol: 25 % (v/v).
17. TEMED, commercial product (Sigma) (*see Note 2*).

3 Methods

3.1 Plant Material

Holm oak mature acorns (with brown pericarp) were used in our research. Once harvested, acorns were selected, cleaned and stored in airtight polyethylene bags at $4 \pm 1^\circ\text{C}$. Acorns were scarified with a knife by making transversal and longitudinal cuts, thus permitting the pericarp to be rapidly removed. Once peeled off, cotyledons (including embryo axes) were triturated in an analytical mill to obtain a homogeneous fine flour (Fig. 1). The obtained powder was weighed and stored in a desiccator at $4 \pm 1^\circ\text{C}$, until protein extraction [18].

3.2 Protein Extraction (See Note 3)

3.2.1 Sample Delipidation

1. Grind acorn flour (10 g) in liquid nitrogen with a precooled mortar and pestle until a fine powder is formed (see Note 4).
2. Add 60 mL of diethylether and shake the mixture for 1 h at 4 °C (see Note 5).
3. Centrifuge at 15,000×*g* for 15 min (4 °C) and decant the supernatant (see Note 6).
4. Repeat steps 2 and 3 two more times.
5. Add 60 mL of acetone containing 0.07 % (w/v) DTT to the above pellet, and shake the mixture for 10 min at 4 °C (see Note 7).
6. Centrifuge at 15,000×*g* for 15 min (4 °C) and decant the supernatant (see Note 8).
7. Allow the pellet to dry at room temperature until no traces of acetone remained.

3.2.2 Solubilization of the Albumin Fraction

1. Add 50 mL of ES1 and 1 g of PVPP to the pellet obtained in the previous step.
2. Sonicate the mixture (while keeping on ice) for 15 s (50 W, amplitude 60), and shake in an orbital shaker for 2 h at 4 °C.
3. Centrifuge at 15,000×*g* for 15 min (4 °C) and transfer the supernatant to a new tube (see Note 9).
4. Precipitate proteins in the supernatant by adding four volumes (approximately 100 mL for each tube) of 0.07 % (w/v) DTT containing cold acetone, and keep 1 h at 20 °C.
5. Centrifuge at 15,800×*g* for 15 min (4 °C) and decant the supernatant.
6. Dry the pellet at room temperature and store it at -80 °C if not immediately used (see Notes 10 and 11).

3.2.3 Solubilization of the Globulin Fraction

1. Add 50 mL of ES2 to the pellet obtained in step 3 (Subheading 3.2.2).
2. Sonicate the mixture (while keeping on ice) for 15 s (50 W, amplitude 60), and shake in an orbital shaker for 2 h at 4 °C.
3. Centrifuge at 15,000×*g* for 15 min (4 °C) and transfer the supernatant to a new tube (see Note 9).
4. Repeat steps 4–6 of the albumin fraction solubilization to obtain the globulin fraction (see Note 12).

3.2.4 Solubilization of the Prolamin Fraction

1. Add 50 mL of ES3 to the pellet obtained in step 3 (Subheading 3.2.3).
2. Sonicate the mixture (while keeping on ice) for 15 s (50 W, amplitude 60), and shake in an orbital shaker for 2 h at 4 °C.

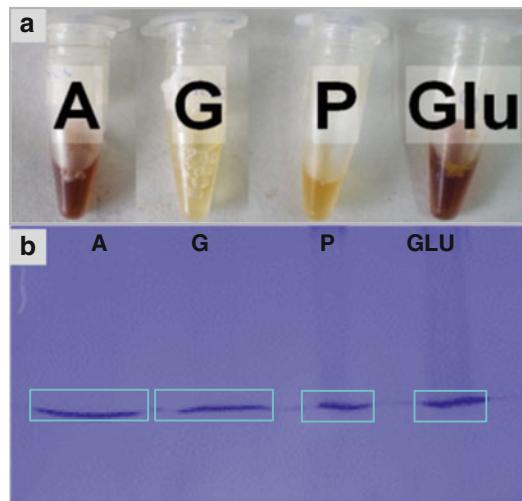


Fig. 2 (a) Aspects of each protein fraction obtained. *A* Albumins, *G* Globulins, *P* Prolamins, *Glu* Glutelins. (b) Gel precleaning. 100 μ g of each fraction (Albumins, *A*; Globulins, *G*; Prolamins, *P*; Glutelins, *Glu*) obtained with the protein were cleaned in a 5 % SDS-PAGE that was further stained with Coomassie Brilliant Blue. Labeled proteins bands were cut from the gel and analyzed by LC coupled to MS/MS

3. Centrifuge at 15,000 $\times g$ for 15 min (4 °C), and transfer the supernatant to a new tube (*see Note 9*).
4. Repeat **steps 4–6** of the albumin fraction solubilization to obtain the prolamin fraction.
1. Add 50 mL of ES4 to the pellet obtained in the **step 3** (Subheading [3.2.4](#)).
2. Sonicate the mixture (while keeping on ice) for 15 s (50 W, amplitude 60), and shake in an orbital shaker for 2 h at room temperature.
3. Centrifuge at 15,000 $\times g$ for 15 min (4 °C), and transfer the supernatant to a new tube (*see Note 9*).
4. Repeat **steps 4–6** of the albumin fraction solubilization to obtain the pellet corresponding to the glutelin fractions (*see Note 13*). (*see Note 14*)

3.3 Protein Solubilization

In all cases dried pellets resulting from each protein fraction precipitation were weighed and suspended in 7 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 2 % (v/v) Triton X-100, 100 mM DTT at a (w/v) ratio of 1:2 (*see Note 15*). Then, they were shaken for 1 h to facilitate protein solubilization. Once the pellet was solubilized and the insoluble material eliminated by centrifugation (*see Note 16*) (Fig. [2a](#)), the protein content was quantified by the Bradford methods [19], using bovine albumin as standard.

3.4 Protein Sample Preparation for LC-MS

One hundred µg of proteins corresponding to the different fractions were subjected to one-dimensional, denaturing SDS-PAGE according to Laemmli [20], with minor modifications. We used 7 cm stacking (5 % polyacrylamide) and 3 cm resolving (12.5 % polyacrylamide) gels, these run at 80 V in a Mini PROTEAN II cell (*see Note 17*) (Fig. 2b). Gel was stained by colloidal Coomassie [21].

Bands were excised using a clean scalpel, transferred to Multiwell 96 plates and digested with modified porcine trypsin by using a ProGest (Genomics Solution) digestion station. The digestion protocol used is that of Schevchenko et al. [22], with minor variations. Gel plugs were destained by incubation (twice for 30 min) with 200 mM ammonium bicarbonate in 40 % ACN at 37 °C, then subjected to three consecutive dehydration–rehydration cycles with pure acetonitrile and 25 mM ammonium bicarbonate in 40 % ACN, respectively; and, finally, dried at room temperature for 10 min. Trypsin, at a final concentration of 12.5 ng/µL in 25 mM ammonium bicarbonate, was added to the dry gel pieces and the digestion proceeded at 37 °C for 12 h. Peptides were extracted from gel plugs with 0.5 % TFA, 50 % ACN (15 min incubation), dried under vacuum and resuspended in 0.1 % TFA.

3.5 LC-MS and Protein Identification

Tryptic peptides (obtained according Subheading 3.4) were injected into the LC system (i.e., a Finnigan Surveyor HPLC system). Peptides were detected in a LTQ-Orbitrap equipped with a nanoelectrospray ion source (nESI). The acquired data can be analyzed with Proteome Discoverer v1.3 software (Thermo Fisher Scientific, USA) and MASCOT (<http://www.matrixscience.com>) or SEQUEST (<http://fields.scripps.edu/sequest/>) algorithms. Protein identification was conducted by combining search (MS plus MS/MS) to the entries of a nonredundant protein Viridiplantae database downloaded from the National Center for Biotechnology Information. Search parameters were the enzyme trypsin, taxonomy restrictions to Viridiplantae, ±50 ppm peptide mass tolerance in MS, ±0.8 Da for MS/MS data and, carbamidomethyl (C) as a fixed modification, whereas methionine oxidation as a variable modification (*see Note 18*).

3.6 Main Results

Considering the particular composition of the acorn flour, rich in non-protein material [23, 24], the main challenge was to obtain protein extracts of enough quality for the subsequent analyses and identification of proteins by MS. In Holm oak acorns, glutelin was the major fraction, corresponding approximately to 80 % of the total protein extracted with this protocol.

The analysis by LC coupled with MS/MS allowed the identification of 509 protein species, belonging to the following functional categories: storage, energy production and metabolism, defense, antioxidant system, signaling and gene expression regulation,

DNA and protein metabolism, cell wall biogenesis, cytoskeleton and, cellular cycle and translocation. The percentage of total proteins identified which were present in one fraction was: 25, 23, 15, and 23 % for the albumin, globulin, prolamin, and glutelin fractions, respectively. We have established the first *Q. ilex* acorn reference proteome that can be used in subsequent analysis, namely, the identification of protein markers to discriminate between populations, studies on the germination process, etc.

The main limitation of this approach is that the genus *Quercus* is an orphan plant species poorly characterized at the genetic and physiological levels. The low number of well annotated genomic sequences available for genus *Quercus*, and for *Q. ilex* in particular, in public databases limits the confident protein identification. Moreover, most software tools have been developed to search in a non-error tolerant way against a database of known proteins [10, 23]. A higher number of identifications which are more accurate would be obtained by searching against EST sequences [25, 26]. These sequences can be obtained from the following EST public database (<http://compbio.dfci.harvard.edu/tgi/>; <http://www.fagaceae.org/>; <http://www.ncbi.nlm.nih.gov/nuclest/>); currently, our group is constructing an EST database for *Quercus*.

4 Notes

1. References to specific companies, equipment or pieces of equipment are not mandatory and do not represent an endorsement by the authors
2. Standard reagents and items of current use, such as Bovine serum albumin (BSA), liquid nitrogen and polyvinylpyrrolidone, are not specified.
3. Protein extraction and fractionation procedures were based on the differential solubility criteria described by Osborne (1924), and later adapted to *Quercus* subsp. by Fonseca et al. (1997) and Martin et al. (2009) [2, 27, 28].
4. The extraction is performed with 10 g of fresh tissue because fractions of albumins, globulins and prolamins are minority proteins, with 1 g of tissue the small amount of protein isolated in the mentioned fraction is not enough to analyse.
5. This step is to reduce lipid content, which allows the obtaining of a higher quality extract for subsequent analysis.
6. Be careful that the pellet does not stick tightly to the bottom of the tube.
7. This step is to remove traces of diethylether.

8. Remove the supernatant with a pipette to avoid losing the pellet. Be careful: do not move the pellet because it is very easy to disintegrate
9. Remove the supernatant with a pipette to avoid disaggregating the pellet and carry contaminant.
10. The pellet of the albumins fraction appearance may be viscous; this aspect is normal, indicating that it may contain the remainders of lipids and polysaccharides [29].
11. Dry the pellet to ensure that the acetone was completely removed, if it lasts for a long time is difficult to resuspend the protein.
12. The pellet of the globulins fraction has two phases, namely, one solid and another oily. Remove the last one and use the solid phase to resuspend proteins.
13. The glutelins solubilization is performed at room temperature to prevent the insolubilization of SDS.
14. Steps 1–3 of each fraction solubilization should be repeated in order to make sure that the maximum amount of proteins corresponding to each fraction are solubilized.
15. The volume of buffer to resuspend proteins depends on the amount of precipitate obtained. It is recommended that samples are well concentrated.
16. The insoluble material is abundant in the case of the glutelin fraction.
17. The extracts are cleaned in this step, removing contaminants that may interfere with the identification by MS such a salt, detergents, etc.
18. For more details you can consult the following review [30].

Acknowledgments

This work has been financed by the Spanish Ministry of Innovation and Science, project AGL2009-12243-C02-02 “Variability, cataloguing, responses to stresses and clonal propagation of holm oak (*Quercus ilex* subsp. *ballota*) (DECOVA)” (FEDER Co-financed) and the Andalusian Government (Junta de Andalucía). MC Romero Rodríguez has been supported by a grant from the “Fundación Parque Tecnológico Itaipu-Binacional” and “Centro Multidisciplinario de Investigaciones Tecnológicas - Dirección General de Investigación Científica y Tecnológica-Universidad Nacional de Asunción - Paraguay”.

References

1. Osborne TB (1908) Our present knowledge of plant proteins. *Science* 28:417–427
2. Osborne TB (1924) The vegetable proteins. Logmans, London
3. Shewry PR, Napier JA, Tatham AS (1995) Seed storage proteins: structures and biosynthesis. *Plant Cell* 7:945–956
4. Catusse J, Pestsova E, Bechettoile C et al (2007) Betaomics: a combined proteome and transcriptome profiling approach to characterize seed germination and vigour in sugarbeet seeds. In: Adkins SW, Ashmore S, Navie SC (eds) *Seeds: biology, development and ecology*. CAB International, London, pp 169–178
5. Wilkins M (2009) Proteomics data mining. *Expert Rev Proteomics* 6:599–603
6. Wilkins MR, Pasquali C, Appel RD et al (1996) From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Nat Biotechnol* 14:61–65
7. Catusse J, Rajjou L, Job C et al (2008) Proteome of Seed Development and Germination. In: Agrawal G, Rakwal R (eds) *Plant Proteomics: Technologies, Strategies, and Applications*. John Wiley & Sons, Inc., pp 191–206
8. González-Fernández R, Prats E, Jorrín-Novo JV (2010) Proteomics of plant pathogenic fungi. *J Biomed Biotechnol*. doi:[10.1155/2010/932527](https://doi.org/10.1155/2010/932527)
9. Hossain Z, Nouri MZ, Komatsu S (2012) Plant cell organelle proteomics in response to abiotic stress. *J Proteome Res* 11:37–48
10. Jorrín-Novo JV, Maldonado AM, Echevarría-Zomeño S et al (2009) Plant proteomics update (2007–2008): second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. *J Proteomics* 72:285–314
11. Jorrín-Novo JV, Maldonado AM, Castillejo MA (2007) Plant proteome analysis: a 2006 update. *Proteomics* 7:2947–2962
12. Rossignol M, Peltier JB, Mock HP et al (2006) Plant proteome analysis: a 2004–2006 update. *Proteomics* 6:5529–5548
13. Valero-Galván J, Jorrín-Novo JV, Cabrera A et al (2012) Population variability based on the morphometry and chemical composition of the acorn in Holm oak (*Quercus ilex* subsp. *ballota* [Desf.] Samp.). *Eur J For Res* 131:893–904
14. Agrawal GK, Hajduch M, Graham K et al (2008) In-depth investigation of the soybean seed-filling proteome and comparison with a parallel study of rapeseed. *Plant Physiol* 148: 504–518
15. Hajduch M, Ganapathy A, Stein JW et al (2005) A systematic proteomic study of seed filling in soybean. establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. *Plant Physiol* 137:1397–1419
16. Krishnan HB, Oehrle NW, Natarajan SS (2009) A rapid and simple procedure for the depletion of abundant storage proteins from legume seeds to advance proteome analysis: a case study using *Glycine max*. *Proteomics* 9:3174–3188
17. Miernyk JA, Hajduch M (2011) Seed proteomics. *J Proteomics* 74:389–400
18. Valero-Galván J, Valledor L, Cerrillo RM et al (2011) Studies of variability in Holm oak (*Quercus ilex* subsp. *ballota* [Desf.] Samp.) through acorn protein profile analysis. *J Proteomics* 74:1244–1255
19. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 7: 248–254
20. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685
21. Neuhoff V, Arbold N, Taube D et al (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9:255–262
22. Schevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* 68:850–858
23. Abril N, Gion JM, Kerner R et al (2011) Proteomics research on forest trees, the most recalcitrant and orphan plant species. *Phytochemistry* 72:1219–1242
24. Bonner FT, Vozzo JA (1987) Seed biology and technology of *Quercus*. *Gen Tech Rep* SO-66. New Orleans
25. Choudhary JS, Blackstock WP, Creasy DM et al (2001) Matching peptide mass spectra to EST and genomic DNA databases. *Trends Biotechnol* 19:S17–S22
26. Edwards NJ (2007) Novel peptide identification from tandem mass spectra using ESTs and sequence database compression. *Mol Syst Biol* 3:102. doi:[10.1038/msb4100142](https://doi.org/10.1038/msb4100142)

27. Fonseca PA, Ferreira RB, Teixeira AR (1997) Seed proteins from *Quercus suber*. *J Agric Food Chem* 45:3443–3447
28. Martin MA, Navarro-Cerrillo RM, Ortega P et al (2009) The use of cotyledon proteins to assess the genetic diversity in sweet holm oak. *J For Sci* 55:526–531
29. Görg A, Weiss W, Dunnet MJ (2004) Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 4:3665–3685
30. Schulze WX, Usadel B (2010) Quantitation in mass-spectrometry-based proteomics. *Annu Rev Plant Biol* 61:491–516

Chapter 28

Xylem Sap Proteomics

Thomas Dugé de Bernonville, Cécile Albenne, Matthieu Arlat, Laurent Hoffmann, Emmanuelle Lauber, and Elisabeth Jamet

Abstract

Proteomic analysis of xylem sap has recently become a major field of interest to understand several biological questions related to plant development and responses to environmental clues. The xylem sap appears as a dynamic fluid undergoing changes in its proteome upon abiotic and biotic stresses. Unlike cell compartments which are amenable to purification in sufficient amount prior to proteomic analysis, the xylem sap has to be collected in particular conditions to avoid contamination by intracellular proteins and to obtain enough material. A model plant like *Arabidopsis thaliana* is not suitable for such an analysis because efficient harvesting of xylem sap is difficult. The analysis of the xylem sap proteome also requires specific procedures to concentrate proteins and to focus on proteins predicted to be secreted. Indeed, xylem sap proteins appear to be synthesized and secreted in the root stele or to originate from dying differentiated xylem cells. This chapter describes protocols to collect xylem sap from *Brassica* species and to prepare total and N-glycoprotein extracts for identification of proteins by mass spectrometry analyses and bioinformatics.

Key words Cell wall, Glycoproteomics, Xylem sap

1 Introduction

During land colonization by plants, about 430 million years ago, the multicellular photosynthetic organisms had to undergo a series of evolutionary events to establish the conditions for successful colonization of new environments like improvement of photosynthetic activity and absorption of water and nutriments from the soil. As a consequence, to interconnect all organs and to remain hydrated, terrestrial plants acquired conducting specialized tissues [1]. The vascular network of seed plants is organized in bundles and is composed of two major tissue types, xylem and phloem, originating from the procambium, the vascular meristem. Both tissues typically comprise specialized cell types, including conducting elements, and cells that fulfill nourishing (parenchyma) and mechanical strengthening (fibers) functions. The conducting tracheary elements of xylem allow the hydration of the plant body

by connecting the aerial shoots to the subterranean roots that take up water and mineral salts [2]. The long-distance movement throughout the entire organism of this aqueous fluid, called xylem sap, mainly results from root pressure and leaf water evapotranspiration. The tracheary element differentiation is ending by cell death, leaving a protoplast-free hollow tube surrounded by a lignified secondary cell wall [3]. Then, the xylem sap is considered as part of the apoplast and the xylem vessels form a “super-apoplast” structure allowing supra-cellular transport processes. Some plant pathogens borrow the “super-apoplast” route to move from the primary infection site and colonize distant tissues. Besides, xylem orchestrates the allocation of nutrients by supplying amino acids, organic acids, and carbohydrates within the organism [4]. Furthermore, the xylem network participates in the regulation of plant development by providing long distance transport of hormones like cytokinin and abscissic acid.

The origin and the function of the xylem proteins became the focus of several investigations using proteomic analyses. The xylem sap proteome was investigated in different plant genera, including *Brassica* [5–7], *Cucurbita* and *Cucumis* [5, 8], *Glycine* [9, 10], *Malus* [11], *Oryza* [12], *Populus* [13], *Vitis* [14, 15] and *Zea* [16]. Different strategies were used for protein separation and identification by mass spectrometry (MS) (Table 1). In most cases, proteins were separated by standard 1D- (1D-E) or 2D-electrophoresis (2D-E) prior to tryptic digestion and MS or LC (liquid chromatography)-MS analysis. In one case, peptides were directly analyzed by 2D-LC MS/MS without any protein separation by electrophoresis [12]. In another case, a short run of 1D-E was performed to get only three samples to analyze by LC-MS/MS [7]. LC-MS/MS was used in most cases since the nucleotide sequence data (ESTs or genomic) are scarce for most of the studied plants. The xylem sap proteomes of only four plants were obtained with homologous sequences: *Brassica oleracea* [7], *Glycine max* [9, 10, 22], *Oryza sativa* [12], and *Solanum lycopersicum* [20]. In the other cases, heterologous sequences were used for protein identification. This strategy did not allow the precise identification of genes encoding the identified proteins especially in the case of multigene families.

Subtilisin Ser-proteases, oxido-reductases (among which peroxidases), lipases, and enzymes involved in carbohydrate metabolism are the most represented protein families in analyzed xylem saps. Many xylem sap proteins were found to be secreted [7, 13]. The proportion of proteins predicted to be intracellular may reach one third of the total xylem sap proteome in *Populus* [13], but is less important (<15 %) in annual plants like *Brassica* [6, 7] and *Z. mays* [16]. Intracellular proteins identified in xylem sap could be released during the programmed cell death of precursor cells involved in xylem differentiation. Several of the so-called pathogenesis-related (PR) proteins (thaumatin-like, chitinases)

Table 1
Some representative xylem sap proteomes

| Plant | Reference | Total number of identified proteins ^a | Number of proteins predicted to have a signal peptide ^b | Protein separation ^c /MS strategy |
|---|-----------|--|--|--|
| <i>B. napus</i> | [6] | 40 | 39 | 2D-E LC-MS/MS (ESI-Q-TOF) |
| <i>B. napus</i> | [5] | 10 | 10 | 1D-E LC-MS/MS (ESI-Q-TOF) |
| <i>B. oleracea</i> | [7] | 189 | 164 | Short 1D-E run (three samples) LC-MS/MS (LTQ XL ion trap) |
| <i>B. oleracea</i> | [17] | 24 | 14 | 2D-E LC-MS/MS (LTQ-Orbitrap) |
| <i>B. oleracea</i> | [5] | 10 | 10 | 1D-E LC-MS/MS (ESI-Q-TOF) |
| <i>Cucurbita maxima</i> | [5] | 11 | 11 | 1D-E LC-MS/MS (ESI-Q-TOF) |
| <i>Cucumis sativus</i> | [5] | 9 | 9 | 1D-E LC-MS/MS (ESI-Q-TOF) |
| <i>G. max</i> | [9] | 24 | 18 | 2D-E LC-MS/MS (LCQ-DECA XP Plus ion trap) |
| <i>G. max</i> | [10] | 14 | 11 | 1D-E LC-MS/MS (LTQ-Orbitrap) |
| <i>G. max</i> | [22] | 17 | 11 | 2D-E LC-MS/MS (ESI-Q-TOF) |
| <i>O. sativa</i> | [12] | 118 | 71 | 1D-E LC-MS/MS (ion trap) 2D-LC-MS/MS (ion trap) |
| <i>P. trichocarpa</i> x <i>P. deltoides</i> | [13] | 97 | 33–24 (NC) | 1D-E LC-MS/MS (ESI-Q-TRAP) |
| <i>S. lycopersicum</i> | [19] | 5 | 5 | 1D-E MALDI-TOF MS |
| <i>S. lycopersicum</i> | [20] | 20 | 20 | 2D-E MALDI-TOF MS LC-MS/MS (ESI-Q-TOF) |
| <i>Z. mays</i> | [16] | 59 | 58–1 (NC) | 2D-E LC-MS/MS (ESI-Q-TOF) |
| <i>Z. mays</i> | [18] | 26 | 22 | 2D-E LC-MS/MS (ESI-Q-TOF) |

^aOnly non redundant proteins are considered. Numbers in bold correspond to proteins identified using EST or genomic data of the studied plant

^bSignal peptides are predicted using TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) or SignalP (<http://www.cbs.dtu.dk/services/SignalP/>). When available, the number of nonclassically secreted proteins (NC) determined using SecretomeP is indicated (<http://www.cbs.dtu.dk/services/SecretomeP/>)

^c1D-E and 2D-E respectively mean one- and two-dimensional electrophoresis

were found in xylem sap proteomes of stressed plants: (1) during abiotic stress in *B. oleracea* (salt stress) [17] and *Z. mays* (drought stress) [18], (2) in *Fusarium*-infected *S. lycopersicum* [19, 20] and *Glycine* [21] and (3) during nitrogen-fixing symbiosis in soybean [22]. Several PR-proteins were also identified constitutively in *Populus* [13], *Glycine* [9], *Brassica*, *Cucumis* and *Cucurbita* [5]. Interestingly, analysis of the *N*-glycoproteome of the *B. oleracea* xylem sap indicates that *N*-glycosylation is a major post-translational modification of xylem sap proteins since near half of the proteins was found to display *N*-glycosylations [7]. Taken together, these data indicate that xylem sap is a dynamic fluid undergoing specific changes in response to environmental challenges. This emphasizes the need to analyze the xylem sap proteome to better understand its physiology.

However, studying the xylem sap proteome often encounters methodological problems. Since protein concentrations are quite low in xylem sap, relatively important volumes must be collected to recover enough material for identification of proteins by mass spectrometry and bioinformatics. Xylem sap collection is often achieved through root pressure-driven exudation. This method is by far more suitable for non woody plant families like *Brassicaceae* or *Cucurbitaceae*, since their stem diameter is large enough to collect important volumes of xylem sap. It should be noted that the small model plant *Arabidopsis* is unsuitable because of inefficient xylem sap harvesting. To enhance xylem sap collection, an external pressure may be applied to the roots to shorten sampling times, e.g. with a Scholander chamber or by applying a hydraulic pressure [10, 17, 18]. The cutting of the plant stem and eventually the application of an external pressure may lead to excessive intracellular protein release. However, this artifact is difficult to avoid and homogeneous sap collection methods may help the comparison of xylem sap proteomes between plants. Collection methods on woody plants like *Populus* and *Vitis* require more manipulations, like external bark removal to avoid phloem contamination [13, 15].

In this chapter, methods to collect xylem saps from herbaceous plants and to analyze their proteomes will be described. Cabbage will be taken as an example.

2 Materials

All materials should be carefully cleaned before use to avoid contamination with keratins which can prevent protein identification using MS. Gloves are required. Chemicals should be for proteomic use only. PlusOne™ (GE Healthcare, Orsay, France) products should be used for SDS-PAGE. All solutions and buffers should be filtered through 0.22 µm pore size filters. Glass plates for electrophoresis should be cleaned with ethanol.

2.1 Plant Production

1. *Brassica oleracea* var Bartolo F1 calibrated seeds treated with a fungicide (Bejo, France).
2. Universal compost for seedlings and cuttings, containing peat moss, sphagnum peat brown, clay, and perlite (catalog number A16805, AGRI Garonne, Castelginest, France).
3. Peat pots composed of 80 % wood fiber and 20 % peat (catalog number 53803, AGRI Garonne).
4. Universal fertilizer (Universel ALGOFLASH NPK 6/6/6).
5. Greenhouse 22–27 °C, 16 h of daylight, 10,000 lx, 55–75 % of relative humidity.

2.2 Xylem Sap Harvesting

1. Mini-greenhouses (Stewart) 52 × 42 × 24 cm.
2. Rasor blades.

2.3 Preparation of the Protein Extract

1. Dialysis tubes: Mega GeBAflex-tubes MWCO 12–14 kDa (catalog number D080-6/10-10, Gene Bio-Application L.T.D., Yavne, Israel) (see Note 1).
2. 1 M Tris–HCl pH 7.4 stock solution. Weigh 12.1 g of Tris (catalog number 26.128-3094-B, EUROMEDEX, Mundolsheim, France). Add water to a volume of 90 mL. Mix and adjust pH with HCl. Make up to 100 mL with water. Store at 4 °C.
3. 5 M NaCl stock solution. Weigh 29.2 g of NaCl. Add water to a volume of 100 mL. Store at 4 °C.
4. 10 mM MgCl₂/MnCl₂/CaCl₂ stock solution. Weigh 0.244 g MgCl₂ hexahydrate (catalog number M9272, Sigma-Aldrich, Saint-Louis, MO), 0.237 g MnCl₂ tetrahydrate (catalog number 221279, Sigma-Aldrich) and 0.176 g CaCl₂ dihydrate (catalog number 223506, Sigma-Aldrich) (see Note 2). Add water to a volume of 250 mL. Store at 4 °C.
5. Dialysis buffer B1: 20 mM Tris–HCl pH 7.4, 150 mM NaCl, 1 mM MgCl₂/MnCl₂/CaCl₂. Mix 20 mL of 1 M Tris–HCl pH 7.4, 30 mL of 5 M NaCl, 100 mL of 10 mM MgCl₂/MnCl₂/CaCl₂ stock solutions. Add water to a volume of 1 L (see Note 3). Store at 4 °C.
6. Desalting columns: Econo-Pac 10 DG columns (catalog number 732-2010, Bio-Rad, Hercules, CA)
7. Desalting buffer B2: 0.1 M ammonium formate. Weigh 6.3 g of ammonium formate (catalog number 156264, Sigma-Aldrich). Add water to a volume of 1 L. Store at 4 °C.
8. Bradford method reagent: Coomassie Protein assay Kit (catalog number UPF8640, Uptima, Interchim, Montluçon, France).
9. Lyophilizator Christ Alpha 1–4 LSC (Fischer Scientific, Illkirch, France).

2.4 Affinity chromatography on Concanavalin A

1. Concanavalin A (ConA) agarose conjugate (catalog number C7555, Sigma-Aldrich) (see Note 4).
2. ConA prewashing buffer B3: 20 mM Tris-HCl pH 7.4, 1 M NaCl, 3.3 mM MgCl₂/MnCl₂/CaCl₂. Mix 2 mL of 1 M Tris-HCl pH 7.4, 20 mL of 5 M NaCl, 33 mL of 10 mM MgCl₂/MnCl₂/CaCl₂ stock solutions. Add water to a volume of 100 mL. Store at 4 °C.
3. ConA equilibration and washing buffer B1: 20 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM MgCl₂/MnCl₂/CaCl₂. Mix 20 mL of 1 M Tris-HCl pH 7.4, 30 mL of 5 M NaCl, 100 mL of 10 mM MgCl₂/MnCl₂/CaCl₂ stock solutions. Add water to a volume of 1 L. Store at 4 °C.
4. ConA elution buffer B4: 20 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM MgCl₂/MnCl₂/CaCl₂, 1 M methyl- α -D-glucopyranose (catalog number M9376, Sigma-Aldrich). Weigh 9.7 g of methyl- α -D-glucopyranose. Add buffer B1 to a volume of 50 mL. Store at 4 °C.
5. Desalting columns: Econo-Pac 10 DG columns (catalog number 732-2010, Bio-Rad).
6. Desalting buffer B2: 0.1 M ammonium formate. Weigh 6.3 g of ammonium formate and add water to a volume of 1 L. Store at 4 °C.
7. Lyophilizator Christ Alpha 1-4 LSC (Fischer Scientific).

2.5 Separation of Proteins by SDS-PolyAcrylamide Gel Electrophoresis

1. Mini-PROTEAN® Tetra Cell (Bio-Rad).
2. Resolving gel buffer stock solution: 1.5 M Tris-HCl pH 8.8. Weigh 36.3 g of Tris (PlusOne™, catalog number 17-1321-01, GE Healthcare). Add water to a volume of 180 mL. Mix and adjust pH with HCl. Make up to 200 mL with water. Store at 4 °C.
3. Stacking gel buffer stock solution: 1.25 M Tris-HCl pH 6.8. Weigh 15 g of Tris (PlusOne™, catalog number 17-1321-01, GE Healthcare). Add water to a volume of 90 mL. Mix and adjust pH with HCl. Make up to 100 mL with water. Store at 4 °C.
4. PlusOne™ acrylamide 40 % solution (catalog number 17-1303-01, GE Healthcare).
5. PlusOne™ *N,N'*-methylenebisacrylamide 2 % solution (catalog number 17-1306-01, GE Healthcare).
6. 10% SDS solution. Weigh 5 g of SDS (PlusOne™, catalog number 17-1313-01, GE Healthcare). Add water to a volume of 30 mL. Mix until total dissolution (see Note 5). Complete to a final volume of 50 mL. Store at room temperature.
7. Ammonium persulfate (APS) 10 % solution. Weigh 1 g of APS (PlusOne™, catalog number 17-1311-01, GE Healthcare). Add water to a volume of 10 mL. Store 1 mL aliquots at -20 °C.

8. Tetramethyl-ethylenediamine (TEMED) (PlusOne™, catalog number 17-1312-01, GE Healthcare).
9. Resolving gel composition E1: 0.375 M Tris–HCl, pH 8.8, 0.1 % SDS, 12.5 %/0.33 % acrylamide/bis-acrylamide, 0.075 % APS, 0.5 % TEMED. For one gel, mix 3.04 mL of 40 % acrylamide solution, 1.67 mL of 2 % bisacrylamide solution, 2.5 mL of 1.5 M Tris–HCl pH 8.8, 100 µL of 10% SDS, 2.64 mL of water, 50 µL of 10% APS and 5 µL TEMED (*see Note 6*).
10. Stacking gel composition E2: 0.125 M Tris–HCl, pH 6.8, 0.1 % SDS, 4 %/0.11 % acrylamide/bis-acrylamide, 0.05 % APS, 1 % TEMED. For one gel (10×12×0.15 cm), mix 486 µL of 40 % acrylamide solution, 268 µL of 2 % bisacrylamide solution, 0.5 mL of 1.25 M Tris–HCl pH 6.8, 50 µL of 10 % SDS, 3.68 mL of water, 25 µL of 10 % APS and 5 µL TEMED (*see Note 6*).
11. Electrophoresis sample buffer E3: 0.062 M Tris–HCl, pH 6.8, 2 % SDS, 10% glycerol, 5 % β-mercaptoethanol. To prepare a 4× buffer, mix 4 mL of 1.25 M Tris–HCl pH 6.8, 0.8 g SDS (PlusOne™, catalog number 17-1313-01, GE Healthcare), 1.18 mL of 85 % glycerol solution (PlusOne™, catalog number 17-1325-01, GE Healthcare), 2 mL of β-mercaptoethanol (PlusOne™, catalog number 17-1317-01, GE Healthcare) and 0.25 mL of 0.05 % (w/v) bromophenol blue (catalog number B0126, Sigma-Aldrich).
12. Electrophoresis migration buffer E4: 0.025 M Tris, 0.192 M glycine, 0.1 % SDS. To prepare a 10× buffer, weigh 30.3 g of Tris (PlusOne™, catalog number 17-1321-01, GE Healthcare), 145 g of glycine (PlusOne™, catalog number 17-1323-01), 10 g of SDS (PlusOne™, catalog number 17-1313-01, GE Healthcare). Add water to a final volume of 1 L.
13. Colloidal blue gel staining: PageBlue™ Protein Staining Kit (catalog number R0571, Fermentas, Villebon sur Yvette, France).
14. Silver nitrate gel staining solutions:
 - Protein fixation solution S1: 45 % ethanol, 5 % acetic acid.
 - Washing solution S2: 30 % ethanol.
 - Sensitization solution S3: 0.02 % (w/v) thiosulfate ($\text{Na}_2\text{S}_2\text{O}_3 \cdot 5\text{H}_2\text{O}$) (catalog number S8503, Sigma-Aldrich).
 - Silver impregnation solution S4: 0.1 % (w/v) silver nitrate (NO_3Ag) (catalog number S0139, Sigma-Aldrich).
 - Washing solution S5: 2.5 % (w/v) Na_2CO_3 (catalog number 223530, Sigma-Aldrich).
 - Development solution S6: 0.04 % formaldehyde, 2.5 % (w/v) Na_2CO_3 (catalog number 223530, Sigma-Aldrich).
 - Staining arrest solution S7: 1 % acetic acid.
15. Gel storage solution S7: 1 % acetic acid.

2.6 Bioinformatic Analysis of Proteins Identified by MS

1. TargetP: <http://www.cbs.dtu.dk/services/TargetP/> [23].
2. Predotar: <http://urgi.versailles.inra.fr/predotar/predotar.html> [24].
3. PSORT: <http://psort.hgc.jp/form.html>.
4. GPIsom: <http://gpi.unibe.ch/> [25].
5. BigPI: http://mendel.imp.ac.at/sat/gpi/gpi_server.html [26].
6. PROSITE: <http://prosite.expasy.org/> [27].
7. PFAM: <http://pfam.sanger.ac.uk/> [28].
8. InterProScan: <http://www.ebi.ac.uk/Tools/pfa/iprscan/> [29].
9. BLAST: <http://blast.ncbi.nlm.nih.gov/Blast.cgi> [30].
10. CD-search: <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> [31].
11. ProTerNyc: <http://www.polebio.lrsv.ups-tlse.fr/ProTerNyc/> [32].

3 Methods

The overall strategy from xylem sap harvesting to protein identification by MS and bioinformatics analysis is schematized on Fig. 1.

3.1 Plant Production

1. Sow seeds on loam. Place pots in the greenhouse. After 15 days, repot seedlings individually in peat pots containing loam. Place pots in the greenhouse for 3–4 weeks.
2. Water plants regularly (*see Note 7*). Give fertilizer once a week.

3.2 Xylem Sap Harvesting (*See Note 8*)

1. Put 5–6 week-old plants (3–4 pairs of leaves) in mini-greenhouses placed in the lab at room temperature (*see Notes 9–10*) for one night. Pots must be saturated with water.
2. Cut stems horizontally above the first pair of leaves with a razor blade (*see Note 11*). Rinse thoroughly with water and dry with paper towels.
3. Collect regularly drops of xylem sap exuding on the surface of the cut with a Micropipette (Fig. 1) (*see Note 12*). Store sap at –20 °C if rapidly used or at –80 °C for long conservation.
4. Harvest during 10–12 h. Number each fraction. Discard the two first fractions and pool other fractions (*see Note 13*). Routinely, 0.7 mL/plant/day of xylem sap was collected.

3.3 Preparation of the Protein Extract

1. After harvesting of the xylem sap, dialyze the samples (about 20 mL) against buffer B1 in a Mega GeBAflex-tube. Dialysis is performed against a large volume of buffer B1 (100–1,000-fold that of the sample) for 3 h at 4 °C. Change dialysis buffer B1 and continue dialysis for another 3 h.

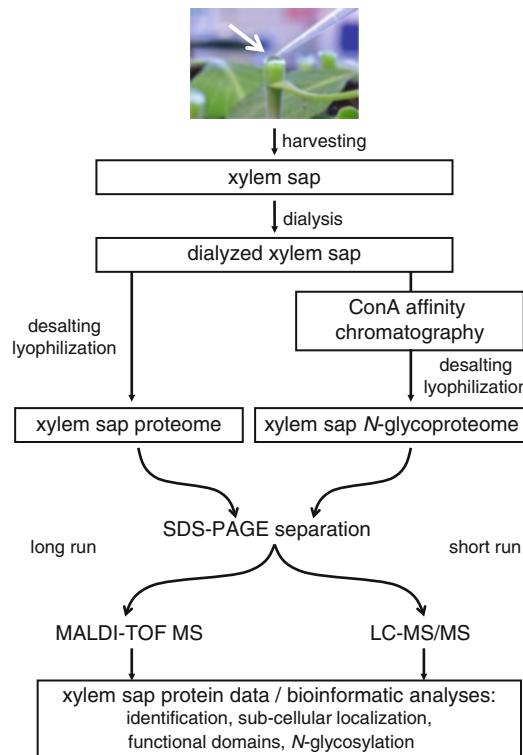


Fig. 1 Strategy for xylem sap proteomic analysis from harvesting of xylem sap to protein identification by mass spectrometry and bioinformatics analysis

2. Desalt half of this dialyzed sample using Econo-Pac 10 DG columns to get the “xylem sap proteome.” Equilibrate the columns with 20 mL of buffer B2. Load 3 mL of dialyzed xylem sap sample onto each column. Elute proteins with 4 mL of buffer B2. Pool elution fractions and lyophilize resulting desalting sample overnight.
3. Quantify proteins by the Bradford method following the Uptima Coo Protein assay Kit (*see Note 14*).
4. Use the second half of the dialyzed sample to get the “xylem sap N-glycoproteome” (*see Subheading 3.4*).

3.4 Affinity Chromatography on ConA (See Note 15)

1. Perform ConA chromatography in batch mode to capture xylem sap N-glycoproteins (*see Note 16*).
2. Prewash the resin (0.6 mL) with 20-fold volumes of buffer B3 (*see Note 17*) and equilibrate with tenfold volume of buffer B1. Mix the dialyzed xylem sap sample (10 mL) with the matrix in batch for 1 h at 4 °C under mild shaking. Wash the resin three times with 1.5 mL of buffer B1 after flow-through removal. Elute proteins with 1.5 mL of buffer B4. Repeat elution twice.

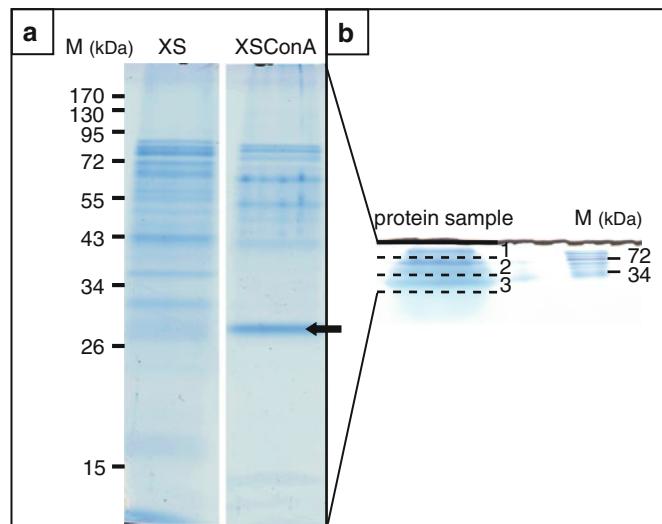


Fig. 2 Separation of proteins by SDS-PAGE prior to MS analysis. **(a)** A long run is performed prior to sampling for MALDI-TOF MS analysis. XS stands for xylem sap proteins and XSConA for xylem sap proteins eluted from a ConA affinity column. The arrow points at ConA leaking from the column. Each stained protein band is cut out of the polyacrylamide-SDS gel and prepared for MS analysis. **(b)** A short run is performed prior to sampling for LC-MS/MS analysis. The thick stained band is cut in three parts delimited by dotted lines (1–3) and each of them is prepared for MS analysis. The polyacrylamide gels are stained with colloidal blue. Molecular mass markers (M) are indicated

3.5 Separation of Proteins by SDS-PAGE

3. Combine first and second elution fractions. Desalt using Econo-Pac 10 DG columns and lyophilize as described (see Subheading 3.3, step 2). The sample obtained is the “xylem sap *N*-glycoproteome.”
1. Resuspend the two dried protein samples (“xylem sap proteome” and “xylem sap *N*-glycoproteome”) in 300 μ L and 100 μ L of UHQ water respectively. Load 50 μ L of each sample on a 10 \times 12 \times 0.15 cm SDS-polyacrylamide gel with stacking gel and resolving gel at a concentration of 4 %/0.11 % and 12.50 %/0.33 % of acrylamide/bisacrylamide respectively.
2. For MALDI-TOF (Matrix-Assisted Laser Desorption Ionization-Time Of Flight) MS analyses, perform separation by SDS-PAGE at 20 mA per gel for the stacking gel and at 40 mA per gel for the resolving gel, until the front line reaches the bottom of the gel (total migration time of about 75 min) (long run, Fig. 2a).
3. For LC (Liquid Chromatography)-MS/MS analyses, perform separation by SDS-PAGE at 20 mA per gel for the stacking and the resolving gels until the proteins migrate about 6 mm in

the resolving gel (total migration time of about 50 min) (short run, Fig. 2b) (see Note 18).

4. Stain the gels with colloidal blue or silver nitrate (see Note 19).
5. For colloidal blue staining, wash gel three times 15 min with water before staining overnight using the PageBlue™ Protein Staining solution. Destain with water. Store gels in storage solution S7.
6. For silver nitrate staining (see Note 20), incubate gels overnight in fixation solution S1, then rinse threefold 10 min in washing solution S2 and threefold 10 min in water. Perform sensitization for maximum 2 min using solution S3. Rinse the gels threefold 10 min in water. Carry out silver impregnation 30 min at 4 °C using solution S4. Rinse the gels threefold 10 min in water and once with washing solution S5. Achieve development using solution S6. Stop staining with solution S7. Store gels in storage solution S7.

3.6 Bioinformatic Analysis of Proteins Identified by MS

1. Predict subcellular localization with different software like TargetP, Predotar, and PSORT (see Note 21). Each of them is working with a different rationale. TargetP and Predotar are neural-based tools trained on experimentally proven signal sequences. PSORT is based on amino acid sequence properties, e.g. PSORT looks for a N-terminal basically-charged region and a central hydrophobic region to predict a signal sequences. Predict the presence of a GPI anchor site with GPIsom and BigPI software.
2. Predict functional domains with PROSITE and PFAM. PROSITE software can predict short patterns with a high probability of occurrence as well as motifs and patterns (see Note 22). PFAM predicts two types of domains: PFAM-A entries are of high quality and are manually curated. PFAM-B entries are automatically generated and only give indications when no PFAM-A domain is found. InterProScan collects the results of predictions of several software such as PROSITE and PFAM and proposes IPR superfamilies grouping domains predicted by these software. When no results are obtained with these tools, look for proteins sharing common sequences using the BLAST tool. In addition to proteins showing sequence homology to the protein of interest, BLAST also provides a library of conserved domains (CD).
3. Predict N-glycosylation sites and composition of N-glycans attached to the predicted sites using MALDI-TOF MS data with *ProTerNyc*. *ProTerNyc* uses (1) PROSITE, since N-glycosylation sites are short patterns of four amino acids (N-[P]-[ST]-[P], where N (Asn) is the N-glycosylation site) and (2) literature data describing the structures of N-glycans [33].

4 Notes

1. Mega GeBAflex-tubes were used for dialysis to reduce loss of material. Different tube caps can be used depending on the volume to dialyze (3–20 mL).
2. Due to their hygroscopic nature, $MgCl_2$ and $MnCl_2$ salts must be kept in closed bottles in a dry atmosphere.
3. Buffer B1 should be prepared from stock solutions because $MnCl_2$ might precipitate when powders are introduced all together.
4. ConA agarose resin is supplied as a suspension. Take a twofold volume of this suspension to get the volume required for chromatography.
5. When preparing the SDS solution, warm it in a 60 °C water bath to facilitate SDS dissolution. If precipitation occurs during storage, the same procedure should be carried out.
6. APS and TEMED being polyacrylamide polymerizing agents, they need to be introduced last during resolving and stacking gel preparation. Gels should be poured immediately.
7. Prevent drying of the soil. It should remain wet.
8. Avoid xylem sap harvesting during the summer. Plants can be stressed by high temperatures, especially cabbages. The quantity of sap collected is quite low.
9. Room temperature should be 20–25 °C.
10. Avoid placing the mini-greenhouses in ventilated areas.
11. The cut should be clean.
12. Collection takes place every 30 min to 1.5 h. Monitor and collect the drops before they fall off the cut surface. Quickly close the mini-greenhouses after each sampling to prevent drying of the sections.
13. Total hexose analysis suggests that the two first fractions are probably contaminated with phloem sap or cytoplasm of cut cells (T Dugé de Bernonville, unpublished data).
14. Protein content in the xylem sap protein extract can be too low to be quantified using the Bradford method. The same problem can be encountered after ConA chromatography. However, protein detection is possible after protein separation by SDS-PAGE and staining with colloidal blue or silver nitrate.
15. ConA lectin affinity chromatography is specific for Man residues and allows specific capture of *N*-glycoproteins [34]. This approach can be carried out for xylem sap proteomics since most proteins are *N*-glycosylated through the secretion pathway. It allows enrichment in minor proteins and consequently enlargement of the proteome coverage.

16. ConA chromatography is performed in batch mode to reduce material loss and to control the time of contact between the protein sample and the resin. After incubation with the different buffers, the resin is separated from the buffer by low speed centrifugation to avoid bound-protein loss.
17. A part of the ConA lectin conjugated to the agarose resin was shown to leak during chromatographic steps. Consequently, a pre-washing step is added to remove loosely bound ConA and limit the leakage of this lectin in elution fractions.
18. Proteins are briefly separated by SDS-PAGE to get three samples in order to concentrate proteins and to increase the efficiency of the tryptic digestion.
19. A silver nitrate staining can be carried out to detect low abundant proteins not revealed by colloidal blue staining.
20. A silver nitrate staining protocol compatible with mass spectrometry analyses need to be performed [35].
21. It is important to infer the subcellular localization of the proteins identified by subcellular proteomics. It gives information about the quality of the protein extract and allows a more reliable interpretation of experimental results. Whatever the care taken to perform the preparation of a fraction enriched in a subcellular compartment, there are always contaminant proteins originating from other cell compartments. A predicted subcellular localization should be validated only if the results of two predictions are consistent.
22. Using software dedicated to prediction of functional domains rather than BLAST homology search is usually more reliable. Indeed, there might be wrong annotations due to partial homology to a previously annotated protein [36].

Acknowledgments

The authors are grateful to the Université Paul Sabatier-Toulouse 3, France, CNRS, and INRA for support. A grant for the experiments and the postdoctoral position of TDDB were provided by the French *Agence Nationale de la Recherche* (Grant ANR-08-BLAN-0193-01). The authors wish to thank Drs Benoît Valot and Michel Zivy at the Plateforme d'Analyse Protéomique de Paris Sud-Ouest (*PAPPSSO*) for fruitful collaboration. Thibaut Douché is acknowledged for his contribution to proteomics developments in the lab.

References

1. Raven JA (1993) The evolution of vascular plants in relation to quantitative functioning of dead water-conducting cells and stomata. *Biol Rev* 68:337–363
2. Sattelmacher B (2001) The apoplast and its significance for plant mineral nutrition. *New Phytol* 149:167–192
3. Fukuda H (1996) Xylogenesis: initiation, progression, and cell death. *Annu Rev Plant Physiol Plant Mol Biol* 47:299–325
4. Satoh S (2006) Organic substances in xylem sap delivered to above-ground organs by the roots. *J Plant Res* 119:179–187
5. Buhtz A, Kolasa A, Arlt K et al (2004) Xylem sap protein composition is conserved among different plant species. *Planta* 219:610–618
6. Kehr J, Buhtz A, Giavalisco P (2005) Analysis of xylem sap proteins from *Brassica napus*. *BMC Plant Biol* 5:11
7. Ligat L, Lauber E, Albenne C et al (2011) Analysis of the xylem sap proteome of *Brassica oleracea* reveals a high content in secreted proteins. *Proteomics* 11:1798–1813
8. Masuda S, Kamada H, Satoh S (2001) Chitinase in cucumber xylem sap. *Biosci Biotechnol Biochem* 65:1883–1885
9. Djordjevic M, Oakes M, Li et al (2007) The *Glycine max* xylem sap and apoplast proteome. *J Proteome Res* 6:3771–3779
10. Krishnan H, Natarajan S, Bennett J et al (2011) Protein and metabolite composition of xylem sap from field-grown soybeans (*Glycine max*). *Planta* 233:921–931
11. Biles C, Abeles F (1991) Xylem sap proteins. *Plant Physiol* 96:597–601
12. Aki T, Shigyo M, Nakano R et al (2008) Nano scale proteomics revealed the presence of regulatory proteins including three FT-like proteins in phloem and xylem saps from rice. *Plant Cell Physiol* 49:767–790
13. Dafoe N, Constabel C (2009) Proteomic analysis of hybrid poplar xylem sap. *Phytochemistry* 70:856–863
14. Aguero C, Thorne E, Ibanez A et al (2008) Xylem sap proteins from *Vitis vinifera* L. Chardonnay. *Am J Enol Vitic* 59:306–311
15. Basha S, Mazhar H, Vasanthaiah H (2010) Proteomics approach to identify unique xylem sap proteins in Pierce's disease-tolerant *Vitis* species. *Appl Biochem Biotechnol* 160:932–944
16. Alvarez S, Goodger J, Marsh E et al (2006) Characterization of the maize xylem sap proteome. *J Proteome Res* 5:963–972
17. Fernandez-Garcia N, Hernandez M, Casado-Vela J et al (2011) Changes to the proteome and targeted metabolites of xylem sap in *Brassica oleracea* in response to salt stress. *Plant Cell Environ* 34:821–836
18. Alvarez S, Marsh E, Schroeder S et al (2008) Metabolomic and proteomic changes in the xylem sap of maize under drought. *Plant Cell Environ* 31:325–340
19. Rep M, Dekker H, Vossen J et al (2002) Mass spectrometric identification of isoforms of PR proteins in xylem sap of fungus-infected tomato. *Plant Physiol* 130:904–917
20. Houterman P, Speijer D, Dekker H et al (2007) The mixed xylem sap proteome of *Fusarium oxysporum*-infected tomato plants. *Mol Plant Pathol* 8:215–221
21. Li S, Hartman G, Lee B-S et al (2000) Identification of a stress-induced protein in stem exudates of soybean seedlings root-infected with *Fusarium solani* f. sp. *glycines*. *Plant Physiol Biochem* 38:803–809
22. Subramanian S, Cho U-H, Keyes C et al (2009) Distinct changes in soybean xylem sap proteome in response to pathogenic and symbiotic microbe interactions. *BMC Plant Biol* 9:119
23. Emanuelsson O, Brunak S, von Heijne G et al (2007) Locating proteins in the cell using TargetP, SignalP, and related tools. *Nat Protoc* 2:953–971
24. Small I, Peters N, Legeai F et al (2004) Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4:1581–1590
25. Fankhauser N, Mäser P (2005) Identification of GPI anchor attachment signals by a Kohonen self-organizing map. *Bioinformatics* 21: 1846–1852
26. Eisenhaber B, Wildpaner M, Schultz CJ et al (2003) Glycosylphosphatidylinositol lipid anchoring of plant proteins. Sensitive prediction from sequence- and genome-wide studies for *Arabidopsis* and rice. *Plant Physiol* 133: 1691–1701
27. Sigrist CJA, Cerutti L, de Castro E et al (2010) PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res* 38(D):161–166
28. Punta M, Coggill PC, Eberhardt RY et al (2012) The Pfam protein families database. *Nucleic Acids Res* 40(D):290–301
29. McDowall J, Hunter S (2011) InterPro protein classification. *Methods Mol Biol* 694: 37–47
30. Altschul SF, Madden TL, Schäffer AA et al (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402

31. Marchler-Bauer A, Lu S, Anderson JB et al (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* 39(D):225–229
32. Albenne C, Canut H, Boudart G et al (2009) Plant cell wall proteomics: mass spectrometry data, a trove for research on protein structure/function relationships. *Mol Plant* 2: 977–989
33. Faye L, Boulafous A, Benchabane M et al (2005) Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. *Vaccine* 23:1770–1778
34. Minic Z, Jamet E, Négroni L et al (2007) A sub-proteome of *Arabidopsis thaliana* mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. *J Exp Bot* 58:2503–2512
35. Shevchenko A, Wilm M, Vorm O et al (1996) Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels. *Anal Chem* 68:850–858
36. San Clemente H, Pont-Lezica R, Jamet E (2009) Bioinformatics as a tool for assessing the quality of sub-cellular proteomic strategies and inferring functions of proteins: plant cell wall proteomics as a test case. *Bioinform Biol Insights* 3:15–28

Chapter 29

Suspension-Cultured Plant Cells as a Tool to Analyze the Extracellular Proteome

Ana B. Sabater-Jara, Lorena Almagro, Sarai Belchí-Navarro, María J. Martínez-Esteso, Sabry M. Youssef, Juan Casado-Vela, Juan C. Vera-Urbina, Susana Sellés-Marchart, Roque Bru-Martínez, and María A. Pedreño

Abstract

Suspension-cultured cells (SCC) are generally considered the most suitable cell systems to carry out scientific studies, including the extracellular proteome (secretome). SCC are initiated by transferring friable callus fragments into flasks containing liquid culture medium for cell biomass growth, and they are maintained in an orbital shaker to supply the sufficient oxygen that allows cell growth. SCC increase rapidly during the exponential phase and after 10–20 days (depending on the cell culture nature), the growth rate starts to decrease due to limitation of nutrients, and to maintain for decades these kinds of cell cultures is needed to transfer a portion of these SCC into a fresh culture medium. Despite the central role played by extracellular proteins in most processes that control growth and development, the secretome has been less well characterized than other subcellular compartments, meaning that our understanding of the cell wall physiology is still very limited. Useful proteomic tools have emerged in recent years to unravel metabolic network that occurs in cell walls. With the recent progress made in mass spectrometry technology, it has become feasible to identify proteins from a given organ, tissue, cells, or even a subcellular compartment. Compared with other methods used to isolate cell wall proteins, the spent medium of SCC provides a convenient, continuous, and reliable and unique source of extracellular proteins. Therefore, this biological system could be used as a large-scale cell culture from which these proteins can be secreted, easily separated from cells without cell disruption, and so, without any cytosolic contamination, easily recovered from the extracellular medium. This nondestructive cell wall proteome approach discloses a set of proteins that are specifically expressed in the remodelling of the cell wall architecture and stress defense.

Key words Suspension-cultured cells, Extracellular proteome, Liquid chromatography, Mass spectrometry-based protein identification, De novo sequencing, Functional annotation

1 Introduction

Plant cells can be cultivated under sterile conditions in liquid (as suspension-cultured cells, SCC) or solid (as callus) culture media maintaining an undifferentiated cell state and showing an

unlimited growth. Usually, any part of the plant which contains undifferentiated cells can be used to initiate a plant cell culture although firstly the plant tissue must be sterilized since the most severe problem during the establishment of a plant cell culture is the high degree of microbial contamination of wild-grown plants. Explants (selected part of the sterilized plant tissue) are transferred to a solid culture medium where dedifferentiated cells begin to grow producing a mass of cells known as callus culture. Usually, callus formation is observed at the cut surface of the explants although sometimes callus formation also starts inside of the explants and breaks through the surface of this selected plant tissue. Callus cultures are aggregates of undifferentiated cells usually grown on solidified nutrient media which maintain their state of undifferentiated growth by adding to the culture media, exogenous phytohormones, mainly a low balanced mixture of auxins and cytokinins. In these conditions, the callus cultures show slow growth and cellular heterogeneity that can be reduced by subsequent subculturing (every 3–6 weeks) which in turn maintain these cultures for decades.

The callus cultures are often used to obtain SCC. In fact, SCC is initiated by transferring friable callus fragments into flasks containing liquid culture medium for cell biomass growth, and they are maintained in an orbital shaker to supply the sufficient oxygen that allows cell growth. SCC usually proliferate more rapidly than callus cultures and they have three different phases of growth. During the lag phase, the rate of cell growth is slow. Then, it increases rapidly during the exponential phase and after 10–20 days (depending on the cell culture nature), the growth rate starts to decrease due to limitation of nutrients, and to maintain for decades these cultures are needed to transfer a portion of these SCC into fresh culture medium. These cultures are generally considered the most suitable cell systems to carry out scientific studies.

During closing decades SCC has been extensively used to study cell growth and metabolism since they are able to synthesize a wide variety of secondary metabolites often of pharmaceutical importance. In particular, cell culture systems represent a potential renewable source of valuable compounds, flavors, fragrances, colorants, and proteins, which cannot be produced by microbial cells or chemical synthesis [1]. In plants, these secondary metabolites do not perform vital physiological functions, but they are produced to ward off potential predators, attract pollinators, or combat infectious diseases. The principal advantage of SCC is that it may provide continuous, reliable source of plant pharmaceuticals and could be used for the large-scale culture of plant cells from which these metabolites can be extracted. Other advantages of such systems for the production of these compounds are the maintenance of the uniform quality of the natural products, the independence of geographic location and climatic changes, free of diseases,

the stability of supply, and a closer relationship between supply and demand. In some cases, the production of secondary metabolites by using large-scale cell cultures is technically feasible, when the generation and maintenance of cell biomass are achieved together with a high yield in the production of these metabolites. However, in some cases, the use of SCC is limited by the economy of the process determined by the low yields of secondary metabolite production obtained. To increase the production of secondary metabolites using SCC as cell factories, different approaches have extensively been used such as the screening and selection of high-producing cell lines, cell immobilization, optimization of culture conditions and feeding, and elicitation. Also, genetic modification or metabolic engineering are promising biotechnological approaches to improve the production of these compounds [2, 3].

Similarly, SCC has been extensively used to study the signaling pathways involved in plant defense against several elicitors. In fact, the use of SCC has allowed the identification of numerous early events and proteins or genes involved in the signaling cascades immediately downstream of elicitor perception [4]. Early events, such as protein phosphorylation or activation of plasma membrane proteins, mobilize or generate directly or indirectly diverse signaling molecules (such as free calcium, nitric oxide, and active oxygen species) which regulate many processes, interconnecting branch pathways that amplify and specify the physiological response through transcriptional and metabolic changes [5].

In the same way, SCC has been used as potential production systems of a wide range of recombinant proteins [6]. Like microbes, SCC is inexpensive to grow and maintain, and they can be cultivated easily in large-scale bioreactors [7, 8] and because they are higher eukaryotes they can carry out many of the posttranslational modifications that occur in human cells. Also, they can synthesize complex proteins and glycoproteins, such as immunoglobulins [9, 10] and interleukins [11]. Perhaps, the most important advantage of SCC over the use of the whole plant for the production of recombinant proteins is the much simpler procedure for product isolation and purification [12–14] especially when the product is secreted into the culture medium. This means that good manufacturing practice could be implemented throughout the production pipeline [6]. Although no recombinant proteins have yet been produced commercially using SCC, there have been many proof-of-principle studies and several companies are investigating the commercial feasibility of such production systems.

On the other hand, plant cells secrete a variety of proteins into the extracellular space or bind to the cell wall [15–17]. The function of extracellular space is crucial for plant life since it includes processes related with growth, tissue structure, defense against biotic and abiotic stress factors, transport, osmoregulation, cell adhesion, and gas exchange [18]. Therefore, the cell wall is

considered a dynamic compartment that plays a critical role not only in determining cell shape and growth but also in interacting with environmental factors including those required for nutrition and defense against abiotic stress and pathogen attacks [19, 20]. Despite the central role of extracellular proteins in most processes that control growth, development [21], and even defense responses, the extracellular proteome (secretome) is less well characterized than other subcellular compartments. In fact, the *Arabidopsis* genome shows that about 17 % of the genome (i.e., 5,000 genes) encodes for proteins with a predicted signal peptide that targets them towards the cell wall [22], but the function of most remains unknown. Different proteomic tools are used in order to investigate and understand the functions of cell wall proteins. Thus, the use of two-dimensional electrophoresis, the availability of complete genome sequences, and the recent progress made in mass spectrometry technology allow the complete identification of entire extracellular proteome and the assignment of new functions. Compared with other methods for isolating extracellular proteins (cell wall isolation, intercellular washing fluid extraction, amongst others), the spent medium of an SCC provides a convenient, continuous, and unique source of extracellular proteins, easily obtained without cell disruption and so without any cytosolic contamination [23–26]. In fact, the spent medium of SCC constitutes a useful model system of reduced complexity that contains all proteins involved in the different aspects of cell wall metabolism, including those involved in defense. Indeed, pathogenesis-related-proteins (PR-proteins), specifically PR-1 and -5, endochitinases, chitinases, peroxidases, and other hydrolytic enzymes, are often presented in the extracellular proteome of SCC from *Zinnia elegans*, *Cycas revoluta*, *Taxus baccata* [27], *Capsicum* sp. [25, 28], *Nicotiana tabacum* [24], and *Solanum lycopersicum* [26]. Most of these proteins are involved in the dynamic modification of the cell wall architecture during cell culture growth and in the prevention of cell wall degradation by microbial attacks.

For these reasons, the main goal of this chapter is focused on the description of all procedures needed, from the initiation of an SCC, obtaining and purifying proteins from extracellular medium until the analysis and identification of the extracellular proteome.

2 Materials

2.1 Plant Material

Immature berries from *Vitis vinifera*.

2.2 Equipment for Plant Cell Cultures

1. Laminar air-flow chamber.
2. Autoclave.
3. Plant growth chamber.

4. Rotary shaker.
5. Sterile scalpels, forceps, and Petri dishes.
6. Filtration equipment (Buchner funnel and flask for filtration under reduced pressure).

2.3 Sterilizing Solutions

1. Ethanol solution: 70 % (v/v) in water.
2. Calcium hypochlorite solution: 7 % (w/v) in water.
3. Tween 20.
4. Sterile distilled water.

2.4 Plant Culture Medium

1. Murashige and Skoog (MS) basal salt mixture (Duchefa or Sigma) [29].
2. Gamborg B₅ basal salt mixture (Duchefa or Sigma) [30].
3. Kinetin (Sigma) 200 mg/L. Dissolve 0.02 g kinetin in 90 mL water. Add, while stirring, drop by drop, 1 M NaOH. Add water to a volume of 100 mL. Mix and store at 4 °C.
4. 1-Naphthaleneacetic acid (NAA) (Sigma) 200 mg/L. Dissolve 0.02 g NAA in 90 mL water. Add, while stirring, drop by drop, 1 M NaOH. Add water to a volume of 100 mL. Mix and store at 4 °C.
5. Morel vitamins. Prepare 100 mL of a 1,000× stock of Morel vitamins [31] in water. While stirring, add 10 mg calcium pantothenate, 1 mg of biotin, 10 mg of nicotinic acid, 10 mg of thiamine, 10 mg of pyridoxine, and 10 g of myoinositol. Store at 4 °C.
6. Acid (1 N HCl) and alkali (1 M NaOH or KOH) solutions to adjust the pH of the media.
7. Plant Agar® (Duchefa).
8. Induction medium. Add 4.3 g MS basal salt mixture, 30 g sucrose, and 0.25 g casein hydrolysate to 1 L distilled water (see Note 1). Stir and then add 1 mL 200 mg/L kinetin, 0.5 mL 200 mg/L NAA, and 1 mL of 1,000× Morel vitamins. While stirring, adjust the medium to pH 6.0 using HCl, NaOH, or KOH. For solid media, add 7.0 g of agar per liter of media and autoclave for 20 min at 121 °C (see Note 2). After that, distribute the medium in sterilized Petri dish (see Note 3).
9. Growth medium. Add 3.05 g Gamborg B₅ basal salt mixture, 20 g sucrose, and 0.25 g casein hydrolysate to 1 L distilled water (see Note 1). Stir and then add 1 mL 200 mg/L kinetin, 0.5 mL 100 mg/L NAA, and 1 mL of Morel vitamins [31]. While stirring, adjust the medium to pH 6.0 using HCl, NaOH, or KOH and then distribute the medium in 250 mL Erlenmeyer flasks which contains 100 mL of Gamborg B₅. For solid media, add 0.7 g of agar in each, and autoclave them for 20 min (see Note 2) at 121 °C.

2.5 Components for Protein Extraction: Trichloroacetic Acid Precipitation

1. Extracellular medium from *Vitis vinifera* SCC.
2. 2 % (w/v) polyvinylpolypyrrolidone (PVPP): Add 2 g PVPP to 100 mL of extracellular medium.
3. Trichloroacetic acid (TCA) solution: 24 % in water. Store at 4 °C.

2.6 Protein Quantification

1. Bio-Rad Protein Assay (dye reagent concentrate) (Bio-Rad) (*see Note 4*).
2. Protein standard: Bovine serum albumin (BSA) at a concentration of 0.2 mg/mL in distilled water is used as a stock solution. Store at -20 °C (*see Note 5*).

2.7 Sodium Dodecyl Sulfate-Polyacrylamide Gel

Prepare the gel electrophoresis using a MiniProtean® 3 Cell electrophoresis kit (Bio-Rad).

1. Stacking gel buffer: 0.5 M Tris-HCl, pH 6.8. Store at 4 °C.
2. Resolving gel buffer: 1.5 M Tris-HCl, pH 8.8. Weigh 181.5 g Tris. Add water to a volume of 1,000 mL. Mix and adjust pH with 1 M HCl. Store at 4 °C.
3. 30 % Acrylamide/Bisacrylamide solution, 37.5:1 (Bio-Rad).
4. SDS solution: 10 % (w/v) in water. Store at room temperature.
5. Ammonium persulfate (APS) solution: 10 % (w/v) in water. Store at room temperature (*see Note 6*).
6. *N,N,N',N'*-Tetramethylethylenediamine (TEMED, Sigma). Store at 4 °C (*see Note 7*).
7. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) running buffer: 25 mM Tris-HCl, pH 8.3, 192 mM glycine, 0.1 % SDS. Store at room temperature (*see Note 8*).
8. SDS-PAGE loading buffer: 0.5 M Tris-HCl, pH 6.8, 10 % glycerol, 10 % SDS, 0.5 % bromophenol blue, 5 % β -mercaptoethanol. Store at room temperature.
9. Prestained SDS-PAGE standards, broad range (Bio-Rad).
10. PowerPac 300 (Bio-Rad).

2.8 Two-Dimensional Gel Electrophoresis (IEF/SDS-PAGE)

1. Immobilized pH gradient (IPG) strips of pH range 3–10 and 13 cm length (GE Healthcare).
2. Rehydration solution with IPG buffer: 7 M Urea, 2 M thiourea, 4 % (w/v) CHAPS, IPG buffer pH 3–10 0.5 % (v/v), 50 mM dithiothreitol (DTT), and 0.002 % bromophenol blue (*see Note 9*). Weigh 0.84 g urea and add water to a volume of 1.2 mL. Dissolve and add 0.30 g thiourea, 0.08 g CHAPS, 0.016 g DTT, 10 μ L IPG buffer, and a few amount of bromophenol blue and complete to 2 mL with water.
3. Reagents for SDS-PAGE are the same as described in Subheading 2.7 (items 2–7).

4. SDS equilibration buffer: 50 mM Tris pH 8.8, 6 M urea, 30 % (v/v) glycerol, 2 % (w/v) SDS.
5. Agarose solution: Suspend 0.5 % (w/v) agarose in SDS-PAGE running buffer and dissolve by heating.

2.9 Silver Gel Staining

1. Fixing solution: 30 % (v/v) methanol, 10 % (v/v) glacial acetic acid.
2. Incubation solution 1: 5 % (v/v) methanol.
3. Incubation solution 2: 0.02 % (w/v) sodium thiosulfate.
4. Incubation solution 3: 0.2 % (w/v) silver nitrate.
5. Developing solution: 3 % (w/v) sodium carbonate, 37 % (w/v) formaldehyde, 2 mL incubation solution 2.
6. Stop solution: 1.4 % (w/v) Na₂-EDTA.
7. Destaining solution: 0.2 % (w/v) potassium ferricyanide. Add 0.2 g of potassium ferricyanide in 100 mL of 0.02 % (w/v) sodium thiosulfate.

2.10 Coomassie Brilliant Blue G Staining

1. Coomassie staining solution: 1 % (w/v) Coomassie Brilliant Blue G in water, filter to remove any insoluble compounds; 10 % (v/v) of this solution is used to prepare the staining solution with 10 % (v/v) glacial acetic acid and 40 % (v/v) methanol.
2. Coomassie destaining solution: 30 % (v/v) methanol, 10 % (v/v) glacial acetic acid.

2.11 Colloidal Coomassie Blue G-250 Staining

1. Fixation solution: 10 % (v/v) acetic acid, 40 % (v/v) methanol.
2. Coomassie blue G-250 solution: (5 mL, 5 % (w/v) in water). Weigh 0.25 g Coomassie blue G-250, dissolve with 5 mL of water, and stir solution for a few minutes to disperse the dye. The dye will not dissolve completely.
3. Staining solution: Weigh 25 g ammonium sulfate, add 3 mL of 85 % phosphoric acid, and 5 mL of 5 % Coomassie blue G-250 solution and add water to a volume of 250 mL. Finally add 62.5 mL of methanol.

2.12 In-Gel Protein Digestion

1. Soaking solution: 25 mM ammonium bicarbonate.
2. Reducing solution: 60 mM DTT in 25 mM ammonium bicarbonate.
3. Alkylation solution: 100 mM iodoacetamide (freshly prepared, immediately before using) in 25 mM ammonium bicarbonate.
4. Sequencing-grade modified porcine trypsin (Promega).
5. Formic acid (FA) solution: 1 % in water.

2.13 HPLC Tandem Mass Spectrometry

1. Nano-HPLC system.
2. Mass spectrometer equipped with a nano-electrospray ion source (see Note 10).
3. FA solution: 0.1 % in water.
4. Reverse-phase (RP) chromatography buffer A (RPB-A): 0.1 % FA.
5. RP chromatography buffer B (RPB-B): 0.1 % FA (v/v) in 90 % (v/v) acetonitrile (ACN).
6. Reverse-phase chromatography trap column: Zorbax 300SB-C18, 5 mm × 0.3 mm, 5 µm particle size (Agilent Technologies).
7. RP chromatography analytical column: Zorbax 300SB-C18 analytical column, 150 mm × 75 µm, 3.5 µm particle size (Agilent Technologies).
8. Search engines for LC-MS/MS protein identification (see Note 10).

2.14 MALDI-TOF Mass Spectrometry

1. Spot picker (Gel Company).
2. Desalting cartridges Zip Tips C18 (Agilent Technologies).
3. Peptide calibration standard (Bruker Daltonics) that consists of a combination of peptides that provides a good calibration across a typical mass range between 1,000 and 3,500 Da.
4. 0.1 % Trifluoroacetic acid (TFA).
5. Matrix solution: 4.7 mg/mL α-cyano-4-hydroxycinnamic acid (CHCA, Sigma-Aldrich) in 70 % ACN.
6. AnchorChip MALDI target (Bruker Daltonics).
7. Autoflex mass spectrometer (Bruker Daltonics).
8. MASCOT: Search engine for peptide mass fingerprint protein identification.

2.15 Peptide De Novo Sequencing

1. Peaks Studio v4.5 SP2 (www.bioinformaticssolutions.com).

3 Methods

3.1 Plant Cell Cultures and Culture Conditions (Establishment of Callus Cultures and Suspension-Cultured Cells)

We describe here how a plant cell culture (see Note 11) is initiated and what the culture conditions are, in order to use the extracellular medium for analyzing the extracellular proteome (secretome).

1. Disinfect immature fruits from *Vitis vinifera* by immersing first in 70 % ethanol solution for 1 min and then into 7 % (w/v) calcium hypochlorite solution with 0.1 % Tween 20 for 15 min (see Note 12).

- Once the fruits are disinfected, and working in the laminar air-flow chamber, remove calcium hypochlorite. Add sterile distilled water and wash three times for 1 min each time.
- Divide the immature fruit berries in four portions with the help of a scalpel, working on sterile Petri dish.
- Transfer the explant into calli-induction medium, and place Petri dishes at 25 °C in darkness (*see Note 13*). Subculture cultures every 3 weeks until they reach approximately 3 cm in size.
- After callus formation, the different callus tissues are separated from each explant and transferred to 250 mL sterile Erlenmeyer flasks containing 100 mL of solid growth medium. Both explants and calli are maintained at 25 °C in darkness, and subcultured for 2–3 weeks.
- Initiate grapevine SCC by transferring 20 g of fresh-weight friable calli into 250 mL Erlenmeyer flasks containing 100 mL of liquid growth medium. Maintain SCC in a rotary shaker (110 rpm) at 25 °C in darkness. Subculture SCC every 14 days by diluting with one volume of liquid growth medium and then distributing into two flasks.
- Separate cells from the extracellular medium by filtration using a gentle vacuum. This extracellular medium is used for protein precipitation.

3.2 Protein Sample Preparation

- SCC is filtrated through a glass vacuum filter and the extracellular medium is kept for further analysis.
- The cell-free medium is frozen at –20 °C overnight. Then, the extracellular medium is thawed and centrifuged at a medium speed (8,000 $\times g$ (Sorvall RC-5B plus, SS-34 rotor)) in order to remove polysaccharides.
- Add ethyl acetate to the supernatant (1:4, v/v) to remove stilbenes (*see Note 14*). Stir for 30 min and centrifuge at 6,700 $\times g$.
- After removing the organic phase, the aqueous phase is supplemented with 2 % (w/v) PVPP and incubated for 60 min with shaking. The slurry is centrifuged at 8,000 $\times g$ for 15 min to remove the insoluble PVPP and the supernatant is kept for protein extraction (*see Note 15*).
- The proteins are recovered by precipitation adding TCA to a final concentration of 6 % (w/v) and pelleted at 13,400 $\times g$ during 10 min (*see Notes 16 and 17*).
- The pellets are washed in ice-cold methanol which means that the pellet is resuspended and then centrifuged at 13,400 $\times g$. Repeat the sequence twice.
- The pellet is then washed in ice-cold acetone three times as described above.

8. The pellet is now dried and solubilized in an appropriated sample buffer for further analysis (*see Note 18*). Alternatively, store the pellet at -20°C , for longer term preservation. Samples are ready to use in protein quantitation and electrophoretic assays.

3.3 Protein Quantification

1. Calibration curve: Prepare several dilutions of standard protein (BSA) containing volumes of 10, 20, 40, 60, 80, 100, and 120 μL of BSA (0.2 mg/mL) into test tubes and make each up to 800 μL with distilled water. Use 800 μL of distilled water as control.
2. To determine the sample concentration, assay a range of dilutions (1, 1:10, 1:100, 1:1,000). Prepare duplicates of each sample.
3. Add 200 μL of dye protein reagent (Subheading 2.6) to each tube and mix well by gentle vortex-mixing (avoid excess foaming).
4. Incubate the samples with the protein reagent at room temperature for 30 min in darkness.
5. Measure absorbance at 595 nm of each sample and standards.

3.4 One-Dimensional Electrophoresis

1. Add 10 μL SDS-PAGE loading buffer to the protein sample (2–10 μg of extracellular protein) in a final volume of 30 μL and heat samples at 95°C for 5 min.
2. Prepare the resolving gel buffer, with 12 % (w/v) acrylamide gel final concentration, by mixing 2.5 mL resolving buffer, 4 mL 30 % acrylamide/Bis solution, 0.1 mL SDS solution, and 3.4 mL distilled water. Degas with vacuum for 15 min. Add 50 μL APS and 5 μL TEMED, and pour the gel within a 7.25×10 cm, 1 mm thick gel cassette. Pour approximately 5 mL of this solution into the gel cassette and completely cover the solution surface with isobutanol or water to obtain a flat front on top of the resolving gel. Leave at room temperature until the solution is completely polymerized. Gently dry the gel surface with blotting paper.
3. Prepare the solution for the stacking gel by mixing 1.25 mL of stacking gel buffer with 0.67 mL acrylamide/Bis solution, 0.05 mL SDS solution, and 3.05 mL distilled water. Degas with vacuum for 15 min. Add 25 μL APS and 10 μL TEMED to the solution of the stacking gel and gently stir to obtain a uniform solution. Pour about 2.5 mL above the resolving gel and insert the well-forming comb into this solution (*see Note 19*). Polymerize the gel for at least 45 min at room temperature to allow complete polymerization.
4. When polymerization is completed, remove the comb from the stacking gel. Remove any spacer from the bottom of the gel cassette, and assemble the cassette in the electrophoresis

tank. Fill the top reservoir with SDS-PAGE running buffer. Check for leaks from the top into the bottom compartment. Rinse out any non-polymerized acrylamide solution from the wells using electrophoresis buffer.

5. Apply the sample by using a syringe and add carefully up to 30 μ L of sample to the bottom of each well. The volume and protein concentration of the sample should be sufficient to give at least 10 μ g of each protein. Apply 15 μ L of the molecular weight standards to one or two wells, preferably in an asymmetric position.
6. To perform the electrophoresis running, close the apparatus with the lid and connect the wires to the power supply unit (*see Note 20*). Apply 200 V for ~45 min in this system. Run the gel until the blue dye front reaches the bottom of the gel (*see Note 21*). Disconnect the electrophoresis unit from the power supply, remove the lid, and discard the SDS-PAGE running buffer. Separate the gel from plates with the help of a spatula, discard the stacking gel, and wash the separating gel with distilled water to remove traces of SDS-PAGE running buffer.

3.5 Gel Staining

The selection of Coomassie blue or silver gel staining is dependent on the quantity of protein that samples contain. Silver gel staining is usually used when samples contain nanograms of proteins.

All steps should be performed with gentle shaking of the staining tray at room temperature.

1. For silver gel staining (*see Note 22*), gels are stained according to the method of Blum et al. 1987 [32]. Fix gel for at least 30 min in fixative solution. Transfer gel to incubation solution 1 for 15 min. Decant carefully the fixer and rinse the gel three times for 5 min with deionized water to remove the fixative solution. Transfer gel to incubation solution 2 for 2 min. Decant thiosulfate carefully and rinse the gel three times for 30 s with deionized water to remove the incubation solution. Transfer gel to incubation solution 3 for 25 min. Decant silver nitrate carefully and rinse the gel three times for 1 min with deionized water to remove the incubation solution. Transfer gel to developing solution and leave for a maximum of 10 min. Transfer gel to the stop staining solution for 10 min. Rinse repeatedly gel with water to remove residual stain.

For silver gel destaining, cut the interest bands from gels and destain with destaining solution until no bands are visible. The gel will have a yellow hue. Rinse gel repeatedly (five times, 5 min) with water until gel is transparent and has no background color (Fig. 1). Place bands in different tubes containing distilled water until their processing.

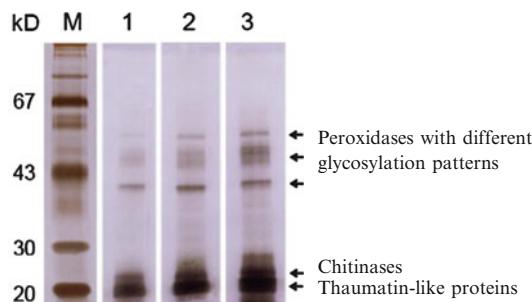


Fig. 1 SDS-PAGE followed by silver protein staining of the elicited extracellular medium of *V. vinifera*. Lane M, Molecular weight markers; Lanes 1–3, extracellular medium obtained from treated grapevine cells with methyl β -cyclodextrin (MBCD) alone (lane 3) or in combination with methyl jasmonate (MeJA) at different concentrations (450 μ M, lane 2 and 270 μ M, lane 1)

2. For Coomassie brilliant blue staining (see Note 23), place the gel in a square Petri dish containing 100 mL Coomassie staining solution. Incubate the gel for 45 min to 3 h in the staining solution. Destain the gel by adding 100 mL Coomassie destaining solution. To remove the background, destain the gel with several changes of Coomassie destaining solution.
3. Keep gels in distilled water for short-term preservation. It might be convenient to transfer carefully the gel to a heat-sealable bag for longer term storage.
4. Excise the bands of interest from gel, and place them in different tubes containing distilled water until their processing.

3.6 In-Gel Protein Digestion

Once the interest bands are cut from gels, deposit in 96-well plates and process automatically in a Proteineer DP (Bruker Daltonics). The digestion protocol used is based on that described by Schevchenko et al. [33] with minor modifications. Wash gel plugs with 25 mM ammonium bicarbonate to remove dye and SDS impurities. Reduce the samples with 60 mM DTT and alkylate cysteines by adding an excess of iodoacetamide followed by digestion with porcine trypsin (Promega) at 37 °C for 6 h. Extract peptides by washing in 25 mM ammonium bicarbonate, then in 70 % ACN, and then in 1 % FA. Finally dry tryptic peptides using a speed-vacuum centrifuge.

3.7 LC-MS/MS Identification and Database Searches

1. Resuspend tryptic peptides obtained from in-gel digestion (Subheading 3.6) in 100 μ L of 0.1 % FA and 3 % ACN, suitable for LC-MS/MS analysis.
2. Concentrate and desalt the tryptic peptides on a Zorbax 300SB-C18 cartridge and separate on an analytical Zorbax RP C18 column with an Agilent 1200 HPLC system.

3. Elute peptides at 300 nL/min flow rate by using a 40-min linear gradient from 3 to 60 % ACN with 0.1 % FA.
4. Scan and fragment the peptides with a mass spectrometer equipped with a nano-electrospray ion source (*see Note 10*). The three most intense precursor ions, ranging from 400 to 2,000 *m/z*, are scanned and measured in the mass spectrometer at a 60,000 resolution at *m/z* 400, and the corresponding fragment ions generated are measured in the mass spectrometer after collision-induced dissociation (CID) fragmentation using 35–40 % normalized collision energy.
5. Perform the database searches against NCBIInr (non-redundant) database (<ftp://ftp.ncbi.nih.gov/blast/db/FASTA/nr.gz>) combining both MASCOT [34] and SEQUEST [35] search engines implemented in Proteome Discoverer software (*see Note 10*).
6. Use the following parameters for the searches: trypsin as the proteolytic enzyme, allowing for one missed cleavage; carboxyamidomethyl cisteine as fixed modifications; oxidation of methionine as variable modification.
7. Filter the resulting peptides to show the list of proteins identified with less than 1% false discovery rate. Consider positive identifications only when two or more peptides are matched, and their score is >20 for MASCOT and >2.5 for SEQUEST.

3.8 Peptide De Novo Sequencing

1. Tryptic peptides generated from in-gel protein digestion are analyzed by reverse-phase HPLC as described above (Subheading 3.7). Both precursor and fragment ions are accurately measured in the mass spectrometer and fragment ions are generated by CID fragmentation.
2. Fragment spectra are de novo sequencing using Peaks Studio v4.5 SP2 (www.bioinformaticssolutions.com). Parental and fragment mass error tolerances are 0.01 and 0.02 Da, respectively. Only those sequences with peptide scores ≥ 50 and quality values ≥ 0.75 and containing ≥ 4 amino acid strings with confidence ≥ 99.5 are considered as valid candidates. Fragment spectra are also manually validated and the proposed sequences searched using the BLASTP software [36].

3.9 Two-Dimensional Gel Electrophoresis: Isoelectrofocusing Plus SDS-PAGE

1. For sample in-gel rehydration [37], directly solubilize 500 μ g of protein sample (pellet) in 250 μ L rehydration solution with IPG buffer 3–10 (*see Note 24*).
2. Select the Strip Holders corresponding to the IPG strip length chosen for the experiment and 13 cm IPG strip pH 3–10 NL (*see Note 25*), pipette 250 μ L of sample-containing rehydration solution into each strip holder base, and remove any large bubbles.

3. Remove the protective cover from the surface of the IPG strips (*see Note 26*) and slowly lower the IPG strip (gel slide down) onto the rehydration solution, without trapping air bubbles. Then, cover the IPG strip with 1–2 mL of IPG DryStrip Cover Fluid (which minimizes evaporation and prevents urea crystallization) and apply the plastic cover.
4. Apply the IPGphor strip holders onto the electrode contact area of the IPGphor (*see Note 27*), close the safety lid and pressure blocks on the underside of the cover, and ensure that the IPG strip keeps in good contact with the electrodes as the gel swells.
5. Apply low voltage (50 V) during rehydration for 12 h at 20 °C for improving the entry of high molecular proteins [38].
6. After active rehydration, start isoelectric focusing at 20 °C using the protocol recommended by the manufacturer: firstly, 1-h step-and-hold at 500 V, next 1-h gradient at 1,000 V, followed by 2 h-and-30 min gradient steps at 8,000 V, and finally step-and-hold at 8,000 V until an accumulated voltage of 18,000 V is reached.
7. For IPG strip equilibration, place the focused IPG strips into individual test tubes (250 mm long; 20 mm internal diameter). Dissolve 0.1 g of DTT in 10 mL of SDS equilibration buffer and add to each tube. Seal the tubes with parafilm and equilibrate the IPG strips for 15 min in a shaker to ensure the equilibration of the strips. Rinse IPG strips with deionized water and place in other test tube. Dissolve 0.25 g of iodoacetamide in 10 mL of SDS equilibration buffer and add to each tube. Seal the tubes with parafilm, equilibrate them for an additional 15 min in a shaker, and finally, rinse the IPG strips with deionized water prior to the second dimension.
8. For SDS-PAGE, set up the gel cassette and prepare the gel sandwich (16×16 cm) using the Hoefer™ SE 600 vertical set. Mix 7.5 mL of resolving gel buffer, 12.5 mL of acrylamide/Bis solution, and 9.5 mL water in a 50 mL conical flask. Add 0.3 mL SDS 10 %, 150 µL of APS, and 125 µL of TEMED. Pipette the solution into one corner of the sandwich, taking care not to introduce any air bubbles. Fill solution until 1 cm below the top of the glass plate to allow space for overlaying each gel with isopropanol (*see Note 28*). Allow the gel to polymerize for a minimum of 8 h or overnight.
9. Remove the overlay by rinsing the top of the gel several times with deionized water. Invert the tray to allow drain. To ensure a seamless contact remove residual liquid by blotting one corner with a lint-free tissue. Use a thin plastic ruler to slide the IPG strip into the gap between the two glass plates.

10. Add 2 mL of hot (75 °C) agarose solution and continue to slide the strip down onto the surface of the SDS gel until good contact is achieved. Ensure that no air bubbles are trapped between the IPG strip and the slab gel surface.
11. Apply the molecular weight markers soaking a filter paper pad with 5 µL, let it dry, and apply to the left of the IPG strip. Allow the agarose to solidify for at least 5 min.
12. Attach the gel sandwich to the upper chamber and place it in the casting unit. Fill the lower (2–4 L) and upper (450–600 mL) chamber with running buffer and place the safety lid on the unit.
13. Apply a constant current of 25 mA per gel until the dye reaches the front of the gel. Then turn off the power supply, disconnect the leads, and remove the safety lid.
14. Following electrophoresis pull out the upper buffer chamber assembly, slide away the spacers, and separate the plates with the use of a spatula. The gel remains on one of the glass plates. Invert the plate and position the gel low over the container with 200 mL of fixative solution.

3.10 Coomassie Colloidal Staining

Gels are stained according to the method of Neuhoff et al. [39]. For each step use a volume of 300 mL per two-dimensional gel and perform all steps under shaking. Fix the gel for at least 60 min or overnight in fixative solution. Decant the fixer carefully and rinse the gel three times for 10 min with deionized water to remove the fixative solution. Transfer gel to staining solution and leave it overnight. Rinse the gel repeatedly with warm water (45–55 °C) to remove residual stain (Fig. 2a).

3.11 Image Analysis Using Progenesis SameSpots v3.0 (Nonlinear Dynamics)

1. Scan gel images using a transmission-light scanner (Image Scanner, GE Healthcare) with a resolution of 300 dpi and digitized on 16 bits.
2. Select a good image as a reference with a clear and representative spot pattern which has a minimum of distortion.
3. Ignore the edges of the gel drawing a mask over these areas.
4. Align each of the images to the chosen reference: Select between 15 and 20 prominent spots to manually assign and use the automatic vector tool to add additional vectors.
5. After automatic spot detection and matching, revise manually the spots with edition tools for correct detection.
6. Establish gel groups according to the experimental design and normalize spot volume intensity ratios for each spot.
7. List all the spots and their normalized volume.
8. Select the statistically significant differentiated spots across the experiment based on ANOVA (p -value >0.05).

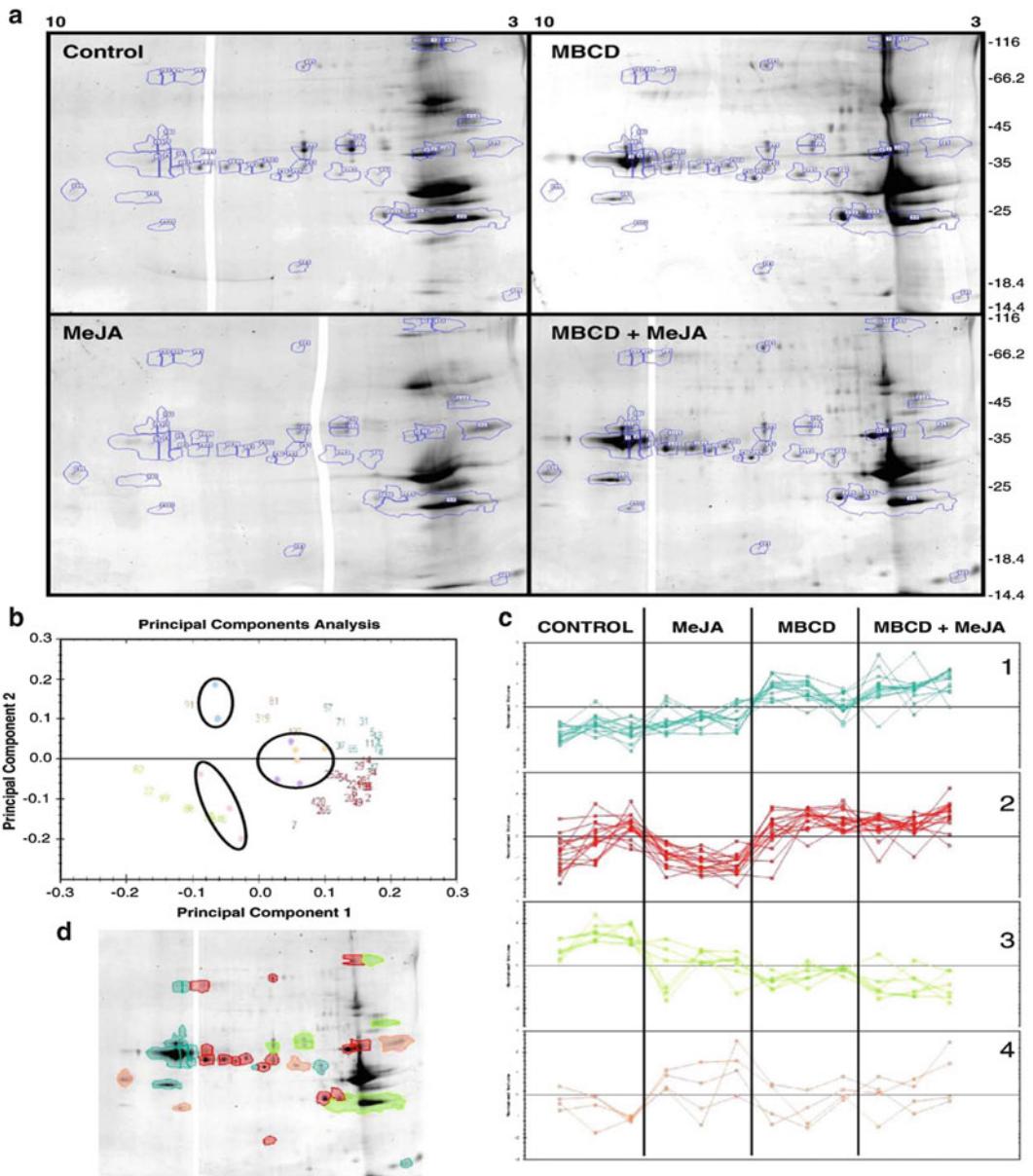


Fig. 2 Two-dimensional electrophoresis analysis of elicited extracellular medium obtained from grapevine cell cultures (cv. Gamay). **(a)** 2D gel sample images of the control and the three elicitation treatments, showing the set of selected spots in the experiment. At the top to the right, the *p*l and *Mw* (kDa) values are provided for reference. **(b)** PCA bi-plot of the two first principal components. Together they both explained 71 % of the selected spot variability, and the third principal component only explained an additional 5 %. Colored dots and numbers represent the gels and spots, respectively. **(c)** Standardized expression profiles separated into expression pattern clusters. Each line represents the standardized abundance of a spot across all the gels and belongs to one of the four clusters generated by hierarchical cluster analysis implemented in Progenesis SameSpots. **(d)** 2D gel sample image showing the location of the selected spots and its pertaining to a cluster by a color code. Dots: pink (control), blue (methyl jasmonate, MeJA), purple (methyl β -cyclodextrin, MBCD), light brown (MBCD+MeJA). Numbers, clusters, and spot boundaries: light blue (cluster 1 spots), red (cluster 2 spots), light green (cluster 3 spots), light brown (cluster 4 spots). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this chapter.) (Reproduced from [23] with permission from Elsevier) (color figure online)

9. Automatically statistical analysis is followed within SameSpots software. A principal component analysis (PCA) [40], correlation analysis (CA), and power analysis are implemented in SameSpots which allow studying the protein abundance trends in the experiment (Fig. 2b–d).
10. Choose a picking image for further analysis by mass spectrometry.

3.12 Protein Identification and Database Searches

Mass spectrometry analysis in the case of protein identification from spots is usually performed by a sequential strategy which consists in the identification of peptides obtained after in-gel digestion by peptide mass fingerprint (PMF) in a MALDI-TOF instrument followed by identification of those unidentified spots by tandem mass spectrometry (MS/MS) in an LC-MS/MS instrument. In the case of samples containing complex protein mixtures, the identification of proteins is directly performed after in-solution protein digestion followed by LC-MS/MS.

1. *MALDI-TOF mass spectrometry.* Spots or bands of interest are excised with the aid of a spot picker. The peptides for mass spectrometry analysis are obtained after in-gel digestion [33]. The tryptic fragments are desalted using Zip Tips C18 according to the manufacturer's instructions. Eluted peptides are completely dehydrated in a vacuum centrifuge and resuspended in 0.1 % TFA (see Note 29). A droplet of 0.5 μ L is spotted onto the AnchorChip MALDI target. After add 0.5 μ L of the CHCA matrix solution to the droplets and allow to air-dry at room temperature (see Note 30). External mass calibration was performed using a peptide calibration mixture at 100 fmol/ μ L concentration. A droplet of 0.5 μ L is spotted onto the AnchorChip MALDI target. After add 0.5 μ L of the CHCA matrix solution to the droplets and allow to air-dry at room temperature (see Note 31). Delayed-extraction MALDI-TOF spectra are recorded using an Autoflex (Bruker Daltonics) equipped with a nitrogen laser. Spectra were obtained in the positive-ion reflection mode. \geq 60 laser shots are acquired within each spot after manual identification of areas with strong signals. Summed spectral signal acquired in each spot is stored and used for further database searches and subsequent protein identification (see Note 32).

Protein identification is performed by peptide mass mapping and database searching using the in-house licensed search engine MASCOT (Matrix science). The following parameters are set for searches using MASCOT: enzyme: trypsin; fixed modifications: carbamidomethyl cysteine; variable modifications: oxidation of methionine; peptide tolerance: \pm 0.1 Da; peptide charge: 1H⁺ monoisotopic; number of missed cleavage sites: up to 1 missed cleavage site; decoy database searches

were allowed. The searches can be done using public databases such as NCBInr or SwissProt (see Note 33). A protein is considered identified when the minimal following criteria are fulfilled: a score of 65, four matched peptides, and sequence coverage of 25 %.

2. *MS/MS*. Protein extracted from extracellular media can be analyzed directly by mass spectrometry without a previous protein gel-based separation. In this case, a protein in-liquid digestion is performed [41]. Alternatively, in order to decrease sample complexity a peptide separation by strong cation/anion exchange chromatography (SCX/SAX) can be performed previously to MS/MS analysis which is known as MuDPIT approach [42].

The tryptic fragments are analyzed by LC-MS/MS using an Agilent 1100 series nano-HPLC system lined on an XCT plus ion trap mass spectrometer (Agilent technologies) equipped with a nano-ESI source. Sample concentration and desalting are performed on a Zorbax 300SB-C18 trap column at 3 μ L/min of RPB-A while peptide separation was achieved on a Zorbax 300SB-C18 analytical column using a 30-min linear gradient of 5–35 % RPB-B at a constant flow rate of 0.3 μ L/min. MS and MS/MS spectra are acquired in the standard enhanced mode (26,000 m/z per second) and the ultra-scan mode (8,100 m/z per second), respectively. Mass spectrometer settings for MS/MS analyses included an ionization potential of 1.8 kV and an ICC smart target (number of ions in the trap before scan out) of 400,000 or 150 ms of accumulation. MS/MS analyses are performed using automated switching with a preference for doubly charged ions and a threshold of 10^5 counts and 1.3 V fragmentation amplitude. Each MS/MS spectral dataset (ca. 1,200 spectra/run) is processed to determine monoisotopic masses and charge states, to merge MS/MS spectra with the same precursor ($\Delta m/z < 1.4$ Da and chromatographic $\Delta t < 15$ s) and to select high-quality spectra with the extraction tool of the Spectrum Mill Proteomics Workbench (SMPW). A two-step search is performed. The reduced dataset is searched against a suitable protein database, e.g., NCBInr, in the identity mode with the MS/MS Search tool of the SMPW using the following parameters: trypsin, up to two missed cleavages, fixed modification S-carbamidomethyl, variable modifications Met-oxidation, Asn- and Gln-deamidation, and a mass tolerance of 2.5 Da for the precursor and 0.7 Da for product ions. A sequence tag length minimum of three and four minimum peaks detected is selected and calculate reversed database scores. Peptide hits are validated first in the peptide mode and then in the protein mode according to the manufacturer's recommended score settings. Validated files are summarized in the protein mode to assemble peptides into

proteins. Manually all peptides are revised removing those whose score is equal to the forward-reverse scores. A protein was considered identified with a minimum of two different peptides and a score above 20 marked in the SMPW search engine (Fig. 3).

3.13 Functional Annotation by Blast2GO Software

The list of proteins identified in a proteomic experiment can be submitted into Blast2GO (B2G) software [43] in order to achieve information mainly about protein description of identified sequences and its functional annotation based on Gene Ontology (GO) vocabulary. This functional annotation may predict a putative function for uncharacterized proteins. Basically, B2G uses Blast searches to find similar sequences up to several hundred input sequences at once. Then, the program extracts the GO terms associated to each of the obtained protein hits and returns an evaluated annotation for the query sequences. Indeed, B2G is a useful tool which allows to visualize GO annotations as a reconstructed structure showing the relationships within Kyoto Encyclopedia of Genes and Genomes (KEGG) maps.

1. Export protein summary top hits for identified proteins in an excel file.
2. Create a comma-separated value text file (.csv file) with the accession number (AN) of the identified proteins in one row. Edit the file in a text processor to be sure that ANs are separated by “,”; otherwise replace automatically ‘;’ for ‘,’ and substitute carriage return features if any by ‘,’ in order to have all ANs codes in a single line and separated by commas. Save this file as plain text.
3. Retrieve sequences in NCBInr using Batch ENTREZ tool (<http://www.ncbi.nlm.nih.gov/sites/batchentrez>) in FASTA format using the created file.
4. Add *.fasta extension to the retrieved file previously to B2G procedure.
5. Run Blast2GO via Java Web start (<http://www.blast2go.com/b2glaunch/startblast2go>) using 500 MB memory (see tutorial <http://www.blast2go.com/b2glaunch>).
6. Load every FASTA file (.fasta) by selecting protein sequence file. The unique sequences of the list will be read by the software adding information about the sequence length.
7. Run Blast step with default features (*e*-value cutoff 1×10^{-50} , number of Blast hits: 20) and selecting Blastp search in NCBInr database (see Note 34).
8. Once Blast step is finished run GO-mapping step. Automatically B2G tool retrieves GO terms to the loaded sequences.
9. In the next step, annotation is performed automatically using default parameters: *e*-value Hit-Filter of 1×10^{-6} , an Hsp-Hit

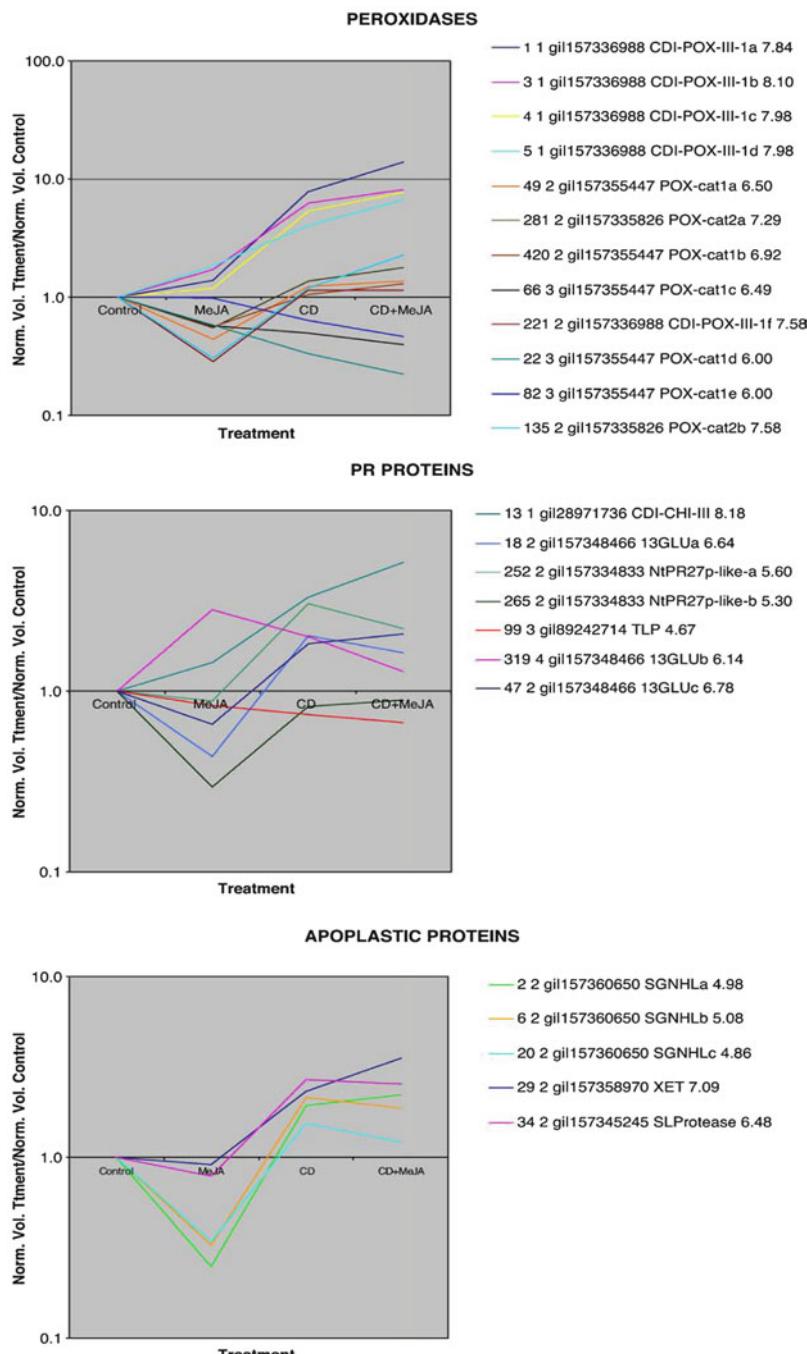


Fig. 3 Functional groups of identified proteins differentially expressed in elicited grapevine cell culture (cv. Gamay) extracellular medium. The identified proteins were distributed into three groups according to their known function: peroxidases, PR proteins other than peroxidases, and apoplastic proteins. Expression profiles are normalized in relation to control cultures and presented in log 10 scale. Legends on the right provide the following data of the protein: spot number, cluster number, gene index number, acronym, and pI (reproduced from [23] with permission from Elsevier)

coverage cutoff of 0, an annotation cutoff of 55, and a GO weight of 5 although this can be adjusted to the analysis requirements [44]. Run InterProScan and merge InterProScan GOs to annotation. Then, run GO-enzyme code mapping. At this stage, protein sequences are annotated and ready to apply functional analysis tool using B2G software (*see Note 35*).

4 Notes

1. Select a container twice the size of the final volume you want to prepare. Measure out approximately 90 % of the final required volume of water, e.g., 900 mL for a final volume of 1,000 mL, and while stirring the water, add the solid components of the medium and stir until completely dissolved and after adjusting pH, add additional water to bring the medium to the final volume.
2. Plant tissue culture media are generally sterilized by autoclaving at 121 °C and 1.05 kg/cm² (15–20 psi). The time required for sterilization depends upon the volume of medium in the vessel since the time required for the liquid volume to reach the sterilizing temperature (121 °C) is different.
3. Distribute the medium in Petri dishes (20 ml in each) when its temperature is about 45 °C and when the medium solidifies sealing the plates with Parafilm® to avoid dehydration.
4. Please check the list of compatible chemicals and potential interfering chemicals typically found in the protein extraction buffer.
5. Do not freeze vials with stock solution more than once.
6. It is best to prepare this fresh each time.
7. TEMED accelerates the decomposition of APS molecules into sulfate-free radicals and these, in turn, initiate the polymerization.
8. Simple method of preparing running buffer: Prepare 10× native buffer (0.25 M Tris, 1.92 M glycine, and 1 % SDS). Weigh 30.3 g Tris, 144 g glycine, and 10 g of SDS, mix, and make it to 1 L with water. The pH of this solution should not be adjusted and store at room temperature. Dilute 100 mL of 10× native buffer with 900 mL of distilled water.
9. The choice of the most appropriate rehydration solution for the sample depends on its specific protein solubility requirements. Urea solubilizes and denatures proteins, unfolding them to expose internal ionizable amino acids. For solubilization of the more hydrophobic proteins, use thiourea/urea [45].
10. LTQ Orbitrap (Thermo Fischer Scientific); XCT ion trap (Agilent Technologies). Search engines: Proteome Discoverer

software (Thermo Fischer Scientific), Spectrum Mill Proteomics Workbench (SMPW) (Agilent Technologies).

11. All plant organs including roots, stems, leaves, flowers, and embryos can be used as explants. The ability to form callus depends upon the organs and their developmental stages.
12. The disinfection process of plant material, which will be used as explant source for initiating in vitro culture, varies depending on the tissue and the nature of the explants. It is therefore necessary to perform a disinfection process optimization by varying not only the proportion of the disinfecting agent but also the time necessary for disinfection, to ensure success in the establishment of in vitro culture. 70 % ethanol improves the disinfected process since this increases the contact between the plant material and the disinfected agent. The disinfection process also improves by adding 0.1 % Tween 20 to 7 % calcium hypochlorite solution.
13. Sealing Petri dishes with Parafilm® to avoid dehydration.
14. In the case of grapevine (*Vitis vinifera* L.) SCC due to the extracellular accumulation of resveratrol and related stilbenes, secondary metabolites which are produced upon elicitation, a step prior to protein extraction is needed in order to remove these metabolites from the extracellular medium. The extracellular medium is treated with ethyl acetate to extract these stilbenes.
15. Once extracellular medium is depleted in stilbenes, residual polyphenol compounds are removed by incubation with PVPP. In general, plant tissues and plant cell cultures contain phenolic compounds which can modify proteins through an enzyme-catalyzed oxidative reaction interfering with protein extraction. PVPP or polyvinylpyrrolidone (PVP) removes phenolic compounds by adsorption.
16. In case the aim of the study is to preserve the enzymatic activity, protein precipitation should be performed in native conditions; thus TCA protein precipitation method is not appropriate for such analysis. Ammonium sulfate precipitation is a commonly used method for protein purification based on the alteration of the protein solubility. In the presence of high salt concentrations, the proteins tend to aggregate and precipitate out of solution. However, differential precipitation of proteins using ammonium sulfate presents some limitations that should be kept in mind as several contaminants will remain in solution (e.g., nucleic acids) and many proteins remain soluble at high salt concentration. This is the reason why this method is not recommended when total protein representation is desired being particularly helpful for protein purification based on the specific enrichment or pre-fractionation for a particular protein

population. Ammonium sulfate precipitation is performed as follows: dialyze 100 mL of the extracellular medium with 50 mM sodium acetate buffer pH 5.0 overnight at 4 °C to remove salts remaining in the culture medium. Add 65 g (NH₄)₂SO₄ (up to 95 % saturation) to 100 mL of extracellular medium and maintain for 1 h at 4 °C on the stirrer. The percent of ammonium saturation should be adjusted according to precipitate a particular protein fraction [46]. Centrifuge at 3,500×*g* for 20 min at 4 °C. Discard the supernatant, and resuspend the precipitated protein in 50 mM sodium acetate buffer pH 5.0. A last step for dialyzing protein sample should be introduced in order to remove salts which will interfere with isoelectrofocusing (IEF). Dialysis is carried out overnight at 4 °C in 50 mM sodium acetate buffer pH 5.0.

17. TCA is a very effective precipitant compound. TCA-based precipitation is a popular method for sample preparation of both one- and two-dimensional gel electrophoresis, because it can concentrate samples, remove salts and polysaccharides, and denature endogenous proteases. Limitations of TCA precipitation include difficulties in the protein resolubilization that can be solved by the aid of a sonicator probe. Special care should be paid to wash protein pellets with acetone or methanol to remove residual TCA prior to electrophoretic analysis or mass spectrometric analysis due to the potential protein degradation or modification under extended exposure to this low pH solution.
18. Do not overdry protein precipitate in order to facilitate its resolubilization. Protein precipitate is resuspended either in one- and two-dimensional electrophoresis sample buffer if the following analysis is performed by one- or two-dimensional electrophoretic analysis, respectively. For one-dimensional electrophoresis, resuspend the protein precipitate in lysis buffer (7 M urea, 2 M thiourea, 30 mM Tris-HCl, pH 9.0, and 4 % CHAPS). Store at 4 °C, and for two-dimensional electrophoresis use rehydration solution with IPG buffer (Subheading 2.8).
19. To avoid air bubble entrapment, insert the well-forming comb inclined at one edge.
20. Check the electrical connections on the cell to ensure that solution is not in contact with either banana plug, and connect the anode to the (+) terminal on the power supply, and the cathode to the negative terminal.
21. Preclude excessive heating by placing the chamber in a cold room.
22. For staining proteins a highly sensitive silver staining technique can be used which permits detection of polypeptides in polyacrylamide gels at concentrations 100-fold lower than Coomassie brilliant blue R-250.

23. Coomassie Brilliant Blue G is a sensitive stain for protein detection in PAGE gels. Coomassie staining gives blue bands on a background, with a sensitivity of 50–100 ng/band. All steps are done on a shaker with gentle mixing.
24. The optimal quantity of protein to load varies widely depending on factors such as sample complexity, the length (7, 11, 13, 18, and 24 cm) and pH range (from very acidic proteins at pH 3 to extremely basic proteins at pH 11) of the Immobiline DryStrip gel, and the method of visualizing the two-dimensional gel separations. The volume of rehydration solution used to solubilize the sample depends on the sample loading method (sample in-gel rehydration or cup-loading) and the length of the Immobiline DryStrip gel used for the first-dimension separation.
25. For increased resolution between pH 5 and pH 7, use a non-linear gradient pH 3–10 strip (pH 3–10 NL) to distribute the proteins more evenly over the strip.
26. Remove the protective cover from the surface of the IPG strip starting at the acidic (+) to prevent damage to the basic end of the gel, which is generally softer.
27. The pointed end of the holder is over the anode (+) (pointed to the back of the unit) and the blunt end over the cathode (−). Guide marks along the sides of the platform show approximate positioning for each strip holder size (7, 11, 13, 18, and 24 cm). Check that each of the two external electrode contacts makes metal-to-metal contact.
28. Overlay the resolving gel with water for gels having acrylamide concentration lower than 8 % and use isobutanol or isopropanol saturated with water for gels of 10 % or greater. This overlay prevents contact with atmospheric oxygen (which inhibits acrylamide polymerization) in addition to helping to level the resolving gel solution.
29. If necessary, store either tryptic peptides before or after desalting step at -20°C (<24 h) or -80°C (for long periods).
30. Prepare a fresh matrix solution each time or a stored solution at room temperature no longer than 48 h. Perform always a spin centrifugation to remove the non-dissolved matrix.
31. An additional control is put onto the MALDI target consisting of a trypsin digest from BSA protein. Calibration is performed according to a quadratic regression curve taking into account four or five points that bracketed the mass range of interest.
32. It is recommended to examine different regions of the spot to find those giving superior signal-to-noise ratios.
33. Database searching can be speeded up by using a restricted database. This database can be created with a subset of National

Center for Biotechnology Information non-redundant (NCBInr) protein database (<http://www.ncbi.nlm.nih.gov/>).

Sequences are retrieved using the keyword search enquiry “organism,” adding potentially contaminating proteins retrieved using the keyword search enquiry “trypsin” OR “keratin.”

34. Check manually one by one each protein description one to one all the loaded sequences within the project. The descriptions can be modified by clicking right button of the mouse over the protein entry and selecting “change annotation and description.”
35. Revise manually the GO terms associated to all the sequences loaded to the project. Once the Blast, mapping, and annotation processes are finished, the application offers the possibility of direct statistical analysis on protein function information. A common analysis is the statistical assessment of GO term enrichments in a group of interesting proteins when compared with a reference group [47]. This enrichment analysis is based on the Fisher’s Exact Test applying a robust false discovery rate (FDR) correction for multiple testing and returns a list of significant GO terms ranked by their corrected or one-test *p*-values. B2G performs various statistical charts which allow visualization of the results obtained at blasting, mapping, or annotation.

Acknowledgements

This work has been supported by research grants from MICINN-FEDER (BIO2008-2941, BFU2009-08151, and BIO2011-29856-C02-02). ABSJ and AL held a grant from the Ministerio de Ciencia e Innovación. MJME acknowledges a grant from Fundación Caja Murcia-University of Alicante. University of Alicante Proteomics Facility is a member of Proteored (ISCIII). JCV acknowledges support from Red Temática de Investigacion Cooperativa de Cáncer (RTICC).

References

1. Mulabagal V, Tsay HS (2004) Plant cell cultures – an alternative and efficient source for the production of biologically important secondary metabolites. *Int J Appl Sci Eng Technol* 2:29–48
2. Verpoorte R, Contin A, Memelink J (2002) Biotechnology for production of plant secondary metabolites. *Phytochem Rev* 1:13–25
3. van der Heijden R, Jacobs DI, Snoeijer W et al (2004) The *Catharanthus* alkaloids: pharmacognosy and biotechnology. *Curr Med Chem* 11:1241–1253
4. Almagro L, Bru R, Pugin A et al (2011) Integrated study of early signalling network in tobacco cell cultures elicited with methyl jasmonate and cyclodextrins. *Plant Physiol Biochem* 51:1–9
5. García-Brunner A, Lamotte O, Vandelle E et al (2006) Early signaling events induced by elicitors of plant defenses. *Mol Plant Microbe Interact* 19:711–724
6. Hellwig S, Drossard J, Twyman RM et al (2004) Plant cell cultures for the production of

- recombinant proteins. *Nat Biotechnol* 22: 1415–1422
7. Schlatmann JE, ten Hoopen HJG, Heijnen JJ (1996) Large-scale production of secondary metabolites by plant cell cultures. In: DiCosmo F, Misawa M (eds) *Plant cell culture secondary metabolism—toward industrial application*, 1st edn. CRC, Boca Raton, FL, pp 11–52
8. Vera-Urbina JC, Sellés-Marchart S, Martinez-Esteso MJ et al (2011) Production of grapevine cell biomass (*Vitis vinifera* L. Gamay) and resveratrol in custom and commercial bioreactors using cyclodextrins and methyl jasmonate elicitors. In: Delmas D (ed) *Resveratrol: source, production, and health benefits*. Nova, Hauppauge, NY
9. Drossard J, Nähring JM, Hellwig S et al (1997) Production of engineered antibodies in tobacco plants and cell suspension cultures. In: *Antibody engineering, New technology, application and commercialization*, vol. 2, IBC's 8th annual international conference on antibody engineering, pp 69–93 (IBC Library Services, Southborough, Massachusetts)
10. Fischer R, Liao YC, Drossard J (1999) Affinity-purification of a TMV-specific recombinant full-size antibody from a transgenic tobacco suspension culture. *J Immunol Methods* 226:1–10
11. Kwon TH, Seo JE, Kim J (2003) Expression and secretion of the heterodimeric protein interleukin-12 in plant cell suspension culture. *Biotechnol Bioeng* 81:870–875
12. Doran PM (2000) Foreign protein production in plant tissue cultures. *Curr Opin Biotechnol* 11:199–204
13. Fischer R, Liao YC, Hoffmann K et al (1999) Molecular farming of recombinant antibodies in plants. *Biol Chem* 380:825–839
14. Fischer R, Emans N, Schuster F et al (1999) Towards molecular farming in the future: using plant-cell-suspension cultures as bioreactors. *Biotechnol Appl Biochem* 30:109–112
15. Okushima Y, Koizumi N, Kusano T et al (2000) Secreted proteins of tobacco cultured BY2 cells: identification of a new member of pathogenesis-related proteins. *Plant Mol Biol* 42:479–488
16. Dani V, Simon WJ, Duranti M et al (2005) Changes in the tobacco leaf apoplast proteome in response to salt stress. *Proteomics* 5: 737–745
17. Tran HT, Plaxton WC (2008) Proteomic analysis of alterations in the secretome of *Arabidopsis thaliana* suspension cells subjected to nutritional phosphate deficiency. *Proteomics* 8:4317–4326
18. Sakurai N (1998) Dynamic function and regulation of apoplast in the plant body. *J Plant Res* 111:133–148
19. Pignocchi C, Kiddie G, Hernandez I et al (2006) Ascorbate oxidasedependent changes in the redox state of the apoplast modulate gene transcript accumulation leading to modified hormone signaling and orchestration of defense processes in tobacco. *Plant Physiol* 141:423–435
20. Seifert GJ, Blaukopf C (2010) Irritable walls: the plant extracellular matrix and signaling. *Plant Physiol* 153:467–478
21. Fry SC (2004) Primary cell wall metabolism: tracking the careers of wall polymers in living plant cells. *New Phytol* 161:641–675
22. Jamet E, Canut H, Boudart G et al (2006) Cell wall proteins: a new insight through proteomics. *Trends Plant Sci* 11:33–39
23. Martinez-Esteso MJ, Sellés-Marchart S, Vera-Urbina JC et al (2009) Changes of defense proteins in the extracellular proteome of grapevine (*Vitis vinifera* cv. Gamay) cell cultures in response to elicitors. *J Proteomics* 73: 331–341
24. Lippmann R, Kaspar S, Rutten T et al (2009) Protein and metabolite analysis reveals permanent induction of stress defense and cell regeneration processes in a tobacco cell suspension culture. *Int J Mol Sci* 10:3012–3032
25. Sabater-Jara AB, Almagro A, Belchí-Navarro S et al (2010) Induction of sesquiterpenes, phytosterols and extracellular pathogenesis-related proteins in elicited cell cultures of *Capsicum annuum*. *J Plant Physiol* 167: 1273–1281
26. Briceño Z, Almagro L, Sabater-Jara AB et al (2012) Enhancement of phytosterols, taraxasterol and induction of extracellular pathogenesis-related proteins in cell cultures of *Solanum lycopersicum* cv Micro-Tom elicited with cyclodextrins and methyl jasmonate. *J Plant Physiol*. doi:10.1016/j.jplph.2012.03.008
27. Novo-Uzal E, Gómez-Ros LV, Hernández JA et al (2009) Analysis of the soluble cell wall proteome of gymnosperms. *J Plant Physiol* 166:831–843
28. Sabater-Jara AB, Almagro L, Belchí-Navarro S et al (2011) Methyl jasmonate induces extracellular pathogenesis-related proteins in cell cultures of *Capsicum chinense*. *Plant Signal Behav* 6:440–442
29. Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol Plant* 15:473–497
30. Gamborg OL, Miller RA, Ojima K (1968) Nutrient requirements of suspension cultures

- of soybean root cells. *Exp Cell Res* 50: 151–158
31. Morel G, Wetmore RM (1951) Fern callus tissue culture. *Am J Bot* 38:141–143
 32. Blum H, Beier H, Gross HJ (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* 8:93–99
 33. Shevchenko A, Wilm M, Vorm O et al (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal Chem* 68:850–858
 34. Perkins DN, Pappin DJ, Creasy DM et al (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20:3551–3567
 35. Yates JR, Eng JK, McCormack AL et al (1995) Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal Chem* 67:1426–1436
 36. Lopez-Llorca LV, Gómez-Vidal S, Monfort E et al (2010) Expression of serine proteases in egg-parasitic nematophagous fungi during barley root colonization. *Fungal Genet Biol* 47: 342–351
 37. Rabilloud T, Valette C, Lawrence JJ (1994) Sample application by in-gel rehydration improves the resolution of two-dimensional electrophoresis with immobilized pH gradients in the first dimension. *Electrophoresis* 15: 1552–1558
 38. Görg A, Obermaier C, Bought G et al (2000) The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 21:1037–1053
 39. Neuhoff V, Arold N, Taube D et al (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9:255–262
 40. Udina F (2005) Interactive biplot construction. *J Stat Softw* 13:1–16
 41. Kinter M, Sherman N (2000) Protein sequencing and identification using tandem mass spectrometry. Wiley, New York, pp 166–206
 42. Link AJ, Eng J, Schieltz DM et al (1999) Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* 17:676–682
 43. Conesa A, Götz S, Garcia-Gomez JM et al (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676
 44. Götz A, García-Gómez JM, Terol J et al (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36:3420–3435
 45. Rabilloud T, Adessi C, Giraudel A et al (1997) Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 18: 307–316
 46. Englard S, Seifter S (1990) Precipitation techniques. *Methods Enzymol* 182:285–300
 47. Martínez-Esteso MJ, Casado-Vela J, Sellés-Marchart S et al (2011) iTRAQ-based profiling of grape berry exocarp proteins during ripening using a parallel mass spectrometric method. *Mol Biosyst* 7:749–765

Chapter 30

Pollen Cultivation and Preparation for Proteomic Studies

Heidi Pertl-Obermeyer and Gerhard Obermeyer

Abstract

The quality of the collected experimental data very much depends on the quality of the biological starting material. Especially the proteome analysis of a highly dynamic system like the germinating and tube-growing pollen grain needs several precautions which allow an accurate and acceptable interpretation of the obtained results. Optimized protocols for pollen collection, storage, and in vitro culture as well as pollen organelle separations are described which help to obtain well-defined and reproducible experimental conditions for the subsequent proteomic analysis.

Key words Allergy, Germination, *Lilium longiflorum*, Organelle isolation, Marker enzyme, Pollen grain, Pollen tube, Tip growth

1 Introduction

For several reasons the male gametophytes of plants (the pollen) are interesting subjects for proteome research. Firstly, they are the source of allergens affecting human health. Secondly, the germination of pollen grains and the subsequent pollen tube growth are rapid and highly dynamic processes during plant reproduction which, in turn, is the basis for food production for a growing human population. Finally, the special growth process known as tip growth, by which the pollen tubes elongate, is a complex network of interacting proteins and signal transduction cascades. Table 1 presents some of the important pollen proteome studies. No proteomic studies on pollen allergens are included because the majority of these reports use commercially available pollen grains that are hardly able to germinate anymore and mainly focus on the allergy-causing properties of proteins but not on their physiological role.

To investigate the various biological questions related to pollen allergy and pollen physiology, the source and the way how the biological material is prepared are essential prerequisites for the accuracy of the subsequent analysis. Here we describe several methods to obtain and prepare defined protein fractions from

Table 1
Relevant physiological pollen proteome studies

| Species | Developmental state | | Type of proteomic study | Reference |
|--------------------------------|--------------------------------|--|-------------------------------|--|
| | Pollen fraction | | | |
| <i>Arabidopsis thaliana</i> | Mature pollen | | 2-D gels, spot identification | Noir S et al. (2005) Biochem Biophys Res Comm 337: 1257–1266 |
| | Total protein | | | |
| <i>Arabidopsis thaliana</i> | Mature pollen | | 2-D gels, spot identification | Holmes-Davies R et al. (2005) Proteomics 5: 4864–4884 |
| | Total protein | | | |
| <i>Arabidopsis thaliana</i> | Pollen grains and tubes | | 2-D gels, spot identification | Zhou J et al. (2009) J Integrative Plant Biol 51: 438–455 |
| | Total protein | | | |
| <i>Arabidopsis thaliana</i> | Mature pollen | | Shot gun | Grobei MA et al. (2009) Genome Res 19: 1786–1800 |
| | Total protein | | | |
| <i>Arabidopsis thaliana</i> | Pollen grains and tubes | | 2-D gels, spot identification | Ge W et al. (2011) BBA 1814: 1964–1973 |
| | Pollen coat proteins | | | |
| <i>Oryza sativa</i> | Mature pollen | | 2-D gels, spot identification | Dai S et al. (2006) Proteomics 6: 2504–2529 |
| | Total and pollen coat proteins | | | |
| <i>Oryza sativa</i> | Pollen grains and tubes | | 2-D gels, spot identification | Dai S et al. (2007) Mol Cell Proteomics 6: 207–230 |
| | Total protein | | | |
| <i>Lycopersicon esculentum</i> | Mature pollen | | 2-D gels, spot identification | Sheoran IS (2007) J Exp Bot 58: 3525–3535 |
| | Total protein | | | |
| <i>Lycopersicon esculentum</i> | Mature pollen | | Shot gun | Lopez-Casdo G et al. (2012) Proteomics 12: 761–774 |
| | Total protein | | | |
| <i>Canola (Brassica napus)</i> | Mature and germinated | | 2-D gels, spot identification | Sheoran IS (2009) Planta 230: 779–793 |
| | Total and secreted proteins | | | |
| <i>Lilium longiflorum</i> | Young, mature, and germinated | | 2-D gels, spot identification | Miki-Hirosige H et al. (2004) Sex Plant Reprod 16: 209–214 |
| | Total protein | | | |
| <i>Lilium longiflorum</i> | Mature and germinated | | Shot gun | Pertl et al. [1] |
| | Membrane proteins | | | |
| <i>Lilium davidii</i> | Mature and germinated | | Shot gun | Han B et al. (2010) J Integrative Plant Biol 52: 1043–1058 |
| | Plasma membrane proteins | | | |
| <i>Picea meyeri</i> | Germinated pollen | | 2-D gel, spot identification | Chen T et al. (2009) Plant Physiology 149: 1111–1126 |
| | Total protein | | | |
| <i>Picea meyeri</i> | Germinated pollen | | 2-D gel, spot identification | Chen Y et al. (2006) Plant J 47: 174–195 |
| | Total protein | | | |
| <i>Quercus ilex</i> | Mature pollen | | 2-D gel, shot gun | Galvan JV et al. (2012) J Proteomics 75: 2736–2744 |
| | Total protein | | | |

pollen grains and tubes used for proteomic analyses. The protocols were initially optimized for *Lilium* pollen [1, 2] but can be easily adapted to other pollen species.

2 Materials

The pollen viability, namely, their ability to germinate in appropriate artificial media, depends not only on the plant growth conditions during pollen development but also very much on the collection time (state of pollen development) and storage conditions. Note that optimal conditions may vary between plant species and should be tested. For instance, the lily pollen is stored at -80°C whereas tobacco pollen grains (*Nicotiana tabacum*) “survive” for 1–2 years at -20°C . How pollen grains are stored also depends on the subsequent studies, e.g., for the identification of pollen allergens, pollen grains may be stored at room temperature whereas for studies on pollen germination long-term storage in a freezer is necessary.

2.1 Collection and Storage of Pollen Grains for In Vitro Cultures

1. Collect mature, fully dehydrated anthers (see Note 1) from flowers with a pair of tweezers and freeze them immediately in a tube filled with liquid nitrogen.
2. Carefully close the lid (see Note 2) and store the shock-frozen pollen grains at -80°C (*Lilium longiflorum*) or -20°C (*N. tabacum*, *Arabidopsis thaliana*).

2.2 Collection and Storage of Pollen Grains for Allergen Studies

1. Collect pollen grains from the plants during a dry weather period with almost no wind into an appropriate plastic centrifugation tube (see Note 3).
2. Sieve the pollen grains through a nylon mesh of appropriate mesh width to separate the grains from other plant material. Nylon meshes with standard widths are offered by general laboratory suppliers or can be obtained from special companies (e.g., Pieper Filter GmbH, Bad Zwischenahn, Germany, www.pieper-filter.de).
3. Store pollen grains at room temperature in a dry, cool, and dark place. Avoid direct sunlight as well as temperature and humidity changes as they may damage the pollen grains (see Note 4).

2.3 Germination Media

The following media were tested in the authors’ lab for germination and tube growth of pollen grains from various species. Notice that pollen grains of *A. thaliana* are difficult to cultivate in vitro and often require specific attention concerning the plant ecotype, developmental state of the flower, the temperature regime during plant growth, and others [3]. So far, the best results regarding the total germination frequency, synchrony of germination, and

Table 2
Stock solutions for testing various media for pollen in vitro cultures

| Solution | MW (g/l) | Grams per 50 ml |
|--|-----------------|------------------------|
| 1 M KCl | 74.56 | 3.7280 |
| 100 mM CaCl ₂ | 147.02 | 0.7351 |
| 100 mM Ca(NO ₃) ₂ | 236.15 | 1.1801 |
| 100 mM MgCl ₂ | 203.31 | 1.0166 |
| 100 mM MgSO ₄ | 246.48 | 1.2324 |
| 100 mM H ₃ BO ₃ | 61.83 | 0.3092 |
| 100 mM KH ₂ PO ₄ | 136.09 | 0.6805 |

amount of biological material which are important for pollen proteomic studies were obtained for lily pollen.

When testing various media conditions for optimizing pollen grain germination and tube growth, prepare stock solutions of the media ingredients (Table 2). The culture media listed below have been tested in the authors' lab. However, other media compositions are reported in the literature.

1. Medium B (lily): 10 % (w/v) sucrose, 1.6 mM H₃BO₃, 1 mM KCl, 0.1 mM CaCl₂. Adjust to pH 5.6 with tiny amounts of MES or Tris (*see Note 5*).
2. Medium M (lily): 300 mM mannitol, 1.6 mM H₃BO₃, 1 mM KCl, 0.1 mM CaCl₂, 25 mM MES adjusted with BTP to pH 5.6.
3. Medium P (lily): 20 % (w/v) PEG (3350), 1.6 mM H₃BO₃. Adjust to pH 5.6 with tiny amounts of Tris or BTP.
4. Arabidopsis pollen medium (solidified, [4]): 18 % (w/v) sucrose, 0.01 % boric acid, 1 mM MgCl₂, 1 mM CaCl₂, 1 mM Ca(NO₃)₂, 0.5 % (w/v) agar, pH 7 adjust with HCl or KOH.
5. Arabidopsis pollen medium (solidified, [3]): 10 % (w/v) sucrose, 0.01 % (w/v) boric acid, 5 mM CaCl₂, 5 mM KCl, 1 mM MgSO₄, pH 7.5 adjust with ca. 30 µl of 0.1 M NaOH (non-buffered), 1.5 % (w/v) low-melting agarose (Sigma, type VII, A4018). Add sucrose after pH was adjusted.
6. Arabidopsis pollen medium (liquid, [3]): 0.01 % boric acid, 5 mM CaCl₂, 5 mM KCl, 1 mM MgSO₄, 10 % (w/v) sucrose, pH 7.5 (non-buffered).
7. Arabidopsis pollen medium (solidified, [5]): 20 % (w/v) sucrose, 0.001 % (w/v) H₃BO₃, 20 mM CaCl₂, 0.1 mM KH₂PO₄, 12 % (w/v) PEG 4000, 0.5 % (w/v) agar, pH 7.

8. Arabidopsis pollen medium (liquid, [6]): 18 % (w/v) sucrose, 1 mM KCl, 0.49 mM H_3BO_3 , 1 mM $MgSO_4$, 2 mM $CaCl_2$, 2 mM $Ca(NO_3)_2$, pH 7. Germinate in 3 ml medium in a 25 ml Erlenmeyer.
9. Arabidopsis pollen medium (solidified, [6]): 18 % (w/v) sucrose, 1 mM KCl, 0.49 mM H_3BO_3 , 1 mM $MgSO_4$, 2 mM $CaCl_2$, 2 mM $Ca(NO_3)_2$, pH 7, 1 % (w/v) agar.
10. Tobacco pollen medium [7]: 6 % (w/v) sucrose, 1.6 mM H_3BO_3 , 200 μM $CaCl_2$, 1 mM MES adjusted with NaOH to pH 5.5.

2.4 Buffers and Solutions for Membrane Preparations

1. Homogenization buffer: 330 mM sucrose, 100 mM KCl, 5 mM DTT, 1 mM EDTA, 50 mM Tris adjusted with MES to pH 7.2. Protease inhibitors are added from stock solution to the ice-cold homogenization buffer just before use to give the final concentrations of 10 μM leupeptin, 1 μM pepstatin A, 1 mM PMSF, and 2 μM E-64 (see Note 6).
2. Centrifugation buffer: 1 mM $MgSO_4$, 1 mM Tris adjusted with MES to pH 7.2.
3. Sucrose solutions: 18 % (w/w); weight in 1.8 g sucrose and fill up with centrifugation buffer to 10 g. The other sucrose solutions (25, 30, 34, 38, 45 % (w/w)) are prepared the same way.

2.5 Solutions for ATP Hydrolysis Assays

1. Solution A: Dissolve subsequently the following chemicals in 750 ml distilled water in a glass beaker: 4.2 g ammonium molybdate, 28.6 ml concentrated H_2SO_4 , and 20 g sodium dodecyl sulfate (SDS). Fill up to 1 l with distilled water (see Note 7).
2. Solution B: 10 % (w/v) ascorbic acid. Prepare fresh before starting the experiment.
3. Calculate the needed total volume of the phosphate reagent for an experiment and prepare the reagent by mixing 6 volumes of solution A and 1 volume of solution B. Prepare fresh before the experiment.

2.6 Reaction Buffers for ATP Hydrolysis

1. Vanadate-sensitive plasma membrane H^+ ATPase (P-ATPase): 25 mM K_2SO_4 , 50 mM MES/Tris pH 6.8, 0.1 mM EDTA, 4 mM $MgSO_4$, 100 μM Na-molybdate, 0.005 % Triton X100; add directly to cuvettes 100 nM baflomycin, 1 mM sodium azide, $\pm 200 \mu M$ sodium vanadate (see Note 8).
2. Baflomycin-sensitive vacuolar H^+ ATPase (V-ATPase): 25 mM K_2SO_4 , 50 mM Tris/MES pH 7.5, 0.1 mM EDTA, 4 mM $MgSO_4$, 100 μM Na-molybdate, 0.005 % Triton X100; add directly to cuvettes 200 μM sodium vanadate, 1 mM sodium azide, $\pm 100 nM$ baflomycin.

3. Azide-sensitive mitochondrial H⁺ ATPase (F-ATPase): 25 mM K₂SO₄, 50 mM Tris/MES pH 8.5, 0.1 mM EDTA, 4 mM MgSO₄, 100 µM Na-molybdate, 0.005 % Triton X100; add directly to cuvettes 100 nM bafilomycin, 200 µM sodium vanadate, ±1 mM sodium azide.

2.7 Substrate Stock Solution

100 mM sodium ATP: Prepare the ATP stock solution by dissolving sodium ATP in a buffer of 25 mM BTP of pH 8.0 (see Note 9).

3 Methods

3.1 Pollen Culture

In general, no differences in germination frequency, tube length, or tube morphology can be observed between fresh and frozen pollen grains which are stored under optimal conditions. However, some pollen species need a gentle “warm-up” when taken from the -80 °C freezer and a few minute incubation in a humid chamber (imbibition) helps to obtain high germination frequencies. For some pollen species seasonal changes in the germination frequency and the tube length can be observed. In lily pollen, the seasonal variation in their germination capability during the year is still conserved in the frozen state at -80 °C (Fig. 1a). Lily pollen grains collected and aliquots frozen in June/July germinate quite reasonable in July, August, and September when incubated in culture medium. The germination rate declines during the winter months and, surprisingly, increases again from April to June! Furthermore, in the winter months no reproducible pattern of the expression of the plasma membrane H⁺ ATPase could be observed. The variations of the time-dependent expression were really high (Fig. 1b) and during these months no reproducible data could be obtained neither by immunodetection nor by identification of peptides via mass spectrometry analysis.

3.1.1 Lily Pollen Culture

1. Take fresh or frozen lily pollen grains (*L. longiflorum*) from one anther and resuspend in 6–12 ml lily pollen germination medium (see Note 10).
2. Incubate for up to 4 h. Usually, pollen tubes emerge after 40–60 min (Fig. 2). Incubation for longer times than 4 h may yield in a high amount of bursting pollen tubes which will falsify the results (see Note 11).

3.1.2 *Arabidopsis* Pollen Culture

1. Collect flowers from *Arabidopsis* plants and dehydrate for 1–2 h at room temperature [4, 6].
2. Imbibe anthers, entire flowers, or frozen pollen grains for 30 min at RT in a humid chamber.
3. Dip flowers onto the solidified agar germination medium or dust pollen grains onto the agar surface. In case of liquid cultures resuspend pollen grains in germination medium (40 flowers

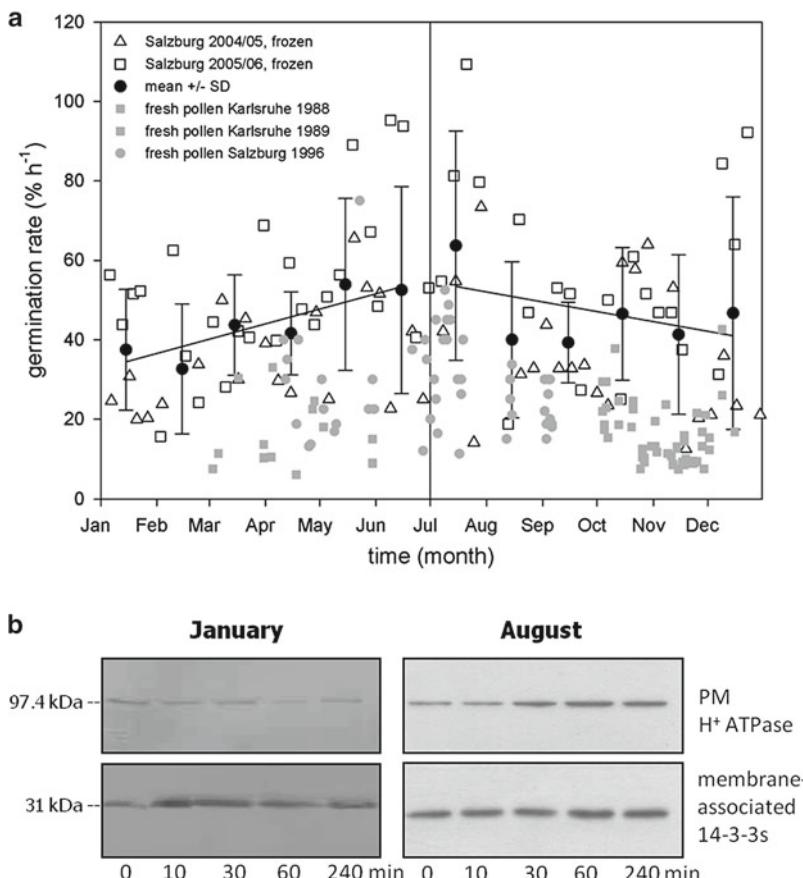


Fig. 1 Seasonal changes of lily pollen germination. (a) Mature pollen grains of *Lilium longiflorum* were either cultivated immediately in medium B (fresh) or stored according to the described method at -80°C (frozen). Germination data were collected during the seasons and at locations as indicated (Karlsruhe, Germany, or Salzburg, Austria). In seasons 2004–2005 and 2005–2006, the pollen grains were frozen in liquid nitrogen during June and stored at -80°C . Germination assays were performed once in a week until next year June. A regression line drawn during the monthly average germination rate (filled circle) which was calculated from both seasons clearly shows a higher germination frequency during April to September than in the months October to March. (b) Detection of the plasma membrane H⁺ ATPase and membrane-associated 14-3-3 proteins in microsomal fractions prepared from lily pollen grains. The time-dependent increase in the amount of PM ATPases was only detectable and reproducible during April to September. In all other months with low germination capability the pattern of PM ATPase expression varied irreproducibly

per 250 μl in liquid medium [3] or germinate in 3 ml medium in a 25 ml Erlenmeyer [6] (see Note 12)).

4. Incubate for 6–16 h at RT (22°C [3]). First pollen tubes become visible after 1–2 h (Fig. 3).
- 3.1.3 Tobacco Pollen Culture**
1. Resuspend tobacco pollen grains from five flowers in 3–6 ml germination medium in a small Petri dish (see Note 13).
 2. Incubate for 3–4 h. Pollen tubes can be observed after 30–60 min.

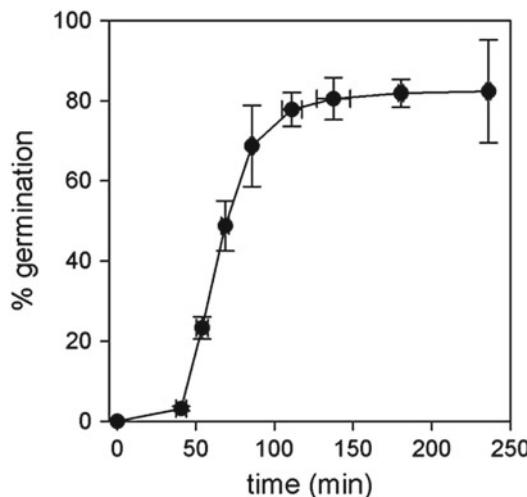


Fig. 2 Germination of lily pollen grains in Med M. For each time point 200–300 pollen grains were counted. Five independent repetitions, mean \pm S.D.

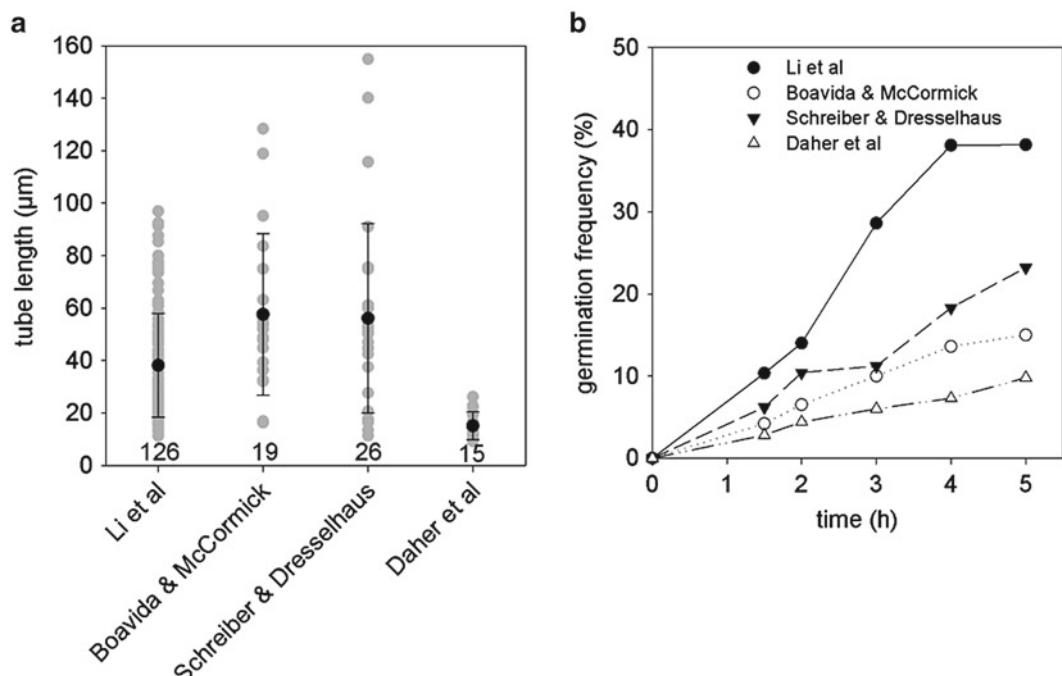


Fig. 3 In vitro culture of *Arabidopsis thaliana* pollen in different germination media. Fresh pollen grains were collected from flowers in which the anthers were still shorter than the pistils and media were prepared according to [3–6]. (a) Tube length after 5-h incubation in the media. Numbers indicate the number of pollen tubes measured. (b) Germination frequency of *Arabidopsis* pollen grains in the given media

3.1.4 *Medicago truncatula* Pollen Culture

1. Spread *Medicago truncatula* pollen grains onto the solidified *Arabidopsis* germination medium described by Li et al. [4] and incubate for up to 5 h. Pollen tubes can be seen after 45 min.

3.2 Preparation of Pollen Exudates

To identify allergenic pollen proteins, pollen grains are often incubated in solutions to release the allergenic proteins. It has to be noted that the fractions obtained by this method have to be denoted as “pollen exudates” or “pollen diffusates” to distinguish them from pollen extracts which are prepared by damaging the pollen grains. Indeed, damage of pollen grains has to be avoided when allergens should be identified in subsequent proteomic experiments because otherwise the high amounts of intracellular proteins would cover the less abundant allergens.

1. Resuspend 100 mg pollen grains in 1 ml medium (*see Note 14*).
2. Gently shake for 20 min at RT or use an end-over-end rotator.
3. Pellet larger pollen grains, e.g., lily pollen, by centrifugation at $3,000 \times \text{g}$ for 15 min at RT. Collect the supernatant (=pollen exudates [8]). Smaller pollen grains (e.g., pollen grains from *Arabidopsis*, birch, mugwort) should be filtered through 5 μm centrifugal filter units (e.g., UFC30SV00, Millipore, Billerica, MA, USA) using $1,000 \times \text{g}$ and 5 min at RT.
4. Dialyze the pollen exudates against water or 50 mM NH_4CO_3 and lyophilize the solution for long-term storage at -20°C .

3.3 Preparation of Homogenates, Cytosolic and Membrane Fractions from Pollen

The quality of the preparation and, thus, the quality of the investigated proteins very much depend on performance during the preparation steps. All steps should be performed in an ice bath ($0\text{--}4^\circ\text{C}$) and speedily. The following procedure was optimized for lily pollen to investigate the membrane proteome [1, 9], but can be adapted to other pollen species. The starting material is sufficient to allow organelle membrane isolations and to perform several electrophoresis and immunodetection assays.

1. Resuspend pollen grains from 25 flowers in germination medium or water. Pellet the pollen grain ($1,000 \times \text{g}$, 15 min, 4°C) and resuspend in ice-cold homogenization buffer.
2. Homogenize the pollen grains with a teflon pistil (Potter-Elvehjem type, Potter S from Sartorius, Germany) by circa 20 down- and upward strokes (*see Note 15*).
3. Filter the homogenate through a 10–30 μm nylon mesh (*see Note 16*). Keep the filtrate.
4. Centrifuge the filtrate at $7,500 \times \text{g}$, for 15 min, at 4°C . Decant the supernatant carefully into a high-speed centrifugation tube (*see Note 17*). Try to keep the lipids which float on top of the medium in the tube.
5. Centrifuge the supernatant at $48,000 \times \text{g}$ for 75 min at 4°C to pellet all membranes.
6. Decant the supernatant carefully without the floating lipids. This supernatant represents the cytosolic fraction which contains all soluble proteins including those from organelle lumens.

7. Remove the lipids from the inner tube walls carefully with a paper towel. Do not touch the pellet!
8. Resuspend the pellet in centrifugation buffer (*see Note 18*). Usually, $2 \times 500 \mu\text{l}$ are sufficient. This is the microsomal or the membrane fraction which contains all endomembranes (ER, Golgi, mitochondria, vacuole) and the plasma membrane. Determine the protein concentration of the fractions (e.g., Lowry assay compatible with DTT!) and store it at -20°C (1 month) or -80°C (1 year).

3.4 Isolation of Organelle Membranes on a Discontinuous Sucrose Density Gradient

Although this isolation method is quite old, the big advantage is the possibility to isolate vacuolar, ER, Golgi, mitochondrial, and plasma membrane vesicles in one preparation step simultaneously and from the same pool of pollen grains. Aqueous two-phase partitioning may get organelle fractions of higher purity but can be optimized just for the isolation of one membrane fraction. A preparative free-flow electrophoresis system also allows the simultaneous separation and isolation of different organelle vesicles but is extremely expensive and labor intensive.

1. Prepare the sucrose solutions the day before and store them at 4°C .
2. Prepare the sucrose step gradient in a high-speed centrifugation tube (36 ml) by carefully pipetting layers of sucrose solutions. Start with the highest sucrose concentration: 4 ml of 45 % followed by 7 ml of 38 %, 34 %, 30 %, 4 ml of 25 %, and finally 4 ml of 18 % sucrose solution (*see Note 19*). Experienced researchers may add the sucrose solutions manually using a normal pipette but sucrose layers should not be mixed!
3. Layer up to 1 ml MF on top of the sucrose gradient. Homogenize the MF after it has been thawed up on ice with a small glass homogenizer. Keep the MF cool.
4. Tare the centrifugation tubes carefully to at least 1 mg tolerance and centrifuge in a swing-out rotor at $100,000 \times g$ for 2 h 15 min at 4°C (Sorvall rotor AH-629, ca. 28,000 rpm).
5. Collect the interphases (Fig. 4a) with a Pasteur glass pipette whose opening has been bended by 90° , into reaction tubes. Keep interphases on ice.
6. Determine the sucrose concentration of the interphase fractions using a refractometer.
7. Dilute the interphase fractions to at least 10 % sucrose and pellet the membranes by centrifugation at $145,000 \times g$ for 60 min at 4°C in a fixed-angle rotor.
8. Discard supernatant and resuspend the pellet carefully in centrifugation buffer (*see Note 18*).
9. The fractions can be stored at -80°C for up to 3 months without severe loss of enzyme activities.

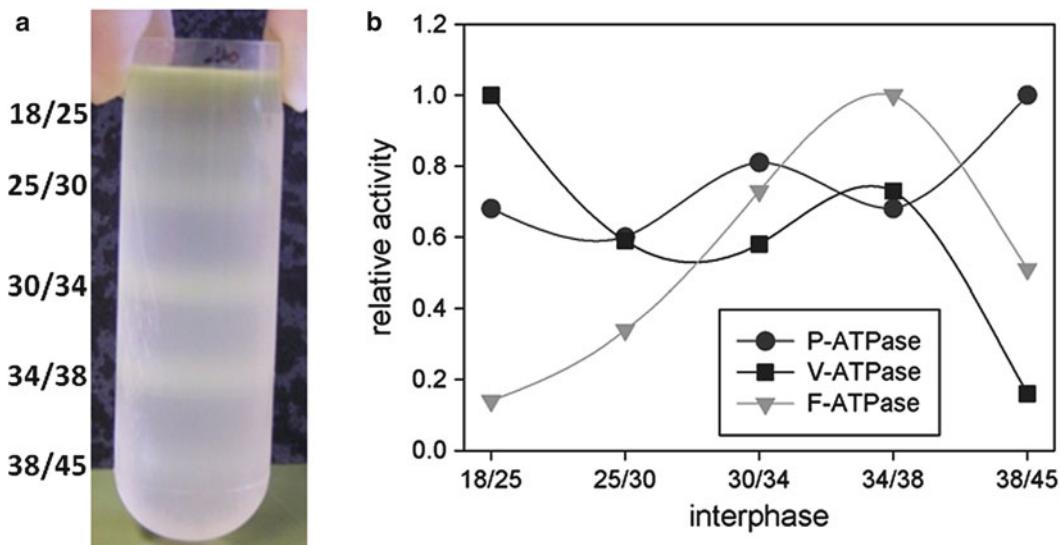


Fig. 4 Sucrose step gradient after centrifugation. The organelles from a microsomal fraction obtained from pollen tubes grown for 240 min (MF_{240}) were separated. (a) The interphases which correspond to the enriched organelle fractions as indicated are clearly visible. (b) An example of an ATP hydrolysis assay analysis using the P-, V-, and F-type ATPases as marker enzymes for the tonoplast, the inner mitochondrial membrane, and the plasma membrane, respectively

3.5 Marker Enzymes and ATP Hydrolysis Assay

The marker enzyme concept is based on the fact that certain enzyme activities co-localize with just one organelle membrane. However, due to membrane and protein trafficking some proteins or enzyme activities can be detected in many organelles but the main signal is always specific for one organelle. An excellent description of the marker enzyme concept, its critical evaluation, and detailed protocols for measuring enzyme activities are given [10–12]. Here, we present an ATP hydrolysis assay to distinguish between the P-type H^+ ATPase of the plasma membrane, the V-type H^+ ATPase of the tonoplast, and the F-type H^+ ATPase of the inner mitochondrial membrane. The ATP hydrolysis assay is based on the detection of the released inorganic phosphate and was first described by Fiske and Subbarow in 1925 [13] with modifications [14, 15]. Due to the fact that the different ATPase types can be inhibited by specific inhibitors, the ATP hydrolysis assay can be performed in the absence and the presence of a specific inhibitor to determine the amounts of vanadate-sensitive (P-type ATPase), bafilomycin-sensitive (V-type), and azide-sensitive (F-ATPase) ATP hydrolysis.

1. Label plastic half-micro cuvettes. Perform the assay in triplicates and add cuvettes for the following blank values: buffer blank (no substrate (ATP), no protein), substrate blank (with substrate, no protein), and vesicle blank (no substrate, protein).
2. Add an appropriate volume of reaction buffer to the cuvettes. The total volume of the reaction is 300 μ l. Consider the volumes

of added inhibitors, vesicle fractions, substrate, etc. Add the inhibitors, e.g., 200 μ M vanadate and 100 nM baflomycin for measuring the azide-sensitive ATP hydrolysis. Add the specific inhibitor to one-half of the cuvettes, in this example 1 mM azide, to determine the difference of ATP hydrolysis activity in the presence and absence of azide (azide-sensitive ATP hydrolysis activity, *see Note 8*).

3. Add equivalent volume of the vesicle/interphase fractions to achieve between 5 and 10 μ g protein per cuvette. Mix well with the pipette tip and incubate for 30 min at RT.
4. Start the reaction with 9 μ l 100 mM ATP stock solution (*see Note 20*). Mix well and incubate the reaction for 60 min (*see Note 21*).
5. Stop the reaction by addition of 700 μ l phosphate reagent. Incubate for color development for 60 min and measure the absorption at 820 nm.
6. The absorption value ($A_{vesicle}$) is corrected for blank values to obtain the amount of liberated phosphate only by enzyme activity as follows:

$$A_{corr} = A_{vesicle} - A_{bufferblank} - (A_{vesicleblank} - A_{bufferblank}) - (A_{substrateblank} - A_{bufferblank}) \\ = A_{vesicle} - A_{vesicleblank} - A_{substrateblank} + A_{bufferblank}$$

The corrected absorbance value can be used to calculate phosphate concentration using a calibration curve.

4 Notes

1. Mature anthers can be identified by their exposed pollen grains. Adjust the tube size to the amount of pollen grains you need to store: e.g., lily pollen are best stored in 1.5 ml reaction tubes for single anthers, 15 ml centrifugation tubes for up to 5 flowers, and 50 ml centrifugation tubes can be used to store anthers from 25 to 50 flowers. Sometimes, anthers can be dried at room temperature in a desiccation vessel before freezing.
2. To avoid exploding of the tube, pinch a small hole into the lid and let the liquid nitrogen completely evaporate before closing the lid. Wear protection gloves and goggles!
3. Keep in mind that most allergenic pollen grains originate from wind-pollinating plants which means that you hardly collect any pollen grains on a windy day. For best amounts of pollen allergens, pollen grains should be dry. Otherwise they have already lost most of their proteins from the pollen coat. Use a pierced lid to close the tube. Always avoid humidity.

4. We stored birch pollen grains for more than 1 year on a lab bench without any loss of allergens or changes in the allergen release kinetics.
5. A solution of sucrose and boric acid is a weak buffer system and usually has a pH around 5.6 at room temperature. However, by addition of small crystals of MES or Tris the pH can be adjusted. Whenever pH is an important parameter, use medium M or add 10–25 mM MES to medium B and adjust with Tris.
6. The following stock solutions of the protease inhibitors were prepared: 10 mM pepstatin A in DMSO, 10 mM leupeptin in distilled water, 1 M PMSF in DMSO, and 1 mM E-64 in distilled water.
7. Be sure that all chemicals used do not contain traces of phosphate as this will not only reduce the assay's sensitivity but can also make the detection impossible. The phosphate reagent will turn blue immediately after preparation! To remove traces of phosphate from the glassware, all beakers, cylinders, magnetic stirrer bars, etc. were washed with methanol, followed by 1:5 diluted HCl and methanol and three times exhaustive rinsing with distilled water. We use plasticware which is also treated as described above and used only for the phosphate reagents. Never use phosphate-containing detergent for cleaning lab ware when working on phosphate determination!
8. The following inhibitor stock solutions are used: 10 mM sodium vanadate (Na_3VO_4) in distilled water, 15 μM baflomycin A1 in DMSO, and 100 mM sodium azide (NaN_3) in distilled water.
9. ATP may hydrolyze at acidic pH values and the released phosphate will decrease the test's sensitivity. Autohydrolysis of ATP during the assay time is monitored in the substrate blank cuvette.
10. Fresh lily pollen grains stick to each other by their pollenkitt. One anther can be placed in 1 ml germination medium. Fierce and intense shaking of the tube helps to wash the pollen grains from the anther. The anther can be removed and the 1 ml pollen suspension can be transferred to a new tube. Washing with 5 % ethanol solution also helps to remove the pollenkitt and to resuspend pollen grains, but may affect the germination rate.
11. Highest germination frequencies (>90 %) were obtained by resuspending 30 mg lily pollen in 20 ml germination medium M which were transferred to a large Petri dish (145 mm diameter, Greiner, Kremsmünster, Austria). As a thumb rule, one lily anther contains 12 mg (\approx 40,000) pollen grains.
12. We never succeed in germinating useful amounts of *Arabidopsis* pollen in liquid media for biochemical analyses.

13. Approximately 50 tobacco flowers are needed for 20 mg pollen grains.
14. A large variety of solutions and media has been used for preparation of pollen exudates (water, germination media, phosphate-buffered saline (PBS), or NH_4CO_3). Note that the choice of solution often selects already the type of proteins that are released [8]. Especially the ionic strength and the pH of the solution affect the release of proteins from the pollen coat because the cell wall acts as an ion exchange matrix.
15. The upward movement of the pistil creates a negative pressure and causes the pollen grains to burst. Carefully move the pistil without sucking too much air into the solution. Other ways of breaking up pollen grains are possible, e.g., ultrasonication, freezing in liquid nitrogen, and homogenization by pistil and mortar. However, these methods generated a large number of mixed vesicles which cannot be separated anymore by means of marker enzyme analysis. Mixed vesicles occur during the homogenization step by vesicle bursting followed by resealing or fusion with other burst vesicles.
16. The filter retains all exine shell, partially broken pollen grains, and some large organelles (nuclei, generative cells). Smaller mesh width (5 μm) can be used without any severe loss of material.
17. The pellet contains heavy organelles like starch grains, nuclei, generative or sperm cells, and mitochondria.
18. Carefully resuspend the microsomal pellet on ice by pipetting the centrifugation buffer over the pellet until it resolves by itself. This can take up to 10 min! Never pierce into the pellet with the pipette tip or try to mix the pellet with the pipette tip. This will create large MF lumps but not a homogeneous fraction which is important for reproducible results in the following experiments, e.g., equal loading of SDS-PAGE gels, ATP hydrolysis assays, or H^+ transport experiments.
19. It has to be noted that the sucrose step gradient has to be optimized for every system to obtain optimal organelle vesicle separation. First experiments are run with a continuous sucrose gradient and the distribution of organelle membranes is monitored with marker enzyme assays, by proteomic identification of marker enzyme peptides, or by immunodetection of organelle markers along the gradient. A step gradient centrifugation is much faster than the centrifugation of a continuous gradient (ca. 16 h) and the sucrose steps can be adapted to the special needs of the experiments (e.g., separation of all organelles or separation of one or two organelle fractions with a higher purity). Parameters like the sucrose concentration, the volume of the sucrose steps, the number of sucrose steps, and the centrifugation speed as well as time have to be optimized.

20. A convenient time pattern can be achieved by starting a reaction in the cuvette every 15–20 s. All additions of solutions and the determination of the OD at 820 nm follow this time regime.
21. Depending on the protein concentration the reaction time can be extended up to 3 h. It should be noticed that starting 60 cuvettes in a 15-s time pattern takes 15 min and therefore is the minimum incubation time.

Acknowledgements

The work in G.O.s lab is partially financed by the Austrian Science Fund (P21298-B03) and the University Priority Programme “Biosciences and Health.” The technical assistance of Aniela Sommer with the *Arabidopsis* pollen germination is appreciated.

References

1. Perl H, Schulze WX, Obermeyer G (2009) The pollen organelle membrane proteome reveals highly spatial-temporal dynamics during germination and tube growth of lily pollen. *J Proteome Res* 8:5142–5152
2. Perl H, Rittmann S, Schulze WX et al (2011) Identification of lily pollen 14-3-3 isoforms and their subcellular and time-dependent expression profile. *Biol Chem* 392:249–262
3. Boavida LC, McCormick S (2007) Temperature as a determinant factor for increased and reproducible *in vitro* pollen germination in *Arabidopsis thaliana*. *Plant J* 52:570–582
4. Li H, Lin Y, Heath R et al (1999) Control of pollen tube tip growth by a Rop GTPase-dependent pathway that leads to tip-localised calcium influx. *Plant Cell* 11:1731–1742
5. Schreiber DN, Dresselhaus T (2003) *In vitro* pollen germination and transient transformation of *Zea mays* and other plant species. *Plant Mol Biol Rep* 21:31–41
6. Daher FB, Chebli Y, Geitmann A (2009) Optimization of conditions for germination of cold-stored *Arabidopsis thaliana* pollen. *Plant Cell Rep* 28:347–357
7. Michard E, Dias P, Feijo JA (2008) Tobacco pollen tubes as cellular models for ion dynamics: improved spatial and temporal resolution of extracellular flux and free cytosolic concentration of calcium and protons using pHluorin and YC3.1 CaMeleon. *Sex Plant Reprod* 21:169–181
8. Hoidn C, Puchner E, Perl H et al (2005) Non-diffusional release of allergens from pollen grains of *Artemisia vulgaris* and *Lilium longiflorum* depends mainly on the type of the allergen. *Int Arch Allergy Immunol* 137: 27–36
9. Perl H, Gehwolf R, Obermeyer G (2005) The distribution of membrane-bound 14-3-3 proteins in organelle-enriched fractions of germinating lily pollen. *Plant Biol* 7: 140–147
10. Quail PH (1979) Plant cell fractionation. *Annu Rev Plant Physiol* 30:425–484
11. Widell S, Larsson C (1990) A critical evaluation of markers used in plasma membrane purification. In: Larsson C, Moller IM (eds) *The plant plasma membrane*. Springer, Heidelberg, pp 16–43
12. Briskin DP, Leonard R, Hodges TK (1987) Isolation of the plasma membrane: membrane markers and general principles. *Methods Enzymol* 148:542–558
13. Fiske CH, Subbarow Y (1925) The colorimetric determination of phosphorous. *J Biol Chem* 66:375–400
14. Ames BN (1962) Assays of inorganic phosphate, total phosphate and phosphatases. *Methods Enzymol* 8:115–118
15. Lanzetta PA, Alvarez LJ, Reinach PS et al (1979) An improved assay for nanomole amounts of inorganic phosphate. *Anal Biochem* 100:95–97

Part V

Subcellular Proteomics

Chapter 31

Isolation of the Plant Cytosolic Fraction for Proteomic Analysis

Gonzalo M. Estavillo, Yves Verhertbruggen, Henrik V. Scheller, Barry J. Pogson, Joshua L. Heazlewood, and Jun Ito

Abstract

The cytosol is the fluid portion of the cell that is not partitioned by membranes. It contains a highly diverse collection of substances and is central to many essential cellular processes ranging from signal transduction, metabolite production and transport, protein biosynthesis and degradation to stress response and defense. Despite its importance, only a few proteomic studies have been performed on the plant cytosol. This is largely due to difficulties in isolating relatively pure samples from plant material free of disrupted organelle material. In this chapter we outline methods for isolating the cytosolic fraction from *Arabidopsis* cell cultures and seedlings and provide guidance on assessing purity for analysis by mass spectrometry.

Key words Cytosol, Isolation, Seedling, Cell culture, *Arabidopsis*, Proteomics

1 Introduction

The plant cytosol is a dynamic and complex intracellular fluid. It is the site for numerous biological reactions and often serves as a channel for metabolic interactions and communication between membrane-bound organelles. Its aqueous environment is highly crowded with dissolved ionic solutes, small molecule metabolites, and macromolecules. Essential processes known to occur in the plant cytosol include glycolysis [1], part of the pentose phosphate pathway [2], protein biosynthesis and degradation [3, 4], signal transduction [5], primary and secondary metabolite biosynthesis and transportation [6], stress response signaling [7], and accumulation of enzymes for defense and detoxification [8]. It has been estimated that over 50 % of eukaryotic proteins are present in the cytosol [9].

A large-scale proteomic analysis of cytosolic fractions from cell suspensions of the model plant, *Arabidopsis thaliana* produced a set of 1,071 proteins [10]. Functional classification of this set confirmed many of these proteins have well known cytosol-related

roles. This includes the 80S ribosome, tRNA catabolic enzymes, the ubiquitin-proteasome pathway, glycolysis, and associated sugar metabolism pathways, phenylpropanoid biosynthesis, vitamin and nucleotide metabolism, signaling and stress-responsive molecules, and NDP-sugar biosynthesis. In addition, the identification of hundreds of proteins with unknown functions can be used to further build on our understanding of how metabolism in the cytosol influences plant function.

Cell fractionation methods in plants have been used for decades to isolate different cell organelles. Methods for the isolation of intact chloroplasts [11–14], mitochondria [15, 16] and nuclei [17, 18] have been described. Most of these methods rely on the mechanical disruption of plants by mincing tissue with blades, blending, or with a mortar and pestle. However, multiple studies have established that higher yields of organelles can be obtained when starting with protoplasts [11, 19, 20].

Plant cell protoplasts are produced by using fungal enzymes (e.g., cellulases and pectinases) to degrade the cell wall. Many procedures have long been described for the production of protoplasts from a variety of plant species and organ types [21–24]. The degradation of cell walls is performed in the presence of an osmotic, usually a sugar such as mannitol, sucrose, or sorbitol. This prevents protoplast disruption due to differences in osmolality between the cell interior and the digestion medium. Protoplasts have been extensively used to characterize many physiological processes in plant cells, such as photosynthesis [24], guard cell signaling [25], and protein localization [26] to name a few, and more recently for the isolation of subcellular fractions [10, 19, 27].

Separation of the cytosolic fraction from plant samples requires a method that can break apart or remove plant cell walls, while also incurring minimal damage to organelles to prevent release of organelle proteins into the cytosolic fraction. One such method used to isolate cytosolic fractions from *Arabidopsis* cell cultures was based on earlier fractionation studies undertaken on subcellular organelles [10]. These organelle studies had demonstrated that superior yields of intact organelles could be obtained from protoplasts and their disruption by gentle pressure with a Potter-Elvehjem tissue grinder resulted in maintenance of organelle integrity [19, 20]. This gentle homogenization method is followed by removal of organelles using differential centrifugation. Extrapolating such approaches to plant tissues, as opposed to cell cultures, requires several extra processing steps to ensure success. When attempting the isolation of the cytosolic fraction from whole *Arabidopsis* seedlings, it is critical to start with a pure sample of intact protoplasts, as many break during digestion. Isolation of intact protoplasts from plant tissues can be achieved by exploiting their differential flotation density in sugar gradients [28]. However, when not using a heterotrophic plant cell culture, an additional difficulty to isolate the cytosolic fraction from

plant tissues is to minimize the cross-contamination of chloroplasts with the intact protoplasts. Density gradients have been demonstrated to prevent such contamination [26].

Here, we present two techniques for the purification of cytosolic enriched fractions from *Arabidopsis*. The methods are based on published procedures, which have been adapted to isolate cytosolic enriched fractions from either *Arabidopsis* cell cultures [19, 20] or seedlings [23, 28]. These procedures have been successfully used for large scale proteomic characterization of the cytosolic fraction from cultured cells [10] or for the study of protein localization between cytosol and chloroplasts of *Arabidopsis* seedlings [26]. The isolated fractions can be cleared of other cellular contaminants via differential centrifugation and sugar gradients. In addition to the cytosolic fraction of seedlings, chloroplasts can also be isolated using a discontinuous Percoll™ gradient after protoplastation [12]. The final fractions can be analyzed for organelle contamination by immunoblotting or mass spectrometry-based selected reaction monitoring against known subcellular protein markers. We also discuss how to optimize the procedure to varying conditions that affect the purity of the fractions (e.g., the age and the quality of the starting material or the type of homogenizer and frequency of strokes).

2 Materials

Prepare all solutions using ultrapure water (prepared by purifying deionized water to attain a sensitivity of $18\text{ M}\Omega\text{ cm}$ at $25\text{ }^{\circ}\text{C}$) and analytical grade reagents. Prepare reagents at room temperature. Perform all centrifugation steps at $4\text{ }^{\circ}\text{C}$.

2.1 Purification of the Cytosolic Fraction from Plant Cell Cultures

1. Plant material: *Arabidopsis* cell culture.
2. Enzyme Buffer: 0.4 M mannitol, 3.5 mM MES-NaOH buffer, pH 5.7, 0.4 % (w/v) cellulase RS (Yakult Pharmaceutical, Tokyo, Japan) and 0.05 % (w/v) pectolyase Y-23 (Yakult Pharmaceutical, Tokyo, Japan). Prepare immediately before use (see Note 1).
3. Wash Buffer: 0.4 M mannitol, 3.5 mM MES-NaOH, pH 5.7.
4. Homogenization Buffer: 0.4 M sucrose (osmotic), 3 mM EDTA, 50 mM Tris-HCl buffer and 2 mM dithiothreitol [DTT]. Add reducing agent DTT just prior to homogenization (see Note 2).
5. Glass-Teflon Potter-Elvehjem Tissue Grinders (30–50 mL capacity), cooled on ice. For gentle disruption of protoplasts, the space between the Teflon pestle and the tube should be approximately 100 μm .
6. Miraclot (Merck KGaA, Germany).

7. Preparative centrifuge with rotors capable of processing 2×500 mL samples at $800 \times \text{g}$ and 4×50 mL samples at $10,000 \times \text{g}$.
8. Ultracentrifuge with fixed angle rotor capable of processing 4×12 mL samples at $100,000 \times \text{g}$.
9. 5 kDa Ultrafree centrifugal filter devices (EMD Millipore, MA, USA) for protein concentration.
10. Tandem mass spectrometer (MS/MS) coupled with online liquid chromatography (LC) (nanoflow or capillary flow rates) capable of data dependent acquisitions.
11. Tandem mass spectrometer (MS/MS) coupled with online liquid chromatography (LC) capable of undertaking selected reaction monitoring techniques (e.g., triple quadrupole mass spectrometer).
12. Collection of plant organelle marker antibodies (Agrisera AB, Vännäs, Sweden).

2.2 Purification of the Cytosolic Fraction from Plant Seedlings

1. Plant material: 14-day-old *Arabidopsis* grown in 2.2 g/L Murashige and Skoog (MS) medium, 1 % (w/v) sucrose, and 0.7 % (w/v) agar at $100 \mu\text{mol}$ photons/ m^2 s for a 12 h photo-period at 24°C .
2. Digestion Medium: 1.5 % (w/v) cellulase (Yakult Pharmaceutical, Tokyo, Japan), 0.4 % (w/v) macerozyme (Yakult Pharmaceutical, Tokyo, Japan), 0.5 M sucrose, 20 mM KCl, 10 mM CaCl₂, and 20 mM MES-KOH, pH 5.7. Dissolve enzymes in the solution until it becomes clear light brown and filter through 0.45 μm syringe filter device (see Note 3).
3. Wash Buffer: 0.5 M sucrose, 20 mM KCl, 10 mM CaCl₂, and 20 mM MES-KOH, pH 5.7.
4. Flotation Medium I (FMI): 0.5 M Sucrose, 1 mM MgCl₂, and 5 mM HEPES, pH 7.0 (see ref. 28).
5. Flotation Medium II (FMII): 0.4 M sucrose, 0.1 M sorbitol, 1 mM MgCl₂, and 5 mM HEPES, pH 7.0 [28].
6. Flotation medium III (FMIII): 0.5 M sorbitol, 1 mM MgCl₂, and 5 mM HEPES, pH 7.0 [28].
7. Chloroplast Buffer: 50 mM HEPES-KOH, pH 8.0, 5 mM EDTA, 5 mM EGTA, 330 mM sorbitol, 5 mM cysteine, and 5 mM ascorbic acid.
8. Percoll Solution: 95 % (w/v) PercollTM, 3 % (w/v) PEG 6000, 1 % (w/v) FicollTM [12] (see Note 4).
9. Gradient Mixture: 25 mM HEPES-NaOH (pH 8.0), 10 mM EDTA, 5 % (w/v) sorbitol [12].
10. Glass-Teflon Potter-Elvehjem Tissue Grinders (10–20 mL capacity), cooled on ice.

11. Miraclot (Merck KGaA, Germany).
12. Pasteur pipettes (glass) with long neck.
13. 5 kDa Ultrafree centrifugal filter devices (EMD Millipore, MA, USA) for protein concentration.
14. Preparative centrifuge with swing out rotors capable of processing 6×30 mL samples at up to $800 \times g$ and 4×150 mL samples at up to $3,000 \times g$.
15. Ultracentrifuge with fixed angle rotor capable of processing 4×10 mL samples at $100,000 \times g$.
16. Vacuum pump, such as Vacuubrand model MZ 2C or similar (Wertheim, Germany), capable of vacuum 9.0 mbar and maximal pumping speed of $1.7/2.0 \text{ m}^3/\text{h}$.
17. Corex centrifuge tubes, 15 and 30 mL capacity.
18. Collection of plant organelle marker antibodies (Agrisera AB, Vännäs, Sweden).

3 Methods

3.1 Purification of the Cytosolic Fraction from Plant Cell Cultures

Carry out all procedures at room temperature unless otherwise specified.

1. Protoplast production
 - (a) Collect cells and remove cell culture medium by filtering through two pieces of Miraclot.
 - (b) Use a ratio of 1:5 plant cells fresh weight (FW) to Enzyme Buffer (i.e., 10 g cells to 50 mL buffer) to generate protoplasts from *Arabidopsis* cells culture (see Note 5).
 - (c) Resuspend filtered cells in Enzyme Buffer and incubate in a wide-base conical flask for 3 h with gentle orbital rotation (~85 rpm) in the dark (see Note 6).
 - (d) Harvest protoplasts by centrifuging at $800 \times g$ for 10 min at 4°C . Gently resuspend the pellet (protoplasts) with Wash Buffer using a 1:5 ratio of cells to buffer. Centrifuge at $800 \times g$ for 10 min at 4°C and repeat for a total of two washes with Wash Buffer. Proceed with homogenization (see Note 7).
2. Protoplast homogenization
 - (a) Resuspend protoplasts in Homogenization Buffer and disrupt them with about five strokes of a Potter-Elvehjem homogenizer at 4°C . Use a ratio of 1:2 cells FW to Homogenization Buffer. This step is preferably done in a cold room (see Note 8).

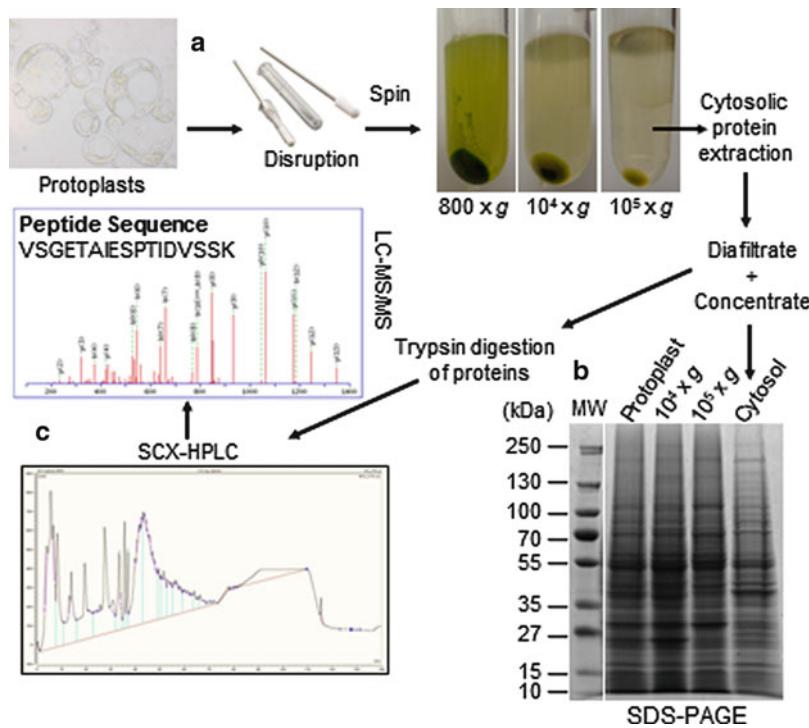


Fig. 1 Isolation of the cytosolic fraction from *Arabidopsis* cell cultures for proteomic analysis. (a) Protoplasts from 7-day-old cell cultures are gently disrupted with a Potter-Elvehjem homogenizer. Broken protoplasts are subjected to three successive centrifugation steps of: 800 $\times g$ for 15 min, 10,000 $\times g$ for 15 min, and 100,000 $\times g$ for 1 h. These remove unbroken cells, large cellular debris, organelles, and small vesicular components from the cytosolic fraction (Cytosol). (b) After diafiltration and concentration of cytosolic proteins using 5 kDa cut-off spin columns, the sample is separated by SDS-PAGE. Note the banding pattern of the cytosolic fraction differs to that of intact protoplasts, the 10,000 $\times g$ and 100,000 $\times g$ pellet fractions. These gels are also subjected to immunological analysis to evaluate organelle contamination in the cytosolic fraction (see Fig. 2). (c) Isolated cytosolic proteins are digested with trypsin, desalting and fractionated by strong cation exchange (SCX) liquid chromatography to create multiple fractions for peptide identification by LC-MS/MS analysis (offline MudPIT). Peptides are detected by a UV detector set at a wavelength of 214 nm and collected in 15–16 fractions, with most peptides eluting from the SCX column within the first half of the 140 min run

3. Isolation of the cytosolic fraction

- Centrifuge the homogenate at 800 $\times g$ for 15 min at 4 °C and discard the pellet containing unbroken cells, large cellular debris, and nuclei (Fig. 1).
- Centrifuge the supernatant at 10,000 $\times g$ for 15 min at 4 °C and discard the pellet, which is enriched in mitochondria and plastids (Fig. 1).
- Centrifuge the supernatant at 100,000 $\times g$ for 1 h at 4 °C and collect the supernatant (cytosolic fraction). The discarded pellet contains secretory pathway components, plasma membrane, Golgi, and ER (see Note 9 and Fig. 1).

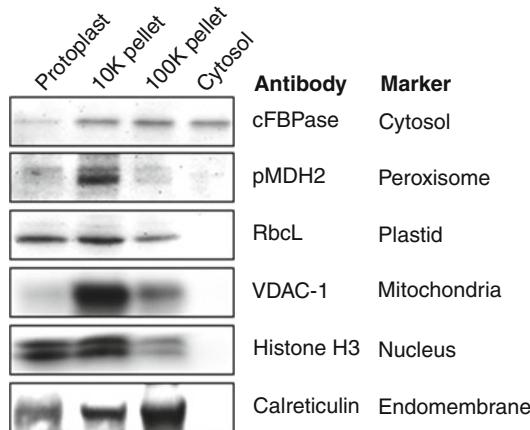


Fig. 2 Immunological analysis of the cytosolic enrichment method using 5 μ g protein from each fraction. Western blotting of subcellular protein markers from *Arabidopsis* cell culture protoplasts (protoplast), 10,000 \times g crude mixed organelle pellet (10K pellet), 100,000 \times g crude mixed organelle pellet (100K pellet) and cytosolic fraction (cytosol) after SDS-PAGE. Polyclonal antibodies rose against *Arabidopsis* cFBPase (cytosol), pMDH2 (peroxisome), RbcL (plastid), VDAC-1 (mitochondria), histone H3 (nucleus), and calreticulin (endomembrane system) are applied to the samples to determine their enrichment of subcellular compartments. The lack of organelle protein contaminants in the cytosolic fraction is clearly seen

Diafiltrate the cytosolic fraction with 100 % H_2O and concentrate sample using 5 kDa Ultrafree centrifugal filter device (see Note 10).

- Perform total protein quantification with Bradford or Lowry assay. Immediately store protein samples in 2 mg aliquots at -80°C .

4. Purity assessment

- Purity of the cytosolic fraction can be assessed by immunoblotting with commercially available antibodies against plant organelle protein markers (e.g., mitochondria, plastids, nuclei, the ER, plasma membranes, peroxisomes, and cytosol) (Fig. 2).
- In addition, mass spectrometry-based selected reaction monitoring (SRM) can be used to quantify organelle protein contamination in the cytosolic fraction as a supplement to immunoblotting. SRM is a highly specific and sensitive tandem mass spectrometry technique for quantifying an individual precursor ion and its fragment ion (transition) in a complex biological mixture. A selection of SRM transitions have been established for profiling the cytosolic fraction after enrichment [10].

3.2 Purification of the Cytosolic Fraction from *Arabidopsis* Seedlings

All steps are performed at 4 °C except the infiltration and incubation in digestion medium.

1. Protoplast isolation

- (a) Harvest approximately 20 g of seedlings in the morning before the light period starts (see **Note 11**).
- (b) In a beaker, add to the seedlings 30 mL Digestion Medium and vacuum infiltrate the plant material for 30 min (see **Note 12**). Incubate for an additional 2.5 h in the *dark*. As explained in **Note 5**, protoplast generation may vary with the plant origin and its development. So, we optimized this step using a ratio of 1:1.5 or 1:2 seedlings fresh weight (FW) to Digestion Medium.
- (c) Filter the mixture through two layers of Miracloth pre-soaked with Wash Buffer. Collect the filtrate if Digestion Medium is to be reused (see **Note 13**).
- (d) In a beaker, release the protoplast from undigested tissue by resuspension in 60–90 mL Wash Buffer. The volume of Wash Buffer is two to three times higher than the volume of Digestion Medium used in **step 1b**. Release of protoplast is achieved by swirling the mixture on ice for several minutes. Collect the dark-green eluent in a glass beaker on ice. All subsequent steps up to **step 1i** must be performed at 4 °C.
- (e) Split extract into 30 mL Corex tubes, centrifuge for 5 min at $500 \times g$ in a fixed angle or swing out rotor and discard the supernatant.
- (f) Repeat Wash **steps 1d** and **1e** (see **Note 14**).
- (g) Resuspend pellets in 15 mL Flotation Medium I (FMI). Then, overlay 7.5 mL of FMII followed by 3 mL FMIII (see **Note 15**).
- (h) Centrifuge the gradient at $250 \times g$ for 5 min at 4 °C in a *swing-out* rotor (*brake off*).
- (i) Recover the intact protoplasts from the FMIII/FMII interface (Fig. 3) and continue with homogenization. Keep the pellet for chloroplast purification (see **step 4**).

2. Protoplast homogenization

- (a) Disrupt the intact protoplasts using three to five strokes of a prechilled Potter-Elvehjem homogenizer at 4 °C, preferably in a cold room (see **Note 16**).
- (b) Transfer the solution containing the broken protoplasts to a 15 mL Corex tube. Using a long-neck glass Pasteur pipette, insert a 500 µL cushion of 85 % (v/v) Percoll at the bottom of the tube followed by a 1 mL layer of FMIII on top of the protoplast homogenate (see **Note 17**).

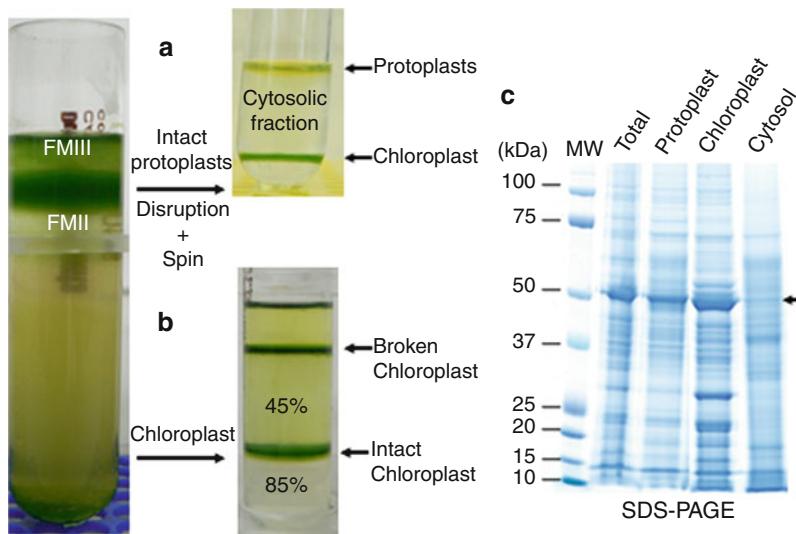


Fig. 3 Method for simultaneous isolation of cytosolic and chloroplast fractions from *Arabidopsis* seedlings. (a) Protoplasts from 16-day-old seedlings are produced by digestion of cell walls using a buffer containing 1.5 % (w/v) cellulase, 0.4 % (w/v) macerozyme, 0.5 M sucrose. Intact protoplasts are then isolated on a density gradient between FMII (0.4 M sucrose, 0.1 M sorbitol) and FMIII (0.5 M sorbitol). Intact protoplasts are disrupted with six strokes of a Potter-Elvehjem homogenizer on ice. Broken protoplasts are then centrifuged in the presence of a bottom cushion of 85 % (v/v) Percoll and top layer of FMIII to clear them from contaminating chloroplasts and unbroken protoplasts. (b) The chloroplast rich fraction is resuspended in buffer containing 0.33 M sorbitol, placed on top of a 45/85 % (v/v) Percoll gradient and centrifuged. Intact chloroplasts are recovered in the 45/85 % (v/v) Percoll interface. (c) Total protein from seedlings (Total), intact protoplasts, chloroplasts, and cytosol is extracted in the presence of 10 % (w/v) TCA in acetone, separated by SDS-PAGE (4–12 %) and visualized after staining. Representative protein profiles are shown for each isolated fraction, with enrichment of RuBisCO large subunit (arrow) in chloroplast samples and its depletion in cytosolic fraction

3. Isolation of the cytosolic fraction

- Centrifuge at $2,000 \times g$ for 10 min at 4 °C in a swing-out rotor (*brake off*). Unbroken protoplasts are in the FMIII/sample interface, while chloroplasts should sit on top of the 85 % (v/v) Percoll™ cushion (Fig. 3).
- Very carefully, recover the middle phase containing the cytosolic fraction using a clean long-neck glass Pasteur pipette, and centrifuge this fraction at $100,000 \times g$ for 30 min at 4 °C. This will remove any contaminating organelles and membranes.
- Concentrate the supernatant (cytosolic fraction) using a 5 kDa Ultrafree centrifugal filter device (Millipore) (see Note 10).

4. Isolation of chloroplasts

- Using a long-neck Pasteur pipette, recover the material on top of the 85 % (v/v) Percoll cushion from step 3a and gently resuspend in 10 mL of Chloroplast Buffer.

- (b) Centrifuge at $250 \times g$ for 5 min at 4 °C in a swing-out rotor.
- (c) Remove as much of the supernatant as possible and gently resuspend the pellet in a small volume (0.5–1 mL) of the same buffer.
- (d) Using a cut-off pipette tip, carefully load the resuspended chloroplast mixture onto a 45/85 % (v/v) Percoll™ gradient (see Note 18), and centrifuge at $3,000 \times g$ 15 min at 4 °C in a swing-out rotor (*brake off*).
- (e) With a long-neck Pasteur pipette, collect intact chloroplasts from the 45/85 % (v/v) interface (see Note 19) (Fig. 3) and transfer them in a new tube.
- (f) Prior to centrifugation at $800 \times g$ for 15 min in a swing-out rotor at 4 °C, wash with 10 volumes of Chloroplast Buffer to dilute the Percoll™.
- (g) Finally, resuspend the intact, clean chloroplasts in 500 mL of Chloroplast Buffer, or 10 % (w/v) TCA in acetone if a total protein extraction is required.

5. Purity Assessment

- (a) Assess purity of both cytosolic and chloroplast fractions by immunoblotting with commercially available antibodies against plant organelle protein markers (e.g., mitochondria, plastids, nuclei, the ER, plasma membranes, peroxisomes, and cytosol (Agrisera AB, Vännäs, Sweden)) (Fig. 2).

4 Notes

1. It is critical to freshly dissolve the plant cell wall degrading enzymes (cellulase and pectolyase) in enzyme buffer immediately prior addition to plant cells. The enzyme buffer solution contains: mannitol (acting as osmoticum to maintain isotonic conditions between the medium and plant cells), MES buffer (to achieve optimal pH for *Arabidopsis* cell suspension) and cellulase and pectolyase (a mix of enzymes that degrade the cellulosic and pectic components of plant cell walls). Wash Buffer is required for subsequent wash steps to remove residual cellulase and pectolyase from the protoplasts.
2. The osmoticum maintains the integrity of organelle membranes. The buffer counteracts acidification from ruptured vacuoles, EDTA chelates divalent metal ions required by phospholipases and various proteases and the reductant DTT prevents damage from oxidizing agents.
3. It is preferable to freshly prepare the Digestion Buffer. However, Digestion Buffer can be used if stored at -20 °C for up to 6 months.

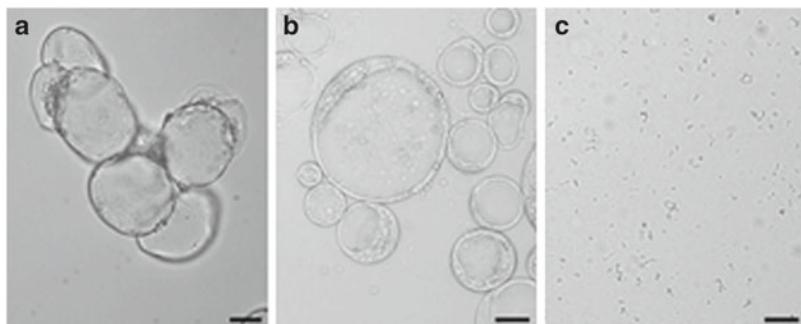


Fig. 4 Protoplasts from *Arabidopsis thaliana* cell culture under light microscope. (a) Seven-day-old cells and (b) intact protoplasts produced by degradation of the cell wall using a buffer containing 0.4 % (w/v) cellulase, 0.05 % (w/v) pectolyase, 0.4 M mannitol. (c) Released cellular contents of disrupting protoplasts with five strokes of a Potter-Elvehjem homogenizer on ice. Scale bars: 20 μ m

4. Note that the original recipe by [12] includes 1 % (w/v) bovine serum albumin (BSA). However, this may interfere with protein quantification if not adequately removed by washing. An alternative is to include BSA in this buffer and wash with Chloroplast Buffer without BSA.
5. The ratio of plant cells fresh weight (FW) to enzyme buffer solution for protoplast generation may require some adjustment as it depends on the plant origin and its growth. The optimal ratio can be determined by the amount of released cellular contents during the homogenization step (Fig. 4). For our *Arabidopsis* cell cultures, the optimal ratio range is in between 1:4 and 1:5 plant cells (FW) to enzyme buffer solution. Ratios below this range are generally ineffective at producing protoplasts.
6. For the successful production of protoplasts from cell cultures, two critical factors are the use of a wide-base conical flask and the speed of the orbital shaker during digestion. The depth of Enzyme Buffer solution containing *Arabidopsis* cells in the conical flask should be relatively shallow (i.e., no more than ~3 cm) and the shaker should be set at a speed in where cells are maintained in suspension. Setting the shaker speed too slow will allow cells to settle to the bottom of the flask during incubation and likely result in inefficient enzymatic digestion of cell walls. Conversely, setting the shaker speed too high may cause protoplasts to rupture from higher energy collisions, by either hitting each other or the flask.
7. After protoplasts are removed from the Enzyme Buffer and washed free of cellulase and pectolyase, they should be homogenized as soon as possible.

8. It is important for effective disruption of protoplasts, that a vacuum is formed with each stroke by pulling the Teflon pestle to create an air bubble in the solution under the pestle. Generally, homogenizers with a space of approximately 50–100 μm between the Teflon pestle and the tube are the most effective for disrupting protoplasts from *Arabidopsis* cells, with minimal disruption of organelles. To assess the effectiveness of protoplast disruption with a given Potter-Elvehjem homogenizer, take aliquots of protoplast samples before and after homogenization and observe them using a light microscope. Successful disruption of protoplasts in the post-homogenized sample will be apparent by the presence of released cellular contents (Fig. 4). If this is not the case, then adjustments to factors (see Notes 1 and 2) may be necessary.
9. Isolation of the cytosolic fraction should be performed as quickly as possible while maintaining a temperature of 4 °C and preventing prolonged periods of incubation between centrifugation steps. The time between homogenization and isolation of the cytosolic fraction significantly affects sample integrity and yield.
10. The removal of Homogenization Buffer from the isolated cytosolic fraction is critical as its constituents are not compatible with protein precipitation or digestion, and downstream LC-MS/MS analysis. We have investigated the following methods to remove the homogenization buffer from the cytosolic protein fraction: dialysis, size exclusion FPLC, C4 hydrophobic interaction (C4) HPLC and diafiltration 5 kDa cut-off spin column. We found that the diafiltration spin column is the most effective method. It is easy to use, has a high protein retention rate (~95 %) and is compatible with LC-MS/MS. To remove most of the homogenization buffer, the samples are concentrated, diluted with 100 % H₂O using a maximum spin column volume and reconcentrated. This is repeated a total of three times. Samples can then be used for analysis by immunoblotting or LC-MS/MS.
11. Preferably, place the seedlings in the dark for 24 h. This promotes degradation of starch grana, which can break chloroplasts during centrifugation.
12. The seedlings do not need to be chopped. Vacuum infiltration is very efficient to promote organ imbibition with the Digestion Medium. Proper infiltration will result in tissue turning very dark green.
13. The Digestion Medium can be reused after removing cellular debris by centrifugation (10 min at 3,000 $\times g$) and storing at -20 °C.

14. These washing steps are critical to eliminate contaminant proteins from the Digestion Medium such as the degrading enzymes. Keep an aliquot of Digestion Medium and analyze its content by SDS-PAGE to evaluate the level of contamination in the cytosolic fraction.
15. Overlaying the extract with floating mediums II and III, requires extreme care as the interfaces between the two solutions can be easily disrupted. To avoid this, we recommend overlaying the two mediums at 4 °C by placing a tube at an inclined angle on ice and directly above it, the needle of an upright syringe attached to a stand making contact with the upper inside wall of the tube. Floating medium II is placed in the syringe body, which is then slowly transferred into the resting tube by gravity. Repeat this step with floating medium III. Note that this is important to achieve a sharp interface.
16. The intensity and number of strokes may need to be optimized. We found that three to six strokes results in a good quality fraction, while more than six may provoke disruption of organelles and result in cross-contamination.
17. The goal of this step is to remove chloroplasts and unbroken protoplasts. The 85 % Percoll cushion prevents chloroplast breakage during centrifugation, which would result in the contamination of the middle phase. The objective of the top layer (FMIII) is to maintain unbroken protoplasts on the top interface.
18. Percoll gradients must be freshly made. Add the Percoll Solution and Gradient Mixture in a separate tube to achieve the desired Percoll Solution final concentration.
19. The intact chloroplasts should sediment at the 45/85 % (v/v) Percoll interface. The layer between the Chloroplast Buffer and the 45 % (v/v) phase mostly consists of broken chloroplasts.

Acknowledgements

This work (J.I., J.L.H., H.V.S. and Y.V.) was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. B.J.P. and G.E. are supported by the ARC Centre of Excellence in Plant Energy Biology (CE0561495).

References

- Plaxton WC (1996) The organization and regulation of plant glycolysis. *Annu Rev Plant Physiol Plant Mol Biol* 47:185–214
- Schnarrenberger C, Flechner A, Martin W (1995) Enzymatic evidence for a complete oxidative pentose phosphate pathway in chloroplasts and an incomplete pathway in the cytosol of spinach leaves. *Plant Physiol* 108: 609–614
- Vierstra RD (2009) The ubiquitin-26S proteasome system at the nexus of plant biology. *Nat Rev Mol Cell Biol* 10:385–397
- Carroll AJ, Heazlewood JL, Ito J et al (2008) Analysis of the *Arabidopsis* cytosolic ribosome proteome provides detailed insights into its components and their post-translational modification. *Mol Cell Proteomics* 7:347–369
- Lecourieux D, Ranjeva R, Pugin A (2006) Calcium in plant defence-signalling pathways. *New Phytol* 171:249–269
- Lunn JE (2007) Compartmentation in plant metabolism. *J Exp Bot* 58:35–47
- Cazale AC, Clement M, Chiarenza S et al (2009) Altered expression of cytosolic/nuclear HSC70-1 molecular chaperone affects development and abiotic stress tolerance in *Arabidopsis thaliana*. *J Exp Bot* 60: 2653–2664
- Dixon DP, Hawkins T, Hussey PJ et al (2009) Enzyme activities and subcellular localization of members of the *Arabidopsis* glutathione transferase superfamily. *J Exp Bot* 60: 1207–1218
- Lahav M, Schoenfeld N, Epstein O et al (1982) A method for obtaining high recovery of purified subcellular fractions of rat liver homogenate. *Anal Biochem* 121:114–122
- Ito J, Batth TS, Petzold CJ et al (2010) Analysis of the *Arabidopsis* cytosolic proteome highlights subcellular partitioning of central plant metabolism. *J Proteome Res* 10: 1571–1582
- Fitzpatrick LM, Keegstra K (2001) A method for isolating a high yield of *Arabidopsis* chloroplasts capable of efficient import of precursor proteins. *Plant J* 27:59–65
- Aronsson H, Jarvis P (2002) A simple method for isolating import-competent *Arabidopsis* chloroplasts. *FEBS Lett* 529:215–220
- Schulz A, Knoetzel J, Scheller HV et al (2004) Uptake of a fluorescent dye as a swift and simple indicator of organelle intactness: import-competent chloroplasts from soil-grown *Arabidopsis*. *J Histochem Cytochem* 52: 701–704
- Seigneurin-Berny D, Salvi D, Joyard J et al (2008) Purification of intact chloroplasts from *Arabidopsis* and spinach leaves by isopycnic centrifugation, vol 40, *Current protocols in cell biology*. Wiley, New York
- Day D, Neuburger M, Douce R (1985) Biochemical characterization of chlorophyll-free mitochondria from pea leaves. *Funct Plant Biol* 12:219–228
- Rödiger A, Baudisch B, Bernd Klösgen R (2010) Simultaneous isolation of intact mitochondria and chloroplasts from a single pulping of plant tissue. *J Plant Physiol* 167: 620–624
- Calikowski TT, Meier I (2006) Isolation of nuclear proteins. In: Salinas J, Sanchez-Serrano JJ (eds) *Arabidopsis protocols*, 2nd edn. Humana Press Inc., Totowa, NJ, pp 393–402
- Folta KM, Kaufman LS (2006) Isolation of *Arabidopsis* nuclei and measurement of gene transcription rates using nuclear run-on assays. *Nat Protoc* 1:3094–3100
- Eubel H, Meyer EH, Taylor NL et al (2008) Novel proteins, putative membrane transporters, and an integrated metabolic network are revealed by quantitative proteomic analysis of *Arabidopsis* cell culture peroxisomes. *Plant Physiol* 148:1809–1829
- Meyer EH, Millar AH (2008) Isolation of mitochondria from plant cell culture. *Methods Mol Biol* 425:163–169
- Takebe I, Labib G, Melchers G (1971) Regeneration of whole plants from isolated mesophyll protoplasts of tobacco. *Naturwissenschaften* 58:318–320
- Kanai R, Edwards GE (1973) Separation of mesophyll protoplasts and bundle sheath cells from maize leaves for photosynthetic studies. *Plant Physiol* 51:1133–1137
- Yoo S-D, Cho Y-H, Sheen J (2007) *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat Protoc* 2:1565–1572
- Zhang Y, Su J, Duan S et al (2011) A highly efficient rice green tissue protoplast system for transient gene expression and studying light/chloroplast-related processes. *Plant Methods* 7:30
- Zhao Z, Zhang W, Stanley BA et al (2008) Functional proteomics of *Arabidopsis thaliana* guard cells uncovers new stomatal signaling pathways. *Plant Cell* 20:3210–3226
- Estavillo GM, Crisp PA, Pornsiriwong W et al (2011) Evidence for a SALL1-PAP chloroplast retrograde pathway that functions in drought and high light signaling in *Arabidopsis*. *Plant Cell* 23:3992–4012

27. Parsons HT, Christiansen K, Knierim B et al (2012) Isolation and proteomic characterization of the *Arabidopsis* Golgi defines functional and novel components involved in plant cell wall biosynthesis. *Plant Physiol* 159:12–26
28. Gardestrom P, Wigge B (1988) Influence of photorepiration on ATP/ADP ratios in the chloroplasts, mitochondria, and cytosol, studied by rapid fractionation of barley (*Hordeum vulgare*) protoplasts. *Plant Physiol* 88:69–76

Chapter 32

Quantitative Proteomic Analysis of Intact Plastids

Takeshi Shiraya, Kentaro Kaneko, and Toshiaki Mitsui

Abstract

Plastids are specialized cell organelles in plant cells that are differentiated into various forms including chloroplasts, chromoplasts, and amyloplasts, and fulfill important functions in maintaining the overall cell metabolism and sensing environmental factors such as sunlight. It is therefore important to grasp the mechanisms of differentiation and functional changes of plastids in order to enhance the understanding of vegetality. In this chapter, details of a method for the extraction of intact plastids that makes analysis possible while maintaining the plastid functions are provided; in addition, a quantitative shotgun method for analyzing the composition and changes in the content of proteins in plastids as a result of environmental impacts is described.

Key words Chloroplast, iTRAQ, Quantitative shotgun proteomics, Percoll density gradient centrifugation, Protein extraction

1 Introduction

Plastids, of which chloroplasts are a representative type, exhibit a variety of functions in plant cells. In addition to carbon, nitrogen, and sulfur assimilation, plastids perform the synthesis of important chemical compounds in the plant including storage materials such as carbohydrates and lipids. The decline and loss of plastid functions, which can be caused by various environmental stresses including dehydration, salinity, changes in temperature, and so on, lead to aging of cells and cell death. Thus, the normal expression of plastid functions is essential to maintain the living activity of cells.

The origin of the term *plastid* lies in the Greek word *plastikos*, which means “molded.” It is important to gain a systematic understanding of the “molded” mechanism of these variable organelles. In former times, plastid research was mainly conducted through research on photosynthesis in chloroplasts. However, studies have been actively carried out in recent years on various aspects of the forms of differentiation, such as control of the differentiation of

root etioplasts, the mechanism of mass transport in the amyloplasts of reserve tissue, and the features of chromoplasts observed in fruit and petals [1]. Proteomic technology is essential in these analyses, and there are many cases in which the use of intact plastids for proteomic analysis is indispensable.

Spinach, pea, and wheat leaves have been frequently used as experimental materials in photosynthetic research on isolated chloroplasts, because (1) these materials can be readily obtained, (2) the leaves are soft, and (3) there are few compounds to interfere with organelle isolation. With the rapid advances in genome analysis, however, the need for organelle isolation from model plants and crops, such as rice, maize, tomato, *Arabidopsis thaliana*, and *Lotus japonicus*, has greatly increased [2–9]. Although variations in the size and shape of chloroplasts are observed in algae, such variations are less pronounced in the case of higher plants, with typical dimensions of 5–10 μm in diameter and 3–4 μm in thickness [1]. When isolating intact chloroplasts from higher plants, it is necessary to consider the tissue organization and cell components specific to the vegetable species rather than the size and shape of the chloroplasts.

Here, a typical procedure for the isolation of intact plastids, mainly focusing on chloroplasts, as well as a method of troubleshooting are described. We also outline a well-established comprehensive shotgun proteomic analysis technique using the iTRAQ labeling method to characterize the characteristics of intact plastids.

2 Materials

2.1 Sample Materials

Fresh and soft tissue should be selected as the material. The mixing of starch and inorganic substances should be avoided at the step of plastid isolation (see Notes 1–3).

2.2 Plastid Isolation Buffer

- 1 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)-KOH, pH 7.5: Dissolve 23.8 g of HEPES in about 80 mL of water and adjust the pH with KOH (10N). Make up to 100 mL with water. Store at 4 °C (see Note 4).
- 1 M ethylenediaminetetraacetic acid (EDTA): Dissolve 37.2 g of EDTA 2Na·2H₂O in about 80 mL of water and adjust pH to 8 with KOH (10N). Make up to 100 mL with water. Store at 4 °C (see Note 5).
- 1 M MgCl₂: Dissolve 20.3 g of MgCl₂·6H₂O and make up to 100 mL with water. Store at 4 °C.
- 1 M MnCl₂: Dissolve 12.6 g of MnCl₂ and make up to 100 mL with water. Store at 4 °C.
- 1 M sodium ascorbate: Dissolve 19.8 g of sodium L-ascorbate and make up to 100 mL with water. Store at 4 °C.

6. 1 M DTT: Dissolve 1.54 g of DTT (DL-dithiothreitol) in 10 mL water. Dispense 50 μ L and store at -30°C .
7. Dissolve 5.47 g of sorbitol in about 80 mL (see Note 6), and mix with 5 mL of 1 M HEPES-KOH, pH 7.5, 0.5 mL of 1 M EDTA, 0.5 mL of 1 M MgCl_2 , 0.5 mL of 1 M MnCl_2 , and 5 mL of 1 M sodium ascorbate [10] or 50 μ L of 1 M DTT (see Notes 7 and 8). Make up to 100 mL with water. Stand on ice. The prepared buffer should be used immediately.

2.3 Plastid Dilution Buffer

Dissolve 0.2 g of PVP with 5 mL of 1 M HEPES-KOH, pH 7.5, 0.5 mL of 1 M EDTA, and 0.5 mL of 1 M MgCl_2 , and then, make up to 100 mL with water. Store at 4°C .

2.4 Percoll

1. 80 % Percoll: Mix 100 % Percoll with plastid isolation buffer in the ratio of 4:1 by volume. Store at 4°C .
2. 40 % Percoll: Mix 100 % Percoll with plastid isolation buffer and distilled water in the ratio of 2:1:2 by volume. Store at 4°C .

2.5 Protein Extraction Buffer

1. 1 M Tris-HCl, pH 8: Dissolve 60.6 g of Tris in 400 mL water and adjust the pH with HCl (6N). Make up to 500 mL with water. Store at 4°C .
2. 50 % Glycerol: Mix 50 g of glycerol with water, and make up to 100 mL with water.
3. 8 M urea: Dissolve 48.05 g of urea in 100 mL of water.
4. Dissolve 1.52 g of thiourea and 0.3 g of 3-[(cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS) in about 2 mL water and mix with 1 mL of 1 M Tris-HCl, pH 8, 2 mL of 50 % glycerol, 0.2 mL of Triton X-100, 0.2 mL of 1 M DTT, and 3.75 mL of 8 M urea (see Note 9). Make up to 10 mL with water. Stand on ice. The prepared buffer should be used immediately.

2.6 MS Analysis

1. iTRAQ Reagents kit (AB SCIEX, CA, USA).
2. iCAT Cation Exchange Buffer Pack (Applied Biosystems, CA, USA).
3. Cartridge holder (Applied Biosystems, CA, USA).
4. C-18 column (Sep-Pak C18 cartridges, Waters, Milford, MA, USA).
5. 0.4 % formic acid (FA): Mix 80 μ L of formic acid with LC/MS-grade water and make up to 20 mL in glass bottle.
6. 0.1 % FA: Mix 7.5 mL of 0.4 % FA and 22.5 mL of LC/MS-grade water in glass bottle.
7. 75 % acetonitrile (ACN)/0.1 % FA: Mix 7.5 mL of 0.4 % FA and 22.5 mL of ACN in glass bottle.

3 Methods

Although the extraction method outlined here is for chloroplasts, it can also be used for any other type of plastid. All of the procedures are performed at 4 °C.

3.1 Tissue Disruption

1. Wash 5–10 g of plant tissues (*see Note 10*) with distilled water and immediately put those on ice (*see Notes 11–13*).
2. Add 30 mL of precooled (4 °C) plastid isolation buffer to a stainless steel bath that is placed on ice. Cut the tissues into small sizes (about 3 mm) using a razor or a dissecting scissors washed with acetone or ethanol (*see Note 14*).
3. Wrap the tissues in two layers of Miracloth and crush them using a mortar and pestle.

3.2 Filtration

1. Filter the homogenate through two layers of Miracloth and collect the filtrate.
2. Homogenize the remaining tissue in the Miracloth again and collect the filtrate.
3. Pool the first and second filtrates together and filter through four layers of Miracloth.
4. Centrifuge the filtrate at $200 \times g$ for 4 min at 4 °C to remove the nuclei and cellular debris. Collect the supernatant.
5. Centrifuge the supernatant at $2,000\text{--}2,500 \times g$ for 4 min at 4 °C to precipitate the chloroplasts (*see Note 15*). Discard the supernatant.
6. Resuspend the pellet containing chloroplasts in 3 mL of plastid isolation buffer. Vortex gently (*see Note 16*).

3.3 Percoll Density Gradient Centrifugation and Intact Plastid Isolation

1. Layer the chloroplast suspension slowly onto a discontinuous Percoll density gradient solution containing 1 mL of 80 % and 3 mL of 40 % Percoll (*see Note 17*).
2. Centrifuge at $4,000 \times g$ for 10 min at 4 °C using a swing bucket rotor (P40ST, Hitachi).
3. After centrifugation, intact chloroplasts will be on the interface of the 40 and 80 % Percoll (Fig. 1). The band on the upper layer will contain broken chloroplasts and thylakoid membranes. Use a Pasteur pipette to collect the fraction that contains intact chloroplasts.
4. Add five times the volume of plastid isolation buffer to the collected fraction and wash gently.
5. Centrifuge at $4,000 \times g$ for 10 min at 4 °C. Discard the supernatant.

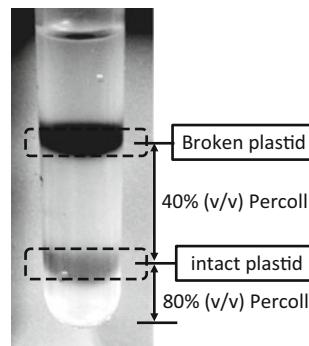


Fig. 1 Isolation of intact plastid from rice seedling

6. Suspend the pellet with an appropriate volume of plastid dilution buffer. Check the chloroplast integrity/quality under a light microscope (*see Note 18*). The suspension can be preserved at -80°C (*see Note 19*). For proteomics application, dissolve the pellet in five times the volume of protein extraction buffer.

3.4 Protein Extraction

The method described here is for quantitative proteomic analysis.

1. Suspend the samples in protein extraction buffer and sonicate.
2. Centrifuge the homogenate at $12,000 \times g$ for 5 min at 4°C and transfer the supernatant to a new tube.
3. Add 1/10 volume of 100 % trichloroacetic acid (TCA) to the supernatant and mix. Incubate the solution on ice for 15 min to precipitate proteins.
4. Centrifuge the homogenate at $10,000 \times g$ for 5 min at 4°C and discard the supernatant.
5. Add 1 mL of ice-cold acetone to the pellet and mix. Centrifuge the homogenate at $10,000 \times g$ for 10 min at 4°C . Discard the supernatant.
6. Repeat step 5 three to five times (*see Note 20*).
7. After drying the pellet, resuspend it in 20 μL of dissolution buffer and 1 μL of denaturant (2 % SDS). Both buffers are supplied in the iTRAQ Reagents kit (*see Note 21*).
8. Measure the protein concentration. 5–50 μg of protein is desirable for iTRAQ labeling.

3.5 iTRAQ Labeling

The peptide labeling with iTRAQ is in accordance with the manufacturer's protocol (<http://www.absciex.com/downloads/mass-spectrometry-literature>). The strategy of quantitative proteomic analysis using iTRAQ is summarized in Fig. 2.

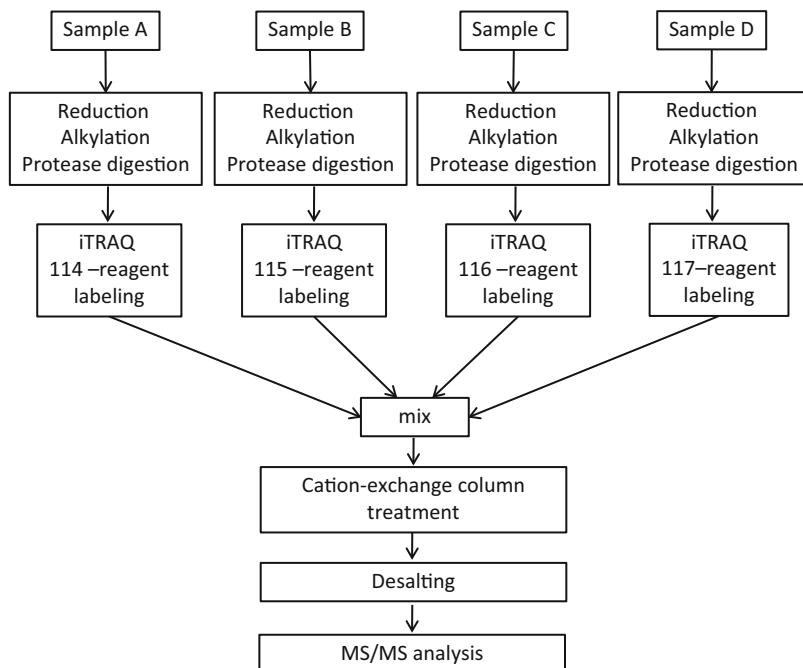


Fig. 2 Flowchart of the quantitative proteomic analysis with iTRAQ labeling

1. All reagents and buffers are contained in the iTRAQ Reagents kit.
2. Add 2 μ L of reducing reagent to each protein sample and incubate for 1 h at 37 °C (reduction) (see Notes 22 and 23).
3. After spin-down, add 1 μ L of cysteine blocking reagent and incubate for 10 min at room temperature (alkylation).
4. After spin-down, add 10 μ L of trypsin (0.5 μ g/ μ L) and incubate overnight at 37 °C (protease digestion).
5. Bring each vial of iTRAQ reagents to room temperature. Add 70 μ L of ethanol to each iTRAQ reagent and vortex for 1 min. After spin-down, transfer the contents of one iTRAQ reagent vial to one sample tube. Mix and incubate for 1 h at room temperature.
6. After adding 900 μ L of LC/MS-grade water, combine the iTRAQ-labeled peptide samples into one new tube.

3.6 Cation Exchange Column Treatment

To increase the efficiency of protein identification in shotgun proteomics analysis using an LC-MS/MS system, the mixture of iTRAQ-labeled peptides is separated into several fractions using iCAT Cation Exchange column [11].

1. Check the pH of the mixture of iTRAQ-labeled peptides using a pH test paper. The desired pH is between 2.5 and 3.5. If the pH exceeds 3.5, add a small amount of FA for adjustment.

2. All buffers for cation exchange chromatography are contained in the iCAT Cation Exchange Buffer Pack.
3. Assemble the holder and set the cation exchange column into it.
4. Wash the column by injecting 1 mL of cation exchange buffer—clean at the rate of 1 drop/s.
5. Inject 2 mL of cation exchange buffer—load at the rate of 1 drop/s.
6. Inject the iTRAQ-labeled peptide mixture into the column at the rate of 1 drop/s and collect the eluate.
7. Inject 1 mL of cation exchange buffer—load at the rate of 2 drops/s to wash the column, and collect the eluate.
8. Prepare 25, 50, 75, 100, 200, and 350 mM KCl solutions by diluting cation exchange buffer—elute (350 mM KCl) with load buffer (without KCl). Inject 0.5 mL of each solution into the column at the rate of 1 drop/s in increasing order of concentration. Collect each eluate in a new 1.5 mL tube (a total of six fractions).
9. Inject 1 mL of cation exchange buffer—clean (1 M KCl buffer) at the rate of 1 drop/s. Collect the eluate in a new 1.5 mL tube.
10. Evaporate the seven eluates (from **steps 6** and **7**) down to 50 μ L using a centrifugal vacuum concentrator.

3.7 Desalting Using a C-18 Column

1. Add 500 μ L of 0.4 % FA to each of the concentrated peptide mixtures.
2. Wash a C-18 column by injecting 2 mL of 75 % ACN/0.1 % FA at the rate of 2 drops/s.
3. Inject 1.5 mL of 0.1 % FA at the rate of 2 drops/s.
4. Inject the iTRAQ-labeled peptide mixture into the column at the rate of 1 drop/s. Collect the eluate and reinject. Collect the second eluate in a new tube.
5. Inject 2 mL of 0.1 % FA at the rate of 1 drop/s to wash the column.
6. Inject 0.6 mL of 75 % ACN/0.1 % FA at the rate of 1 drop/s and collect the eluate in a new 1.5 mL tube.
7. Repeat **steps 2–6** for each fraction.
8. Evaporate all of the samples to dryness using a centrifugal vacuum concentrator.
9. Redissolve each of the iTRAQ-labeled peptide mixtures with a small amount of loading buffer used in the LC-MS/MS system. Transfer the samples into desiccated vials.

3.8 MS Analysis

1. Load iTRAQ-labeled peptides on the column (75 μ m internal diameter, 5 cm; HiQ sil C-18 W-3, KYA TECH Corporation, Tokyo, Japan) using a DiNa-A auto injection system (KYA TECH).
2. Elute peptides with a linear gradient from 0 to 33 % B for 240 min, 33 to 100 % B for 10 min, and back to 0 % B for 15 min at a flow rate of 300 nL/min. Eluting buffers are 0.1 % FA and 5 % ACN (A) and 0.1 % FA and 80 % ACN (B). Peptides eluted from the column are introduced directly into an LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific Inc., Bremen, Germany) at a spray voltage of 4.5 kV.
3. The mass spectrometer is operated using Xcalibur software (Thermo Fisher Scientific). The mass range selected for MS scan sets to 350–1,600 m/z and the top three peaks subject to MS/MS analysis. Perform full MS scan in the Orbitrap, and MS/MS scans in the linear ion trap and Orbitrap: normalized collision energy, 35 eV for CID and 45 eV for HCD; resolution, 60,000.
4. Carry out protein identification by comparing the obtained spectra against data in *Oryza sativa* proteins downloaded from NCBI (<http://www.ncbi.nlm.nih.gov/>) and UniProt (<http://www.uniprot.org/>) database using Proteome Discoverer 1.3 software (Thermo Fisher Scientific). The MS data analysis parameters are as follows: enzyme, trypsin; missed cleavages, 2; MS tolerance, 10 ppm; MS/MS tolerance, 0.8 Da; static modification, carbamidomethylation; and dynamic modification, oxidation (Met). False discovery rates for peptide identification are less than 5 %.
5. Quantitation of iTRAQ reporter ions is performed using Proteome Discoverer 1.3 software.

3.9 A Case Study

A case study of identification and quantitation of chloroplast protein is shown in Fig. 3. The chloroplast proteins isolated from rice seedlings grown in hot (33 °C, 12-h light/28 °C, 12-h dark) and control (28 °C, 12-h light/23 °C, 12-h dark) conditions were subjected to the quantitative shotgun proteomic analysis with iTRAQ labeling. Identification and quantitation of superoxide dismutase [Cu-Zn] protein were carried out, indicating that the expression of chloroplast superoxide dismutase was up-regulated more than threefold under the hot condition.

4 Notes

1. It is important to determine whether the leaf material to be used is of the amylophyll (starchy leaf) or saccharophyll (sugary leaf) type [7, 12]. When handling the starchy leaves of

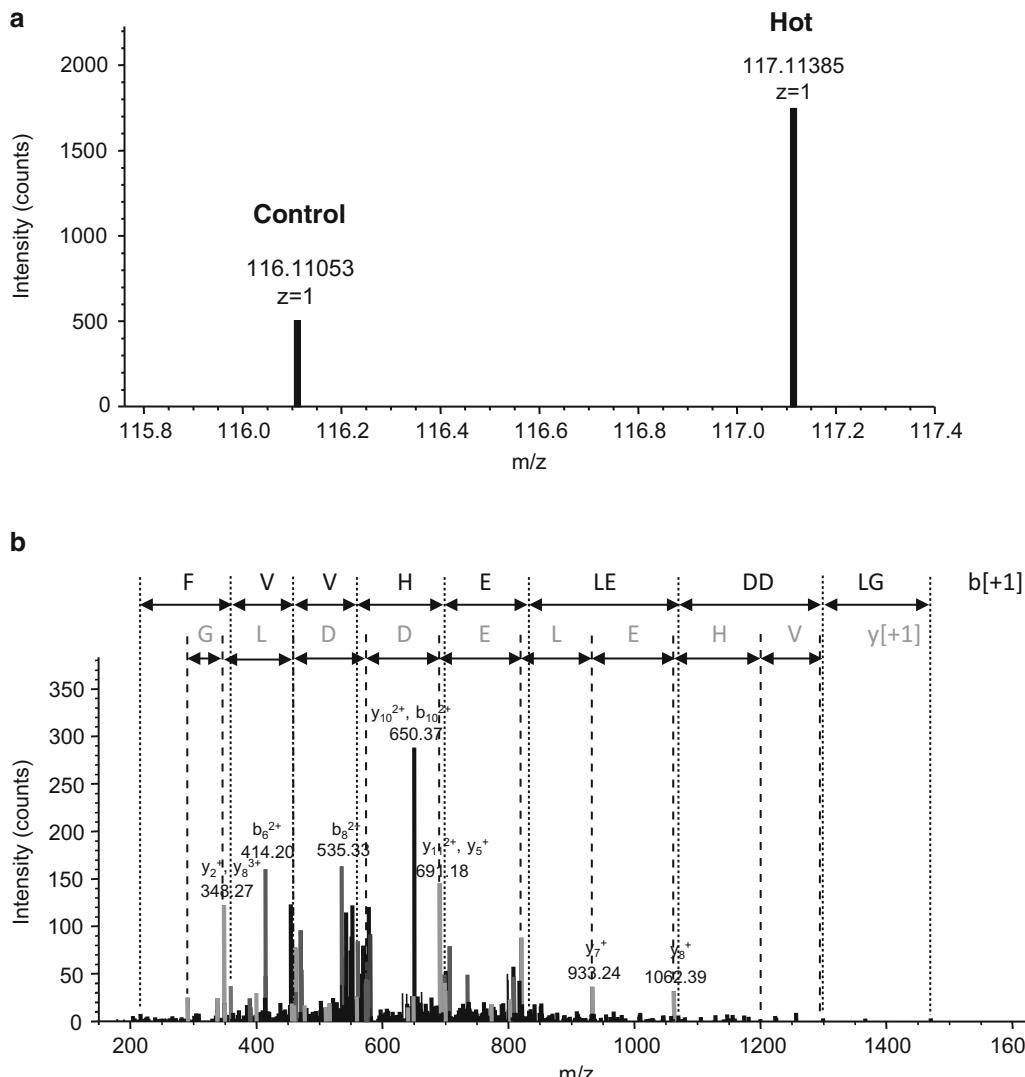


Fig. 3 A case study of superoxide dismutase [Cu-Zn] protein in chloroplasts of rice seedlings. The chloroplast proteins isolated from seedlings grown in hot (33 °C, 12-h light/28 °C, 12-h dark) and control (28 °C, 12-h light/23 °C, 12-h dark) conditions were labeled with iTRAQ-116 and -117 reagents, respectively. (a) Quantitation of iTRAQ-116 and -117 reporter ions for FVHELEDDLG peptide. (b) MS/MS spectrum of a representative superoxide dismutase [Cu-Zn] peptide (FVHELEDDLG) detected in the iTRAQ labeling experiments

dicotyledons such as spinach, tobacco, potato, and *Arabidopsis*, the starch granules, which have high density, often break through and escape from the enveloping membrane of the chloroplasts during centrifugation. Moreover, it is difficult to prevent the contamination of starch granules during chloroplast preparation. It is therefore necessary to consider measures to reduce the starch content. The starch content can be reduced by allowing the plant to grow for 2–3 days in darkness

or weak light (100–150 $\mu\text{mol photons}/\text{m}^2 \text{ s}$ in the case of *Arabidopsis*) [13]. In some cases, the starch content of the leaves may decrease during incubation in water overnight. Most of the monocotyledons including rice and wheat are of the saccharophyll type. Since the starch is accumulated in the leaf sheath (the culm in the case of rice), the mixing of starches can be suppressed if the leaf blades are used. Incubation in darkness for approximately 3 days before extraction is also effective in the preparation of rice chloroplasts [7].

2. The formation of crystallized inorganic substances (phytoliths) causes a decrease in the efficiency of extraction of intact chloroplasts. This is because the crystals physically destroy the enveloping membrane of the plastids. Such crystals, which are known to be formed in various plants, include silicic acid (rice, cactus, *Solanaceae*, *Osmundaceae*, pteridophyte plants), calcium carbonate (*Urticaceae*, *Moraceae*, *Ulmaceae*), and calcium oxalate (bitter vegetables). To successfully isolate intact plastids from such plants, the most effective approach is to increase the amount of sample.
3. High contents of polyphenols and water-soluble proteins cause the chloroplasts and their membranes to aggregate, thus impeding solubilization by sodium dodecyl sulfate (SDS). It should be noted that tobacco contains a large amount of polyphenols [7].
4. Solutions that contain HEPES must be kept in darkness to prevent the formation of hydrogen peroxide [14].
5. A high concentration of EDTA is important to keep the chloroplasts intact and active [15].
6. In rice, the efficiency of plastid extraction may increase when it is performed with a high concentration of sucrose (0.6 M).
7. Sodium ascorbate inhibits myrosinase, an enzyme that hydrolyzes glucosinolates, which cause cell-organelle damage. It is necessary to note, however, that myrosinase is active at low concentrations (0.1–10 mM) of sodium ascorbate [10].
8. Mg^{2+} is crucial for the stability of biological membranes. However, thylakoid membranes from broken chloroplasts are easily aggregated under a high concentration of Mg^{2+} .
9. The addition of cetyltrimethylammonium bromide (CTAB) improves the efficiency of protein extraction from starchy tissues (plastids) by decreasing the viscosity of the extract.
10. For example, rosette leaves harvested from *Arabidopsis* plants grown for 10–14 days under a short-day regime (10-h light/14-h darkness or 8-h light/16-h darkness) to suppress flower stalk development (bolting) [10, 15].

11. In the case of monocotyledons, remove the midrib and veins of the leaf beforehand [16].
12. Perform the steps from extraction to centrifugation quickly for best results. If the sample turns brown, add 5–20 mM of a reducing agent such as DTT or ascorbic acid to the extraction solution.
13. It is desirable to perform all of the steps in low-light condition to reduce the risk of plastid destruction.
14. For plants such as rice, due to the formation of calcium carbonate crystals, it is necessary to perform tissue disruption gently. Chopping with a razor blade may increase the efficiency of extraction of intact chloroplasts. The use of a chilled blender can increase the yield of chloroplasts, but with lower purity.
15. Increasing the centrifugal force (maximum $8,000 \times g$) results in a higher yield of chloroplasts but also increases contamination from broken organelles.
16. The pelleted chloroplasts are easily damaged. Thus, the buffer solution should be applied to the pellet using a pipette tip with the end cut, allowing the suspension to be formed by the pressure of the applied solution.
17. Nycodenz (Axis-Shield PoC AS, Oslo, Norway) can also be used instead of Percoll [17]. Nycodenz has a high separation capacity, low cellular toxicity, and no enzyme activity inhibition.
18. Intact chloroplasts can be easily observed under a light microscope. An intact plastid has a globular shape with glittering surroundings, whereas a broken chloroplast has a distorted shape with a dark green interior.
19. Chloroplasts easily break and their intactness cannot be maintained for a prolonged period even on ice. They are also destroyed by freezing and thawing. It is recommended that isolated chloroplasts be kept in a sucrose solution (0.5–1 M).
20. A large amount of pigment affects the efficiency of protein identification by LC-MS/MS. Therefore, it should be removed completely by TCA/acetone precipitation beforehand.
21. Urea is effective in dissolving protein aggregation, but its final concentration should be kept to less than 1 M.
22. The protein samples should be adjusted with dissolution buffer to an equal concentration.
23. If the sample contains no urea, incubate at 60 °C.

Acknowledgement

This research was supported by Grants-in-Aid for Scientific Research (B) (22380186) and Scientific Research on Innovative Areas (22114507) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1. Lopez-Juez E, Pyke KA (2005) Plastids unleashed: their development and their integration in plant development. *Int J Dev Biol* 49:557–577
2. Mizusawa N, Yamamoto N, Miyao M (1999) Characterization of damage to the D1 protein of photosystem II under photoinhibitory illumination in non-phosphorylated and phosphorylated thylakoid membranes. *J Photochem Photobiol B Biol* 48:97–103
3. Perry SE, Li HM, Keegstra K (1991) In vitro reconstitution of protein transport into chloroplasts. *Methods Cell Biol* 34:327–344
4. Waegemann K, Soll J (1991) Characterization of the protein import apparatus in isolated outer envelope of chloroplast. *Plant J* 261:89–92
5. Nanjo Y, Oka H, Ikarashi N et al (2006) Rice plastidial N-glycosylated nucleotide pyrophosphatase/phosphodiesterase is transported from the ER-golgi to the chloroplast through the secretory pathway. *Plant Cell* 18:2582–2592
6. Blomqvist LA, Ryberg M, Sundqvist C (2008) Proteomic analysis of highly purified prolamellar bodies reveals their significance in chloroplast development. *Photosynth Res* 96:37–50
7. Miyao M (2009) Isolation of intact chloroplasts, thylakoids and photosystem II membranes from higher plants. *Low Temp Sci* 67:197–203
8. Kley J, Heil M, Muck A, Svatos A, Boland W (2010) Isolating intact chloroplasts from small *Arabidopsis* samples for proteomic studies. *Anal Biochem* 398:198–202
9. Shao J, Zhang Y, Yu J, Guo L, Ding Y (2011) Isolation of thylakoid membrane complexes from rice by a new double-strips BN/SDS-PAGE and bioinformatics prediction of stromal ridge subunits interaction. *PLoS One* 6:e20342. doi:[10.1371/journal.pone.0020342](https://doi.org/10.1371/journal.pone.0020342)
10. Schulz A, Knoetzel J, Scheller HV, Mant A (2004) Uptake of a fluorescent dye as a swift and simple indicator of organelle intactness: import-competent chloroplasts from soil-grown *Arabidopsis*. *J Histochem Cytochem* 52:701–704
11. Fukao Y, Ferjani A, Tomioka R et al (2011) iTRAQ analysis reveals mechanisms of growth defects due to excess zinc in *Arabidopsis*. *Plant Physiol* 155:1893–1907
12. Sakai T, Hayashi K (1973) Studies on the distribution of starch and sugary leaves in monocotyledonous plants. *J Plant Res* 86:13–25
13. Lilly JW, Havey MJ, Jackson SA, Jiang J (2001) Cytogenomic analyses reveal the structural plasticity of the chloroplast genome in higher plants. *Plant Cell* 13:245–254
14. Zigler JS Jr, Lepe-Zuniga JL, Vistica B, Gery I (1985) Analysis of the cytotoxic effects of light-exposed HEPES-containing culture medium. *In Vitro Cell Dev Biol* 21:282–287
15. Kunst L (1998) Preparation of physiologically active chloroplasts from *Arabidopsis*. In: Martinez-Zapater JM, Salinas J (eds) *Methods in molecular biology: *Arabidopsis* protocols*, vol 82. Humana, Totowa, NJ, pp 43–48
16. Oldenburg DJ, Bendich AJ (2004) Changes in the structure of DNA molecules and the amount of DNA per plastid during chloroplast development in maize. *J Mol Biol* 344:1311–1330
17. von Zychlinski A, Kleffmann T, Krishnamurthy N et al (2005) Proteome analysis of the rice etioplast: metabolic and regulatory networks and novel protein functions. *Mol Cell Proteomics* 4:1072–1084

Chapter 33

Shotgun Proteomics of Plant Plasma Membrane and Microdomain Proteins Using Nano-LC-MS/MS

Daisuke Takahashi, Bin Li, Takato Nakayama, Yukio Kawamura, and Matsuo Uemura

Abstract

Shotgun proteomics allows the comprehensive analysis of proteins extracted from plant cells, subcellular organelles, and membranes. Previously, two-dimensional gel electrophoresis-based proteomics was used for mass spectrometric analysis of plasma membrane proteins. In order to get comprehensive proteome profiles of the plasma membrane including highly hydrophobic proteins with a number of transmembrane domains, a mass spectrometry-based shotgun proteomics method using nano-LC-MS/MS for proteins from the plasma membrane proteins and plasma membrane microdomain fraction is described. The results obtained are easily applicable to label-free protein semiquantification.

Key words Plasma membrane, Detergent-resistant membrane, Microdomain, Nano-LC-MS/MS, Shotgun proteomics, Label-free semiquantification, In-gel digestion, In-solution digestion

1 Introduction

Comprehensive protein identification consists of solubilization and pre-separation of proteins, peptide digestion and fragmentation using trypsin, and separation and detection of each peptide with liquid chromatography-tandem mass spectrometer (LC-MS/MS) [1–4]. Compared with soluble protein proteomics, the solubilization and pre-separation steps for proteomics of cellular membranes including the plasma membrane (PM) is difficult because of a large number of the highly hydrophobic properties of the proteins and its highly hydrophobic lipid environments [5–7]. Although membrane proteomics has been performed by two-dimensional gel electrophoresis (2-DE)-based proteomics [5, 7, 8], PM proteins are particularly difficult to solubilize and 2-DE-based proteomics requires a large amount of valuable PM proteins [9]. Microdomains, which are considered to exist as extremely hydrophobic compartments in the PM because of the enrichment of specific lipids and

are isolated as detergent-resistant membrane (DRM) fractions, have been studied because of their involvement in important cellular process in PM areas: microdomains are even more intratable to comprehensive proteomics study [10–15]. These factors together have challenged researchers to obtain a comprehensive view of the PM and microdomain proteomic profiles. Liquid chromatography and mass spectrometry technologies have advanced rapidly, providing higher resolution and reliable results for a huge amount of peptides [16, 17]. In particular, nano-flow reverse-phase liquid chromatography allows the separation of proteins without pre-separation using 1-DE or 2-DE [17]. Here, we describe PM and DRM protein preparation methods that are adapted to nano-LC-MS/MS-based shotgun proteomics using two sample preparation methods, “in-gel” and “in-solution” peptide digestion. We used the aerial parts of oat plants as an example and the methods described are applicable to any other plants such as rye [18], *Arabidopsis* [19], and *Brachypodium distachyon* (data not shown). In the “in-gel digestion protocol,” solubilized PM and DRM proteins were applied to sodium dodecyl sulfate (SDS) in a polyacrylamide gel to remove non-proteinaceous materials and subsequently subjected to tryptic digestion. In the “in-solution digestion protocol,” we used an MPEX PTS reagents kit (GL science Inc., Tokyo, Japan), which has been widely used for solubilization of proteins in mammalian and bacteria PMs, such as HeLa cells and *Escherichia coli* cells, respectively [20]. Using these methods, thousands of PM and DRM proteins were consistently identified, including highly hydrophobic proteins with 20+ transmembrane domains. Furthermore, these data can be used for protein quantification.

2 Materials

Prepare all solutions using ultrapure water (prepared by purifying deionized water to attain a sensitivity of $18.2\text{ M}\Omega\text{ cm}$ at $24\text{ }^\circ\text{C}$) and analytical grade reagents. Prepare and store all reagents at room temperature (unless indicated otherwise). Diligently follow all waste disposal regulations when disposing of waste materials.

2.1 Plasma Membrane Purification Components

Several items, including 2 L of ultrapure water, a Polytron homogenizer, centrifuge rotors, and ultracentrifuge rotors, should be chilled at $4\text{ }^\circ\text{C}$.

1. Homogenizing medium: 0.5 M sorbitol, 50 mM Mops-KOH, pH 7.6, 5 mM EGTA (pH 8.0), 5 mM EDTA (pH 8.0), 5 % (w/v) polyvinylpyrrolidone-40 (molecular weight 40,000), 0.5 % (w/v) BSA, 2.5 mM phenylmethanesulfonyl fluoride (PMSF), 4 mM salicylhydroxamic acid (SHAM), 2.5 mM 1,4-dithiothreitol (DTT). Store at $4\text{ }^\circ\text{C}$ (see Note 1).

2. Microsome (MS)-suspension medium: 10 mM KH₂PO₄/K₂HPO₄ (K-P) buffer (pH 7.8), 0.25 M sucrose. Store at 4 °C (*see Note 2*).
3. NaCl medium: Add 1.17 g of NaCl to 180 mL MS-suspension medium and stir moderately using a stirring bar. Add MS-suspension medium up to 200 mL with a graduated cylinder. Store at 4 °C (*see Note 3*).
4. Plasma membrane (PM)-suspension medium: 10 mM Mops-KOH (pH 7.3), 2 mM EGTA (pH 8.0), 0.25 M sucrose. Store at 4 °C (*see Note 4*).
5. Two-phase partition medium: Weigh 1.45 g of polyethylene glycol 3350 and 1.45 g dextran in a 40 mL centrifuge tube. Add 9.3 mL MS-suspension medium and 7.3 mL NaCl medium to the centrifuge tube and mix well by shaking. Incubate at 4 °C overnight to completely dissolve the polymers (*see Note 5*). Prepare three tubes per sample.
6. BioRad Protein Assay Kit (BioRad Laboratories, CA, USA): Store at 4 °C.

2.2 Detergent-Resistant Membrane Extraction Components

Ultracentrifuge rotors should be prechilled at 4 °C.

1. TED buffer: 50 mM Tris-HCl (pH 7.4), 3 mM EGTA (pH 8.0), 1 mM DTT. This buffer should be freshly prepared.
2. 10 % (w/v) Triton X-100 buffer: Add 1 g of Triton X-100 to TED buffer and then adjust to 10 mL volume. Shake the Triton X-100 buffer with a shaker for 3 h to completely dissolve Triton X-100. This buffer should be freshly prepared.
3. 65, 48, 35, 30, and 5 % (w/w) sucrose solution (in TED buffer): Weigh 65, 24, 17.5, 15, and 2.5 g of sucrose and dissolve in 35, 26, 32.5, 35, and 47.5 g of TED solution, respectively. These solutions should be freshly prepared.

2.3 In-Gel Tryptic Digestion

2.3.1 SDS Polyacrylamide Gel Components

If more contaminant-free samples are to be analyzed, precast gels (PAGEL NPU-10L; ATTO Corporation, Tokyo, Japan) should be used instead of making hand-cast gels from the following components.

1. Running gel solution: 1.5 M Tris-HCl, pH 8.8. Add approximately 900 mL water to a 1 L glass beaker. Add 181.7 g Tris and stir moderately using a stirring bar. After Tris is completely dissolved, adjust pH with HCl using a pH meter. Add water up to 1 L with a graduated cylinder (*see Note 6*).
2. SDS sample buffer (2×): 2 % (w/v) SDS, 50 mM Tris-HCl (pH 6.8), 6 % (v/v) β-mercaptoethanol, 10 % (w/v) glycerol, a spatula of bromophenol blue (BPB). Store at 4 °C for current use or at -30 °C for long-term storage (*see Note 7*).

3. TGS running buffer: 0.025 M Tris, 0.188 M glycine, 0.1 % (w/v) SDS (*see Note 8*).
4. 30 % (w/v) acrylamide solution: Weigh 29 g of acrylamide monomer and 1 g methylenebisacrylamide. Add to 50 mL water in 100 mL glass beaker and stir moderately using a stirring bar. After the solids have completely dissolved, add water up to 100 mL with a graduated cylinder and store at 4 °C, with protection from light using a light-shielding bottle or wrapping with aluminum foil (*see Note 9*).
5. 10 % (w/v) SDS: Add 10 g of SDS to 50 mL water and stir moderately using a stirring bar. Add water up to 100 mL with a graduated cylinder.
6. 10 % (w/v) ammonium persulfate: Add 1 g of ammonium persulfate to 8 mL water and stir moderately using a stirring bar. Add water up to 10 mL with a graduated cylinder. Store at 4 °C for current use or at -30 °C for long-term storage.
7. *N,N,N,N'*-Tetramethyl-ethylenediamine (TEMED) (Wako Pure Chemical Industries, Tokyo, Japan). Store at 4 °C.

2.3.2 Tryptic Digestion Components

All of these processes must be carefully performed at a clean bench with gloves and clean lab coat throughout to avoid contamination by keratin, dust, and other exogenous proteinaceous materials.

1. Fixation solution: Mix 50 mL of water, 40 mL of ethanol, and 10 mL of acetic acid.
2. 0.1 M ammonium bicarbonate: Weigh 3.95 g of ammonium bicarbonate and add to 400 mL of water in a glass beaker. Stir moderately using a stirring bar and add water up to 500 mL with a graduated cylinder. Store at 4 °C (*see Note 10*).
3. Acetonitrile (LC-MS grade) (Wako Pure Chemical Industries, Tokyo, Japan). Store at 4 °C (*see Note 10*).
4. 25 mM ammonium bicarbonate/50 % (v/v) acetonitrile: Mix 50 mL of acetonitrile (LC-MS grade), 25 mL of 0.1 M ammonium bicarbonate, and 25 mL of water. Store at 4 °C (*see Note 10*).
5. Reduction buffer: Weigh 7.7 mg DTT and add to 5 mL of 0.1 M ammonium bicarbonate in a conical tube (*see Note 11*).
6. 55 mM iodoacetamide (IAA)/0.1 M ammonium bicarbonate: Weigh 51 mg IAA and add to 5 mL of 0.1 M ammonium bicarbonate in a conical tube (*see Note 11*).
7. Protease solution: Add 2 mL of 0.1 M ammonium bicarbonate into a vial containing 20 µg of trypsin (Sequence grade modified, Promega KK, Tokyo, Japan). Mix well and aliquot into twenty 0.5 mL microtubes. Store at -20 °C.

8. 5 % (v/v) trifluoroacetic acid (TFA)/50 % (v/v) acetonitrile: Mix 450 μ L of water and 500 μ L of acetonitrile in a 1.5 mL microtube. Quickly add 50 μ L of TFA into the solution and mix well (*see Note 11*).
9. 0.1 % (v/v) TFA: Quickly add 1 μ L of TFA into 999 μ L of water and mix well (*see Note 11*).

2.4 In-Solution Tryptic Digestion

All of these processes must be carefully performed at a clean bench with gloves and a clean lab coat throughout to avoid contamination by keratin, dust, and other exogenous proteinaceous materials.

1. MPEX PTS reagents kit (GL science Inc., Tokyo, Japan): Make solution B, DTT solution, IAA solution, and trypsin solution according to the manufacturer's instruction manual. Only solution B should be stored at 4 °C. Prepare fresh DTT solution, IAA solution, and trypsin solution immediately before use.
2. 5 % (v/v) acetonitrile/0.1 % (v/v) TFA: Add 50 μ L of acetonitrile and 1 μ L of TFA into 949 μ L of water and mix well (*see Note 11*).
3. Pierce BCA protein assay kit (Thermo Fisher Scientific, MA, USA): Store at room temperature.

2.5 Peptide Purification Components

1. SPE C-TIP T-300 (Nikkyo Technos Co., Ltd., Tokyo, Japan).
2. 1.5 mL microtubes: Make a hole of 3 mm in diameter on the cap with a soldering iron. Prepare two tubes per sample.
3. Solution A: Add 800 μ L of acetonitrile and 5 μ L of TFA into 195 μ L of water and mix well (*see Note 11*).
4. Solution B: Add 40 μ L of acetonitrile and 5 μ L of TFA into 955 μ L of water and mix well (*see Note 11*).
5. 0.1 % (v/v) TFA: Quickly add 1 μ L of TFA into 999 μ L of water and mix well (*see Note 11*).

3 Methods

Wear gloves and a clean lab coat throughout the processes to avoid contamination by keratin, dust, and other exogenous proteinaceous materials. It is preferable to use low-protein-absorption microtubes at all stages.

3.1 Plasma Membrane Purification

Perform all steps on crushed ice (unless indicated otherwise). Schematic outline of the procedure is described in Fig. 1.

1. Cut out the aerial parts of oat seedlings and weigh the samples (10–70 g in fresh weight is suitable for the plasma membrane purification). Put the harvested plants on a plastic container

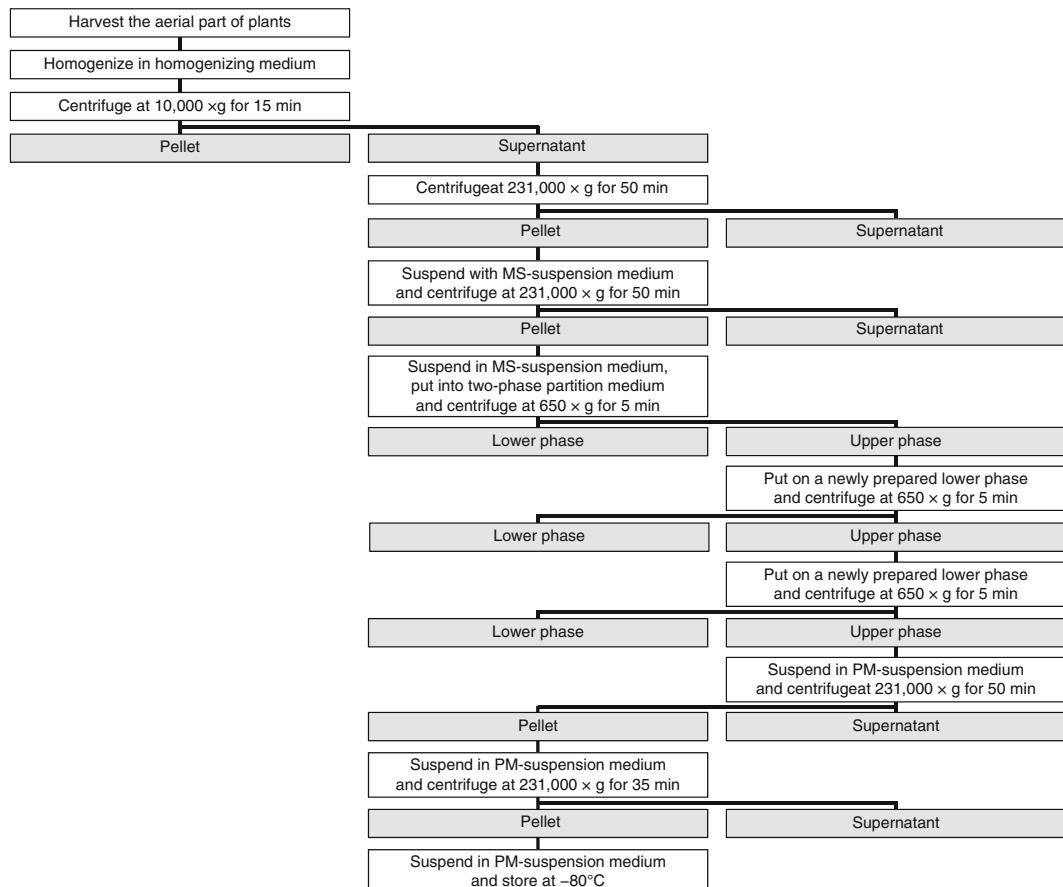


Fig. 1 Schematic outline of the plasma membrane preparation. All steps from harvesting leaves to suspending the purified plasma membrane fractions were described

and wash with 500 mL of chilled water. Wash twice. Drain on a paper towel and put on crushed ice.

2. Cut into small pieces with razor blades. Immediately, put into four volumes of chilled homogenizing medium and mix well with a spatula. The homogenizing medium containing the samples should be again cooled on crushed ice (*see Note 12*).
3. Homogenize with a chilled Polytron generator (PT10SK, Kinematica Inc., Lucerne, Switzerland) until the samples are broken down into tiny pieces (speed 6 for 60–90 s). Filter the homogenates through four layers of gauze and squeeze tightly. Put the filtrates into 40 mL centrifuge tubes and balance them in pairs. Centrifuge at $10,000 \times g$ for 15 min with a chilled rotor to remove debris and heavy membrane fractions (*see Note 13*).

4. Transfer the supernatants into ultracentrifuge tubes by decantation. Centrifuge at $231,000 \times g$ for 50 min with a chilled ultracentrifuge rotor to precipitate the microsome fractions. Discard supernatants by decantation.
5. Add an appropriate quantity of MS-suspension medium to each tube and homogenize the pellets with a Teflon-glass homogenizer. Collect the microsomal suspensions with a large Pasteur pipette into ultracentrifuge tubes. Balance ultracentrifuge tubes in pairs with MS-suspension medium.
6. Ultracentrifuge at $231,000 \times g$ for 50 min as described in **step 4**. Put 5 mL of MS-suspension medium in a Teflon-glass homogenizer and mark the water surface on the glass homogenizer as an indication of 5 mL volume. After centrifugation, discard the supernatant with an aspirator.
7. Put 2 mL of MS-suspension medium and break up the precipitated pellet with a glass rod. Transfer into a Teflon-glass homogenizer using a large Pasteur pipette. Put 2 mL of MS-suspension medium into the tube and pipette up and down to break up the remaining pellet. Transfer into a Teflon-glass homogenizer and add MS-suspension medium to 5 mL. Homogenize well with an electric Teflon-glass homogenizer (moving up and down five times) with cooling on ice (*see Note 14*).
8. Put all of the homogenate in a centrifuge tube containing two-phase partition medium (tube A). Put 5 mL of MS-suspension medium to other two-phase partition systems (tubes B and C). Chill on crushed ice for 10 min. During this time, mix well every 2 min.
9. Centrifuge tubes A and B at $650 \times g$ for 5 min in a chilled rotor. Two phases should be observed to have settled in tubes A and B. Discard the upper phase of tube B with a Pasteur pipette and transfer the upper phase of tube A into tube B. Chill on crushed ice for 10 min. During this time, mix well every 2 min (*see Note 15*).
10. Centrifuge tubes B and C at $650 \times g$ for 5 min in a chilled rotor. Discard the upper phase of tube C with a Pasteur pipette and transfer the upper phase of tube B into tube C. Balance tube C with another centrifuge tube filled with water. Chill on crushed ice for 10 min. During this time, mix well every 2 min (*see Note 15*).
11. Centrifuge at $650 \times g$ for 5 min and split the resultant upper phase of tube C into two ultracentrifuge tubes. Fill up the tubes with PM-suspension medium and balance them. Ultracentrifuge at $231,000 \times g$ for 50 min, as described in **step 4** (*see Note 15*).

12. Discard the supernatant with an aspirator. Add an appropriate quantity of PM-suspension medium to each tube and homogenize the pellets with a Teflon-glass homogenizer. Collect the plasma membrane suspensions with a large Pasteur pipette into ultracentrifuge tubes. Balance ultracentrifuge tubes in pairs with PM-suspension medium. Ultracentrifuge again at $231,000 \times g$ for 35 min.
13. Discard the supernatant with an aspirator. Add a minimal quantity of PM-suspension medium to the plasma membrane pellets. Homogenize the pellets with a glass rod. Transfer into a Teflon-glass homogenizer and homogenize well using an electric Teflon-glass homogenizer (moving up and down five times) with cooling on ice. Transfer into a 1.5 mL microtube.
14. Measure the protein content using the Bradford assay (BioRad Protein Assay Kit). Use 10 μ g of protein for tryptic digestion and LC-MS/MS analysis. The remaining PM fractions should be frozen in liquid nitrogen immediately and stored at -80°C .

3.2 Detergent-Resistant Membrane Extraction

Perform all steps on crushed ice (unless indicated otherwise).

1. Prepare PM with approximately 2.5 mg protein and dilute with PM-suspension medium in an ultracentrifuge tube. After balancing the tubes in pairs, ultracentrifuge at $231,000 \times g$ for 35 min (*see Note 16*).
2. Add 2,000 μ L of PM-suspension medium in an ultracentrifuge tube and grind pellets with a glass rod. Transfer into a Teflon-glass homogenizer and homogenize well using an electric homogenizer (moving up and down five times). Measure the protein content by Bradford assay and place PM samples with 2 mg of protein into a 35 mL swing rotor tube. Adjust the volume to 2.7 mL by adding PM-suspension medium.
3. Add 300 μ L of 10 % (w/v) Triton X-100 buffer and mix well (at this point, protein: detergent ratio is 1:15). Incubate for 30 min.
4. Add 12 mL of 65 % (w/w) sucrose solution and mix well (at this point, the final concentration of sucrose is 52 %). Overlay 5 mL of 48, 35, 30, and 5 % (w/w) sucrose solution slowly in sequence (*see Note 17*).
5. Balance the swing rotor tubes in pairs by adding 5 % (w/w) sucrose solution and ultracentrifuge in a swing rotor at $141,000 \times g$ for 20 h.
6. DRMs will be visible as a white layer at the interface of the 35 %/48 % (w/w) sucrose solution. Recover the white layer and place it in an ultracentrifuge tube. Dilute with TED buffer

and balance them in pairs. Ultracentrifuge (w/w) at $231,000 \times g$ for 35 min (*see Note 18*).

7. Discard the supernatant. Add an appropriate quantity of PM-suspension medium to each tube and homogenize the pellets with a Teflon-glass homogenizer. After balancing in pairs, ultracentrifuge at $231,000 \times g$ for 35 min.
8. Discard the supernatant with an aspirator. Add a minimal quantity of PM-suspension medium to the sample tube. Break up the pellets with a glass rod, transfer into a Teflon-glass homogenizer, and homogenize well using an electric Teflon-glass homogenizer (moving up and down five times). Transfer into a 1.5 mL microtube. The DRM fraction should be frozen in liquid nitrogen immediately and stored at -80°C .

3.3 In-Gel Tryptic Digestion

3.3.1 SDS Polyacrylamide Gel Electrophoresis

1. Mix 2.5 mL running gel solution, 3.35 mL 30 % (w/v) acrylamide solution, and 3.95 mL water in a conical flask. Degas with a vacuum pump for 5 min. Add 100 μL of 10 % (w/v) ammonium persulfate, 100 μL of 10 % (w/v) SDS, and 5 μL of TEMED. Cast gel into a 90 mm (W) \times 83 mm (H) \times 1 mm (T) gel cassette immediately. Insert a 14-well comb without introducing air bubbles. Incubate at room temperature for 1 h (*see Note 19*).
2. Mix 5 μg of PM or DRM protein samples (within 10 μL) and equal volume of SDS sample buffer. Vortex and centrifuge tubes briefly. Heat at 95°C for 5 min. Centrifuge and cool to room temperature.
3. Wash out the wells by pipetting up and down. Slowly load the samples onto the gel. Electrophorese at 100 V until the upper end of sample dye band enters 2 mm from the well (*see Note 20*).
4. Pry the gel plates open with a knife. Cut out the gel slice from the well to 2 mm in front of the BPB dye with a scalpel on a glass plate. Cut the gel slice into four equal pieces (Fig. 2) and put into 1.5 mL microtubes (*see Note 21*).
5. Add 200 μL of fixation solution and agitate for 10 min. Centrifuge briefly and discard the supernatant. Repeat these steps twice.

3.3.2 In-Gel Tryptic Digestion for Nano-LC-MS/MS

All of these procedures must be performed at room temperature unless otherwise specified.

1. Add 200 μL of water and agitate for 10 min. Centrifuge briefly and discard the supernatant.
2. Add 400 μL of 25 mM ammonium bicarbonate/50 % (v/v) acetonitrile and agitate for 10 min. Centrifuge briefly and discard the supernatant.

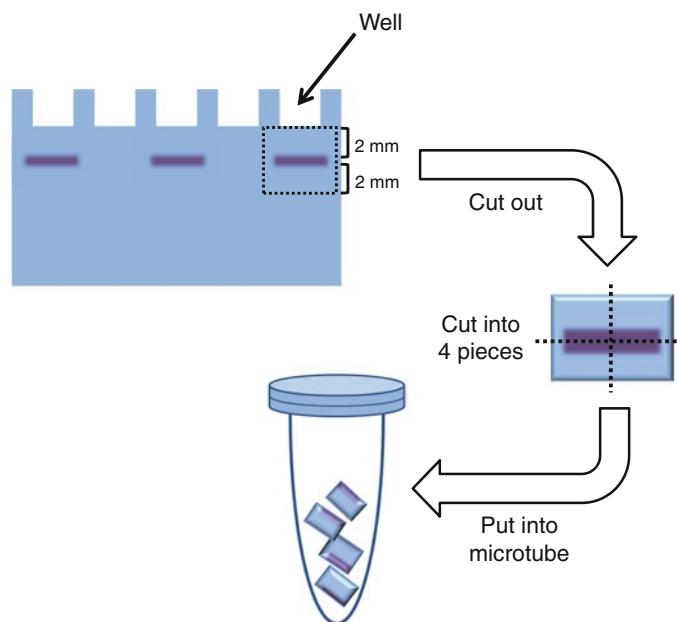


Fig. 2 Excision of a protein band from gel. Wells are separated to avoid contamination of different samples during loading and electrophoresis. After the BPB dye migrates into a gel (about 2 mm), a 4 mm of gel piece centered on the BPB dye band is cut out. Subsequently, the gel piece is cut into four equal pieces. Each of the gel pieces is separately put into 1.5 mL microtubes

3. Add 200 μ L of acetonitrile and incubate at room temperature for 5 min. Centrifuge briefly and discard the supernatant (*see Note 22*).
4. Add 100 μ L of 0.1 M ammonium bicarbonate and centrifuge briefly. Incubate at room temperature for 5 min (*see Note 23*).
5. Add 100 μ L of acetonitrile and centrifuge briefly. Incubate at room temperature for 15 min. Centrifuge briefly and discard the supernatant (*see Note 24*).
6. Dry out the gel samples using a centrifugal concentrator for 45 min (*see Note 25*).
7. Add 100 μ L of reduction buffer and centrifuge briefly. Incubate at 56 $^{\circ}$ C for 45 min. Discard the supernatant.
8. Add 100 μ L of 55 mM IAA/0.1 M ammonium bicarbonate and centrifuge briefly. Incubate in the dark at room temperature for 30 min. Discard the supernatant.
9. Add 200 μ L of water and agitate for 10 min. Centrifuge briefly and discard the supernatant.

10. Add 400 μ L of 25 mM ammonium bicarbonate/50 % (v/v) acetonitrile and agitate for 10 min. Centrifuge briefly and discard the supernatant.
11. Add 200 μ L of acetonitrile and incubate at room temperature for 5 min. Centrifuge briefly and discard the supernatant (*see Note 22*).
12. Add 100 μ L of 0.1 M ammonium bicarbonate and centrifuge briefly. Incubate at room temperature for 5 min (*see Note 23*).
13. Add 100 μ L of acetonitrile and centrifuge briefly. Incubate at room temperature for 15 min. Centrifuge briefly and discard the supernatant (*see Note 24*).
14. Dry out the gel samples using centrifugal concentrator for 45 min (*see Note 25*).
15. Cool to room temperature and put on ice. Add 25 μ L of protease solution to each tube and centrifuge the tubes briefly. Incubate on ice for 45 min (*see Note 23*).
16. Discard the supernatant and add 100 μ L of 0.1 M ammonium bicarbonate. Centrifuge the tubes briefly and incubate at 37 °C for 20 h.
17. Agitate for 15 min and add 100 μ L of acetonitrile. Agitate for 15 min and collect the supernatant (*see Note 24*).
18. Add 5 % (v/v) TFA/50 % (v/v) acetonitrile and agitate for 15 min. Centrifuge the tubes briefly. Collect the supernatant in tubes, as described in **step 17** (*see Note 24*).
19. Dry out the collected supernatant using a centrifugal concentrator for 1 h. Add 30 μ L of 0.1 % TFA. Store at -30 °C (*see Note 26*).

3.4 In-Solution Tryptic Digestion

All of these procedures must be performed at room temperature unless otherwise specified.

1. Precipitate 100 μ g of PM or DRM protein using an ultracentrifuge (231,000 $\times g$, 4 °C, 50 min).
2. Add solution B to the centrifuge tubes and homogenize with a Teflon-glass homogenizer. Transfer to 1.5 mL microtubes.
3. Solubilize samples and measure the protein concentration with a Pierce BCA protein assay kit according to the instruction manual from the manufacturer.
4. Transfer 5 μ g of PM protein to another 1.5 mL microtube. Make up to 20 μ L with solution A.
5. Perform reductive alkylation and tryptic digestion according to the instruction manual, and store at -30 °C (*see Note 26*).

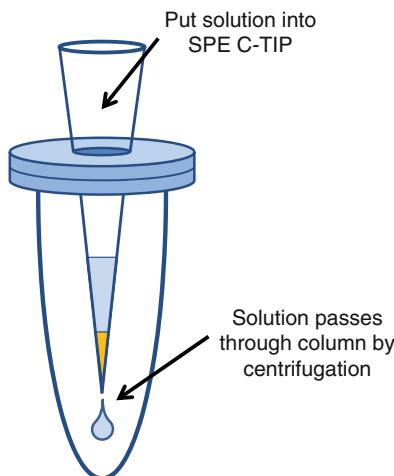


Fig. 3 Peptide purification assembly. A C-TIP is inserted in a hole drilled with a soldering iron in a 1.5 mL microtube cap. Solution A, solution B, and the sample solution are put into the C-TIP, in that order, and then the assembly is centrifuged. The mixed solution passes through a C18 column and peptides are absorbed on, or eluted from, the C-TIP

3.5 Peptide Purification

All of these procedures must be performed at a clean bench whenever possible and at room temperature unless otherwise specified.

1. Insert a SPE C-TIP into the 3 mm hole in the microtube top (Fig. 3).
2. Add 30 μ L of solution A to the upper side of the SPE C-TIP for preconditioning. Centrifuge at $1,000 \times g$ for 30 s to get solution A through the tip column.
3. Add 30 μ L of solution B from upper side of SPE C-TIP for preconditioning. Centrifuge at $1,000 \times g$ for 30 s to get solution B through the tip column.
4. After confirming that the column is moist, add the entire trypsin-digested peptide sample to the upper side of the SPE C-TIP for column absorption. Centrifuge at $1,000 \times g$ for 30 s to get the sample solution through the tip column.
5. Add 30 μ L of solution B from upper side of SPE C-TIP for cleaning. Centrifuge at $1,000 \times g$ for 30 s to get solution B through the tip column.
6. Put a vial insert for each LC-MS/MS sampler into another holed microtube. Transfer the SPE C-TIP into the microtube.
7. Add 30 μ L of solution A to the upper side of the SPE C-TIP for elution. Centrifuge at $1,000 \times g$ for 30 s to get solution A through the tip column. Discard the SPE C-TIP.
8. Dry out the eluted samples using a centrifugal concentrator for 15 min. Add 15 μ L of 0.1 % (v/v) TFA. Put the vial insert into the vial and close the lid. Store at -30°C (see Note 23).

3.6 Nano-LC-MS/MS Analysis

Separate digested and purified peptide solutions with a C18 column by nano-flow LC. Make a linear gradient of acetonitrile (from 5 % [v/v] to 45 % [v/v]) at a flow rate of 500 nL/min for 100 min. Detect and analyze the separated and ionized peptides in a mass spectrometer. Examples of analyzed results using 5 µg of oat PM and DRM proteins are shown in Figs. 4, 5, and 6. You can see detailed results in figure legends.

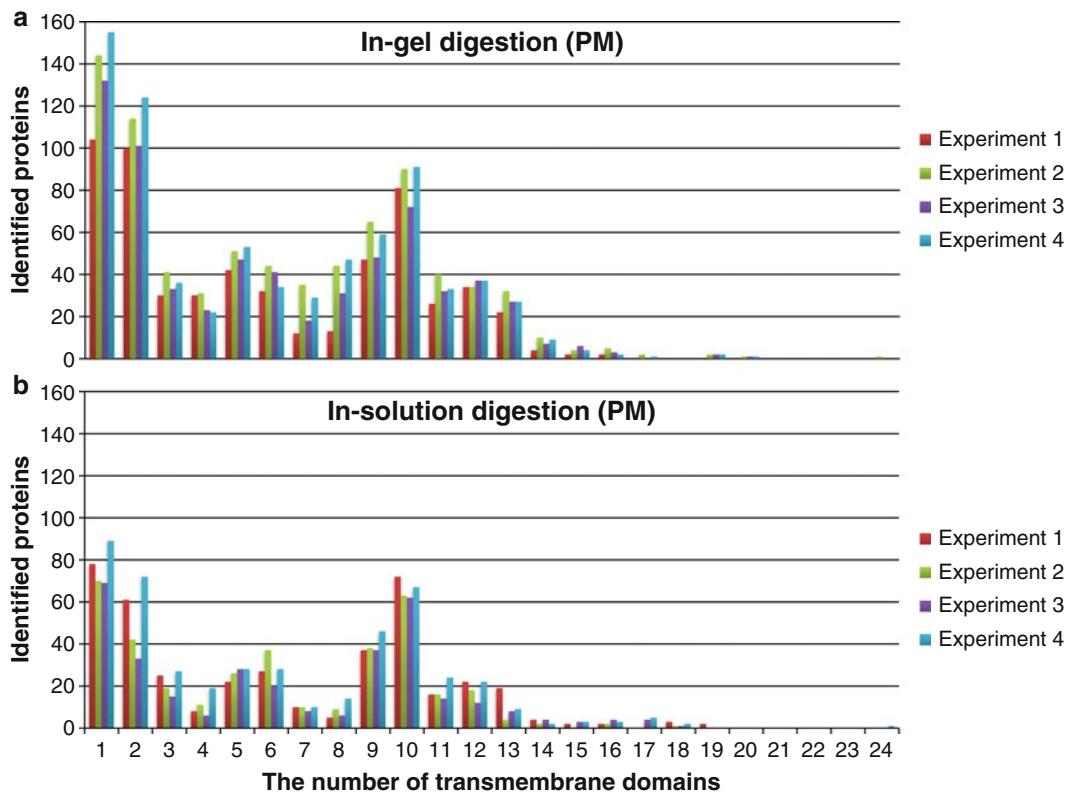


Fig. 4 The number of transmembrane domains in oat PM proteins separated and identified following in-gel or in-solution tryptic digestion. Peptide sequences were searched against the NCBI database (version 20120216, comprising 17,282,984 sequences), taxonomy viridiplantae. Transmembrane domains were estimated by SOSUI engine ver. 1.10 (<http://bp.nuap.nagoya-u.ac.jp/sosui/>). (a) Proteins with up to 24 transmembrane domains were identified in oat PM by in-gel digestion in four biological replicates. On average, 700 proteins with transmembrane domains were identified in the four replicates. (b) Proteins with up to 24 transmembrane domains were identified in oat PM by in-solution digestion in four biological replicates. On average, 397 proteins with transmembrane domains were identified in the four replicates

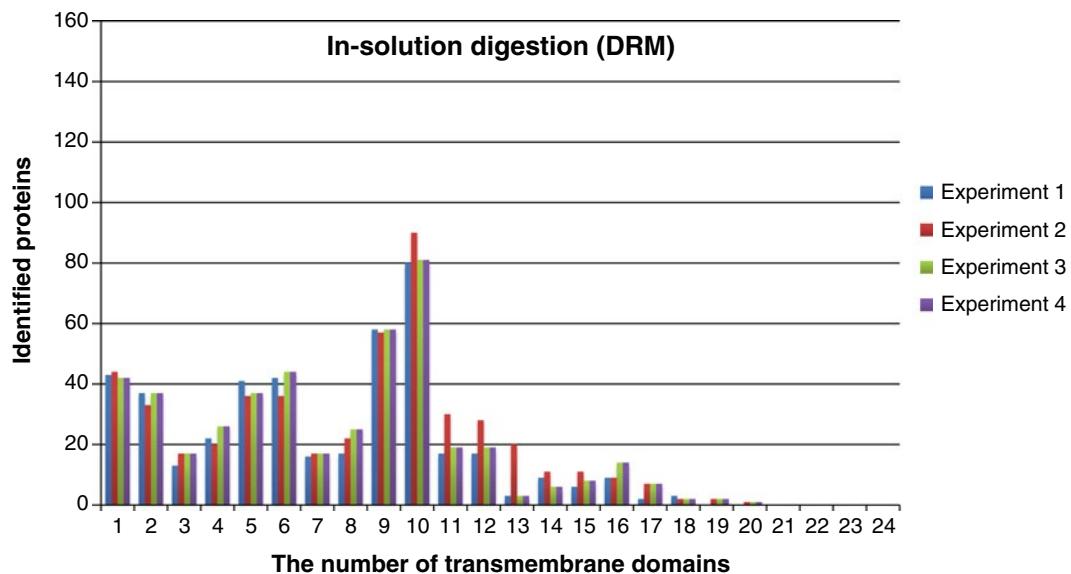


Fig. 5 Distribution of the number of transmembrane domains in oat DRM proteins identified using the “in-solution digestion” protocol. Peptides were analyzed as described in Fig. 4. Compared to PM (Fig. 4b), in DRM, more transmembrane proteins, especially those containing more than five transmembrane domains, could be identified

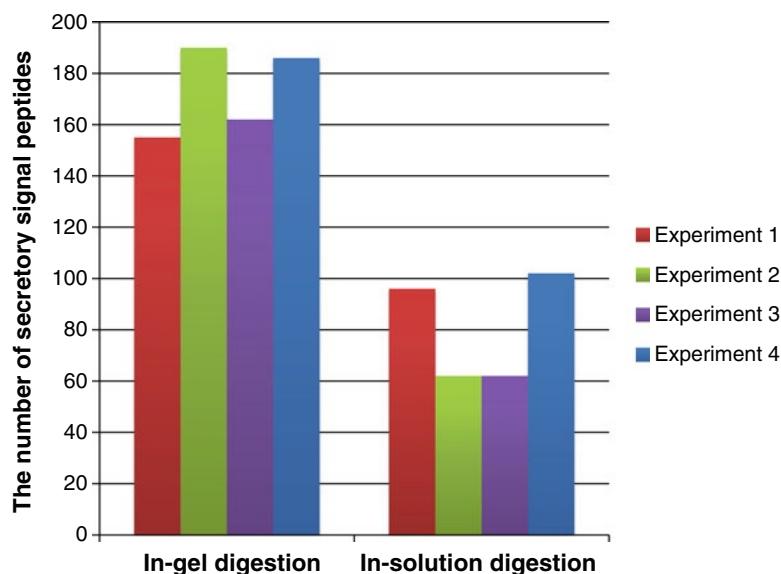


Fig. 6 Estimated number of proteins with secretory signal peptides in oat PM. 173 and 81 proteins with signal peptides were identified from four replicates of in-gel or in-solution tryptic digests of 5 μ g oat PM proteins, respectively

4 Notes

1. Mops-KOH (pH 7.6), EGTA (pH 8.0), and EDTA (pH 8.0) should be prepared as 0.5 M stock solutions and stored at 4 °C. The pH of EGTA and EDTA should be adjusted using NaOH. When BSA is dissolved, BSA powder should be preset at room temperature. PMSF and SHAM should be prepared as 1 and 1.6 M stock solutions in DMSO, respectively, and stored at 4 °C. DTT should be stored at -30 °C as a 1 M stock solution. PMSF, SHAM, and DTT should be diluted only as needed just before use. If you prepare *Arabidopsis* PM fraction, the homogenizing medium should consist of 0.5 M sorbitol, 50 mM Mops-KOH (pH 7.6), 5 mM EGTA (pH 8.0), 5 mM EDTA (pH 8.0), 1.5 % (w/v) polyvinylpyrrolidone-40 (molecular weight 40,000), 0.5 % (w/v) BSA, 2 mM PMSF, 4 mM SHAM, and 2.5 mM DTT.
2. KH₂PO₄/K₂HPO₄ (K-P) buffer (pH 7.8) should be prepared as a 0.5 M stock solution and diluted to make the MS-suspension medium. First, 200 mL of 0.5 M K₂HPO₄ and 30 mL of 0.5 M KH₂PO₄ are prepared. The pH of the 0.5 M K₂HPO₄ is adjusted to 7.8 by adding 0.5 M KH₂PO₄, monitored by a pH meter. If you prepare *Arabidopsis* PM fraction, the MS suspending medium should contain 10 mM KH₂PO₄/K₂HPO₄ (K-P) buffer (pH 7.8) and 0.3 M sucrose.
3. If you prepare *Arabidopsis* PM fraction, the final concentration of NaCl should be adjusted to 100 mM in the MS-suspension medium.
4. Mops-KOH (pH 7.3) and EGTA (pH 8.0) should be prepared as a 0.5 M stock solution and stored at 4 °C. If you prepare *Arabidopsis* PM fraction, the PM suspending medium consists of 10 mM Mops-KOH (pH 7.3), 1 mM EGTA (pH 8.0), and 0.3 M sucrose.
5. For *Arabidopsis* PM preparation, you should weigh 1.4 g of polyethylene glycol 3350 and 1.4 g dextran and add to 9.4 mL MS-suspension medium and 7.3 mL NaCl medium in a 40 mL centrifuge tube.
6. When the pH of the Tris buffer is adjusted, the buffer solution should be at room temperature. The pH of Tris can be affected by the temperature of the solution. Addition of HCl results in an increase in temperature by heat of neutralization and dilution. To avoid a temperature increase of the solution, add the HCl slowly and intermittently.
7. β-mercaptoethanol is a reducing agent and should be added to the sample buffer just before use.

8. TGS buffer is normally made as a 10× stock solution. First, make 1 L of 10× TGS buffer consisting of 30.3 g Tris, 141.4 g glycine, and 10 g SDS. Just before use, dilute 100 mL of 10× TGS buffer with 900 mL of water.
9. Unpolymerized acrylamide is neurotoxic. Acrylamide powder requires careful handling. Wear gloves, a clean lab coat, and a mask, and pay attention to people around you when weighing acrylamide. Store at 4 °C. Add polymerization agent before discarding any spare acrylamide solution.
10. These solutions should be dispensed into a small volume and sealed with parafilm to prevent contamination and evaporation. Store at 4 °C and use within 1 month of preparation.
11. DTT and IAA can be easily modified in solution for a short period and TFA evaporates quickly. Solutions including DTT, IAA, and TFA should be freshly prepared just immediately before use.
12. For *Arabidopsis* PM preparation, plants must be put in homogenizing medium directly and immediately after harvest and washing. Subsequently, plants should be cut with clean scissors in the medium.
13. For *Arabidopsis* PM preparation, the homogenates should be centrifuged at $5,000 \times \text{g}$ for 10 min.
14. In this step, homogenization should not be too long or too vigorous because harsh homogenization can severely disrupt membrane integrity.
15. Two-phase partitioning is the most important step for preparing highly purified PM. When the upper phase of the two-phase partition medium is removed, the Pasteur pipette should be moved from left to right near the boundary of the two phases to prevent taking lower phase. For *Arabidopsis* PM preparation, the two phases should be centrifuged at $440 \times \text{g}$ for 5 min.
16. The yield of the PM preparation is expected to be 2.5 mg protein from 70 to 100 g (FW) of oat leaves.
17. One of the keys to making a good step gradient with sucrose solutions is pouring the solution slowly along the inner wall of the tube.
18. In this step, the upper portions of the white band should be discarded first and then the DRM layer should be collected carefully.
19. All parts of the gel cassette should be wiped with ethanol or acetone on cleaning tissue to prevent contamination with other proteins including keratin.

20. Wells are separated between samples to prevent mixing up samples and electrophoresis. Electric power supply is turned on constant voltage mode.
21. Be careful not to mix with the next sample bands. Illuminate the glass plate from below with a fluorescent lamp to see the gels easily.
22. At this stage, dehydrated, compressed, and completely bleached gels should be observed. If the gels do not change, repeat this step twice.
23. At this stage, rehydrated and swollen gels should be observed. If the gels do not change, repeat this step twice.
24. Gels are sometimes partly bleached, but this is acceptable.
25. At this stage, dehydrated and compressed gels are easily lost by electrostatic force. Take extra care.
26. Digested and purified peptides should be analyzed by nano-LC-MS/MS within 1 week.

Acknowledgements

This work was supported in part by Grants-in-Aid for Scientific Research (#22120003 and #24370018) from MEXT, Japan, to Y.K. and M.U.

References

1. Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. *Nature* 422:198–207
2. Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. *Science* 312:212–217
3. Kersten B, Burkle L, Kuhn E-J et al (2002) Large-scale plant proteomics. *Plant Mol Biol* 48:133–141
4. van Wijk KJ (2001) Challenges and prospects of plant proteomics. *Plant Physiol* 126:501–508
5. Santoni V, Kieffer S, Desclaux D et al (2000) Membrane proteomics: use of additive main effects with multiplicative interaction model to classify plasma membrane proteins according to their solubility and electrophoretic properties. *Electrophoresis* 21:3329–3344
6. Luche S, Santoni V, Rabilloud T (2003) Evaluation of nonionic and zwitterionic detergents as membrane protein solubilizers in two-dimensional electrophoresis. *Proteomics* 3:249–253
7. Gorg A, Weiss W, Dunn M-J (2004) Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 4:3665–3685
8. Rabilloud T (2009) Membrane proteins and proteomics: love is possible, but so difficult. *Electrophoresis* 30:174–180
9. Rabilloud T (2002) Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains. *Proteomics* 2:3–10
10. Simons K, Ikenen E (1997) Functional rafts in cell membranes. *Nature* 387:569–572
11. Peskan T, Westermann M, Oelmüller R (2000) Identification of low-density Triton X-100-insoluble plasma membrane microdomains in higher plants. *Eur J Biochem* 267:6989–6995
12. Mongrand S, Morel J, Laroche J et al (2004) Lipid rafts in higher plant cells. *J Biol Chem* 279:36277–36286
13. Bhat RA, Panstruga R (2005) Lipid rafts in plants. *Planta* 223:5–19
14. Martin SW, Glover BJ, Davies JM (2005) Lipid microdomains: plant membranes get organized. *Trends Plant Sci* 10:263–265
15. Grennan AK (2007) Lipid rafts in plants. *Plant Physiol* 143:1083–1085
16. Thakur S-S, Geiger T, Chatterjee B et al (2011) Deep and highly sensitive proteome

- coverage by LC-MS/MS without prefractionation. *Mol Cell Proteomics* 10:M110.003699
- 17. Matros A, Kasper S, Witzel K et al (2011) Recent progress in liquid chromatography-based separation and label-free quantitative plant proteomics. *Phytochemistry* 72:963–974
 - 18. Takahashi D, Kawamura Y, Yamashita T et al (2011) Detergent-resistant plasma membrane proteome in oat and rye: similarities and dissimilarities between two monocotyledonous plants. *J Proteome Res* 11:1654–1665
 - 19. Li B, Takahashi D, Kawamura Y et al (2012) Comparison of plasma membrane proteomic changes of *Arabidopsis* suspension cells (T87 line) after cold and abscisic acid treatment in association with freezing tolerance development. *Plant Cell Physiol* 53: 543–554
 - 20. Masuda T, Tomita M, Ishihama Y (2008) Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *J Proteome Res* 7:731–740

Chapter 34

Plant Mitochondrial Proteomics

Shaobai Huang, Richard P. Jacoby, A. Harvey Millar, and Nicolas L. Taylor

Abstract

Mitochondria are responsible for a number of major biochemical processes in plant cells including oxidative phosphorylation and photorespiration. Traditionally their primary role has been viewed as the oxidation of organic acids via the tricarboxylic acid cycle and the synthesis of ATP coupled to the transfer of electrons to O₂. More recently its role in the synthesis of many metabolites such as amino acids, lipids, and vitamins has been revealed. They also contain large number of transporters including members of the mitochondrial carrier substrate family (MCSF) that allow the exchange of metabolites with the cytosol. Mitochondria also contain their own genome and actively transcribe and translate a set of proteins that are coordinated with proteins encoded by the nuclear genome to produce large multisubunit enzymes. To reveal the full diversity of metabolism carried out by mitochondria significant efforts have sought to uncover the protein profile of mitochondria from both crops and model plants. Successful proteomic analysis depends on the preparation of high-quality isolated mitochondria, coupled to high-resolution proteomic techniques for identification, quantitation, and assessment of the degree of contamination by other organelles and cellular compartments. Here we outline a mitochondrial isolation protocol that can be applied to a range of plant tissues, and detail methods of assessing the quality and purity of the resultant sample, including calculations of respiratory control ratio, marker enzyme assays, differential in-gel electrophoresis, and quantitative gel-free mass spectrometry.

Key words Proteomics, Mitochondria, Mass spectrometry, Gradient purification, 2D-Polyacrylamide gel electrophoresis, 2D-Differential in-gel electrophoresis, Spectral counting

1 Introduction

Plant mitochondria are well known as the site of production of the mobile energy currency ATP, for their role in the recycling of phosphoglycolate in photorespiration and the oxidation of organic acids by the TCA cycle. In addition they also have well-studied roles in plant development, fertility, abiotic stress tolerance, disease susceptibility and defense, senescence, and programmed cell death. Recent advances in peptide mass spectrometry coupled with the complete sequencing and extensive annotation of the nuclear and organelle genomes of a variety of plant species have been impor-

tant for identifying the 1,000–1,500 proteins thought to be present in plant mitochondria. Mitochondria have also provided an excellent model system through which a number of novel proteomic analyses have been developed. They are relatively discrete membrane-bound organelles that are found in significant numbers in most plant tissues and typically they represent 2–5 % of total cellular protein. Procedures for mitochondrial isolations from different plants are well established and are capable of producing milligrams of mitochondrial proteins from 10 to 100 g of fresh plant material. The majority of the protein complement is within the dynamic range of standard proteomic techniques, although low-abundance and hydrophobic proteins still represent challenging targets.

A variety of studies have now begun to uncover the proteome of mitochondria from *Arabidopsis* [1–24], rice [5, 25–29], wheat [30, 31], maize [32], barley [33], pea [34, 35], soybean [36, 37], *medicago* [38], and *Chlamydomonas* [39]. Gel-based protein separation of proteins followed by mass spectrometry identification has been the most widely used strategy to date to determine proteome composition. Isoelectric focusing (IEF)/SDS-PAGE is valued for its reproducible and well-resolved separation of soluble proteins, while blue native (BN)/SDS-PAGE has been deployed to define the protein components of the large protein complexes I–V of the respiratory chain, and has proven particularly successful for visualizing hydrophobic proteins not amenable to IEF [8, 18, 19]. Gel-bound proteins are routinely identified by a wide range of tandem mass spectrometry techniques, while traditional approaches such as Edman degradation and peptide mass fingerprinting (PMF) [1, 2, 40] have been largely superseded by ESI-MS/MS [3, 12, 19, 20, 24] and MALDI-TOF/TOF [41]. Larger scale analyses have digested complex mitochondrial samples in the liquid phase, and used high-performance liquid chromatography (HPLC) of peptides coupled to tandem mass spectrometry to provide many hundreds of identification of mitochondrial proteins [6, 7, 24, 28]. Critical to studies of mitochondria is the deduction of which proteins are of mitochondrial origin and which are non-mitochondrial. Contaminants can co-localize with mitochondrial fractions during isolation due to similarities in density (as with Percoll® centrifugation) or charge (as with free flow electrophoresis (FFE)) or may be connected or associated biochemically such as components of glycolysis [42, 43] or ascorbic acid (vitamin C) biosynthesis [44]. The use of quantitative proteomics platforms to compare protein abundance in enriched versus depleted mitochondrial fractions provides a high-throughput basis for determining the subcellular localization of proteins [9, 17, 24]. The following techniques will outline a basic procedure for isolating plant mitochondria, investigating the proteome using 2D-SDS-PAGE-DIGE or LC-MS/MS of peptides, and defining enrichment or contamination by differential

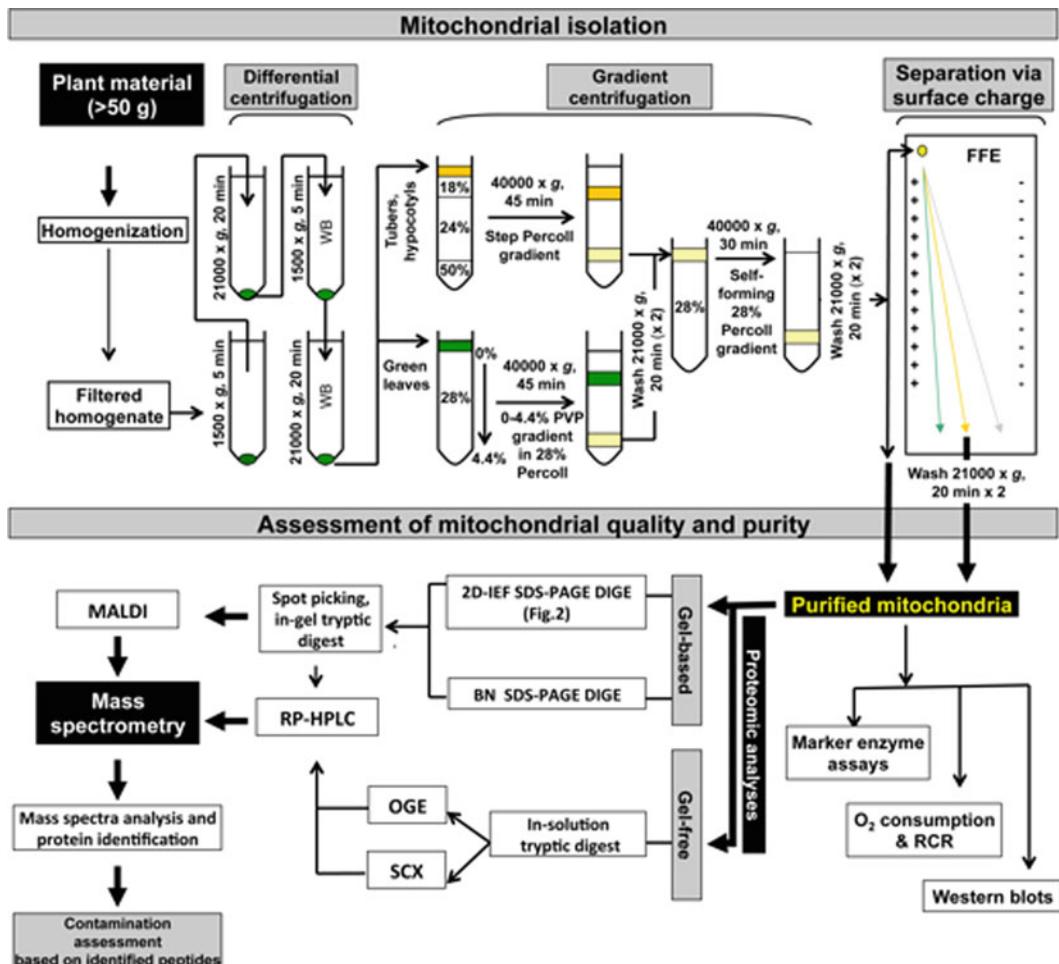


Fig. 1 Overview of the workflow for the purification of plant mitochondria and the determination of their quality and purity. Plant material is homogenized in extraction buffer and the homogenate is filtered. The resulting homogenate is centrifuged first at low speed and then at high speed, which is repeated to generate a crude organelle pellet. The organelles are further separated by gradient centrifugation using step or continuous gradients depending on the original plant material. The resulting mitochondria fraction is washed and collected as purified mitochondria. Further purification of mitochondria can be carried out using free flow electrophoresis (FFE) depending on the accessibility of hardware and the objective of experiment. The quality and purity of isolated mitochondria can be determined by specific marker enzyme assays, O₂ consumption and RCR calculations and western blots. For proteomic analysis, the purified mitochondria can be separated using gel-based or gel-free methods. For gel-based method, BN-SDS-PAGE-DIGE and 2D-IEF-SDS-PAGE-DIGE are normally used. The gel spots containing mitochondria proteins are collected and digested with trypsin for further mass spectrometry analysis. For gel-free method, whole mitochondria are digested with trypsin and then can be separated by off-gel electrophoresis (OGE) or strong cation exchange (SCX) prior to RP-HPLC and mass spectrometry analysis

in-gel electrophoresis (DIGE) or spectral counting approaches. An overview of the workflow for the isolation of plant mitochondria, fractionation of proteins/peptides, and determination of contaminants is presented in Fig. 1.

2 Materials

The procedure for the isolation of plant mitochondria contains steps of homogenization, differential centrifugation, and isopycnic centrifugation, while in some cases FFE is applied as a second dimension of purification based on surface charge (Fig. 1). In this chapter, the FFE procedure is not described as the hardware required is not widely available and we have chosen to focus on techniques accessible to most laboratories (for further information on FFE isolation of mitochondria, *see* refs. 9, 17, 28).

2.1 Preparation of Density Gradients for Mitochondrial Isolation

1. Gradient buffer 1: 20 mM TES-KOH, pH 7.5, 0.6 M mannitol, 0.2 % (w/v) BSA.
2. Gradient buffer 2: 20 mM TES-KOH, pH 7.5, 0.6 M sucrose, 0.2 % (w/v) BSA.
3. Percoll® (GE Healthcare Life Sciences).
4. Ice bucket with 2 L of ice.
5. 25 mL syringe with a 19-gauge hypodermic needle.
6. Rack for the syringe.
7. Two-chamber gradient pourer and peristaltic gradient pump (for isolation of mitochondria from green tissues).

2.2 Mitochondrial Isolation

1. Plant tissue—Mitochondria have been isolated to varying degrees of purity from a wide variety of plants and plant tissues. When choosing plant material for mitochondrial isolations it is important to choose a tissue that can be produced in large quantities (>50 g), is amenable to homogenization by either blending techniques (i.e., Polytron or Waring Blender) or grinding in a mortar and pestle, and contains low levels of interfering compounds such as phenolics; particular tissue types are advantageous in this regard:
 - Nonphotosynthetic tissues, such as roots, hypocotyls, tubers, and cell cultures, can be used for large-scale, but low-percentage-yield, mitochondrial isolation (approximately 300 mg mitochondrial protein from 5 to 10 kg fresh weight (FW), 30–60 µg/g FW) [1, 2, 6, 7, 21, 24].
 - Etiolated seedling tissues such as dark-grown cotyledons or shoots produce high yields of functional mitochondria because they are free of dense chloroplast membranes with lower phenolic content than green tissues (approximately 20 mg mitochondrial protein from 100 g FW, 200 µg/g FW) [28, 45].
 - Green tissues, such as leaves, are used for mitochondria isolation for photosynthetic-related function analysis, although the high abundance of thylakoid membranes and phenolics

presents particular challenges (5–10 mg mitochondrial protein from 50 to 100 g FW, 100 µg/g FW) [2, 12, 35, 46].

- Certain tissues, such as stems, seeds, succulents, embryos, and fruits, are often more difficult to isolate mitochondria from although it is possible, but specialized techniques are required and these are beyond the scope of this chapter.

The power of proteomics is largely dictated by the availability of protein sequences from the species being studied, so therefore species with high-quality publicly available genome sequences such as *Arabidopsis* and rice are preferred for proteomics investigations due to the ease of analysis and interpretation of mass spectrometry data. However species with un-sequenced genomes can be matched against EST databases or matched across species to plant databases (i.e., NCBInr—Viridiplantae), although the limitations of these approaches should be fully appreciated before proceeding [47].

2. Homogenization solution: 0.4 M sucrose (*see Note 1*), 5 mM EGTA (*see Note 2*), 50 mM sodium pyrophosphate–KOH, pH 7.5 (*see Note 3*), 0.5 % (w/v) BSA (*see Note 4*), 10 mM ascorbate (*see Note 5*), and 1 % (w/v) PVP-40 (*see Note 6*). Prepare this media 1 day before the isolation and store it overnight at 4 °C (*see Note 7*).
3. Wash solution (WS): 10 mM TES-KOH, pH 7.2, 0.3 M sucrose, and 0.1 % (w/v) BSA.
4. Buchner funnel.
5. Sterile distilled water.
6. Perspex vessel (45 × 60 × 220 mm).
7. Polytron® homogenizer equipped with a 200 mm long × 20 mm diameter dispersing head (*see Note 8*).
8. Miraclot (Merck Millipore).
9. Muslin.
10. 50 mL centrifuge tubes.
11. Centrifuge with fixed-angle rotor (Beckman Coulter Avanti J-26XP with JA-25.50 rotor or an equivalent).
12. Soft-haired artist's paintbrush.
13. 28 % (v/v) Percoll®, 0–4.4 % (w/v) PVP-40 gradients (*see Subheading 3.2.3*).
14. 3 mL disposable pipette.
1. Mitochondria isolated according to Subheading 3.2.
2. Clark-type electrode, such as those manufactured by Hansatech (Kings Lynn, UK) or Rank Brothers (Bottisham, UK).

2.3 Determining Mitochondria Quality

3. Mitochondrial reaction medium: 0.3 M mannitol, 10 mM TES-KOH (pH 7.5), 3 mM MgSO₄·7H₂O, 10 mM NaCl, 5 mM KH₂PO₄, 0.1 % (w/v) BSA.
4. Stock solutions of respiratory substrates: 100 mM NADH; 1 M succinate; 50 mM ATP; 500 mM pyruvate; 50 mM malate; 30 mM NAD⁺; 10 mM thiamine pyrophosphate; 12 mM coenzyme A. Coenzyme A should be made up in 1 % (w/v) cysteine and all other components should be made up in 500 mM TES-KOH, pH 7.5.
5. 100 mM Adenosine 5'-diphosphate (ADP) made up in 500 mM TES-KOH, pH 7.5.

2.4 Determining Mitochondria Purity

2.4.1 Marker Enzyme Assays of Mitochondria

Fumarase

1. Visible wavelength spectrophotometer.
2. 0.1 M KH₂PO₄–NaOH, pH 7.7.
3. 10 % (v/v) Triton X-100.
4. 1 M Malate.
5. Mitochondria isolated according to Subheading 3.2.
6. Reaction master mix: 70 mM KH₂PO₄–NaOH, pH 7.7, 0.05 % (v/v) Triton X-100. 900 µL master mix per assay (i.e., to fill a 1 mL spectrophotometric cuvette after subsequent additions of sample and substrate) (see Note 9).

Aconitase

1. Visible wavelength spectrophotometer.
2. 0.1 M HEPES–NaOH, pH 7.5.
3. 10 % (v/v) Triton X-100.
4. 20 mM Nicotinamide adenine dinucleotide phosphate (NADP).
5. 0.1 M MnCl₂.
6. 2,000 U/mL NADP-isocitrate dehydrogenase (ICDH) (porcine heart).
7. 0.2 M Aconitate.
8. Mitochondria isolated according to Subheading 3.2.
9. Reaction master mix: 80 mM HEPES–NaOH, pH 7.5, 0.05 % (v/v) Triton X-100, 0.5 mM NADP, 0.5 M MnCl₂, 2 U NADP-ICDH. 900 µL master mix per assay (i.e., to fill a 1 mL spectrophotometric cuvette after subsequent additions of sample and substrate) (see Note 9).

2.4.2 Marker Enzyme Assay of Chloroplast Contamination

Phosphoribulokinase

1. Mitochondria isolated according to Subheading 3.2.
2. 2× reaction buffer: 200 mM Tris–HCl (pH 7.8), 20 mM MgCl₂, 40 mM KCl, 20 mM DTT.
3. 100 mM adenosine 5'-triphosphate (ATP) disodium salt in 1× reaction buffer.

4. 100 mM phospho(enol)pyruvic acid monosodium salt hydrate in 1× reaction buffer.
5. 10 mM NADH in 1× reaction buffer, prepare only on the day of use.
6. Pyruvate kinase (e.g., rabbit muscle), 400–800 U/mg protein.
7. Lactate dehydrogenase (e.g., *Lactobacillus leichmannii*), 150–300 U/mg protein.
8. H₂O.
9. Visible wavelength spectrophotometer.

2.4.3 Marker Enzyme Assay of Peroxisome Contamination

Catalase

1. Clark-type electrode, such as those manufactured by Hansatech (Kings Lynn, UK) or Rank Brothers (Bottisham, UK).
2. Reaction medium: 10 mM TES-KOH, pH 7.2, 0.3 M sucrose, 5 mM KH₂PO₄, 10 mM NaCl, 2 mM MgSO₄, and 0.1 % (w/v) BSA.
3. Sodium hydrosulfite.
4. 1 % (v/v) H₂O₂ solution.
5. Mitochondria isolated according to Subheading 3.2.

2.4.4 Differential In-Gel Electrophoresis of Purified Mitochondrial Proteins

There are two main gel-based methods to separate the plant mitochondrial proteins for DIGE analysis (Fig. 1). Blue-native (BN) SDS-PAGE DIGE is well suited to profile the composition of individual ETC complexes, and detailed procedures of BN-SDS-PAGE [48] and BN-SDS-PAGE-DIGE [49] are available and are not covered here. 2D-IEF-SDS-PAGE is a highly reproducible platform that is commonly used for mitochondrial proteomic analysis, and the combination of DIGE and 2D-IEF-SDS-PAGE (2D-IEF-SDS-PAGE-DIGE) provides a powerful tool to define contaminants when comparing two mitochondrial isolation procedures.

Mitochondrial Precipitation and Resuspension

1. Acetone.
2. Microcentrifuge such as Eppendorf 5430 or an equivalent.
3. Lysis buffer 1: 8 M Urea, 4 % (w/v) CHAPS, 40 mM Tris (2-amino-2-hydroxymethyl-propane-1,3-diol) (see Note 10).

CyDye Labelling

1. Cy-Dyes Cy5, Cy3, and Cy2 (GE Healthcare Life Sciences).
2. Dimethylformamide (DMF) (water-free).
3. 10 mM Lysine.
4. Vortex and microfuge.
5. Lysis buffer 2: 8 M Urea, 4 % (w/v) CHAPS, 40 mM Tris (2-amino-2-hydroxymethyl-propane-1,3-diol), 72 mM DTT (see Note 10).

Isoelectric Focusing

1. Rehydration buffer: 8 M Urea, 2 % (w/v) CHAPS, 0.05 % (w/v) bromophenol blue. Just prior to use add 0.5 % (v/v) IPG buffer (pH 3–10 NL, GE Healthcare Life Sciences) and 18 mM DTT (*see Note 10*).
2. Vortex and microcentrifuge such as Eppendorf 5430 or an equivalent.
3. Ettan ceramic strip holder or equivalent (GE Healthcare Life Sciences) for IEF equipment being used.
4. Immobiline DryStrips 3–10 NL 240 mm (GE Healthcare Life Sciences) or equivalent.
5. IPG cover fluid (GE Healthcare Life Sciences) or mineral oil equivalent.
6. Ettan IPGphor 3 (GE Healthcare Life Sciences) or an equivalent.

Transferring and Setting IEF Strips onto SDS-PAGE

1. 12 % SDS-PAGE gel solution: 375 mM Tris–HCl, pH 8.8, 0.1 % (w/v) SDS, 0.36 % (w/v) bisacrylamide, 12 % (w/v) acrylamide, with 0.1 % (w/v) AMPS, 0.04 % (v/v) TEMED being added just before use.
2. Forceps.
3. Equilibration buffer: 50 mM Tris–HCl, pH 8.8, 6 M urea, 2 % (w/v) SDS, 26 % (v/v) glycerol, 0.05 % (w/v) bromophenol blue. DTT (65 mM) or IAA (135 mM) are added freshly before use.
4. Rinse solution: 1.5 M Tris–HCl pH 8.8, 1 % (w/v) SDS.
5. IPG fixing solution: 375 mM Tris–HCl, pH 8.8, 1 % (w/v) SDS, 0.5 % (w/v) agarose, 0.5 % (w/v) bromophenol blue.

Gel Scanning and Software Analysis

1. Software analysis packages such as DeCyder 2D Software (GE Healthcare Life Sciences) or DeCodon Software (DECODON).

Extraction and Digestion of Proteins That Differ in Abundance on DIGE

1. V-bottom 96-well plate made of polypropylene such as Greiner Bio One (651201) and sealing film such as Nunc (232702) or equivalent.
2. Destain solution: 50 % (v/v) ACN, 10 mM NH₄HCO₃.
3. Orbital shaker.
4. Dry block heater.
5. Digestion solution: 12 µg/mL Trypsin, 10 mM NH₄HCO₃, 0.0012 % TFA (*see Notes 11 and 12*).
6. 37 °C heating oven.

Identification of Proteins That Differ in Abundance on DIGE

- A. Method using MALDI-MS/MS.
 1. Peptides extracted from a protein spot as outlined in Subheading 3.4.4.
 2. MALDI mass spectrometry calibration standards.

3. HPLC-grade acetonitrile (ACN), water, trifluoroacetic acid (TFA), and ammonium phosphate (monobasic) ($\text{NH}_4\text{H}_2\text{PO}_4$).
4. Machine-specific MALDI plate.
5. Saturated α -cyano-4-hydroxycinnamic acid (CHCA) matrix solution. Take 150 μL of 90 % (v/v) ACN and 0.1 % (v/v) TFA, add a small spatula of CHCA, vortex, and then sonicate for 15 min. If all matrix is solubilized, then add more CHCA and repeat sonication until undissolved matrix is visible (saturated). Centrifuge at 10,000 $\times g$ for 5 min to pellet undissolved matrix (*see Note 13*).
6. Matrix solution (600 μL). This contains 516 μL 95 % (v/v) ACN, 0.1 % (v/v) TFA, 27 μL saturated CHCA solution, 6 μL 10 % (v/v) TFA, and 6 μL 100 mM $\text{NH}_4\text{H}_2\text{PO}_4$.

B. Method using ESI-MS/MS.

1. Peptides extracted from a protein spot as outlined *above*.
2. HPLC-grade ACN, water, and formic acid (FA).
3. C18 column (*see Note 14*).

2.4.5 Gel-Free Peptide Fractionation and Identification to Determine Mitochondrial Purity

Sample Preparation

1. Resuspension solution: 8 M Urea, 50 mM NH_4HCO_3 , 5 mM DTT, pH 8.0.
2. Vortex.
3. IAA stock solution: 100 mM Iodoacetamide.
4. Dilution buffer: 50 mM NH_4HCO_3 , pH 8.0.
5. Digestion solution: 1 $\mu\text{g}/\mu\text{L}$ trypsin in 1 mM CaCl_2 .
6. Formic acid.
7. C18 spin columns (such as Pierce C18 spin columns, Thermo Scientific) or C18 embedded tips (such as ZipTips, Merck Millipore).
8. Charging and elution solution: 70 % (v/v) ACN, 0.1 % (v/v) FA (HPLC grade).
9. Equilibration and washing solution: 2 % (v/v) ACN, 0.1 % (v/v) FA (HPLC grade).
10. Vacuum centrifuge.

Sample Fractionation and Mass Spectrometry

1. Peptides extracted from a protein spot as outlined in Subheading 3.4.5.
2. HPLC-grade ACN, water, and formic acid.
3. C18 column (*see Note 15*).

Data Analysis

1. Microsoft Excel is useful but not essential as is statistical package such as Analyze-it (www.analyze-it.com) add-on for Microsoft Excel.

3 Methods

3.1 Preparation of Density Gradients

1. The different concentrations of Percoll are prepared in the appropriate gradient buffer. Gradients should be prepared in advance and stored on ice until required.
2. For isolation of mitochondria from nongreen tissues, two different gradients are used. The first is a step gradient consisting of 5 mL of 40 % (v/v) Percoll overlayered with 20 mL 24 % (v/v) Percoll and then 10 mL 18 % (v/v) Percoll, all made up using gradient buffer 1. The layers can be conveniently poured over one another by allowing the solution to run through a 19-gauge hypodermic needle placed against the inside of the centrifuge tube held at a 45° angle. The second gradient consists of 30 mL of 28 % (v/v) Percoll made up in gradient buffer 2. A sigmoidal gradient self-forms during centrifugation due to the sedimentation of the Percoll poly-dispersed silica colloid.
3. For isolation of mitochondria from green tissues, the first gradient is a single linear polyvinylchloride (PVP-40) gradient in 28 % (v/v) Percoll. Preparation of this gradient requires a two-chamber gradient pourer which consists of a Perspex block with two cylindrical chambers connected by a small pipe, with only one of the chambers having an outflow. Both the connecting pipe and outflow are metered by taps. The outflow tube is run through a peristaltic gradient pump set to ~20 rpm, and is then secured against the inside of a 50 mL centrifuge tube held at a 45° angle (ensure that tube outflow is above the final level to which the gradient solution will reach in the tube). It is important that the pump and the tube are held below the level of the gradient maker so that solution will flow by gravity. Two solutions are prepared: a “heavy” solution (in this case 28 % (v/v) Percoll and 4.4 % (v/v) PVP-40 (from a 20 % (w/v) stock) made up in gradient buffer 2) and a “light” solution (28 % (v/v) Percoll made up in gradient buffer 2). The “light” solution (15 mL) is placed in the chamber with no outflow. The tap connecting the two chambers should be briefly opened to displace air in the connecting pipe and ensure flow. The “heavy” solution (15 mL) is placed into the chamber with the outflow. A magnetic stir bar is placed in the chamber with outflow and the solution is rapidly stirred. The taps on both the connecting pipe and the outflow are then opened simultaneously. As the gradient pours, the level of solution in both chambers should drop at the same rate. The second gradient is identical to that used for nongreen tissues.

3.2 Mitochondrial Isolation

The isolation of high-purity intact mitochondria is vital to the success of plant mitochondrial proteome studies. The use of differential and isopycnic centrifugation for mitochondrial isolations from

different tissues results in preparations typically with 92–98 % purity on a protein basis [6]. The intensity and duration of homogenization are the critical factors in determining the trade-off between intactness versus quantity of the isolated mitochondria.

3.2.1 Homogenization

The choice of homogenization technique depends upon the rigidity of the tissue under study. Generally, the more rigid the tissue, the more shearing forces required for tissue disruption and this is best provided by mechanical devices such as a Polytron® homogenizer. This suits fibrous tissues such as rice or wheat shoots whereas pliable tissues such as *Arabidopsis* rosettes and spinach leaves can be homogenized by hand using a mortar and pestle. The optimal ratio of homogenization medium to plant tissue is generally 1 g of tissue to 4–6 mL of homogenizing solution. Altering this ratio may reduce quality or yield. All buffers and equipment should be ice-cold before beginning the isolation.

1. If using mortar and pestle, grind the tissue in homogenization medium using a mixture of downward and circular strokes. The addition of acid-washed sand may enhance the rupturing of tissue.
2. If using Polytron® homogenizer, blend the tissue in a Perspex chamber with homogenization medium at 50 % full speed with a succession of 1-s bursts. We generally find that three to five bursts give the appropriate amount of homogenization. Too many bursts will increase yield, but may also decrease the quality of the isolated mitochondria.

3.2.2 Filtering and Differential Centrifugation

Before mitochondria can be enriched by differential centrifugation, particulate matter and cell debris need to be removed by filtration. A wide range of fabrics are employed for filtration; we routinely use between two and four layers of Miracloth (see Note 16). For centrifugation any centrifuge capable of the desired relative centrifugal force with a fixed-angle rotor with sufficient capacity (~400 mL) for the homogenate will suffice. We generally use a Beckman Coulter J-26XP centrifuge with a Beckman Coulter JA-25.50 rotor (see Note 17). It is important to conduct these initial filtration and centrifugation steps as quickly as possible, to reduce the time that mitochondria are exposed to the damaging vacuole compounds released during homogenization and present in the initial lysate.

1. The homogenate is filtered through four layers of Miracloth via a funnel into conical flask. The speed and yield of this process are increased by gently wringing the cloth into the funnel. This is best performed in a 4 °C cold room.
2. The filtrated homogenate is transferred into 50 mL centrifuge tubes, and centrifuged in a precooled rotor for 5 min at 1,500 $\times g$ at 4 °C.

3. The supernatant is gently transferred into another set of centrifuge tubes without disturbing the pellet containing starch, nuclei, and cell debris. The supernatant is centrifuged at $21,000 \times g$ for 20 min at 4 °C.
4. Discard supernatant. The pellet containing crude mitochondria is resuspended in ~4.0 mL of wash solution with the aid of a clean soft-bristle artist's paintbrush. The resuspended organelles are transferred to 50 mL centrifuge tubes, the volume adjusted to 40 mL with more WS, and centrifuged at $1,500 \times g$ for 5 min at 4 °C.
5. The supernatants are transferred into another set of tubes and centrifuged at $21,000 \times g$ for 20 min at 4 °C. The resulting pellet containing crude organelles can be uniformly resuspended in a small volume (~2.0 mL) of WS with the aid of a clean soft-bristle paintbrush.

3.2.3 Density Gradient Centrifugation

This crude mitochondrial fraction is still heavily contaminated and remains unsuitable for proteomics. The mitochondria are further enriched using step or continuous gradient centrifugation (Fig. 1). We employ Percoll® step gradients for non-photosynthetic tissues and Percoll® continuous gradients with 0–4.4 % (v/v) PVP-40 for photosynthetic tissues. Once the mitochondrial fraction has been collected from the gradient, two wash steps are required to dilute the concentrations of Percoll® and PVP by at least tenfold. This is followed by a second self-forming 28 % Percoll® gradient and washes to further purify mitochondria.

1. The washed mitochondria are layered over 35 mL of the chosen gradient solution in a 50 mL centrifuge tube (see Note 18) and then centrifuged at $\sim 40,000 \times g$ for 45 min at 4 °C with the brake off during deceleration (see Note 19).
2. After centrifugation, the less dense layers of the gradient containing broken thylakoid membranes are discarded by aspiration. The mitochondria should form a yellow-brown band toward the bottom of the tube. Collect this band with a Pasteur pipette, taking care not to collect the very dense matter pooled at the bottom of the tube. The collected mitochondria are diluted with at least 4 volumes of WS without BSA, and centrifuged at $21,000 \times g$ for 15 min at 4 °C in 50 mL tubes.
3. The resulting loose pellet is resuspended in WS without BSA and centrifuged again at $21,000 \times g$ for 15 min. The mitochondrial pellet is resuspended in a small volume of WS (~1 mL) without BSA.
4. This pellet is then layered on top of a 28 % (v/v) Percoll® solution for a second gradient step in 50 mL centrifuge tube and then centrifuged at $\sim 40,000 \times g$ for 30 min at 4 °C with the brake off during deceleration (see Note 19).

- After centrifugation, the mitochondria should form a yellow-brown band toward the top of the tube. Collect this band with a Pasteur pipette and dilute with at least 4 volumes of WS without BSA, and centrifuge at $21,000 \times g$ for 15 min at 4 °C in 50 mL tubes.
- The resulting loose pellet is resuspended in WS without BSA and centrifuged again at $21,000 \times g$ for 15 min. The mitochondrial pellet is resuspended in a small volume of WS (~1 mL) without BSA.
- This will yield a concentration of ~5–20 mg mitochondrial protein/mL (see Note 20).

3.3 Determining Mitochondria Quality

Respiratory rates of isolated mitochondria can be measured as substrate-dependent oxygen consumption using an oxygen electrode. The dependence of respiration rate upon added ADP can be used to calculate the respiratory control ratio that gives an indication of the extent to which the mitochondria are coupled/intact.

- Set up the electrode according to the manufacturer's instructions using 50 % (w/v) saturated KCl as an electrolyte and calibrate between air-saturated water (253 nmol O₂/mL at 25 °C) and zero (established by adding a few crystals of sodium hydrosulfite to the water) (see Note 21).
- The assay should be conducted at 25 °C. To 1 mL of mitochondrial reaction medium add 5–20 µL (~100 µg of mitochondrial protein) of mitochondrial suspension (see Note 22). Determine the background rate of oxygen consumption [rate (a)]. Add one of the following substrate combinations (all concentrations are final), 1 mM NADH; 10 mM succinate and 0.5 mM ATP; 5 mM pyruvate, 0.5 mM malate, 0.3 mM NAD⁺, 0.1 mM thiamine pyrophosphate and 12 µM CoA and determine the rate of oxygen consumption. Add 1 µL of 100 mM ADP. The rate of oxygen consumption should increase: measure the initial linear rate [rate (b)]. After a few minutes the ADP will be depleted and the rate of oxygen consumption will be reduced to [rate (c)]. The respiratory control ratio is given by

$$\text{RCR} = \frac{[\text{rate (c)} - \text{rate (a)}]}{[\text{rate (b)} - \text{rate (a)}]}$$

3.4 Determining Mitochondrial Purity

During the isolation of mitochondria most contamination occurs as co-purification of similarly dense chloroplasts and peroxisomes. By profiling the activity of certain enzymes that are known to be localized in these other cellular compartments, we can determine the purity of isolated mitochondria.

- 3.4.1 Marker Enzyme Assays of Mitochondria**
- Fumarase
1. Add mitochondrial protein sample (10–100 µg protein) to 900 µL reaction master mix.
 2. Add malate to start reaction to a final concentration of 50 mM.
 3. Progression of the reaction is measured directly at 340 nM ($\epsilon=2.55\text{ mM}^{-1}$).
- Aconitase
1. Add mitochondrial protein sample (10–100 µg protein) to 900 µL reaction master mix.
 2. Add aconitate to start reaction to a final concentration of 8 mM (see Note 23).
 3. Progression of the reaction is measured as NADP reduction to NADPH at 340 nM ($\epsilon=6.22\text{ mM}^{-1}$).
- 3.4.2 Marker Enzyme Assay of Chloroplast Contamination**
- Phosphoribulokinase
1. Pipette 0.5 mL 2× reaction buffer and mitochondrial protein sample (10–100 µg protein).
 2. 2× reaction buffer: 200 mM Tris–HCl (pH 7.8), 20 mM MgCl₂, 40 mM KCl, and 20 mM DTT.
 3. Add 2 mM ATP, 2 mM phosphoenol pyruvate, 0.2 mM NADH, 3.5 U/mL pyruvate kinase, and 5 U/mL lactate dehydrogenase.
 4. Add H₂O to 975 µL and mix.
 5. Monitor the $A_{340\text{nm}}$ until constant.
 6. Perform a baseline correction.
 7. Add 25 µL 20 mM ribulose-5-phosphate (Ru5P) (final concentration 0.5 mM) to initiate the reaction.
 8. Activities are assayed by following the oxidation of NADH at 340 nm ($\epsilon=6.22\text{ mM}^{-1}$).
- 3.4.3 Marker Enzyme Assay of Peroxisome Contamination**
- Catalase
1. Catalase activity can be measured in an oxygen electrode, by monitoring the production of oxygen after the addition of H₂O₂.
 2. Set up the electrode according to the manufacturer's instructions using 50 % (w/v) saturated KCl as an electrolyte and calibrate between air-saturated water (253 nmol O₂/mL at 25 °C) and zero (established by adding a few crystals of sodium hydrosulfite to the water) (see Note 21).
 3. The assay should be conducted at 25 °C. To 1 mL of mitochondrial reaction medium add 5–20 µL of mitochondrial suspension (see Note 22).
 4. Add 4 µL of 1 % (v/v) H₂O₂ and monitor oxygen evolution. Rates of oxygen evolution should be calculated as nmol O₂/min mg protein. Typical rate for highly purified peroxisomes is ~14 µmol O₂/min mg protein.

Western blotting of the sample and probing available antibodies of known subcellular localized proteins such as RubisCO as a chloroplast marker, KAT2 as a peroxisome marker, and porin as a mitochondrial marker can also assess the purity of isolated mitochondria.

3.4.4 2D-IEF-SDS

Differential In-Gel Electrophoresis of Purified Mitochondrial Proteins

2D-IEF-SDS-DIGE can be used to compare the purity of mitochondrial fractions by identifying the proteins that differ in abundance between two samples that were prepared using different isolation procedures. We have previously used 2D-IEF-SDS-DIGE to show that FFE purification of mitochondria results in lower abundance of proteins from contaminating organelles (Fig. 2, [9, 28]). Compared to other methods of assessing mitochondrial purity such as Western blot, the main advantage of 2D-IEF-SDS-DIGE is that a far higher number of proteins can be quantified in a single experiment (Fig. 2).

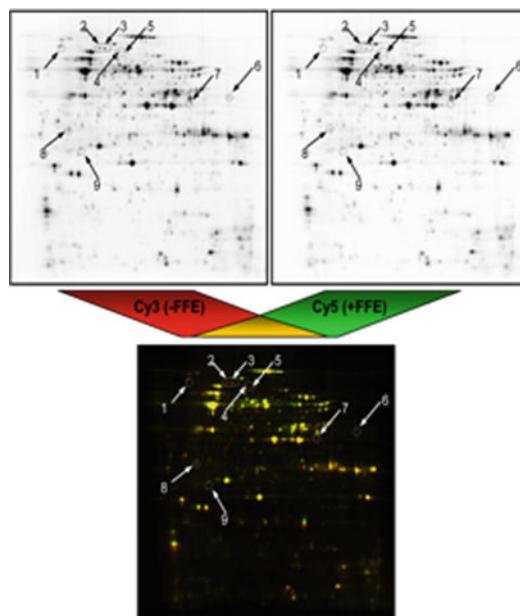


Fig. 2 Example of 2D-IEF-SDS-DIGE of isolated rice mitochondria before and after free flow electrophoresis purification to determine the degree of contamination [28]. Samples before FFE treatment (−FFE; labelled with Cy3, shown in red) and after FFE treatment (+FFE; labelled with Cy5, shown in green) were compared. The *top panels* are gel images of each fluorescence signal, and the *bottom panel* is a combined fluorescence image electronically overlaid using ImageQuant TL software (GE Healthcare Life Sciences). *Yellow spots* represent proteins of equal abundance before and after FFE purification. Spots that are more abundant in samples before FFE purification are *red*, and those more abundant in samples after FFE purification are *green*. The *numbered arrows* indicate proteins with statistically significantly decreased abundance after FFE purification ($n=3$, $P>0.05$), which were chloroplast or peroxisome proteins when identified by mass spectrometry

Mitochondrial Precipitation and Resuspension

1. Purified mitochondria (50 µg) are acetone precipitated by adding 9× the sample volume of cold acetone and placing them at -20 °C overnight or at -80 °C for 2 h. This step removes salts, lipids, chlorophyll, and cellular components that could potentially interfere with CyDye labelling and the IEF step.
2. The precipitated sample is spun at 20,000×*g* for 15 min at 4 °C. The resulting supernatant is discarded and the pellets dried at room temperature for 15 min. The pellets (mitochondrial proteins) are then resuspended in 10 µL of lysis buffer 1 at a protein concentration ~5 µg/µL (see Note 24).
3. The resuspended samples are spun at 20,000×*g* for 10 min at 4 °C to remove any insoluble material. Transfer the supernatant to a new tube.

CyDye Labelling

When running a DIGE gel, it is recommended to prepare three samples: the two samples undergoing direct comparison are labelled with either Cy3 or Cy5, while a pooled mixture of all the samples in the experiment (internal standard) is labelled with Cy2. When running multiple biological replicates of the same sample, it is important that dye allocation is swapped between replicates. Routinely we use minimal labelling, wherein CyDyes bind to ~1–2 % of all Arg or Lys residues in sample. Samples are incubated with the appropriate label on ice for 30 min, and the excess dye is then quenched by the addition of 1 mM lysine chloride (see Note 25).

1. Dilute CyDyes to 400 pmol/µL in water-free DMF.
2. Add 1 µL of the diluted CyDye solution to each of the three samples, vortex, and spin down. Incubate on ice and in the dark for 30 min.
3. Add 1 µL of a freshly prepared 10 mM lysine solution to each sample to stop the labelling reaction. Vortex and spin down. Incubate on ice for 10 min.
4. Add 12 µL of lysis buffer 2. Vortex and spin down.

Isoelectric Focusing

Samples are pooled and diluted in rehydration buffer, which contains ampholytes to enable IEF (see Note 26). The IEF parameters are based on the use of an Ettan IPGphor 3 electrophoresis unit and Immobiline DryStrips 3–10 NL 240 mm (GE Healthcare Life Sciences).

1. Pool the different samples of the set you want to compare. Add rehydration buffer to yield a final volume suitable for the IPG strips you intent to use (450 µL for 240 mm strips). Vortex and spin at 20,000×*g* for 20 min. Transfer the supernatant to a new tube, discarding insoluble pellet.
2. Distribute the sample evenly with the well of the ceramic strip holder and cover with an IPG strip with “gel side” facing downwards. Completely cover the strip with 1 mL of IPG cover fluid, and place the lid onto the ceramic strip holder.

3. Place the ceramic strip holder on the IPGphor system aligned in the direction indicated by anode and cathode regions. Run the IEF parameters for 24 h run as follows: 30 V for 12 h (step), 500 V for 1 h (step), 1,000 V for 2 h (gradient), 8,000 V for 2 h (gradient), and 8,000 V for 6 h (step).

Transferring and Setting IEF Strips onto SDS-PAGE

This step involves reducing and alkylating the strip-bound proteins, and transferring the IPG strip onto a preprepared 12 % (w/v) acrylamide gel without a stacking gel encased in low-fluorescent glass plates (*see Note 27*).

1. Incubate the IPG strip in equilibration buffer containing 65 mM DTT for 15 min in the dark with gentle rocking (*see Note 28*).
2. Incubate the IPG strip in equilibration buffer containing 135 mM IAA for 15 min in the dark with gentle rocking.
3. Rinse the strip for 5–10 s in 1.5 M Tris–HCl pH 8.8 containing 1 % (w/v) SDS.
4. Place the strip on the top of a 12 % (w/v) acrylamide gel, and then add about 5 mL of warm IEF fixing solution to secure the strip (*see Note 29*).
5. Once the IEF fixing solution has solidified, assemble the gel apparatus and run the gels (*see Notes 30 and 31*).

Gel Scanning and Software Analysis

Once SDS-PAGE step is completed, the different proteomes under comparison are visualized by scanning the gel at three wavelengths (633, 532, and 488 nm) with a fluorescent scanner according to the manufacturer's instructions. The generated files are analyzed by commercial software packages such as DeCyder 2D Software (GE Healthcare Life Sciences) and DeCodon Software (DECODON). Generally the literature [12, 28, 45, 50] considers a protein spot to differ significantly between the two samples if the abundance differs by greater than 1.5-fold with a *P*-value of less than 0.05 following Student's *t*-test.

Extraction and Digestion of Proteins That Differ in Abundance on DIGE

The DIGE analysis reveals which protein spots display differing abundance, but peptide mass spectrometry is required to identify these proteins. As CyDye-labelled proteins cannot be seen with the naked eye and DIGE gels are relatively of low abundance it is advantageous to use preparative gels for mass spectrometry identification (*see Note 32*).

1. Match the spots of interest carefully from the DIGE gel with its corresponding spot on the preparative gel, excise the spots, and place them into the wells of the 96-well plate.
2. To each well add 50 μ L of destain solution to each excised gel plug, cover the 96-well plates, and shake on orbital rocker at 700 rpm for 30–45 min.

3. Remove destain solution and discard, and add another 50 μ L of detaining solution. Cover the 96-well plate and shake on orbital rocker at 700 rpm for 30–45 min.
4. Remove destain solution and discard. Dry gel pieces at 50 °C for 20 min with the lid open (*see Note 33*).
5. Add 15 μ L of digestion solution to each dried gel spot (visually check to ensure that each gel spot is immersed in solution). Cover and incubate at 37 °C for 12–16 h for digestion.
6. Add 10–15 μ L 100 % (v/v) acetonitrile to each well and cover and shake on orbital rocker at 750 rpm for 15 min.
7. Take supernatant from each sample and place into a new 96-well plate.
8. Add 10–15 μ L of extraction solution to gel pieces. Cover and shake on orbital rocker at 750 rpm for 15 min.
9. Take supernatant from each sample and place into the same well of the 96-well plate from **step 7** above.
10. Repeat **steps 4** and **5**.
11. Dry down samples in the new plate using a vacuum centrifuge.

Identification of Proteins That Differ in Abundance Following DIGE

The identification of proteins that differ in abundance following the isolation of mitochondria by two differing isolation techniques allows us to determine the degree of contamination from a number of likely sources including other organelles or cellular compartments. Recently we showed that the top six sources of contaminating protein of mitochondria following two-density gradient isolation were the chloroplast (45 % of contaminating proteins), plasma membrane (35 % of contaminating proteins), cytoplasm (12 % of contaminating proteins), vacuole (5 % of contaminating proteins), extracellular proteins (2 % of contaminating proteins), and peroxisome (1 % of contaminating proteins) [24]. Although protein tandem mass spectrometry is a highly variable process depending on both the type of ionization and configuration of mass spectrometry used, we have outlined a general procedure for both MALDI-MS/MS and ESI-MS/MS to the point of sample ionization following which machine-specific parameters/knowledge are required for each hardware setup.

A. Method using MALDI-MS/MS.

Peptide samples derived from trypsin digestion of excised gel spots to be analyzed by MALDI-MS/MS are “spotted” onto a MALDI plate with an appropriate matrix using a dried droplet method.

1. Appropriate machine-specific calibration standards should always be spotted close to the samples being analyzed.

2. The extracted peptides are resuspended in 1–2 μ L 5 % (v/v) acetonitrile and 0.1 % (v/v) TFA.
3. Pipette 1–2 μ L of resuspended peptides on the MALDI plate and let dry until the volume has reduced by ~50 %.
4. Add 1–2 μ L of matrix solution on the sample and mix (*see Note 34*).
5. Let the spot dry.
6. The sample plate is then placed in the mass spectrometer and MS and MS/MS spectra obtained.
7. The resulting MS and MS/MS spectra can then be interpreted by standard proteomics approaches using software packages such as Mascot (Matrix Sciences), Seaquest (Thermo Scientific), or X!tandem (www.thegpm.org/tandem/).

B. Method using ESI-MS/MS.

Peptide samples derived from trypsin digestion of excised gel spots to be analyzed by ESI-MS/MS require fractionation to allow separation of the various peptides in time to allow identification. This is crucial to allow the assessment of peptide abundance (MS mode), selection of a precursor peptide (MS mode), and its fragmentation and detection of fragment ions (MS/MS mode). If peptides are not fractionated then the duty cycle of the mass spec would not allow the identification of sufficient peptides to confirm identification. This fractionation is typically achieved by reverse-phase (RP)-HPLC using a C18 column (*see Note 14*).

1. The extracted peptides are resuspended in 10 μ L 5 % (v/v) acetonitrile and 0.1 % (v/v) formic acid.
2. The resuspended sample is loaded into the HPLC flow (5 % (v/v) acetonitrile, 0.1 % (v/v) formic acid) prior to the in-line C18 column by direct injection by use of an HPLC sampler (*see Note 35*).
3. Once the sample has bound the C18 column a gradient of 5 % (v/v) acetonitrile and 0.1 % (v/v) formic acid to 60 % (v/v) acetonitrile and 0.1 % (v/v) formic acid is run to sequentially elute bound peptides directly into the mass spectrometer and MS and MS/MS spectra collected.
4. The resulting MS and MS/MS spectra can then be interpreted by standard proteomics approaches using software packages such as Mascot (Matrix Sciences), Seaquest (Thermo Scientific), or X!tandem (www.thegpm.org/tandem/).
5. The column is then washed briefly at 80 % (v/v) acetonitrile and 0.1 % (v/v) formic acid and re-equilibrated with 5 % (v/v) acetonitrile and 0.1 % (v/v) formic acid prior to the next sample.

3.4.5 Gel-Free Peptide Fractionation and Identification to Determine Mitochondrial Purity

Increasingly peptide-level separations of digested protein extracts are becoming the preferred means of sample fractionation in proteomics. These approaches are of higher throughput, overcome many of the physiochemical biases inherent in gel-based approaches, and typically identify many more proteins including lower abundance proteins that are not stained on gels. However, they lack visual representation of the sample achieved by gel separation and the knowledge gained of a protein's Mr and isoelectric point, which can add an additional level of confidence to the protein identifications obtained. Gel-free techniques can also be used to determine the mitochondrial or non-mitochondrial origin of a protein, by harnessing the high-throughput capability of direct RP-HPLC-ESI-MS/MS. To do this we produce triplicate samples of low-quality, mid-quality, and high-quality mitochondrial isolations. Typically these have been collected pre gradients (for low quality), post gradients and washes (for med quality), and following FFE (for high quality) [9, 17, 51]. These samples are then quantitatively analyzed to determine quantitative enrichment (QE) of mitochondrial proteins and quantitative depletion (QD) of non-mitochondrial proteins in the increasingly purified mitochondrial samples (low → med → high). Thus any proteins seen to increase in abundance when analyzing our increasingly pure mitochondria sample are of mitochondrial origin and any protein seen to decrease in abundance is a contaminant. Here we outline methods for determining the mitochondrial origin of proteins using a quantitative enrichment/depletion (QED) approach.

Sample Preparation

Tryptic digestion and peptide mass spectrometry can be impaired by the presence of salts, lipids, cellular components, and particularly charged species. To remove these mitochondrial proteins are acetone precipitated as outlined in Subheading 3.4.4.

1. Following precipitations add resuspension buffer to the dried pellets (1 μ L per 10 μ g of mitochondrial protein) and incubate at room temperature for 45–60 min with occasional vortexing.
2. Once the sample is in solution add IAA to 10 mM and incubate in the dark for 30 min.
3. Dilute the sample to 1 M urea with dilution buffer.
4. Add digestion solution, and incubate at 37 °C for 12–16 h.

The high concentrations of urea, DTT, IAA, and NH_4HCO_3 in the sample are incompatible with ionization/mass spectrometry. These are removed by a reverse-phase-solid-phase extraction (RP-SPE) step using C18 spin columns or C18 embedded pipette tips.

5. Add 2.5 μ L of FA to the sample.
6. Charge the C18 column (or tip) by loading with 750 μ L (10 μ L, pipette once) of 70 % (v/v) ACN and 0.1 % (v/v) FA and spinning at 150 $\times g$ for 2 min, and discard the flow-through.

7. Equilibrate the C18 column by loading with 750 μ L (10 μ L, pipette twice) of 2 % (v/v) ACN and 0.1 % (v/v) FA and spinning at $150 \times g$ for 3 min, and discard the flow-through.
8. Load the C18 column (or tip, pipette ten times from well) with sample, spin at $150 \times g$ for 3 min, and discard the flow-through.
9. Wash the sample by loading the C18 column with 400 μ L (10 μ L, pipette once) of 2 % (v/v) ACN and 0.1 % (v/v) FA and spinning at $150 \times g$ for 3 min, and discard the flow-through.
10. Repeat the previous step.
11. Elute your sample by loading the C18 column with 750 μ L (2–10 μ L, pipette ten times into well) of 70 % (v/v) ACN and 0.1 % (v/v) FA and spinning at $150 \times g$ for 2 min, and retain the flow-through.
12. Repeat the previous step, and pool the two flow-throughs together.
13. Dry down sample in vacuum centrifuge.

Sample Fractionation and Mass Spectrometry

Complex peptide mixtures such as those derived from trypsin digestion of whole mitochondria can simply be fractionated and analyzed by RP-HPLC-ESI-MS/MS; however to increase the number of proteins identified from the sample other fractionation techniques can be implemented prior to the RP. Two common approaches are strong cation exchange (SCX) where peptides are fractionated by their net surface charge and off-gel electrophoresis (OGE) where peptides are fractionated by their total net charge. Protocols for these approaches will not be covered in this review, but both techniques generate many fractions and are directly amenable to subsequent RP-HPLC-ESI-MS/MS.

1. The extracted peptides are resuspended in 2 μ L 5 % (v/v) acetonitrile and 0.1 % (v/v) formic acid.
2. The resuspended sample is loaded into the HPLC flow (5 % (v/v) acetonitrile, 0.1 % (v/v) formic acid) prior to the in-line C18 column by direct injection by use of an HPLC sampler (*see Notes 15 and 36*).
3. Once the sample has bound the C18 column a gradient of 5 % (v/v) acetonitrile and 0.1 % (v/v) formic acid to 60 % (v/v) acetonitrile and 0.1 % (v/v) formic acid is run to sequentially elute bound peptides directly into the mass spectrometer and MS and MS/MS spectra collected.
4. The resulting MS and MS/MS spectra can then be interpreted by standard proteomics approaches using software packages such as Mascot (Matrix Sciences), Seaquest (Thermo Scientific), or X!tandem (www.thegpm.org/tandem/).
5. The column is then washed briefly at 80 % (v/v) acetonitrile and 0.1 % (v/v) formic acid and re-equilibrated with 5 % (v/v) acetonitrile and 0.1 % (v/v) formic acid prior to the next sample.

Data Analysis

To determine the mitochondrial or non-mitochondrial origin of a protein we use a quantitative measure of protein abundance across the differentially enriched mitochondria isolations. In its simplest form this involves spectral counting, in which the number of peptides identified for a given protein is used as a measure of its abundance. For example if for a particular protein 6 peptides were matched in the low-purity sample, 9 in the medium-purity sample, and 12 in the high-purity sample, then this protein would be confirmed as a mitochondrial protein. As the number of peptides identified for this protein increases in abundance with increasing mitochondrial purity we can confirm its mitochondrial origin. A further extension of this method is a statistical analysis of the changes observed for a particular protein. To do this using standard statistical such as the *t*-test, a number of statistical assumptions must be met. To meet these Zybalov et al. [52] developed the normalized spectral abundance factor (NSAF), which includes a parameter that also compensates for variations in different protein lengths. They showed that the natural log of each NSAF results in a Gaussian distribution of a dataset, permitting analysis using the *t*-tests. However it becomes essential to avoid $\ln(0)$ and thus zero spectral counts need to be replaced by a fractional value empirically derived for each data set analyzed. This aim is to find the smallest value that maintains a normal distribution of the $\ln(\text{NSAF})$ data [52] and replace all zero values with this fractional value. Once these statistical parameters are implemented it becomes possible to determine the statistical significance of a particular protein enrichment/depletion, confirming or refuting its mitochondrial origin. However in some cases, difficulties with abundant mitochondrial proteins can occur where all possible peptides for a protein are found in all three samples and thus no enrichment is observed; fortunately in most cases we have encountered these proteins to be generally well-known mitochondrial proteins that have had their subcellular localization previously confirmed by other techniques.

4 Notes

1. 0.3–0.4 M of osmoticum is required using either sucrose or mannitol to maintain mitochondrial structure by preventing excessive swelling and rupturing of membranes.
2. 2–5 mM of divalent cation chelator (EDTA or EGTA) acts to inhibit the activity of various phospholipases and proteases.
3. 25–50 mM of a basic pH buffer system (MOPS, TES, or Na-pyrophosphate) prevents media acidification during vacuole rupture.

4. 0.2–1 % (w/v) bovine serum albumin (BSA) acts as an accessible sacrificial protein target for released proteases and to scavenge free fatty acids.
5. Reductant (cysteine, ascorbate, dithiothreitol (DTT), or β -mercaptoethanol) is added freshly prior to homogenization to a final concentration of 5–20 mM to protect against oxidants produced or released during homogenization.
6. For green tissues the addition of 1–2 % (w/v) polyvinylpyrrolidone (PVP-40) helps to remove phenolic compounds liberated from the vacuole that can damage organelles in the initial homogenate.
7. The ascorbate (Na L-ascorbate, Mr 198.11) and cysteine (Mr 137.7) are added to the chilled homogenization solution immediately before use, correcting the pH after addition.
8. Available from Kinematica (<http://www.kinematica-inc.com/>).
9. It is recommended to make up only about 10 mL of reaction master mix and use for a series of eight to ten assays immediately.
10. Aliquots of lysis and rehydration buffers are stored indefinitely at -20°C .
11. TFA should be prepared at the concentration of 10 % (v/v) and diluted to the required concentration before use.
12. Trypsin typically comes in aliquots of 25 μg that is resuspended in 250 μL of 0.01 % (v/v) TFA (100 $\mu\text{g}/\text{mL}$ trypsin) and made up to final concentration of 12.5 $\mu\text{g}/\text{mL}$ and 10 mM NH_4HCO_3 using 200 mM NH_4HCO_3 and water.
13. Saturated CHCA matrix solution can be used for 7 days; all other solutions should be made immediately prior to use.
14. Many varieties of commercial columns with many variations in column architecture and stationary phases exist; we typically use homemade Microsorb C18 0.5 mm \times 100 mm, 5 μm , 100 \AA columns.
15. Many varieties of commercial columns with many variations in column architecture and stationary phases exist; we typically use Agilent Zorbax 300SB C18 75 μm \times 150 mm, 5 μm , 300 \AA columns.
16. Available from Merck Millipore (<http://www.merckmillipore.com/>).
17. Available from Beckman Coulter (<http://www.beckmancoulter.com/>).
18. Care must be taken when layering the washed mitochondria on top of the prepared gradient solution to avoid disturbing the gradient. This is best achieved by running the extract down the side of the tube as opposed to adding it drop by drop.

19. If the brake is engaged during deceleration the gradient fractionation achieved during centrifugation will be lost due to sample mixing.
20. We confirm protein concentration at this stage prior to further analysis using the Bradford protein assay [53].
21. The performance of oxygen electrodes deteriorates over time due to electrochemical deposition of chloride and oxide salts on the silver anode. It is therefore necessary to periodically clean the anode using aluminum oxide polishing paste.
22. It is important not to add too much mitochondrial suspension; otherwise the rate of oxygen consumption will exceed the response time of the oxygen electrode. It is therefore advisable to test several different concentrations of mitochondrial suspension and choose one in which the oxygen consumption rate in the presence of substrate gives an appropriate rate.
23. Aconitase activity can be determined through the measurement of isocitrate production from citrate. Isocitrate production rate is measured by activity of an isocitrate- and NADP-dependent enzyme. Aconitase contains a Fe-S center that is readily damaged by H_2O_2 -inhibiting activity of the protein, and the protein itself has been shown to be degraded during prolonged oxidative stress.
24. Full solubilization should occur in around 20 min using a pipette and vortex mixer to disrupt pellet. When using whole-tissue extracts, the resuspension step can be notoriously problematic, but typically mitochondria proteins resuspend relatively easily.
25. The time the CyDyes are exposed to light should be minimized, as the dyes are liable to bleaching.
26. Some proteins are insoluble in this buffer, so we have found that a centrifugation step prior to IEF removes insoluble proteins that cause vertical streaking on the final image.
27. Low-fluorescent glass plates are available from GE Healthcare Life Sciences.
28. We use an orbital rocker set to ~150 rpm, ensuring that the “gel side” of the strip is facing upwards, to prevent bound proteins from rubbing against the base of the container.
29. We use a 0.75 mm thick spacer to press the strip against the top of the gel, and take care to minimize the number of air bubbles between the strip and the gel. It is important to heat the overlay solution to around 50 °C to melt the agarose, but do not overheat, as this might strip the proteins off the IEF strip.
30. We use the Ettan Dalt 6 system, and in our hands, a typical gel requires 270 mA h of electrophoresis. For convenience, we usually run this step overnight (i.e., 15 mA per gel for 18 h).

31. It is important to cover the entire gel apparatus with aluminum foil to minimize the exposure of CyDyes to light.
32. Preparative gels typically contain 500–1,000 µg of protein and are otherwise run identically to the DIGE gels. The resulting gels are stained with colloidal Coomassie blue for visualization.
33. Dried gel pieces are lightweight and can acquire static charge; take care to ensure that gel pieces do not “jump” out of wells.
34. Care must be taken to avoid introducing air bubbles and touching the MALDI plate with the pipette tip.
35. We typically run a 15-min gradient at 10 µL/min.
36. We typically run a 60-min gradient at 300 nL/min.

Acknowledgments

This work was supported by the ARC Centre of Excellence for Plant Energy Biology (CE0561495) and A.H.M. as an ARC Future Fellow. R.P.J. is supported by a Grains Research and Development Corporation (GRDC) PhD scholarship.

References

1. Millar AH, Sweetlove LJ, Giege P et al (2001) Analysis of the *Arabidopsis* mitochondrial proteome. *Plant Physiol* 127:1711–1727
2. Kruft V, Eubel H, Jansch L et al (2001) Proteomic approach to identify novel mitochondrial proteins in *Arabidopsis*. *Plant Physiol* 127:1694–1710
3. Sweetlove LJ, Heazlewood JL, Herald V et al (2002) The impact of oxidative stress on *Arabidopsis* mitochondria. *Plant J* 32:891–904
4. Millar AH, Heazlewood JL (2003) Genomic and proteomic analysis of mitochondrial carrier proteins in *Arabidopsis*. *Plant Physiol* 131:443–453
5. Millar AH, Trend AE, Heazlewood JL (2004) Changes in the mitochondrial proteome during the anoxia to air transition in rice focus around cytochrome-containing respiratory complexes. *J Biol Chem* 279:39471–39478
6. Heazlewood JL, Tonti-Filippini JS, Gout AM et al (2004) Experimental analysis of the *Arabidopsis* mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. *Plant Cell* 16:241–256
7. Brugiere S, Kowalski S, Ferro M et al (2004) The hydrophobic proteome of mitochondrial membranes from *Arabidopsis* cell suspensions. *Phytochemistry* 65:1693–1707
8. Perales M, Eubel H, Heinemeyer J et al (2005) Disruption of a nuclear gene encoding a mitochondrial gamma carbonic anhydrase reduces complex I and supercomplex I+III2 levels and alters mitochondrial physiology in *Arabidopsis*. *J Mol Biol* 350:263–277
9. Eubel H, Lee CP, Kuo J et al (2007) Free-flow electrophoresis for purification of plant mitochondria by surface charge. *Plant J* 52:583–594
10. Winger AM, Taylor NL, Heazlewood JL et al (2007) Identification of intra- and intermolecular disulphide bonding in the plant mitochondrial proteome by diagonal gel electrophoresis. *Proteomics* 7:4158–4170
11. Livaja M, Palmieri MC, von Rad U et al (2008) The effect of the bacterial effector protein harpin on transcriptional profile and mitochondrial proteins of *Arabidopsis thaliana*. *J Proteomics* 71:148–159
12. Lee CP, Eubel H, O'Toole N et al (2008) Heterogeneity of the mitochondrial proteome for photosynthetic and non-photosynthetic *Arabidopsis* metabolism. *Mol Cell Proteomics* 7:1297–1316
13. Koh S, Andre A, Edwards H et al (2005) *Arabidopsis thaliana* subcellular responses to compatible *Erysiphe cichoracearum* infections. *Plant J* 44:516–529

14. Ito J, Taylor NL, Castleden I et al (2009) A survey of the *Arabidopsis thaliana* mitochondrial phosphoproteome. *Proteomics* 9:4229–4240
15. Taylor NL, Tan YF, Jacoby RP et al (2009) Abiotic environmental stress induced changes in the *Arabidopsis thaliana* chloroplast, mitochondrial and peroxisome proteomes. *J Proteomics* 72:367–378
16. Cui J, Liu J, Li Y et al (2011) Integrative identification of *Arabidopsis* mitochondrial proteome and its function exploitation through protein interaction network. *PLoS One* 6:e16022
17. Duncan O, Taylor NL, Carrie C et al (2011) Multiple lines of evidence localize signaling, morphology, and lipid biosynthesis machinery to the mitochondrial outer membrane of *Arabidopsis*. *Plant Physiol* 157:1093–1113
18. Klodmann J, Senkler M, Rode C et al (2011) Defining the protein complex proteome of plant mitochondria. *Plant Physiol* 157:587–598
19. Klodmann J, Sunderhaus S, Nimtz M et al (2010) Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*. *Plant Cell* 22:797–810
20. Lee CP, Eubel H, Millar AH (2010) Diurnal changes in mitochondrial function reveal daily optimisation of light and dark respiratory metabolism in *Arabidopsis*. *Mol Cell Proteomics* 9(10):2125–2139
21. Lee CP, Eubel H, O'Toole N et al (2011) Combining proteomics of root and shoot mitochondria and transcript analysis to define constitutive and variable components in plant mitochondria. *Phytochemistry* 72:1092–1108
22. Skirycz A, Memmi S, De Bodt S et al (2011) A reciprocal ¹⁵N-labeling proteomic analysis of expanding *Arabidopsis* leaves subjected to osmotic stress indicates importance of mitochondria in preserving plastid functions. *J Proteome Res* 10:1018–1029
23. Tan YF, O'Toole N, Taylor NL et al (2010) Divalent metal ions in plant mitochondria and their role in interactions with proteins and oxidative stress-induced damage to respiratory function. *Plant Physiol* 152:747–761
24. Taylor NL, Heazlewood JL, Millar AH (2011) The *Arabidopsis thaliana* 2-D gel mitochondrial proteome: refining the value of reference maps for assessing protein abundance, contaminants and post-translational modifications. *Proteomics* 11:1720–1733
25. Heazlewood JL, Howell KA, Whelan J et al (2003) Towards an analysis of the rice mitochondrial proteome. *Plant Physiol* 132:230–242
26. Kristensen BK, Askerlund P, Bykova NV et al (2004) Identification of oxidised proteins in the matrix of rice leaf mitochondria by immunoprecipitation and two-dimensional liquid chromatography-tandem mass spectrometry. *Phytochemistry* 65:1839–1851
27. Chen X, Wang Y, Li J et al (2009) Mitochondrial proteome during salt stress-induced programmed cell death in rice. *Plant Physiol Biochem* 47:407–415
28. Huang S, Taylor NL, Narvaez R et al (2009) Experimental analysis of the rice mitochondrial proteome, its biogenesis, and heterogeneity. *Plant Physiol* 149:719–734
29. Taylor NL, Howell KA, Heazlewood JL et al (2010) Analysis of the rice mitochondrial carrier family reveals anaerobic accumulation of a basic amino acid carrier involved in arginine metabolism during seed germination. *Plant Physiol* 154:691–704
30. Li CZ, Huang HQ, Yin FY et al (2008) [The effect of *Haynaldia villosa* v chromosome on the mitochondrial proteome of wheat-H. villosa chromosome substitution line and translocation line]. *J Mol Cell Biol* 41:150–154
31. Jacoby RP, Millar AH, Taylor NL (2010) Wheat mitochondrial proteomes provide new links between antioxidant defense and plant salinity tolerance. *J Proteome Res* 9:6595–6604
32. Hochholdinger F, Guo L, Schnable PS (2004) Cytoplasmic regulation of the accumulation of nuclear-encoded proteins in the mitochondrial proteome of maize. *Plant J* 37:199–208
33. Shahpuri A, Svensson B, Finnie C (2009) From proteomics to structural studies of cytosolic/mitochondrial-type thioredoxin systems in barley seeds. *Mol Plant* 2:378–389
34. Bardel J, Louwagie M, Jaquinod M et al (2002) A survey of the plant mitochondrial proteome in relation to development. *Proteomics* 2:880–898
35. Taylor NL, Heazlewood JL, Day DA et al (2005) Differential impact of environmental stresses on the pea mitochondrial proteome. *Mol Cell Proteomics* 4:1122–1133
36. Hoa le TP, Nomura M, Kajiwara H et al (2004) Proteomic analysis on symbiotic differentiation of mitochondria in soybean nodules. *Plant Cell Physiol* 45:300–308
37. Yin G, Sun H, Xin X et al (2009) Mitochondrial damage in the soybean seed axis during imbibition at chilling temperatures. *Plant Cell Physiol* 50:1305–1318
38. Dubinin J, Braun HP, Schmitz U et al (2011) The mitochondrial proteome of the model legume *Medicago truncatula*. *Biochim Biophys Acta* 1814:1658–1668

39. Atteia A, Adrait A, Brugiere S et al (2009) A proteomic survey of *Chlamydomonas reinhardtii* mitochondria sheds new light on the metabolic plasticity of the organelle and on the nature of the alpha-proteobacterial mitochondrial ancestor. *Mol Biol Evol* 26:1533–1548
40. Millar AH, Eubel H, Jansch L et al (2004) Mitochondrial cytochrome c oxidase and succinate dehydrogenase complexes contain plant specific subunits. *Plant Mol Biol* 56:77–90
41. Li L, Nelson CJ, Solheim C et al (2012) Determining degradation and synthesis rates of *Arabidopsis* proteins using the kinetics of progressive ¹⁵N labeling of 2D gel-separated protein spots. *Mol Cell Proteomics* 11(6):M111.010025
42. Giege P, Heazlewood JL, Roessner-Tunali U et al (2003) Enzymes of glycolysis are functionally associated with the mitochondrion in *Arabidopsis* cells. *Plant Cell* 15:2140–2151
43. Graham JW, Williams TC, Morgan M et al (2007) Glycolytic enzymes associate dynamically with mitochondria in response to respiratory demand and support substrate channeling. *Plant Cell* 19:3723–3738
44. Millar AH, Mittova V, Kiddie G et al (2003) Control of ascorbate synthesis by respiration and its implications for stress responses. *Plant Physiol* 133:443–447
45. Shingaki-Wells RN, Huang SB, Taylor NL et al (2011) Differential molecular responses of rice and wheat coleoptiles to anoxia reveal novel metabolic adaptations in amino acid metabolism for tissue tolerance. *Plant Physiol* 156:1706–1724
46. Eubel H, Heazlewood JL, Millar AH (2007) Isolation and subfractionation of plant mitochondria for proteomic analysis. *Methods Mol Biol* 355:49–62
47. Habermann B, Oegema J, Sunyaev S et al (2004) The power and the limitations of cross-species protein identification by mass spectrometry-driven sequence similarity searches. *Mol Cell Proteomics* 3:238–249
48. Eubel H, Braun HP, Millar AH (2005) Blue-native PAGE in plants: a tool in analysis of protein-protein interactions. *Plant Methods* 1:11
49. Heinemeyer J, Scheibe B, Schmitz UK et al (2009) Blue native DIGE as a tool for comparative analyses of protein complexes. *J Proteomics* 72:539–544
50. Jacoby RP, Taylor NL, Millar AH (2011) The role of mitochondrial respiration in salinity tolerance. *Trends Plant Sci* 16:614–623
51. Tan Y-F, Millar AH, Taylor NL (2012) Components of mitochondrial oxidative phosphorylation vary in abundance following exposure to cold and chemical stresses. *J Proteome Res* 11(7):3860–3879
52. Zybalov B, Mosley AL, Sardiu ME et al (2006) Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *J Proteome Res* 5:2339–2347
53. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254

Chapter 35

Separation of the Plant Golgi Apparatus and Endoplasmic Reticulum by Free-Flow Electrophoresis

Harriet T. Parsons, Susana M. González Fernández-Niño, and Joshua L. Heazlewood

Abstract

Free-flow electrophoresis (FFE) is a technique for separation of proteins, peptides, organelles, and cells. With zone electrophoresis (ZE-FFE), organelles are separated according to surface charge. The plant Golgi and endoplasmic reticulum (ER) are similar in density and are therefore separated with difficulty using standard techniques such as density centrifugation. Purification of the ER and Golgi apparatus permits a biochemical and proteomic characterization which can reveal the division of processes between these compartments. Here we describe complete separation between the ER and more negatively charged Golgi compartments using ZE-FFE. We also describe techniques for assigning proteins to partially separated ER and the less negatively charged Golgi compartments.

Key words Free-flow electrophoresis, Golgi apparatus, *Arabidopsis*, Endoplasmic reticulum, Proteomics

1 Introduction

Zone electrophoresis employing a free-flow electrophoresis system (ZE-FFE) has existed for over 50 years. Its potential application to subcellular fractionation can be fully realized when the process is combined with modern mass spectrometry analytical techniques. Its resurgence as a technique across multiple research domains is therefore relatively recent considering its availability for the past half century.

ZE-FFE is typically applied to samples enriched in the subcellular compartment of interest, usually via density centrifugation. This technique has advanced subcellular proteomics in plants through isolation of organelles at appreciably higher purity levels than achieved previously with density centrifugation techniques alone [1, 2]. The separation of mitochondria from peroxisomes [3, 4] and purification of Golgi compartments from a complex background of

contaminants [5] have demonstrated the benefit of surface charge purification of organelles with similar density to other subcellular compartments.

During ZE-FFE, a sample is loaded into a separation chamber in which separation buffers are moving under laminar flow. A voltage is applied across the separation chamber, perpendicular to buffer flow, whilst opposing buffer flow at the top of the chamber directs the separated sample into 96 collection tubes [6]. Extensive parameter optimization is possible when using FFE; multiple variations on the buffer system described here exist in the literature and chamber height (0.5 mm for standard ZE-FFE) can be adjusted by spacers of varying thickness. The diameter of pump tubing delivering buffers can, within certain constraints, be altered. Sample and carrier buffer flow rates are adjustable, as is the voltage. The flexibility and adaptability inherent in this technique have proved essential in enhancing the separation of organelles with similar charge and density such as mitochondria and peroxisomes [3, 4] and the Golgi/other secretory compartments [5]. In the specific case of the endomembrane, although there has been a long history of separating this system using density gradients [7], the limited degree of resolution achieved has necessitated involved analytical techniques to adequately tease apart these integrated membrane systems [8].

Here we describe a protocol for preparation of purified plant Golgi and highly enriched ER samples for proteomic analysis from an endomembrane-enriched sample using ZE-FFE. We also describe the initial pre-FFE enrichment and a brief example of proteomic data processing. These methods were developed using suspension-cultured *Arabidopsis* cells; however the same protocol has been applied to cultured cells of other species and other sources of plant tissue.

2 Materials

Prepare all solutions using ultrapure water (prepared by purifying deionized water to attain a sensitivity of $18 \text{ M}\Omega \text{ cm}$ at 25°C) and analytical grade reagents. Prepare all reagents at room temperature. Perform all centrifugation steps at 4°C . Unless otherwise stated, prepare all buffers the day before and store at 4°C .

2.1 *Arabidopsis* *Protoplast Preparation*

1. Temperature-controlled shaking incubator (23°C , 120 rpm) with constant light ($100 \mu\text{E}$).
2. *Arabidopsis* cell culture medium: 2 % (w/v) sucrose, α -naphthaleneacetic acid (0.5 mg/L), kinetin (0.05 mg/L), 1× Murashige and Skoog basal salt mixture [9]. Prepare media

and adjust to pH 5.7 with potassium hydroxide (KOH), autoclave for 20 min at 121 °C, and store at 4 °C (*see Note 1*).

3. Miracloth (Merck KGaA, Germany).
4. Digestion buffer: 500 mM Mannitol, 5 mM 2-(*N*-morpholino) ethanesulfonic acid (MES), adjust to pH 5.7 with KOH. Store at 4 °C. Just prior to use, add 0.4 % (w/v) Cellulase “Onozuka” R-10 and 0.05 % (w/v) pectolyase Y-23 (Yakult Pharmaceutical, Japan).
5. Variable speed benchtop orbital shaker (*see Note 2*).
6. Large-capacity preparative centrifuge with 2×500 mL tube capacity capable of 800×*g* for sample preparation.

2.2 Protoplast Homogenization and Pre-FFE Fraction

1. Homogenization buffer: 1 % (w/v) dextran (Mw 200,000), 0.4 M sucrose, 10 mM disodium hydrogen phosphate (Na₂HPO₄), 3 mM ethylenediaminetetraacetic acid (EDTA), 0.1 % (w/v) bovine serum albumin (BSA), 5 mM dithiothreitol (DTT) (*see Note 3*), pH to 7.1 with sodium hydroxide (NaOH).
2. 500 mM potassium chloride (KCl), prepared in 0.4 M sucrose, 10 mM Na₂HPO₄, 3 mM EDTA, 1 % (w/v) dextran (Mw 200,000), 0.1 % (w/v) BSA, 5 mM DTT (*see Note 3*), pH 7.1 with NaOH, can store at -20 °C.
3. Glass-Teflon Potter-Elvehjem Tissue Grinders (30–50 mL capacity), cooled on ice.
4. Preparative centrifuge with 4×50 mL tube capacity capable of 5,000×*g* for sample preparation.
5. Light microscope capable of visualizing plant cells and large subcellular structures, minimum 40× objective.
6. Ultracentrifuge and swing-out rotor with 40 mL tube capacity capable of 100,000×*g* for gradients.
7. Gradient buffer 1: 1.6 M Sucrose, 10 mM Na₂HPO₄, 3 mM EDTA, pH 7.1 with NaOH, can store at -20 °C.
8. Gradient buffer 2: 1.0 M sucrose, 10 mM Na₂HPO₄, 3 mM EDTA, dextran Mw 200,000 (1 % w/v), 5 mM DTT (*see Note 3*), pH 7.1 with NaOH, can store at -20 °C.
9. Gradient buffer 3: 0.75 M sucrose, 10 mM Na₂HPO₄, 3 mM EDTA, dextran Mw 200,000 (1 % w/v), 5 mM DTT (*see Note 3*), pH 7.1 with NaOH, can store at -20 °C.
10. Gradient buffer 4: 0.2 M sucrose, 10 mM Na₂HPO₄, 3 mM EDTA, dextran Mw 200,000 (1 % w/v), 5 mM DTT (*see Note 3*), pH 7.1 with NaOH, can store at -20 °C.

2.3 Separation of Enriched Membranes by ZE-FFE

1. FFE system: BD™ FFE System (BD Diagnostics, NJ, USA, for model years 2006–2010) or FFE System (FFE Service GmbH, Germany, <http://www.ffeservice.com/>, for model years 2011–present).
2. Spacers and filters for ZE-FFE including 0.5 mm spacer and 0.8 mm electrode filter papers for ZE-FFE (FFE Service GmbH, Germany).
3. 96-well deep-well plates (2 mL).
4. UV-transparent 96-well plates, e.g., UV-Star (Greiner Bio One, NC, USA).
5. Microplate reader capable of reading absorbance at 280 nm, e.g., Single-Mode Microplate Readers (Molecular Devices, CA, USA).
6. FFE buffer 1: 280 mM sucrose, 10 mM acetic acid, 10 mM triethanolamine, 1 mM EDTA, pH to 7.0 with NaOH (*see Note 4*).
7. FFE buffer 2: 200 mM sucrose, 100 mM acetic acid, 100 mM triethanolamine, 10 mM EDTA, pH to 6.5 with NaOH (*see Note 4*).
8. FFE buffer 3: 100 mM acetic acid, 100 mM triethanolamine, 10 mM EDTA, pH to 6.5 with NaOH (*see Note 4*).
9. Ultracentrifuge and fixed-angle rotor with 10–15 mL tube capacity, capable of $100,000 \times g$ for sample concentration.
10. 10 mM Tris (hydroxymethyl) aminomethane (Tris–HCl), pH 7.5, stored at 4 °C.

2.4 Post-FFE Sample Analysis

1. High-grade trypsin, e.g., Trypsin, from Porcine pancreas (Sigma-Aldrich, MO, USA).
2. SpeedVac concentrator.
3. Ultra-micro SpinColumns with C₁₈ (Harvard Apparatus, MA, USA).
4. ACN1 solution: 80 % acetonitrile (v/v) with 0.1 % trifluoroacetic acid (v/v).
5. ACN2 solution: 2 % acetonitrile (v/v) with 0.1 % trifluoroacetic acid (v/v).
6. Tandem mass spectrometer (MS/MS) with online liquid chromatography (LC) capabilities (nanoflow or capillary flow rates) capable of data-dependent acquisitions.
7. Search engine for analyzing mass spectrometry data to identify proteins, e.g., Mascot (Matrix Science, UK).
8. Proteomic profiling and quantitation software, e.g., Scaffold 3 (Proteome Software, OR, USA).

3 Methods

3.1 Protoplast Preparation of *Arabidopsis* Cell Suspension Cultures

1. Maintain growing *Arabidopsis* cells in 100 mL aliquots, and subculture weekly at 1:10 ratio. Filter 7-day-old cells from approximately ten flasks through Miracloth and firmly squeeze out the remaining culture medium. Weigh out 60–80 g cells if possible; otherwise work with no less than 30 g cells (*see Note 5*).
2. For 80 g cells (fresh weight), use 800 mL of digestion buffer (*see Note 6*). Divide into equal volumes and add 0.4 % (w/v) cellulase and 0.05 % (w/v) pectolyase to 400 mL of digestion buffer in a 4 L wide-bottomed, conical flask. Add cells and ensure that they are adequately suspended in the buffer. Place on an orbital shaker and rotate slowly in the dark (wrapped in foil) for 3 h (*see Note 7*).
3. Protoplasts are harvested by centrifugation at $800 \times g$ for 5 min (rotor with capacity for 250–500 mL), gently resuspended in around 150 mL of the remaining digestion buffer (400 mL, no enzymes), and centrifuged at $800 \times g$ and buffer discarded. Repeating this step two more times ensures complete removal of cellulase and pectolyase from the protoplasts (*see Note 8*).
4. After the final wash in digestion buffer (no enzymes), resuspend the pellet/protoplasts in homogenization buffer using a minimum ratio of 1:1 (w/v) using the original fresh weight of cells to buffer volume.

3.2 Homogenization of Protoplasts and Pre-FFE Enrichment

1. Working at 4 °C, homogenize the protoplasts in homogenization buffer using 4–5 strokes at even pressure with a Potter-Elvehjem homogenizer (*see Note 9*).
2. Compare pre- and post-homogenized protoplasts under a light microscope to ensure adequate disruption.
3. Centrifuge homogenate at $5,000 \times g$ for 15 min (rotor with capacity for 50 mL samples).
4. Transfer the supernatant to an appropriately sized, chilled beaker; slowly add 500 mM KCl drop-wise whilst gently agitating the homogenate, until a final concentration of 50 mM KCl is reached; and then incubate at 4 °C for 5 min.
5. Add 5 mL of gradient buffer 1 to the required number of 40 mL ultracentrifuge tubes. Using a plastic Pasteur pipette, gently layer the homogenate onto the cushion of gradient buffer 1 until all tubes are at least two-thirds full (*see Note 10*).
6. If collecting intact Golgi cisternae use gradient 1 and ultracentrifuge for 1 h at $60,000 \times g$. If collecting Golgi and other secretory membranes use gradient 2 and ultracentrifuge at $100,000 \times g$ (*see Notes 11*).

7. Remove the supernatant such that the cushion surface is disturbed as little as possible (see Note 12). Onto this surface, gently layer 15–20 mL of gradient buffer 2. If enriching for intact Golgi cisternae only, follow this with 8–10 mL of gradient buffer 3 and 5–8 mL of gradient buffer 4 (gradient 1). If enriching for generic secretory membranes, follow gradient buffer 2 with 12–15 mL of gradient buffer 4 (gradient 2), or until tubes are at least two-thirds full (see Note 13).
8. Ultracentrifuge samples at $100,000 \times g$ for 90 min.
9. For gradient 1, two bands should be present. Discard the upper band and remove the lower band with a plastic Pasteur pipette. For gradient 2 a single band should be present (see Note 14).
10. Measure the protein concentration of the collected sample and dilute so that the sucrose concentration is as close to FFE buffer 1 as possible, although a total protein concentration less than 0.75 $\mu\text{g}/\mu\text{L}$ is not recommended. Keep the desired band in a 15 mL tube on ice (see Note 15).

3.3 Separation of Secretory Membranes Using ZE-FFE

1. Ensure that the FFE system is correctly set up for ZE-FFE (Fig. 1), the temperature set to 8 °C, the media pump has been calibrated, and all necessary quality control tests have been undertaken (see Note 16).
2. Load *FFE buffer 1* (inlet tubes 2–6), *FFE buffer 2* (inlets 1 and 7), and *FFE buffer 3* (electrodes) and slowly fill chamber with buffers avoiding air bubbles. Set the media flow rate to 250 mL/h. Switch on electrode circuits; set the voltage, current, and power to 700 V, 150 mA, and 100 W, respectively; and switch on the voltage (see Note 17). Allow the system to equilibrate and stabilize for 20–30 min prior to use.
3. Load the sample at a flow rate of 2,500–3,000 $\mu\text{L}/\text{h}$. The sample should appear homogenous and be visible in the lower half of the separation chamber but is usually too dilute to be seen in the upper half (see Note 18).
4. After about 10 min, start collecting samples in the precooled 2 mL deep-well plates (see Note 19).
5. Throughout the ZE-FEE process, monitor protein distribution by removing 150–250 μL from plates and measuring at 280 nm using UV-transparent 96-well plates. Protein peak positions and profiles should remain constant throughout separation.
6. After the sample has been processed through the FFE, pool fractions of interest from multiple plates and centrifuge at $100,000 \times g$ for 60 min (see Note 20).
7. Remove the supernatant and rinse pellets gently by trickling 4–5 mL of 10 mM Tris–HCl (pH 7.5) down the sides of tubes.

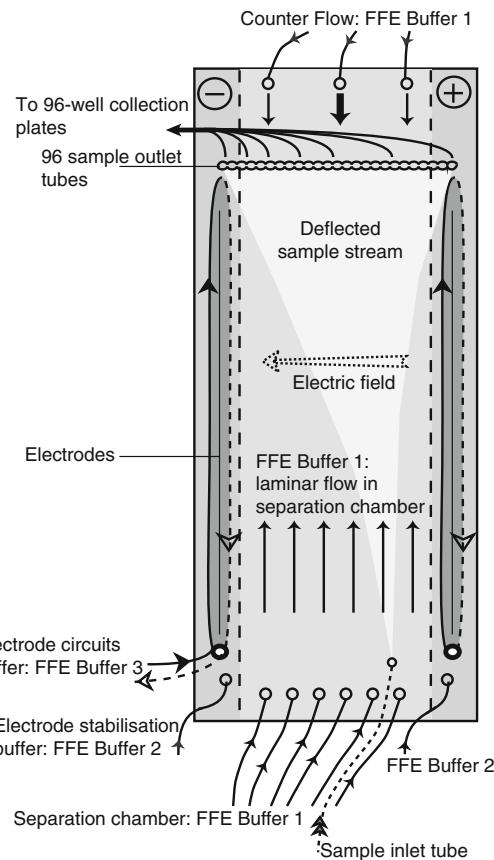


Fig. 1 Schematic diagram of zonal electrophoresis-free flow electrophoresis (ZE-FFE) setup showing the laminar flow direction in the chamber and progressive deflection of the sample in the electric field. In this mode the chamber uses a 0.5 mm spacer and the apparatus is placed horizontally during separation. The electrodes and FFE buffer 3 are separated from the chamber by membranes and filter papers

Resuspend in 20–80 μ L of the 10 mM Tris–HCl (pH 7.5), depending on the pellet size (see Note 21). A final protein concentration of 0.5–1.0 μ g/ μ L is ideal. Store resuspended pellets at –20 °C.

3.4 Tryptic Digestion of Samples and Analysis by Mass Spectrometry

1. Digest at least 10 μ g of protein overnight (37 °C) at a 1:10 trypsin:protein ratio in 40 % (v/v) methanol.
2. Remove methanol in a SpeedVac concentrator until about 1 μ L of sample remains, and then dilute into 25 μ L of the ACN2 solution.
3. Clean and concentrate samples in Ultra-micro SpinColumns (10–25 μ L capacity) after initially hydrating the matrix with H₂O (75 μ L) for 10 min and centrifuge (1,000 \times g, 2 min) as

per the manufacturer's instructions. Wash the SpinColumn with the 50 μ L ACN1 solution, centrifuge (1,000 $\times g$, 2 min), and prime twice with 50 μ L ACN2 solution, centrifuging (1,000 $\times g$, 2 min) after each step.

4. Add peptide sample resuspended in ACN2 solution, centrifuge (1,000 $\times g$, 2 min), and wash twice with 50 μ L ACN2 solution, centrifuging (1,000 $\times g$, 2 min) after each step. Finally elute with 25–50 μ L ACN2 solution by centrifuging at 1,000 $\times g$ for 2 min. Concentrate and remove acetonitrile with a SpeedVac concentrator until 1–5 μ L of peptide solution remains in the tube.
5. Analyze peptide samples by liquid chromatography tandem mass spectrometry (LC-MS/MS) using an automated data-dependent acquisition method optimized for proteomic analyses and characterizations.
6. Data produced by LC-MS/MS of samples can be analyzed using software such as Mascot (Matrix Science, UK) to identify proteins found in each fraction.
7. The proteins identified in each individual fraction or pools of multiple fractions can be compared to proteins associated with published subcellular proteomes, most conveniently done through the *Arabidopsis* SUBcellular database, SUBA [10]. In the example outlined in Fig. 2 we have used marker proteins associated with particular organelles for which multiple experimental assignations were available. Generally, only robustly representative proteins should be included as organelle markers. By monitoring the occurrence of these markers in individual fractions, it is possible to track the relative migration of subcellular compartments (*see Note 22*).
8. Relative quantification of proteins identified in fractions analyzed by LC-MS/MS can be achieved using specific software, e.g., Scaffold 3. In Fig. 2, spectral counts for those proteins deemed to be robust representative markers of the main organelles present in post-ZE-FFE fractions have been aligned, thereby demonstrating subcellular compartment separation.

3.5 Summary and Future Uses of this Technique

Typical data produced and analyzed from FFE-separated endomembrane samples from *Arabidopsis* cell cultures is outlined in Fig. 2. These data demonstrate that the more electronegative Golgi compartments (fractions 24 and under) can be almost completely purified from the ER and other contaminating compartments using this technique. The parameters given here were selected for ER–Golgi separation in a generic endomembrane-enriched sample. Depending on the type of pre-FFE enrichment used (*see Note 23*) and specific adjustment of parameters, considerable scope exists for separation or purification of other compartments. Potential exists

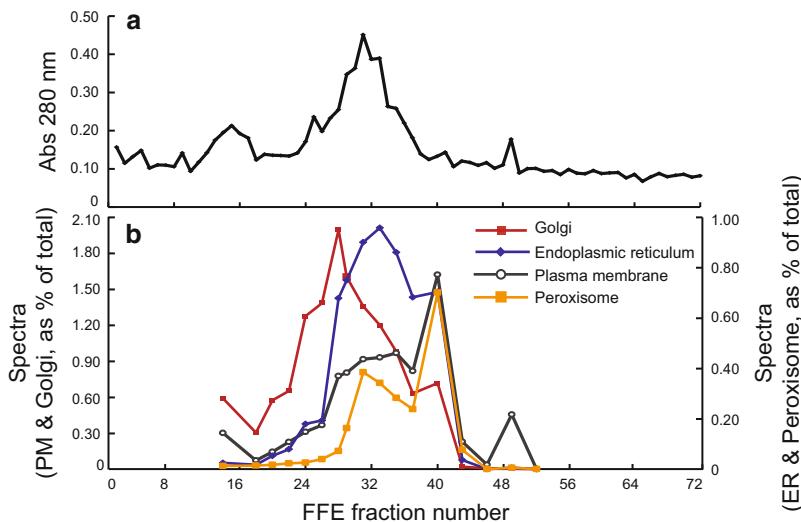


Fig. 2 Distribution of endomembrane fractions after ZE-FFE and analysis by mass spectrometry. **(a)** Total protein after ZE-FFE separation (96-well plate) measured at A_{280} nm from a sample prepared according to outlined procedures. **(b)** The occurrence of relative abundance of organelle marker proteins considered to be robustly representative of particular subcellular compartments in fractions after ZE-FFE. A selection of fractions was analyzed by mass spectrometry and data interrogated against Mascot. Identified proteins were quantified by spectral counting using the Scaffold 3 software. The spectral counts for proteins selected as markers for trafficking and destination compartments (ER, Golgi, PM) were aligned along with contaminating organelles (peroxisome), demonstrating the extent of separation of the ER and Golgi compartments using the techniques and parameters detailed in these methods

for assigning proteins more confidently to compartments that are particularly challenging to separate by coupling FFE, mass spectrometry, and bioinformatics techniques. Such an approach was recently applied to proteins showing a strong pattern of co-migration with known Golgi markers during FFE [5].

4 Notes

1. Media can be dispensed into 250 mL Erlenmeyer flasks as 90 mL aliquots and sealed with foil prior to autoclaving.
2. A variable-speed benchtop orbital shaker with a large orbital throw is optimal (at least 2 cm); this allows slow rotation while maintaining the cells in solution. This permits efficient enzymatic digestion of the cell wall.
3. DTT is added immediately prior to use. A stock solution can be stored at -20°C .
4. All FFE buffers must be used within 24 h and can be stored at 4°C . Triethanolamine should be weighed. A generous volume of FFE buffer 3, e.g., 800 mL for a 1–2-h run, is recommended

so that the electrode buffers are not exhausted. For FFE buffer 1 about 1.5 L is required for a run using the condition described.

5. Cell suspension cultures are ideal for approaches such as subcellular fractionation as a large homogenous starting mass of tissue can be readily obtained. A fine suspension of *Arabidopsis* cells provided the starting material for development of this protocol. However, the same protocol can be applied to cultured cells of other species and other sources of plant tissue. A frequent source of failure when starting with tissue other than cell suspension cultures is insufficient starting material. Keep in mind that yield is highest from suspension-cultured cells, for which 30 g fresh weight after filtering is the minimum workable starting amount for this approach.
6. Maintain approximately this ratio for other starting tissue weights and buffer volumes. For different tissue sources, researchers may want to use a lower homogenization buffer volume: tissue weight and lower FW starting material:enzyme ratios.
7. Rotate at the lowest possible speed at which cells remain in suspension. Enzymes should be added to the digestion buffer immediately before use. Enzymes are easily solubilized by vigorous shaking in a 50 mL aliquot of digestion buffer, prior to their addition to the main solution.
8. The pellet/protoplasts are delicate; decant the supernatant gently to avoid breaking cells.
9. As a guide, the negative pressure on the upstroke should result in an approximate 2 cm space or air bubble between the plunger and the homogenate. For the first few attempts at protoplast disruption, checking the result using a light microscope should reveal rupturing of at least 75 % of the protoplasts. Intact plastids can be used as an indicator for intact Golgi cisternae or other subcellular compartments. The key to this step is ensuring enough mechanical stress to disrupt the protoplasts without destroying subcellular integrity. The number of strokes of the homogenizer should be calibrated using light microscopy.
10. Ideally, use about 20 g fresh weight starting material per gradient. From four gradients expect 8–14 mL of protein at 0.8 $\mu\text{g}/\mu\text{L}$.
11. This step should result in a yellow-colored cushion about 2 mm thick (assuming starting with 60–80 g fresh weight cells).
12. Complete removal of the supernatant is a compromise between the quality of step gradient formation and disturbance of the cushion, with ensuing loss of yield.

13. For Golgi enrichment, it is not recommended to use a sucrose concentration less than 1.0 M (gradient 1). However, if the intention is to enrich for the ER and other secretory compartments at the expense of the Golgi, gradient 2 can be made using a slightly lower concentration (e.g., 0.85–0.95 M gradient buffer 2) and initial spin speed of $100,000 \times g$ rather than $60,000 \times g$ (gradient 1) can be used.
14. Of the two bands present in gradient 1, the lower should be thicker than the upper. A distinctly formed yellow-white band 1–3 mm thick (in both gradients 1 and 2) will give optimized results obtained after FFE.
15. Ideally, **steps 1 and 2** (Subheading 3.2) should be completed as fast as practicably possible. Tubes containing gradient buffer 1 can be prepared in advance. Setup of the FFE and stabilization of the current should have been completed so as to coincide with **step 9**. The current normally takes not more than 20–30 min to stabilize. The conductivity may increase as the sample buffer enters the chamber. This can be rectified by lowering the voltage by, e.g., 20–30 V and waiting for approximately 5 min before starting to collect samples.
16. This is a nontrivial task and some previous experience or instructions in the setup of the system for ZE-FFE has been assumed. Use of fresh filters, membranes, and electrode gaskets can improve separation performance considerably. Great care should be taken that the plastic spacer is exactly centered. Tubing sections under the peristaltic pumps should be checked thoroughly for hairline cracks prior to setup and changed if necessary. The sample inlet tubing should be exchanged if any kinks are present. Excessive indentation from the spacer or the inlet tube on the lower, temperature-cooled plate indicates that a change of plastic cover sheeting is required for optimal separation. All FFE buffers should be made up precisely to the stated pH; otherwise the conductivity will be too high and the current will not stabilize as desired. Much useful material, including references, is available at <http://www.ffeservice.com>.
17. Media flow rate and voltage are the two principal parameters that affect sample migration. For different types of sample preparation a ratio between the two should be optimized; however below a media flow rate of 200 mL/h, diffusion between streams of subcellular compartments may occur. The effect of changing parameters on protein distribution can be instantaneously verified by measuring protein content at 280 nm in UV-transparent 96-well plates.
18. If particulates are present in the sample, pass the sample once through a glass Pasteur pipette. This improves homogeneity and seems to improve separation but the effect on compartment integrity should be checked for each different type of sample.

19. At a flow rate of 250 mL/h, plates will fill in approximately 17 min although shorter fill times will prevent sample warming if a plate-cooling device is not available. Store plates at 4 °C and work at 4 °C for **steps 6** and **7** (Subheading **3.3**).
20. Assuming that eight plates at approximately 1.5 mL per well have been collected, pellets will be visible in all fractions above 0.15 A_{280} (measuring 150 μ L at A_{280}); otherwise it may be necessary to pool fractions. If collecting fractions individually, one fraction from about five plates will fit into a standard 13.5 mL open ultracentrifuge tube. This approach will necessitate multiple rounds of centrifugation for maximum yield from 80 g FW of starting material.
21. A yield of approximately 60–80 μ g of highly purified Golgi membranes split over 2–3 least contaminated fractions can be obtained. Using 2–3 fractions gave over 80 % purity but Golgi membrane yield can be increased substantially by pooling 6–8 fractions in experiments where an approximate 70 % purity is acceptable. Obtaining highly enriched ER using this approach is challenging given the associated migration of contaminating membranes (e.g., mitochondria and plasma membrane) over these same fractions during FFE. Nonetheless FFE migration patterns can be analyzed to distinguish ER and contaminating membrane patterns as was recently undertaken for Golgi proteins [5].
22. Although this is a good method for determining organelle migration, results are dependent upon the number of proteomes available for a particular subcellular compartment, e.g., results for peroxisome migration will be more accurate than those for ER migration. Therefore a low number of representative marker proteins is not always indicative of low abundance. The accuracy of this method is a trade-off between the number of markers used and the quality of markers.
23. Organic solvents are incompatible with the standard tubing used in FFE systems. Salt levels in pre-FFE samples should be minimized as they will interfere with buffer conductivity.

Acknowledgments

This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the US Department of Energy under Contract No. DE-AC02-05CH11231.

References

1. Eubel H, Lee CP, Kuo J et al (2007) Free-flow electrophoresis for purification of plant mitochondria by surface charge. *Plant J* 52: 583–594
2. Zischka H, Weber G, Weber PJA et al (2003) Improved proteome analysis of *Saccharomyces cerevisiae* mitochondria by free-flow electrophoresis. *Proteomics* 3:906–916
3. Eubel H, Meyer EH, Taylor NL et al (2008) Novel proteins, putative membrane transporters, and an integrated metabolic network are revealed by quantitative proteomic analysis of *Arabidopsis* cell culture peroxisomes. *Plant Physiol* 148: 1809–1829
4. Islinger M, Li KW, Loos M et al (2010) Peroxisomes from the heavy mitochondrial fraction: isolation by zonal free flow electrophoresis and quantitative mass spectrometrical characterization. *J Proteome Res* 9:113–124
5. Parsons HT, Christiansen K, Knierim B et al (2012) Isolation and proteomic characterization of the *Arabidopsis* Golgi defines functional and novel targets involved in plant cell wall biosynthesis. *Plant Physiol* 159:12–26
6. Islinger M, Eckerskorn C, Volkl A (2010) Free-flow electrophoresis in the proteomic era: a technique in flux. *Electrophoresis* 31: 1754–1763
7. Mitsui T, Kimura S, Igaue I (1990) Isolation and characterization of Golgi membranes from suspension-cultured cells of rice (*Oryza sativa* L.). *Plant Cell Physiol* 31:15–25
8. Dunkley TPJ, Hester S, Shadforth IP et al (2006) Mapping the *Arabidopsis* organelle proteome. *Proc Natl Acad Sci USA* 103: 6518–6523
9. Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol Plant* 15: 473–497
10. Heazlewood JL, Verboom RE, Tonti-Filippini J et al (2007) SUBA: the *Arabidopsis* subcellular database. *Nucleic Acids Res* 35:D213–D218

Chapter 36

Isolation of Leaf Peroxisomes from *Arabidopsis* for Organelle Proteome Analyses

Sigrun Reumann and Rajneesh Singhal

Abstract

The isolation of cell organelles from model organisms in high purity is important for biochemical analyses of single proteins, entire metabolic pathways, and protein complexes and is absolutely essential for organelle proteome analyses. The efficient enrichment of nearly all cell organelles is more difficult from *Arabidopsis* as compared to traditional model plants and especially challenging for peroxisomes. Leaf peroxisomes are generally very unstable in aqueous solution due to the presence of a single membrane and (para-)crystalline inclusions in the matrix. Leaf peroxisomes from *Arabidopsis* are particularly fragile and, moreover, strongly physically adhere to chloroplasts and mitochondria for largely unknown reasons. Here, we provide a detailed protocol for the isolation of *Arabidopsis* leaf peroxisomes by Percoll followed by sucrose density gradient centrifugation that yields high purity suitable for proteome analyses. Diverse enzymatic and immuno-biochemical methods are summarized to assess purity and intactness.

Key words Peroxisome, *Arabidopsis*, Organelle purity, Proteomics

Abbreviations

| | |
|---------|--|
| CE | Crude extract |
| FW | Fresh weight |
| GB | Grinding buffer |
| HPR | Hydroxypyruvate reductase |
| LP-P1/2 | First/second purified leaf peroxisome fraction |
| TE | Tricine-EDTA |

1 Introduction

Our knowledge of the metabolic and regulatory functions of plant peroxisomes is far from complete because fundamental research is made difficult by (1) high peroxisome fragility, (2) low peroxisome abundance in most tissues and organs, (3) pronounced peroxisome adherence to plastids and mitochondria, particularly in

photosynthetic tissue, and (4) high organ-specific and developmental plasticity in peroxisome functions. In the post-genomic era of plant research, traditional disciplines of physiology, cell biology and biochemistry can efficiently be combined with modern “omics” methods such as genomics, proteomics, and bioinformatics to identify and physiologically characterize unknown proteins, to describe protein-protein interactions in functional protein complexes and to reveal novel metabolic pathways and regulatory functions. However, prerequisite for the application of “omics” methods generally is that model organisms are chosen whose genome has been fully sequenced. Traditionally, plant peroxisomes have been isolated from specific tissue (endosperm, cotyledons), developmental stages and specific species such as spinach, pea, castor bean, and pumpkin [1–4]. The complete genome sequence of *Arabidopsis* has been the first plant genome published [5], predestining this plant species for proteome analyses. Five proteome studies of *Arabidopsis* peroxisomes have been published to date for either leaf peroxisomes [6–8], glyoxysomes [9], or peroxisomes from cell cultures [10] using different methods for organelle purification. Numerous novel metabolic pathways and defense functions indicated by these proteome analyses have been partially or fully validated experimentally thereafter (for review *see* refs. 11, 12). Particular the combination of experimental proteomics with computational prediction of peroxisome-targeted matrix proteins has been proven powerful to uncover the proteome of *Arabidopsis* peroxisomes [13, 14].

Here, we present a detailed protocol for the isolation of *Arabidopsis* leaf peroxisomes from photosynthetic tissue, which is largely based on the method described previously [7, 8] with minor improvements. To obtain relatively high yield and purity, this protocol needs to be followed tightly. Even for an experienced scientist, peroxisome purity and yield differ significantly between different isolations, depending, for instance, on the quality of plant material, carefulness and isolation speed and day-to-day variations. It is recommended to analyze the purity of several leaf peroxisome isolates separately by enzymatic or immune-biochemical methods biochemical to select and pool those isolates of highest purity for proteome analyses by either standard 1D or 2D gel electrophoresis combined with mass spectrometry.

2 Materials

2.1 Components for Growth of *Arabidopsis thaliana* Col-0

1. *Arabidopsis* seeds stored at room temperature in 1.5 mL microfuge tubes.
2. Commercial plant soil mixed with perlite (ratio roughly 3:1).
3. Plastic trays and square pots (ca. 6 cm × 6 cm).
4. Fertilizer if required.

2.2 Components for Leaf Peroxisome Isolation

The components and solutions given below are sufficient for one isolation of Arabidopsis leaf peroxisomes, which requires eight Percoll density gradients (*see Notes 1–3*).

1. Grinding Buffer (GB, 120 mL): 170 mM Tricine, pH 7.5, 1.0 M sucrose, 2 mM EDTA, 1 % (w/v) BSA, 10 mM KCl, 1 mM MgCl₂, 0.5 % (w/v) PVP-40, 5 mM DTT. DTT should be added freshly on the preparation day. Prior to use, the GB is supplemented with protease inhibitors (final concentrations: 0.1 mM PMSF, 0.2 mM Benzamidine, 0.2 mM ϵ -Aminocaproic acid) (*see Note 4*).
2. TE buffer: 20 mM Tricine-KOH, pH 7.5, 1 mM EDTA (*see Note 3*).
3. 15 % (v/v) Percoll solution (30 mL): 4.5 mL Percoll, 750 mM sucrose, 0.2 % (w/v) BSA in 150 mL TE buffer (*see Notes 3, 5 and 6*). For one leaf peroxisome isolation 24 mL (8 \times 3 mL) is required.
4. 38 % (v/v) Percoll solution (100 mL): 38 mL Percoll 750 mM sucrose, 0.2 % BSA in TE buffer (*see Notes 3, 6 and 8*). For one leaf peroxisome isolation 88 mL (8 \times 11 mL) is required.
5. 36 % (w/w) sucrose in TE buffer (ca. 150 mL, ca. 1.22 M sucrose): ca. 62.6 g sucrose in TE buffer. The exact concentration of 36 % (w/w) is adjusted using a refractometer (*see Notes 3, 7 and 8*).
6. 38 % (v/v) Percoll: 36 % (w/w) sucrose = 2:1 (ca. 21 mL): Mix two volumes of 38 % (v/v) Percoll (e.g., 14 mL) with one volume of 36 % (w/w) sucrose (e.g., 7 mL). For one leaf peroxisome isolation 16 mL (8 \times 2 mL) is required.
7. 38 % (v/v) Percoll: 36 % (w/w) sucrose = 1:2 (ca. 21 mL): Mix one volume of 38 % (v/v) Percoll (e.g., 7 mL) with two volumes of 36 % (w/w) sucrose (e.g., 14 mL). For one leaf peroxisome isolation 16 mL (8 \times 2 mL) is required.
8. 60 % (w/w) sucrose in TE buffer (ca. 50 mL ca. 2.26 M): 38.6 g sucrose in 50 mL TE buffer (*see Notes 3 and 7*).
9. Percoll density gradient (1st gradient, Fig. 1): Assemble eight discontinuous Percoll density gradients in SS34 tubes in the following manner (*see Note 9*).
 - (a) Fill 11 mL 38 % (v/v) Percoll solution into an empty SS34 tube.
 - (b) Underlay with 2 mL 38 % (v/v) Percoll: 36 % (w/w) sucrose = 2:1.
 - (c) Underlay with 2 mL 38 % (v/v) Percoll: 36 % (w/w) sucrose = 1:2.
 - (d) Underlay with 3 mL 36 % (w/w) sucrose in TE buffer.
 - (e) Overlay with 3 mL 15 % (v/v) Percoll solution.

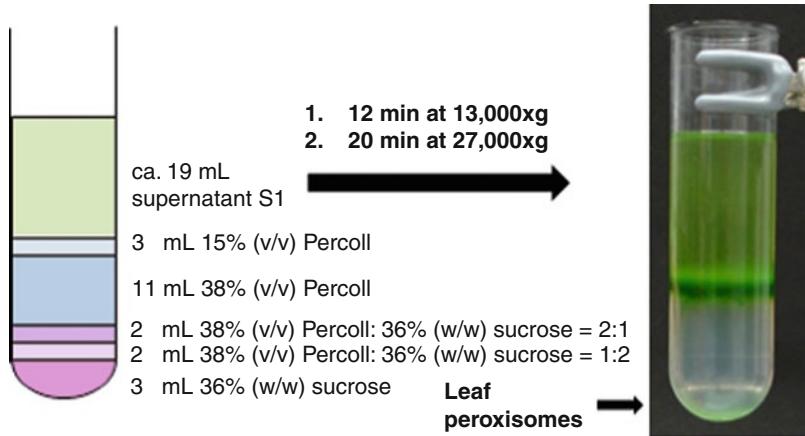


Fig. 1 Leaf peroxisome enrichment from *Arabidopsis* by Percoll density gradient centrifugation (first density gradient). The gradient consists of two different Percoll fractions (15 and 38 % (v/v), in 0.75 M sucrose, *top*), a 36 % (w/w) sucrose cushion (*bottom*) and two intermediate 2-mL fractions comprising mixtures of 38 % (v/v) Percoll and 36 % (w/w) sucrose solution at ratios of 2:1 and 1:2. After centrifugation (12 min at $13,000 \times g$ and 20 min at $27,000 \times g$) the leaf peroxisomes are located at the *bottom* of the centrifugation tube. The figure has been adapted and reprinted from Reumann et al. [7] Suppl. Figure 1

10. Prepare the solutions for the sucrose density gradient in the described manner (Table 1) using the 60 % (w/w) and 36 % (w/w) stock solutions.
11. Sucrose density gradient (2nd gradient): Assemble two discontinuous sucrose density gradients in two ultracentrifuge tubes (e.g., Beckmann SW41 Ti, 13 mL) in the following manner (Fig. 2, *see Note 10*):
 - (a) Fill 1.0 mL 55.2 % (w/w) sucrose (all sucrose solutions in TE buffer, *see Table 1*) into an empty tube.
 - (b) Underlay with 1.0 mL 60 % (w/w) sucrose.
 - (c) Overlay both fractions with 0.5 mL 50.5 % (w/w) sucrose.
 - (d) Overlay with 2.0 mL 48.5 % (w/w).
 - (e) Overlay with 1.0 mL 46.0 % (w/w).
 - (f) Overlay with 1.0 mL 43.7 % (w/w).
 - (g) Overlay with 1.0 mL 41.2 % (w/w).

Mark the fraction interfaces with a water-resistant pen. One gradient is needed for the partially purified leaf peroxisome fraction, while the second gradient is used as a counterbalance. Store on ice until use.
12. Ice buckets containing ice.
13. Miraclot (Calbiochem Ltd., Nottingham, UK).

Table 1
Preparation of different sucrose density gradient fractions from stock solutions

| Final sucrose conc. (% (w/w)) | Final volume (mL) | Volume of 60 % (w/w) sucrose sol. (mL) | Volume of 36 % (w/w) sucrose sol. (mL) |
|-------------------------------|-------------------|--|--|
| 41.2 | 50 | 10 | 40 |
| 43.7 | 50 | 15 | 35 |
| 46.0 | 50 | 20 | 30 |
| 48.5 | 50 | 25 | 25 |
| 50.5 | 50 | 35 | 15 |
| 55.2 | 50 | 40 | 10 |

From two sucrose stock solutions of 60 % (w/w) and 36 % (w/w) different working solutions required for the sucrose density gradient for Arabidopsis leaf peroxisome isolation are prepared in 50-mL Falcon tubes and stored in the refrigerator. Note that mixing of the two stock solutions in the given proportions only roughly yields the desired sucrose concentration of the working solutions. The precise sucrose concentration needs to be adjusted by addition of a small volume of appropriate stock solution using a refractometer

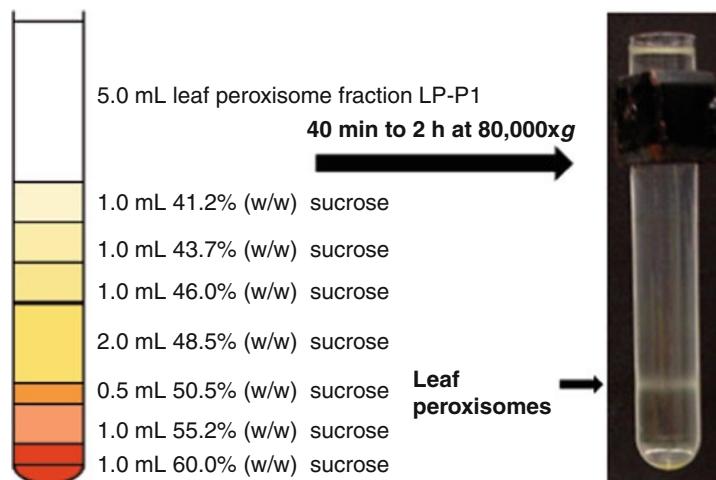


Fig. 2 Leaf peroxisome enrichment from *Arabidopsis* by sucrose density gradient centrifugation (second density gradient). The gradient consists of seven different sucrose fractions (41.2 % *top* to 60.0 % (w/w) *bottom*). After centrifugation (40 min to 2 h at $80,000 \times g$) the leaf peroxisomes are located at the bottom of the centrifugation tube at the interface between 50.5 and 55.2 % (w/w) sucrose. The figure has been adapted and reprinted from Reumann et al. [7] Suppl. Figure 1

14. Sorvall centrifuge RC-5C with SS34 rotor.
15. Ultracentrifuge with swinging bucket rotor (e.g., SW41 Ti).
16. Ultrapure water.
17. 20-mL glass pipettes and Pasteur pipettes (disposable).

18. Potter homogenizer, Wheaton USA.
19. Mortar and pestle.
20. Cold-room (4 °C).
21. Beakers, measuring cylinders and funnel.
22. Vacuum pump.

3 Methods

3.1 Growth of *Arabidopsis* Plants

1. Sow approximately 3–5 seeds of *Arabidopsis thaliana* Col-0 in each pot (15 pots per tray, ca. 1 tray, i.e., approx. 60 plants per isolation) filled up with commercial soil and perlite.
2. Fill the tray ca. 1 cm with water and let the pots soak up the water overnight. Discard excess water.
3. Keep the tray(s) in the cold-room for stratification for 2 days.
4. Transfer the *Arabidopsis* tray(s) to a plant growth room maintained at ~22 °C and light intensity of 100~150 μmol/m²/s in a 16/8 h cycle (long-day).
5. Cover the tray with a plastic dome in the first week until germination to maintain high humidity.
6. Treat the plants with Hoagland nutrient solution [15] once a week, if required.
7. Healthy *Arabidopsis* plants (4–5 weeks old, 60 g fresh weight (FW) are generally required for high leaf peroxisome yield and purity).

3.2 Isolation of Leaf Peroxisomes

1. Harvest the plants in the end of the dark or (preferentially) after an extended dark period (see Note 11).
2. Cut the rosette plants at the top of the root and weigh 60–70 g fresh weight (FW) for a single isolation. Wash the rosette plants to remove all the dirt associated with them. Transfer the leaves to a plastic bag and store between ice crushes until use (generally >2 h) (see Note 11).
3. The isolation procedure should be started in a cold-room (4 °C) or at least on ice using precooled solutions, equipment (e.g., mortar), and centrifuges.
4. Transfer the *Arabidopsis* leaves from one plastic bag into the precooled mortar and add 120 mL of grinding buffer (GB). Grind the leaves with a pestle for ca. 5–10 min until the buffer is dark green and individual leaves are no longer visible. For small mortars it might be advantageous to split the leaves into two equal portions (2× ca. 30 g FW) and grind each with 60 mL GB to increase the grinding efficiency.

Table 2
Biochemical methods established for purity assessment of Arabidopsis leaf peroxisomes

| Organelle | Enzymatic assay | | Immunoblotting | |
|------------------|----------------------|-----------------|--------------------|-------------|
| | Marker enzyme | Reference | Marker protein | Reference |
| Leaf peroxisomes | HPR | [7, 8, 20] | APX | [21] |
| | Catalase | [9, 10, 22–24] | Catalase | [9, 10, 21] |
| | | | KAT2 | [10] |
| Mitochondria | Fumarase | [7, 20, 23, 24] | VDAC | [8] |
| | Cytochrome C oxidase | [9, 21] | NADH dehydrogenase | [22] |
| | | | Cpn10 | [9] |
| Chloroplasts | NADP dependent GAPDH | [7, 20] | PSI-D (thylakoids) | [25] |
| | | | Cpn20 | [9] |
| ER | NADH:Cyt c reductase | [23] | BiP | [9, 22] |
| | | | Calnexin | [22] |

The following acronyms were used: *APX* ascorbate peroxidase, *BiP* binding protein, *Cpn10/20* chaperonin 10/20, *HPR* hydroxypyruvate reductase, *KAT2 thiolase*, *NADP-GAPDH* NADP dependent glyceraldehyde dehydrogenase, *PSI-D* photosystem I subunit D, *SSU* small subunit of RubisCO, *VDAC* 30-kD voltage-dependent anion-selective channel

5. Filter the ground tissue through miracloth (1–3 layers) into an Erlenmeyer flask to obtain the crude extract (CE). Gently squeeze the miracloth to collect the entire CE in the flask. If the CE is to be analyzed (e.g., for marker enzyme activities such as hydroxypyruvate reductase, HPR, for leaf peroxisomes, *see Table 2*), aliquots are taken (*see Notes 12 and 13*).
6. Pour the CE about equally into six SS34 tubes (polypropylene or polycarbonate), balance and centrifuge them for 1 min at $5000 \times g$ (Sorvall SS34) to sediment chloroplasts and nuclei (*see Note 14*).
7. Pour the six supernatants into a precooled 300-mL beaker (*see Note 15*). Layer carefully about 19 mL on top of each of the eight SS34 Percoll gradients (*see item 9 of Subheading 2.2*).
8. Centrifuge the Percoll gradients for 12 min at $13,000 \times g$ at 4°C in a Sorvall SS34 rotor and then increase the speed to $27,000 \times g$ and spin for another 20 min (no slow start, brake on).
9. Carefully suck off the top fractions of the gradients using a short Pasteur glass pipette attached to a vacuum pump. Take off the entire green 15/38 % (v/v) Percoll interface (chloroplasts/thylakoid fraction, Fig. 1) while removing only a minor

portion (ca. 2–3 mL) of the (transparent) 38 % (v/v) Percoll fraction. Carry out this step for *all* eight gradients (stored on ice) before moving on to the next work-up step (**step 10**). In this way residual thylakoids drain from the tube walls downwards and accumulate as a thin light green layer on top of the remaining 38 % (v/v) Percoll fraction.

10. At this step, work up the SS34 tubes one by one until final transfer of the peroxisome fraction to the 100-mL collection beaker. Carefully suck off the residual thylakoid layer, the 38 % (v/v) Percoll fraction and the mixed 38 % (v/v) Percoll: 36 % (w/w) sucrose fractions. Manually rotate the SS34 tube slowly so that each gradient fraction is entirely removed. Only leave about 2–3 mL 36 % (w/w) sucrose solution in each tube including the leaf peroxisomes visible as whitish soft sediment at the bottom of the gradient (Fig. 1). Resuspend the peroxisome pellet gently with a disposable 3-mL plastic pipette of relatively wide opening. Pool the eight peroxisome fractions in a (pre-cooled) 100-mL beaker (ca. 16 mL in total). By this harvest method, a post-centrifugation contamination of the leaf peroxisome fraction by the upper thylakoid fractions can be largely avoided. A significant chloroplast/thylakoid contamination is indicated by a greenish color of the leaf peroxisome fraction.
11. To remove residual Percoll and fully adjust the sucrose concentration to 36 % (w/w), dilute the leaf peroxisome fraction approx. 1:4 very gently by mL-wise adding ca. 65 mL 36 % (w/w) sucrose (in TE buffer) to a final volume of about 80 mL.
12. Pour the diluted peroxisome fraction into four SS34 tubes and centrifuge them at $39,000 \times g$ for 30 min.
13. Suck off the supernatant using a vacuum pump and carefully collect the washed leaf peroxisome fraction (ca. 3–4 mL in total) found at the bottom of the tubes by using disposable 3-mL plastic Pasteur pipettes with a wide opening to reduce shear forces.
14. Homogenize the fraction carefully by approx. five slow strokes using a potter homogenizer. The homogenizer should have moderate space between the pistil and the glass wall to avoid application of damaging shear forces onto the leaf peroxisomes. Transfer the fraction to a 10-mL measuring cylinder and adjust the final volume to 5 mL using 36 % (w/w) sucrose solution. Add the three protease inhibitors at the given concentrations as described above for the GB. This is the first purified leaf peroxisome fraction (referred to as LP-P1 in Fig. 1 [7]), whose purity and peroxisome intactness can be investigated by measuring the activity of the leaf peroxisome marker enzyme HPR (Table 2) (*see Note 13*).
15. Carefully lay the peroxisome fraction on top of the 41.2 % (w/w) sucrose fraction of the preprepared sucrose density

gradient assembled in an ultracentrifuge tube (*see Note 10* and Subheading 2.2 item 11, Fig. 2). Load 5 mL of 36 % (w/w) sucrose solution on the second density gradient needed as a counterbalance for the leaf peroxisome gradient during ultracentrifugation.

16. Place the two gradients in two precooled SW41 Ti buckets. Carefully balance the filled and the empty SW41 Ti buckets and spin for 40 min to 2 h at $80,000 \times g$ (*see Note 16*).
17. After centrifugation a white band of leaf peroxisomes should be visible at the interface of the 50.5 % (w/w) and 55.2 % (w/w) sucrose density gradient fractions (Fig. 2). Suck off the upper fractions with a Pasteur glass pipette attached to a vacuum pump. Harvest the peroxisome band in a volume of approx. 1 mL and adjust the volume in a 1.5-mL Eppendorf tube to 1.5 mL using 36 % (w/w) sucrose solution or TE buffer only (*see Note 17*).
18. Add the following protease inhibitors at the following final concentrations: PMSF (1 mM), Benzamidine (2 mM), ϵ -Aminocaproic acid (2 mM), Aprotinin (1 μ g/mL), Pepstatin (1 μ g/mL), and Leupeptin (1 μ g/mL). Prepare a second Eppendorf tube as blank for protein determination containing approx. the same sucrose concentration in TE buffer (ca. 45 % w/w) and the same final concentrations of the six protease inhibitors.
19. Mix well and prepare a few 50- μ L aliquots for analysis of protein concentration and purity using enzymatic assays such as HPR activity and immunoblotting (Table 2) and freeze all samples at -20°C (*see Note 18*).
20. To concentrate proteins for subsequent proteome studies we recommend chloroform/methanol precipitation [16].

4 Notes

1. If two Sorvall centrifuges are available, one person can carry out two leaf peroxisome isolations in ca. 4–5 h, and one to two persons can carry out four to six leaf peroxisome isolations per day. The average yield of one isolation is ca. 120 μ g protein [7].
2. All solutions are generally prepared on the previous day and stored in the refrigerator. The pH of solutions is adjusted after precooling to ca. 10°C .
3. Stock solutions such as 1.0 M Tricine, 100 mM EDTA, 1.0 M KCl, and 1.0 M MgCl₂ are recommended to facilitate buffer preparation.
4. Appropriate stock solutions such as 100 mM PMSF (in MeOH), 200 mM Benzamidine (in H₂O, freshly prepared),

and 200 mM ϵ -Aminocaproic acid (in H_2O , freshly prepared) are recommended. Instead of this self-mixed protease inhibitor set, alternatives can be used such as the Complete Protease Inhibitor Cocktail Tablets (Roche).

5. Measure exactly the required Percoll volume and the final solution volume because the final Percoll concentration of 15 % (v/v) crucially determines peroxisome yield and purity.
6. If preparing the Percoll solutions directly in TE buffer, pH 7.5, the pH needs to be readjusted back to 7.5 because the pH of Percoll (ca. 9) increases the pH of the TE buffer.
7. The sucrose concentration (in % w/w) needs to be adjusted to highest accuracy at 20 °C using a refractometer.
8. The density of the 38 % (v/v) Percoll solution and the 36 % (w/w) sucrose in TE buffer solution differ only marginally from each other but are prerequisite for proper gradient preparation and leaf peroxisome isolation. Prior to gradient assembly it is recommended to check in an Eppendorf tube whether the density of the 36 % (w/w) sucrose solution is indeed slightly higher than that of the 38 % (v/v) Percoll solution by either underlying or overlaying 0.5 mL of one solution with 0.5 mL of the other. If positioning of the two solutions is difficult to see, one solution aliquot can be stained blue with a grain of Coomassie Brilliant Blue (e.g., R250).
9. Store the gradients on ice for at least 30 min before use.
10. If required, the total gradient volume can be reduced by reducing the volume of the lower fractions, i.e., 0.8 mL 60.0 % (w/w) and 0.8 mL 55.2 % (w/w) sucrose.
11. The plants should best be kept in the dark for about 12 h before starting the peroxisome isolation. This extended dark incubation lowers the physical adherence between leaf peroxisomes, chloroplasts, and mitochondria and significantly reduces the contamination of leaf peroxisomes by chloroplasts and mitochondria. Also the post-harvest leaf incubation on ice appears to reduce inter-organelle adherence.
12. Mix the crude extract by gently shaking the Erlenmeyer flask, determine the total volume of the crude extract (approx. 140 mL), take small aliquots for organelle enrichment analyses (Table 2) (e.g., 3 × 1 mL), and freeze them at -20 °C.
13. In *Arabidopsis*, two additional HPR homologues have recently been shown to account for minor NAD(P)H-dependent HPR activities in the cytosol and chloroplasts, respectively [17, 18], but are not considered to significantly diminish the traditional, easy and reliable use of peroxisomal HPR as a marker enzyme for leaf peroxisomes.
14. Start measuring the 1-min centrifugation when the maximum speed has been reached.

15. The sediment of cell debris, nuclei and chloroplasts is tight, allowing simple decanting of the supernatant. Overlaying the CE supernatant on top of the Percoll gradients shall be done as quickly as possible but also carefully. A second pair of helping hands is advantageous. The CE supernatant fraction shall be quantitatively loaded onto the Percoll gradients to maximize yield.
16. If the leaf peroxisomes are very labile, for instance if deriving from stressed plants or senescent leaves, they tend to lyse during ultracentrifugation and float from their typical density of ca. 1.23 g/mL to the top of the sucrose density gradient. Hence, shorter centrifugation times (e.g., 40 min) are recommended for labile peroxisomes to obtain maximum yield.
17. Collect the peroxisome band in no more than 1.2 mL because the viscosity of the sucrose solutions needs to be reduced to guarantee accurate pipetting and freezing.
18. Protein determination has been carried out according to Lowry [19].

Acknowledgment

Dr. R. Singhal was supported by an YGGDRASIL IS-MOBIL fellowship from the Norwegian Research Council.

References

1. Yu C, Huang AH (1986) Conversion of serine to glycerate in intact spinach leaf peroxisomes: role of malate dehydrogenase. *Arch Biochem Biophys* 245:125–133
2. Lopez-Huertas E, Sandalio LM, Del Rio LA (1995) Integral membrane polypeptides of pea leaf peroxisomes: characterization and response to plant stress. *Plant Physiol Biochem* 33:295–302
3. Cooper TG, Beevers H (1969) Beta oxidation in glyoxysomes from castor bean endosperm. *J Biol Chem* 244:3514–3520
4. Harrison-Lowe N, Olsen LJ (2006) Isolation of glyoxysomes from pumpkin cotyledons. *Curr Protoc Cell Biol* 3(19):1–8
5. The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–815
6. Fukao Y, Hayashi M, Nishimura M (2002) Proteomic analysis of leaf peroxisomal proteins in greening cotyledons of *Arabidopsis thaliana*. *Plant Cell Physiol* 43:689–696
7. Reumann S, Babujee L, Ma C et al (2007) Proteome analysis of *Arabidopsis* leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *Plant Cell* 19:3170–3193
8. Reumann S, Quan S, Aung K et al (2009) In-depth proteome analysis of *Arabidopsis* leaf peroxisomes combined with *in vivo* subcellular targeting verification indicates novel metabolic and regulatory functions of peroxisomes. *Plant Physiol* 150:125–143
9. Fukao Y, Hayashi M, Hara-Nishimura I et al (2003) Novel glyoxysomal protein kinase, GPK1, identified by proteomic analysis of glyoxysomes in etiolated cotyledons of *Arabidopsis thaliana*. *Plant Cell Physiol* 44:1002–1012
10. Eubel H, Meyer EH, Taylor NL et al (2008) Novel proteins, putative membrane transporters, and an integrated metabolic network are revealed by quantitative proteomic analysis of *Arabidopsis* cell culture peroxisomes. *Plant Physiol* 148:1809–1829

11. Kaur N, Reumann S, Hu J (2009) Peroxisome biogenesis and function. In: Somerville CR, Meyerowitz EM (eds) *The Arabidopsis book*. The American Society of Plant Biologists, Rockville, MD, USA, pp 1–41
12. Hu J, Baker A, Bartel B et al (2012) Plant peroxisomes: biogenesis and function. *Plant Cell* 24:2279–2303
13. Lingner T, Kataya AR, Antonicelli GE et al (2011) Identification of novel plant peroxisomal targeting signals by a combination of machine learning methods and *in vivo* subcellular targeting analyses. *Plant Cell* 23:1556–1572
14. Lingner T, Kataya AR, Reumann S (2012) Experimental and statistical post-validation of positive example EST sequences carrying peroxisome targeting signals type 1 (PTS1). *Plant Signal Behav* 7
15. Hoagland DR, Arnon DI (1950) The water-culture method for growing plants without soil. California Agricultural Experiment Station Circular 347:1–32
16. Wessel D, Flugge UI (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* 138:141–143
17. Timm S, Nunes-Nesi A, Parnik T et al (2008) A cytosolic pathway for the conversion of hydroxypyruvate to glycerate during photorespiration in *Arabidopsis*. *Plant Cell* 20: 2848–2859
18. Timm S, Florian A, Jahnke K et al (2011) The hydroxypyruvate-reducing system in *Arabidopsis*: multiple enzymes for the same end. *Plant Physiol* 155:694–705
19. Lowry OH, Rosebrough NJ, Farr AL et al (1951) Protein measurement with the Folin phenol reagent. *J Biol Chem* 193:265–275
20. Ma C, Haslbeck M, Babujee L et al (2006) Identification and characterization of a stress-inducible and a constitutive small heat-shock protein targeted to the matrix of plant peroxisomes. *Plant Physiol* 141:47–60
21. Lisenbee CS, Lingard MJ, Trelease RN (2005) *Arabidopsis* peroxisomes possess functionally redundant membrane and matrix isoforms of monodehydroascorbate reductase. *Plant J* 43: 900–914
22. Lisenbee CS, Heinze M, Trelease RN (2003) Peroxisomal ascorbate peroxidase resides within a subdomain of rough endoplasmic reticulum in wild-type *Arabidopsis* cells. *Plant Physiol* 132:870–882
23. Tugal HB, Pool M, Baker A (1999) *Arabidopsis* 22-kilodalton peroxisomal membrane protein. Nucleotide sequence analysis and biochemical characterization. *Plant Physiol* 120:309–320
24. Liepman AH, Olsen LJ (2001) Peroxisomal alanine: glyoxylate aminotransferase (AGT1) is a photorespiratory enzyme with multiple substrates in *Arabidopsis thaliana*. *Plant J* 25: 487–498
25. Kubis SE, Lilley KS, Jarvis P (2008) Isolation and preparation of chloroplasts from *Arabidopsis thaliana* plants. *Methods Mol Biol* 425:171–186

Part VI

Comparative Proteomics

Chapter 37

Proteomics of Field Samples in Radioactive Chernobyl Area

Katarína Klubicová, Namik M. Rashydov, and Martin Hajduch

Abstract

Two serious nuclear accidents during the past quarter of a century contaminated large agricultural areas with radioactivity. The remediation and possible recovery of radio-contaminated areas for agricultural purposes require comprehensive characterization of plants grown in such places. Here we describe the quantitative proteomics method that we use to analyze proteins isolated from seeds of plants grown in radioactive Chernobyl zone.

Key words Chernobyl, Radioactivity, Plants, Seeds, Protein extraction, Two-dimensional electrophoresis, In-gel digestion

1 Introduction

Nuclear accidents in Ukraine (Chernobyl, 1986) and Japan (Fukushima, 2011) contaminated large agricultural areas with radioactivity. Surprisingly, plants grow and successfully reproduce in these areas. Detailed characterizations of plants grown in radio-contaminated areas provide necessary scientific foundations for agricultural recovery of areas contaminated with radioactivity [1, 2]. Proteomics is method of choice for analysis of plants grown in radio-contaminated areas because is capable to quantify hundreds of proteins in single experiment. Moreover, proteomics visualize metabolic changes regardless their transcriptional, translational, or epigenetic nature [3].

Our group use proteomics approach to investigate developing and matured seeds of plant species grown in radioactive Chernobyl area since 2007. Using combination of phenol protein extraction, protein two-dimensional electrophoresis (2-DE), and tandem mass spectrometry we were able to characterize mature and developing seeds of soybean grown in radioactive Chernobyl area for two subsequent generations [4, 5]. Moreover, using same methodology we investigated flax as suitable crop for agricultural practices in

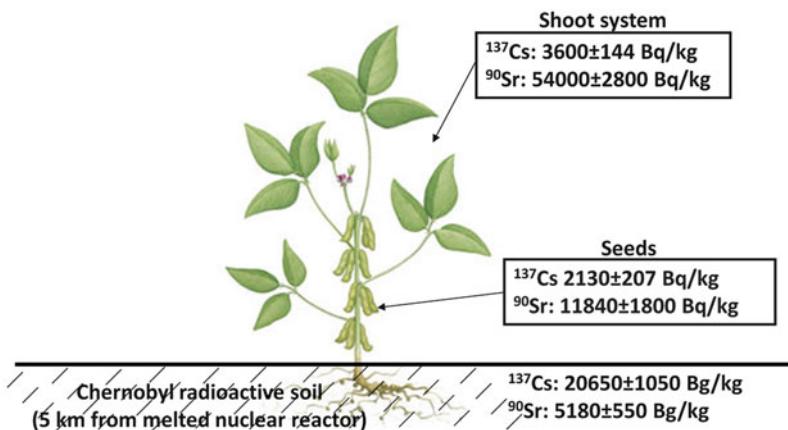


Fig. 1 Schematic view on ^{137}Cs and ^{90}Sr uptake by soybean. Soybean shoot system (soybean plant without roots, seed pods, and seeds) accumulated higher amounts of radioactivity than seeds during our experiments in radioactive Chernobyl zone (Chistogalovka area)

agricultural areas remediated from radioactivity and provided detailed characterization of flax seeds harvested from radioactive and remediated Chernobyl areas [1, 2, 6].

Herein we describe the methodology that we use to analyze seed proteome in radioactive Chernobyl area. Firstly, soils in radioactive areas are often nutritiously poor and thus the fertilization might be necessary. Secondly, selection of plant material determines the level of radioactivity in the experiments. In our experience, the uptake of radioactive elements differs between plants and also between organs/tissues (Fig. 1). For instance, soybean shoot system (soybean without roots, seed pods, and seeds) accumulated $3,600 \pm 144$ Bq/kg of ^{137}Cs and $54,000 \pm 2,800$ Bq/kg of ^{90}Sr , during our experiments in radioactive Chernobyl zone (Chistogalovka area). However, seeds accumulated only $2,130 \pm 207$ Bq/kg of ^{137}Cs and $11,840 \pm 1,800$ Bq/kg of ^{90}Sr [5]. Furthermore, flax seeds accumulated even less radioactivity, 780 ± 39 Bq/kg of ^{137}Cs and $3,550 \pm 360$ Bq/kg of ^{90}Sr at the same area.

2 Materials

2.1 Phenol-Based Protein Extraction

1. Homogenization solution: 50 % (v/v) phenol, 0.45 M sucrose, 5 mM EDTA, 0.2 (v/v) 2-mercaptoethanol, 50 mM Tris-HCl, pH 8.8 (*see Note 1*).
2. 0.1 M ammonium acetate in 100 % methanol (store at -20°C).
3. 80 % acetone (store at -20°C).
4. 70 % ethanol (store at -20°C).

2.2 Isoelectric Focusing (IEF)

1. IEF extraction/solubilization media: 8 M urea, 2 M thiourea, 2 % (w/v) CHAPS, 2 % (v/v) Triton X-100, 50 mM DTT (*see Note 2*).
2. Immobilized pH Gradient (IPG) buffer or Ampholytes.
3. Mineral oil.

2.3 SDS-PAGE

1. 1.5 M Tris-HCl, pH 8.8.
2. 30 % acrylamide/bisacrylamide solution (29:1).
3. 10 % APS in water.
4. TEMED.
5. Isobutanol.
6. SDS equilibration buffer: 1.5 M Tris-HCl, 6 M urea, 30 % (v/v) glycerol, 5 % (w/v) SDS (if used stacking gel) or 6 M urea, 50 mM Tris-HCl, pH 8.8, 2 % (w/v) SDS, 30 % (v/v) glycerol (if stacking gel is not used) with 2 % (w/v) dithiothreitol (DTT) or 2.5 % (w/v) iodoacetamide (IAA) (*see Note 3*).
7. SDS running buffer: 25 mM Tris, 0.192 M glycine, 0.1 % (w/v) SDS.
8. 0.5 % (w/v) agarose in SDS running buffer with traces of bromophenol blue.
9. Colloidal Coomassie Blue staining solution: 20 % (v/v) ethanol, 1.6 % (v/v) phosphoric acid, 8 % (w/v) ammonium sulfate, 0.08 % (w/v) Coomassie Brilliant Blue (CBB) G-250 (*see Note 4*).

2.4 Protein Digestion

1. Wash solution: 50 % acetonitrile, 50 mM ammonium bicarbonate.
2. 100 % acetonitrile.
3. Trypsin in 50 mM ammonium bicarbonate.
4. Extraction solution: 60 % acetonitrile, 1 % formic acid.

3 Methods

3.1 Phenol-Based Protein Extraction

- This extraction method is modified from Hurkman and Tanaka [7].
1. Grind sample (0.5 g) to a fine powder with liquid nitrogen, mortar and pestle.
 2. While still in mortar, immediately resuspend powder in 10 mL of homogenization solution. Homogenate will freeze (*see Note 5*).
 3. Allow homogenate to reach room temperature (RT), transfer to a 50 mL propylene tube and mix slowly on a shaker for 30 min at 4 °C.
 4. Centrifuge at 5,000 × g and 4 °C for 15 min.

5. Remove carefully upper phase (phenol) and precipitate proteins with five volumes of ice cold 0.1 M ammonium acetate in 100 % methanol at -20 °C for 16 h.
6. Centrifuge at 5,000 $\times g$ and 4 °C for 15 min.
7. Wash the protein pellet thoroughly twice in 20 mL of 0.1 M ammonium acetate in 100 % methanol followed with two washes in ice-cold 80 % acetone and final wash 70 % ethanol. Mix for 10–15 min. Centrifuge and repeat washing procedure (see Note 6).

3.2 Isoelectric Focusing

1. Resuspend protein pellet in rehydration solution.
2. Remove insoluble matter by centrifugation for 20 min at 14,000 $\times g$.
3. Determine protein concentration by Bradford assay [8].
4. Required amount of proteins add to 1.5 mL eppendorf tube, add corresponding IPG buffer or Ampholyte and bring volume up to required volume with IEF extraction solution (315 μ L of sample for 17 cm long IPG strips). Vortex and spin for 5 min at max speed.
5. Transfer the protein solution into isoelectrophoretic focusing tray and rehydrate for 1 h at RT.
6. Overlay strips with mineral oil (1 mL).
7. Place strip into isoelectric focusing and start isoelectric focusing under these conditions.
 - (a) Active rehydration (10 h at 50 V).
 - (b) 100 V for 100 Vh.
 - (c) 500 V for 500 Vh.
 - (d) 8,000 V for 99 KVh.
 - (e) Hold at 50 V.

3.3 SDS-PAGE Electrophoresis

1. Remove IPG strips from focusing unit.
2. Incubate IPG strips in SDS equilibration buffer with 2 % (w/v) DTT for 15 min.
3. Incubate IPG strips in SDS equilibration buffer with 2.5 % (w/v) IAA for 15 min.
4. Rinse strips with running buffer and place onto 12 % acrylamide gels.
5. Overlay strips with 0.5 % agarose in running buffer with traces of bromophenol blue.
6. Carry second dimension SDS-PAGE at 2W/gel for 16 h or until dye migrated off the gel.
7. Following SDS-PAGE, wash gels in three times in water for 15 min and stain for at least 12 h with CBB.

3.4 In Gel Protease Digestion

1. Wash gel pieces with 500 μ L of wash solution and incubate at RT for 15 min with gentle agitation.
2. Remove solution with a pipette carefully.
3. Wash gel pieces two more times with 500 μ L of wash solution for 15 min or until the Coomassie dye has been removed.
4. Dehydrate the gel pieces in 100 % acetonitrile for 5 min. Note: When dehydrated, the gel pieces will have an opaque white color and will be significantly smaller in size.
5. Remove acetonitrile with pipette and dry at RT for 5–10 min.
6. While gel pieces are drying prepare protease digestion solution.
7. Rehydrate the gel with a minimal volume of protease digestion solution (see Note 7).
8. Spin down sample by centrifugation.
9. Add 25–50 μ L of extraction solution to gel pieces and agitate gently by vortexing at lowest setting for 10 min.
10. Spin down sample by brief centrifugation (12 kg for 30 s).
11. Transfer the supernatant (containing tryptic peptides) to the tube.
12. Re-extract the gel plugs with an additional 25–50 μ L of extraction solution.
13. Spin down sample and transfer supernatant to tube from step 11.
14. Dry the pooled extracted peptides by centrifugal evaporation to near dryness. Do not use heat. Do not dry for extended time.
15. Store in –80 °C until analyzed by mass spectrometry for protein identification.

4 Mass Spectrometry Strategy and Results Obtained

The approach used for the identification of proteins presented in excised gel pieces was based on tandem mass spectrometry (MS/MS). Briefly, gel pieces were “in-gel” digested with trypsin and separated using nanoAcquity UPLC system (Waters, Milford, MA, USA). This system contained reverse-phase chromatography using a BEH 130 C18 column (100 μ m \times 150 mm, 1.7 μ m particle size; Waters, Milford, MA, USA). Separated tryptic peptides were ionized by nanospray and analyzed by Q-TOF Premier MS/MS instrument (Waters, Milford, MA, USA). For data acquisition, the MS^E method was used that is based on alternate scans at low and high collision energies. The MS data were searched against plant UniProt databases using ProteinLynx Global Server v. 2.4 (Waters, UK).

The results obtained using the protocol described herein were deposited to Web-based database www.chernobylproteomics.sav.sk where the data are available in user-friendly format using interactive 2-DE gels.

5 Notes

1. Homogenization solution: Prepare stock solution containing 0.9 M sucrose, 10 mM EDTA, and 100 mM Tris–HCl (pH 8.8) and store at 4 °C. Before use, add 0.4 % of (v/v) 2-mercaptoethanol and one volume of phenol saturated by Tris–HCl pH 8.8. Invert tube for several times to mix the solution.
2. IEF extraction media/solubilization buffer: prepare solution without DTT and store aliquots at –20 °C. Add DTT before use.
3. SDS equilibration buffer: prepare solution without DTT or IAA and store aliquots at –20 °C. Add DTT or IAA before use.
4. Filter Colloidal Coomassie Blue (CBB) staining solution to avoid interaction of unresolved CBB with 2-DE gel.
5. No peptidase inhibitors are necessary. Freezing of homogenate in mortal after tissue dinging with liquid nitrogen followed by thawing in the presence of phenol avoids protein degradation.
6. After protein precipitation, pellet should be washed thoroughly. It is recommended to resuspend pellet in washing solution until it looks like a milk, without visible solid particles.
7. Use 20 µL for small gel plugs. Add more if gel pieces absorb all liquid. Gel pieces must be hydrated throughout the digest. Digest overnight at 37 °C.

Acknowledgments

This work was supported by the Seventh Framework Program of the European Union—International Reintegration Grant (MIRG-CT-2007-200165), Scientific Grant Agency of the Ministry of Education of the Slovak Republic and the Academy of Sciences (VEGA-2/0126/11), APVV granting agency (APVV-0740-11), and Research & Development Operational Program funded by the ERDF—Centre of Excellence for White-Green Biotechnology (ITMS 26220120054). This paper reflects only the author's views and the Community is not liable for any use that might be made of information contained herein.

References

1. Klubicová K, Berčák M, Danchenko M et al (2011) Agricultural recovery of a formerly radioactive area: I. Establishment of high-resolution quantitative protein map of mature flax seeds harvested from the remediated Chernobyl area. *Phytochemistry* 72:1308–1315
2. Klubicová K, Danchenko M, Skultety L et al (2011) Agricultural recovery of a formerly radioactive area: II. Systematic proteomic characterization of flax seed development in the remediated Chernobyl area. *J Proteomics* 74:1378–1384
3. Hajduch M, Hearne LB, Miernyk JA et al (2010) Systems analysis of seed filling in *Arabidopsis*: using general linear modeling to assess concordance of transcript and protein expression. *Plant Physiol* 152:2078–2087
4. Danchenko M, Skultety L, Rashydov NM et al (2009) Proteomic analysis of mature soybean seeds from the Chernobyl area suggests plant adaptation to the contaminated environment. *J Proteome Res* 8:2915–2922
5. Klubicová K, Danchenko M, Skultety L, et al (2012) Soybeans Grown in the Chernobyl Area Produce Fertile Seeds that Have Increased Heavy Metal Resistance and Modified Carbon Metabolism. *PLoS One* 7(10):e48169
6. Klubicová K, Danchenko M, Skultety L et al (2010) Proteomics analysis of flax grown in Chernobyl area suggests limited effect of contaminated environment on seed proteome. *Environ Sci Technol* 44:6940–6946
7. Hurkman WJ, Tanaka CK (1986) Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol* 81:802–806
8. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254

Chapter 38

Differential Proteome and Secretome Analysis During Rice–Pathogen Interaction

Yiming Wang, Sang Gon Kim, Jingni Wu, Sun Tae Kim, and Kyu Young Kang

Abstract

Substantial evidences implicate that sample preparation and protein extraction in proteomic studies of plant–pathogen interactions are critical to understand cross talk between host and pathogen. Therefore, interest is growing in applying proteomics techniques to investigate simultaneously secreted proteins from rice and pathogen. We have found, however, that most proteins of interest are low abundant so that proper prefractionation or extraction of secreted proteins from extracellular space (ECS) in the rice leaf is required to excavate relevant protein. This chapter describes the preparation of sample and extraction procedure to enrich the proteins interested before separation by 2-DE or LC-MS/MS. This method significantly increases the sensitivity of proteomic comparisons.

Key words Rice proteomics, Rice, *Magnaporthe oryzae*, *Xanthomonas oryzae*, 2-DE

1 Introduction

A number of rapid defense responses in plants are initiated after pathogen attack and are activated rapidly hereafter [1, 2]. In general, an interaction is incompatible when the rice plant recognizes the invading pathogen early enough and activates the host resistance genes, resulting in a hypersensitive response (HR) and the triggering of rapid and effective defense responses including oxidative burst, the production of pathogenesis-related (PR) proteins, and phytoalexins [3]. In contrast, an interaction is compatible when the rice plant responds too late to restrict ingress of the pathogen.

Upon interactions, the extracellular space (ECS) of the rice leaf serves as a front line of battle field of defense against the invading pathogen. Both species secrete a diversity of functional components into the ECS. It is likely that host arranges various molecular events such as strengthening of the cell wall, and producing antimicrobial activity by pathogenesis-related (PR) proteins,

metabolites, ions, and a large variety of hydrolytic enzymes in the ECS for survival [4, 5]. ECS is also a place where pathogens secrete proteins, such as cell wall hydrolytic enzymes, peptidases, toxins, oxidation/reduction, and apoplastic elicitors or effectors, and that play a crucial role in infection and pathogenicity [6–8].

To understand their interaction mechanism, the genome-geared high-throughput technologies, such as transcriptomics, proteomics, and metabolomics, have been conducted [5, 9, 10]. Plant proteomic studies have been performed during its growth and development, and against biotic and abiotic stresses. However, little report is available on *in planta* secretome investigation of rice leaves infected with pathogen. The success of the 2-DE technology for proteome analysis will rely on the ability to display more rare proteins and insoluble proteins. In this chapter, we have reported to provide step-by-step protocol for protein analyses of rice and pathogen during rice–pathogen interaction through phenol, PEG fractionation, and secreted protein extraction method.

2 Materials

2.1 Preparation of *Magnaporthe oryzae*

1. *M. oryzae* (compatible, incompatible race).
2. Petri dish.
3. Sterile glass beaker.
4. Rice bran agar medium (25 g/L rice bran powder, 1 g/L sucrose, and 20 g/L agar).
5. Incubator (28 °C).

2.2 Preparation of *Xanthomonas oryzae*

1. *Xanthomonas oryzae* pv. *oryzae* (compatible, incompatible races).
2. Petri dish.
3. PSA plate; 1 % w/v peptone (Becton and Dickinson, Franklin Lakes, NJ, USA), 1 % w/v sucrose, and 1.5 % w/v agar.
4. PS liquid medium; 1 % w/v peptone, 1 % w/v sucrose.
5. 0.01 % Tween-20 (Amersco, Solon, OH, USA).
6. Incubator (30 °C).

2.3 Preparation of Host Plant

1. Dry mature seeds of rice (*Oryzae sativa* L.) (see Note 1).
2. Washing solution: 70 % ethanol in distilled or Milli Q water.
3. Sterilizing solution: 3 % sodium hypochlorite (Junsei Chemical, Tokyo, Japan) in distilled or Milli Q water.
4. Humidity growth chamber (Daehan, Seoul, Korea) with white fluorescent light (wavelength 390–500 nm, 150 $\mu\text{mol}/\text{m}^2/\text{s}$, 16 h light/8 h dark cycles) and 70 % relative humidity at 28 °C.
5. Soil (Pot mix for rice) (see Note 2).

6. Seedling pot (55×50×55 mm, T. O. Plastics, MN, USA).
7. Fourth or fifth leaf stage of rice seedling (see Note 3).
8. Liquid nitrogen.
9. Sterile/clean mortar and pestle.
10. 40 mL Nalgene centrifuge tubes (Rochester, NY, USA) and 1.5 mL microfuge tubes (Sarstedt, Nümbrecht, Germany).

2.4 Preparation of Total Protein

Extraction: Phenol Extraction

1. Rice leaves (*O. sativa* L. cv. Jinheung) (4th and 5th leaf stages) with/without *M. oryzae* or *X. oryzae* infection.
2. Phenol: Tris-HCl, pH 7.9, saturated (Amresco, Solon, OH, USA).
3. Mg/NP-40 extraction buffer: 0.5 M Tris-HCl, pH 8.3, 2 % v/v Nonidet P (NP)-40 (Amresco, Solon, OH, USA), and 20 mM MgCl₂. This solution can be stored at 4 °C.

Just before extraction, add 2 % (v/v) β-mercaptoethanol (Amresco, Solon, OH, USA), 1 mM phenylmethylsulfonyl fluoride (PMSF; Merck, Whitehouse Station, NJ, USA), and 1 % w/v polyvinylpolypyrrolidone (PVPP; Sigma, St. Louis, MO, USA).

4. Washing solution: 80 % (v/v) acetone (SK Chemical, Ulsan, South Korea) in deionized water. This solution can be stored at -20 °C.
5. Precipitation solution: 100 % methanol (Burdick & Jackson, Morristown, NJ, USA) containing 0.1 M ammonium acetate (Sigma, St. Louis, MO, USA). This solution can be stored at -20 °C.

2.5 Preparation of Polyethylene Glycol (PEG) Fractionation

1. Mg/NP-40 extraction buffer.
2. 50 % PEG 4000 solution (Sigma, St. Louis, MO, USA).
3. 100 % methanol containing 0.1 M ammonium acetate.
4. Tris-HCl, pH 7.9, saturated phenol.

2.6 In Planta Secretome

1. Nalgene 250 mL centrifuge bottle.
2. Extraction buffer (CA buffer): 200 mM CaCl₂ (Sigma, St. Louis, MO, USA), 5 mM Na-acetate (Sigma, St. Louis, MO, USA), pH 4.3.
3. No.2 filter paper (Advantec MFS, Dublin, CA, USA).
4. Tris-HCl, pH 7.9, saturated phenol.
5. Precipitation solution: 100 % methanol containing 0.1 M ammonium acetate.
6. Resuspension solution: 80 % methanol containing 0.1 M ammonium acetate.
7. Washing solution: 80 % acetone.

8. Rehydration buffer: 8 M (w/v) urea (Amresco, Solon, OH, USA), 2 % (w/v) CHAPS (Amresco, Solon, OH, USA), 0.002 % (w/v) bromophenol blue (Amresco, Solon, OH, USA), 20 mM dithiothreitol (DTT; Amresco, Solon, OH, USA), 0.5 % (v/v) pharmalyte (pH 5–8; GE Healthcare Bio-Sciences AB, Uppsala, Sweden).

3 Methods

3.1 Preparation of Synchronous Culture

Inoculum and Infection of *M. oryzae* Fungal Pathogens

1. Fungal conidia were spread on the rice bran agar medium and grew for 3 days.
2. Aerial mycelia were removed by a spreader and incubated under inflorescent light for another 3 days to get synchronously produced conidia for inoculation.
3. Four-week-old rice seedlings were inoculated with conidia suspension (1×10^6 conidia/mL) of incompatible or compatible races of *M. oryzae* using an air sprayer (see Note 4).
4. Inoculated plants were kept in a humidity chamber at 28 °C.
5. Leaves were harvested at 72 h post-inoculation (PI) and processed immediately for isolation of proteins (see Note 5).
6. For transcript profiling, leaf samples were collected at 12, 48, and 72 h PI, frozen in liquid nitrogen, and stored at -70 °C.

3.2 Infection of *Xanthomonas oryzae*

1. Inoculate *X. oryzae* into 5 mL PS liquid medium from glycerol stock or PSA agar medium culture.
2. Culture cells of *X. oryzae* in PS liquid medium at 30 °C incubator with shaking at $150 \times g$ overnight.
3. Transfer the 5 mL cultured cells into 500 mL fresh PS liquid medium and cultured for 6–8 h until OD 0.8–1.0.
4. Collect and wash cells of *X. oryzae* with distilled water twice.
5. Dilute cells to 1×10^8 cfu/mL with 0.01 % tween-20 in distilled water.
6. Cut leaf tip (approx. 2 cm) with sterilized scissors and immerse leaf into the *X. oryzae* solutions for 30 min.
7. Put into the humidity chamber for disease development.

3.3 Total Protein Extraction: Phenol Method

This method is depicted as a flow chart in Fig. 1.

1. Before use, mix equal volume of Mg/NP-40 extraction buffer (see Note 6) and phenol (phenol extraction solution) (see Note 7).
2. Add 10 mL phenol extraction solution to a 50 mL Nalgene centrifuge tube containing 2 g of finely powdered tissue sample. Mix gently at room temperature (RT) for 10 min.

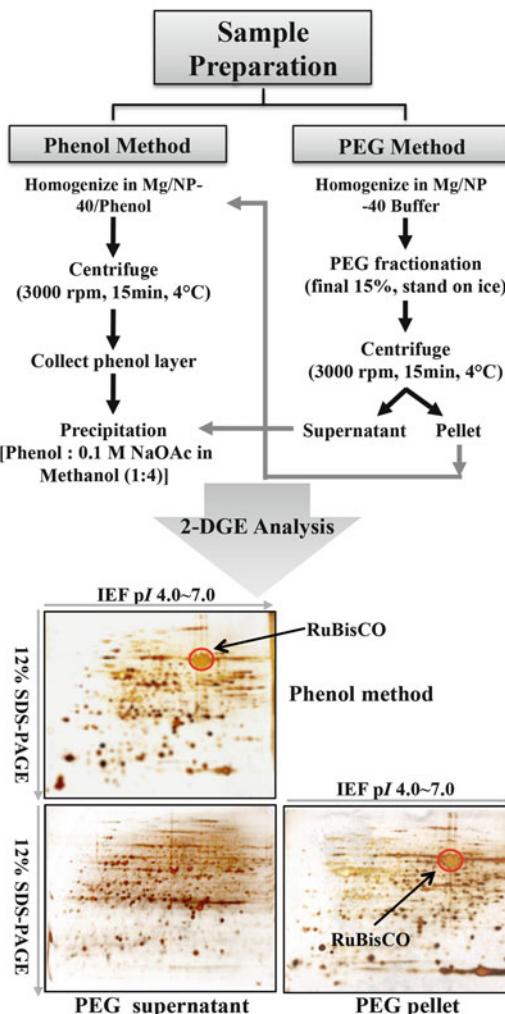


Fig. 1 Extraction procedure for phenol and PEG fractionation method

3. Centrifuge at $12,000 \times g$ for 15 min at 4 °C. Transfer phenol phase to a new 50 mL Nalgene centrifuge tube.
4. Add 4 volume of methanol containing 0.1 M ammonium acetate, vortex, and place at -20 °C for at least 2 h or overnight to precipitate proteins from the phenol.
5. Wash the protein pellet with 10 mL of 0.1 M ammonium acetate in methanol. Repeat this step two times.
6. Discard the solution and wash the protein pellet with 10 mL of 80 % acetone. Repeat this step two more times.
7. Semi-dry the tubes in at RT to remove excess liquid.
8. Resuspend protein pellet in rehydration buffer and proceed for 2-DE analysis (IPG gel system).
9. Proceed for 2-DE analysis (Fig. 1).

3.4 PEG Fractionation for Removal of RuBisCO

This method is depicted as a flow chart in Fig. 1.

1. Incubate finely powdered tissue sample with Mg/NP-40 extraction buffer (5 mL/g tissue).
2. Mix for 30 min on ice and centrifuge at $12,000 \times g$ for 15 min.
3. Transfer the supernatant into a new 50 mL Nalgene centrifuge tube.
4. Add PEG to a final concentration of 15 % (w/v) using a 50 % (w/v) PEG stock solution and incubate on ice for 30 min or more (see Note 8).
5. Centrifuge at $12,000 \times g$ for 15 min at 4 °C.
6. Save the PEG pellet. Transfer supernatant to a new tube and add 4 volumes of 100 % acetone.
7. Place at -20 °C for 2–3 h to precipitate proteins and then centrifuge at $3,000 \times g$ for 15 min.
8. Resuspend the PEG pellets in 10 mL of Mg/NP-40 extraction buffer and vortex at RT.
9. Add an equal volume of water-saturated phenol, vortex, and centrifuge at $3,000 \times g$ for 15 min.
10. Collect the phenolic phase and precipitate proteins by adding 4 volumes of methanol containing 0.1 M ammonium acetate at -20 °C for 3 h followed by centrifugation at $3,000 \times g$ for 10 min.
11. Decant solution carefully and wash protein pellet three times with methanol containing 0.1 M ammonium acetate. Store in 80 % acetone at -20 °C until use.
12. Proceed for 2-DE analysis (see Note 11).

3.5 In Planta Secretome from Rice Leaf Infected with *M. oryzae* or *X. oryzae*

This method is depicted as a flow chart in Fig. 2.

1. Cut rice leaves (length of average 5 cm)-infected with/without *M. oryzae* or *X. oryzae* (approx. 50 g) with scissors (see Note 9).
2. Place cut leaves in a 250 mL centrifuge bottle containing 150 mL of CA buffer (see Note 10) and hold on a constant horizontal shaker at 100 strokes/min for 1 h on ice to extract secreted proteins.
3. Filter the extraction buffer with No.2 filter paper to remove leaf debris.
4. Centrifuge at $2,500 \times g$ for 15 min at 4 °C and transfer the buffer to a new centrifuge tube.
5. Add 50 mL of water-saturated phenol to the extracted supernatant solution, mix, and centrifuge at $5,000 \times g$ for 15 min at RT, and transfer phenolic phases to new tube. Save supernatant solution and repeat this step once more to collect enough protein combine phenol fractions.

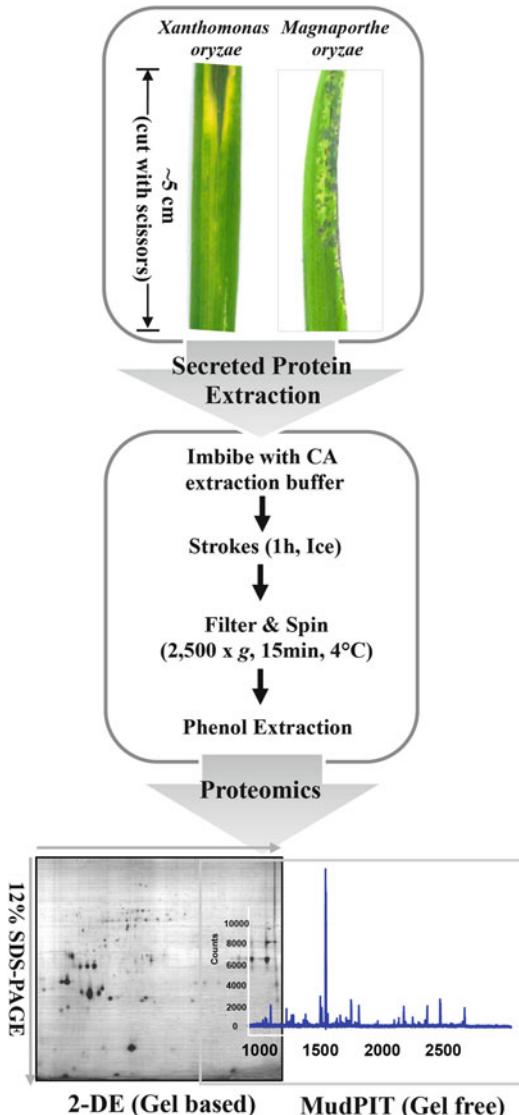


Fig. 2 Workflow of secretome analysis during rice–pathogen interaction

6. Add 1 M sucrose, vortex, and centrifuge at $5,000 \times g$ for 15 min at RT, collect supernatant phenol phase, and transfer new tube.
7. Add 4 volumes of 100 mM ammonium acetate in methanol to precipitate secreted proteins for overnight at -20°C .
8. Centrifuge at $5,000 \times g$ for 10 min, wash and resuspend precipitated protein in 80 % methanol containing 0.1 mM ammonium acetate. The washing step was repeated three times.
9. Wash with 80 % acetone and stored the precipitated secretory proteins in the same solution at -20°C .
10. Proceed for 2-DE (*see Note 11*) or MudPIT analysis (*see Note 12*).

4 Notes

1. Dry mature seeds are stored at 10 °C for long term storage.
2. Any soil can be used if recommended for rice growth and development.
3. Use clean/sterile equipment and gloves at all times for the experiments.
4. To get differential levels of expression from both host and pathogen, preparation of synchronously grown inoculum is critical. High density infection foci are helpful to get more interaction events.
5. Rice leaf samples used at 72 h after inoculation with *M. oryzae* were in transition phase of biotrophic and early necrotrophic stage for minimizing cytoplasmic contamination. Certain protein markers [Glucose-6-phosphate dehydrogenase (G6PDH), malate dehydrogenase (MDH),mannosidase, phosphoenol-pyruvate (PEP) carboxylase, and cytosolic aldolase] can be used for the purity assessment.
6. The Mg/NP-40 extraction systems produced more resolvable spots by fractionation of RuBisCO into PEG pellet, which is about 50 % of total soluble rice leaf protein [11].
7. The phenol extraction method can minimize proteolysis and optimize extraction of membrane proteins. This method results in the best resolution of proteins on 2-DE gels as removing non-protein components that interfered with IEF [12].
8. Prefractionation of protein samples using PEG prior to 2-DE can specifically fractionate RuBisCO into 15 % PEG pellet. This method can enrich low-abundance proteins. Advantages of using PEG are that proteins are fractionated under native conditions and analysis of 2.7 times more well-separated proteins compared with conventional single-step analysis [12].
9. Proteomic studies of secretory proteins have been limited in any organism, including plants. Moreover, it is yet to be determined how many proteins are found in the secretome of a given organism under normal growth and adverse environmental conditions. This *in planta* secretome method serves as a valuable resource toward developing a near complete secretome of rice-pathogen interaction.
10. The methods, which were based on VIC and gravity-extraction, have primarily been used to isolate apoplastic fluid from plant tissue (mainly leaf) for subsequent preparation of secretory proteins with little or no contamination from intracellular proteins [4]. Both methods suffer from the drawback of low recovery of secretory proteins and hence require further improvement in their workflow. To improve these problems,

we tried to use CA buffer as the extraction buffer composed of calcium chloride in acetate buffer.

11. The 2-DE, image analysis of 2-D gels, and trypsin digestion of 2-D protein spots were performed essentially as described previously [13, 14]. Briefly, the colloidal CBB stained 2-D gels were scanned (300 dpi, 16-bit gray scale pixel depth, TIFF file) for image analysis using ImageMaster 2D Platinum imaging software ver. 6.0 (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). The intensity of each spot was normalized as an average of the intensity of spots on the gel. Differentially expressed ($p < 0.05$) protein spots present in three independent biological samples were selected, digested with trypsin (Promega, sequencing grade), and subjected to MS analysis.
12. MudPIT analysis were basically performed as described previously [15] using Partisphere strong cation exchanger (SCX; Whatman, Clifton, NJ, USA) and Polaris C18-A. As peptides were eluted from the microcapillary column, they were electrosprayed into a LTQ linear ion trap mass spectrometer (ThermoFisher, CA, USA) with the application of a 2.3-kV spray voltage applied distally to the waste of the HPLC split.

Acknowledgments

This work was supported by National Research Foundation of Korea Grant funded by the Korean Government (Ministry of Education, Science and Technology) (NRF-2011-355-C00151) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science, and Technology (2011-0027211 and 2011-0018092).

References

1. Liu J, Wang X, Mitchell T et al (2010) Recent progress and understanding of the molecular mechanisms of the rice-*Magnaporthe oryzae* interaction. *Mol Plant Pathol* 11:419–427
2. Talbot NJ (2003) On the trail of a cereal killer: exploring the biology of *Magnaporthe grisea*. *Annu Rev Microbiol* 57:177–202
3. Dixon RA, Lamb CJ (1990) Molecular communication in interactions between plants and microbial pathogens. *Annu Rev Plant Physiol Plant Mol Biol* 41:339–367
4. Agrawal GK, Jwa NS, Lebrun MH et al (2010) Plant secretome: unlocking secrets of the secreted proteins. *Proteomics* 10:799–827
5. Li Q, Chen F, Sun L et al (2006) Expression profiling of rice genes in early defense responses to blast and bacterial blight pathogens using cDNA microarray. *Physiol Mol Plant Pathol* 68:51–60
6. Valent B, Khang CH (2010) Recent advances in rice blast effector research. *Curr Opin Plant Biol* 13:434–441
7. Kamoun SA (2006) Catalogue of the effector secretome of plant pathogenic oomycetes. *Annu Rev Phytopathol* 44:41–60
8. Stergiopoulos I, De Wit PJGM (2009) Fungal effector proteins. *Annu Rev Phytopathol* 47: 233–263
9. Agrawal GK, Rakwal R (2011) Rice proteomics: a move toward expanded proteome coverage to comparative and functional proteomics uncovers the mysteries of rice and plant biology. *Proteomics* 11:1630–1649

10. Jones OAH, Maguire ML, Griffin JL et al (2011) Using metabolic profiling to assess plant-pathogen interactions: an example using rice (*Oryza sativa*) and the blast pathogen *Magnaporthe grisea*. *Eur J Plant Pathol* 129: 539–554
11. Des Francs CC, Thielllement H, Vienne DD (1985) Analysis of leaf proteins by two-dimensional gel electrophoresis: protease action as exemplified by ribulose bisphosphate carboxylase/oxygenase degradation and procedure to avoid proteolysis during extraction. *Plant Physiol* 78:178–182
12. Kim ST, Cho KS, Jang YS et al (2001) Two-dimensional electrophoretic analysis of rice proteins by polyethylene glycol fractionation for protein arrays. *Electrophoresis* 22: 2103–2109
13. Kim ST, Kang YH, Wang Y et al (2009) Secretome analysis of differentially induced proteins in rice suspension-cultured cells triggered by rice blast fungus and elicitor. *Proteomics* 9:1302–1313
14. Kim ST, Kim SG, Hwang DH et al (2004) Proteomic analysis of pathogen-responsive proteins from rice leaves induced by rice blast fungus, *Magnaporthe grisea*. *Proteomics* 4:3569–3578
15. Kim H, Jo S, Song HJ et al (2007) Myelin basic protein as a binding partner and calmodulin adaptor for the BKCa channel. *Proteomics* 7:2591–2602

Chapter 39

Protein Extraction and Gel-Based Separation Methods to Analyze Responses to Pathogens in Carnation (*Dianthus caryophyllus* L)

Harold Duban Ardila, Raquel González Fernández, Blanca Ligia Higuera, Inmaculada Redondo, and Sixta Tulia Martínez

Abstract

We are currently using a 2-DE-based proteomics approach to study plant responses to pathogenic fungi by using the carnation (*Dianthus caryophyllus* L)–*Fusarium oxysporum* f. sp. *dianthi* pathosystem. It is clear that the protocols for the first stages of a standard proteomics workflow must be optimized to each biological system and objectives of the research. The optimization procedure for the extraction and separation of proteins by 1-DE and 2-DE in the indicated system is reported. This strategy can be extrapolated to other plant–pathogen interaction systems in order to perform an evaluation of the changes in the host protein profile caused by the pathogen and to identify proteins which, at early stages, are involved or implicated in the plant defense response.

Key words Carnation, Fusarium, Plant defense, Plant–pathogen interaction proteomics

1 Introduction

Carnation (*Dianthus caryophyllus* L) is one of the most important products in commercial flowers production worldwide. The researches carried out on this plant have been mainly focused on the molecular study of the processes associated to the biosynthesis of pigments and senescence [1, 2]. However, the studies at a biochemical level for this species have been too limited and there are not any established protocols that allow analysis at a proteomic level. Protein analysis using two-dimensional electrophoresis (2-DE) is one of the most used tools for the analysis of proteomes in plants and has allowed the study of a large number of biological processes such as the interactions with pathogen microorganisms [3, 4]. In recent years, the descriptive analysis of proteins involved in interactions with pathogens, using 2-DE and subsequent mass spectrometry (MS) identification, has become an important tool in

phytopathology [5]. Nevertheless, its application is limited to obtain a quality protein extract free from non-protein contaminants, such as polysaccharides and/or polyphenol compounds, which can interact in a nonspecific way with proteins, affecting their migration and therefore the resolution in the protein separation. Taking into account that the presence of such contaminants is determined partly by the correct combination of plant tissue and extraction method [6, 7], it is necessary to evaluate the different protocols that lead to find the best conditions for protein extraction when a proteomics experiment in general is carried out for the first time. It is important that the methodology used provides a high yield in the protein extraction and a high sensitivity in the staining process, because of the proteins associated to plant defense mechanisms are generally low abundant, and require the highest possible sensitivity to be detected with alternative types of staining different than the commonly used [8]. The study about defense responses in plants against a determined pathogen can be performed comparing the biochemical responses among genotypes with resistance differences to the disease [9]. Thus, the evaluation of proteins associated with defense can be done in carnation by comparing those ones that are differentially induced between susceptible and resistant varieties during the inoculation with the pathogen of interest, in this case *Fusarium oxysporum* f. sp. *dianthi* (*Fod*), the agent responsible of vascular wilting. This chapter discusses the conditions for the protein extraction and separation required to carry out the proteome analysis using gel-based proteomic techniques (one-dimensional electrophoresis (1-DE) and 2-DE), in carnation stems and roots during the infection with *Fod*. The employed workflow is described and discussed, and may be applied in other plant species.

2 Materials

Mention of specific companies or pieces of equipment does not represent an endorsement by the authors.

2.1 Plant Material

Rooted cuttings of carnation cultivars Candy and Tasman, resistant and susceptible respectively to the vascular wilting caused by *Fod*, were used for this study (supplied by the company Grupo Chia, Colombia).

2.2 Fungal Isolate

Fusarium oxysporum f. sp. *dianthi* race 2 was isolated from stems of a susceptible carnation plant with typical symptoms of vascular wilt. Typical conidia morphology was evaluated microscopically and the specie [10] and race [11] were verified using molecular markers.

2.3 Equipment

1. Lyophilizer (Labconco).
2. Analytical Balance (Mettler Toledo).
3. Vortex (Velp Scientifica).
4. Microcentrifuge (Eppendorf).
5. Ultrasonic homogenizer (Cole Palmer).
6. Spectrophotometer (Eppendorf Bio Photometer).
7. Criterion Stain Free Gel Imaging System (Bio-Rad).
8. Isoelectric Focusing System (PROTEAN IEF cell, Bio-Rad).
9. Electrophoresis systems (PROTEAN II XL, Bio-Rad).
10. FX Pro Plus Multiimager Fluorescence Scanner (Bio-Rad).
11. GS-800 Calibrated Densitometer (Bio-Rad).
12. Quantity One and PDQuest software (Bio-Rad).

2.4 Reagents, Solutions, and Buffers

Analytical grade reagents must be used, unless other grades are specified. The prepared solutions should be kept at 4 °C or at -20 °C if indicated. Reagents and solutions must be discarded once used, according to current regulations. It is mandatory to use a fume cupboard when working with volatile and dangerous compounds. Personal protection elements such as robes, gloves and glasses are of mandatory use.

1. Phenol solution equilibrated with 10 mM Tris-HCl, pH 8 (P4557, Sigma).
2. Bradford reagent (Bio-Rad).
3. ReadyStrip™ IPG strips, 11 cm, pH 3–10 L, and 17 cm, pH 3–10 L and NL (Bio-Rad).
4. Mineral oil (Bio-Rad).
5. Criterion TGX Stain-Free Precast Gels 4–20 % (Bio-Rad).
6. Precision Plus Protein Standards (Bio-Rad).
7. 30 % Acrylamide/Bis-acrylamide solution (Bio-Rad).
8. SDS-PAGE Broad Range markers 6.5–200 KDa (Bio-Rad).
9. SYPRO Ruby reagent (Bio-Rad).
10. Specialty Standard IEF (Bio-Rad).
11. Total protein extraction buffer: 0.5 M Tris-HCl, pH 8.0, 5 % (w/v) SDS, 15 % (v/v) glycerol, 100 mM dithiothreitol (DTT), 1 mM Phenylmethylsulfonyl Fluoride (PMSF).
12. 10% (w/v) trichloroacetic acid (TCA), 0.07% DL-dithiothreitol DTT (w/v) in 80 % (v/v) acetone.
13. 100 % (w/v) trichloroacetic acid (TCA) in water.
14. 0.1 M ammonium acetate in 100 and 80 % (v/v) methanol solutions.

15. Lysis solution: 30 % sucrose (w/v), 2 % SDS (w/v), 5 % β -mercaptoethanol (v/v).
16. Lysis-saturated phenol (1:1 ratio) solution: mix 1 volume of 10 mM Tris–HCl, pH 8.0 buffer-saturated phenol (Sigma[®]) with 1 volume of lysis solution.
17. Solubilization solution: 7 M urea, 2 M thiourea, 4 % (w/v) 3[(3-Cholamidopropyl)-dimethylammonium]-propanesulfonic acid (CHAPS), 0.5 % (w/v) Triton X-100 and 20 mM DTT.
18. Bovine serum albumin (BSA) solution (1 mg/mL).
19. IPG strip rehydration solution: 7 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 0.01 % (w/v) bromophenol blue, 100 mM DTT, and 0.2 % (v/v) of 3–10 ampholytes (Bio-Rad[®]).
20. IPG strip equilibrium solution: 375 mM Tris–HCl, pH 8.8, 6 M urea, 20 % (v/v) glycerol, 2 % (w/v) SDS.
21. 10 % (w/v) sodium dodecyl sulfate (SDS) solution.
22. 10 % (w/v) ammonium persulfate (APS) solution.
23. Electrophoresis stacking gel buffer: 1.5 M Tris–HCl, pH 8.8.
24. Electrophoresis running buffer: 50 mM Tris–HCl, pH 8, 192 mM Glycine, 1 % (w/v) SDS.
25. Electrophoresis sample buffer: 0.5 M Tris–HCl, pH 6.8, 10 % (v/v) glycerol, 10 % (w/v) SDS, 0.5 % (w/v) bromophenol blue, 5 % (v/v) β -mercaptoethanol.
26. Sealing solution: 0.5 % (w/v) agarose and 0.01 % (w/v) bromophenol blue.
27. Gel staining Coomassie colloidal suspension: 0.06 M $(\text{NH}_4)_2\text{SO}_4$, 20 % (v/v) methanol, 1.9 % (w/w) H_3PO_4 , 0.1 % (w/v) Coomassie G-250.
28. Gel destaining solution A: 0.1 M Tris- H_3PO_4 , pH 6.5.
29. Gel destaining solution B: 25 % (v/v) methanol.
30. Gel destaining solution C: 20 % (w/v) ammonium sulfate.
31. Gel fixing solution: 10 % (v/v) methanol, 7 % (w/v) acetic acid.
32. Gel washing solution: 20 mM ammonium bicarbonate, 50 % (v/v) acetonitrile.

3 Methods

The workflow in Fig. 1 is stated in order to choose the conditions to evaluate proteins associated to carnation defense against the pathogen responsible for vascular wilting by 2-DE. This workflow may be applied to other plant species in similar studies.

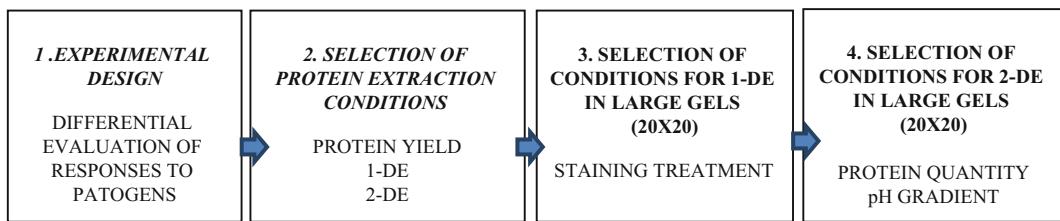


Fig. 1 Workflow to select protein extraction and separation conditions from carnation roots and stems in proteomics research

3.1 Experimental Design

The proteomic comparative analysis, as any biological experiment, must start with the experimental design. In the present work, carnation varieties with different levels of resistance to *F. oxysporum* f. sp. *dianthi*, both inoculated and non-inoculated plants, with samples collected at different times after inoculation have been used (see Note 1). The plant material must be collected in a homogeneous way (see Note 2). The infection and pathogen localization must be confirmed (see Note 3).

3.2 Making the Protein Extract

3.2.1 Procedure Based on Extraction with Buffer and Subsequent TCA Precipitation

Three different protein extraction protocols reported in the literature were tested [6]. They include the use of treatments with TCA and acetone to remove non-protein compounds and for protease inactivation. The procedures applied to carnation stems and roots are the following.

1. Transfer 20 mg of lyophilized plant material to a 2 mL Eppendorf tube (see Note 4).
2. Add 1 mL of total protein extraction buffer and shake for 30 min at 4 °C.
3. Centrifuge at 16,000×*g* for 10 min at 4 °C and transfer the supernatant to a new 2 mL tube.
4. Add 1 mL of 100 % (w/v) TCA and mix well using vortex. Stand for 2 h at 4 °C. Centrifuge at 16,000×*g* for 5 min at 4 °C. Discard the supernatant.
5. Wash the pellet (three times) with 80 % acetone.
6. Centrifuge at 16,000×*g* for 5 min at 4 °C and discard the supernatant.
7. Dry the pellet to remove acetone residues in a fume cupboard.
8. Dissolve the pellet in 200 µL of the solubilization solution and shake for 4 h at 4 °C.
9. Quantify proteins using the Bradford method [12].
10. Store the extract at -20 °C for further analysis.

3.2.2 Procedure Based**on TCA-Acetone****Precipitation**

1. Transfer 20 mg of lyophilized plant material to a 2 mL Eppendorf tube (*see Note 4*).
2. Add 500 μ L of 10 % (w/v) TCA, 0.07 % (w/v) DTT in 80 % (v/v) acetone, and sonicate (three times \times 10 s at 50 W, amplitude 60) at 4 °C. Keep on ice for 1 min.
3. Mix vigorously by vortexing.
4. Fill the tube with 10 % (w/v) TCA, 0.07 % (w/v) DTT in 80 % (v/v) acetone, and vortex.
5. Allow to precipitate overnight at -20 °C.
6. Centrifuge at 16,000 $\times g$ for 10 min at 4 °C and discard the supernatant.
7. Wash the pellet (three times) with 80 % acetone.
8. Centrifuge at 16,000 $\times g$ for 5 min at 4 °C and discard the supernatant.
9. Air-dry the pellet to completely remove acetone.
10. Dissolve the pellet in 200 μ L of the solubilization solution for 4 h at 4 °C.
11. Quantify proteins using the Bradford method [12].
12. Store the extract at -20 °C for further analysis.

3.2.3 Procedure Based**on TCA-Acetone****Precipitation and
Subsequent Extraction
with Phenol**

1. Transfer 20 mg of lyophilized plant material to a 2 mL Eppendorf tube (*see Note 4*).
2. Fill the tube with 10 % (w/v) TCA in 80 % (v/v) acetone and sonicate (three times \times 10 s at 50 W, amplitude 60, each) at 4 °C. Keep on ice for 1 min. Vortex vigorously.
3. Centrifuge at 16,000 $\times g$ for 5 min and discard the supernatant.
4. Fill the tube with 0.1 M ammonium acetate in 80 % (v/v) methanol, and mix well by vortexing.
5. Centrifuge at 16,000 $\times g$ for 5 min at 4 °C and discard the supernatant.
6. Fill the tube with 80 % (v/v) acetone and mix well by vortexing.
7. Centrifuge at 16,000 $\times g$ for 5 min at 4 °C and remove the supernatant.
8. Air-dry the pellet at room temperature to fully eliminate acetone.
9. Fill the tube with the lysis-saturated phenol (1:1 ratio) solution, homogenize using a pestle, and keep on ice for 5 min.
10. Centrifuge at 16,000 $\times g$ for 5 min at 4 °C and transfer the upper phenolic phase into a new 2 mL tube.

11. Fill the tube with 0.1 M ammonium acetate in 100 % (v/v) methanol, mix and allow precipitating overnight at -20 °C.
12. Centrifuge at 16,000×*g* for 5 min at 4 °C and discard the supernatant.
13. Wash the pellet with 100 % (v/v) methanol and mix well by vortexing.
14. Centrifuge at 16,000×*g* for 5 min at 4 °C and remove the supernatant.
15. Wash the pellet with 80 % (v/v) acetone and mix well by vortexing.
16. Centrifuge at 16,000×*g* for 5 min at 4 °C and discard the supernatant.
17. Air-dry the pellet to remove acetone.
18. Dissolve the pellet in 200 µL of the solubilization solution, and shake for 4 h at 4 °C.
19. Quantify proteins using the Bradford method [12].
20. Store the extract at -20 °C for further analysis.

Selection of the best extraction procedure will depend on protein yield, and number of bands or spots visualized and resolved in 1-DE or 2-DE gels. Protein yield was determined as µg of BSA equivalents/mg of lyophilized plant material (*see Notes 4 and 5*).

3.3 One-Dimensional Electrophoresis (SDS-PAGE) on Small Gels

Arrange the complete electrophoresis system for small gels (*see Note 6*). In this case, the Criterion Stain Free Gel system was used.

1. Remove the comb from the gels carefully and insert the system in the electrophoresis chamber according to the manufacturer's instructions. For instance, pre-cast polyacrylamide gels (4–20 % gradient) can be used (*see Note 7*).
2. Add enough electrophoresis buffer to the electrophoresis chamber to the top of the gel and wash the wells with the same solution.
3. Prepare the samples using the different extracts obtained from different methods. Take a volume containing 15 µg of protein and adjust with water to a volume of 10 µL. Add 5 µL of loading buffer and mix by vortexing.
4. Simultaneously, prepare the molecular weight markers (Precision Plus Protein Standards).
5. Heat the sample and markers for 5 min at 100 °C, and load them in each of the wells carefully.
6. Close the electrophoresis chamber and set at 150 V for running. Unplug the power supply and remove the gel carefully when the electrophoresis is finished.

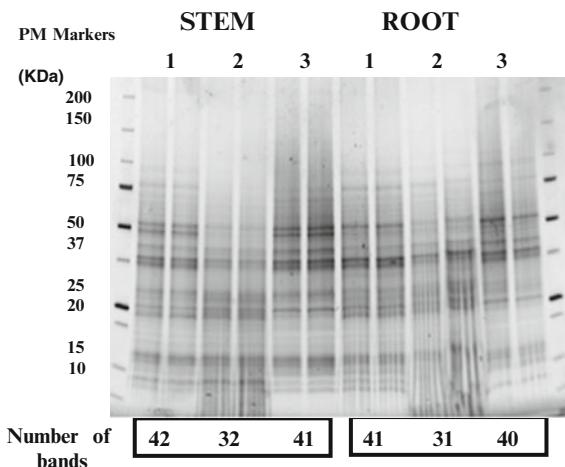


Fig. 2 One-DE of protein extracts obtained using different extraction protocols from carnation stems and roots infected with *Fod*. Criterion Free Stain electrophoresis system Bio-Rad® was used. Extraction and precipitation with TCA (1); TCA-acetone method (2); TCA-acetone-phenol method (3)

7. Take the gel image using Criterion Stain Free Gel Imaging System (see Note 8).
8. Evaluate the number of bands for each of the tested treatments and in each plant organ under study, using the Quantity One® software from Bio-Rad (Fig. 2). Compare the obtained results in number of bands and its resolution (see Note 9).

3.4 Two-Dimensional Gel Electrophoresis on Short Strips

Perform 2-D electrophoresis of different amounts of protein in the range of 100–200 µg for 11 cm-long IPG strips (pH 3–10). The isoelectric focusing (IEF) (first dimension) is carry out in the PROTEAN® IEF cell system. For the second dimension, pre-cast gels can be used with the Criterion system (see Note 10).

1. Add protein extract obtained from the tissue under study in independent 1.5 mL tubes, in equivalent amounts of 100 and 200 µg of protein. Add IPG strip rehydration solution to a final volume of 185 µL. Shake vigorously by vortexing (see Note 11).
2. Pour gently the previously prepared mixture creating a line in one of the rails of the IPG strip focusing tray, avoiding the formation of bubbles.
3. Place an 11 cm-long IPG strip with a 3–10 pH range, ensuring that it is properly placed to avoid the entry of air bubbles between the strip and the sample. The position of the strip must allow its end marks to coincide with the position of the (+, anode) and (–, cathode) electrodes.

4. Add 1–2 mL of mineral oil to prevent dehydration during the rehydration process and cover with the plastic faceplate of the system.
5. Perform the rehydration process placing the IPG strip focusing tray in the focusing system for 16 h applying 50 V at 20 °C (active re-hydration).
6. Once rehydration is finished, place paper strips moistened with deionised water between the strip and the electrode using clamps, towards the end corresponding to the anode (+).
7. Program the PROTEAN IEF cell system for the isoelectric focusing according to the instructions of the system. For these samples, a program at 20 °C with a final voltage of 8,000 V (30,000 V·h) was used. The maximum current per strip must be 50 µA.
8. Once the isoelectric focusing is finished, remove the strips from the system and keep them vertically to dispose of most of the mineral oil. Strips must be placed keeping the gel on the top in the strip-holding system (rehydration/equilibration tray) with the same order in which they were removed from the system (*see Note 12*).
9. At the time of performing the 2-DE, gently remove the comb from the gels and insert the system in the electrophoresis chamber.
10. Add enough running buffer to the electrophoresis chamber to the top of the gel and wash the well where the strip will be placed.
11. Wash the strip twice using the running buffer directly on the strip-holding system (rehydration/equilibration tray).
12. Equilibrate each strip with 2–3 mL of a IPG equilibrium solution containing 2 % (w/v) DTT for 15 min while shaking. Subsequently, perform the same procedure but using an equilibrium buffer with 2 % (w/v) iodoacetamide. Finally, wash with enough running buffer to remove the remnant equilibrium solution.
13. Place the strip on top of the gel using clamps, avoiding accumulation of bubbles between the strip and the gel. This procedure must be done preventing tearing of the strip and allowing the pH 3 end to be located on the left end of the gel.
14. Add 2 µL of weight marker (Precision Plus Protein Standards) in each of the gel wells chosen for the markers.
15. Seal the system gently using the sealing solution and avoiding the formation of bubbles. In order to do so, heat the solution in a microwave oven and add 1–2 mL of it on the strip, previously positioned, until reaching the top of the glass. Use a plastic Pasteur pipette.
16. Close the electrophoresis chamber and perform the electrophoresis procedure at 150 V.

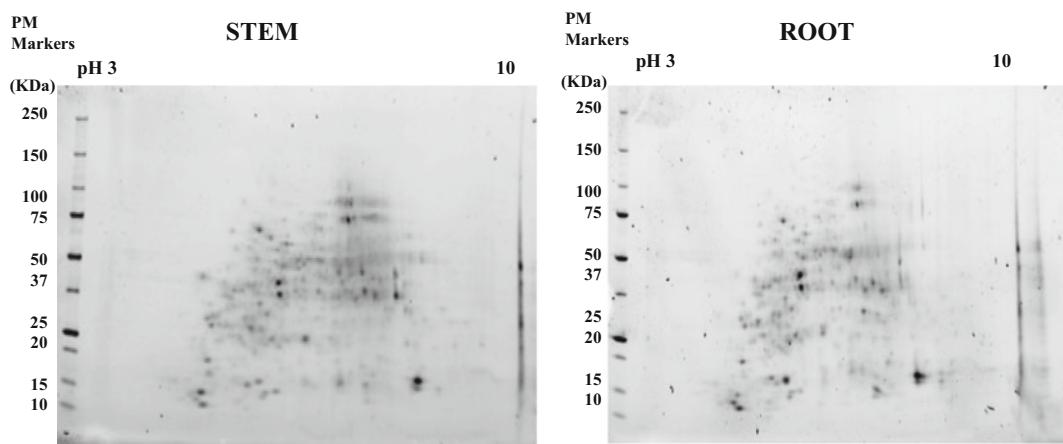


Fig. 3 Two-DE of protein extracts obtained using TCA–acetone–phenol protocol from carnation stems and roots infected with *Fod* using Criterion-Free Stain system Bio-Rad®

17. Switch off the power supply and remove the gel carefully, when the electrophoresis is finished.
18. Perform the visualization using the Criterion Stain Free Gel Imaging System and capture the corresponding image.
19. Evaluate the gel image in terms of number of bands, intensity and resolution (Fig. 3). The presence of striking indicates the existence of non-protein contaminants which are affecting the separation (see Note 13).

3.5 SDS-PAGE on Large Gels

The selection of a staining method for obtaining the best sensitivity in bands detection is necessary when a 1-DE in large gels (20×20 cm) analysis for a plant extract is done for the first time. A comparative analysis using different gel staining methods must be done, once the protein extraction method had been selected.

1. Arrange the complete electrophoresis system for 20×20 gels. In this case, the PROTEAN II XL Bio-Rad system was used (see Note 6).
2. Prepare the separation gel (see Note 14).
3. Prepare the concentration gel (see Note 15).
4. Place the comb into the gel cassette assembly and when ready to pour the gel, transfer the gel solution between the glass plates using a pipette.
5. Once the gel polymerization has finished place the gel cassette assembly into the electrophoresis chamber according to the manufacturer's instructions.
6. Gently remove the comb out of the gel and add enough running buffer into the electrophoresis chamber. Using a pipette, wash the wells with the same solution.

7. Prepare the samples using a sample volume containing 100 µg of protein and add water to a volume of 20 µL. At this point, add 10 µL of loading buffer and mix by vortexing.
8. Simultaneously, prepare the molecular weight markers (SDS-PAGE Broad Range markers). Heat the samples and markers for 5 min at 100 °C and load them in each of the wells carefully.
9. Close the electrophoresis chamber and set at 100 V for running overnight at 4 °C. Unplug the power supply and remove the gel carefully when the electrophoresis is finished.
10. Stain the gels according to the protocols available in your laboratory. In this case two protocols were evaluated: SYPRO Ruby staining and SYPRO Ruby–Coomassie colloidal staining (*see Note 16*).

3.5.1 SYPRO Ruby Staining

1. Wash the gel for 30 min with the fixing solution.
2. Gently remove the fixing solution and add SYPRO® Ruby Bio-Rad. Shake overnight in the dark (16–18 h).
3. Wash the gel with the fixing solution for 30–60 min.
4. Wash the gel using lots of water.
5. Capture the image in the Fluorescence Scanner device (FX Pro Plus Multiimager, Bio-Rad).

3.5.2 SYPRO Ruby–Coomassie Colloidal Staining

1. Remove the SYPRO Ruby from the gel. In order to do so, wash it 3 to 4 h with the gel washing solution.
2. Wash with ddH₂O for 30 min.
3. Add Coomassie colloidal solution and keep the gel for 60 h in continuously shaking.
4. Wash the gel using destaining solution A for 3 min.
5. Wash using destaining solution B for 1 min.
6. Wash using destaining solution C at least 24 h.
7. Capture the image in the calibrated densitometer (GS-800 Calibrated Densitometer, Bio-Rad).

3.5.3 Gel Image Analysis

Compare the number of bands for each of the tested treatments using the Quantity One® software from Bio-Rad.

1. Load the images in the Quantity One® Bio-Rad software format. Images must share the same format and resolution.
2. Perform the automatic detection of bands on the different gels under the same sensitivity conditions. Verify that each band detected corresponds to one band properly.
3. Compare the number of detected bands and selected the more sensitive method (*see Note 17*) (Fig. 4).

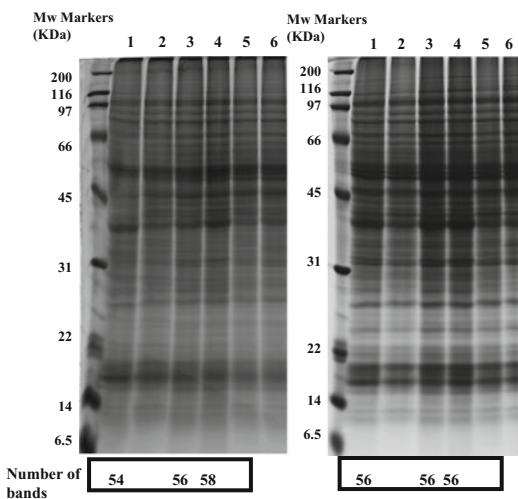


Fig. 4 Staining comparison protocols. SYPRO Ruby (*left*) and SYPRO Ruby + Coomassie colloidal (*right*). Analysis in 1-DE (20×20 cm) for protein extracts from carnation roots infected with *Fod* in different hours post-inoculation (hpi). Resistant control 0 hpi (1), Susceptible control 0 hpi (2), Resistant control 6 hpi (3), Resistant inoculated 6 hpi (4), Susceptible control 6 hpi (5), Susceptible inoculated 6 hpi (6). The number of bands for treatment 1, 3, and 4 is shown in the *bottom* of the graph

3.6 Two-Dimensional Gel Electrophoresis on Large Strips

3.6.1 Amount of Protein

Perform 2-DE of different amounts of protein in the range of 200–400 μ g for 17 cm-long IPG strips (pH 3–10). In order to evaluate the separation resolution for the first dimension, compare lineal and no lineal IPG strips (pH 3–10) (*see Note 18*).

1. Add protein extract obtained from the tissue under study in equivalent amounts of 200 and 400 μ g of protein in different 1.5 mL tubes. Add IPG strip rehydration solution to a final volume of 400 μ L. Shake vigorously by vortexing (*see Note 11*).
2. Pour gently the previously prepared mixture creating a line in one of the rails of the IPG strip focusing tray, avoiding the formation of bubbles. Repeat the procedure for each sample.
3. Place in each rail a 17 cm-long IPG strip with a 3–10 pH linear range, ensuring that it is properly placed to avoid the entry of air bubbles between the strip and the sample. The position of the strip must allow its end marks to coincide with the position of the (+, anode) and (–, cathode) electrodes.
4. Add 1–2 mL of mineral oil to prevent dehydration during the rehydration process and cover with the plastic faceplate of the system.
5. Perform the rehydration process placing the IPG strip focusing tray in the focusing system for 16 h applying 50 V at 20 °C (active rehydration).

6. Once rehydration is finished, place paper strips moistened with deionised water between the strip and the electrode using clamps, towards the end corresponding to the anode (+).
7. Program the PROTEAN IEF cell system for the isoelectric focusing according to the instructions of the system. For these samples, a program at 20 °C with a final voltage of 10,000 V (60,000 V·h) was used. The maximum current per strip must be 50 µA.
8. Once the isoelectric focusing is finished, remove the strips from the system and keep them vertically to dispose of most of the mineral oil. Strips must be placed keeping the gel on the top in the strip-holding system (rehydration/equilibration tray) with the same order in which they were removed from the system (*see Note 12*).
9. At the time of performing the 2-DE, wash the strip twice using the electrophoresis running buffer directly on the strip-holding plate (rehydration/equilibration tray).
10. Equilibrate each strip with 5 mL of IPG equilibrium solution containing 2 % (w/v) DTT for 15 min with gentle agitation. Subsequently, perform the same procedure but using an equilibrium buffer with 2 % (w/v) iodoacetamide. Finally, wash with a running buffer to remove the remnant equilibrium solution.
11. Prepare a 12 % resolving polyacrylamide gel (*see Note 14*).
12. Insert the gel cassette assembly to the electrophoresis chamber according to the manufacturer's instructions and add enough electrophoresis buffer to the top of the gel.
13. Place the strip previously equilibrated in **steps 9–11**, on top of the gel using clamps, avoiding the accumulation of bubbles between the strip and the gel. This procedure must be done preventing tearing of the strip and allowing the pH 3 end to be located on the left end of the gel.
14. Gently moisten a piece of filter paper (10 × 5 mm) with 5 µL of weight marker (SDS-PAGE Broad Range markers Bio-Rad®) and place it on the left side of the strip.
15. Seal the system gently using the sealing solution and avoiding the formation of bubbles. In order to do so, heat the solution in a microwave oven and add 1–2 mL of it on the strip, previously positioned, until reaching the top of the glass. Use a plastic Pasteur pipette.
16. Close the electrophoresis chamber and perform the electrophoresis procedure at 100 V.
17. Switch off the power supply and remove the gel carefully, when the electrophoresis is finished.
18. Perform the staining procedure according to the previously selected protocol. For carnation extracts, SYPRO Ruby–Coomassie colloidal staining must be used (Subheading **3.3**).

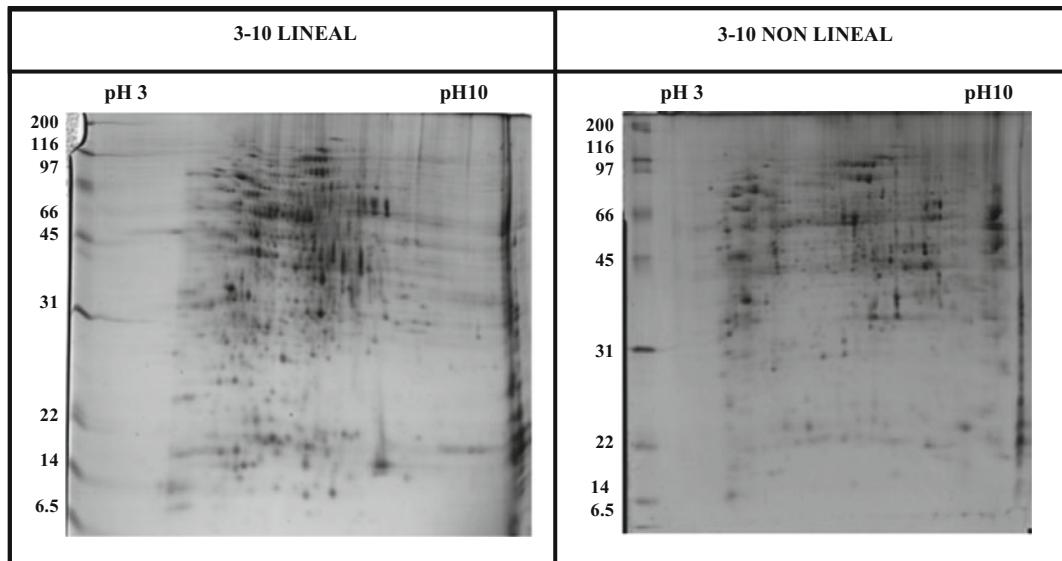


Fig. 5 Two-DE in large gels (20 × 20 cm) of extracts obtained from carnation stems using different pH gradient in the first dimension separation: *left* (Lineal) and *right* (non lineal)

19. Capture the image using the calibrated densitometer (GS-800 Calibrated Densitometer, Bio-Rad).
20. Compare visually the resolution and the number of present spots for both amounts of evaluated proteins, in each organ of the plant (*see Note 19*).

3.6.2 pH Gradient

1. Add in different 1.5 mL tubes extract equivalent to 400 µg of protein and complete to 400 µL using IPG strip rehydration solution. Shake vigorously by vortexing (*see Note 11*).
2. Pour gently the previously prepared mixture creating a line in one of the rails of the IPG strip focusing tray, avoiding the formation of bubbles. Repeat using a clean rail for each sample.
3. Place for each sample, 17 cm-long IPG strips with different pH range gradient. In this case, pH 3–10 IPG strips with lineal and no lineal gradient were used (*see Note 20*). Avoid trapping air bubbles between the strip and the sample. The position of the strip must allow its end marks to coincide with the position of the (+, anode) and (–, cathode) electrodes.
4. Repeat the procedure previously described from **steps 4 to 19** in Subheading **3.6.1**, for IPG strip rehydration, electro focusing, second dimension separation and staining.
5. Compare the resolution achieved in the separation between the evaluated gradients. At this point, it is noticed that the best separation for extracts obtained from carnation tissues is obtained when a gradient from 3 to 10 NL is used (Fig. 5).

6. Additional assays must be done for isoelectric point assignation when the best separation is achieved with the nonlinear gradient. Add 5 μ L of the commercial markers (Bio-Rad® Specialty Standard IEF) (see Note 21), and complete to 400 μ L using IPG strip rehydration solution. Shake vigorously by vortexing.
7. Pour gently the previously prepared mixture creating a line in one of the rails of the IPG strip focusing tray, avoiding the formation of bubbles. Repeat the procedure previously described from step 3 to 7 in Subheading 3.6.1 for rehydration and isoelectrofocusing.
8. Once the isoelectric focusing is finished, remove the strip from the system and keep it vertically to dispose most of the mineral oil. Add coomassie colloidal solution to the IPG strip directly according to numeral 3–6 in Subheading 3.5.2.
9. Measure using a rule, the position from the left end mark (pH 3) to each band separated on the previously stained strip. Assign the isoelectric point for each band according to this measured position and the marker composition.
10. Compare the position of these bands and the position of most intense spots in a 2-D gel obtained using the same type of IPG strip NL. Assign isoelectric points for these representative spots according to its displacement from the pH 3 end mark on the gel and the position of the bands in the strip.
11. Assign the pI for the most intense and representative spots according to the described comparison during the PDQuest Analysis. The assignment of pI for the other proteins is made automatically.

The extraction and separation (1-DE and 2-DE) conditions described above were required to carry out the proteome analysis in carnation stems and roots during the infection with *Fod*, using gel-based proteomic techniques, as different authors have previously described in other plant-pathogen interactions [4, 5, 13, 14].

3.7 Downstream Steps in the Workflow and Main Results

MS analysis and protein identification is currently carried out at the SCAI (Servicio Central de Apoyo a la Investigación, Universidad de Córdoba).

1. Spots are cut from the gel using a station ProPic Investigator (GenomicSolution).
2. The resulting gel fragments are subjected to digestion with trypsin and the resulting peptides analyzed by MALDI-TOF/TOF.
3. The identification is based on comparing the MS fragmentation patterns and corresponding MS/MS, on the NCBIInr database using as search engine MASCOT (MatrixScience), and limiting the taxonomic category to plants (*Viridiplantae*).

With the above procedure were identified 25 from 78 variable spots. This low rate of identification is due to the lack of reported sequences for carnation in the NCBInr. A specific species database construction using results of large-scale sequencing of RNA and DNA, as well as its corresponding annotation, is necessary in order to development researches in plant species poorly studied at proteomic level as carnation. Despite that, we have identified some constitutive proteins associated with the resistant phenotype, such as a class III peroxidase and a NB-ARC resistance protein (Nucleotide Binding domain shared by Apaf-1, certain R gene products, and CED-4 fused to C-terminal leucine-rich repeats). Likewise, a differential protein accumulation associated to carbohydrate metabolism reorganization was found in the resistant cultivar during pathogen infection at root level. The next step will be focused to develop a specific database for identifying those non-identified remaining proteins which are differentially accumulated in the resistant cultivar, and finally elucidate the principal earliest resistant mechanisms acting in carnation against *Fusarium oxysporum* f. sp. *dianthi*.

4 Notes

1. The most appropriate experimental approach should be selected according to the type of pathogen and previous studies about the plant–pathogen interaction model. Considering previous studies with *F. oxysporum* f. sp. *dianthi*, an in vivo inoculation assay with the pathogen was chosen [15].
2. At least two cultivars differing in terms of resistance to the pathogen must be selected for the in vivo inoculation assay. A certain number of plants must be used to obtain, at least, biological triplicates for each treatment during each sampling [16].
3. Inoculation conditions and sampling timing must be selected considering the nature of the pathogen and the type of disease. In this case, the previously reported conditions are followed [15], but sampling is done at early times within the first 4 days (0, 6, 12, 24, 48, and 96 post-inoculation hours). It is important to consider the local or systemic nature of the host plant response. A completely randomized experimental design using treatments including no inoculated controls becomes a suitable alternative for this type of studies in plants. In order to do this, it is recommended to use the following treatments: T1, Resistant Control; T2, Resistant Inoculated; T3, Susceptible Control; T4, Susceptible Inoculated.
4. Non-lyophilized material can be used for extraction. In this case, report the results as mg of protein per g of fresh vegetable material, using amounts ten times larger than for lyophilized material due to the high moisture level in vegetable samples (around 90 %). Different vegetal material amounts (lyophilized

or not lyophilized) can be assayed according to the vegetable material available.

5. Compare particular results obtained from your system. According to the evaluation performed in carnation tissues, TCA–acetone and TCA–acetone–phenol extraction methods have the best results in terms of protein yield. However, it is necessary to evaluate the quality of the extracted proteins using 1-DE. Therefore, this technique can show whether non-protein contaminant extraction has occurred or not.
6. Clean all the parts of the electrophoresis chamber, glasses, splitters and combs carefully.
7. For other electrophoretic system, general conditions to perform the 1-DE are similar to the ones introduced here, but following particular specifications provided by the particular electrophoretic system provider. If gels must be prepared, a discontinuous system with 12 % acrylamide in the separation gel is recommended.
8. For other electrophoretic systems with no visualization system, a Coomassie colloidal staining described later in Subheading 3.5.2 from step 2 is recommended.
9. A rapid evaluation can be performed visually by counting bands on each rail. Likewise, the presence of contaminants in the extracts can be evaluated by means of visual inspection of the gels. For this particular case, the TCA–acetone–phenol method provides the best results in terms of extraction, number of bands, and extracts cleaning.
10. The first dimension separation using IPG strips is a critical point during 2-DE, so the presence of non-protein contaminants affects its resolution and efficiency. Therefore, 2-DE separation is a decisive factor for selecting an extraction method for proteomic approaches.
11. Final volume varies based on the dimensions of the strips reported by the manufacturer.
12. Strips can be kept at –20 °C before their use.
13. In this point, the pI range distribution for the separated proteins from a particular extract can be evaluated preliminary. For extracts obtained from carnation, proteins are distributed in a 4–10 range principally. According to these results, the TCA–acetone–phenol method for protein extraction is adequate for proteomic approaches in carnation tissues.
14. Mix 20.2 mL of ddH₂O, 24 mL of acrylamide-bisacrylamide Bio-Rad, 15 mL of Tris–HCl pH 8.8 buffer, 600 µL of 10 % (w/v) SDS, and 312 µL of 10 % (w/v) APS. Stir the corresponding solution with a magnetic mixer and add 31 µL of TEMED to start the polymerization process. Quickly, add this solution to the cassette assembly to 2 cm below the top of the smaller glass. Leave

some mL in the beaker for polymerization control. Add some 2-propanol mL to straighten the level of the gel and remove unwanted air bubbles. Once polymerization is finished, remove the isopropanol on the top of the gel carefully and wash with lots of water. Try to remove the residual water excess using filter paper.

15. Mix 10.5 mL of ddH₂O, 3 mL of acrylamide-bisacrylamide Bio-Rad, 4.5 mL of Tris-HCl pH 8.8 buffer, 180 μ L of 10 % (w/v) SDS, and 90 μ L of 10 % (w/v) APS. Stir the corresponding solution with a magnetic mixer and add 18 μ L of TEMED to start the polymerization process.
16. Coomassie colloidal staining shows less sensitivity than SYPRO Ruby–Coomassie colloidal staining. A higher number of detected bands with the same amount of protein appeared with this combined method.
17. SYPRO Ruby and SYPRO Ruby–Coomassie colloidal staining shows similar sensitivity when compared to the number of detected bands for the same rail. However, SYPRO Ruby® Bio-Rad staining is unstable and can be only visualized with the system based on fluorescence detection. By contrast, SYPRO Ruby–Coomassie colloidal staining has more stability and allows the evaluation of bands in a visual way. Under these criteria, the combined SYPRO Ruby® Bio-Rad–Coomassie colloidal staining is selected for carnation extract analysis.
18. The protein amount and the pH gradient on the IPG strips used in the IEF are determinant factors affecting sensitivity and resolution in 2-DE. Generally, a greater amount of protein allows the visualization of a higher number of spots when low-abundance proteins are present in the extract. Furthermore, increasing the resolution in the first pI dimension allows noticing a higher number of spots.
19. A high amount of protein does not necessarily lead to better results. For some vegetal tissues, when using a high amount of protein, the threshold of contaminants can be exceeded and the resolution can be affected during first dimension separation. Considering, the number of spots and the resolution, the amount of protein for 2-DE selected is 400 μ g.
20. No additional gradients were evaluated considering the presence of spots in the whole range (pH 3–10), according to 2-D gels electrophoresis on short strips (Subheading 3.4). Other separation ranges can be used (for example, pH 5–8) for increase the resolution in the first dimension separation when other tissues or species are studied.
21. Other commercial protein IEF markers or pure reference proteins can be used. This is the case of the standard 2D-SDS PAGE Bio-Rad®, which permits the assignment of isoelectric points in a complex sample when the marker is simultaneously separated in a 2-D gel under the same conditions.

Acknowledgments

This work was supported by the University of Córdoba (AGR-0164: Agroforestry and Plant Biochemistry and Proteomics Research Group), the Regional Government of Andalusia (Junta de Andalucía), Universidad Nacional de Colombia (Bogotá), and COLCIENCIAS (Colombia).

References

1. Fukui Y, Tanaka Y, Kusumi T et al (2003) A rationale for the shift in colour towards blue in transgenic carnation flowers expressing the flavonoid 3',5'-hydroxylase gene. *Phytochemistry* 63:15–23
2. Tanase K, Onozaki T, Satoh S et al (2008) Differential expression levels of ethylene biosynthetic pathway genes during senescence of long-lived carnation cultivars. *Postharvest Biol Technol* 47:210–217
3. Thurston G, Regan S, Rampitsch C et al (2005) Proteomic and phosphoproteomic approaches to understand plant–pathogen interactions. *Physiol Mol Plant P* 66:3–11
4. Palomares-Rius J, Castillo P, Navas-Cortés J et al (2011) A proteomic study of in-root interactions between chickpea pathogens: the root-knot nematode *Meloidogyne artiellia* and the soil-borne fungus *Fusarium oxysporum* f. sp. *ciceris* race 5. *J Proteomics* 74:2034–2051
5. Quirino BF, Candido ES, Campos PF et al (2010) Proteomic approaches to study plant–pathogen interactions. *Phytochemistry* 71:351–362
6. Wang W, Tai F, Chen S (2008) Optimizing protein extraction from plant tissues for enhanced proteomics analysis. *J Sep Sci* 31: 2032–2039
7. Maldonado AM, Echevarría-Zomeño S, Jean-Baptiste S et al (2008) Evaluation of three different protocols of protein extraction for *Arabidopsis thaliana* leaf proteome analysis by two-dimensional electrophoresis. *J Proteomics* 71:461–472
8. Lopez MF, Berggren K et al (2000) A comparison of silver stain and SYPRO Ruby Protein Gel Stain with respect to protein detection in two-dimensional gels and identification by peptide mass profiling. *Electrophoresis* 20:3673–3683
9. Wise RP, Moscou MJ, Bogdanove AJ et al (2007) Transcript profiling in host-pathogen interactions. *Annu Rev Phytopathol* 45: 329–369
10. Abd-Elsalam KA, Aly I, Abdel-Satar MA et al (2003) PCR identification of *Fusarium* genus based on nuclear ribosomal-DNA sequence data. *Afr J Biotechol* 2:82–85
11. Chiocchetti A, Bernardo I, Daboussi MJ et al (2009) Detection of *Fusarium oxysporum* f. sp. *dianthi* in carnation tissue by PCR amplification of transposon insertions. *Phytopathology* 89:1169–1175
12. Ramagli LS, Rodriguez LV (1985) Quantitation of microgram amounts of protein in two-dimensional polyacrylamide gel electrophoresis sample buffer. *Electrophoresis* 6:559–563
13. Gonzalez-Fernandez R, Prats E, Jorrin-Novo J (2010) Proteomics of plant pathogenic fungi. *J Biomed Biotechnol* 2010:932527. doi:[10.1155/2010/932527](https://doi.org/10.1155/2010/932527)
14. Gonzalez-Fernandez R, Jorrin-Novo JV (2012) Contribution of proteomics to the study of plant pathogenic fungi. *J Proteome Res* 11:3–16
15. Higuera BL, de Montes GV (1996) Contribution of HPLC to the study of the defence mechanisms acting in carnation (*Dianthus caryophyllus* L.) roots on infection with *Fusarium oxysporum* f. sp. *dianthi*. *J High Resolut Chromatogr* 19:706–708
16. Valledor L, Jorrin-Novo J (2011) Back to the basics: Maximizing the information obtained by quantitative two dimensional gel electrophoresis analyses by an appropriate experimental design and statistical analyses. *J Proteomics* 74:1–18

Part VII

Posttranslational Proteomics

Chapter 40

StageTip-Based HAMMOC, an Efficient and Inexpensive Phosphopeptide Enrichment Method for Plant Shotgun Phosphoproteomics

Hirofumi Nakagami

Abstract

Phosphopeptide enrichment is the most critical step for successful LC-MS/MS-based shotgun phosphoproteomics. Recent technological improvements have made selective phosphopeptide enrichment from non-fractionated whole cell lysate digests with a single-step procedure possible. Here, a handy protocol is described for phosphopeptide enrichment from plant materials using hydroxy acid-modified metal oxide chromatography (HAMMOC) with a stop-and-go-extraction tip (StageTip).

Key words Shotgun phosphoproteomics, Plant, Phosphorylation, Phosphopeptide, HAMMOC, Hydroxy acid-modified metal oxide chromatography, Titania, StageTip, Stop-and-go-extraction tip

Abbreviations

| | |
|----------|--|
| ACN | Acetonitrile |
| DTT | Dithiothreitol |
| FASP | Filter-aided sample preparation |
| HAMMOC | Hydroxy acid-modified metal oxide chromatography |
| IAM | Iodoacetamide |
| IMAC | Immobilized metal-ion affinity chromatography |
| LC-MS | Liquid chromatography-mass spectrometry |
| LysC | Lysyl endopeptidase |
| MOC | Metal oxide chromatography |
| PTM | Posttranslational modification |
| PTS | Phase-transfer surfactant |
| SCX | Strong cation exchange chromatography |
| StageTip | Stop-and-go-extraction tip |
| TFA | Trifluoroacetic acid |

1 Introduction

Posttranslational modifications (PTMs) represent the most common mechanism by which protein functions can be altered. Cellular signaling networks usually utilize PTMs to transmit signals. To completely understand the molecular mechanisms of signaling pathways, and to isolate signaling factors, it is necessary to monitor the PTM status of proteins during signal transduction. Proteomics is one of the best available tools for studying PTMs. Therefore, it is well suited for the analysis of unknown signaling pathways.

Among the several PTMs described thus far, phosphorylation is the most extensively studied, and it has been shown to play a role in a wide range of physiological responses [1]. Advances in liquid chromatography-mass spectrometry (LC-MS)-based techniques, together with complete genome sequencing, have paved the way for high-throughput, large-scale, proteomics analysis (shotgun proteomics). However, the current proteomics technology based on large-scale analysis of PTMs requires several improvements. Typically, in shotgun proteomics, trypsinized peptides derived from crude protein samples are analyzed, but the number of peptides with a single PTM represents only a very small fraction of the total number of peptides produced, so very few peptides with PTMs are identified. One of the key requirements for successful PTM-oriented proteomics (modified proteomics) is to establish efficient enrichment methods for posttranslationally modified peptides or proteins (Fig. 1).

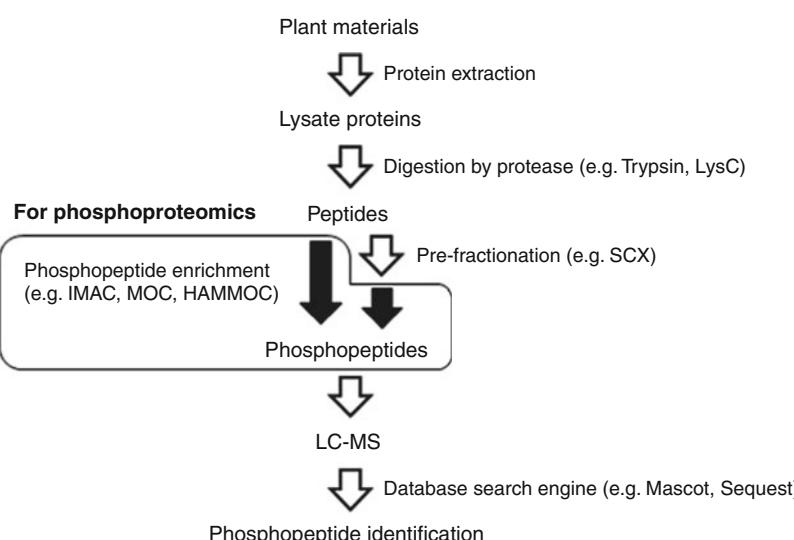


Fig. 1 Typical workflow in shotgun proteomics

Table 1
Shotgun phosphoproteomics studies in plants

| Plant material | Identified phosphopeptide | Phosphopeptide enrichment | Reference |
|--------------------------------------|---------------------------|---------------------------|-----------|
| Arabidopsis cell (membrane fraction) | 283 | SCX-IMAC | [21] |
| Arabidopsis cell (membrane fraction) | 1,172 | SCX-MOC | [22] |
| Arabidopsis cell (membrane fraction) | ? | SCX-IMAC | [23] |
| Arabidopsis cell | 2,597 | HAMMOC | [12] |
| Arabidopsis plant | 3,029 | SCX-IMAC | [24] |
| | | SCX-MOC | |
| Medicago root | 3,457 | SCX-IMAC | [25] |
| Arabidopsis cell | ca. 500 | MOC | [26] |
| Rice cell | 6,919 | HAMMOC | [11] |
| Arabidopsis plant | 4,675 | HAMMOC | [27] |
| Arabidopsis plant | 3,589 | IMAC | [28] |
| | | MOC | |
| Arabidopsis root | 849 | HAMMOC | [29] |
| | | IMAC | |
| Soybean seed | 2,001 | HAMMOC | [30] |
| Rapeseed seed | | IMAC | |
| Arabidopsis seed | | | |
| Medicago root | 15,335 | SCX-IMAC | [31] |
| Arabidopsis pollen | 598 | SCX-IMAC | [32] |
| | | SCX-MOC | |
| Maize leaf | 3,664 phosphosite | SCX-IMAC | [33] |
| Soybean root hair | 1.625 | Ni-NTA | [34] |

Phosphopeptide enrichment is one of the most rapidly developing methods in the field of modified proteomics. The development of phosphopeptide enrichment methods enabled shotgun proteomics-based phosphoproteomics (shotgun phosphoproteomics), by which thousands of phosphoproteins and their modified sites could be identified simultaneously (Table 1) [2]. Currently, immobilized metal-ion affinity chromatography (IMAC) and metal oxide chromatography (MOC), which are based on the affinity of metals for phosphate, are the most widely accepted approaches for phosphopeptide enrichment [3, 4]. Original protocols for IMAC

and MOC had the serious handicap of contamination with non-phosphorylated acidic peptides that also have affinity for the metals [5–7]. Therefore, strong cation exchange chromatography (SCX) was often used in combination with IMAC or MOC, because SCX provides a crude separation of phosphopeptides from non-phosphopeptides [8]. Recently, improved IMAC or MOC with enhanced specificity against phosphopeptides allowed for the efficient enrichment of phosphopeptides by IMAC or MOC alone, without the SCX pre-fractionation. The advent of a simple, single-step method brought phosphoproteomics within reach for people who are otherwise unfamiliar with shotgun proteomics. Still, pre-fractionation is very useful strategy for covering the phosphoproteome in depth. Selectivity of the different phosphopeptide enrichment methods differs significantly, and thus some combination of methodology was effective for improving coverage [9]. For tyrosine phosphoproteomics, antibodies which specifically recognize tyrosine-phosphorylated peptides are available [10]. Combination approaches are favored for greater coverage, but more LC-MS runs and larger samples are required for analysis. Thus, the choice of an appropriate enrichment strategy requires careful consideration.

Hydroxy acid-modified metal oxide chromatography (HAMMOC) has been successfully applied for analyzing phosphoproteomes of various organisms, including plants [11–15]. The principle of HAMMOC is to reduce non-phosphorylated acidic peptide binding, and is similar to the blocking step of Western blot analysis in that it reduces nonspecific antibody binding [14]. Hydroxy acids binding to metal oxides is weaker than to a phosphate group, but stronger than to the carboxylic groups of non-phosphorylated acidic peptides. Therefore, addition of the hydroxy acids during affinity binding steps can suppress binding of non-phosphorylated acidic peptides, but not phosphopeptides, to metal oxides. Titania and zirconia with lactic acid and β -hydroxypropanoic acid, respectively, are effective combinations for selective phosphopeptide enrichment [14].

In general, phosphopeptide enrichment methods are very simple, and are routinely utilized in laboratories where these methods were developed. However, very often these methods do not work well in other laboratories for unknown reasons. Here, a protocol for simple and robust phosphopeptide enrichment from plant materials based on titania utilized HAMMOC is described in detail so that it can be reproduced in any laboratory. The handling of titania beads is the key consideration for successful phosphopeptide enrichment. The use of self-producible stop-and-go-extraction tips (StageTip) makes phosphopeptide sample preparation for LC-MS/MS analysis economically affordable [16–18].

2 Materials

Prepare all solutions using ultrapure water and reagents with highest purity or grade. Use low-binding tips and tubes to minimize peptide loss caused by adsorption to the plastic, but do not use surface coating materials (e.g., siliconized) which may contaminate samples. Take maximum care to avoid keratin contamination.

2.1 Protein Extraction and Digestion

1. Mortar and pestle, or bead mill.
2. Liquid nitrogen.
3. Tris buffer: 10 mM Tris–HCl, pH 9.0. Store at 4 °C.
4. Extraction buffer (Prepare just before use): 8 M Urea in Tris buffer, phosphatase inhibitors (*see Note 1*). For 100 ml buffer preparation, weigh 48.048 g of urea and add the Tris buffer to a final volume of 100 ml. Add phosphatase inhibitors.
5. Protein quantification reagents (e.g., BCA Protein Assay Kit; Pierce).
6. Ammonium bicarbonate (NH_4HCO_3) (Prepare just before use): For 100 ml of 50 mM buffer, weigh 0.3953 g of NH_4HCO_3 and dissolve in 100 ml water.
7. Dithiothreitol (DTT): Dissolve DTT in water to a concentration of 1 M. Divide into small aliquots and store at –20 °C.
8. Reduction buffer (Prepare just before use): 10 mM DTT in the 50 mM NH_4HCO_3 . Dilute the 1 M DTT with the 50 mM NH_4HCO_3 .
9. Iodoacetamide (IAM): Dissolve IAM in water to a concentration of 1 M. Divide into small aliquots and store at –20 °C.
10. Alkylation buffer (Prepare just before use): Dilute the 1 M IAM with 50 mM NH_4HCO_3 to a final concentration of 50 mM IAM.
11. Lysyl Endopeptidase (LysC), MS grade.
12. Trypsin, MS grade.

2.2 StageTip Column and Spin Adaptor Preparation

Some types of StageTips and spin adaptors are commercially available.

1. Empore C18 extraction disk (Part Number 2215, 3 M).
2. Empore C8 extraction disk (Part Number 2214, 3 M).
3. Methanol.
4. 16 gauge, Kel-F Hub NDL, 2 in., point style 3 (Part Number 90516, Hamilton) needle to punch out small disks for preparing 200 μl pipet tip used in StageTip columns.

5. 13 and 14 gauge, Kel-F Hub NDL, 2 in., point style 3 (Part Numbers 7750-04 and 7750-05, Hamilton) needles to punch out small disks for preparing 1 ml pipet tip used in StageTip columns.
6. Plunger Assemblies (Part Numbers 1122-01 and 1162-01, Hamilton): 1122-01 and 1162-01 are for 16 gauge and 13/14 gauge needles, respectively.
7. Low binding 200 μ l and 1 ml pipet tips.
8. Lids of micro-centrifuge tubes.
9. Lids of 15 ml conical tubes.
10. Sharp-ended tweezers to drill a hole in the lids of micro-centrifuge tubes.
11. Soldering iron or drill to drill a hole in 15 ml conical tube lids.

2.3 Desalt the Digests with C18 1 ml StageTip Column

1. C18 1 ml StageTip: 1 ml pipet tip packed with four C18 disks with the 14 gauge needle and two C18 disks with the 13 gauge needle. Can be kept at room temperature.
2. 15 ml Conical tube.
3. The spin adaptor for 15 ml conical tube.
4. 2 % trifluoroacetic acid (TFA): 2 % TFA in water.
5. Solution A: 80 % acetonitrile (ACN) and 0.1 % TFA in water.
6. Solution B: 5 % ACN and 0.1 % TFA in water.
7. SpeedVac.

2.4 HAMMOC StageTip Column Preparation

Some types of StageTips are commercially available.

1. C8 200 μ l StageTip: 200 μ l pipet tip packed with a single C8 disk made with the 16 gauge needle. Can be kept at room temperature.
2. Titania (TiO_2) beads: Titansphere Bulk Media, 10 μ m (Part number 5020-75010, GL science). Store under dry conditions to keep specificity against phosphopeptides (see Note 2). The specificity varies to some extent from one production lot to another.
3. Methanol.

2.5 Phosphopeptide Enrichment by HAMMOC

1. Solution C: 300 mg/ml lactic acid in solution A.
2. HAMMOC StageTip: The C8 200 μ l StageTip with 3 mg TiO_2 beads. Prepare just before use (see Note 2).
3. Micro-centrifuge tube.
4. The spin adaptor for micro-centrifuge tube.
5. Solution A: 80 % ACN and 0.1 % TFA in water.

6. 20 % Phosphate: Dilute *ortho*-phosphoric acid in water to 20 %.
7. 5 % Ammonia: Dilute ammonia solution in water to 5 %.
8. 5 % Piperidine: Dilute piperidine in water to 5 %.

2.6 Desalt the Phosphopeptides with C18 200 μ l StageTip Column

1. C18 200 μ l StageTip: packed with a single C18 disk made with the 16 gauge needle. Can be kept at room temperature.
2. Micro-centrifuge tube.
3. The spin adaptor for micro-centrifuge tube.
4. Solution A: 80 % ACN and 0.1 % TFA in water.
5. Solution B: 5 % ACN and 0.1 % TFA in water.
6. SpeedVac.

3 Methods

3.1 Protein Extraction and Digestion (See Note 3)

1. Grind frozen plant material to a fine powder using a mortar and pestle, or apparatus such as bead mills (see Note 4). Make sure that the samples are constantly frozen after harvesting until suspended in the extraction buffer in the next step. Proteins can be rapidly dephosphorylated upon cell disruption without denaturants or phosphatase inhibitors.
2. Add denaturing extraction buffer to the powder and mix thoroughly (see Note 1). Make sure that all the powder comes in contact with the extraction buffer before the powder is thawed. Transfer the sample to a centrifuge tube if necessary. Do not heat the sample over 37 °C to avoid unwanted artificial protein modifications by a urea derivative.
3. Spin down cell debris and collect supernatant.
4. Determine protein concentration of the extract.
5. Add reduction buffer to the extract (1 μ l buffer for every 50 μ g protein (see Note 5)) and incubate for 30 min at room temperature (25 °C).
6. Add alkylation buffer to the reduced sample (1 μ l buffer for every 50 μ g protein (see Note 5)) and incubate in the dark for 20 min at room temperature (25 °C).
7. Add LysC to the alkylated sample (1 μ g for every 50 μ g protein (see Note 6)) and incubate for 3 h to overnight (see Note 7) at room temperature (25 °C).
8. Dilute the LysC digest with 4 volumes 50 mM NH_4HCO_3 (see Note 8).
9. Add Trypsin to the diluted sample (1 μ g for every 50 μ g protein) and incubate for overnight at room temperature (25 °C).

3.2 StageTip and Spin Adaptor Preparation

This protocol is adapted from [18].

1. Place the Empore C8 or C18 extraction disks on a clean disposable plastic petri dish. The C8 disk is for a HAMMOC StageTip. The C18 disk is for a desalting StageTip.
2. Wet the membrane disks with methanol.
3. Punch out small disks using 90° bevel needles with proper inner diameters (*see Note 9*).
4. Push the disk out of the needle and fix it into tapering of a pipet tip using plunger assemblies (*see Note 10*).
5. Drill a hole in the lids of micro-centrifuge tubes and 15 ml conical tubes, and use as spin adaptors for the 200 μ l and 1 ml StageTips, respectively.

3.3 Desalt the Digests with C18 1 ml StageTip (See Note 11)

1. Prepare a C18 1 ml StageTip.
2. Place the C18 1 ml StageTip on a 15 ml conical tube using the spin adaptor.
3. Add an equal volume of 2 % TFA to acidify the prepared digest (*see Note 12*). Acidification is required for peptide binding to the C18 membrane disks.
4. Conditioning of the C18 1 ml StageTip with Solution A. Load 1 ml Solution A onto the C18 1 ml StageTip. Spin the C18 1 ml StageTip to force Solution A through.
5. Conditioning of the C18 1 ml StageTip with Solution B. Load 1 ml Solution B onto the C18 1 ml StageTip. Spin the C18 1 ml StageTip to force Solution B through.
6. Load the acidified sample onto the conditioned C18 1 ml StageTip (*see Note 13*). Spin the C18 1 ml StageTip to bind peptides to the C18 disks and force the rest through.
7. Washing of the C18 1 ml StageTip with Solution B. Load 1 ml Solution B on the C18 1 ml StageTip. Spin the C18 1 ml StageTip to force Solution B through.
8. Elution of the bound peptides from the C18 disks with Solution A. Replace the waste conical tube with a new tube. Load 800 μ l Solution A on the C18 1 ml StageTip. Spin the C18 1 ml StageTip to elute peptides into a clean tube.
9. Dry the eluent in a SpeedVac. Dried peptide sample can be stored at -20 °C.

3.4 HAMMOC StageTip Preparation

1. Prepare a C8 200 μ l StageTip (*see Note 14*).
2. Weigh TiO₂ beads and add to the disk in the C8 200 μ l StageTip (*see Note 15*). To prepare many columns, TiO₂ beads can be suspended in methanol to load on C8 200 μ l StageTips (*see Note 16*). Most importantly, TiO₂ beads have to be kept under dry conditions to keep specificity against phosphopeptides (*see Note 2*).

3.5 Phosphopeptide Enrichment by HAMMOC

1. Dissolve desalted peptides in 200 μ l Solution C.
2. Place the HAMMOC StageTip on a micro-centrifuge tube using the spin adaptor.
3. Conditioning of the HAMMOC StageTip with Solution A. Load 80 μ l Solution A onto the HAMMOC StageTip. Spin the HAMMOC StageTip to force Solution A through.
4. Conditioning of the HAMMOC StageTip with Solution C. Load 80 μ l Solution C onto the HAMMOC StageTip. Spin the HAMMOC StageTip to force Solution C through.
5. Repeat step 4.
6. Replace the waste micro-centrifuge tube with a new tube. Load the dissolved sample onto the conditioned HAMMOC StageTip. Spin the HAMMOC StageTip to force the sample through.
7. Collect the flow-through and load onto the same HAMMOC StageTip again. Spin the HAMMOC StageTip to force the sample through.
8. Washing of the HAMMOC StageTip with Solution C. Load 80 μ l Solution C onto the HAMMOC StageTip. Spin the HAMMOC StageTip to force Solution C through.
9. Washing the HAMMOC StageTip with Solution A. Load 80 μ l Solution A onto the HAMMOC StageTip. Spin the HAMMOC StageTip to force Solution A through. This step is necessary to wash out non-phosphopeptides which may be bound to the C8 disk.
10. Add 100 μ l 20 % phosphate to a 1.5 ml micro-centrifuge tube into which phosphopeptides will be eluted.
11. Elution of the phosphopeptides from the TiO_2 beads with Elution buffer 1 (5 % Ammonia). Replace the waste micro-centrifuge tube with a tube containing 20 % phosphate. Load 50 μ l Elution buffer 1 on the HAMMOC StageTip. Spin the HAMMOC StageTip to elute phosphopeptides into the tube. Phosphopeptides are unstable in elution buffers under basic conditions, and eluents have to be acidified immediately.
12. Elution of remaining phosphopeptides from TiO_2 beads with Elution buffer 2 (5 % Piperidine) (see Note 17). Load 50 μ l Elution buffer 2 onto the HAMMOC StageTip. Spin the HAMMOC StageTip to elute phosphopeptides into the tube.

3.6 Desalt the Phosphopeptides with C18 200 μ l StageTip

1. Prepare a C18 200 μ l StageTip.
2. Place the C18 200 μ l StageTip on a 15 ml micro-centrifuge tube using the spin adaptor.
3. Conditioning of the C18 200 μ l StageTip with Solution A. Load 40 μ l Solution A onto the C18 200 μ l StageTip. Spin the C18 200 μ l StageTip to force Solution A through.

4. Conditioning of the C18 200 μ l StageTip with Solution B. Load 40 μ l Solution B onto the C18 200 μ l StageTip. Spin the C18 200 μ l StageTip to force Solution B through.
5. Load the acidified phosphopeptide enriched sample onto the conditioned C18 200 μ l StageTip. Spin the C18 200 μ l StageTip to bind peptides to the C18 disks and force the rest through.
6. Washing of the C18 200 μ l StageTip with Solution B. Load 40 μ l Solution B on the C18 200 μ l StageTip. Spin the C18 200 μ l StageTip to force Solution B through.
7. Elution of the bound peptides from the C18 disks with Solution A. Replace the waste conical tube with a new tube. Load 40 μ l Solution A on the C18 200 μ l StageTip. Spin the C18 200 μ l StageTip to elute peptides into the clean tube (*see Note 18*).
8. Dry the eluent in a SpeedVac. Dried peptide sample can be stored at -20 °C.
9. Dissolve the desalted peptides in buffer for LC-MS/MS analysis.

4 Notes

1. We add 1 ml buffer to the tube and vortex immediately. Phosphatase inhibitors may not be required in the denaturing extraction buffer. If alternative extraction buffers without denaturants are used, don't forget to add phosphatase inhibitors.
2. The specificity of TiO₂ beads against phosphopeptides is reduced by water absorption when it is kept without desiccation. Specificity can be recovered by heating the beads in a drying oven at 130 °C for 30 min.
3. Other methods such as filter-aided sample preparation (FASP) and phase-transfer surfactant (PTS) protocol may help recover more membrane proteins [19, 20].
4. We usually prepare 2 ml micro-centrifuge tubes suitable for bead mill apparatus with 4 zirconia beads (3 mm diameter) for each tube, and harvest ca. 200 mg plant materials directly into the tubes. Samples are frozen in liquid nitrogen and can be kept at -80 °C.
5. Not necessary to be a precise ratio.
6. For a sample intended for trypsin digestion, we use 0.5 μ g LysC for every 50 μ g protein. It is not necessary to digest with LysC prior to trypsin. However, LysC digestion under dena-

turing conditions greatly increases efficiency of the reaction, and can inactivate unwanted enzymes which would otherwise be active in trypsin-compatible buffers.

7. Incubate overnight when a sample will not proceed to the trypsin digestion and a LysC digest is used for phosphopeptide enrichment.
8. Trypsin retains most of its activity in 2.0 M urea.
9. We use Kel-F hub (KF) needles with a point style3 from Hamilton. A 16 gauge needle for 200 μ l StageTip preparation and 13/14 gauge needles are best for 1 ml StageTip preparation.
10. We use plunger assemblies from Hamilton. A plunger (part number 1122-01) for the 16 gauge needle and a plunger (part number 1162-01) for the 13/14 gauge needles.
11. Any C18 columns can be used for desalting the digests.
12. When different buffers were used for extraction and digestion steps, check pH with pH test paper.
13. We use a single C18 1 ml StageTip for desalting 400 μ g protein digest.
14. The membrane disk is only for keeping TiO₂ beads.
15. Specificity and capacity of TiO₂ beads against phosphopeptides may vary with samples. Test different ratios of TiO₂ to protein digest amounts and find the optimal conditions for your application. We usually use 3 mg TiO₂ beads for a single column and apply 400 μ g of protein digests from *Arabidopsis*.
16. Keep stirring the beads and dispense an equal amount of well-suspended solution onto the C8 columns.
17. Pyrrolidine can be used instead of piperidine [15].
18. It is better to elute peptides into a vial or a 96-well plate for LC-MS to minimize sample loss from transfers.

References

1. Dissmeyer N, Schnittger A (2011) The age of protein kinases. *Methods Mol Biol* 779:7–52
2. Nakagami H, Sugiyama N, Ishihama Y et al (2012) Shotguns in the front line: phosphoproteomics in plants. *Plant Cell Physiol* 53:118–124
3. Macek B, Mann M, Olsen JV (2009) Global and site-specific quantitative phosphoproteomics: principles and applications. *Annu Rev Pharmacol Toxicol* 49:199–221
4. Thingholm TE, Jensen ON, Larsen MR (2009) Analytical strategies for phosphoproteomics. *Proteomics* 9:1451–1468
5. Stensballe A, Andersen S, Jensen ON (2001) Characterization of phosphoproteins from electrophoretic gels by nanoscale Fe(III) affinity chromatography with off-line mass spectrometry analysis. *Proteomics* 1:207–222
6. Ikeguchi Y, Nakamura H (1997) Determination of organic phosphates by column-switching high performance anion-exchange chromatography using on-line preconcentration on titania. *Anal Sci* 13:479–483
7. Larsen MR, Thingholm TE, Jensen ON et al (2005) Highly selective enrichment of phosphorylated peptides from peptide mix-

- tures using titanium dioxide microcolumns. *Mol Cell Proteomics* 4:873–886
8. Beausoleil SA, Jedrychowski M, Schwartz D et al (2004) Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci USA* 101:12130–12135
 9. Bodenmiller B, Mueller LN, Mueller M et al (2007) Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat Methods* 4:231–237
 10. Bergstrom Lind S, Molin M, Savitski MM et al (2008) Immunoaffinity enrichments followed by mass spectrometric detection for studying global protein tyrosine phosphorylation. *J Proteome Res* 7:2897–2910
 11. Nakagami H, Sugiyama N, Mochida K et al (2010) Large-scale comparative phosphoproteomics identifies conserved phosphorylation sites in plants. *Plant Physiol* 153:1161–1174
 12. Sugiyama N, Nakagami H, Mochida K et al (2008) Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in *Arabidopsis*. *Mol Syst Biol* 4:193
 13. Ravichandran A, Sugiyama N, Tomita M et al (2009) Ser/Thr/Tyr phosphoproteome analysis of pathogenic and non-pathogenic *Pseudomonas* species. *Proteomics* 9:2764–2775
 14. Sugiyama N, Masuda T, Shinoda K et al (2007) Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. *Mol Cell Proteomics* 6:1103–1109
 15. Kyono Y, Sugiyama N, Imami K et al (2008) Successive and selective release of phosphorylated peptides captured by hydroxy acid-modified metal oxide chromatography. *J Proteome Res* 7:4585–4593
 16. Ishihama Y, Rappaport J, Mann M (2006) Modular stop and go extraction tips with stacked disks for parallel and multidimensional peptide fractionation in proteomics. *J Proteome Res* 5:988–994
 17. Rappaport J, Ishihama Y, 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
 18. Rappaport J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2:1896–1906
 19. Masuda T, Tomita M, Ishihama Y (2008) Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *J Proteome Res* 7:731–740
 20. Wisniewski JR, Zougman A, Nagaraj N et al (2009) Universal sample preparation method for proteome analysis. *Nat Methods* 6:359–362
 21. Nuhse TS, Stensballe A, Jensen ON et al (2004) Phosphoproteomics of the *Arabidopsis* plasma membrane and a new phosphorylation site database. *Plant Cell* 16:2394–2405
 22. Benschop JJ, Mohammed S, O'Flaherty M et al (2007) Quantitative phosphoproteomics of early elicitor signaling in *Arabidopsis*. *Mol Cell Proteomics* 6:1198–1214
 23. Nuhse TS, Bottrill AR, Jones AM et al (2007) Quantitative phosphoproteomic analysis of plasma membrane proteins reveals regulatory mechanisms of plant innate immune responses. *Plant J* 51:931–940
 24. Reiland S, Messerli G, Baerenfaller K et al (2009) Large-scale *Arabidopsis* phosphoproteome profiling reveals novel chloroplast kinase substrates and phosphorylation networks. *Plant Physiol* 150:889–903
 25. Grimsrud PA, den Os D, Wenger CD et al (2010) Large-scale phosphoprotein analysis in *Medicago truncatula* roots provides insight into in vivo kinase activity in legumes. *Plant Physiol* 152:19–28
 26. Chen Y, Hoehenwarter W, Weckwerth W (2010) Comparative analysis of phytohormone-responsive phosphoproteins in *Arabidopsis thaliana* using TiO_2 -phosphopeptide enrichment and mass accuracy precursor alignment. *Plant J* 63:1–17
 27. Kline KG, Barrett-Wilt GA, Sussman MR (2010) In planta changes in protein phosphorylation induced by the plant hormone abscisic acid. *Proc Natl Acad Sci USA* 107:15986–15991
 28. Reiland S, Finazzi G, Endler A et al (2011) Comparative phosphoproteome profiling reveals a function of the STN8 kinase in fine-tuning of cyclic electron flow (CEF). *Proc Natl Acad Sci USA* 108:12955–12960
 29. Lan P, Li W, Wen TN et al (2012) Quantitative phosphoproteome profiling of iron-deficient *Arabidopsis* roots. *Plant Physiol* 159:403–417
 30. Meyer LJ, Gao J, Xu D et al (2012) Phosphoproteomic analysis of seed maturation in *Arabidopsis*, rapeseed, and soybean. *Plant Physiol* 159:517–528
 31. Rose CM, Venkateshwaran M, Volkening JD et al (2012) Rapid phosphoproteomic and transcriptomic changes in the rhizobia-legume symbiosis. *Mol Cell Proteomics* 11:724–744
 32. Mayank P, Grossman J, Wuest S et al (2012) Characterization of the phosphoproteome of mature *Arabidopsis* pollen. *Plant J* 72:89–101

33. Bonhomme L, Valot B, Tardieu F et al (2012) Phosphoproteome dynamics upon changes in plant water status reveal early events associated with rapid growth adjustment in maize leaves. *Mol Cell Proteomics* 11:957–972
34. Nguyen TH, Brechenmacher L, Aldrich J et al (2012) Quantitative phosphoproteomic analysis of soybean root hairs inoculated with *bradyrhizobium japonicum*. *Mol Cell Proteomics* 11(11):1140–1155

Chapter 41

Simultaneous Identification and Quantification of Nitrosylation Sites by Combination of Biotin Switch and ICAT Labeling

Abasse Fares, Claude Nespolous, Michel Rossignol, and Jean-Benoît Peltier

Abstract

S-nitrosylation is a widespread modification of proteins. In plants, most information available to date regarding this modification was obtained using nitric oxide donors and concerned the proteins but not the identification of cysteine residues specifically modified in the proteins or their quantification. Here, we describe a method for the identification of endogenously nitrosylated cysteines in *Arabidopsis* and, simultaneously, the measurement of relative change in their abundance within binary comparisons.

Key words Protein S-nitrosylation, Biotin-switch, Isotope-coded-affinity tag, Nitrosothiols, Quantitation

Abbreviations

| | |
|-----------|--|
| CID | Collision-induced dissociation |
| DMF | Dimethylformamide |
| EDTA | Ethylenediaminetetraacetic acid |
| ESI-MS/MS | Electrospray ionization-tandem mass spectrometry |
| FDR | False discovery rate |
| HPLC | High-performance liquid chromatography |
| ICAT | Isotope-coded-affinity tag |
| MMTS | Methyl methanethiosulfonate |
| Q-TOF | Quadrupole-time of flight |
| SDS | Sodium dodecyl sulfate |
| Th | Thomson is a unit of mass-to-charge ratio |

1 Introduction

S-nitrosylation, the modification of the thiol group of cysteines by nitric oxide (NO) constitutes a wide-spread post-translational modification of crucial importance in a large array of processes, including in plants [1]. At the protein level, the “biotin-switch” method constituted a breakdown for the identification of nitrosylated proteins [2]. This method consists in blocking non-nitrosylated cysteines, and then reducing nitrosylated ones, prior to conjugation with a thiol reagent bound to biotin, thus allowing the selection of proteins containing formerly nitrosylated cysteines by avidin–biotin interaction. Very recently, at the peptide and amino acid level, the use of a thiol-specific isotope-coded affinity tag (ICAT) was reported to permit comparative quantification of differently nitrosylated peptides [3]. In both cases, however, nitrosylation was favored using a NO donor.

In the plant area, very similar progresses were noticed. Notably, the use of the biotin-switch allowed first to determine sets of nitrosylated proteins and then to identify proteins differentially nitrosylated in response to various stresses [2, 4–6]. However, as for other organisms, NO donors were used, thus preventing to decipher naturally occurring nitrosylation events, and all comparative analysis were based on quantification of nitrosylated proteins, without information about the concerned nitrosocysteines. Moreover, no approach aiming at large-scale identification of nitrosylated residues was published to date.

We detail here (Fig. 1) a combination of the biotin-switch method and ICAT labeling that was proven [7] to allow, for *Arabidopsis* proteins, the large-scale identification of endogenous nitrosothiols, in the absence of NO donor.

2 Materials

Prepare all solutions using ultrapure water (prepared by purifying deionized water to attain a resistivity of $18.2\text{ M}\Omega\text{ cm}$ at $25\text{ }^\circ\text{C}$) and analytical grade reagents. Prepare and store all reagents at $4\text{ }^\circ\text{C}$ (unless indicated otherwise). Diligently follow all waste disposal regulations when disposing waste materials. Do not add sodium azide to the reagents. Wear appropriate personal protective equipment (safety glasses, gloves, protective clothing) when handling chemicals such as MMTS, DMF, SDS, NaOH, ICAT reagents and buffers and use them only with adequate ventilation (e.g., fume hood). All the experiments from the cell grinding to the ICAT labeling are done in the darkness to protect nitrosothiols from light. All the 1.5 ml microcentrifuge tubes used in this experiment are LoBind Eppendorf tubes (Eppendorf 022431081) reducing adsorption to the tube walls.

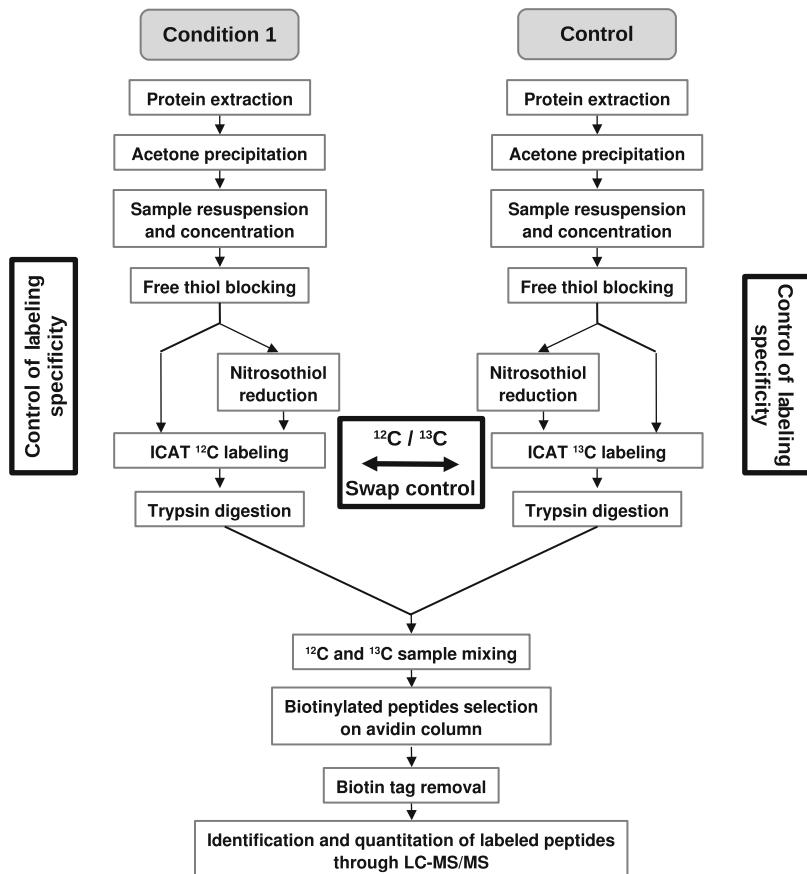


Fig. 1 Workflow for the identification and quantitation of nitrosothiols using ICAT labels. Three types of control are done: (1) the usual control related to condition 1, (2) a control of the labeling specificity where the nitrosothiol reduction is omitted and (3) a control to verify the absence of bias between ICAT 12C/13C labeling where the two labels are exchanged (swap)

2.1 Cell Culture

1. Culture medium: Culture medium is based on a modified Murashige and Skoog (MS) basal salt mixture (Sigma M2909) supplemented with vitamins (Sigma M3900) (*see Note 1*). The full description of the culture medium preparation is beyond the scope of this paper but the detailed protocol is available upon request. Different nutrients and chemicals have been added or reinforced comparing to MS basal salt mixture such as thiamine HCl, casein hydrolysate, 1-naphthaleneacetic acid, KNO_3 , $FeSO_4$, phosphate buffer (NaH_2PO_4/Na_2HPO_4), and sucrose for a final concentration of 1 μM , 0.1 g/L, 1 μM , 19 mM, 2.5 μM , 0.1 mM, and 43 mM respectively. Liquid culture media are autoclaved at 120 °C during 30 min and stored at room temperature in the dark until use.

2. Cell subculturing: *Arabidopsis* cells (ecotype Columbia, Col0) are subcultured every 10 days. 1/10 of the previous culture is diluted in 250 mL of fresh MS medium. Cell culture is maintained at 23 °C under continuous light and shaking (127 rpm/min) in an incubation shaker.

Check instructions

2.2 Cell Grinding

1. SDS stock solution: Prepare 250 mL of 20 % SDS solution (288.38 g/mol) by dissolving 50 g of SDS in 190 mL of water. Mix with stir bar/magnetic stirrer but do not vortex. Heat (80 °C) the solution in the microwave to speed up the dissolving and bring the volume up to 250 mL with water. Store preferentially the solution at 25 °C to avoid crystallization.
2. Hepes buffer: To prepare 1 L of 1 M Hepes (238.3 g/mol) buffer, pH 7.7, add 238.3 g of Hepes powder to 850 mL of water. Adjust the pH with NaOH pellets or saturated solution at first and use 2 M NaOH when pH gets close to the required pH. Bring the volume up to 1 L with water.
3. EDTA solution: To prepare 0.5 L of a 0.5 M EDTA solution, add 146.12 g of EDTA (292.24 g/mol) in 400 mL of water. Adjust immediately to pH 8 with NaOH (add few NaOH pellets at first and use NaOH 2 M when pH gets close to the required pH). Finally, bring the volume up to 500 mL with water.
4. HENS buffer, pH 7.7: 150 mM Hepes, 5 mM EDTA, 0.5 mM neocuproine 1 % SDS. To prepare 1 L of HENS buffer pH 7.7, mix 150 mL of Hepes 1 M pH 7.7 (from stock) with 100 mL of 0.5 M EDTA (from stock). Add 0.104 g of neocuproine (208.269 g/mol) and 50 mL of 20 % SDS (from stock). Bring the volume up to 1 L with water. Mix and verify the pH (7.7).
5. One Büchner funnel glass filter.
6. A blender (IKA Analytical mill A10—Fisher Scientific MPF-400-010J) or a mortar (see Note 2).
7. A 38 µm sieve.
8. A vacuum pump and a funnel.
9. 1 L of acetone placed at -20 °C.
10. 50 mL centrifuge tubes.
11. A centrifuge with a rotor for 50 mL centrifuge tubes.
12. 2D Quant kit assay (GE healthcare 80-6483-56) (see Note 3).

2.3 BS-ICAT Labeling

1. MMTS stock solution: To prepare a 2 M MMTS solution, mix 212 µL of MMTS (9.43 M is the initial concentration in the bottle) and 788 µL of DMF. Prepare fresh and store at room temperature in the dark until use.

2. A pack of filter devices (Amicon Ultra-0.5 centrifugal filter unit UFC503096, Amicon Ultra-4 UFC803096, and Amicon Ultra-15 UFC903096) (*see Note 4*).
3. HENU buffer, pH 7.7: 150 mM Hepes, 5 mM EDTA, 0.5 mM neocuproine (208.269 g/mol), 6 M Urea (60.06 g/mol): To prepare 250 mL of HENU buffer, mix 37.5 mL of 1 M Hepes, pH 7.7 (from stock), with 25 mL of 0.5 M EDTA (from stock). Add 0.026 g of neocuproine and 90.1 g of urea. Bring the volume up to 250 mL with water. Mix and verify the pH (7.7).
4. HEN buffer, pH 7.7: 150 mM Hepes, 5 mM EDTA, 0.5 mM neocuproine: To prepare 250 mL of HEN buffer, mix 37.5 mL of 1 M Hepes, pH 7.7 (from stock), with 25 mL of 0.5 M EDTA (from stock). Add 0.026 g of neocuproine. Bring the volume up to 250 mL with water. Mix and verify the pH (7.7).
5. A tabletop centrifuge (for 1.5 mL centrifuge tube).
6. SNO reducing solution: 0.5 M ascorbic acid (176.12 g/mol) and 100 mM CuCl₂ · 2 H₂O (170.48 g/mol). To prepare 0.5 mL of 0.5 M ascorbic acid and 0.1 M CuCl₂ 2 H₂O; weigh 44 mg of ascorbic acid and 8.6 mg of CuCl₂. Add 5 µL to 500 µL of sample (*see Subheading 3.2, step 2*) Prepare fresh and store at room temperature in the dark until use.
7. Acetonitrile.
8. ICAT labeling kit including buffer, labels, avidin column and cleaving reagents (ABSCIEX 4337336).
9. Ammonium carbonate solution: To prepare 1 M ammonium bicarbonate solution, dissolve 0.79 g of ammonium carbonate (79.06 g/mol) in 8 mL of water. Vortex to mix and bring the volume up to 10 mL.
10. Trypsin.
11. A vacuum concentrator.
12. An incubator.
13. A syringe-pump.
14. 5 mL syringes.

2.4 Peptide Cleaning with C18 Desalting Microcolumns

1. C18 desalting microcolumns (*see Note 5*).
2. Formic acid, acetonitrile as above (*see Subheading 2.3*).
3. Binding buffer: 0.1 % trifluoroacetic acid (vol/vol) in H₂O.
4. Elution buffer: 50 % acetonitrile (vol/vol), 0.5 % formic acid (vol/vol) in H₂O.

2.5 Mass Spectrometry Analysis

1. Mass spectrometer coupled to ESI-nanoflow HPLC (*see Note 6*).
2. Convenient software for quantitative analysis of MS spectra and database searching (*see Note 6*).

3 Methods

Carry out all procedures at room temperature unless otherwise specified.

3.1 Sample Preparation

1. Filter cells (5 g) through a Büchner funnel glass filter and wash them twice with 250 mL of water. Freeze immediately the cells in liquid nitrogen before storage at -85 °C until use.
2. Grind cells in a blender or in a mortar (precooled with liquid N₂) and dissolve the resulting powder in 30 mL of preheated (45 °C) extraction buffer (HENS) made of 150 mM Hepes buffer pH 7.7, 5 mM EDTA, 0.5 mM neocuproine, and 1 % SDS (preheating is used to avoid SDS precipitation). Filter quickly the solution through a 38 µm sieve using a vacuum pump and a funnel. Split the filtrate (25 mL) in 5×5 mL in 50 mL centrifuge tubes and dilute the filtrate (ratio 1:9) in cold acetone (-20 °C) (*see Note 7*). Keep at -20 °C for at least 30 min to precipitate proteins. Spin the 50 mL centrifuge tubes at 6,000×*g* for 15 min and collect the pellet. Allow pellet to dry under a hood for 15 min. Resuspend proteins in 5 mL of HENS buffer and quantify the protein concentration using the 2D Quant kit assay. Take out 2 mg of proteins and keep a protein concentration below 0.8 mg/mL (dilute your sample in HENS buffer if necessary).

3.2 BS-ICAT Labeling

1. Incubate the sample (2 mg of proteins) in a 1.5 mL centrifuge tube with 100 mM MMTS (20 min, 50 °C) to block free thiols (Fig. 1). Do not exceed 0.8 mg/mL of protein to keep the blocking step as efficient as possible. After the MMTS blocking step, dilute your sample with HENU buffer to keep your SDS concentration below 0.1 % (*see Note 8*). MMTS excess is removed by filtration using filter devices. To condition the filter devices (Amicon Ultra-15) add 15 mL of water and spin for 10 min at 5,000×*g* (fixed angle). Repeat this step twice and replace water by the HENU buffer for another step conditioning. Load the sample to the filter device and spin for 10 min at 5,000×*g*. Wash the sample three times (ratio 1:100) with HENU buffer on filter devices and a fourth time with HEN to dilute the urea at the end of the washing process. Make sure that the solution (wash buffer and sample) are well homogenized by pipetting up and down the solution in the filter device before centrifugation. Transfer the sample to a 0.5 mL filter device (Amicon Ultra-0.5) after filter device conditioning and add HEN buffer to the sample to bring the volume up to 500 µL.
2. To reduce nitrosothiols, add 5 mM ascorbic acid and 1 mM CuCl₂ (an enhancer of the reaction) and incubate in the filter device for 1 h at RT (Fig. 1; *see Note 9*).

3. Prepare ICAT reagents as described in the ICAT reagent kit manual (Subheading 6.2.2 in the manual). Bring to room temperature a vial of Cleavable ICAT Reagent Light and a vial of Cleavable ICAT Reagent Heavy. Centrifuge the reagents to bring all powder to the bottom of each vial. Add 40 μ L of acetonitrile to each reagent vial. Vortex to mix and then centrifuge for a few seconds to bring all solution to the bottom of the tube. All of the ICAT reagent may not dissolve. Transfer the entire contents of the control sample to the vial of the Light reagent. Transfer the entire contents of the test sample to the vial of the Heavy reagent. Vortex to mix and then centrifuge for a few seconds to bring all solution to the bottom of the tube. All of the ICAT reagent should dissolve. Incubate for 2 h at 37 °C in an incubator. As a control, it is highly recommended to swap the label in a parallel experiment to check that there is no bias towards one of the labeling reagent (Fig. 1). Excess of ICAT reagent is removed through centrifugation on filter devices (spin at 14,000 \times g for 7 min) as shown previously for MMTS and use 50 mM ammonium carbonate buffer for the three washing steps.
4. Dissolve lyophilized trypsin (20 μ g per vial) in 50 μ L of 50 mM of ammonium carbonate buffer. Digest proteins directly in the filter devices with trypsin. Keep trypsin/protein ratio between 1:50 and 1:100. Incubate overnight at 37 °C. Place a new collecting centrifuge tube on the filter device to make sure that the trypsinized peptides cannot be contaminated by the previous steps. Spin the filter tubes at 7,500 \times g for 5 min and collect the filtrate. It is recommended to wash the filter device with 100 μ L of 50 % acetonitrile followed by a second washing step with 500 mM NaCl. Putative adsorbed peptides are collected through centrifugation for each washing step and pooled with the primary filtrate. Mix heavy- and light-labeled peptides to the same tube.
5. Clean up the labeled peptides (*see Note 5*) using C18 desalting microcolumns (Sep-Pak kit) and follow the manufacturer's recommendations: the size of the cartridge should be selected on the basis of the amount of starting protein, considering their capacities are about 5 % (w/w) of the packing material's weight. For example, for 20 mg of a protein digest, a SepPak with 500 mg of C18 beads is recommended. In this protocol, we assume the use of 500 mg of SepPak (500-mg bulk material, 3 or 6 mL, 800- μ L bed volume). Volumes should be adapted accordingly for different sizes. A vacuum manifold can be used to increase solvent flow rates through the cartridge, or alternatively, air pressure is recommended for high-capacity SepPak, as it provides more uniform peptide loading and elution. Wash and condition the cartridge using 9 mL of acetonitrile followed

by 3 mL of 50 % acetonitrile and 0.5 % formic acid. Equilibrate with 9 mL of 0.1 % trifluoroacetic acid. Load sample in 0.4 % trifluoroacetic acid. Wash/desalt with 9 mL of 0.1 % trifluoroacetic acid. Wash (to remove trifluoroacetic acid) with 900 μ L of 0.5 % formic acid. Elute with 5 mL of 50 % acetonitrile and 0.5 % formic acid, and collect eluate in a 15 mL conical tube.

6. Set the avidin column as recommended by the manufacturer: mark the inlet and outlet ends of the cartridge (or mark with a directional arrow) for future use. Use the same flow direction in all runs to prevent particles that may accumulate at the cartridge inlet from clogging the outlet tubing. Insert the avidin cartridge into the cartridge holder. Inject 2 mL of the Affinity Buffer-Elute using a syringe pump. Divert to waste (*see Note 10*). Inject 2 mL of the Affinity Buffer-Load. Divert to waste. Neutralize each cation-exchange fraction by adding 500 μ L of the Affinity Buffer-Load. Check the pH using pH paper. If the pH is not 7, adjust by adding more Affinity Buffer-Load. Vortex to mix and then centrifuge for a few seconds to bring all solution to the bottom of the tube. Remove an optional 1 μ L process-monitoring aliquot and label as “pre-avidin.” Label three fraction-collection tubes: #1 (Flow-Through), #2 (Wash), and #3 (Elute), then place in a rack. Slowly inject (~1 drop/s) the neutralized fraction onto the avidin cartridge and collect the flow-through into tube #1 (Flow-Through). Inject 500 μ L of Affinity Buffer-Load onto the cartridge and continue to collect in tube #1. (Keep tube #1 until you confirm that loading on the avidin cartridge is successful. If loading fails, you can repeat loading using tube #1 after you troubleshoot the cause of the loading failure. To reduce the salt concentration, inject 1 mL of Affinity Buffer-Wash 1. Divert the output to waste. To remove nonspecifically bound peptides, inject 1 mL of Affinity Buffer-Wash 2. Collect the first 500 μ L in tube #2. Divert the remaining 500 μ L to waste. Inject 1 mL of water. Divert to waste. Fill a syringe with 800 μ L of the Affinity Buffer-Elute. To elute the labeled peptides, slowly inject (~1 drop/s) 50 μ L of the Affinity Buffer-Elute and discard the eluate. Inject the remaining 750 μ L of Affinity Buffer-Elute and collect the eluate in tube #3 (Elute). Vortex to mix and then centrifuge for a few seconds to bring all solution to the bottom of the tube. Remove an optional 1 μ L process-monitoring aliquot after eluting from the avidin cartridge, and label as “post-avidin.” If you have additional cation-exchange fractions, repeat the steps described. Evaporate each affinity-eluted fraction to dryness in a centrifugal vacuum concentrator. In a fresh tube, prepare the final cleaving reagent by combining Cleaving Reagent A and Cleaving Reagent B in a 95:5 ratio. You need ~90 μ L of final cleaving reagent per sample. Vortex to mix and then centrifuge for a few seconds to bring all solution to the bottom of the tube. Transfer ~90 μ L of

freshly prepared cleaving reagent to each sample tube. Vortex to mix and then centrifuge for a few seconds to bring all solution to the bottom of the tube. Incubate for 2 h at 37 °C. Centrifuge the tube for a few seconds to bring all solution to the bottom of the tube. Evaporate the sample to dryness in a centrifugal vacuum concentrator (~30 to 60 min). Resuspend the peptides in 0.1 % formic acid and desalt them on C18 desalting microcolumns (see Note 5) prior to MS analysis.

3.3 Mass Spectrometry and Data Analysis

1. Peptides were resolved on an HPLC coupled to a mass spectrometer using C18 PepMap100 phase for pre-column and column (75 µm diameter, 250 mm length) and a 2–27 % solvent B (90 % acetonitrile, 0.1 % formic acid) linear gradient, eluted into a Q-TOF mass spectrometer and fragmented by CID [8, 9] (see Note 6). For peptide identification, raw data were processed using the DataAnalysis software (Bruker) and the UniProtKB database was searched through the Mascot (Matrix Science, v.2.2.04) engine, both in target and decoy mode, using 20 ppm and 0.05 Da mass tolerance for peptides and fragments respectively, methionine oxidation and ICAT light or heavy as variable modification and allowing one missed cleavage. Significant matches ($p < 0.05$) were selected according to Mascot score. The FDR within a replicate was below 2 %, with an average lower than 1 % over all experiments [10, 11] (see Note 11). For quantitative analysis, extracted ion chromatograms were calculated (DataAnalysis software) and processed (Warp-LC software, Bruker) to determine the ratios of intensity of light and heavy ICAT labeled peptides (Fig. 2). Ratio data were normalized to the mean of all ratios within each replicate and averaged for each peptide across replicates.

3.4 Example of Results

Saline stress has been applied on *Arabidopsis* suspension cells. 10 days old cells were treated with 100 mM NaCl (or water as control) 30 min prior to be harvested. Three replicates have been made. 123 peptides were quantified with an average standard deviation of 10 %. Three peptides were found over-nitrosylated (left side of the sigmoid, log ratio >2.5) and 14 showed a de-nitrosylation (right side of the sigmoid, ratio ≤ 0.25) (Figs. 2, 3, and 4). Ratios close to 1 could represent constitutive nitrosylation because related peptides are labeled in both conditions (treated and control) and at the same level. This would implicate a fairly stable and widespread basic nitrosylation level in many proteins. Additional controls (see Note 9) show that this labeling is likely induced by an incomplete blocking of free SH by MMTS.

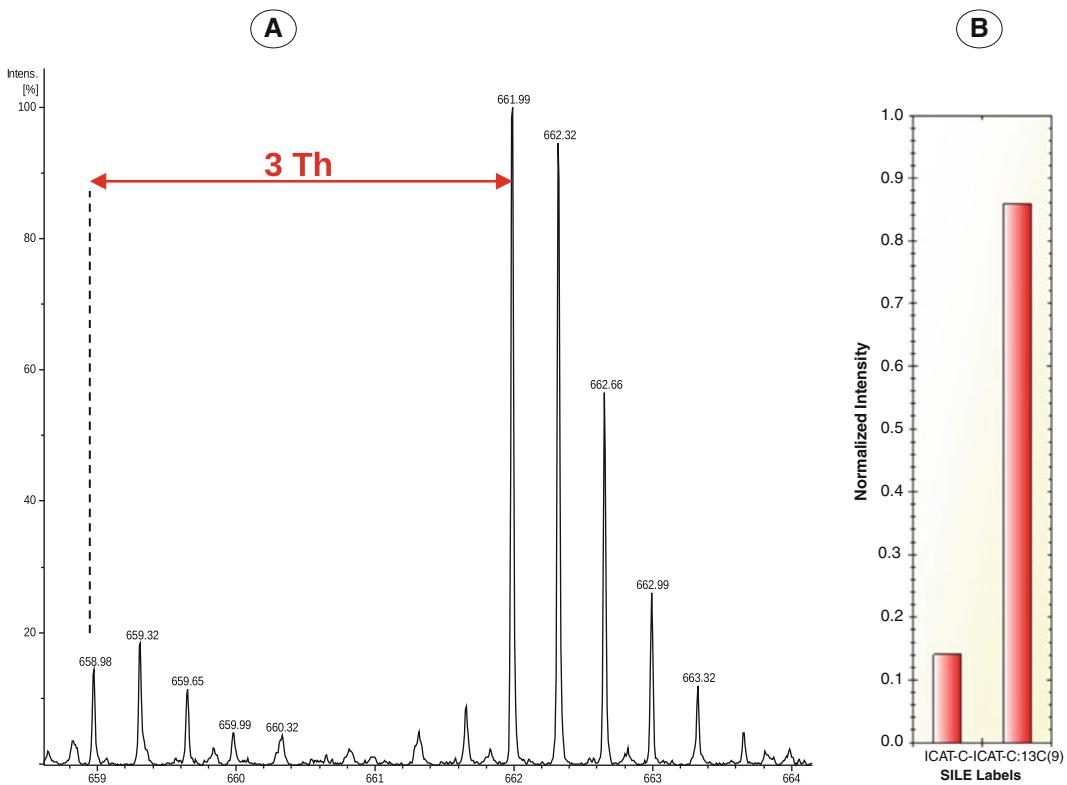


Fig. 2 ICAT labeling of endogenous nitrosothiols. A saline stress was applied on *A. thaliana* cells and nitrosylated proteins were labeled (ICAT-13C(9)). In parallel, proteins from control cells were labeled with (ICAT-12C(9)) and the two samples were mixed after labeling. **(a)** Example of a peptide pair from *Arabidopsis* Blue Copper protein (AT5G20230), separated by 3 Th due to charge state ($z=3$). **(b)** The ratio between the two peak intensities gives a relative quantitation of the peptides abundance under the two conditions

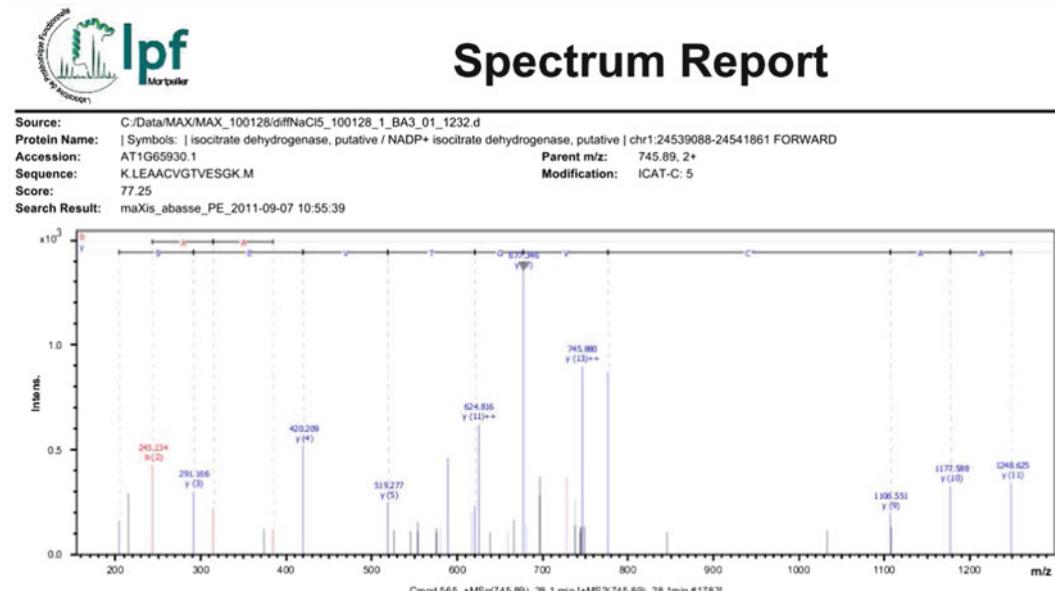


Fig. 3 MS/MS spectrum for an identified nitrosylated peptide

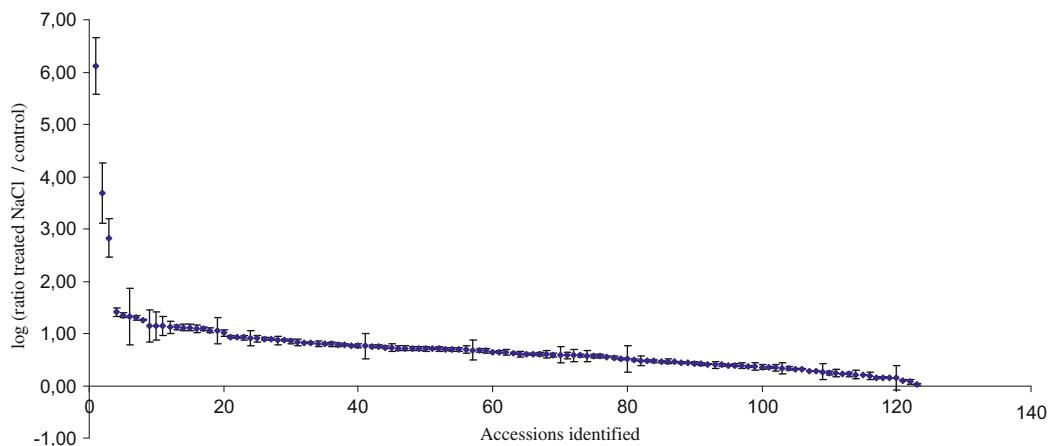


Fig. 4 Quantitation of nitrosothiols labeled with ICAT reagents in *Arabidopsis* cells treated with NaCl during 30 min. Three replicates have been made and only peptides present in all three replicates were quantified. Ratio mentioned in the figure represent treated cells/control. Every single point in the figure represents an accession and the standard deviation is indicated for each accession

4 Notes

1. Cells are quite sensitive to any changes and we do prefer to stick to the same medium provider.
2. This type of blender is recommended for its high efficiency on frozen material.
3. 2D Quant Kit is able to deal with high concentration of SDS and reductant.
4. We tried different filter devices and Amicon Ultra gave us the best compromise yield/speed of the concentration process.
5. C18 desalting microcolumns can be used at different times to get rid of impurities in peptide samples but some reagents like SDS are incompatible with C18. Depending of the sample volume, Zip-Tip pipette tips from Millipore (ZTC18M096) can be used (vol <150 μ L) or Sep-Pak from Waters WAT036820 (bigger volumes).
6. The procedure described here used a Q-TOF mass spectrometer (Maxis, Bruker), coupled to an Ultimate 3000 HPLC (Dionex), and corresponding Bruker software (DataAnalysis, Warp-LC).
7. The 50 mL centrifuge tubes are acetone resistant and can be spun at 6,000 $\times g$. 45 mL of cold acetone are added to 5 mL of sample in each 50 mL centrifuge tube.
8. Depending of the sample concentration, two filter devices can be used (Amicon Ultra-4 or -15). Moreover, when filter devices

are used, SDS concentration cannot exceed 0.1 % SDS. Consequently, samples must be diluted with HENU before loading on the filter devices. Urea in HENU buffer is used to facilitate SDS removal.

9. A control of the labeling specificity is absolutely necessary. Indeed, potential incomplete free thiol blocking needs to be evaluated. It is then recommended to introduce an additional control where the nitrosothiol reduction by ascorbate/CuCl₂ is omitted.
10. Injecting the Elute buffer before loading sample is required to free up low-affinity binding sites on the avidin cartridge.
11. If TP is true positive matches and FP is false positive matches, the number of matches in the target database is TP+FP and the number of matches in the decoy database is FP. The quantity that is reported is the False Discovery Rate (FDR)=FP/(FP+TP).

Acknowledgments

A.F. benefited from a PhD grant of the Région Languedoc-Roussillon. The authors acknowledge the help of the INRA Mass Spectrometry Proteomics Platform from the Pole Protéome de Montpellier in performing the analysis.

References

1. Astier J, Rasul S, Koen E et al (2011) S-nitrosylation: an emerging post-translational protein modification in plants. *Plant Sci* 181: 527–533
2. Jaffrey SR, Erdjument-Bromage H, Ferris CD et al (2001) S-nitrosylation: a physiological signal for neuronal nitric oxide. *Nat Cell Biol* 3: 193–197
3. Wu C, Parrott AM, Liu T et al (2011) Distinction of thioredoxin transnitrosylation and denitrosylation target proteins by the iCAT quantitative approach. *J Prot* 74: 2498–2509
4. Lindermayr C, Saalbach G, Durner J (2005) Proteomic identification of S-nitrosylated proteins in *Arabidopsis*. *Plant Physiol* 137: 921–930
5. Tanou G, Job C, Rajjou L et al (2009) Proteomics reveals the overlapping roles of hydrogen peroxide and nitric oxide in the acclimation of citrus plants to salinity. *Plant J* 60: 795–804
6. Abat JK, Deswal R (2009) Differential modulation of S-nitrosoproteome of *Brassica juncea* by low temperature: change in S-nitrosylation of Rubisco is responsible for the inactivation of its carboxylase activity. *Proteomics* 9: 4368–4380
7. Fares A, Rossignol M, Peltier JB (2011) Proteomics investigation of endogenous S-nitrosylation in *Arabidopsis*. *Biochem Biophys Res Commun* 416:331–336
8. Wells JM, McLuckey SA (2005) Collision-induced dissociation (CID) of peptides and proteins. *Methods Enzymol* 402:148–185
9. Cox J, Mann M (2011) Quantitative, high-resolution proteomics for data-driven systems biology. *Annu Rev Biochem* 80: 273–299
10. Choi H, Nesvizhskii AI (2008) False discovery rates and related statistical concepts in mass spectrometry-based proteomics. *J Proteome Res* 7:47–50
11. Reiter L, Claassen M, Schrimpf SP et al (2009) Protein identification false discovery rates for very large proteomics data sets generated by tandem mass spectrometry. *Mol Cell Proteomics* 8:2405–2417

Chapter 42

Tandem Metal-Oxide Affinity Chromatography for Enhanced Depth of Phosphoproteome Analysis

Gerold J.M. Beckers, Wolfgang Hoehenwarter, Horst Röhrlig, Uwe Conrath, and Wolfram Weckwerth

Abstract

In eukaryotic cells many diverse cellular functions are regulated by reversible protein phosphorylation. In recent years, phosphoproteomics has become a powerful tool to study protein phosphorylation because it allows unbiased localization, and site-specific quantification, of *in vivo* phosphorylation of hundreds of proteins in a single experiment. A common strategy to identify phosphoproteins and their phosphorylation sites from complex biological samples is the enrichment of phosphopeptides from digested cellular lysates followed by mass spectrometry. However, despite the high sensitivity of modern mass spectrometers the large dynamic range of protein abundance and the transient nature of protein phosphorylation remained major pitfalls in MS-based phosphoproteomics. Tandem metal-oxide affinity chromatography (MOAC) represents a robust and highly selective approach for the identification and site-specific quantification of low abundant phosphoproteins that is based on the successive enrichment of phosphoproteins and -peptides. This strategy combines protein extraction under denaturing conditions, phosphoprotein enrichment using Al(OH)_3 -based MOAC, tryptic digestion of enriched phosphoproteins followed by TiO_2 -based MOAC of phosphopeptides. Thus, tandem MOAC effectively targets the phosphate moiety of phosphoproteins and phosphopeptides and, thus, allows probing of the phosphoproteome to unprecedented depth.

Key words Phosphoproteomics, Tandem-MOAC, LC-MS, Orbitrap

1 Introduction

Reversible protein phosphorylation is a multifunctional posttranslational modification which serves to regulate enzyme catalytic activity, localization, stability, and interactions. With 2–4 % of eukaryotic proteomes accounting for kinases and phosphatases, it is estimated that one-third of all proteins are phosphorylated. Current advances for determining the phosphosites, stoichiometries, and functional roles of each of these *in vivo* phosphorylations

Gerold J. M. Beckers and Wolfgang Hoehenwarter contributed equally to this work.

emerged from mass spectrometry (MS)-based phosphoproteomics strategies. To facilitate protein phosphorylation analysis, a number of enrichment methods have been applied to separate phosphorylated proteins or peptides from non-phosphorylated ones. A collection of metal-based affinity methods such as immobilized metal affinity chromatography (IMAC), metal oxide affinity chromatography (MOAC), and strong cation exchange (SCX) are among the most commonly used. These enrichment methods, in combination with MS, now allow detection and quantification of hundreds of *in vivo* phosphorylation sites from complex biological samples. Today, enrichment of phosphorylated peptides using TiO_2 upon proteolysis digestion of the protein extract has become popular and is applied most frequently. Besides enrichment strategies that target the phosphate moiety of phosphopeptides, methods for the enrichment of phosphoproteins have also been described. However, dynamics of protein abundance in a given cell or tissue, the transient nature and low-stoichiometry of the phosphorylation event remain major challenges in phosphoproteomics.

In this chapter we describe a cost-efficient workflow for the consecutive enrichment of phosphoproteins and peptides and their subsequent analysis by LC/MS. We combine denatured protein extraction with highly selective Al(OH)_3 -MOAC enrichment of phosphoproteins [1, 2], proteolytic cleavage of enriched phosphoproteins, and TiO_2 -MOAC of phosphopeptides. In our hands, this approach ensured reproducible enrichment of transiently modified peptides allowing direct identification and site-specific quantification of phosphorylation of many low abundant phosphoproteins such as transcription factors, kinases, and phosphatases [3].

2 Materials

Prepare all solutions using ultrapure water (bi-distilled or double distilled, deionized) and analytical/ultra HPLC grade reagents.

2.1 Total Protein Extraction

1. 10 % (w/v) Trichloroacetic acid (TCA) in acetone.
2. 10 % (w/v) TCA in dH_2O .
3. Dense SDS buffer: 100 mM tris(hydroxymethyl)aminomethane (Tris) pH 8.0, 30 % (w/v) sucrose, 2 % (w/v) sodium dodecyl sulfate (SDS), 5 % (v/v) β -mercaptoethanol.
4. Phenol, Tris-saturated pH 8.0.
5. 100 mM Ammonium acetate in methanol (MeOH).
6. Incubation buffer A (IB/A): 30 mM 2-(*N*-morpholino)-ethanesulfonic acid (MES) pH 6.1, 0.25 % (w/v) 3-(3-cholamidopropyl)dimethylammonio-1-propanesulfonate (CHAPS), 7 M urea, 2 M thiourea.

2.2 *Al(OH)₃-Based MOAC Enrichment of Phosphoproteins*

1. Incubation buffer B (IB/B): 30 mM MES pH 6.1, 0.25 % (w/v) CHAPS, 8 M urea, 200 mM sodium glutamate, 200 mM potassium aspartate, 30 mM imidazole.
2. Incubation buffer C (IB/C): Mix 1 vol IB/A with 2 vol IB/B.
3. Incubation buffer 200 (IB200): 30 mM MES pH 6.1, 0.25 % (w/v) CHAPS, 8 M urea, 200 mM sodium glutamate, 200 mM potassium aspartate, 20 mM imidazole.
4. Elution buffer (EB): 200 mM potassium pyrophosphate pH 9.0, 8 M urea.
5. 2 % (w/v) sodium deoxycholate (DOC).
6. 100 % (w/v) TCA in dH₂O.
7. 25 % (w/v) TCA in dH₂O.
8. 80 % (v/v) acetone in 50 mM Tris pH 7.5 (stored at -20 °C until use).

2.3 *Phosphoprotein Digestion and Peptide Desalting*

1. 100 mM ammonium bicarbonate (AmBiC), 8 M urea.
2. 100 mM AmBiC.
3. 20 % (v/v) acetonitrile (ACN), 100 mM AmBiC.
4. 10 % (v/v) ACN, 25 mM AmBiC.

2.4 *TiO₂-Based MOAC Enrichment of Phosphopeptides*

1. Buffer 1 (B1): 50 % (v/v) ACN, 2.5 % (v/v) trifluoroacetic acid (TFA) in dH₂O, phthalic acid-saturated (~500 mg of phthalic acid saturates 20 mL B1).
2. Buffer 2 (B2): 50 % (v/v) ACN, 0.1 % (v/v) TFA in dH₂O.
3. Buffer 3 (B3): 0.1 % (v/v) TFA in dH₂O.
4. Buffer 4 (B4): 5 % (v/v) ammonia in dH₂O.

3 Methods

3.1 *Total Protein Extraction*

All steps during total protein extraction should be carried out using ice-cold buffer solutions except for dense SDS buffer and Tris-saturated phenol.

1. Grind 5–10 g (FW) *Arabidopsis* tissue to a fine powder in liquid nitrogen and fill two prechilled 50 mL tubes up to the 17.5 mL mark. If not processed immediately after sampling, store at -80 °C until analysis.
2. Pre-warm Tris-saturated phenol at RT and freshly prepare dense SDS buffer.
3. Add 30 mL ice-cold acetone to each sample, vortex vigorously to resuspend the tissue powder and spin down at 3,000×*g* for 5 min at 4 °C.

4. Discard supernatant and repeat **step 3**.
5. Discard the supernatant and resuspend each pellet in 30 mL ice-cold 10 % (w/v) TCA in acetone.
6. Incubate the samples in an ultrasound water bath for 10 min at 4 °C. Add crushed ice to keep the samples cold.
7. Recentrifuge and wash each pellet twice with 30 mL ice-cold 10 % (w/v) TCA in acetone.
8. Wash each pellet once with 30 mL ice-cold 10 % (w/v) TCA in dH₂O and then twice with 30 mL ice-cold 80 % (v/v) acetone.
9. Resuspend each tissue powder pellet in 24 mL freshly prepared dense SDS buffer.
10. Add 24 mL tris-saturated phenol pH 8.0 and vortex vigorously for ~1 min.
11. Separate phases by centrifugation at RT for 30 min at 3,000 $\times \text{g}$.
12. Transfer each upper phenolic phase to a centrifuge bottle (phenol resistant, e.g., Nalgene® 175 mL conical-bottom polypropylene copolymer (PPCO) centrifuge bottles).
13. Add 5 volumes of ice-cold 100 mM ammonium acetate in MeOH, vortex and precipitate proteins for 1 h at -20 °C.
14. Collect precipitated proteins by centrifugation at 7,500 $\times \text{g}$ for 10 min at 4 °C.
15. Discard the supernatants and resuspend the protein pellets in 30 mL ice-cold 100 mM ammonium acetate in MeOH.
16. Recentrifuge protein samples at 7,500 $\times \text{g}$ for 10 min at 4 °C.
17. Wash the protein samples once again with 30 mL ice-cold 100 mM ammonium acetate in MeOH and twice with 30 mL ice-cold 80 % acetone.
18. Discard as much acetone as possible, air-dry the protein pellets shortly on ice and dissolve each protein pellet in 6 mL IB/A by head-over-head incubation at 10 °C overnight.
19. Alternatively, dry protein pellets can be stored at -20 °C.

3.2 Al(OH)₃-Based MOAC Enrichment of Phosphoproteins

1. Pool two protein pellets each dissolved in 6 mL IB/A in a 50 mL tube.
2. Spin down undissolved proteins by centrifugation at 7,500 $\times \text{g}$ for 10 min at 10 °C.
3. Transfer the supernatant to a clean 50 mL tube and determine protein concentration by Bradford assay.
4. Adjust the protein concentration to 3 mg/mL using IB/A.
5. Transfer 12 mL of the 3 mg/mL protein solution in IB/A to a clean 50 mL tube.

6. Add 2 vol IB/B and immediately mix by head-over-head incubation.
7. Clear protein solution by centrifugation at $7,500 \times g$ for 10 min at 10 °C.
8. Meanwhile transfer 2.88 g Al(OH)₃ to a 50 mL round-bottom tube.
9. Resuspend in 36 mL IB/C to equilibrate the chromatography media.
10. Spin down Al(OH)₃ for 3 min at $3,000 \times g$ and remove supernatant using a 25 mL pipette.
11. Repeat **steps 8–10**.
12. Discard as much IB/C from the pelleted Al(OH)₃ as possible using a gel-loader pipet-tip and add 36 mL cleared dissolved protein sample.
13. Allow phosphoproteins to bind the Al(OH)₃ by head-over-head incubation at 10 °C for 1 h.
14. Spin down by centrifugation at $3,000 \times g$ for 5 min at 10 °C.
15. Either discard non-bound protein fraction or store for further analysis.
16. Wash Al(OH)₃-bound phosphoproteins six times by resuspension in 36 mL IB200. Invert the tubes carefully and do not vortex Al(OH)₃-bound phosphoproteins. To spin down Al(OH)₃-bound phosphoproteins, centrifuge at $3,000 \times g$ for 5 min at 10 °C.
17. After the last washing step, remove as much IB200 as possible using a gel-loader pipette tip.
18. Resuspend the Al(OH)₃-bound phosphoproteins in 24 mL EB and allow them to elute from the Al(OH)₃ for 30 min by head-over-head incubation at RT.
19. Pellet Al(OH)₃ by centrifugation at $3,000 \times g$ for 5 min at RT.
20. Transfer supernatant containing the eluted phosphoproteins to a clean 50 mL conical bottom tube.
21. Centrifuge samples at $7,500 \times g$ for 10 min at RT to pellet as much Al(OH)₃ as possible.
22. Apply 12 mL supernatant to an Amicon Ultra-15 centrifugal filter unit with a 10 kDa cut-off and centrifuge at $3,000 \times g$ for 30 min at RT.
23. Repeat **step 22** to concentrate protein sample to a minimal volume of 250 µL.
24. Transfer as much retentate as possible to a 1.5 mL Eppendorf Protein LoBind® tube and dilute sample by adding 4 vol of dH₂O.

25. Add 12.5 μ L 2 % (w/v) DOC solution, vortex vigorously and incubate for 5 min at RT.
26. Add 125 μ L ice cold 100 % (w/v) TCA in dH₂O, vortex and incubate on ice for 2 h to precipitate phosphoproteins.
27. Pellet precipitated phosphoproteins by centrifugation at 14,000 $\times g$ for 10 min at 4 °C.
28. Discard supernatant and resuspend pellet once in 1 mL ice-cold 25 % (w/v) TCA in dH₂O. Fine dispersal of the phosphoprotein pellet is important to dissolve all residual Al(OH)₃ in the acidic washing buffer.
29. Recentrifuge samples and wash pellet with 1 mL ice-cold 80 % acetone in 50 mM Tris pH 7.5.
30. Wash pellets with 1 mL ice-cold 100 % acetone and remove supernatant with a pipette.
31. Shortly air-dry phosphoproteins on ice, and store at -20 °C or immediately continue with protein digestion.

3.3 Protein Digestion and Peptide Desalting

3.3.1 Protein Digestion

1. Dissolve phosphoproteins by resuspension in 400 μ L 100 mM AmBiC, 8 M urea and incubation at RT with slight agitation for 1 h.
2. Pellet non-dissolved proteins by centrifugation at 14,000 $\times g$ for 10 min at RT.
3. Transfer supernatant to a clean 1.5 mL Eppendorf Protein LoBind® tube and determine protein concentration using Bradford protein assay.
4. To 500 μ g phosphoproteins add 1 vol 20 % ACN, 100 mM AmBiC.
5. *Optional:* Add 5 μ g endoprotease LysC enzyme (sequencing grade) and incubate at 37 °C in the dark with slight agitation for 5 h.
6. Dilute protein sample 1:1 with 10 % (v/v) ACN, 25 mM AmBiC.
7. Add 10 μ L Poroszyme immobilized modified Trypsin slurry (sequencing grade).
8. Incubate with agitation overnight at 37 °C in the dark.

3.3.2 Peptide Desalting

For peptide desalting we routinely use 96-well C-18 Solid Phase Extraction Column plates (Agilent) in a vacuum manifold system, however other supports like for example C-18 Stage tips can be used as well. When using the SPEC plates in a vacuum manifold to dry the disk membranes of the columns we apply minimum vacuum and release gently to avoid damaging the columns. The collection/waste tray is emptied after each washing/equilibration step. All steps are performed at RT.

1. Add 400 μ L MeOH to each well to activate and wash the SPEC plate and repeat once (use one well for each sample).
2. Equilibrate four times with 400 μ L dH₂O.
3. Meanwhile spin down trypsin beads by centrifugation at 14,000 $\times g$ for 10 min at RT.
4. Load 500 μ L peptides-containing supernatant to the center of the well and allow the column to absorb the peptides by incubating for 1 min.
5. Repeat **step 4** until each sample has been completely loaded onto a single well.
6. Wash four times with 400 μ L dH₂O.
7. Change collection plate and use an Eppendorf Protein LoBind 96-deepwell plate to collect peptides.
8. Elute peptides with 200 μ L MeOH by incubating for 5 min and repeat once.
9. Transfer the peptide solution (eluate) to a 1.5 mL Eppendorf Protein LoBind tube using a glass Pasteur pipette.
10. Dry peptides completely in a speedvac.

3.4 *TiO₂-Based MOAC Enrichment of Phosphopeptides*

1. Dissolve peptides in 100 μ L B1 and centrifuge for 10 min at 12,000 $\times g$.
2. For each sample containing 500 μ g peptides weigh 12.5 mg TiO₂ into a spin column with a polyethylene filter of 10 μ m pore size, screw cap and press-in bottom plug.
3. Equilibrate TiO₂ by adding 250 μ L B1 to each column and incubating for 5 min.
4. Spin column at 700 $\times g$ for 2 min.
5. Add peptide mixture of **step 1** to the column and allow phosphopeptides to bind the TiO₂ chromatography media for 15 min by closing the column and incubating it head-over-head.
6. Open the column and place it in a clean tube. Spin at 700 $\times g$ for 2 min, the flow-through fraction containing non-phosphorylated peptides can be desalted and stored at -20 °C for further analysis.
7. The column is then washed six times in total using 250 μ L buffer for each washing step: 2 \times B1, 2 \times B2 and then 2 \times B3.
8. Phosphopeptides are eluted with 100 μ L B4 by incubating the closed column for 5 min head-over-head.
9. Collect the eluted phosphopeptides in a clean 1.5 mL Eppendorf Protein LoBind tube by centrifugation at 700 $\times g$ for 2 min.
10. Dry peptides completely in a speedvac.

3.5 LC-MS (See Note 1)

1. Calibrate and prepare/tune the LC-MS system according to the manufacturer's instructions. We highly recommend the use of UPLC chromatography systems and HR/AM mass spectrometers for optimal results. We use a LTQ-OrbiTrap XL mass spectrometer both from Thermo Scientific.
2. Dissolve the peptides in 8 μ L 5 % (v/v) ACN, 0.1 % (v/v) FA in dH₂O. Pellet insoluble debris by centrifugation at 14,000 \times g for 10 min. Carefully transfer 7 μ L of clear phosphopeptide solution to the sample container appropriate for the LC auto sampler.
3. Program the LC-MS method to inject 5 μ L of phosphopeptide solution. For chromatography we use a Chromolith CapRod monolithic column with a C18 stationary phase with a length of 150 mm and an inner diameter of 0.1 mm. We use 0.1 % (v/v) FA in dH₂O and 90 % (v/v) ACN, 0.1 % (v/v) FA in dH₂O for gradient elution and reverse phase separation of phosphopeptides.
4. Program the LC-MS method to deliver a suitable flow rate and elution gradient. When using packed columns, small particle diameters will allow higher flow rates without adversely affecting resolution allowing shorter analysis time. Longer columns will achieve higher resolution and have a higher peak capacity however the inlet pressure will be limiting. We used a linear gradient from 5 to 35 % organic mobile phase content in 120 min and a flow rate of 500 nl/min.
5. Equilibrate the column delivering an isocratic flow with the same mobile phase composition as the phosphopeptide solution, in our case 5 % (v/v) ACN, 0.1 % (v/v) FA in dH₂O for 2 or more column volumes.
6. Set up the LC-MS junction via an appropriate emitter and ESI source and establish a stable electrospray. This step can be difficult. Adjust the spray voltage and emitter position, replace the emitter, check the ion optics settings and ensure suitable potential gradients and RF amplitudes for the ions to be able to overcome possible potential barriers. If problems persist check the calibration, run instrument diagnostics and if necessary clean and tune the system according to the manufacturer's instructions.
7. Program the LC-MS method to acquire MS data throughout the duration of phosphopeptide elution from the LC system. As mentioned above many mass spectrometry methods can be applied to the analysis of phosphopeptides dependent on the capabilities of the LC-MS system and the desired results. We used data-dependent analysis in which the five most intense ions recorded in a HR/AM full scan of the total ion population in the

OrbiTrap mass analyzer were isolated with an isolation width of 2 Th in the linear trap quadrupole (LTQ) mass analyzer and fragmented applying collision induced dissociation (CID). The fragment ions were then recorded in low resolution MS/MS spectra. Electron transfer dissociation (ETD) fragmentation is recommended when available because it leaves the phosphate-ester bond intact thereby preserving the phosphate moiety in peptide primary structure and generates extensive fragmentation of the peptide backbone for MS/MS spectra rich in sequence information [4]. Automatic gain control (AGC) for the Orbitrap was set to 5e+05. The Maximum injection time (max IT) was set to 500 ms, and injection wave forms were enabled. Full scan mass spectra were internally calibrated on the fly using lock mass for an average mass error of less than 1 ppm. Exclusion duration for data-dependent selection of MS/MS precursor ions was set to 30 s and exclusion width was \pm 10 ppm. Exclusion duration of 30 s may seem low but we thought it suitable to increase the number of MS/MS spectra acquired for low abundant phosphopeptides. Charge state screening was enabled and single-charged ions were excluded from MS/MS acquisition. Multistage activation/pseudoMS3 was enabled with a neutral loss mass list of 293.91, 195.94, 97.97, 48.999, 32.66, and 24.49 Da to generate concatenated MS/MS spectra containing peptide backbone fragments when initial CID resulted in release of the phosphate moiety. The AGC for MS/MS acquisition in the LTQ was set to 3e+04; the max IT was set to 100 ms. Three microscans were acquired per MS/MS spectrum.

8. Control all LC-MS system and method parameters, the sample list, the vial position in the auto sampler and start the analysis.
1. Load raw files. MaxQuant accepts .raw files from Thermo Scientific mass spectrometers. MaxQuant version 1.3.0.5 accepts mzXml files so it is compatible with other instrument vendors. Convert the mass spectrometer's output files to mzXml format if necessary. There are several tools available on line such as RAW Xtractor as part of the Census package from the Scripps Institute, DTA Supercharge integrated into MSQuant or MM File Conversion as part of the MassMatrix database search engine for this purpose. Some difficulties configuring these software to individual systems may arise.
2. Generate an experimental design template and then load it into MaxQuant under the tab "Identification & quantification." We recommend declaring each LC-MS analysis including replicate analyses as an individual experiment. Grouping LC-MS analyses will produce a summed intensity over all grouped analyses for each identified phosphopeptide and will preclude estimation of variability between individual analyses.

3.6 Phosphopeptide Identification, Quantification, and Phosphosite Mapping (See Note 2)

3. Upload the database to be searched, adjust default settings as desired. Add Phospho (STY) to the tolerated protein modifications, relax Peptide and Protein FDRs if desired. We recommend keeping the Site FDR at 0.01 for high quality mapping of phosphosites. Check label-free quantification if applicable.
4. Run MaxQuant. If sufficient computing power is available increase the number of threads. On 64 bit machines with 4 or 8 Gb of RAM most small to medium scale jobs should be completed within a day. If problems arise uncheck protein quantification options. MaxQuant will no longer quantify phosphopeptide intensities but will still map phosphosites.
5. Open Viewer to analyze search results. The modifications tab contains information on all identified phosphopeptides including q-values and posterior error probabilities (PEPs) as well as likelihood estimates of phosphorylated residues (phosphorylation sites) and intensities. The peptides tab contains information on all peptide spectral matches (PSMs) so redundant PSMs for each phosphorylated peptide and possible unphosphorylated counterparts can be found here.

Another good option for phosphopeptide identification and probabilistic phosphosite mapping is phosphoRS, integrated into Proteome Discoverer versions 1.3 and higher. Results similar to MaxQuant are achieved.

4 Notes

1. Liquid chromatography (LC) on-line with high resolution accurate mass (HR/AM) mass spectrometry (MS) is the preferred method for large scale analysis of phosphopeptides. The precise analytical strategy in the framework of LC-MS however is dependent on the instrumental capabilities at hand, the focus of the study, the researcher's expertise and experience, and other parameters, so a wide range of applications are possible downstream of tandem MOAC enrichment of phosphopeptides. In our research we have performed large scale identification and quantification of site-specific phosphorylation using a shotgun proteomics approach modified especially for phosphopeptide identification but targeted strategies aimed at specific phosphopeptides or phosphorylation sites are also conceivable. Keeping this in mind we describe our approach commenting on possible alterations.
2. Mass spectrometry is a powerful application for the large scale mapping of phosphorylation sites to protein primary structure that also allows quantification of site specific phosphorylation stoichiometries.

A complete unbiased approach for quantification independent from database search is mass accuracy precursor alignment (MAPA) using the software ProtMax [5, 6] (<http://www.univie.ac.at/mosys/software.html>). A step-by-step tutorial can be found at [6]. After enrichment techniques phosphopeptides can be quantified in a very convenient way in combination with multivariate statistical analysis [3, 7]. Multivariate statistics is of paramount importance to distinguish constitutively phosphorylated peptides from phosphopeptides showing a clear response to the treatment or regulatory effects [3, 7]. We routinely perform these analyses using COVAIN, a Matlab®-based statistical toolbox that allows a deep statistical analysis and data preprocessing in a user-friendly environment [8] (<http://www.univie.ac.at/mosys/software.html>).

Further we recommend MaxQuant for phosphopeptide identification, phosphosite mapping, and quantification of phosphorylation stoichiometries. MaxQuant operation has been described previously [9] and an excellent online user forum is available via the software's homepage at <http://maxquant.org/>. We describe some modifications to the default settings for phosphopeptides.

Acknowledgments

We thank present and former colleagues from our labs, especially Dennis Hopkins, Bastian Minkenberg, Matthias Nagler, Ella Nukarinen, and Martin Thomas who contributed to the development and optimization of the strategies described in this manuscript.

References

1. Wolschin F, Weckwerth W (2005) Combining metal oxide affinity chromatography (MOAC) and selective mass spectrometry for robust identification of *in vivo* protein phosphorylation sites. *Plant Methods* 1:9
2. Wolschin F, Wienkoop S, Weckwerth W (2005) Enrichment of phosphorylated proteins and peptides from complex mixtures using metal oxide/hydroxide affinity chromatography (MOAC). *Proteomics* 5:4389–4397
3. Hoehenwarter W, Thomas M, Nukarinen E et al (2013) Identification of novel *in vivo* MAP kinase substrates in *Arabidopsis thaliana* through use of tandem metal oxide affinity chromatography. *Mol Cel Proteomics* 12:369–380
4. Molina H, Horn DM, Tang N et al (2007) Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc Natl Acad Sci USA* 104:2199–2204
5. Hoehenwarter W, van Dongen JT, Wienkoop S et al (2008) A rapid approach for phenotype-screening and database independent detection of cSNP/protein polymorphism using mass accuracy precursor alignment. *Proteomics* 8:4214–4225
6. Egelhofer V, Wolfgang H, Lyon D et al (2013) Using ProtMAX to create high mass accuracy precursor alignments from label-free quantitative mass spectrometry data generated in large shotgun proteomics experiments. *Nat Protoc* 8(3):595–601
7. Chen Y, Hoehenwarter W, Weckwerth W (2010) Comparative analysis of phytohormone-responsive phosphoproteins in *Arabidopsis thaliana* using TiO₂-phosphopeptide enrichment.

- ment and mass accuracy precursor alignment. *Plant J* 63:1–17
8. Sun X, Weckwerth W (2012) COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics* 8:81–93
9. Cox J, Matic I, Hilger M et al (2009) A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc* 4:698–705

Chapter 43

***N*-Glycoprotein Enrichment by Lectin Affinity Chromatography**

Eliel Ruiz-May, Carmen Catalá, and Jocelyn K.C. Rose

Abstract

Lectins are proteins that bind to sugars with varying specificities and several have been identified that show differential binding to structurally variable glycans attached to glycoproteins. Consequently, lectin affinity chromatography represents a valuable tool for glycoproteome studies, allowing enrichment of glycoproteins in samples prior to their identification by mass spectrometry (MS). From the perspective of plant scientists, lectin enrichment has proven useful for studies of the proteomes of the secretory pathways and cell wall, due to the high proportion of constituent proteins that are glycosylated. This chapter outlines a strategy to generate samples enriched with glycoproteins from bulk plant tissues prior to further characterization by MS, or other techniques.

Key words Glycoprotein, Lectins, Affinity chromatography, Concanavalin A

1 Introduction

Glycosylation is a highly complex posttranslational modification associated with many eukaryotic proteins, involving the attachment of oligosaccharide moieties and their subsequent modification by a large battery of glycan modifying enzymes. This results in structurally diverse pool of glycoproteins and glycoforms [1, 2]. In plants, these glycoproteins can be classified in *N*-glycoproteins, where *N*-glycans are covalently linked to asparagine in the sequon N-X-(S/T), where X can be any amino acid except proline [3], and *O*-glycoproteins, where in the glycan is attached to the hydroxyl group of serine, threonine, or hydroxylated proline residues [4–7], with no specific sequon. Several analytical platforms have been developed for systematic studies of glycoproteins from bacteria, yeast and animals [8–14], but there are not yet an analogous system for plant glycoproteomes. A typical workflow might comprise front-end enrichment of glycoproteins/glycopeptides, identification of peptide sequence, determination of glycosylation sites and

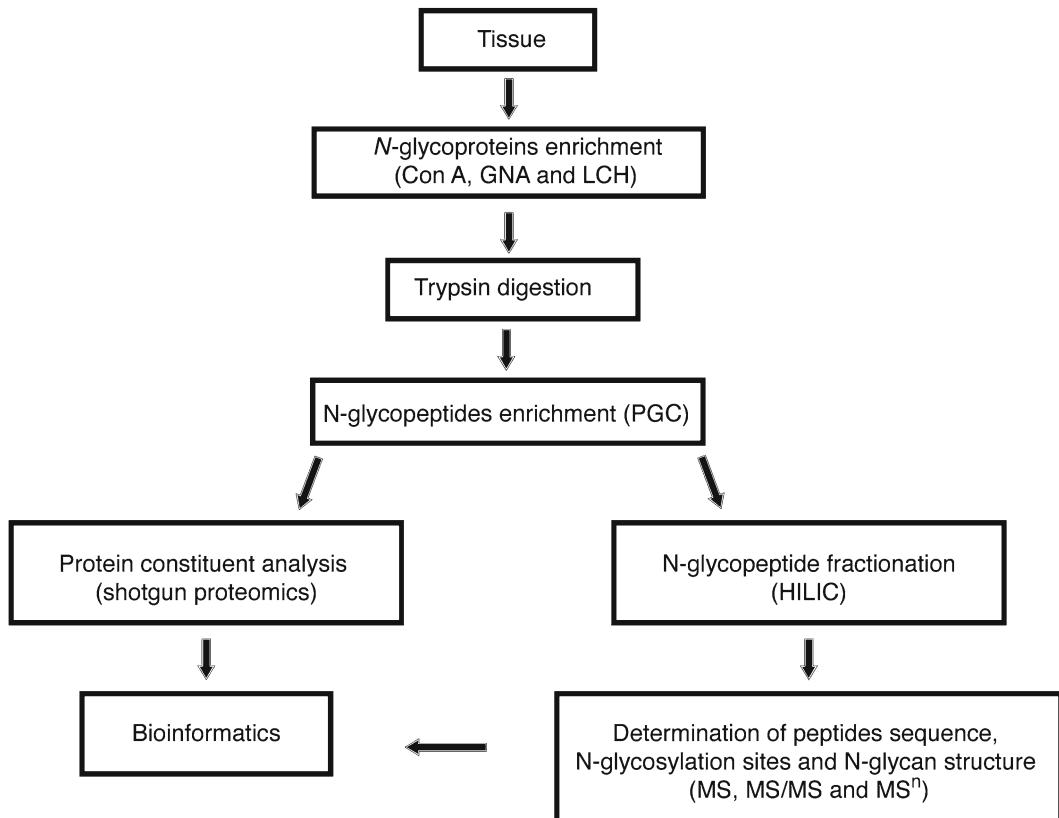


Fig. 1 Strategy workflow for systematic study of *N*-glycoproteins

site occupancy, and interpretation of glycan structure and glycoforms (Fig. 1). However, several factors can complicate the glycoprotein analysis, including the complexity of biological samples and low protein abundance. Therefore, the reduction of the complexity and enrichment of glycoproteins can represent an effective first step. This can be accomplished by lectin affinity chromatography [15–17] and chemical methods, such as hydrazine chemistry [18] and boronic acid [19]. However, lectin affinity chromatography with Concanavalin A has been more commonly used for studies of plant glycoproteins [15–17] and a number of glycoproteomic analyses have been reported in the last few years employing lectin affinity as an early enrichment step [20], including those that use multiple lectins to increase the population of captured glycoproteins [17]. To this end, a range of lectins with different affinities is now commercially available (Table 1). However, the chemical methods should be considered as complementary analytical approaches for a systematic study [14, 21], although they are not further discussed here.

In this chapter we present a protocol to enrich for *N*-glycoproteins from plant tissues using lectin affinity chromatography, and to prepare the samples for downstream experiments designed to identify the proteins using mass spectrometry (MS).

Table 1
Commercially available glycan-binding lectins for the enrichment of glycoproteins

| Lectin type | Name | Source | Affinity |
|--|---|--------------------------------|---|
| Mannose binding lectins | Con A (concanavalin A) | <i>Canavalia ensiformis</i> | High-mannose, hybrid and biantennary complex type N-glycans [30–33] |
| | LCH (Lentil lectin) | <i>Lens culinaris</i> | Fucosylated core region of bi- and triantennary complex type N-glycans [34] |
| | GNA (Snowdrop lectin) | <i>Galanthus nivalis</i> | α -1,3 and α -1,6 link high mannose structure [35] |
| Fucose binding lectins | UEA (<i>Ulex europaeus</i> agglutinin) | <i>Ulex europaeus</i> | Fuc α 1-2Gal-R [36] |
| | AAL (<i>Aleuria aurantia</i>) | <i>Aleuria aurantia</i> | Fuc α 1-2Gal β 1-4(Fuc α 13/4) Gal β 1-4GlcNAc; R ₂ -GlcNAc β 1-4(Fuc α 1 6) GlcNAc-R ₁ [37] |
| Galactose/ <i>N</i> -acetylgalactosamine binding lectins | RCA (<i>Ricinus communis</i> Agglutinin) | <i>Ricinus communis</i> | Gal β 1-4GlcNAc β 1-R [38, 39] |
| | PNA (Peanut Agglutinin) | <i>Arachis hypogaea</i> | Gal β 1-3GalNAc α 1-Ser/Thr (T-Antigen) [40–42] |
| | AIL (Jacalin) | <i>Artocarpus integrifolia</i> | (Sia)Gal β 1-3GalNAc α 1-Ser/Thr (T-Antigen) [43] |
| | VVL (Hairy vetch lectin) | <i>Vicia villosa</i> | GalNAc α -Ser/Thr (Tn-Antigen) [44, 45] |
| Sialic acid/ <i>N</i> -acetylglucosamine binding lectins | WGA (Wheat Germ agglutinin) | <i>Triticum vulgaris</i> | GlcNAc β 1-4GlcNAc β 1-4GlcNAc, Neu5Ac (sialic acid) [46, 47] |
| | SNA (Elderberry lectin) | <i>Sambucus nigra</i> | Neu5Ac α 2-6Gal(NAc)-R [48] |
| | MAL (<i>Maackia amurensis</i> lectin) | <i>Maackia amurensis</i> | Neu5Ac/Gc α 2-3Gal β 1-4GlcNAc β 1-R [49] |

2 Materials

Note: mention of specific companies or pieces of equipment does not represent an endorsement by the authors.

2.1 Lectin Resins

Several lectins coupled to Sepharose beads, magnetic beads [22, 23] or to beads packed into chromatography columns/cartridges are commercially available (e.g., Qiagen) to facilitate the enrichment of a diverse population of glycoproteins (Table 1). Of these, Concanavalin A (Con A) is the most frequently used. It is possible to work with several lectins sequentially [24, 25], in parallel [8, 14, 26]

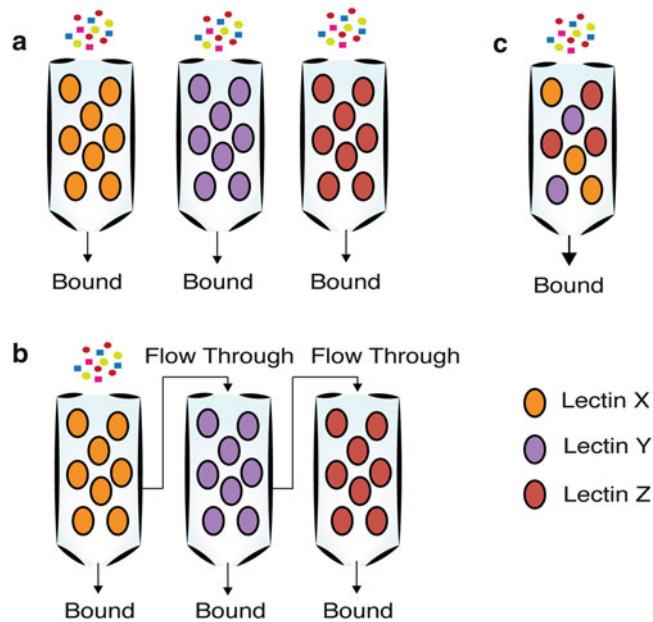


Fig. 2 Different modalities of lectin affinity chromatography. Lectin enrichment can be done in parallel (a), sequentially (b), and using a mix of several lectins (c). The type of lectins used can be tailored to the kind of glycoprotein that is being targeted

and as mixtures [27], as illustrated in Fig. 2. If the main goal is to enrich for *N*-glycoproteins an effective option is the use of mannose binding lectins (Table 1).

2.2 Porous Graphitic Carbon (PGC) Columns

Prepacked PGC columns (1 mL, brand name Hypersep Hypercarb) can be purchased from Thermo Scientific.

2.3 Buffer Solutions

Buffer solutions must be prepared with bi-distilled water, filtered with 0.45 μm filters, degassed, and precooled to 4 °C before starting the main protocol.

1. Stock buffers: prepare 1 M Tris, pH 7.0 (Solution A) and 5 M NaCl (Solution B) solutions. Dissolve 121.4 g Tris in 900 mL water, and adjust the pH with HCl, and the final volume to 1 L with water. Dissolve 292 g NaCl in 500 mL water and adjust the volume to 1 L.
2. Protein Extraction buffer: 25 mM Tris, pH 7.0, 0.5 M NaCl, 0.2 M CaCl₂, and 20 $\mu\text{L}/\text{g}$ fresh weight protease inhibitor cocktails. For 100 mL of protein extraction buffer mix: 2.5 mL A solution, 10 mL B solution, 2.94 g CaCl₂, and 20 $\mu\text{L}/\text{g}$ fresh weight protease inhibitor cocktails and adjust the volume to 100 mL (see Note 1).
3. Binding buffer (Lectin): 20 mM Tris–HCl, 0.5 M NaCl, 1 mM CaCl₂, 1 mM MnCl₂, and 1 mM MgCl₂. Mix 2 mL Solution

A, 10 mL Solution B, 0.0147 g CaCl₂, 0.0197 g MnCl₂, and 0.0203 g MgCl₂ and adjust the volume to 100 mL.

4. Elution buffer (Lectin): binding buffer plus 0.5 M α -methyl-D-mannopyranoside. Dissolve 9.71 g α -methyl-D-mannopyranoside in 100 mL binding buffer.
5. 100 mM Ammonium bicarbonate (Na₂CO₃): dissolve 1.05 g Na₂CO₃ in water and adjust the final volume to 100 mL.
6. Resuspension buffer: 8 M urea, 100 mM Na₂CO₃. Dissolve 0.961 g urea in 2 mL 100 mM Na₂CO₃. Prepare this solution fresh immediately prior to use.
7. 2 M Dithiothreitol (DTT): dissolve 0.154 g DTT in 100 mM Na₂CO₃, adjust the final volume to 500 μ L, divide into small aliquots (e.g., 5 μ L) and store at -80 °C until use.
8. 257 mM Iodoacetamide: dissolve 0.023 g iodoacetamide in 100 mM Na₂CO₃ and adjust the final volume 500 μ L. Prepare this solution fresh immediately prior to use.
9. 1 M NaOH: dissolve 4 g NaOH in water and adjust the volume to 100 mL.
10. 30 % Acetic acid: mix 30 mL glacial acetic acid with 70 mL water. This solution should be prepared in a fume hood to avoid potentially toxic fumes.
11. Porous graphitic Carbon (PGC) wash solvent: 5 % acetonitrile, 0.1 % formic acid (v/v) in water. Mix 5 mL acetonitrile and 100 μ L formic acid and adjust the volume to 100 mL with water. This solution should be prepared in the fume hood.
12. PGC Elution solvent: 50 % acetonitrile, 0.1 % formic acid (v/v) in water. Mix 50 mL and 100 μ L formic acid and adjust the volume to 100 mL with water. This solution should be prepared in the fume hood.

3 Methods

3.1 Tissue Collection

1. Weigh out 3 g of plant material (fresh or flash frozen in liquid nitrogen and stored at -80 °C). Replicated biological samples should also be considered, depending on the experimental goals.

3.2 Protein Extraction

1. Powder the samples in liquid nitrogen using a pestle and mortar and add a tenth mass of polyvinylpolypyrrolidone (PVPP; 1 g/10 g fresh weight) to help remove phenolic compounds.
2. Homogenize the material in three volumes of extraction buffer (45 mL) with a tissue homogenizer (e.g., Polytron, Kinematica) for 15 s, and then filter with Miracloth (Calbiochem). Shake the crude extract at 5 rpm on a rocking platform for 2 h at 4 °C.

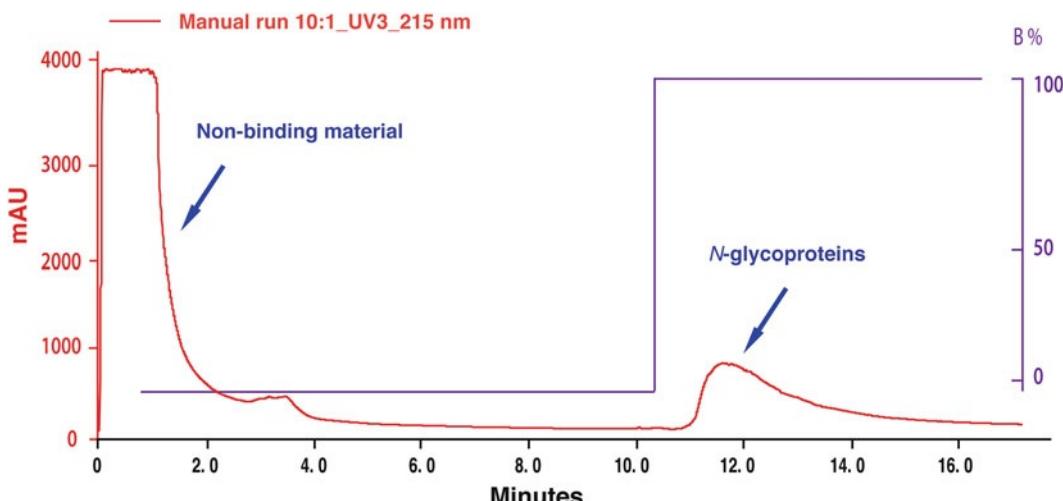


Fig. 3 A typical UV absorbance chromatogram of the *N*-glycoprotein enriched fraction from a crude protein extract

3. Centrifuge the crude extract at $15,000 \times g$ for 30 min. Recover the supernatant and filter through a $0.45 \mu\text{m}$ syringe filter. Set aside 1 mL for a subsequent protein quantification assay and visualization by SDS-PAGE analysis.
4. Protein quantification: use bicinchoninic acid (BCA, 28) or the Bradford assay [29] for protein quantification, with bovine serum albumin (BSA) to generate a standard curve.

3.3 *N*-Glycoproteins Enrichment Using Lectin Cartridges

1. Equilibration step: use 10 column volumes of binding buffer (50 mL of binding buffer in the case of 5 mL chromatography cartridges) at a flow rate of 0.08 mL/min (see Note 2).
2. Loading step: samples should be loaded slowly to increase binding of the glycoproteins to the lectins (see Note 3). Collect the flow through for subsequent analysis by SDS-PAGE.
3. Washing step: wash the column with ten column volumes of binding buffer, or until the absorbance at 280 nm returns to the baseline value.
4. Elution step: elute the bound protein with five column volumes of elution buffer at a flow rate of approximately 0.75 mL/min. Collect 1 mL fractions. A typical elution chromatogram of glycoproteins is shown in Fig. 3.

3.4 *N*-Glycoprotein Batch Enrichment

1. Equilibration step: Mix 150 μL of Con A slurry (commercially obtained lectin resins are typically shipped as a suspension with 20 % ethanol as a preservative) with 5 mL of binding buffer in a 15 mL tubes. Briefly shake the suspension and centrifuge at $1,000 \times g$ for 2 min. Discard the supernatant and repeat the process twice more for a total of three incubations/washes.

2. Loading step: add the protein extract from Subheading 3.2, step 3 to the resin and shake on a rocking platform for 1 h at 4 °C. Centrifuge at $1,000 \times g$ for 2 min and recover the supernatant for subsequent analysis by SDS PAGE.
3. Washing step: Add 15 mL of binding buffer to the Con A resin with the bound protein sample and mix thoroughly for 1 min. Centrifuge and recover supernatant/flow through. Repeat this step twice more for a total of three washes. The spectrophotometric absorbance (OD 280 nm) value can be used to approximate the protein concentration in the supernatant and used as an indication of when no more protein is being eluted from the resin. In general, three washes are sufficient to remove most of the nonspecifically bound protein, but additional washes can be used if significant amounts of proteins are still being eluted after three washes.
4. Elution step: Add 500 μ L of elution buffer to the resin and mix for 1 min. Centrifuge at $1,000 \times g$ for 2 min and recover the supernatant for analysis. Repeat this step twice more. The glycoproteins will be present in these eluted fractions.

3.5 Concentration and Dialysis

1. Pool the three fractions from the elution step and reduce the volume of the sample by 50 % in a centrifugal concentrator using a 5 kDa cutoff centrifugal concentrator (Amicon Ultra-15, Millipore, Billerica MA). Next perform a solvent exchange step by first adding an equal volume of 100 mM ammonium bicarbonate (e.g., 4 mL of glycoprotein extract plus 4 mL 100 mM ammonium bicarbonate) then again reducing the volume by 50 % by centrifugation. Repeat this solvent exchange step at least three times.
2. Lyophilize the final sample containing the glycoproteins extract and resuspend the sample in 300 μ L resuspension buffer, vortex for 5 min and place in a sonicating water bath for 5 min.
3. Centrifuge the suspension at $13,000 \times g$ for 3 min. Recover the supernatant and set aside 50 μ L for subsequent protein quantification and analysis by SDS-PAGE, and store the remainder at -80 °C if necessary.

3.6 Protein Digestion

1. Mix an aliquot of the sample containing 100 μ g of glycoprotein with resuspension buffer to a final volume of 200 μ L.
2. Add 1 μ L 2 M DDT (the final concentration will be 10 mM) and incubate for 1 h at room temperature. Do not heat the samples above room temperature (see Note 4).
3. Add 20 μ L 257 mM iodoacetamide (the final concentration will be 25 mM) and incubate for 30 min in the dark at room temperature.

4. Add 1.6 mL 100 mM ammonium bicarbonate to achieve a final concentration of 0.99 M urea.
5. Add trypsin to achieve a final trypsin–glycoprotein sample ratio of 1:20 (5 µg trypsin–100 µg glycoprotein) and incubate for 16 h at 37 °C

3.7 Affinity Purification of N-Glycopeptides Using a Porous Graphitic Carbon (PGC) Column

Equilibration step: Pass the following solutions sequentially through the PGC column and discard the flow through:

1. 1 mL 1 M NaOH.
2. 2 mL water.
3. 1 mL 30 % acetic acid.
4. 2 mL water.
5. 1 mL elution solvent (50 % acetonitrile, 0.1 % formic acid [v/v] in water).
6. 1 mL wash solvent (5 % acetonitrile, 0.1 % formic acid [v/v] in water).

Do not allow air to enter the column.

Loading step:

1. Adjust the pH of the samples to 5.0 with 0.1 % trifluoroacetic acid (TFA).
2. Slowly load the sample (approximately 1 drop/s) onto the column followed by 1 mL water. Collect the flow through for subsequent analysis by MS if needed to evaluate the unbound peptides.

Desalting step:

1. Pass 1 mL wash solvent through the column and recover the flow through for subsequent analysis by MS if needed. This sample can then be desalted using a conventional reverse phase C18 solid phase extraction prior to analysis by MS (see Note 5).

Elution step:

1. Pass 1 mL elution solvent through the cartridge bed and collect the flow through.
2. Gently pass air through the column to elute all the solvent into a collection tube.
3. Dry the sample in a rotary evaporator (e.g., Savant) prior to downstream MS analysis.

After the glycoprotein digestion a pool of peptides and glycopeptides will be present in the enriched samples. After the desalting and cleanup steps various downstream experiments are possible. These include protein sequence identification, characterization of glycopeptide sequences, determination of the glycosylation sites and interpretation of *N*-glycan structure in the *N*-glycopeptides.

4 Notes

1. Phenylmethanesulfonylfluoride (PMSF) can be added to the protease inhibitor cocktail (1 mM final concentration) for better protease inhibition.
2. It is important to note that cartridges used for the first time can leak some of the lectin during the glycoprotein elution step. Therefore, it is recommended to wash and equilibrate the columns with more than 10 volumes of binding buffer prior to application of the sample.
3. When the volume of the sample is large and the protein concentration is low, it may help to perform the loading step in a cold room with a peristaltic pump. It is important to load the samples at as a low a rate as possible. Alternatively, a fast protein liquid chromatography (FPLC; GE Healthcare) system with a sample loop can be used.
4. Protein samples in buffers containing urea must not be heated because this can cause carbamylation: urea in solution is in equilibrium with ammonium cyanate and the isocyanic acid reacts with protein amino groups. This results in considerable charge heterogeneity, which complicates subsequent MS analysis.
5. The majority of the glycopeptides bind to the PGC matrix; however, some peptides and glycopeptides do not bind. If an objective is to identify non-glycosylated peptides that were derived from the original glycoproteins, then analysis of the flow through may also be included at this point.

Acknowledgments

Funding to JKCR for research in this area is provided by the NSF Plant Genome Research Program (DBI-0606595) and the New York State Office of Science, Technology and Academic Research (NYSTAR).

References

1. Marino K, Bones J, Kattla JJ et al (2010) A systematic approach to protein glycosylation analysis: a path through the maze. *Nat Chem Biol* 6:713–723
2. An HJ, Froehlich JW, Lebrilla CB (2009) Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Curr Opin Chem Biol* 13:421–426
3. Pless DD, Lennarz WJ (1977) Enzymatic conversion of proteins to glycoproteins. *Proc Natl Acad Sci USA* 74:134–138
4. Matsuoka K, Watanabe N, Nakamura K (1995) O-glycosylation of a precursor to a sweet potato vacuolar protein, sporamin, expressed in tobacco cells. *Plant J* 8:877–889
5. Cho YP, Chrispeels MJ (1976) Serine-O-galactosyl linkages in glycopeptides from carrot cell-walls. *Phytochemistry* 15: 165–169
6. Showalter AM (2001) Arabinogalactan-proteins: structure, expression and function. *Cell Mol Life Sci* 58:1399–1417

7. Velasquez SM, Ricardi MM, Dorosz JG et al (2011) *O*-glycosylated cell wall proteins are essential in root hair growth. *Science* 332: 1401–1403
8. Zielinska DF, Gnad F, Wisniewski JR et al (2010) Precision mapping of an *in vivo* *N*-glycoproteome reveals rigid topological and sequence constraints. *Cell* 141:897–907
9. Kaji H, Kamiie J, Kawakami H et al (2007) Proteomics reveals *N*-linked glycoprotein diversity in *Caenorhabditis elegans* and suggests an atypical translocation mechanism for integral membrane proteins. *Mol Cell Proteomics* 6:2100–2109
10. Wollscheid B, Bausch-Fluck D, Henderson C et al (2009) Mass-spectrometric identification and relative quantification of *N*-linked cell surface glycoproteins. *Nat Biotechnol* 27: 378–386
11. Gundry RL, Raginski K, Tarasova Y et al (2009) The mouse C2C12 myoblast cell surface *N*-linked glycoproteome: identification, glycosite occupancy, and membrane orientation. *Mol Cell Proteomics* 8:2555–2569
12. Liu T, Qian WJ, Gritsenko MA et al (2005) Human plasma *N*-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. *J Proteome Res* 4:2070–2080
13. Bunkenborg J, Pilch BJ, Podtelejnikov AV et al (2004) Screening for *N*-glycosylated proteins by liquid chromatography mass spectrometry. *Proteomics* 4:454–465
14. Lee A, Kolarich D, Haynes PA et al (2009) Rat liver membrane glycoproteome: enrichment by phase partitioning and glycoprotein capture. *J Proteome Res* 8:770–781
15. Minic Z, Jamet E, Negroni L et al (2007) A sub-proteome of *Arabidopsis thaliana* mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. *J Exp Bot* 58:2503–2512
16. Catala C, Howe KJ, Hucko S et al (2011) Towards characterization of the glycoproteome of tomato (*Solanum lycopersicum*) fruit using Concanavalin A lectin affinity chromatography and LC-MALDI-MS/MS analysis. *Proteomics* 8:1530–1544
17. Zhang Y, Giboulot A, Zivy M et al (2010) Combining various strategies to increase the coverage of the plant cell wall glycoproteome. *Phytochemistry* 10:1109–1123
18. Zhang H, Li XJ, Martin DB et al (2003) Identification and quantification of *N*-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* 21:660–666
19. Sparbier K, Koch S, Kessler I et al (2005) Selective isolation of glycoproteins and glycopeptides for MALDI-TOF MS detection supported by magnetic particles. *J Biomol Tech* 16:407–413
20. Nilsson CL (2011) Lectin techniques for glycoproteomics. *Curr Proteomics* 8:248–256
21. McDonald CA, Yang JY, Marathe V et al (2009) Combining results from lectin affinity chromatography and glycocapture approaches substantially improves the coverage of the glycoproteome. *Mol Cell Proteomics* 8: 287–301
22. Choi E, Loo D, Dennis JW et al (2011) High-throughput lectin magnetic bead array-coupled tandem mass spectrometry for glycoprotein biomarker discovery. *Electrophoresis* 32:3564–3575
23. Yang G, Cui T, Chen Q et al (2012) Isolation and identification of native membrane glycoproteins from living cell by concanavalin A-magnetic particle conjugates. *Anal Biochem* 421:339–341
24. Yamamoto K, Tsuji T, Osawa T (1998) Analysis of asparagine-linked oligosaccharides by sequential lectin-affinity chromatography. *Methods Mol Biol* 76:35–51
25. Yamamoto K, Tsuji T, Osawa T (1995) Analysis of asparagine-linked oligosaccharides by sequential lectin affinity chromatography. *Mol Biotechnol* 3:25–36
26. Yang Z, Harris LE, Palmer-Toy DE et al (2006) Multilectin affinity chromatography for characterization of multiple glycoprotein biomarker candidates in serum from breast cancer patients. *Clin Chem* 52:1897–1905
27. Li L, Wang L, Zhang W et al (2004) Correlation of serum VEGF levels with clinical stage, therapy efficacy, tumor metastasis and patient survival in ovarian cancer. *Anticancer Res* 24:1973–1979
28. Smith PK, Krohn RI, Hermanson GT et al (1985) Measurement of protein using biinchoninic acid. *Anal Biochem* 150:76–85
29. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248–254
30. Hardman KD, Ainsworth CF (1972) Myo-inositol binding site of concanavalin A. *Nature* 237:54–55
31. Hardman KD, Ainsworth CF (1972) Structure of concanavalin A at 2.4-A resolution. *Biochemistry* 11:4910–4919
32. Sumner JB (1919) The globulins of the jack bean, *Canavalia ensiformis*. *J Biol Chem* 37:137–142
33. Sumner JB, Howell SF (1936) Identification of hemagglutinin of jack bean with concanavalin A. *J Bacteriol* 32:227–237
34. Kornfeld K, Reitman ML, Kornfeld R (1981) The carbohydrate-binding specificity of pea

- and *lentil lectins*. Fucose is an important determinant. *J Biol Chem* 256:6633–6640
- 35. Shibuya N, Goldstein IJ, Van Damme EJ et al (1988) Binding properties of a mannose-specific lectin from the snowdrop (*Galanthus nivalis*) bulb. *J Biol Chem* 263: 728–734
 - 36. Onozaki K, Homma Y, Hashimoto T (1979) Purification of an L-fucose binding lectin from *Ulex europeus* by affinity column chromatography. *Experientia* 35:1556–1557
 - 37. Kochibe N, Furukawa K (1980) Purification and properties of a novel fucose-specific hemagglutinin of *Aleuria aurantia*. *Biochemistry* 19:2841–2846
 - 38. Nicolson GL, Blaustein J, Etzler ME (1974) Characterization of two plant lectins from *Ricinus communis* and their quantitative interaction with a murine lymphoma. *Biochemistry* 13:196–204
 - 39. Nicolson GL, Blaustein J (1972) The interaction of *Ricinus communis* agglutinin with normal and tumor cell surfaces. *Biochim Biophys Acta* 266:543–547
 - 40. Farrar GH, Uhlenbruck G, Holz G (1980) Comparison of isolated peanut agglutinin receptor glycoproteins from human, bovine and porcine erythrocyte membranes. *Biochim Biophys Acta* 603:185–197
 - 41. Lotan R, Skutelsky E, Danon D et al (1975) The purification, composition, and specificity of the anti-T lectin from peanut (*Arachis hypogaea*). *J Biol Chem* 250:8518–8523
 - 42. Novogrodsky A, Lotan R, Ravid A et al (1975) Peanut agglutinin, a new mitogen that binds to galactosyl sites exposed after neuraminidase treatment. *J Immunol* 115:1243–1248
 - 43. Bunn-Moreno MM, Campos-Neto A (1981) Lectin(s) extracted from seeds of *Artocarpus integrifolia* (jackfruit): potent and selective stimulator(s) of distinct human T and B cell functions. *J Immunol* 127:427–429
 - 44. Grubhoffer L, Ticha M, Kocourek J (1981) Isolation and properties of a lectin from the seeds of hairy vetch (*Vicia villosa* Roth). *Biochem J* 195:623–626
 - 45. Kimura A, Wigzell H, Holmquist G et al (1979) Selective affinity fractionation of murine cytotoxic T lymphocytes (CTL). Unique lectin specific binding of the CTL associated surface glycoprotein, T 145. *J Exp Med* 149:473–484
 - 46. Nagata Y, Burger MM (1974) Wheat germ agglutinin. Molecular characteristics and specificity for sugar binding. *J Biol Chem* 249: 3116–3122
 - 47. Nagata Y, Burger MM (1972) Wheat germ agglutinin. Isolation and crystallization. *J Biol Chem* 247:2248–2250
 - 48. Shibuya N, Goldstein IJ, Broekaert WF et al (1987) The elderberry (*Sambucus nigra* L.) bark lectin recognizes the Neu5Ac(alpha 2-6) Gal/GalNAc sequence. *J Biol Chem* 262: 1596–1601
 - 49. Yamamoto K, Konami Y, Irimura T (1997) Sialic acid-binding motif of *Maackia amurensis* lectins. *J Biochem* 121:756–761

Chapter 44

Rapid and High-Throughput *N*-Glycomic Analysis of Plant Glycoproteins

**Kentaro Kaneko, Takeshi Shiraya, Toshiaki Mitsui,
and Shin-ichiro Nishimura**

Abstract

Glycoprotein is a major element in higher organisms including mammals and plants. It is widely accepted that variation in cellular *N*-glycome is related to modulation in dynamic cellular mechanisms such as cell-cell adhesion, cell activation, and malignant alterations in mammalian cells. However, the physiological importance of glycan modification of glycoproteins in plant cells is still a matter of dispute. Therefore, a comprehensive and high-throughput analysis of *N*-glycome in plant glycoproteins is needed. Here, an application of the glycoblottting-mass spectrometry technique to plant glycoprotein research is described.

Key words AmyI-1, BlotGlyco beads, Glycoblottting, *N*-glycome, Plant glycoprotein

1 Introduction

Most of the proteins in higher animals and plants are glycoproteins bearing an *N*-linked oligosaccharide side chain. The consensus sequence Asn-X-Ser/Thr is found in this sugar chain binding-site of plant glycoprotein, and the reducing group of the sugar chain binds to the nitrogen atom of the asparagine residue [1, 2]. *N*-linked oligosaccharide chains conjugated to plant glycoproteins typically have a common core structure, and are classified into high mannose- and complex-type glycan chains. The structure of the high mannose-type glycan chain is $\text{Man}_{(5-9)}\text{GlcNAc}_2$, while the complex-type glycan chain shows a diverse structure, with the terminal sugars *N*-acetylglucosamine, xylose, fucose, and galactose added to the core structure. The conjugation of α 1,3-fucose and β 1,2-xylose residue to the core structure is peculiar to plant glycoprotein [2–4]. Moreover, the small oligosaccharide chains, called paucimannosidic-type glycans, in which the xylose and/or fucose residue is conjugated to the core structure of $\text{Man}_{(2-3)}\text{GlcNAc}_2$, are frequently observed in plant cells [5]. The plant glycans containing sugar with electric charge,

such as sialic acid residues, which are detected in mammalian glycoproteins, have not been registered in the KEGG Glycan [6], GlycoMod Tool [7], and GlycoSuite [8] databases.

In higher plant cells, the transfer of $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ precursor glycan to the nascent polypeptide, and the modification of the glycan chain, occurs through a conserved biosynthetic pathway [9]. A set of glucosidase and mannosidase converts $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ to $\text{Man}_8\text{GlcNAc}_2$ sequentially in the endoplasmic reticulum, and the $\text{Man}_8\text{GlcNAc}_2$ glycan is then trimmed to $\text{Man}_5\text{GlcNAc}_2$ by *cis*-Golgi-resident mannosidase I [10], and further modified into complex chains by Golgi mannosidase II and glycosyltransferases. It has been reported that *N*-acetylglucosaminyltransferase I and *N*-acetylglucosaminyltransferase II localize in the *cis*-Golgi to medial-Golgi compartments; and β 1,2-xylosyltransferase, α 1,3-fucosyltransferase, α 1,4-fucosyltransferase and β 1,3-galactosyltransferase localize in the medial-Golgi to *trans*-Golgi compartments [9]. Interestingly, recent investigations have provided evidence on the trafficking of glycoproteins [11–15] to the plastids through the secretory pathway. There is probably considerable membrane traffic between the endomembrane system and the plastids.

However, the physiological importance of glycan modification of glycoproteins in plants remains obscure. To clarify this, it is necessary to employ a rapid, efficient, sensitive and high-throughput method for analyzing plant *N*-linked oligosaccharide chains. Here, we describe an application of the glycoblotting-mass spectrometry method [16–20] that has made comprehensive *N*-glycome analysis possible in recent years.

2 Materials

Milli-Q water is used in all solutions containing water. The methanol and acetonitrile (ACN) are LC-MS grade.

2.1 Sample Preparation

1. 50 mM citrate-phosphate buffer, pH 5: Dissolve 0.48 g of citric acid and 0.39 g of sodium dihydrogen phosphate dehydrate in about 50 mL of water, and adjust pH with NaOH (1 N). Make up to 50 mL with water. Store at room temperature.
2. 100 mM (\pm)-dithiothreitol (DTT): Dissolve 1.54 g of DTT in 10 mL of water to prepare 1 M solution. Store at -20 °C. Dilute to 100 mM before use.
3. 100 mM 2-iodoacetamide: Dissolve 0.18 g of 2-iodoacetamide in 1 mL of water. Prepare before use.

2.2 Hydrolysis of *N*-Glycans

1. Protease K (0.6 U/ μ L, Roche, Basel, Switzerland): Store at 4 °C.
2. Glycopeptidase A (50 μ U/ μ L, Roche): Store at 4 °C.

2.3 Glycoblotting

1. MultiScreen Solvent filter plate (Millipore, Billerica, USA) and vacuum manifold (Waters, Milford, USA).
2. Oligosaccharide Purification Kit BlotGlyco® (SUMITOMO BAKELITE, Tokyo, Japan): BlotGlyco® beads (Hydrazide Functionalized Polymer) [18], $\text{N}\alpha$ -((aminoxy)acetyl)tryptophanylarginine methyl ester (aoWR) [17].
3. 50 μM chitotetraose (GN4): Dissolve 50 mg of GN4 in 1.2 mL of water to prepare a 50 mM solution. Store at -20°C . Dilute to 50 μM before use.
4. 2 % (v/v) acetic acid/ACN: Mix 9.8 mL of ACN and 0.2 mL of acetic acid. Store at room temperature.
5. 2 M guanidine hydrochloride: Dissolve 1.9 g of guanidine hydrochloride in 10 mL of water. Store at room temperature.
6. 1 % (v/v) trimethylamine-methanol: Mix 9.9 mL of trimethylamine and 0.1 mL of methanol. Store at room temperature.
7. 10 % (v/v) acetic anhydride-methanol: Mix 100 μL of acetic anhydride and 900 μL of methanol. Prepare before use.
8. 10 mM HCl: Store at room temperature.

2.4 Removal of Free

Labeling Reagent by Hydrophilic Interaction

Chromatography (HILIC)

1. MassPREP HILIC 96-well Plate (Waters, Milford, USA).
2. 1 % acetic acid/99 % ACN: Mix 9.9 mL of ACN and 0.1 mL of acetic acid.
3. 1 % acetic acid/95 % ACN: Mix 9.5 mL of ACN, 0.1 mL of acetic acid and 0.4 mL of water.
4. 1 % acetic acid/5 % ACN: Mix 0.5 mL of ACN, 0.1 mL acetic acid and 9.4 mL of water.
5. 1 % (v/v) acetic acid/water: Mix 0.1 mL of acetic acid and 9.9 mL of water.

2.5 MS Analysis of N-Glycans

1. 10 mg/mL 2,5-dihydroxybenzoic acid (DHB): Dissolve 10 mg of DHB in 1 mL of 30 % ACN.
2. MTP 384 target plate ground steel TF (Burker Daltonics, Billerica, USA).

3 Methods

3.1 Sample Preparation

1. Freeze-dry 50 μg of glycoproteins (see Note 1).
2. Dissolve the glycoprotein sample in 20 μL of 50 mM citrate-phosphate buffer (pH 5.0) and incubate at 60°C for 10 min.
3. Add 2 μL of 100 mM DTT to the sample and incubate at 60°C for 30 min. After adding 3 μL of 100 mM iodoacetamide, the mixture is incubated in the dark at room temperature for 30 min (see Note 2).

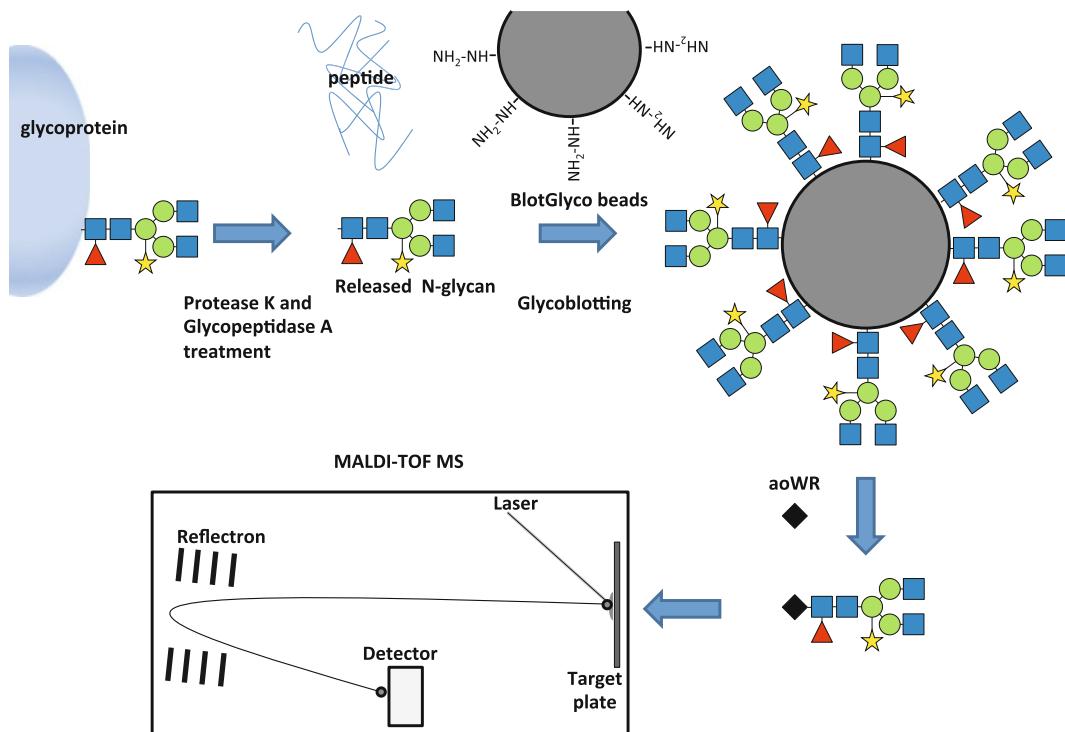


Fig. 1 Schematic illustration of the glycoblotting method

3.2 Release of N-Glycans from Glycoproteins

1. Incubate the protein sample with 2 μ L of protease K (0.6 U/ μ L) at 37 °C for 1 h, and stop the proteolysis reaction by heating at 90 °C for 10 min.
2. After cooling to room temperature, the mixture is incubated with 5 μ L of glycopeptidase A (50 μ U/ μ L) at 37 °C for 16 h (see Note 3).

3.3 Glycoblotting

The “glycoblotting” method involves easily and rapidly obtaining N-glycan derivatives tagged with a variety of compounds from a crude glycoprotein sample contaminated with amino acids, DNA and RNA, etc. The hydrazide ligands ($-\text{NHNH}_2$) of the BlotGlyco beads specifically capture glycans digested from the glycoproteins, and the captured glycans are labeled and released by a labeling reagent such as aoWR (Fig. 1).

1. Pour 500 μ L of BlotGlyco beads suspension onto a well of MultiScreen Solvent filter plate (see Note 4), and remove the suspension solvent by vacuum.
2. Add 25 μ L of the digested mix and 1 μ L of 50 μ M GN4 (see Note 5) to the well.
3. Add 234 μ L of 2 % (v/v) acetic acid/ACN to the well, and dry up at 80 °C for 1 h (see Note 6).

- Wash twice with 200 μ L of 2 M guanidine hydrochloride, twice with 200 μ L of water, and twice with 200 μ L of 1 % (v/v) trimethylamine–methanol, sequentially.
- Add 100 μ L of 10 % (v/v) acetic anhydride–methanol to the well, and incubate at room temperature for 30 min.
- Remove the solvent, and wash twice with 200 μ L of 10 mM HCl, and twice with 200 μ L of methanol.
- Wash with 200 μ L of water.
- Add 20 μ L of aoWR solution and 180 μ L of 2 % (v/v) acetic acid/ACN to the well, then incubate at 80 °C for 1 h (see Note 7).
- Elute and collect the aoWR-labeled glycans with 100 μ L of water. Store at –20 °C (see Note 8).

3.4 Removal of Free Labeling Reagent by HILIC

The HILIC treatment should be carried out just before applying the samples to the MALDI-target plate (see Note 9).

- Dilute the collected glycan sample with 9 volumes of 1 % (v/v) acetic acid/ACN.
- A well of MassPREP HILIC 96-well Plate is washed twice with 200 μ L of 1 % (v/v) acetic acid/water, and equilibrated twice with 200 μ L of 1 % acetic acid/95 % ACN (see Note 4).
- Load the diluted sample onto the equilibrated well of HILIC. After naturally dropping for a few minutes, the well retaining the glycans is washed twice with 200 μ L of 1 % acetic acid/95 % ACN.
- Add 100 μ L of 1 % acetic acid/5 % ACN and collect the eluent from the well.
- Dry up the eluent by Speed Vac.

3.5 MS Analysis of N-Glycans

- Dissolve the dried-up sample with 1 μ L of water, and add 1 μ L of 10 mg/mL DHB (see Note 10).
- Spot aliquots of the sample mixture with matrix (1 μ L) onto two distinct places on MTP 384 target plate ground steel, and dry at ambient temperature.
- The MALDI-TOF-MS spectrum is acquired in a reflector, in positive-ion mode, typically summing 1,000 shots on the Bruker Daltonics Autoflex III or UltraflexIII (see Note 11).
- Pick possible *N*-glycan peaks (m/z) in the spectra using the FlexAnalysis ver. 3 (Bruker Daltonics) software. The glycans' structures are speculated using the GlycoMod Tool (<http://br.expasy.org/tools/glycomod/>) and GlycoSuite Web site (<https://glycosuite.proteomesystems.com/glycosuite/glycodb>) (see Notes 12–14).
- Normalize intensity of the isotopic peak of each glycan with 25 pmol of internal standard (GN4) for each status.

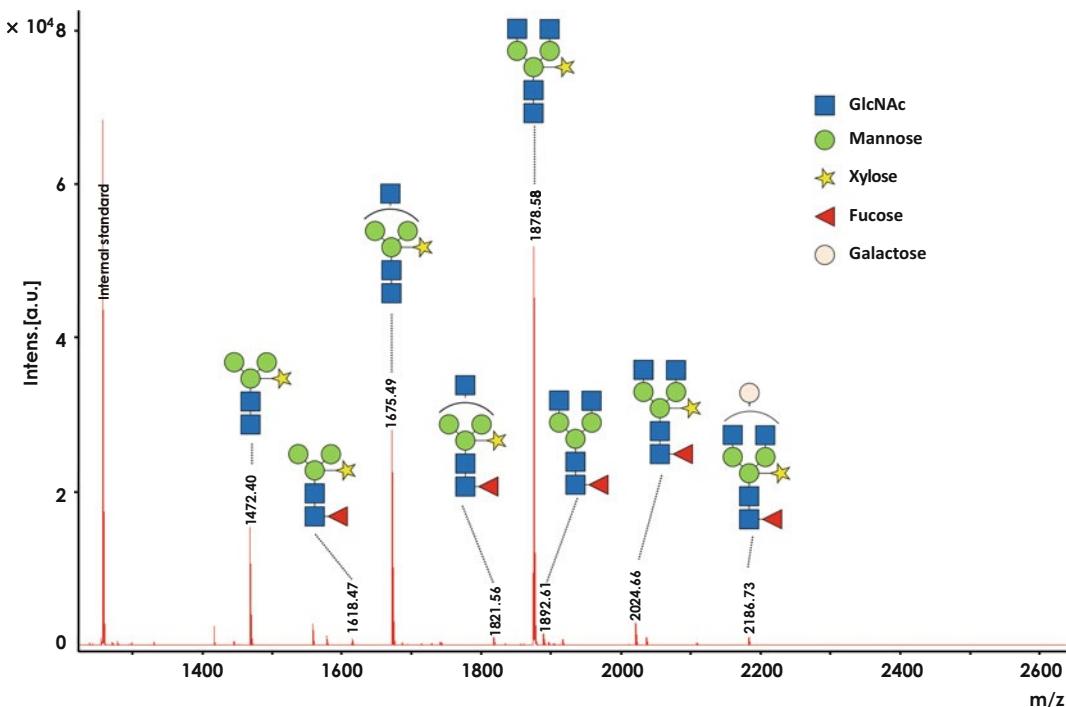


Fig. 2 A typical mass spectrum of *N*-glycans of Amyl-1, obtained by the glycoblotting-mass spectrometry method. All molecular weights represent monoisotopic masses of the respective $[M+aoWR-H_2O+H]^+$ ions of glycan species. $aoWR$, m/z 447.22; H_2O , m/z 18.01. Structural annotation of peaks detected in the MALDI-TOF MS were obtained using the GlycoMod tool online database

3.6 *N*-Glycome in Rice α -Amylase I-1 (Amyl-1)

Major *N*-glycoforms of rice secretory Amyl-1 has been determined using traditional aminopyridine (PA)-derivatization and multidimensional HPLC analysis [3]. The same Amyl-1 glycoprotein sample was subjected to the glycoblotting-mass spectrometry technique, and the obtained results are shown in Fig. 2 and Table 1. Several fucose-containing glycans (i.e., $Man_3GlcNAc_4Fuc_1Xyl_1$, $Man_3GlcNAc_3Fuc_1Xyl_1$, $Man_4GlcNAc_4Fuc_1Xyl_1$, $Man_3GlcNAc_2Fuc_1Xyl_1$, $Man_3GlcNAc_3Fuc_1$) were newly detected using the glycoblotting method.

4 Notes

1. The “glycoblotting” technique is applicable to a wide range of protein amounts (10^{-6} – 10^{-3} g). Use of detergent is allowed, but its concentration must be as low as possible (e.g., Triton X-100 should be less than 0.5 M).
2. The step involving the reduction and alkylation of proteins may be skipped when the protein sample is dissolved in solution.
3. The common *N*-glycan releasing enzyme, PNGase F, should not be used, because this enzyme cannot digest the plant *N*-glycans with conjugation of α 1,3-fucose to the core structure.

Table 1
N-glycome in rice Amyl-1

| <i>m/z</i> ^a [M+H] ⁺ | Δ ^b | Estimated N-glycan structures ^c | pmol/25 μ g protein |
|--|-----------------------|--|-------------------------|
| 831.13 | -0.33 | Internal standard (GN4) | 50.00 |
| 895.15 | -0.19 | Hex ₂ HexNAc ₂ dHex ₁ | 0.01 |
| 911.16 | -0.17 | Hex ₃ HexNAc ₂ | 0.01 |
| 936.21 | -0.16 | Hex ₁ HexNAc ₃ dHex ₁ | 0.02 |
| 1,027.21 | -0.17 | Hex ₂ HexNAc ₂ dHex ₁ Pent ₁ | 0.03 |
| 1,043.19 | -0.18 | Hex ₃ HexNAc ₂ Pent ₁ | 10.01 |
| 1,057.22 | -0.18 | Hex ₃ HexNAc ₂ dHex ₁ | 0.07 |
| 1,114.24 | -0.18 | Hex ₃ HexNAc ₃ | 0.06 |
| 1,189.26 | -0.18 | Hex ₃ HexNAc ₂ dHex ₁ Pent ₁ | 0.54 |
| 1,246.28 | -0.17 | Hex ₃ HexNAc ₃ Pent ₁ | 23.02 |
| 1,260.31 | -0.16 | Hex ₃ HexNAc ₃ dHex ₁ | 0.25 |
| 1,276.37 | -0.10 | Hex ₄ HexNAc ₃ | 0.01 |
| 1,351.35 | -0.14 | Hex ₄ HexNAc ₂ dHex ₁ Pent ₁ | 0.02 |
| 1,392.35 | -0.16 | Hex ₃ HexNAc ₃ dHex ₁ Pent ₁ | 0.76 |
| 1,397.35 | -0.14 | Hex ₆ HexNAc ₂ | 0.01 |
| 1,406.46 | -0.07 | Hex ₃ HexNAc ₃ dHex ₂ | 0.01 |
| 1,408.36 | -0.15 | Hex ₄ HexNAc ₃ Pent ₁ | 0.16 |
| 1,449.37 | -0.16 | Hex ₃ HexNAc ₄ Pent ₁ | 56.02 |
| 1,463.40 | -0.15 | Hex ₃ HexNAc ₄ dHex ₁ | 1.18 |
| 1,559.42 | -0.13 | Hex ₇ HexNAc ₂ | 0.00 |
| 1,595.45 | -0.14 | Hex ₃ HexNAc ₄ dHex ₁ Pent ₁ | 2.26 |
| 1,609.47 | -0.14 | Hex ₃ HexNAc ₄ dHex ₂ | 0.01 |
| 1,652.50 | -0.12 | Hex ₃ HexNAc ₅ Pent ₁ | 0.02 |
| 1,682.48 | -0.15 | Hex ₄ HexNAc ₅ | 0.21 |
| 1,757.52 | -0.13 | Hex ₄ HexNAc ₄ dHex ₁ Pent ₁ | 0.75 |
| 1,771.59 | -0.07 | Hex ₄ HexNAc ₄ dHex ₂ | 0.01 |
| 1,773.57 | -0.08 | Hex ₅ HexNAc ₄ Pent ₁ | 0.01 |
| 1,800.49 | -0.19 | Hex ₃ HexNAc ₃ dHex ₂ Pent ₃ | 0.03 |
| 1,962.57 | -0.16 | Hex ₃ HexNAc ₃ dHex ₄ Pent ₂ | 0.01 |

^a*m/z* of free glycan [M+H]⁺

^bDifference from the theoretical mass value

^cStructural annotation was achieved using GlycoMod tool online database. Hex, hexose; HexNAc, N-acetyl hexosamine; dHex, deoxyhexose; Pent, pentose

4. Up to 96 samples can be treated simultaneously by using a MultiScreen Solvent filter plate and a MassPREP HILIC 96-well Plate.
5. The sample glycans and internal standards should be blended in a balanced manner. It is not recommended to use the internal standards for analyzing the sample without estimating the glycan contents.
6. Dry up until the smell of acetic acid completely disappears.
7. The glycoblotting method can be applied to reductive amination with common fluorescent dyes such as 2-aminobenzamide [20].
8. The sample should be subjected to the HILIC treatment within a week.
9. When the surplus labeling reagent is removed from the sample mixture, the label compound bound to the glycan will be released through chemical disequilibrium.
10. The laser shots should hit the crystals on the periphery of the sample spot with DHM matrix, in order to acquire stable mass spectra.
11. The number, power and gain of the laser shots should be constant for quantitative analysis. Stable isotope-coded derivatization is also available and useful [16–18].
12. Difference (Δ_{std}) from the theoretical mass value of an internal standard ($M+H$ 831.33533 m/z) was determined in each measurement. The identification of glycan was manually performed with the following parameters: Δ value of glycan <2 Da, $\Delta_{\text{std}} + 0.3$ Da $> \Delta > \Delta_{\text{std}} - 0.3$ Da.
13. Acidic sugars such as *N*-acetylneuramic acid, sialic acid, glucuronic acid, sulfate and phosphate-contained saccharides have never been detected in Plant *N*-glycans.
14. It should be noted that artificially methylated glycans (+14 m/z) could be detected in the MS spectrum.

Acknowledgments

This research was supported by Grants-in-Aid for Scientific Research (B) (22380186) and Scientific Research on Innovative Areas (22114507) from the Ministry of Education, Culture, Sports, Science, and Technology, Japan, to TM. KK was a holder of a Research Fellowship of the Japan Society for the Promotion of Science.

References

1. Takahashi N, Hotta T, Ishihara H et al (1986) Xylose-containing common structural unit in N-linked oligosaccharides of laccase from sycamore cells. *Biochemistry* 25:388–395
2. Terashima M, Kubo A, Suzawa M et al (1994) The roles of the N-linked carbohydrate chain of rice α -amylase in thermostability and enzyme kinetics. *Eur J Biochem* 226:249–254
3. Hayashi M, Tsuru A, Mitsui T et al (1990) Structure and biosynthesis of the xylose-containing carbohydrate moiety of rice α -amylase. *Eur J Biochem* 191:287–295
4. Bardor M, Faveeuw C, Fitchette A-C et al (2003) Immunoreactivity in mammals of two typical plant glyco-epitopes, core α (1,3)-fucose and core xylose. *Glycobiology* 13:427–434
5. Lerouge P, Cabanes-Macheteau M, Rayon C et al (1998) *N*-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol Biol* 38:31–48
6. Hashimoto K, Goto S, Kawano S et al (2006) KEGG as a glycome informatics resource. *Glycobiology* 16:63R–70R
7. Cooper CA, Gasteiger E, Packer NH (2001) GlycoMod—a software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics* 1:340–349
8. Cooper CA, Harrison MJ, Wilkins MR, Packer NH (2001) GlycoSuiteDB: a new curated relational database of glycoprotein glycan structures and their biological sources. *Nucleic Acids Res* 29:332–335
9. Saint-Jore-Dupas C, Nebenführ A, Boulaflous A et al (2006) Plant *N*-glycan processing enzymes employ different targeting mechanisms for their spatial arrangement along the secretory pathway. *Plant Cell* 18:3182–3200
10. Kajiura H, Koiwa H, Nakazawa Y et al (2010) Two *Arabidopsis thaliana* Golgi α -mannosidase I enzymes are responsible for plant *N*-glycan maturation. *Glycobiology* 20:235–247
11. Asatsuma S, Sawada C, Itoh K et al (2005) Involvement of α -amylase I-1 in starch degradation in rice chloroplasts. *Plant Cell Physiol* 46:858–869
12. Villarejo A, Burén S, Larsson S et al (2005) Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nat Cell Biol* 7:1224–1231
13. Nanjo Y, Oka H, Ikarashi N et al (2006) Rice plastidial *N*-glycosylated nucleotide pyrophosphatase/phosphodiesterase is transported from the ER-Golgi to the chloroplast through the secretory pathway. *Plant Cell* 18:2582–2592
14. Kitajima A, Asatsuma S, Okada H et al (2009) The rice α -amylase glycoprotein is targeted from the Golgi apparatus through the secretory pathway to the plastids. *Plant Cell* 21:2844–2858
15. Burén S, Ortega-Villasante C, Blanco-Rivero A et al (2011) Importance of post-translational modifications for functionality of a chloroplast-localized carbonic anhydrase (CAH1) in *Arabidopsis thaliana*. *PLoS One* 6:1–15
16. Nishimura S-I, Niikura K, Kuroguchi M et al (2005) High-throughput protein glycomics: Combined use of chemoselective glycol blotting and MALDI-TOF/TOF mass spectrometry. *Angew Chem Int Ed* 44:91–96
17. Uematsu R, Frukawa J, Nakagawa H et al (2005) High throughput quantitative glycomics and glycoform-focused proteomics of murine dermis and epidermis. *Mol Cell Proteomics* 4:1977–1989
18. Furukawa J, Shinohara Y, Kuramoto H et al (2008) Comprehensive approach to structural and functional glycomics based on chemoselective glycol blotting and sequential tag conversion. *Anal Chem* 80:1094–1101
19. Amano M, Yamaguchi M, Takegawa Y et al (2010) Threshold in stage-specific embryonic glycotypes uncovered by a full portrait of dynamic *N*-glycan expression during cell differentiation. *Mol Cell Proteomics* 9:523–537
20. Abe M, Shimaoka H, Fukushima M, Nishimura S-I (2012) A cross-linked polymer possessing a high density of hydrazide groups: high-throughput glycan purification and labeling for high-performance liquid chromatography analysis. *Polym J* 44:267–277

Chapter 45

Assay for Proteasome-Dependent Protein Degradation and Ubiquitinated Proteins

Takeo Sato, Kaori Sako, and Junji Yamaguchi

Abstract

The ubiquitin-26S proteasome system (UPS) plays a crucial role in selective removal of short-lived target proteins, archiving fine-tuning of post-translation levels of the target proteins. Recently a number of ubiquitin ligases (E3) have been reported as essential regulators of various plant developmental cues and stress responses. To clarify the detailed biochemical and physiological function of the E3 proteins, identification of their target proteins is of great importance. A transient expression system with tobacco leaves is a powerful method to evaluate E3 function and target degradation via UPS. Here simple methods to assay proteasome-dependent protein degradation combined with a tobacco transient expression system and detection of accumulation of ubiquitinated proteins are presented.

Key words Ubiquitin, Ubiquitin ligase, 26S proteasome, Tobacco, *Agrobacterium*, Transient expression

1 Introduction

The ubiquitin-26S proteasome system (UPS) controls multiple phenomena in plant growth and development by regulating the stability of specific key target proteins that govern specific cellular events including photomorphogenesis, cell cycle, senescence, defense response, and phytohormone response [1, 2]. The ubiquitin molecule attaches to the target protein via E1-E3 enzymes and the poly-ubiquitinated substrate is then degraded by a multi-subunit protease complex, the proteasome. Ubiquitin ligase (E3) is the key enzyme to specify the target protein for degradation via UPS (Fig. 1).

The *Arabidopsis* genome contains more than 1,200 genes encoding E3 [3]. While identification of target proteins is essential to reveal the details of the function of E3, most targets remain unclear. Recently, improved proteomics approaches have simplified the task of finding proteins that interact with E3, namely, the candidate target proteins. Thus, it is of great importance to evaluate

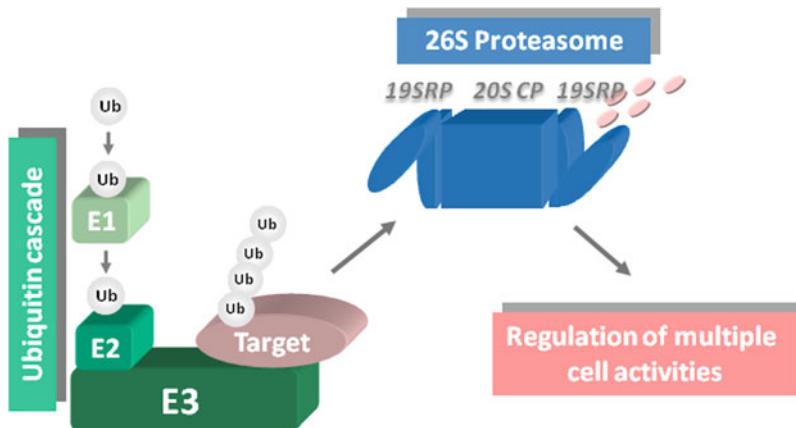


Fig. 1 Schematic model of protein degradation via ubiquitin-proteasome system. E1; Ubiquitin activating enzyme, E2; Ubiquitin conjugating enzyme, E3; Ubiquitin ligase. The ubiquitin (Ub) molecule attaches to the target protein via E1–E3 enzymes, termed the “ubiquitin cascade.” 26S proteasome recognizes the poly-ubiquitin chain attached to the target protein, which is caught and unfolded by 19S regulatory particle (19S RP) and degraded by 20S core particle (20S CP). The ubiquitin-26S proteasome system controls cell activities which are involved in multiple phenomena in plant growth and development

whether a candidate protein is an actual target or not. There are several points to consider in such an evaluation; direct ubiquitination of the protein by specific E3, proteasome-dependent degradation and accumulation in mutants that are deficient in specific E3 functions. In vitro ubiquitination analysis is the most general way to test the direct ubiquitination activity of E3 for the target [4–7], although the success of this approach depends on assay conditions after nature of the E3. Furthermore, the ubiquitination may not be reflected physiologically. On the other hand, it takes a long time to prepare a specific antibody or establish transgenic plants expressing epitope-tagged target protein to investigate *in vivo* degradation of the target protein. Transient expression in tobacco leaves is a powerful method with which to expeditiously prepare and analyze degradation of a target protein via UPS [8, 9]. Recently, the combination of this transient expression and *in vitro* treatment with MG132, a proteasome inhibitor, has been commonly used as a first step to test for UPS-dependent degradation. In this chapter, the use of these procedures to identify 14-3-3 as a target of plant ubiquitin ligase ATL31 is reported [7, 10].

In addition, *in vivo* treatment with MG132 is a convenient way to evaluate the degradation of the target by proteasome and analyze the function of each proteasome subunit under various stress conditions [11, 12]. Here in detail a proteasome-dependent degradation assay with a tobacco transient expression system is described, as well as accumulation of ubiquitinated protein with *in vivo* MG132 treatment.

2 Materials

2.1 Proteasome-Dependent Degradation Assay with Tobacco Transient Expression System

2.1.1 Tobacco and Agrobacterium

2.1.2 Binary Vector

2.1.3 Culture Medium and Infiltration Buffer

2.1.4 Protein Extraction and MG132 Treatment

2.1.5 SDS-PAGE and Western Blotting

1. *Nicotiana benthamiana*.
2. *Rhizobium radiobacter* (GV3101 pMP90).

p35S:ATL31-FLAG (pGWB11), p35S:Myc-14-3-3 \times (pEarley-Gate203), p35S:GFP (pMDC43), p35S-p19 (pBIC) (see Note 1).

1. LB: 0.5 % yeast extract, 1 % Polypeptone, and 1 % NaCl.
2. Antibiotics: Kanamycin.
3. Infiltration buffer: 10 mM MgCl₂ and 10 mM MES-NaOH (pH 5.6).
4. 150 mM Acetosyringone (3,5'-dimethoxy-4'-hydroxy-acetophenone) dissolved in dimethyl sulfoxide (DMSO) (see Note 2).

1. Protein extraction buffer: 50 mM Tris-HCl, pH 7.5, 2 mM ATP, 5 mM MgCl₂, 10 mM 2-mercaptoethanol, and 20 % glycerol (see Note 3).
2. 10 mM MG132 dissolved in DMSO (see Note 4).

1. 2 \times SDS sample buffer: 4 % SDS, 10 % 2-mercaptoethanol, 125 mM Tris-HCl, pH 6.8, 20 % glycerol and 0.002 % bromophenol blue (BPB).
2. 10 % SDS-polyacrylamide gel.
3. SDS electrophoresis running buffer: 50 mM Tris-HCl, pH 8.9, 384 mM glycine, and 0.1 % SDS.
4. Transfer buffer: 39 mM glycine, 48 mM Tris, 0.0375 % SDS, and 20 % methanol.
5. PBS-T buffer: 80 mM Na₂HPO₄, 20 mM NaH₂PO₄, 100 mM NaCl, and 0.1 % Tween-20.
6. Blocking buffer: PBS-T with 5 % skimmed milk.
7. Antibody: anti-Myc antibody from mouse (PL14, MBL, Nagoya, Japan), anti-GFP antibody from mouse (JL-8, Clontech, California, USA), peroxidase-labeled anti-mouse IgG antibody (GE Healthcare, Little Chalfont, UK).
8. Detection solution: Immobilon Western Chemiluminescent HRP Substrate (Millipore, Massachusetts, USA).

2.2 *In Vivo MG132 Treatment and Detection of Ubiquitinated Protein*

2.2.1 Plant Material

2.2.2 Growth Medium

1. MS medium: 4.3 g/L Murashige and Skoog (MS) basal salt mixture, vitamin mixture (nicotinic acid 0.5 mg/L, pyridoxine HCl 0.5 mg/L, thiamine HCl 0.1 mg/L, glycine 2.0 mg/L, myo-inositol 100 mg/L), 2 % sucrose, and 0.25 % gellan gum, adjust pH 5.7 with KOH.
2. 1/2 MS liquid medium: 2.15 g/L Murashige and Skoog basal salt mixture, vitamin mixture, 1 % sucrose, adjust pH 5.7 with KOH.
3. MG132 stock solution: 10 mM MG132 in DMSO.

2.2.3 SDS-PAGE and Western Blotting

1. 2× SDS sample buffer: 4 % SDS, 10 % 2-mercaptoethanol, 125 mM Tris-HCl, pH 6.8, 20 % glycerol and 0.002 % bromophenol blue (BPB).
2. 10 % SDS-polyacrylamide gel.
3. SDS electrophoresis running buffer: 50 mM Tris-HCl, pH 8.9, 384 mM glycine and 0.1 % SDS.
4. Transfer buffer: 39 mM glycine, 48 mM Tris, 0.0375 % SDS and 20 % methanol.
5. PBS-T buffer: 80 mM Na₂HPO₄, 20 mM NaH₂PO₄, 100 mM NaCl and 0.1 % Tween-20.
6. Blocking buffer: PBS-T with 5 % skimmed milk.
7. Antibody: anti-multiubiquitin chains antibody (FK2) (Nippon BioTest laboratories Inc., Tokyo, Japan), peroxidase-labeled anti-mouse IgG antibody (GE Healthcare, Little Chalfont, UK).
8. Detection solution: Immobilon Western Chemiluminescent HRP Substrate (Millipore, Massachusetts, USA).

3 Methods

3.1 Proteasome-Dependent Degradation Assay with Tobacco Transient Expression System

3.1.1 Agrobacterium Culture and Infiltration

1. Grow Agrobacterium transformed with each binary vector in LB medium supplemented with 50 µg/mL Kanamycin up to 1.0 OD₆₀₀ and harvest 1 mL Agrobacterium medium into 1.5 mL tube (see Note 5).
2. Centrifuge the Agrobacterium at 2,000 × g for 3 min in room temperature and remove the supernatant.
3. Resuspend and wash with 1 mL infiltration buffer and centrifuge again under the same conditions.

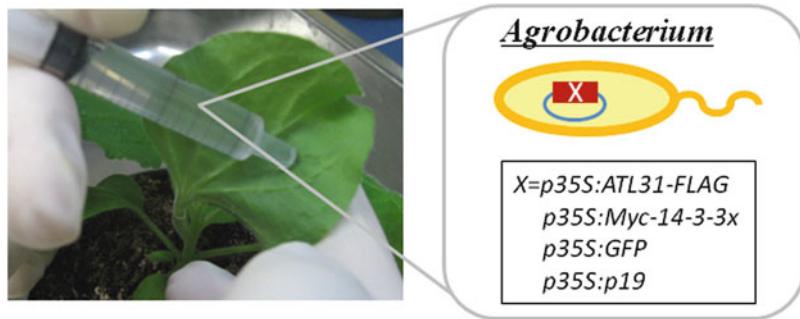


Fig. 2 Agro-infiltration into tobacco leaf. Agrobacterium suspension with 1 mL needle-less syringe is injected into abaxial side of the tobacco leaf. The Agrobacterium contains vectors X having genes coding FLAG-tagged ubiquitin ligase (ATL31-FLAG), Myc-tagged target protein (Myc-14-3-3 χ), GFP as control, and p19 protein. All the constructed genes are constitutively expressed under the *CaMV 35S* promoter

4. Remove the supernatant and resuspend the pellet with 1 mL infiltration buffer (see Note 5) supplemented with 1 μ L Acetosyringone (Final 150 μ M).
 5. Mix the Agrobacterium suspension each with the equal volume.
 6. Inject the Agrobacterium suspension (around 500 μ L/leaf) into the tobacco leaf with a needle-less syringe and incubate for 2–4 days (see Fig. 2, Notes 6 and 7).
- 3.1.2 Protein Extraction and MG132 Treatment**
1. Excise the infiltrated tobacco leaf and grind it with liquid nitrogen.
 2. Add the 500 μ L protein extraction buffer and transfer to the 1.5 mL tube on ice.
 3. Mix and split the extraction into two different tubes and add the MG132 (final 30 μ M) to one tube and DMSO to the other. Incubate the samples at room temperature for 2 h.
 4. Quantify the total amount of protein in the extraction before proceeding further.
 5. Stop the degradation reaction by adding an equal volume of 2 \times SDS sample buffer.
- 3.1.3 SDS-PAGE and Western Blotting**
1. Heat the sample for 5 min at 75 °C and centrifuge by 20,000 \times g for 5 min.
 2. Collect the supernatant and for use with SDS-PAGE (see Note 8).
 3. Load the 3 μ g protein and start the electrophoresis.
 4. Stop the electrophoresis and transfer the protein onto the PVDF membrane.

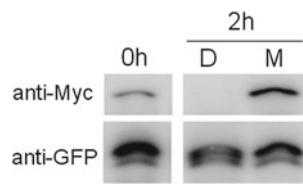


Fig. 3 Inhibition of proteasome-dependent protein degradation with MG132 treatment. The crude extract (0 h) from tobacco leaf was incubated with DMSO (D) or MG132 (M) for 2 h and used for western blotting analysis with anti-Myc and anti-GFP antibody. Overexpression of ubiquitin ligase ATL31 promoted proteasome-dependent degradation of the specific target protein Myc-14-3-3 χ with DMSO treatment for 2 h, whereas MG132 treatment inhibited the degradation of the target. GFP was used as a control for normalization of amount of the expressed proteins

5. After blocking treatment by incubation with blocking buffer, incubate the membrane with 1/5,000-diluted anti-Myc or anti-GFP antibody with PBS-T for 1 h in room temperature with shaking.
6. Remove the antibody and wash the membrane with PBS-T buffer.
7. Incubate the membrane with 1/25,000 HRP-labeled anti-mouse IgG antibody for 1 h in room temperature with shaking.
8. Remove the antibody and wash with PBS-T buffer.
9. Add the detection solution onto the membrane and incubate for 5 min.
10. Detect and quantify the signal with luminescent image analyzer LAS3000 (Fujifilm, Tokyo, Japan) (see Fig. 3 and Note 9).

3.2 In Vivo MG132 Treatment and Detection of Ubiquitinated Protein

3.2.1 Preparing Plants

3.2.2 MG132 Treatment for Plants

1. Sterilize *Arabidopsis thaliana* seeds and sow seeds on petri dish with MS medium.
2. Germinate the seedlings and let grow for 10 days (see Note 10).
1. Add MG132 to 1/2 MS liquid medium to a final concentration of 50 μ M. Also, add DMSO of the same volume of the MG132 to 1/2 MS liquid medium as a negative control.
2. Pour 3 mL, respectively, of 1/2 MS liquid medium containing MG132 or DMSO into a 35 mm petri dish.
3. Float three plants on 1/2 MS liquid medium in each dish. Incubate dishes under normal growth conditions for 15 h (Fig. 4).
4. Wipe attached medium from plants and weigh them.

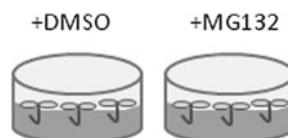


Fig. 4 Diagram of floating method for growth of three plants on liquid medium in each petri dish

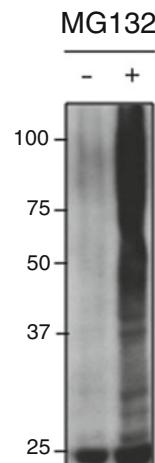


Fig. 5 Detection of polyubiquitinated proteins in *Arabidopsis* Col-0 plants by Western blot. In vivo treatment with MG132 (+) inhibited proteasome activity in plants and thus polyubiquitinated proteins accumulated in a higher molecular range, as detected by anti-ubiquitin antibody

5. Grind plants to a fine powder using liquid nitrogen to extract proteins.
6. Add the powder to 2× SDS sample buffer that has six times the volume of the fresh weight of the sample in a 1.5 mL tube.
7. Centrifuge at 20,000×*g* for 5 min at 4 °C. Transfer the supernatant to a new tube.
8. Perform heat treatment for 3 min at 99 °C.

3.2.3 Detection of Polyubiquitinated Proteins

1. Prepare 10 % SDS-PAGE gel and apply 5 µL samples onto the gel (see Note 11).
2. Separate proteins by SDS-PAGE and carry out western blotting (see Subheading 3.1.3).
3. Incubate the membrane with 1/5,000 anti-FK2 as first antibody and 1/25,000 peroxidase-labeled anti-mouse antibody as second antibody as described above (see Subheading 3.1.3).
4. Detect the chemiluminescent signals for accumulation of polyubiquitinated proteins as described above (see Subheading 3.1.3 and Fig. 5).

4 Notes

1. The p19 protein suppresses silencing of plant genes that function against bacterial infection, thereby promoting sufficient expression of the transformed genes. ATL31 is an *Arabidopsis* ubiquitin ligase and 14-3-3 is its target protein. GFP is used for infection and expression control as it is extremely stable and is not degraded significantly by 26S Proteasome.
2. Acetosyringone is a phenolic compound produced by plants, especially in wound tissue, which activates the *Agrobacterium* infection and increases the transformation efficiency.
3. This buffer allows efficient degradation of poly-ubiquitinated protein by 26S Proteasome. ATP and MgCl₂ are required for the ATPase activity of the RPT subunit in 19S RP.
4. As an alternative to MG132, clasto-Lactacystin β -lactone or Epoxomicin can be used as the proteasome inhibitor.
5. The number and suspended volume of the *Agrobacterium* used for the infiltration assay should be adjusted to increase the efficiency of each protein expression.
6. The incubation time is also adjusted to increase the efficiency of detection of the transformed protein under experimental conditions. Liu et al evaluated the effect of incubation length and expression efficiency [8]. In addition, if the transformed protein has specific functions in terms of plant defense response, the fact that the *Agrobacterium* treatment stimulates plant immune responses should be taken into consideration.
7. Tsuda et al established the efficient transient expression system in *Arabidopsis thaliana* with an effector gene *ArrPto* [13].
8. The protein amount in the extracted sample before reaction (0 h) is quantified and 3 μ g proteins for the SDS-PAGE/WB analysis are used. Use the same volume of digested sample (2 h) as the sample (0 h).
9. GFP intensity with anti-GFP antibody can be used for normalization as a non-degraded protein. In addition, point-mutated ATL31 (ATL31C143S), which abolishes ubiquitin ligase activity is used as a more accurate negative control. Since tobacco plants may have a ubiquitin ligase that is homologous to the transiently expressed protein, the expressed ubiquitin ligase activity can be evaluated against the target protein compared with that of the mutated control, which is deficient in this activity [7].
10. An accumulation of ubiquitinated proteins can be also compared when plants are grown under various growth conditions; for example, under metal deficient conditions [14].

11. Since polyubiquitinated proteins have higher molecular weight, lower concentration polyacrylamide gels (7.5–10 %) are appropriate for the efficient separation of proteins.

Acknowledgments

This work was supported by Grants-in-Aid for Scientific Research (23380198, 24114701) to JY, (24770035) to TS; by Research Fellowships from the Japan Society for the Promotion of Science for Young Scientists (2012–2014) to K.S and also by Cooperative Research Grant of the Gene Research Center, the University of Tsukuba.

References

1. Hershko A, Ciechanover A (1998) The ubiquitin system. *Annu Rev Biochem* 67:425–479
2. Smalle J, Vierstra RD (2004) The ubiquitin 26S proteasome proteolytic pathway. *Annu Rev Plant Biol* 55:555–590
3. Stone SL, Callis J (2007) Ubiquitin ligases mediate growth and development by promoting protein death. *Curr Opin Plant Biol* 10:624–632
4. Qin F, Sakuma Y, Tran LSP et al (2008) Arabidopsis DREB2A-interacting proteins function as RING E3 ligases and negatively regulate plant drought stress-responsive gene expression. *Plant Cell* 20:1693–1707
5. Lee HK, Cho SK, Son O et al (2009) Drought stress-induced Rma1H1, a RING membrane-anchor E3 ubiquitin ligase homolog, regulates aquaporin levels via ubiquitination in transgenic Arabidopsis plants. *Plant Cell* 21:622–641
6. Zhang XR, Garretton V, Chua NH (2005) The AIP2 E3 ligase acts as a novel negative regulator of ABA signaling by promoting ABI3 degradation. *Genes Dev* 19:1532–1543
7. Sato T, Maekawa S, Yasuda S et al (2011) Identification of 14-3-3 proteins as a target of ATL31 ubiquitin ligase, a regulator of the C/N response in Arabidopsis. *Plant J* 60:137–146
8. Liu LJ, Zhang YY, Tang SY et al (2010) An efficient system to detect protein ubiquitination by agroinfiltration in *Nicotiana benthamiana*. *Plant J* 61:893–903
9. Igawa T, Fujiwara M, Takahashi H et al (2009) Isolation and identification of ubiquitin-related proteins from *Arabidopsis* seedlings. *J Exp Bot* 60:3067–3073
10. Sato T, Maekawa S, Yasuda S et al (2009) CNI1/ATL31, a RING-type ubiquitin ligase that functions in the carbon/nitrogen response for growth phase transition in *Arabidopsis* seedlings. *Plant J* 60:852–864
11. Sonoda Y, Sako K, Maki Y et al (2009) Regulation of leaf organ size by the *Arabidopsis* RPT2a 19S proteasome subunit. *Plant J* 60:68–78
12. Book AJ, Smalle J, Lee KH et al (2009) The RPN5 subunit of the 26s proteasome is essential for gametogenesis, sporophyte development, and complex assembly in *arabidopsis*. *Plant Cell* 21:460–478
13. Tsuda K, Qi YP, Nguyen LV et al (2012) An efficient Agrobacterium-mediated transient transformation of *Arabidopsis*. *Plant J* 69:713–719
14. Sakamoto T, Kamiya T, Sako K et al (2011) *Arabidopsis thaliana* 26S proteasome subunits RPT2a and RPT5a are crucial for zinc deficiency-tolerance. *Biosci Biotech Biochem* 75:561–567

Part VIII

Interactomics and Proteinomics

Chapter 46

Separation of Membrane Protein Complexes by Native LDS-PAGE

Janine Arnold, Alexey Shapiguzov, Geoffrey Fucile, Jean-David Rochaix, Michel Goldschmidt-Clermont, and Lutz Andreas Eichacker

Abstract

Gel electrophoresis has become one of the most important methods for the analysis of proteins and protein complexes in a molecular weight range of 1–10⁷ kDa. The separation of membrane protein complexes remained challenging to standardize until the demonstration of Blue Native PAGE in 1991 [1] and Clear Native PAGE in 1994 [2]. We present a robust protocol for high-resolution separation of photosynthetic complexes from *Arabidopsis thaliana* using lithium dodecyl sulfate as anion in a modified Blue Native PAGE (LDS-PAGE). Here, non-covalently bound chlorophyll is used as a sensitive probe to characterize the assembly/biogenesis of the pigment–protein complexes essential for photosynthesis. The high fluorescence yield recorded from chlorophyll-binding protein complexes can also be used to establish the separation of native protein complexes as an electrophoretic standard.

Key words Native LDS-PAGE, Thylakoid membrane, Chlorophyll-binding protein complexes, Fluorescence

Abbreviations

| | |
|-------|---|
| APS | Ammonium persulfate |
| DDM | <i>n</i> -Dodecyl- β -D-maltoside |
| DIG | Digitonin |
| LDS | Lithium dodecyl sulfate |
| PAGE | Polyacrylamide electrophoresis |
| TEMED | Tetramethylethylenediamine |

1 Introduction

Assemblies of proteins constitute molecular machines that make up the functional proteome of a cell [3]. These assemblies can be viewed as a highly dynamic network of protein interactions that reflect the cell's physiological state. It is therefore no wonder that

changes in the assemblies are caused by disease or mutation of structural and regulatory components of a cell [4]. Polyacrylamide gel electrophoresis (PAGE) has become one of the most successful techniques for the separation of proteins [5]. In an electric field, particles like peptides, proteins, and protein display different mobility. These differences can be amplified and stabilized as separation distance by sieving and anti-convective interactions within a porous gel [6]. The technique has been shown to cover the molecular weight range of 1–10⁷ kDa using water soluble as well as membrane integral proteins and protein complexes [1, 7]. Especially, Blue Native (BN) PAGE and Clear Native (CN) PAGE have been demonstrated to provide high resolution for separation of protein complexes from a broad variety of membranes [1, 2, 8–10]. Here, we present a robust variation of BN-PAGE gel electrophoresis for separation of protein complexes. We replace Coomassie against lithium dodecyl sulfate (LDS) and demonstrate white light and fluorescence detection of chlorophyll (Chl)-containing protein complexes at high resolution in a molecular weight range of 2 × 10⁴ to 1.5 × 10⁶ Da using detergent solubilized thylakoid membranes from *Arabidopsis thaliana*.

2 Materials

2.1 Preparation of Thylakoid Membranes

1. Rosette leaves from *Arabidopsis thaliana* grown on soil for 3–4 weeks under cycles of 8 h light (100 μmol/m²/s) and 16 h dark.
2. Extraction buffer: 25 mM Tricine, pH 7.8, 330 mM sorbitol, 1 mM Na-EDTA, 10 mM KCl, 0.15 % (w/v) BSA, 4 mM sodium ascorbate.
3. Lysis buffer: 10 mM Tricine, pH 7.8, 5 mM MgCl₂, 10 mM NaF.
4. Washing Buffer: 25 mM Tricine, pH 7.8, 100 mM sorbitol, 5 mM MgCl₂, 10 mM KCl, 10 mM NaF.
5. Storage buffer: 10 % (v/v) Glycerol, 25 mM Tricine–NaOH, pH 7.8, 100 mM Sorbitol, 5 mM MgCl₂, 10 mM KCl.

All buffers should be freshly prepared and stored at 4 °C.

2.2 Solubilization of Membranes

1. DDM solution: 195.84 mM (10 % w/v) *n*-dodecyl-β-D-maltoside. Store at -20 °C.
2. DIG solution: 81.34 mM (10 % w/v) digitonin, (see Note 1). Store at -20 °C.
3. LDS solution: 183.6 mM (5 % w/v) lithium dodecyl sulfate. Store at -20 °C.
4. Detergent mixture 1: 20 mM Digitonin, 0.073 mM LDS, 20 % (v/v) Glycerol, 0.01 % (w/v) Ponceau S, 15 mM Tricine, 80 mM Bis-Tris (see Note 1).

5. Detergent mixture 2: 9 mM Digitonin, 9 mM *n*-dodecyl- β -D-maltoside, 0.073 mM LDS, 20 % (v/v) Glycerol, 0.01 % (w/v) Ponceau S, 15 mM Tricine, 80 mM Bis-Tris (see Note 1).

2.3 Clear Native LDS-PAGE from *Arabidopsis thaliana* Thylakoid Membranes

1. Thylakoid membranes from *Arabidopsis thaliana*: 1 μ g Chl/ μ l frozen at liquid N₂. Stored at -80 °C (preparation, see Subheading 3.1).
2. LDS-cathode buffer: 80 mM Tricine, 15 mM Bis-Tris, 0.073 mM LDS. Assemble as 10 \times concentrate. Store at 4 °C (see Notes 2 and 3).
3. Anode buffer: 50 mM Bis-Tris/HCl, pH 7.0. Assemble as 10 \times concentrate. Store at 4 °C (see Note 3).
4. Gel buffer: 0.5 M ϵ -amino caproic acid, 50 mM Bis-Tris, titrate with HCl to pH 7.0. Assemble as 6 \times concentrate. Store at 4 °C (see Note 3).
5. Electrophoresis chamber: Hoefer SE400 (vertical slab gel electrophoresis unit). Gel size: 180 mm \times 160 mm.
6. Acrylamide solution: 30 % (w/v) acrylamide/bis acrylamide solution (37.5:1, 2.6 % C), acts as a neurotoxin in unpolymerized state. Store at 4 °C.
7. Glycerol, 50 % (v/v). Store at 4 °C.
8. TEMED: *N,N,N,N'*-tetramethyl-ethylenediamine. Store at 4 °C.
9. APS: 10 % (w/v) ammonium persulfate. Store at -20 °C. Stable at 4 °C for up to 2 weeks.
10. Water-saturated isobutanol: 50 % (v/v) isobutanol. Store at room temperature.
11. Peristaltic tube pump for low viscosity liquid delivery.
12. Power supply like EPS 3501 (GE Healthcare, Buckinghamshire, UK). With ranges of voltage 35–3,500 V, current 1–400 mA, power 1–200 W, time 0:01–500 h, volt hours 1–500,000 Vh, milliampere-hours 1–25,000 mAh.

2.4 Analysis of Thylakoid Membrane Complexes

1. Molecular weight standard: Invitrogen NativeMark, unstained protein standard, IgM Hexamer, IgM Pentamer, Apoferritin band 1, Apoferritin band 2, β -phycoerythrin, Lactate Dehydrogenase, Bovine Serum Albumin, Soybean Trypsin Inhibitor, with molar masses of 1.236 kDa, 1.048 kDa, 720 kDa, 480 kDa, 242 kDa, 146 kDa, 66 kDa, 20 kDa, respectively.
2. White light scanning: Apple Color OneScanner 1200/30.
3. Fluorescence scanning: Li-COR Odyssey, Imaging system (Li-COR Biosciences). Excitation wavelength 680 nm, Emission wavelength 694 nm.
4. Colloidal Coomassie stain, 0.25 % (w/v) Coomassie G250, 50 % (v/v) Methanol, 10 % (v/v) Acetic acid.

3 Methods

This protocol describes how protein assemblies can be separated by a native PAGE approach. The method uses the anionic detergent LDS in the cathode buffer lending the name native LDS-PAGE. This detergent was selected because of its high solubility at 4 °C. In general, native PAGE is intended to separate protein complexes in a state that reflects the physiologically relevant functional state of the protein assembly. It is therefore advisable to minimize protein complex degradation through sample preparation on ice and electrophoresis at 4 °C. Assemblies of chlorophyll-binding thylakoid membrane proteins were found to remain intact in the native LDS-PAGE approach and were easily detected by white light and fluorescence scanning after electrophoresis (Fig. 1). Separation of protein complexes by native PAGE is based on a sieving of the complexes by the acrylamide concentration-dependent pore size of the gel. The method presented here corresponds to a 3.5–12 % (v/v) linear gradient separating gel and a 3 % (v/v) stacking gel (Fig. 1). This concentration facilitates a separation of molar mass standards in the range of 20–1,500 kDa (Fig. 1). Experiments were carried out in a Hoefer SE400 electrophoresis chamber (vertical slab unit).

3.1 *Arabidopsis thaliana* Sample Preparation

Arabidopsis thaliana plants are grown on soil for 3–4 weeks in a growth chamber with ambient white light of 100 μmol/m/s at a light–dark cycle of 8 h-light–16 h-dark.

1. Homogenize leaves 4× 4 s in 150 mL of ice-cold extraction buffer.
2. Filter homogenized leaves through one layer of Miracloth and collect through a funnel in 3× 50 mL conical tubes.
3. Centrifuge for 3 min at 1,800×*g*. Collect the pellet and discard the supernatant.
4. Gently resuspend the pellet in 5 mL extraction buffer using a paintbrush. Wash the brush in 25 mL extraction buffer. Redistribute into two conical tubes and fill to 50 mL with extraction buffer.
5. Centrifuge for 3 min at 1,800×*g*. Collect the pellet and discard the supernatant.
6. Gently resuspend each pellet in 5 mL of lysis buffer with a paintbrush.
7. Combine resuspended material in one conical tube and fill to 50 mL with lysis buffer.
8. Divide sample into two Sorvall HB4 tubes and fill to 25 mL with lysis buffer.
9. Incubate sample 5 min in the dark on ice.

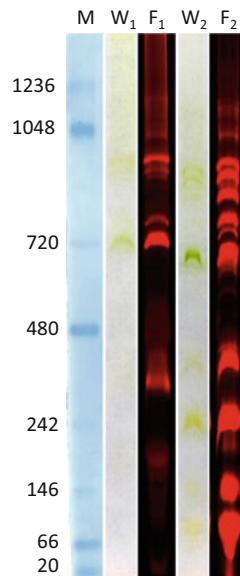


Fig. 1 Separation of thylakoid membrane protein complexes by native LDS-PAGE. Thylakoid membranes corresponding to 8 μ g Chl were solubilized in 20 μ L detergent mixture 1 containing 20 mM Digitonin, 0.073 mM LDS, 20 % (v/v) Glycerol, 0.01 % (w/v) Ponceau S, 15 mM Tricine, 80 mM Bis-Tris (W_1 , F_1) and detergent mixture 2, 9 mM Digitonin, 9 mM *n*-dodecyl- β -D-maltoside, 0.073 mM LDS, 20 % (v/v) Glycerol, 0.01 % (w/v) Ponceau S, 15 mM Tricine, 80 mM Bis-Tris (W_2 , F_2). Non-solubilized material was separated by centrifugation (14,000 $\times g$, 10 min, 4 °C) and discarded. Supernatant containing the solubilized protein complexes was directly loaded onto native LDS-PAGE. Chl-protein complexes were detected by white light scanning (W_1 , W_2) and fluorescence excitation scanning (F_1 , F_2) (Ex 680 nm and Em >694 nm). Standard proteins (M, kD) were stained with colloidal Coomassie G250

10. Centrifuge 5 min at 6,000 $\times g$.
11. Supernatant = chloroplast stroma. Purify stroma from residual thylakoids by centrifugation at 12,000 $\times g$ for 30 min.
12. Pellet = chloroplast thylakoid membrane.
13. Gently resuspend thylakoid membrane pellets from both tubes in 5 mL of wash buffer with a paintbrush and collect in one tube.
14. Collect thylakoids by centrifugation for 5 min at 5,900 $\times g$.
15. Discard the supernatant and keep the thylakoid pellet. Resuspend thylakoids in 3 mL wash buffer using paintbrush.
16. Transfer into glass homogenizer and lever piston three times.
17. Transfer into 5 mL plastic tube
18. Clean homogenizer with 2 mL of wash buffer and homogenize 2 \times and collect in 5 mL tube.

19. For rapid determination of concentration, mix 10 μ L thylakoid membrane extract into 990 μ L 80 % Acetone 4 °C (Dilution, 1/100). Determine absorbance at 652.7 nm ($\epsilon=41$). Calculate total concentration of Chl (μ g/ μ L) as Extinction $_{652.7} \times 2,439$ [11].
20. For storage of thylakoid membranes determine Chl concentration and dilute in storage buffer to a concentration of 1 μ g Chl/ μ L.
21. Freeze aliquots of 100 μ L in liquid N₂ and store tubes at -80 °C (see Note 4).

3.2 Solubilization of Thylakoid Membranes from *Arabidopsis thaliana* for Native LDS-PAGE

1. Use thylakoid membranes corresponding to 8 μ g Chl per lane of the native LDS-PAGE gel.
2. Concentrate thylakoid membranes by centrifugation at 7,500 $\times g$ and 4 °C for 10 min and discard supernatant.
3. Solubilize in 20 μ L of detergent mixture 1 or mixture 2 for 10 min at 4 °C.
4. Separate non-solubilized material by centrifugation at 14,000 $\times g$ and 4 °C for 10 min.
5. Transfer the supernatant containing the solubilized protein complexes to a new micro tube.
6. Load supernatant of different solubilizations onto native LDS-PAGE (see Note 5).

3.3 Electrophoresis of Protein Complexes from *Arabidopsis thaliana* Thylakoids

1. Prepare cathode-, anode- and gel-buffer for LDS-gels. Store at 4 °C.
2. Prepare acrylamide gradient solutions “heavy” and “light” and “stacking” solution. Use evacuation flasks for assembling the solutions. Store at 4 °C (see Note 6).

3.4 Casting of Gradient Gel

1. Clean glass plates and spacers (>1 mm) with ethanol (100 %), assemble the glass plate sandwich, fix it on the casting stand and adjust the assembled casting stand by a spirit level.
2. Use gradient mixer accepting 20–25 mL (see Note 7).
3. Place magnetic stirring rods in both chambers of the gradient mixer and ensure that all ports are closed (see Note 8).
4. Equip gradient mixer outlet with polyethylene tubing of 40 cm length.
5. Get tube pump and assemble tubing connection.
6. Attach a yellow tip at the end of the tubing and tape tip to the center between the glass plates to enable filling of the gradient solutions into the sandwich from the top.
7. Fill solutions in gradient mixer and start the magnetic stirrer. Fill “heavy” solution into the chamber next to the outlet tube (chamber 1). Fill “light” solution in the chamber connected to

Table 1
Volumes for preparation of a native LDS-PAGE gel (Hoefer SE400)

| Solution | “Light” solution, 3.5 % (v/v) | “Heavy” solution, 12 % (v/v) | Stacking gel, 3 % (v/v) |
|---------------------------|----------------------------------|---------------------------------|----------------------------|
| Acrylamide, 30 % (37.5:1) | 1.95 mL | 6.68 mL | 0.5 mL |
| Gel buffer (6×) | 2.784 mL | 2.784 mL | 0.833 mL |
| Glycerol, 50 % | 1.67 mL | 6.68 mL | 0.5 mL |
| Water nanopure | 10.296 mL | 0.556 mL | 3.167 mL |
| Total volume | 16.70 mL | 16.7 mL | 5 mL |
| | Degas 3 min | | |
| APS (10 %) | 57.34 µL | 47.8 µL | 25 µL |
| TEMED | 11.47 µL | 4.78 µL | 5 µL |

chamber 1 (chamber 2). Ensure that the “heavy” solution in chamber 1 is placed at the center of the magnetic stirrer and stirring bar is free to turn (*see Note 9* and Table 1).

- Set the tube-pump speed to 11.5 mL/min (*see Note 10*).
 - Add catalysts APS and TEMED for initiation of the acrylamide polymerization reaction (*see Table 1*).
 - Start pump.
 - Open connection between chamber 1 and 2.
 - Let the solutions run between the glass plates.
 - Overlay the cast gel with water-saturated isobutanol. Allow gel to polymerize for at least 90 min.
 - Rinse the gradient mixer immediately with distilled water to prevent the polymerization of residual gel solution in the tubes and the gradient mixer tubing.
- 3.5 Casting of Stacking Gel**
- After polymerization of the separating gel, remove isobutanol by washing with distilled water.
 - Dry area above separating gel with Whatman paper. Do not touch separation gel surface.
 - Clean comb with denatured ethanol (100 %).
 - Add APS and TEMED to stacking solution and pipette onto separation gel. Fill sandwich to the top of the glass plate.
 - Insert dried comb in stacking solution. Make sure no bubbles are trapped at front of comb or between teeth and align comb teeth ends parallel to separation gel (*see Note 11*).
 - Let gel polymerize for at least 30 min (*see Note 12*).

3.6 Loading and Running of Native LDS Electrophoresis

1. Label the position of the wells on the outer glass plate with a felt pen and remove the comb carefully from the native gel (*see Note 13*).
2. Assemble glass sandwich within electrophoretic apparatus.
3. Fill in cathode buffer into the upper buffer chamber.
4. Rinse the wells with cathode buffer (*see Note 14*).
5. Underlay the samples into the wells using a microsyringe or disposable pipette tips (*see Note 15*).
6. Fill anode buffer in the lower buffer chamber until electrode is immersed.
7. Connect the tubes of the thermostatic circulator to the buffer chamber. Start the thermostatic circulator. Perform electrophoretic run at 4 °C (*see Note 16*).
8. Complete assembly of electrophoresis unit and connect to power supply.
9. Set the power supply to limit voltage. Use 150 V constant. Set mA and W to maximal values. Run overnight for highest resolution. Apply power to electrophoretic set.
10. Stop electrophoresis when the front composed of Ponceau S/ has reached the bottom of the separating gel (*see Note 17*).
11. Stop the thermostatic cooler and disassemble the buffer chamber assembly.
12. After electrophoresis, use glass plate sandwich without disassembly for detection of binding protein complexes by scanning with white light or fluorescence scanners (Fig. 1 W_{1/2} and F_{1/2}).
13. Scan native gels for fluorescence emission by laser excitation at 680 nm and emission wavelength 694 nm using the Odyssey® infrared imaging system (LI-Cor Biosciences) or use alternative scanner like Typhoon Trio (GE Healthcare) by laser excitation at 633 nm and emission of >670 nm using a BP30 emission filter (*see Note 18*).
14. For further analysis of protein complexes like absorbance or mass spectroscopy measurements open sandwich and cut gel bands or stain proteins before further processing of sample of interest.
15. This robust high-resolution separation protocol for membrane protein complexes is free of Coomassie and can be operated at 0 °C. Gels are fully transparent. Fluorescent labeling before and fluorescent readout after the separation has shown excellent results throughout the separation range (Fig. 1). Kinetic studies during biogenesis of membranes and assembly of the proteins can be combined with enzyme assays, gel-blot identifications and mass spectrometry of proteins directly to identify the native complexes within the gel.

4 Notes

1. Digitonin is dissolved by heating the solution to 100 °C. After heating, digitonin can immediately be mixed with other detergents. The rest of the detergent mixture can be stored at -20 °C. Always reheat solution before usage.
2. Add LDS after the buffer reagents have been weighed (pH ~ 6.8 at room temperature equals pH ~ 7.0 at 4 °C).
3. If larger amounts of buffers are prepared, store at -20 °C.
4. To avoid freeze-thaw cycles of membrane samples, freeze 20 µL droplets directly in liquid N₂. Use an open container immersed in the liquid N₂ to gather the droplets. For storage of the droplets, make a hole in the lid of a number of micro tubes. Place a cardboard stand in liquid N₂ and place the micro tubes to precool. Ensure that no liquid N₂ flows into the tubes. Gather frozen droplets into the precooled micro tubes using forceps. Precool the tweezers before picking the droplets. Store tubes at -80 °C. For experimental access to thylakoid membranes, stored droplets can be picked separately with forceps and transferred to a fresh tube.
5. Depending on the organism and membrane that should be solubilized different detergent concentrations and mixtures need to be tested individually.
6. Use glass evacuation flasks for degassing. Shake solutions to initiate gas bubble release. Pay attention that APS and TEMED are added after degassing.
7. Gradient mixer should be filled at least 50 % of each of the chamber volumes.
8. The chamber next to the tube outlet (chamber 1) has to be stirred properly whereas stirring of chamber 2 is not necessary. The stirring bar in chamber 2 acts as balancer for the solution levels in both chambers.
9. Adjust stirrer speed to maintain solution mixed without aeration. Make sure that no air bubbles block the connection between chambers.
10. Measure pump speed by determination of time required to collect a 10 mL fraction.
11. Ensure that the distance between well and separation gel is at least 0.5 cm.
12. If gels are stored overnight, wrap the gel sandwich in wet tissues and keep combs in place.
13. Depending on the electrophoresis apparatus the gel comb has to be removed before or after assembly with the glass sandwich.

14. Fill 50 mL syringe with cathode buffer and use needle with diameter thinner than spacer thickness to rinse wells.
15. If a microsyringe is used, rinse with anode buffer before applying a new sample. If wells remain unused, load detergent mixture 1 or 2 to the empty well.
16. Depending on laboratory condition and available equipment, the run of the native gel electrophoresis could be done in the cold room. In order to control the running behavior, check temperature and pH of cathode buffer in upper buffer chamber.
17. Gel requires about 18 h to complete separation of the solubilized protein complexes (Hoefer SE400, 100 V constant, Gel length = 160 mm).
18. The use of low fluorescent glass plates is recommended.

References

1. Schagger H, von Jagow G (1991) Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal Biochem* 199:223–231
2. Schagger H, Cramer WA, von Jagow G (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal Biochem* 217:220–230
3. Alberts B (1998) The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* 92:291–294
4. Ryan DP, Matthews JM (2005) Protein-protein interactions in human disease. *Curr Opin Struct Biol* 15:441–446
5. Raymond SW, Weintraub L (1959) Acrylamide gel as a supporting medium for zone electrophoresis. *Science* 130:711–712
6. Maure RHR (1971) Basic principles of polyacrylamide gel electrophoresis and some recent advances of the technique. *Ann Biol Clin (Paris)* 29:205–210
7. Strecker V, Wumiai RZ, Wittig I et al (2010) Large pore gels to separate mega protein complexes larger than 10 MDa by blue native electrophoresis: isolation of putative respiratory strings or patches. *Proteomics* 10: 3379–3387
8. Krause F (2006) Detection and analysis of protein-protein interactions in organellar and prokaryotic proteomes by native gel electrophoresis: membrane protein complexes and supercomplexes. *Electrophoresis* 27:2759–2781
9. Reisinger V, Eichacker LA (2006) Analysis of membrane protein complexes by blue native PAGE. *Proteomics* 6(Suppl 2):6–15
10. Jarvi S, Suorsa M, Paakkarinen V et al (2011) Optimized native gel systems for separation of thylakoid protein complexes: novel super- and mega-complexes. *Biochem J* 439:207–214
11. Inskeep WP, Bloom PR (1985) Extinction coefficients of chlorophyll a and b in n, n-dimethylformamide and 80% acetone. *Plant Physiol* 77:483–485

Chapter 47

Identification of Thioredoxin Target Disulfides Using Isotope-Coded Affinity Tags

Per Hägglund, Jakob Bunkenborg, Kenji Maeda, Christine Finnie, and Birte Svensson

Abstract

Thioredoxins (Trx) are small redox proteins that reduce disulfide bonds in various target proteins and maintain cellular thiol redox control. Here, a thiol-specific labeling and affinity enrichment approach for identification and relative quantification of Trx target disulfides in complex protein extracts is described. The procedure utilizes the isotope-coded affinity tag (ICAT) reagents containing a thiol reactive iodoacetamide group and a biotin affinity tag to target peptides containing reduced cysteine residues. The identification of substrates for Trx and the extent of target disulfide reduction is determined by LC-MS/MS-based quantification of tryptic peptides labeled with “light” (¹²C) and “heavy” (¹³C) ICAT reagents. The methodology can be adapted to monitor the effect of different reductants or oxidants on the redox status of thiol/disulfide proteomes in biological systems.

Key words Thioredoxin, Disulfide, Redox proteomics, Thiol, Cysteine, Isotope-coded affinity tag, Iodoacetamide

1 Introduction

Thioredoxin (Trx) is a widely occurring protein disulfide reductase that plays multiple key roles in cellular metabolism. For example, Trx functions as an electron donor for oxidoreductases such as methionine sulfoxide reductase and ribonucleotide reductase, and as a regulator of enzymatic activities, e.g., in the Calvin cycle [1]. Plant Trx contains a redox-active dithiol motif C(G/P)PC and is recycled between an oxidized and reduced state either by NADPH-dependent thioredoxin reductase (NTR) or, in chloroplasts, by ferredoxin thioredoxin reductase. Proteomics techniques have been used for identification of a large number of Trx target proteins [2–13]. These techniques are in general based on one of two different principles: (1) thiol-specific labeling of target proteins or (2) mechanism-based trapping of target proteins using inactive single cysteine (CXXS) Trx. In many cases, however, the identities of the

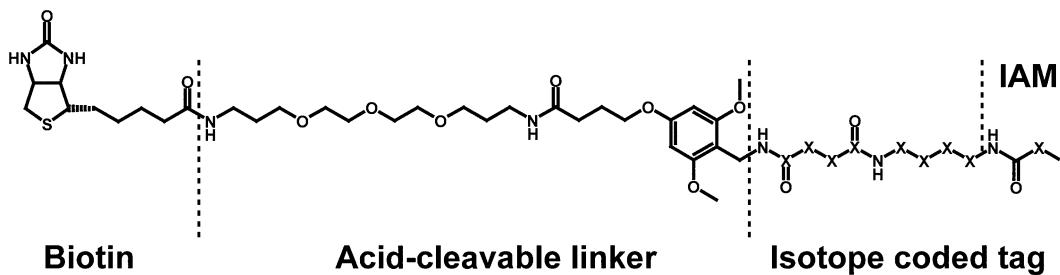


Fig. 1 Structure of acid-cleavable ICAT reagents. The reagent contains four different functionalities: (1) a iodoacetamide group reacting with reduced thiol groups ($-SH$), (2) a tag with nine isotope-coded carbon atoms ($x = {}^{13}C$ (ICAT “heavy”); $x = {}^{12}C$ (ICAT “light”), (3) a biotin tag that serves as an affinity tag for avidin chromatography to enrich labeled peptides and (4) an acid-cleavable linker. Reproduced from ref. 3 with permission from American Chemical Society

specific target disulfide bonds and the extent of their reduction by Trx is unknown. This information is important for an in-depth understanding of the role of Trx in various metabolic pathways. We addressed this issue by developing a quantitative differential thiol labeling procedure using isotope-coded affinity tag (ICAT) reagents [3, 4]. The ICAT reagents contain a thiol-reactive iodoacetamide (IAM) group and isotope coded linkers available in “light” ($ICAT_L$) and “heavy” ($ICAT_H$) forms labeled with ${}^{12}C$ and ${}^{13}C$ stable carbon isotopes, respectively (Fig. 1). The relative abundance of cysteine containing tryptic peptides in two samples reacted with $ICAT_L$ and $ICAT_H$, respectively, can thus be determined by measuring ratios of labeled peptides analyzed by LC-MS/MS. The ICAT reagent also contains a biotin tag for selective enrichment of labeled peptides by avidin affinity chromatography.

ICAT was initially developed in a deuterated form as the first generation of reagents for global quantitative proteomics using stable isotope labeling [14].

ICAT has since been surpassed by alternative reagents (e.g., iTRAQ) that are better suited for general quantitative proteome analysis. Nevertheless, ICAT reagents are still very attractive in the niche of thiol-specific proteomics and the reagents have been successfully applied in various biological systems [15–22]. Since ICAT reacts with reduced thiol groups and not with oxidized forms (e.g., disulfide bonded cysteines) the reagent is well suited to monitor thiol oxidoreduction processes. Here, ICAT reagents are applied for Trx target identification in a multistep, differential labeling procedure (Fig. 2): (1) free thiols in protein extracts are blocked with IAM, (2) parallel samples incubated in the presence or absence of Trx are added IAM to quench Trx and block reduced thiols, (3) remaining protein thiols are fully reduced with tris(2-carboxyethyl) phosphine (TCEP) followed by addition of $ICAT_L$ and $ICAT_H$ to

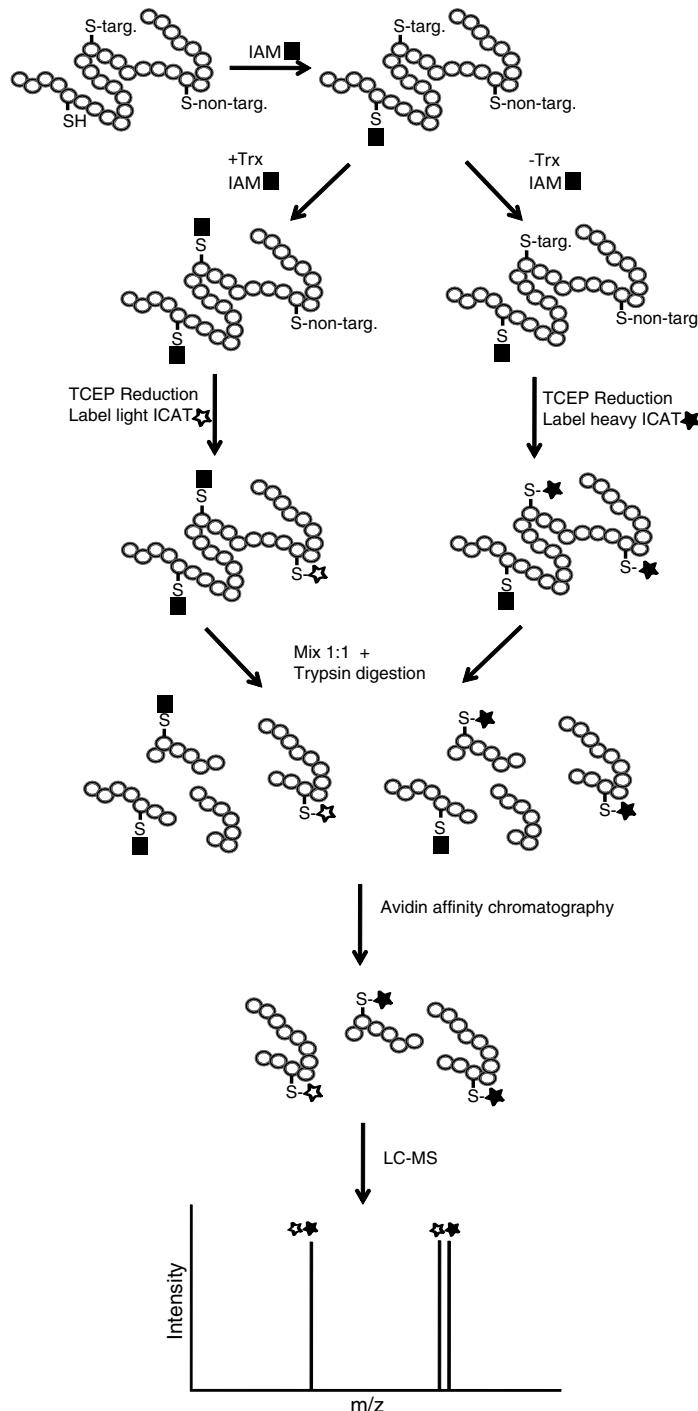


Fig. 2 Workflow for differential ICAT/IAM thiol labeling. Free reduced cysteine thiols (SH) are first blocked with IAM. Samples are incubated +/– Trx for reduction of target disulfides (S-targ.) and treated with IAM to block free thiols released by Trx. Remaining oxidized thiols (S-non-targ.) are then reduced using TCEP and free thiols are labeled with ICAT_L/ICAT_H. The ICAT labeled peptides are enriched by avidin affinity chromatography and analyzed by LC-MS/MS. The peptide identity is established via database searching of the product ion spectra and the relative intensities of ICAT_H/ICAT_L labeled precursor ions are calculated. ICAT_H/ICAT_L peptide ratios of 1 are expected for non-target disulfide bonds and ratios >1 are expected for peptides containing cysteines from disulfide bonds reduced by Trx

the sample exposed to Trx and the control (-Trx), respectively. The two samples are mixed, digested by trypsin and the ICAT labeled peptides are enriched by avidin chromatography. Finally, peptides are analyzed by LC-MS/MS. $\text{ICAT}_H/\text{ICAT}_L$ peptide ratios of 1 are expected for non-target disulfide bonds that have remained intact until the TCEP reduction. In contrast, $\text{ICAT}_H/\text{ICAT}_L$ ratios >1 are expected for peptides containing cysteines from disulfide bonds reduced by Trx that has been blocked by IAM and thus are not available for ICAT_L -labeling.

2 Materials (See Note 1)

2.1 Trx Treatment

Trx incubation buffer: 50 mM Tris/HCl, pH 8.0, 0.2 mM EDTA, 0.72 mM NADPH (see Note 2).

Slide-A-Lyzer dialysis cassettes 3.5K MWCO (Thermo Scientific P/N 66333).

2.2 ICAT Labeling and Trypsin Digestion

Surfactant: 1 % sodium dodecyl sulfate (SDS).

Reducing agent: 50 mM TCEP (see Note 3).

Cleavable ICAT reagent (Applied Biosystems P/N 4339038). Store at -20°C .

N-acetylcysteine (in 500 mM Tris/HCl pH 8.0).

Sequencing grade modified trypsin (Promega P/N V5111). Store at -20°C .

2.3 Purification of ICAT-Labeled Peptides

Cation exchange cartridge 200 μL (Applied Biosystems P/N 4326695).

Avidin affinity cartridge 200 μL (Applied Biosystems P/N 4326694).

Cartridge holder (Applied Biosystems P/N 4326688).

Needle port adaptor (Applied Biosystems P/N 4326689).

Outlet tubing kit (Applied Biosystems P/N 4326690).

Cation exchange loading buffer: 10 mM KH_2PO_4 , 25 % acetonitrile, pH 3.0.

Cation exchange elution buffer: 10 mM KH_2PO_4 , 350 mM KCl, 25 % acetonitrile, pH 3.0.

Cation exchange cleaning buffer: 10 mM KH_2PO_4 , 1 M KCl, 25 % acetonitrile pH 3.0.

Cation exchange storage buffer: 10 mM KH_2PO_4 , 25 % acetonitrile pH 3.0, 0.1 % NaN_3 .

Avidin affinity loading buffer: 20 mM NaH_2PO_4 , 300 mM NaCl, pH 7.2.

Avidin affinity elution buffer: 30 % acetonitrile, 0.4 % trifluoroacetic acid (TFA).

Avidin affinity washing buffer 1: 10 mM NaH₂PO₄, 150 mM NaCl, pH 7.2.

Avidin affinity washing buffer 2: 50 mM NH₄CO₃, 20 % methanol, pH 8.3.

Avidin affinity storage buffer: 20 mM NaH₂PO₄, 300 mM NaCl, pH 7.2, 0.1 % NaN₃.

Cleaving reagent A (Applied Biosystems P/N 4338543) (*see Note 4*).

Cleaving reagent B (Applied Biosystems P/N 4339052) (*see Note 5*). Store at -20 °C.

2.4 C18 StageTip Peptide Purification

Empore C18 Extraction Disk (3M, P/N 2215).

C18 loading buffer: 0.1 % TFA.

C18 elution buffer: 80 % acetonitrile, 0.1 % TFA.

3 Methods

3.1 Trx Treatment

Equilibrate approx. 200–400 µg protein extract (*see Note 6*) in 400 µL Trx incubation buffer and distribute the solution in four eppendorf tubes (A–D; 100 µL/tube). Add 2 µL NTR (final concentration 2 µM) and 9.6 µL Trx (final concentration 6.5 µM) to A and C (*see Note 7*); add 2 µL NTR and 9.6 µL water to B and D. After 1 h incubation at RT, add 12 µL IAM (final concentration 10 mM) to A and B and 12 µL water to C and D. Incubate the samples for 80 min at RT in the dark, cool on ice, add 100 µL 50 mM Tris/HCl pH 8.0, and dialyze against 50 mM Tris/HCl pH 8.0 overnight at 4 °C.

3.2 ICAT Labeling and Trypsin Digestion

Mix samples from Subheading 3.1 (72 µL) with surfactant (8 µL) and reducing agent (2 µL) and incubate at 99 °C for 10 min. Add, 20 µL 11.4 mM cleavable ICAT reagent (in acetonitrile) to A and C (ICAT_L) and B and D (ICAT_H) and incubate 80 min at RT in the dark, followed by quenching of excess ICAT reagent by addition of 10 µL 100 mM N-acetylcysteine (in 500 mM Tris/HCl pH 8.0) and incubate for 10 min at RT. Then mix the samples (A+B and C+D), add 45 µL 100 mM CaCl₂ and adjust the sample volume to 438 µL with water. Finally, add 10 µL sequencing grade modified trypsin (0.1 µg/µL) and incubate the samples at 37 °C overnight.

3.3 Cation Exchange Chromatography

Assemble the cation exchange cartridge into the cartridge holder with the needle port adaptor and the outlet tubing according to the manufacturer's instructions. Inject 2 mL cation exchange loading buffer to condition the cartridge. Transfer samples from

Subheading 3.2 to a 15 mL falcon tube, mix with 4 mL cation exchange loading buffer, and slowly (approx. 1 drop/second) inject onto the cation exchange cartridge. Inject 1 mL cation exchange loading buffer and discard the eluent. Inject 500 μ L cation exchange elution buffer and capture the eluted peptides in an eppendorf tube. Inject 1 mL cation exchange cleaning buffer and discard the flow through. If further samples are to be processed, re-equilibrate the cartridge with 2 mL cation exchange loading buffer and continue as described above, otherwise inject 2 mL cation exchange storage buffer, disassemble the cartridge holder and store the cartridge at 4–8 °C.

3.4 Avidin Affinity Chromatography

Assemble the avidin affinity cartridge into the cartridge holder with the needle port adaptor and the outlet tubing according to the manufacturer's instructions. Inject 2 mL avidin affinity elution buffer followed by 2 mL avidin affinity loading buffer to condition the cartridge. Mix samples from Subheading 3.3 with 500 μ L avidin affinity loading buffer and slowly (approx. 1 drop/second) inject onto the avidin affinity cartridge. Inject 500 μ L avidin affinity loading buffer followed by 1 mL avidin affinity washing buffer 1, 1 mL avidin affinity washing buffer 2 and 1 mL H₂O. Inject 800 μ L avidin affinity elution buffer, where the first 50 μ L eluting from the cartridge is discarded and the remaining 750 μ L containing ICAT-labeled peptides are stored. If further samples are to be processed, re-equilibrate the cartridge with 2 mL avidin affinity elution buffer followed by 2 mL avidin affinity loading buffer and repeat the procedure described above, otherwise inject 2 mL avidin affinity storage buffer, disassemble the cartridge holder and store the cartridge at 4–8 °C.

3.5 ICAT Tag Cleavage

Dry samples from Subheading 3.4 in a SpeedVac centrifuge and incubate with cleaving reagents A and B in a 95:5 ratio (100 μ L total volume) for 2 h at 37 °C to remove the biotin tag from ICAT labeled peptides (Fig. 1). Dry down the samples in a SpeedVac centrifuge.

3.6 C18 StageTip Peptide Purification

StageTips are prepared essentially as described previously [23] by mounting plugs of Empore C18 solid phase extraction disks in Gilson P200 pipette tips (see Note 8). Equilibrate the C18 StageTips by loading 10 μ L C18 elution buffer followed by 10 μ L C18 loading buffer. Dissolve samples from Subheading 3.5 in 10 μ L C18 loading buffer and load onto the stage tip. Wash the StageTips at least twice with 10 μ L C18 loading buffer. Finally, elute peptides with 10 μ L C18 elution buffer, capture the eluate in a 250 μ L eppendorf tube and dry in a SpeedVac centrifuge. Redissolve samples in 10 μ L C18 loading buffer and transfer to a 96-well plate for LC-MS injection.

3.7 LC-MS/MS Analysis

Inject samples from Subheading 3.6 on a nanoflow LC system connected to a fused silica column (75 μ m ID) packed with ReproSil-Pur C18-AQ 3 μ m. Elute peptides with a gradient of 5–40 % acetonitrile in 0.5 % acetic acid directly into an electrospray ionization mass spectrometer equipped with a nano-electrospray ion source. Acquire data in information dependent mode with a cycle of a survey mass spectrum followed by tandem mass spectra of the most intense multiply charged ions that subsequently are placed on a dynamic exclusion list for a suitable period of time.

3.8 Protein Identification and Quantification of ICAT Ratios

Centroided MS/MS data is searched against appropriate protein databases using a search engine such as Mascot (Matrix Science). The following variable modifications should be included: ICAT_H (cysteine), ICAT_L (cysteine), carbamidomethyl (cysteine). ICAT_H/ICAT_L ratios are calculated using software capable of reading the data files of the mass spectrometer used (*see Note 9*). All spectra should be manually inspected and non-quantifiable ICAT-labeled peptide pairs with interfering signals rejected. In general it is also recommended to do a label-swap experiment where heavy and light labels are exchanged. The fraction of the individual target disulfide bonds reduced by Trx can be calculated as:

$$100 \cdot \left[1 - \left(\frac{\text{ICAT}_H}{\text{ICAT}_L} \right)^{-1} \right].$$

4 Notes

1. Unless stated otherwise reagents, buffers and chromatographic material should be stored at 4–8 °C.
2. NADPH should be added immediately before use. Stock solutions of NADPH (e.g., 60 mM) prepared using argon-flushed ddH₂O and protected from light (e.g., in brown eppendorf tubes) can be stored at -20 °C for several months.
3. This solution should be freshly prepared.
4. Caution: contains concentrated TFA.
5. Contains a scavenger that reduces side reactions during the cleaving reaction.
6. In general, care should be taken to avoid oxidation during protein extraction. Buffers should be flushed with argon to remove as much oxygen as possible. If possible, a thiol alkylating agent should be included in the protein extraction buffer to block free thiols. For extraction of proteins from barley seed embryo a buffer containing 5 mM Tris/HCl pH 8.0, 1 mM CaCl₂, Complete® protease inhibitor cocktail, and 10 mM iodoacetamide (IAM) was used [3]. Excess IAM should be removed on

a PD-10 desalting column equilibrated with 50 mM Tris/HCl pH 8.0.

7. Depending on the biological system investigated, appropriate combinations of Trx and NTR are selected. Preparations of Trx and NTR should display electrophoretic homogeneity and enzymatic activity should be confirmed spectrophotometrically using Ellmans reagent as the final electron acceptor [24]. For reduction of target disulfide bonds in barley seed tissue recombinant His-tagged barley HvTrxh1 and HvNTR2 produced in *E. coli* were used [25, 26].
8. A detailed protocol for preparation of StageTips is available [27].
9. ICAT_H/ICAT_L ratios for barley thioredoxin targets were calculated using the software MSQuant (<http://www.msquant.sourceforge.net>). Depending on the MS platform used a number of software tools are available. In our hands ICAT_H/ICAT_L ratios between 0.86 and 2.64 were obtained for 199 labeled tryptic peptides from Trx-treated barley embryo extracts and cysteines in 104 peptides with ICAT_H/ICAT_L >1.22 were classified as Trx targets [3].

Acknowledgments

Karina Jansen and Mette Therese Christensen are acknowledged for technical assistance. The Danish Research Council for Technology and Production Sciences, the Carlsberg Foundation, and the Centre for Advanced Food Studies (LMC) are acknowledged for financial support.

References

1. Buchanan BB, Balmer Y (2005) Redox regulation: a broadening horizon. *Annu Rev Plant Biol* 56:187–220
2. Alkhalfioui F, Renard M, Vensel WH et al (2007) Thioredoxin-linked proteins are reduced during germination of *Medicago truncatula* seeds. *Plant Physiol* 144:1559–1579
3. Hägglund P, Bunkenborg J, Maeda K et al (2008) Identification of thioredoxin protein disulfide targets using a quantitative proteomics approach based on isotope-coded affinity tags. *J Proteome Res* 7:5270–5276
4. Hägglund P, Bunkenborg J, Yang F et al (2010) Identification of thioredoxin target disulfides in proteins released from barley aleurone layers. *J Proteomics* 73:1133–1136
5. Lemaire SD, Guillon B, Le Marechal P et al (2004) New thioredoxin targets in the unicellular photosynthetic eukaryote *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci U S A* 101: 7475–7480
6. Maeda K, Finnie C, Svensson B (2004) Cy5 maleimide labelling for sensitive detection of free thiols in native protein extracts: identification of seed proteins targeted by barley thioredoxin h isoforms. *Biochem J* 378:497–507
7. Maeda K, Finnie C, Svensson B (2005) Identification of thioredoxin h-reducible disulfides in proteomes by differential labelling of cysteines: insight into recognition and regulation of proteins in barley seeds by thioredoxin h. *Proteomics* 5:1634–1644

8. Marchand C, Le Marechal P, Meyer Y et al (2004) New targets of *Arabidopsis* thioredoxins revealed by proteomic analysis. *Proteomics* 4:2696–2706
9. Marchand C, Le Marechal P, Meyer Y et al (2006) Comparative proteomic approaches for the isolation of proteins interacting with thioredoxin. *Proteomics* 6:6528–6537
10. Marx C, Wong JH, Buchanan BB (2003) Thioredoxin and germinating barley: targets and protein redox changes. *Planta* 216:454–460
11. Wong JH, Cai N, Balmer Y et al (2004) Thioredoxin targets of developing wheat seeds identified by complementary proteomic approaches. *Phytochemistry* 65:1629–1640
12. Yamazaki D, Motohashi K, Kasama T et al (2004) Target proteins of the cytosolic thioredoxins in *Arabidopsis thaliana*. *Plant Cell Physiol* 45:18–27
13. Yano H, Wong JH, Lee YM et al (2001) A strategy for the identification of proteins targeted by thioredoxin. *Proc Natl Acad Sci U S A* 98:4794–4799
14. Gygi SP, Rist B, Gerber SA et al (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 17:994–999
15. Fu C, Wu C, Liu T et al (2009) Elucidation of thioredoxin target protein networks in mouse. *Mol Cell Proteomics* 8:1674–1687
16. García-Santamarina S, Boronat S, Espadas G et al (2011) The oxidized thiol proteome in fission yeast – optimization of an ICAT-based method to identify H₂O₂-oxidized proteins. *J Proteomics* 74:2476–2486
17. Leichert LI, Gehrke F, Gudiseva HV et al (2008) Quantifying changes in the thiol redox proteome upon oxidative stress *in vivo*. *Proc Natl Acad Sci U S A* 105:8197–8202
18. Sethuraman M, McComb ME, Huang H et al (2004) Isotope-coded affinity tag (ICAT) approach to redox proteomics: identification and quantitation of oxidant-sensitive cysteine thiols in complex protein mixtures. *J Proteome Res* 3:1228–1233
19. Sethuraman M, McComb ME, Heibeck T et al (2004) Isotope-coded affinity tag approach to identify and quantify oxidant-sensitive protein thiols. *Mol Cell Proteomics* 3:273–278
20. Sethuraman M, Clavreul N, Huang H et al (2007) Quantification of oxidative posttranslational modifications of cysteine thiols of p21ras associated with redox modulation of activity using isotope-coded affinity tags and mass spectrometry. *Free Radic Biol Med* 42:823–829
21. Wu C, Parrott AM, Liu T et al (2011) Distinction of thioredoxin transnitrosylation and denitrosylation target proteins by the ICAT quantitative approach. *J Proteomics* 74:2498–2509
22. Yuan W, Edwards JL (2011) Thiol metabolomics of endothelial cells using capillary liquid chromatography mass spectrometry with isotope coded affinity tags. *J Chromatogr A* 1218: 2561–2568
23. Rappaport J, Ishihama Y, 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
24. Holmgren A (1977) Bovine thioredoxin system. Purification of thioredoxin reductase from calf liver and thymus and studies of its function in disulfide reduction. *J Biol Chem* 252: 4600–4606
25. Maeda K, Häggblund P, Finnie C et al (2006) Structural basis for target protein recognition by the protein disulfide reductase thioredoxin. *Structure* 14:1701–1710
26. Shahpiri A, Svensson B, Finnie C (2008) The NADPH-dependent thioredoxin reductase/thioredoxin system in germinating barley seeds: gene expression, protein profiles and interactions between isoforms of thioredoxin h and thioredoxin reductase. *Plant Physiol* 146: 789–799
27. Rappaport J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2: 1896–1906

Chapter 48

Class III Peroxidases

Sabine Lüthje, Claudia-Nicole Meisrimler, David Hopff, Tim Schütze, Jenny Köppe, and Katrin Heino

Abstract

Class III peroxidases are *heme*-containing proteins of the secretory pathway with an extremely high number of isoenzymes, indicating the tremendous and important functions of this protein family. This chapter describes fractionation of the cell in subproteomes, their separation by polyacrylamide gel electrophoresis (PAGE) and visualization of peroxidase isoenzymes by heme and specific in-gel staining procedures. Soluble and membrane-bound peroxidases were separated by differential centrifugation. Aqueous polymer two-phase partitioning and discontinuous sucrose density gradient were applied to resolve peroxidase profiles of plasma membranes and tonoplast. Peroxidase isoenzymes of subproteomes were further separated by PAGE techniques such as native isoelectric focussing (IEF), high resolution clear native electrophoresis (hrCNE), and modified sodium dodecyl sulfate (modSDS)-PAGE. These techniques were used as stand-alone method or in combination for two-dimensional PAGE.

Key words Cell fractionation, Plasma membrane, Tonoplast, hrCNE, IEF, modSDS-PAGE, Peroxidase in-gel staining

Abbreviations

| | |
|---------|---|
| 2D | Two dimensional |
| AB | Acrylamide/Bis |
| ACA | Aminocaproic acid, 6-aminohexanoic acid |
| APS | Ammonium persulfate |
| BCA | Bicinchoninic acid |
| BLAST | Basic Local Alignment Search Tool |
| BN-PAGE | Blue native PAGE |
| BSA | Bovine serum albumin |
| CHAPS | 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate |
| DAB | 3'3'-Diaminobenzidine |
| DMSO | Dimethyl sulfoxide |
| DTT | Dithiothreitol |
| EDTA | Ethylenediaminetetraacetic acid |
| ESI | Electrospray ionization |

| | |
|-------------|--|
| GB | Gel buffer |
| Guaiacol | 2-Methoxyphenol |
| HEPES | 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid |
| hrCNE | High resolution clear native electrophoresis |
| ICM | Intracellular membranes |
| IEF | Isoelectric focussing electrophoresis |
| LB | Loading buffer |
| MALDI | Matrix-assisted laser desorption/ionization |
| modSDS-PAGE | Modified SDS-PAGE |
| PAGE | Polyacrylamide gel electrophoresis |
| PEG | Poly-ethylene glycol |
| PM | Plasma membrane |
| PMSF | Phenylmethylsulfonyl fluoride |
| PVPP | Polyvinylpolypyrrolidone |
| TEMED | <i>N,N,N',N'</i> -Tetramethyl-ethylenediamine |
| TIFF | Tag Image File Format |
| TMB | Tetramethylbenzidine |
| Tris | Tris(hydroxymethyl)aminomethane 2-amino-2-hydroxymethyl-propane-1,3-diol |

1 Introduction

The superfamily of plant *heme*-containing peroxidases (E.C. 1.11.1.7) is one of the most intensively investigated protein families in plant research. Besides intracellular ascorbate peroxidases (class I peroxidases), peroxidases of the secretory pathway (class III) belong to this protein superfamily [1]. The number of class III peroxidases is extremely high and they are involved in several processes like plant development, cell metabolism, cell elongation, pathogen defense, symbiosis, fruit ripening, senescence and stress responses etc. [2]. The steadily increasing number of identified and predicted peroxidase sequences is documented in the PeroxiBase at <http://peroxibase.toulouse.inra.fr/> [3]. A recent in silico analysis suggests that a significant high number of isoenzymes may be membrane-bound [4]. So far, the majority of investigations studied soluble peroxidases [5–7], whereas much less is known of membrane-bound isoenzymes [8–10].

Here we present a protocol that allows investigation of soluble and membrane-bound class III peroxidases simultaneously in the same sample. After harvesting of plant material, the sample is fractionated to get a higher resolution for the numerous peroxidase isoenzymes. Soluble isoenzymes are prepared by differential centrifugation, concentrated by ammonium sulfate precipitation and used without further fractionation for analysis. In contrast to the soluble isoenzymes, the membrane fraction is further separated. The state of the art for plasma membrane preparation is aqueous polymer two-phase partitioning. This technique was introduced in

the 1980s [11, 12] and is the most widely adopted method to prepare highly enriched plasma membranes from plants [13, 14]. The method is based on surface properties of the membrane [15] that alter during plant development and by biotic or abiotic stress factors due to changes in lipid and protein composition. The overall composition of the plasma membrane can differ between plant species and tissues as well as from one variety to another. Those differences and alterations were clearly demonstrated by proteomic and lipodomic approaches [16–19]. Changes in the membrane surface properties may affect yield and purity of plasma membrane preparations. This has to be kept in mind, if aqueous polymer two-phase partitioning is adopted. An example of this method is described, which has been shown to produce highly enriched plasma membrane fractions for maize (*Zea mays* L.) and wheat (*Triticum aestivum* L.) roots [20–22]. Aqueous polymer two-phase systems for *Arabidopsis thaliana* (L.) Heynh. leaves and *Nicotiana tabacum* L. BY2 cells or for other species and tissues were described elsewhere [15, 23–25].

Sucrose density gradient centrifugation is widely used for tonoplast preparation [26–28]. Tonoplast and plasma membranes are similar in buoyant density but different in surface properties. To avoid contamination by plasma membranes, we use the first lower phase after aqueous polymer two-phase partitioning (depleted in plasma membranes) to separate tonoplast from intracellular membranes (ICM). The latter fraction contains endoplasmic reticulum, Golgi, microbodies, etc. and may be also used for further analysis.

Marker analysis is a crucial point after cell fractionation, because marker abundance may also be changed by stress or other factors [18]. Purity of a membrane preparation can be investigated either by specific staining using transmission electron microscopy, marker enzyme activity or protein-immunoblot analysis [15, 20–22].

To avoid contamination by soluble proteins that are entrapped inside membrane vesicles or attached to the membrane during preparation samples have to be washed [29]. Interaction of attached proteins with the membrane or integral proteins can be destroyed by high salt concentrations. In our studies we use physiological salt conditions for the removal of proteins, which avoids the loss of peripheral membrane proteins. Class III peroxidases appear to interact with membranes by a transmembrane spanning domain [9, 30]. Thus proteins have to be solubilized before polyacrylamide gel electrophoresis (PAGE). Non-ionic detergents or a combination of the zwittergent CHAPS and aminocaproic acid (ACA) have been used for solubilization of plasma membrane-bound peroxidases in the past [4, 18, 31].

After sample preparation, proteins can be separated by native or in-native PAGE to get a higher resolution of the numerous peroxidases. Native isoelectric focussing (IEF) is a standard method to characterize peroxidase isoenzymes [32]. Besides estimation of the

isoelectric points, this method allows calculation of the peroxidase activity. After electrophoresis, the gel is equilibrated in buffer at adequate pH and water insoluble peroxidase substrates are applied. After conversion of the substrate in the presence of H_2O_2 the colored reaction product precipitates inside the gel. After digitalization, intensity of the bands can be estimated using ImageJ (<http://rsbweb.nih.gov/ij/>). Quantification of the colored spots is possible if non-saturating protein concentrations are used and both biological and technical replicates have been done in adequate amounts [18].

High resolution clear native electrophoresis (hrCNE) is a native PAGE that separates proteins according to their native molecular mass [33]. The ionic detergent deoxycholate is used to introduce a negative charge to the solubilized proteins before electrophoresis. Compared to blue native PAGE (BN-PAGE) the advantage of hrCNE is that it is colorless.

Modified SDS-PAGE (modSDS-PAGE) was used to visualize plasma membrane-bound peroxidases [31]. The method is based on the protocol of Thomas et al. [34]. In contrast to SDS-PAGE, modSDS-PAGE works with a reduced concentration of SDS and without reducing agents like dithiothreitol (DTT) or mercaptoethanol in the sample buffer. Although modSDS-PAGE is not a native method *heme*-containing peroxidases can be visualized by staining with tetramethylbenzidine (TMB) and H_2O_2 [31]. Staining is finished within two minutes, because the prosthetic *heme* group is still present but the enzyme is not active any more. This method allows estimation of molecular masses and abundance of the isoenzymes [21]. ModSDS-PAGE can be used either as a first dimension or as a second dimension after native gel-based or gel-free separation methods. Gradient gels may be used to receive a higher resolution of isoperoxidases.

Originally TMB was used for staining of the peroxidase activity of *heme*-containing cytochrome P-450 [34]. It was shown that TMB also stains copper-containing proteins [35]. Thus TMB will react with all *heme*- and copper-containing proteins in a sample. Among these proteins are class I and class III peroxidases. Besides TMB staining, class III peroxidases can be detected by diaminobenzidine (DAB) or more specifically by phenolic substrates that precipitate after conversion in aqueous solutions [36–38]. Guaiacol (2-methoxyphenol) and α -chloro-naphthol are used for visualization of the activity of class III peroxidases. DAB precipitates as an intense brown reaction product, α -chloro-naphthol produces violet bands and the product of guaiacol is orange [32]. For BN-PAGE DAB and guaiacol are the best appropriate peroxidase stains due to their complementary contrast to Coomassie Blue [39], whereas TMB and α -chloro-naphthol can be used additionally for native IEF-PAGE, hrCNE or modSDS-PAGE [10, 31, 32, 40]. After visualization with TMB, peroxidase spots can be picked,

destained, and digested by standard protocols for mass analysis [9, 41]. Membrane-bound peroxidases were identified by matrix-assisted laser desorption/ionization (MALDI) or electrospray ionization (ESI) mass spectrometry in the past [9, 18]. Identified peptides can be used for BLAST searches in any database available in the World Wide Web. PeroxiBase, however, may be the preferred database for peroxidases. Although this chapter is focussed on maize peroxidases, the protocols for PAGE can be applied to other samples and enzyme activity stains too [40].

2 Materials

Prepare all solutions using ultrapure water and analytical grade reagents. Prepare and store all reagents at 4 °C (unless indicated otherwise). All peroxidase activity staining solutions (50 mL) are calculated for one mini-gel (ca. 10 cm × 7 cm).

2.1 Cell Fractionation

A refrigerated centrifuge with angle rotors for 100 and 50 mL tubes is used for preparation of microsomal fractions by differential centrifugation. Plasma membranes and tonoplast are prepared with a swing-out rotor for 50 mL tubes. Membrane fractions are collected by ultracentrifugation with an angle rotor for 22 mL tubes (see Note 1).

1. Homogenization buffer: 250 mM sucrose, 50 mM HEPES-KOH, 1 mM EDTA, pH 7.5. Supplement buffer with 1 mM DTT and 1 % polyvinylpyrrolidone (PVP) directly before use.
2. 100 mM phenylmethylsulfonyl fluoride (PMSF) in 2-propanol.
3. Resuspension medium: 250 mM sucrose and 50 mM HEPES-KOH, pH 7.0.

2.2 Soluble Fraction

1. Ammonium sulfate (Applichem, Germany).
2. Resuspension medium S (for soluble proteins): 50 mM Tris-HCl and 1 mM EDTA, pH 7.5.

2.3 Plasma Membrane Preparation by Aqueous Polymer Two-Phase Partitioning

1. 20 % (w/w) Dextran solution. Weigh 100 g Dextran T500 powder (Pharma Cosmos, Denmark) and 400 g H₂O. Transfer first the Dextran and then the H₂O to a glass bottle and stir it over night. Store the solution in freezer.
2. 40 % (w/w) Poly-ethylene glycol (PEG) solution: Weigh 100 g PEG 3350 powder (Sigma-Aldrich, Germany) and 150 g H₂O. Transfer first the PEG and then the H₂O to a glass bottle and solve it by stirring. Store the solution in freezer.
3. 200 mM PO₄[−] buffer, pH 7.8: Prepare 100 mL Na₂HPO₄ (2.84 g) and KH₂PO₄ (2.72 g) solutions. Adjust Na₂HPO₄ solution to pH 7.8 with the KH₂PO₄ solution.
4. 200 mM KCl solution.

5. Phase buffer: 250 mM sucrose, 5 mM KCl, and 5 mM PO_4^- buffer. Weigh 8.56 g sucrose, transfer it to a 100 mL volumetric flask, add 2.5 mL 200 mM KCl and 200 mM PO_4^- buffer, fill up ad 100 mL with H_2O .

2.4 Tonoplast

1. 1 M Tris solution.
2. 200 mM Tris solution.
3. Tonoplast resuspension medium: 250 mM mannit, 100 mM tricine, 3 mM MgSO_4 , 3 mM EGTA, pH 7.5. Adjust the pH with 1 M Tris-solution.
4. Gradient solution: 28 % sucrose, 5 mM HEPES, and 2 mM DTT, pH 7.5. Adjust the pH with 200 mM Tris solution.
5. Dilution medium: 3 mM MgSO_4 and 50 mM HEPES–NaOH, pH 7.0.
6. Storage medium: 40 % Glycerol, 1 mM EDTA, 10 mM HEPES, pH 7.0. Adjust the pH with 0.1 M NaOH.

2.5 Protein Quantification

1. Bradford [42] reagent: Dissolve 50 mg Coomassie-Brilliant Blue G in 50 mL ethanol, mix it with 100 mL of 85 % phosphoric acid. Pour the acid solution carefully in H_2O and fill it ad 1 L. Filter the solution into a brown bottle and store it at room temperature.
2. Dilutor: 0.9 % NaCl and 0.01 % Triton X-100 in Milli Q water (freshly prepared).
3. Bovine serum albumin (BSA) Standard solution: 0.02 % BSA (5 μL =1 mg protein).
4. UV/VIS-Spectrophotometer.

2.6 Membrane-Washing

1. Wash buffer: 50 mM HEPES–KOH buffer, pH 7.0, 1 mM EDTA, 0.01 % Triton X-100 and 150 mM KCl.

2.7 Electrophoresis

All glass plates used for casting gels should be degreased before use. Mini-Protean unit (Bio-Rad, Germany) and SE 250 vertical unit (GE Healthcare, Germany) are suitable for the electrophoresis types described but the protocols can be also adapted to bigger systems. Power supplies should allow for a voltage around 100–500 V and a current around 200–500 mA. A peristaltic pump and a gradient mixer are required to cast acrylamide gradient gels.

1. Ready-made acrylamide/Bis solution (AB-solution) (Rotiphoresse® Gel 30 (37, 5:1)) (Carl Roth, Germany). Store at 4 °C.
2. Ammonium persulfate (APS): 10 % solution in water (see Note 2).
3. N,N,N,N'-tetramethyl-ethylenediamine (TEMED) (Sigma-Aldrich, Germany). Store at 4 °C.
4. Pre-stained protein ladder (Thermo Scientific, Germany).

2.8 Modified SDS-PAGE (modSDS-PAGE) Components

1. 4× Resolving gel buffer: 1.5 M Tris-HCl with 0.4 % SDS, pH 8.8. Weigh 90.86 g Tris and transfer it to a 500 mL volumetric flask. Add about 350 mL water to the flask and solve the Tris. Weigh 2 g SDS and add it carefully to the Tris solution. When the SDS is solved adjust pH with HCl (see Note 3). Make up to 500 mL with water.
2. 4× Stacking gel buffer: 0.5 M Tris-HCl with 0.4 % SDS, pH 6.8. Weigh 30.2 g Tris and prepare a 500 mL solution as in previous step.
3. 10× SDS-PAGE electrophoresis buffer: 0.25 M Tris, 1.92 M glycine, 1 % SDS.
4. Loading buffer (SDS-LB): 125 mM Tris-HCl, 0.2 % (w/v) SDS, 20 % (w/v) glycerol, and 0.004 % (w/v) bromo-phenol, pH 6.8.
5. Equilibration buffer for the second dimension: 1× stacking gel buffer with 1 % SDS plus 10 % glycerol and a trace of Bromophenol blue.

2.9 First Dimension modSDS-PAGE

1. First dimension modSDS-PAGE: 4 and 18 % acrylamide resolving gel mixture (each 5 mL): Solution I: Mix 0.625 mL AB-solution, 1.25 mL resolving gel buffer (4×), 0.303 g urea, and 3.125 mL water. Solution II: Mix 2.92 mL AB-solution, 1.25 mL resolving gel buffer (4×), 0.303 g urea and fill up with glycerol to 5 mL. Mix well and degas both solutions before use.
2. 4 % Acrylamide stacking gel mixture for the first dimension (5 mL): Mix 0.625 mL AB-solution, 1.25 mL stacking gel buffer (4×), 0.303 g urea and 3.125 mL water, degas it and store cold.

2.10 Second Dimension modSDS-PAGE

1. Second dimension after hrCNE: 4 and 16 % acrylamide resolving gel mixture (each 5 mL): Solution I: Mix 0.625 mL AB-solution, 1.25 mL resolving gel buffer (4×) and 3.125 mL water. Solution II: Mix 2.67 mL AB-solution, 1.25 mL resolving gel buffer (4×) and fill up with glycerol to 5 mL. Degas both solutions before use.
2. Second dimension after native IEF: 4 and 18 % acrylamide resolving gel mixture (each 5 mL): Solution I: Mix 0.625 mL AB-solution, 1.25 mL resolving gel buffer (4×) and 3.125 mL water. Solution II: Mix 2.92 mL AB-solution, 1.25 mL resolving gel buffer (4×) and fill up with glycerol to 5 mL. Degas both solutions before use.
3. 4 % Acrylamide stacking gel mixture for the second dimension (5 mL): Mix 0.625 mL AB-solution, 1.25 mL stacking gel buffer (4×) and 3.125 mL water, degas it and store cold.

2.11 Native IEF-PAGE Components

1. Gel solution: 3 M urea, 2 % CHAPS, 10 % glycerol, 2 % carrier ampholytes pH 3–10 (Serva, Germany), and 7.5 % AB-solution. Solution is ultra sonicated until urea and CHAPS are solved.
2. Cathode buffer: 20 mM lysine and 20 mM arginine (Biorad, Germany).
3. Anode buffer: 10 mM phosphoric acid.
4. 3× loading buffer (3× IEF-LB): 8 % Ampholyte, 8 % CHAPS, 40 % glycerol, 3 M urea.

2.12 High Resolution Clear Native PAGE (hrCNE) Components

1. 3× Gel buffer (3× hrCNE-GB): 75 mM imidazole/HCl and 1.5 M 6-aminocaproic acid (ACA), pH 7.0.
2. Cathode buffer: 50 mM Tricine, 7.5 mM imidazole, 0.05 % deoxycholate, and 0.05 % triton X-100, pH 7.0. The pH is not adjusted.
3. Anode buffer: 25 mM imidazole-HCl, pH 7.0.
4. Loading buffer (hrCNE-LB): 50 mM Imidazole-HCl, 500 mM ACA, 1 mM EDTA, 20 % glycerol, and a trace of Ponceau S.
5. 20 % Digitonin (Sigma-Aldrich, Germany) dissolved in 50 % glycerol by heating up to 95 °C.

2.13 Peroxidase In-Gel Activity Staining Components

Gel staining should be documented by a photo scanner or a digital camera in Tag Image File Format (TIFF) and a resolution of 300–600 DPI for further preparation.

1. 30 % H₂O₂ stock solution (Sigma-Aldrich, Germany) stored at 4 °C.
2. TMB staining solution: Solve 22.5 mg TMB in 15 mL methanol (see Note 4). After the TMB is solved add 35 mL 250 mM Na-acetate buffer pH 5.0.
3. DAB staining solution: 80 mg DAB (Sigma-Aldrich, Germany) are solved in 1 mL dimethyl sulfoxide (DMSO), then add it to 50 mM Na-acetate buffer with 10 mM CaCl₂, pH 5.0 and mix it.
4. α-Chloro-naphthol-staining solution: 0.01 % α-chloro-naphthol in 0.1 M Na₂HPO₄ buffer, pH 6.5.
5. Guaiacol-staining solution: 0.5 % (v/v) Guaiacol in 50 mM Na-acetate buffer and 10 mM CaCl₂, pH 5.0.

3 Methods

3.1 Cell Fractionation

General cell fractionation is exemplified for 5 days old maize roots but can be applied similar to other maize root systems. For other plants and tissues slight modifications might be needed for plasma membrane and tonoplast preparation. All work steps are accomplished at 4 °C, if not mentioned otherwise.

1. 80–130 g maize roots are harvested within 30 min and homogenized with 200 mL homogenization buffer (Subheading 2.1), supplemented with 1 mM DTT and 1 % insoluble PVPP, using a warring blender (3×15 s with a break of 20 s between each homogenizing step). The crude extract is then filtered through a nylon net (125 μ m mesh; Hydrbios, Germany) and 2 mL 100 mM PMSF are added.
2. Intact cell organelles and cell walls (pellet) are separated from the soluble protein fraction and membrane protein fraction (supernatant) by centrifugation at $9,000 \times g$ for 10 min.
3. The supernatant containing soluble protein fraction and microsomes are separated at $50,000 \times g$ for 30 min.
4. The achieved microsomal pellet is resuspended in phase buffer (Subheading 2.3), but the supernatant (soluble proteins) is also kept. Microsomes and soluble fractions are stored below -70 °C until further use.

3.2 Soluble Proteins

1. All soluble proteins (achieved in Subheading 3.1) are precipitated with ammonium sulfate, which is added slowly, under stirring, to the solution (see Note 5) until the concentration reaches 90 % (662 g/L) (see Note 6).
2. The precipitation is accomplished over night at 4 °C under continues stirring.
3. Centrifuge the solution at $50,000 \times g$ for 30 min.
4. Resuspend the pellet in 25–50 mL (see Note 7) resuspension medium S (Subheading 2.2) and desalt 500 μ L sample using centrifugal filter units cut off 10,000 MWCO (Millipore, Germany). After desalting and concentration determine the protein concentration (see Note 8). The final protein concentration should be adjusted to 1–5 mg/mL by filling with resuspension medium S. Store the sample till use below -70 °C.

3.3 Plasma Membrane Preparation by Aqueous Polymer Two-Phase Partitioning

1. Plasma membranes are purified from microsomal fractions (about 80 mg protein) by six steps of aqueous polymer two-phase partitioning using a 36 g phase system consisting of 6.5 % (w/w) dextran T500, 6.5 % (w/w) PEG 3350, 250 mM sucrose, 5 mM KCl, 5 mM phosphate buffer, pH 7.8 (Fig. 1) (see Note 9).
2. 9 g of microsomal fraction (see Subheading 3.1, step 4; keep a part of the microsomal sample for the later analysis) are added to a 27 g phase mixture. The phase system is then mixed thoroughly by 20–30 inversions of the tube and centrifuged at $1,000 \times g$ for 5 min in a swing-out rotor.
3. The resulting upper phase is transferred to the next lower phase, resulting in a 36 g phase system (see Note 10), mixed and centrifuged as explained in the previous step. This procedure is

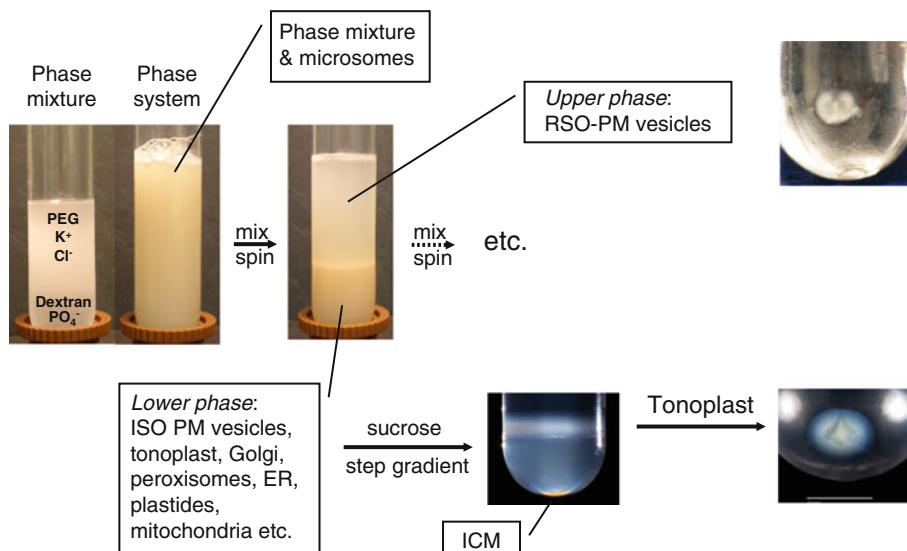


Fig. 1 Membrane fractionation. Aqueous solutions of polyethylene glycol (PEG 3350) and Dextran T-500 separate in two phases in the cold [16]. Density of Dextran is higher compared to PEG. Phosphate has a higher affinity for Dextran, whereas potassium is enriched in the PEG phase. Chloride has a higher affinity for PEG compared to potassium. Thus the hydrophobic PEG phase becomes positive and the hydrophilic Dextran phase becomes negative. After addition of the microsomal fraction to the phase mixture, the phase system is complete and reaches its final concentrations. The system is mixed and separated after centrifugation. The upper phase is transferred to a fresh lower phase, mixed and centrifuged. This step can be repeated 1–5 times. Right-side-out (RSO) plasma membranes (PM) with a negative surface charge are enriched in the PEG phase, whereas inside-out (ISO) PM vesicles, endoplasmatic reticulum (ER), tonoplast, Golgi vesicles, peroxisomes, plastids, and mitochondria are enriched in the first lower phase. After a washing step, the pellet of the lower phase is applied onto a sucrose step gradient to separate the tonoplast from other intracellular membranes (ICM) by discontinuous sucrose density centrifugation

repeated five times, always with a new lower phase. Keep the first lower phase for the tonoplast preparation (explained in Subheading 3.4).

4. The last upper phase (PM) is washed twice with resuspension medium (Subheading 2.1) and pelleted both times by ultracentrifugation at $100,000 \times g$ for 45 min (supernatant is discarded). The final pellet is resuspended in resuspension medium and stored below -70°C until further use.
1. The resulting first lower phase of the plasma membrane preparation is diluted with 200 mL resuspension medium (Subheading 2.1), pelleted for 1 h at $50,000 \times g$, resuspended and washed with 200 mL resuspension medium followed by pelletation for 30 min at $50,000 \times g$.
2. Discard the supernatant and resuspend the pellet in 27 mL tonoplast resuspension medium (Subheading 2.4).

3.4 Tonoplast

- Transfer the 27 mL sample (6–11 mg protein per tube) in a centrifuge tube (50 mL) and underlay it carefully with 8 mL of the gradient medium, using a peristaltic pump (4 mL/min). Centrifuge at $10,000 \times g$ for 3 h at 4 °C in a swing-out rotor.
- The opaque tonoplast enriched band is localized between the sucrose cushion and the supernatant (Fig. 1). It is transferred carefully into a centrifugation tube and filled up to 35 mL with dilution medium and centrifuged at $50,000 \times g$ for 30 min at 4 °C. The pellet on the bottom of the tube presents the intracellular membranes (Endoplasmic reticulum, Golgi, micro-bodies, etc.), which is resuspended in dilution medium and centrifuged similar to the tonoplast sample.
- The final pellets are resuspended in the storage medium and stored below –70 °C until use.

3.5 Protein Quantification

- Add $x\mu\text{L}$ of the sample corresponding to 2–10 µg of protein to a 3 mL cuvette.
- Add dilutor up to 500 µL and mix well.
- Add 2.5 mL Bradford reagent, mix and wait for 5–10 min.
- Measure with the spectrophotometer at 595 nm and compare to a standard curve prepared with BSA (2–12 µg of protein) (see Note 8).

3.5.1 Electrophoresis

The different protein fractions are first isolated as explained in Subheadings 3.1 to 3.4 and followed by modSDS-PAGE or native electrophoretic techniques (hrCNE and native IEF). The used electrophoretic methods are soft enough to allow the detection of peroxidases with different staining methods.

All methods have in common that no reducing agents are used and samples are not heated. The modSDS-PAGE, which is also applied for the second dimension, contains 0.1 % SDS in all buffer systems. However, SDS concentrations up to 1 % are possible for the detection of peroxidases. The enzymes are still specifically detectable by this procedure as they stay intact. The two other electrophoretic methods, IEF and hrCNE, do work without SDS.

3.6 Preparation of Polyacrylamide Slab Gels

3.6.1 Modified SDS-Gel

The modSDS-PAGE is used for the first dimension as well as for the second dimension after IEF and hrCNE. For the modSDS-PAGE in the first dimension 1 M urea is added to the solutions for a better resolution, whereas no urea is used for the second dimension. To get a better resolution of the proteins modSDS-PAGE is casted as a gradient gel: 4–18 % for first dimension or second dimension after IEF and 4–16 % for second dimension after hrCNE.

- Prepare 4 and 16 % acrylamide resolving gel mixture for second dimension after hrCNE (Subheading 2.10) or 4 and 18 % acrylamide resolving gel mixture for second dimension after

native IEF (Subheading 2.10) and first dimension modSDS-PAGE (Subheading 2.9).

2. Prepare the 4 % acrylamide stacking gel mixture (Subheadings 2.9 and 2.10)
3. Mount the gel cassette (1 mm spacer glass plate for the second dimension or 0.75 mm for the first dimension).
4. For the 1 mm gels add 2.5 mL (0.75 mm; 2 mL) solution I and solution II to the two chambers of a gradient mixer with the connecting channel closed (*see Note 11*).
5. Add 13 μ L 10 % APS (0.75 mm; 11 μ L) and 1.3 μ L TEMED (0.75 mm; 1.1 μ L) to each acrylamide mixture in the gradient mixer and agitate using a magnetic stirrer.
6. Open the connecting channel of the gradient mixer and pump the solution in the gel cassette until the mixer is empty (4 mL/min).
7. Overlay the gel with 2-propanol and let the gel polymerize at room temperature (~30 min).
8. Remove the 2-propanol from the top of the gel after polymerization and wash it few times with water and remove the water.
9. Add 35 μ L 10 % APS and 5 μ L TEMED to the 4 % acrylamide stacking gel mixture. Pour the solution on top of the polymerized resolving gel. For the second dimension fill gel solution on the top of the resolving gel (~0.5 cm) and overlay with 2-propanol. For the first dimension cast up to the top and apply a comb. (polymerization ~15 min or over night at room temperature).
10. Remove the 2-propanol, wash with water and wrap gel in water soaked paper and foil. Store the gel at 4 °C until use.

3.6.2 Native IEF-Gel

The native IEF-polyacrylamide gel should be prepared the day before use. Do not keep it in storage.

1. Prepare the 7.5 % acrylamide gel mixture (7 mL): Mix 1.75 mL AB-solution, 0.14 g CHAPS, 1.26 g urea, 350 μ L 40 % ampholyte solution pH 3–10 and solve the components in a ultrasonic bath then fill up with water to 7 mL (*see Note 12*).
2. Mount two gel cassettes (2× spacer glass plates 0.75 mm, 2× front glass plates).
3. Add 10 % APS (70 μ L) and TEMED (7 μ L), to the 7.5 % IEF gel solution, mix it by pivoting and fill it into the casting cassette until the liquid reaches the top and insert the comb.
4. Cover the gel with foil and let the gel polymerize over night at RT.

3.6.3 hrCNE Gradient Gel

Casting one 4–16 % hrCNE gradient gel (0.075 cm × 8 cm × 10 cm) is exemplified.

1. Prepare the 4 % acrylamide gel mixture (5 mL): Mix 0.625 mL AB-solution, 1.67 mL hrCNE-GB (3×) (Subheading 2.12, item 1) and 2.68 mL water.
2. Prepare the 16 % acrylamide gel mixture (5 mL): Mix 2.67 mL AB-solution, 1.67 mL hrCNE-GB (3×), fill up to 5 mL with glycerol.
3. Mount the gel cassette using two glass plates (1 mm × 0.75 mm spacer glass plate, 1× front glass plate).
4. Add 1.75 mL 4 and 16 % acrylamide mixture to the two chambers of a gradient mixer with the connecting channel closed (see Note 11).
5. Add 10 % APS (10 µL) and TEMED (1 µL), to each acrylamide mixture in the gradient mixer and agitate using a magnetic stirrer.
6. Open the connecting channel of the gradient mixer and pump the solution in the gel cassette.
7. Overlay the gel with 2-propanol and let the gel polymerize at room temperature (~30 min).
8. Remove the 2-propanol from the top of the gel after polymerization and wash it few times with water and remove the water.
9. Add 17.5 µL 10 % APS and 1.75 µL TEMED to the rest of the 4 % acrylamide gel mixture (see step 1), pour this solution on top of the polymerized gradient resolving gel, and insert the comb.
10. The sample gel polymerizes within 20 min.
11. Remove the comb and overlay with water and wrap it in water soaked paper and foil. Store the gel at 4 °C until use.

3.7 Sample Preparation for Electrophoresis

3.7.1 Membrane Washing Before Solubilization (Optional Step)

1. Pellet membrane samples at 16,000×*g* for 90 min, remove supernatant and resuspend the pellet in washing buffer. Shake samples at 4 °C for 60 min.
2. Pellet washed membranes at 16,000×*g* for 90 min and remove supernatant.

3.7.2 Solubilization and Preparation for Electrophoresis

1. Prepare three aliquots of each samples (soluble fraction, microsomes, plasma membranes and tonoplast) in 1 mL tubes (see Note 8): 20 µg for the soluble fraction and 50 µg protein for membrane fractions for first dimension PAGE; 50–75 µg of soluble proteins and 250 µg protein for membrane fractions for separation in 2D-PAGE.
2. Pellet membranes at 16,000×*g* for 90 min at 4 °C and remove supernatant.

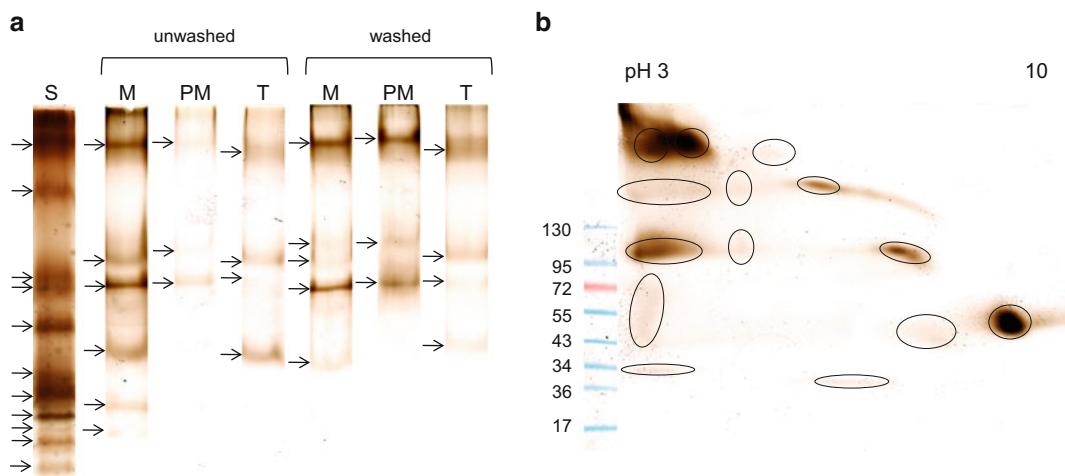


Fig. 2 ModSDS-PAGE as first and second dimension. (a) Peroxidases were separated by modSDS-PAGE (4–18 %) with 1 M urea for soluble (S), microsomal fractions (M), plasma membrane (PM) and tonoplast (T). On each lane 20 µg total protein were loaded. Membrane fractions were used either unwashed or washed with 150 mM KCl in the presence of 0.01 % Triton X-100. (b) 2D-PAGE separation of soluble peroxidases (20 µg protein) in native IEF-PAGE, pH 3–10 followed by second dimension modSDS-PAGE (4–18 %). On the left prestained marker (Fermentas, Germany) is shown with corresponding molecular weight in kD. Peroxidases were stained with DAB (a, b)

3. Resuspend the membranes in 20 µL IEF-LB (for native IEF and modSDS-PAGE) or hrCNE-LB (both 1× concentration).
4. Add digitonin to the hrCNE samples in the protein to detergent ratio of 1:5.
5. Solubilize samples for 1 h on ice with interim mixing.
6. Centrifuge samples at 16,000× \bar{g} for 60 min at 4 °C. Load supernatant for IEF and hrCNE on the corresponding gel. For modSDS-PAGE supernatant is desalting, using ZebaTM Spin Desalting Columns, 7K MWCO (Thermo Scientific, Germany). Desalting samples are mixed 1:1 with (2× SDS-LB).
7. Soluble proteins (20 µg) can be directly mixed with the sample buffers and loaded on the gel.

3.8 Electrophoresis

3.8.1 ModSDS-PAGE

1. Put the gel system together. Fill both chambers with running gel buffer.
2. Transfer the samples (Subheading 3.7) in the gel pockets. Add one marker lane (Fig. 2).
3. Electrophoresis is carried out with 30 mA per gel up to 200 V at 4 °C until the blue front reaches the bottom of the gel.

4. After electrophoresis, gels are transferred to in-gel staining buffers (Subheading 3.9).

3.8.2 Native IEF and Second Dimension modSDS-PAGE

1. Put the gel system together. Fill the upper chamber with IEF cathode buffer and the lower chamber with IEF anode buffer.
2. Transfer the samples in the gel pockets (Subheading 3.7). One lane should be loaded with IEF-LB for pH measurements.
3. Electrophoresis is carried out over night. 12 h at 30 V, for 2 h at 100 V, 1.5 h at 250 V and 1 h at 300 V at 4 °C.
4. pH determination: Prepare 10× 2 mL tubes and fill them with 1 mL water. Slice out the free gel lane only loaded with IEF-LB and cut it in ten similar pieces (~0.6–0.7 cm). Vortex the tubes and keep them over night at 4 °C. Keep the tubes at room temperature for 1 h and vortex each tube for 2 min directly before measurement with a pH-electrode. Results are used to calculate the pH-function of the IEF-gel.
5. Transfer the rest of the gel to the according in-gel staining solution (Subheading 3.9) or continue with the second dimension.
6. Gel lane(s) for the second dimension is/are sliced out and incubated in equilibration buffer for 45 min on a shaker. The gel lane is transferred to the second dimension modSDS-PAGE.
7. Fill both chambers with electrophoresis buffer. Electrophoresis is carried out at 30 mA per gel up to 200 V at 4 °C until the blue front is running down.
8. When the electrophoresis is finished the gel is transferred into the according buffer (Subheading 3.9).

3.8.3 hrCNE and Second Dimension modSDS-PAGE

1. Put the gel system together. Fill the upper chamber with hrCNE cathode buffer and the lower chamber with hrCNE anode buffer.
2. Transfer the samples (Subheading 3.7) in the gel pockets. Add one marker lane.
3. Electrophoresis is carried out for 30 min at 100 V, 1 h up to 500 V constricted to 10 mA per gel at 4 °C.
4. When the electrophoresis is finished the gel system is disassembled. The part of the gel that should be used for the in-gel staining in the first dimension is sliced out and transferred into the according buffer (Subheading 3.9).
5. Gel lane(s) for the second dimension is/are sliced out and incubated in equilibration buffer for 45 min on a shaker. The gel lane is transferred to the second dimension modSDS-PAGE.

6. Fill both chambers with running gel buffer. Electrophoresis is carried out at 30 mA per gel up to 200 V at 4 °C until the blue front is running down.
7. When the electrophoresis is finished the gel is transferred into the according buffer (Subheading 3.9).

3.9 Differential In-Gel Staining for Peroxidases

3.9.1 TMB-Staining

The equilibration steps and stainings are accomplished on a horizontal shaker. All in-gel stainings are scanned before the bands become saturated (Figs. 3 and 4).

1. The gel is transferred into 50 mL TMB-staining solution and the box is covered in aluminum foil (TMB is light sensitive). Keep the gel at a horizontal shaker for ~60 min.
2. Reaction is started by the addition of 0.1 % H_2O_2 (180 μ L). After few seconds turquoise TMB bands appear. When the desired band intensity is reached, the reaction is stopped by the buffer exchanged against 30 % 2-propanol/70 % 250 mM sodium acetate buffer, pH 5.0.

3.9.2 DAB-Staining

1. The gel is equilibrated for 10 min in 50 mL 50 mM Na-acetate, 10 mM $CaCl_2$, pH 5.0.
2. 80 mg DAB are solved in 1 mL DMSO and added to 49 mL 50 mM Na-acetate, 10 mM $CaCl_2$, pH 5.0. The staining solution is mixed.
3. The equilibration solution is exchanged against the staining solution and the gel is equilibrated for 10 min in the solution.
4. The reaction is started by the addition of 0.5 % H_2O_2 (900 μ L). Brown bands appear after a few minutes. When the desired band intensity is reached, the gel is scanned.

3.9.3 α -Chloro-Naphthol-Staining

1. The gel is equilibrated in 50 mL 0.1 M Na_2HPO_4 buffer, pH 6.5 for 15 min.
2. The buffer is exchanged against α -chloro-naphthol-staining solution and again equilibrated for 10 min.
3. The reaction is started with 0.1 % H_2O_2 (180 μ L). Violet bands appear after a few minutes. When the desired band intensity is reached, the gel is scanned.

3.9.4 Guaiacol-Staining

1. The gel is transferred into 50 mL 50 mM Na-acetate, pH 5.0 with 10 mM $CaCl_2$ and equilibrated for 10 min.
2. 0.5 % Guaiacol (250 μ L) is added to the buffer. Equilibrate again 10 min.
3. 0.15 % H_2O_2 (250 μ L) is added to the staining solution. After few seconds the first orange bands become visible. When the desired band intensity is reached, the gel is scanned.

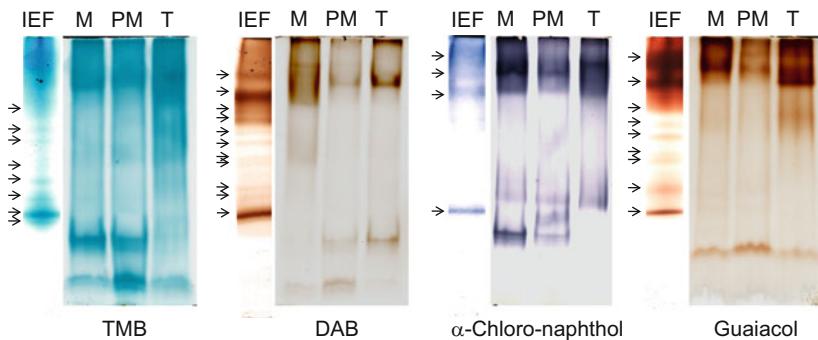


Fig. 3 In-gel staining of peroxidases after separation of membranes by native PAGE. Peroxidase activities were visualized by “specific” stains after separation of microsomal proteins (50 µg) by native IEF-PAGE (pH 3–10) or after separation of washed membrane fractions (50 µg protein) by hrCNE. TMB detect heme- and copper-containing proteins, whereas DAB and the phenolic substrates α -chloro-naphthol and guaiacol react with class III peroxidases. Time of development was 2–3 min for all stains. A shorter time of incubation with substrates may increase the resolution in the upper part of the TMB or guaiacol lanes and reveal additional protein bands in this range, but in that case protein bands with lower abundance or activity will disappear. *M* microsomal fraction, *PM* plasma membrane, *T* tonoplast

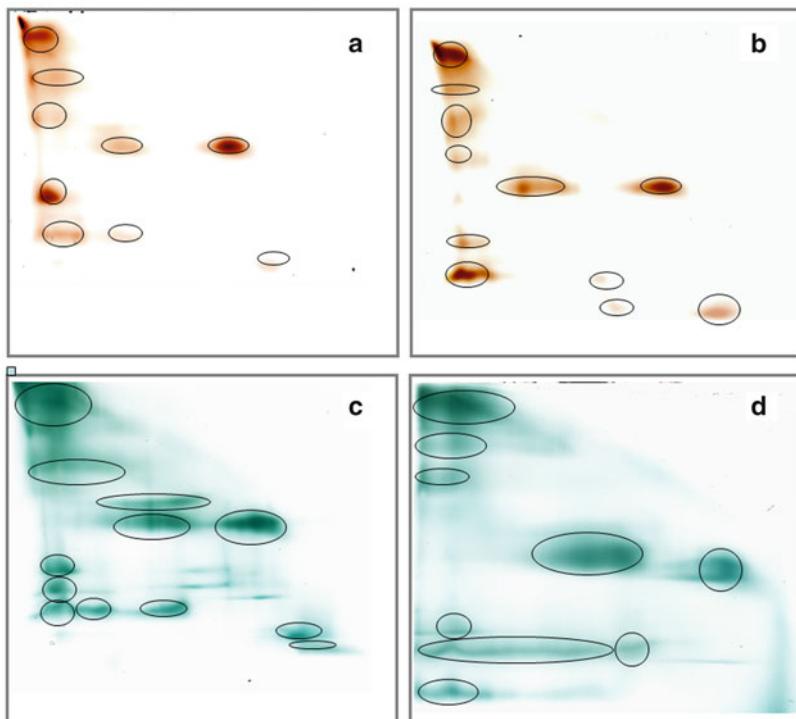


Fig. 4 2D-PAGE (hrCNE/modSDS-PAGE) of membrane fractions. Proteins (250 µg) of washed microsomal fractions (**a**, **c**) and plasma membranes (**b**, **d**) were separated by hrCNE in the first dimension and by modSDS-PAGE in the second dimension. Guaiacol (**a**, **b**) or TMB (**c**, **d**) were used to visualize peroxidases, heme and copper containing proteins. For an easier identification of spots on the gel images may be inverted

4 Notes

1. Alternatively membranes can be pelleted with a refrigerated centrifuge at $50,000 \times g$ for 90 min at 4 °C
2. APS can be prepared as 10 % stock solution and kept as small aliquots until use. Avoid multiple freeze-thawing.
3. Use concentrated HCl (12 N) and careful adjust the pH. Avoid pH equalization after exceeding pH 6.8 with base.
4. Prepare directly before use. Be sure that the TMB is well solved before the addition of the buffer.
5. If salt is added to fast, proteins can be denatured.
6. A salt gradient (e.g., 30 %, 60 %, 90 %) can be used for a further fractionation of the soluble proteins, if needed.
7. The volume depends on the size of the pellet and the used fresh weight.
8. For quantification of proteins a Bradford assay was used (optional BCA, 660 nm (Pierce) or other methods can be used).
9. First weight in the sucrose, KCl, phosphate buffer and water and solve the sucrose then add dextran and PEG and mix well. Weight precise, because all changes in the concentration have an effect on the purity and yield of the plasma membrane.
10. The lower phase is achieved by the addition of 9 g phase buffer to the phase mixture, mixing the solution. The phase system is kept at 4–8 °C over night. The upper phase is removed and the lower phase is used for phase partitioning.
11. If the gel is casted from the top, the higher concentrated solution is filled in the mixing chamber and the lower concentrated solution in the reservoir chamber of the gradient mixer.
12. The pH gradient can be varied by the use or combination of different ampholytes.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG Lu 668/4-4) and the University of Hamburg (Young Researcher Initiative grant to C.N. Meisrimler and PhD student grant to D. Hopff).

References

1. Duroux L, Welinder KG (2003) The peroxidase gene family in plants: a phylogenetic overview. *J Mol Evol* 57:397–407
2. Passardi F, Cosio C, Penel C et al (2005) Peroxidases have more functions than a Swiss army knife. *Plant Cell Rep* 24:255–265
3. Passardi F, Theiler G, Zamocky M et al (2007) PeroxiBase: the peroxidase database. *Phytochemistry* 68:1605–1611
4. Lüthje S, Meirimler CN, Hopff D et al (2011) Phylogeny, topology, structure and functions of membrane-bound class III peroxidases in vascular plants. *Phytochemistry* 72: 1124–1135
5. Almagro L, Gómez Ros LV, Belchi-Navarro S et al (2009) Class III peroxidases in plant defence reactions. *J Exp Bot* 60:377–390
6. Welinder KG, Justesen AF, Kjærsgård IVH et al (2002) Structural diversity and transcription of class III peroxidases from *Arabidopsis thaliana*. *Eur J Biochem* 269:6063–6081
7. De Gara L (2004) Class III peroxidases and ascorbate metabolism in plants. *Phytochem Rev* 3:195–205
8. Martín M, Casano LM, Zapata JM et al (2004) Role of thylakoid Ndh complex and peroxidase in the protection against photo-oxidative stress: fluorescence and enzyme activities in wild-type and *ndhF*-deficient tobacco. *Physiol Plant* 122:443–452
9. Mika A, Buck F, Lüthje S (2008) Membrane-bound class III peroxidases: identification, biochemical properties and sequence analysis of isoenzymes purified from maize (*Zea mays* L.) roots. *J Proteomics* 71:412–424
10. Costa MMR, Hilliou F, Durate P et al (2008) Molecular cloning and characterization of a vacuolar class III peroxidase involved in the metabolism of anticancer alkaloids in *Catharanthus roseus*. *Plant Physiol* 146:403–417
11. Albertsson PA (1971) Partition of cell particles and macromolecules, 2nd edn. Wiley, New York
12. Dolowy K (1984) Bioelectrochemistry of cell surfaces. *Prog Surface Sci* 15:245–368
13. Widell S, Lundborg T, Larsson C (1982) Plasma membranes from oats prepared by partition in an aqueous polymer two-phase system. *Plant Physiol* 70:1429–1435
14. Yoshida S, Uemura M, Niki T et al (1983) Partition of membrane particles in aqueous two-polymer phase system and its practical use for purification of plasma membranes from plants. *Plant Physiol* 72:105–114
15. Larsson C (1985) Plasma membranes. In: Linskens HF, Jackson JF (eds) Modern methods of plant analysis, cell components, new series, vol 1. Springer-Verlag, Berlin, pp 85–104
16. Yamazaki T, Kawamura Y, Uemura M (2009) Extracellular freezing-induced mechanical stress and surface area regulation on the plasma membrane in cold-acclimated plant cells. *Plant Signal Behav* 4:231–233
17. Minami A, Fujiwara M, Furuto A et al (2009) Alterations in detergent-resistant plasma membrane microdomains in *Arabidopsis thaliana* during cold acclimation. *Plant Cell Physiol* 50:341–359
18. Meirimler CN, Planchon S, Renaut J et al (2011) Alteration of plasma membrane-bound redox systems of iron deficient pea roots by chitosan. *J Proteomics* 74:437–449
19. Takahashi D, Kawamura Y, Yamashita T et al (2011) Detergent-resistant plasma membrane proteome in oat and rye: similarities and dissimilarities between two monocotyledoneous plants. *J Proteome Res* 11:1654–1665
20. Lüthje S, Van Gestelen P, Córdoba-Pedregosa MC et al (1998) Quinones in plant plasma membranes—a missing link? *Protoplasma* 205:43–51
21. Mika A, Boenisch MJ, Hopff D et al (2010) Membrane-bound guaiacol peroxidases from maize (*Zea mays* L.) roots are regulated by methyl jasmonate, salicylic acid, and pathogen elicitors. *J Exp Bot* 61:831–841
22. Bérczi A, Larsson C, Widell S et al (1989) On the presence of inside-out plasma membrane vesicles and vanadate-inhibited K^+/Mg^{2+} -ATPase in microsomal fractions from wheat and maize roots. *Physiol Plant* 77:12–19
23. Alexandersson E, Gustavsson N, Bernfur K et al (2008) Purification and proteomic analysis of plant plasma membranes. *Methods Mol Biol* 432:161–173
24. Aidemark M, Andersson CJ, Rasmussen AG et al (2009) Regulation of callose synthase activity in situ in alamethicin-permeabilized *Arabidopsis* and tobacco suspension cells. *BMC Plant Biol* 9:27
25. Bérczi A, Lüthje S, Asard H (2001) b-Type cytochromes in plasma membranes of *Phaseolus vulgaris* hypocotyls, *Arabidopsis thaliana* leaves, and *Zea mays* roots. *Protoplasma* 217:50–55
26. Ratajczak R, Hinz G, Robinson DG (1999) Localization of pyrophosphatase in membranes of cauliflower inflorescence cells. *Planta* 208:205–211
27. Oberbeck K, Drucker M, Robinson DG (1994) V-ATPase and pyrophosphatase in endomembranes of maize roots. *J Exp Bot* 45:235–244

28. Whiteman SA, Nühse TS, Ashford D et al (2008) A proteomic and phosphoproteomic analysis of *Oryza sativa* plasma membrane and vacuolar membrane. *Plant J* 56:146–156
29. Bérzci A, Asard H (2003) Soluble proteins, an often overlooked contaminant in plasma membrane preparations. *Trends Plant Sci* 8:250–251
30. Costa MMR, Hilliou F, Durate P et al (2008) Molecular cloning and characterization of a vacuolar class III peroxidase involved in the metabolism of anticancer alkaloids in *Catharanthus roseus*. *Plant Physiol* 146:403–417
31. Mika A, Lüthje S (2003) Properties of guaiacol peroxidase activities isolated from corn root plasma membranes. *Plant Physiol* 132:1489–1498
32. Kukavica BM, Veljović-Jovanović SD, Menckhoff L et al (2012) Cell wall bound cationic and anionic class III isoperoxidases of pea root—biochemical characterisation and function in root growth. *J Exp Bot* 63:4631–4645
33. Wittig I, Karas M, Schägger H (2007) High resolution clear native electrophoresis for in-gel functional analysis and fluorescence studies of membrane protein complexes. *Mol Cell Proteomics* 6:1215–1225
34. Thomas PE, Ryan D, Levin W (1976) An improved staining procedure for the detection of the peroxidase activity of cytochrome P-450 on sodium dodecyl sulfate polyacrylamide gels. *Anal Biochem* 75:168–176
35. Miller DJ, Nicholas DJD (1984) 3',5,5'-Tetramethylbenzidine/H₂O₂ staining is not specific for heme proteins separated by gel electrophoresis. *Anal Biochem* 140: 577–580
36. Herzog V, Fahimi HD (1973) A new sensitive colorimetric assay for peroxidase using 3,3'-diaminobenzidine as hydrogen donor. *Anal Biochem* 55:554–562
37. Guilbault GC, Kramer DN (1964) 4-Methoxy- α -naphthol as a spectrophotometric reagent substrate for measuring peroxidatic activity. *Anal Chem* 36:2494–2496
38. Bergmeyer HU, Gawehn K, Grassl M (1974) Enzymes as biochemical reagents. In: Bergmeyer HU (ed) *Methods of enzymatic analysis*, 2nd ed, vol 1. Academic, New York, NY, pp 495–496
39. Führs H, Götze S, Specht A et al (2009) Characterization of leaf apoplastic peroxidases and metabolites in *Vigna unguiculata* in response to toxic manganese supply and silicon. *J Exp Bot* 60:1663–1678
40. Lüthje S, Hopff D, Schmitt A et al (2009) Hunting for low abundant redox proteins in plant plasma membranes. *J Proteomics* 73: 475–483
41. Wilm M, Shevchenko A, Houthaeve T et al (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 379:466–469
42. Bradford MM (1976) A rapid and sensitive method for quantification of microgram quantities of protein dye binding. *Anal Biochem* 72:248–254

Part IX

Translational Proteomics

Chapter 49

Proteotyping of Holm Oak (*Quercus ilex* subsp. *ballota*) Provenances Through Proteomic Analysis of Acorn Flour

José Valero Galván, Raquel González Fernández, Luis Valledor, Rafael Ma. Navarro Cerrillo, and Jesus V. Jorrín-Novo

Abstract

Proteomics has become a powerful tool to characterize biodiversity and natural variability in plant species, as well as to catalogue and establish phylogenetic relationships and distances among populations, provenances or ecotypes. In this chapter, we describe the standard proteomics workflow that we currently use in cataloguing Holm oak (*Quercus ilex* subsp. *ballota* [Desf.] Samp.) populations. Proteins are extracted from acorn flour or pollen by TCA/acetone or TCA/acetone-phenol methods, resolved by one- or two-dimensional gel electrophoresis, and gel images are captured and analyzed by appropriate software and statistical packages. Quantitative or qualitative variable bands or spots are subjected to MS analysis in order to identify them and correlate differences in the protein profile with the phenotypes or environmental conditions.

Key words Holm oak proteomics, Plant biodiversity, Plant proteotyping

1 Introduction

Electrophoresis has proven to be the most important and effective tool in analyzing cellular protein profiles both from a quantitative and qualitative point of view. Moreover, it is one of the most convenient methods for characterizing, cataloguing, and establishing phylogenetic relationships and distances among populations, provenances, ecotypes, or genotypes [1–8]. The workflow of a standard 1- or 2-DE-based proteomics experiment includes the following steps: experimental design, sampling, protein extraction, protein separation, protein analysis by mass spectrometry (MS), statistical analysis of the data, and protein identification, using proper bioinformatics software and tools [9–11] (Fig. 1). The experimental design is the key step to extracting the maximum information from an experiment. A good experimental design considers the impact of different sources of variation and the minimum number of replicates to be made in the context of a particular minimum detectable

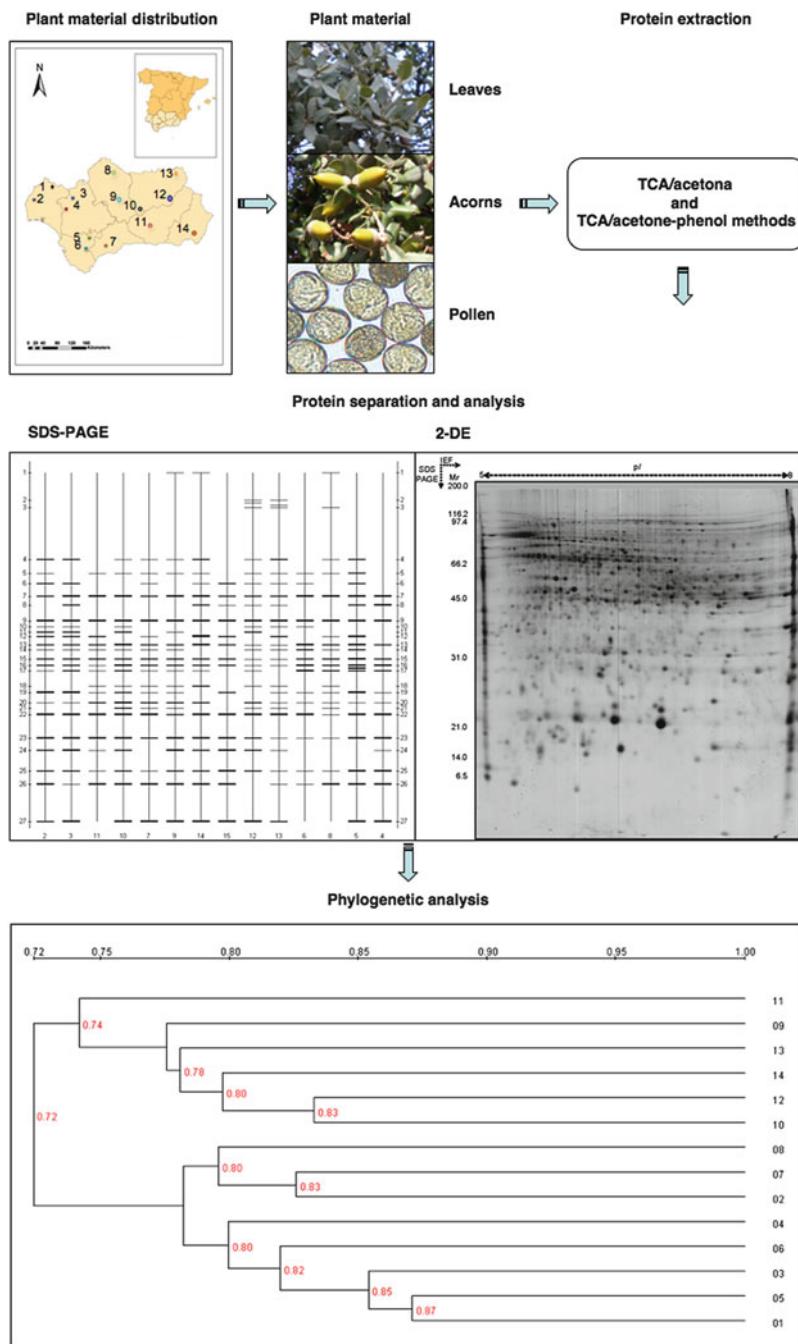


Fig. 1 Overview of the Holm oak proteotyping workflow

difference that one is seeking to achieve [12]. The most appropriate protocols for protein extraction should be developed for each specific species, and they must be optimized for the biological systems (i.e., plant species, organ, tissue, cells), as well as for the

research objectives [13]. In our experience, protein extraction using the TCA/acetone and TCA/acetone–phenol protocols provided the best results with a large variety of plant species [13–17]. Protein solubilization is a critical step. Detergents as CHAPS and chaotropic agents such as urea and thiourea must be used in the solubilization solutions to help in the hydrophobic protein solubilization, obtaining a higher protein yield [9, 16–19]. Protein separation with SDS-PAGE is a quick and accurate technique for plant proteins, especially in the case of comparative proteomics with large numbers of samples to be analyzed. Using appropriate software, SDS-PAGE is a simple, reliable technique for finger-printing crude extracts, and it is especially useful in the case of hydrophobic and low molecular-weight proteins [20]. Furthermore, SDS-PAGE is a good approach to obtain preliminary results before performing 2-DE analysis [15, 16]. Two-DE has been used for separating and displaying the components of large protein complexes, and it has been a reliable tool to study natural variability in several plants species [1, 3–5, 21, 22]. After staining (we currently use colloidal Coomassie staining [23]), images are digitized and analyzed with appropriate software [24]. Quantitative proteomics data are classically assessed by univariate statistics (*t*-test, Mann–Whitney, ANOVA, Kruskal–Wallis), but these methods increase the possibility of false positives, are negatively affected by the raw structure of proteomics data, and they cannot detect trends and protein relations [25–28]. On the other hand, the analysis employing multivariate approaches (i.e., Principal Components, Self Organizing Maps) are described to be more effective, because of its capacity to reduce the complexity of the data, predict trends and also for being less affected by data structure [26–28]. Furthermore, data analysis can be used to discriminate and establish phylogenetic relationships among populations and genotyping, as well as to correlate the profile with edaphoclimatic characteristics and morphometric parameters. Finally, major differential bands or spots are excised from gels and subjected, after tryptic digestion, to MS analysis and their identification [29, 30].

2 Materials

Mention of specific companies or pieces of equipment is not mandatory, and it does not represent an endorsement by the authors.

All the chemicals should be of analytical grade.

2.1 Reagents, Solutions and Buffers

2.1.1 Protein Extraction

1. Liquid Nitrogen.
2. Trichloroacetic acid (TCA) (10 % w/v)/acetone (80 % v/v) solution. Store at -20 °C and use directly from the freezer.
3. 0.1 M Ammonium acetate/methanol (100 % and 80 % v/v) solution. Store at -20 °C and use directly from the freezer.

4. Acetone (80 % v/v) solution. Store at -20 °C and use directly from the freezer.
5. Phenol solution equilibrated with 10 mM Tris-HCl, pH 8 (Sigma, Chemical Company). Store at 4 °C.
6. Sodium dodecyl sulfate (SDS) buffer: 0.1 M Tris-HCl, pH 8, 30 % (w/v) sucrose, 2 % (w/v) SDS, 5 % (v/v) β-mercaptoethanol. Store at 4 °C and temper prior to use. Add the β-mercaptoethanol just before use.
7. Solubilization solution: 9 M urea, 2 M thiourea, 4 % (w/v) 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS), 0.5 % (v/v) Tritón-X100, 20 mM dithiothreitol (DTT). Store in 1 mL aliquots at -20 °C. Add the DTT just before use.

2.1.2 Protein Quantification

1. Bradford solution (Sigma, Chemical Company) (*see Note 1*). Store at 4 °C.
2. Protein standard: bovine serum albumin (BSA) at a concentration of 1 mg/mL in distilled water is used as a stock solution (*see Note 2*). Store at -20 °C.

2.1.3 One-Dimensional Gel Electrophoresis

1. 1.5 M Tris-HCl, pH 8.8. Store at 4 °C.
2. 0.5 M Tris-HCl, pH 6.8. Store at 4 °C.
3. 30 % Acrylamide/Bisacrylamide solution, 37.5:1 (Bio-Rad). Store at 4 °C.
4. Sodium dodecyl sulfate solution (SDS): 10 % (w/v) in water. Store at room temperature.
5. Ammonium persulfate solution (APS): 10 % (w/v) in water. Store at -20 °C (*see Note 3*).
6. *N,N,N',N'*-Tetramethylethylenediamine (TEMED) (Sigma) (*see Note 4*). Store at room temperature.
7. Running buffer: 25 mM Tris-HCl, pH 8.3, 192 mM glycine, 0.1 % (w/v) SDS. Store at room temperature (*see Note 5*).
8. Laemmli buffer: 62.5 mM Tris-HCl, pH 6.8, 25 % (v/v) glycerol, 2 % (w/v) SDS, 0.01 % (w/v) bromophenol blue, 5 % (v/v) β-mercaptoethanol. Store at room temperature. Add the β-mercaptoethanol just before use.
9. Standard broad range molecular weight markers (Bio-Rad). Store at 4 °C.

2.1.4 Two-Dimensional Gel Electrophoresis

1. Immobilized pH gradient strips of pH range 5–10, and 17 cm length (Bio-Rad).
2. Rehydration solution: 7 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 0.5 % (v/v) anfolites, 20 mM DTT, and 0.01 % (w/v) bromophenol blue. Store in 1 mL aliquots at -20 °C. Add the DTT just before use.

3. SDS-PAGE reagents are the same as described in Subheading 2.1.3.
4. Equilibration buffer I: 50 mM Tris-HCl, pH 8.8, 6 M urea, 20 % (v/v) glycerol, 2 % (w/v) SDS, and 2 % (w/v) DTT. Store at 4 °C and temper prior to use. Add the DTT before use.
5. Equilibration buffer II: 50 mM Tris-HCl, pH 8.8, 6 M urea, 20 % (v/v) glycerol, 2 % (w/v) SDS, and 135 mM iodoacetamide. Store at 4 °C and temper prior to use. Add iodoacetamide before use.
1. Staining solution: Weigh 80 g ammonium sulfate, add 22.5 mL of 85 % phosphoric acid, and add 700 mL of water. Dissolve 1 g of Coomassie blue G-250 in 22 mL of water. Mix the two solutions, and then add 200 mL of methanol. Finally add water until 1,000 mL. Store at room temperature.
2. 0.1 M Tris-H₃PO₄. Store at 4 °C.
3. 25 % (v/v) methanol. Store at room temperature.
4. 20 % (p/v) ammonium sulfate. Store at room temperature.

2.1.5 Colloidal

Coomassie Blue G-250

Staining

2.1.6 Protein

Identification

Protein identification was carried out according to the protocols of the Proteomics Service of SCAI at the University of Córdoba.

1. Differential spots.
2. Porcine trypsin (Promega).
3. Desalting cartridges Zip Tips C18 (Agilent Technologies).
4. Peptide calibration standard (Bruker Daltonics), consisting of a combination of peptides that provides a good calibration across a typical mass range between 1,000 and 3,500 Da.
5. 0.1 % (v/v) trifluoroacetic acid (TFA).
6. Ammonium bicarbonate.
7. Acetonitrile (ACN).
8. Trifluoroacetic acid (TFA).
9. α -cyano hydroxycinnamic acid.
10. MALDI mass spectrometry calibration standards.
11. Matrix solution: 4.7 mg/mL α -cyano-4-hydroxycinnamic acid (Sigma) in 70 % (v/v) ACN.
12. AnchorChip MALDI target (Bruker Daltonics).
1. Airtight polyethylene bags.
2. A knife.
3. Blade mill (Moulinex AD56 42).
4. Microsieve (\varnothing : 15 cm, 1 mm).

2.2 Equipment

and Software

5. Mortar and pestle.
6. Vortex.
7. Ultrasonic homogenizer.
8. Microcentrifuge.
9. Disposable microcentrifuge tubes: 1.5 and 2 mL.
10. PROTEAN II Cell and Protean Dodeca Cell (Bio-Rad, Hercules, USA).
11. Protean IEF Cell system (Bio-Rad, Hercules, USA).
12. Micro-tubes mixer.
13. Gel shaker.
14. PowerPac 300 Power Supply (Bio-Rad, Hercules).
15. GS-800™ Calibrated Imaging Densitometer (Bio-Rad).
16. Quantity One® 1-D Analysis software (Bio-Rad Hercules).
17. PD-Quest software v8.1 (Bio-Rad).
18. ProGest digestion station (Genomics Solution).
19. MALDI plates.
20. Automatic ProMs station (Genomic Solution).
21. Resin C18 microcolumn (ZipTip, Millipore).
22. 4800 Proteomics Analyzer (Applied Biosystems).
23. Autoflex mass spectrometer (Bruker Daltonics).
24. MASCOT search engine (Matrix Science Ltd., London; <http://www.matrixscience.com>).

3 Methods

3.1 Plant Material Collection and Storage

Plant material will be collected from different provenances or genotypes. It is important to do a large sampling in order to have the greatest diversity of populations or genotypes possible. Undamaged, homogeneous mature acorns or pollen are collected from at least ten different trees for each population. Once harvested, plant material is put in airtight polyethylene bag, and then stored at 4 ± 1 °C during no more than 12 h.

3.1.1 Acorns

1. Immediately after arriving to the laboratory, and previously to protein extraction, a pool of 20 acorns per tree are scarified with a knife by making transversal and longitudinal cuts, permitting the rapid removal of the pericarp (*see Note 6*).
2. After being peeled out, their embryos (including cotyledons) are crushed in a blade mill (Moulinex AD56 42) until obtain a fine powder (flour).

3. The final powder is weighted and stored in a desiccator at 4 ± 1 °C until protein extraction.
4. For SDS-PAGE protein extraction, flour of ten independent samples per population are used; whereas for 2-DE protein extraction flour representing all studied accessions of each population are crushed together and its proteins are extracted.

3.1.2 Pollen

1. Immediately after arriving to the laboratory, pollen grains are isolated from freshly open flowers by shaking the anthers on a glass slide (see Note 6).
2. Flower debris is removed using a microsieve, and pollen is examined under a light microscope.
3. Pollen is either used immediately, or stored at -70 °C after freezing in liquid nitrogen (see Note 7).

3.2 Protein Extraction from Acorns by the TCA/Acetone Method

To extract protein from acorns, we propose to use the procedure suggested by Damerval et al. [14]. It is particularly suitable for the extraction of proteins from seeds, but also for leaves as reported by Jorge et al. [31] and Maldonado et al. [13]. The method described here has been optimized to leaves and seeds from *Q. ilex* subsp. *ballota* [15], although these procedures can be applied to plant proteomic analysis in general.

1. The powder (flour) (100 mg) is transferred into a 2 mL tube with 1 mL of a solution of 10 % (w/v) TCA/acetone with 0.07 % (w/v) DTT. Mix well using a micropestle and then by vortexing (see Notes 6 and 8).
2. Sonicate 3× 10 s (50 W, amplitude 60) at 4 °C.
3. The proteins are left to precipitate overnight and then centrifuged at $15,000 \times g$ at 4 °C for 15 min. Discard the supernatant.
4. The pellet is washed twice with 1 mL of a solution of acetone with 0.07 % (w/v) DTT, and then centrifuged at $15,000 \times g$ at 4 °C for 15 min. Discard the supernatant.
5. The pellet obtained is dried in the air in order to remove residual acetone (see Note 9).
6. Proteins are dissolved in a solubilization solution for 2 h, by shaking in a microtube mixer at 4 °C (see Note 10).
7. Proteins are quantified using the Bradford method [32]. Prepare the calibration curve using several dilutions of bovine serum albumin protein, containing concentration of 0, 1, 3, 5, 10, 15, and 20 μ L of BSA (1 mg/mL) into 1.5 mL tubes, and make all up to 500 μ L with distilled water. Add 500 μ L of Bradford reagent to each tube and mix well by vortexing gently for thorough mixing. Use 800 μ L of distilled water as control. Incubate the samples with the protein reagent at room

temperature for 30 min in darkness. Measure the absorbance at 595 nm of each sample and standards. Store the protein extracts at -20°C for further analysis.

3.3 Protein Extraction from Pollen by the TCA/Acetone–Phenol Method

The TCA/acetone–phenol protocol provided the best results in terms of spot focusing, resolved spots, spot intensity, unique spots detected, and reproducibility [13]. We propose to use it to extract proteins from leaves and pollen [16, 31], although these procedures can also be applied to plant proteomic analysis in general.

1. The pollen (100 mg) is transferred into a 2 mL tube with 1 mL of a solution of 10 % (w/v) TCA/acetone. Mix well using a micropestle and then by vortexing (*see Notes 6 and 8*).
2. Sonicate 3 \times 10 s (50 W, amplitude 60) at 4°C .
3. Fill the tube with the solution of 10 % (w/v) TCA/acetone. Mix well by vortexing and centrifuge at $16,000 \times g$ at 4°C for 5 min. Remove the supernatant by decanting.
4. Fill the tube with 0.1 M ammonium acetate in 80 % (v/v) methanol. Mix well by vortexing and centrifuge at $16,000 \times g$ at 4°C for 5 min. Discard the supernatant.
5. Fill the tube with a solution of 80 % (v/v) acetone. Mix well by vortexing and centrifuge at $16,000 \times g$ at 4°C for 5 min. Discard the supernatant.
6. Air-dry the pellet at room temperature to remove residual acetone (*see Note 9*).
7. Add 1.2 mL of 1:1 phenol (pH 8, Sigma)/SDS buffer. Mix well using a pipette and by vortexing. Incubate for 5 min on ice and centrifuge at $16,000 \times g$ for 5 min. Transfer the upper phenol phase into a new 1.5-mL tube (*see Note 11*).
8. Fill the tube with a solution of 0.1 M ammonium acetate in 100 % (v/v) methanol, mix well and complete the precipitation overnight at -20°C .
9. Centrifuge at $16,000 \times g$ at 4°C for 5 min and discard the supernatant (a white pellet should be visible).
10. Wash the pellet with 100 % (v/v) methanol and mix by vortexing. Centrifuge at $16,000 \times g$ at 4°C for 5 min and discard the supernatant.
11. Wash the pellet with 80 % (v/v) acetone and mix by vortexing. Centrifuge at $16,000 \times g$ at 4°C for 5 min and discard the supernatant.
12. Dry the pellet at room temperature.
13. Dissolve the proteins in the solubilization solution for 2 h, shaking in a microtube mixer at 4°C (*see Note 10*).
14. Procedures for quantify proteins are the same as described in Subheading 3.2 (**steps 7 and 8**). Store the protein extracts at -20°C for further analysis.

3.4 One-Dimensional Gel Electrophoresis

Prepare the 13 % polyacrylamide gel electrophoresis using a PROTEAN II cells (Bio-Rad, Hercules, USA) electrophoresis kit.

1. Prepare the resolving gel solution by mixing 27.08 mL of 30 % of acrylamide, 15.6 mL of Tris–HCl, pH 8.8, 0.625 mL of SDS, 18.8 mL of distilled water, 0.31 μ L of APS and 31.2 μ L of TEMED. Pour the solution into the gel cassette and cover completely the solution surface with isopropanol to obtain a flat layer on top of the resolving gel. Leave at room temperature until 20 min.
2. When polymerization is completed, prepare the stacking gel. Mix 1.63 mL of 30 % acrylamide, 2.5 mL of Tris–HCl, pH 6.8, 0.1 mL of SDS, 6.75 mL distilled water, 50 μ L APS and 10 μ L TEMED and gently stir to obtain a uniform solution. Pour the resolving gel and transfer the well-forming comb into this solution. Polymerize the gel for at least 30 min at room temperature to allow complete polymerization.
3. The comb is removed from the stacking gel and place the gel in the electrophoresis tank. The gel is covered with running buffer, and 70 μ L of sample is applied to the bottom of each well. The volume and protein concentration of the sample should be sufficient to give at least 50 μ g of each protein. Apply 10 μ L of the molecular weight standards to one or two wells, preferably in an asymmetric position.
4. Connect the wires to the power supply unit and apply 100 V until the blue dye front reaches the bottom of the gel. Disconnect the electrophoresis unit from the power supply, remove the lid and discard the running buffer. Remove the gel from plates with a spatula, discard the stacking gel, and wash the separated gel with distilled water to remove traces of running buffer.
5. Prepare the Coomassie brilliant blue G-250 staining 1 day before the staining process [23]. Place the gel in a tray containing 500 mL of staining solution. Incubate overnight the gel in the staining solution. Once the gel is stained, discard the staining solution and cover the gel with 0.1 M Tris–H₃PO₄. Then, shake for 1–3 min. Discard the solution and cover the gel with 25 % (v/v) of methanol. Then, shake for 1 min. Remove the methanol and wash the gel with 20 % (w/v) of ammonium sulfate for 24 h.
6. Images are digitized using a GS-800 Calibrated Densitometer (Bio-Rad, Hercules, USA) and analyzed with Quantity One software (Bio-Rad, Hercules, USA).
7. Excise the bands of interest from the gel, and place them in different tubes containing distilled water until their processing.

3.5 Two-Dimensional Electrophoresis

Preliminary 2-DE experiments can be carried out with the Mini-Protean 3 system (Bio-Rad, Hercules, USA), using 7 cm pH 3–10 linear gradient strips (Bio-Rad, Hercules, USA) and 13 % polyacrylamide gels, to examine the *pI* range where the proteins are concentrated. In our case, the most protein spots were located in a *pI* range between 5 and 8.

1. Prepare a mix with 300 µg of proteins in 250 mL of rehydration solution on 1.5 mL tube. Load the samples in each lane of the 17 cm strip holders.
2. Remove the protective cover from the surface of the IPG strips and slowly lower the IPG strip (gel slide down) onto the rehydration solution, without trapping air bubbles. Then cover the IPG strip with 1–2 mL of mineral oil and apply the plastic cover.
3. Apply a low voltage (50 V) during rehydration for 12 h at 20 °C for improving the entry of high molecular weight proteins [33].
4. After active rehydration, start isoelectric focusing at 20 °C using the following parameters 250 V for 2 min, followed by 150 min linear gradient from 250 to 10,000 V, and finally focus on up to 40,000 V at 10,000 Vh.
5. After IEF, the strips are immediately reduced and alkylated according to [33]. IPG strips are equilibrated in two steps. Firstly, it is performed with 2 % (w/v) DTT in equilibration buffer I for 10 min in agitation at room temperature; secondly, it is carried out with 2.5 % (w/v) iodoacetamide in equilibration buffer II for 10 min in agitation at room temperature.
6. The second dimension is performed on 13 % polyacrylamide gels using the Protean Dodeca Cell (Bio-Rad, Hercules, USA). The gels can be run at 150 constant volts until the dye reaches the bottom of the gel.
7. The gels are stained employing the colloidal Coomassie method [23]. Soak the gel in a tray containing 50 mL of staining solution. Incubate overnight the gel in the staining solution. Once the gel is stained, discard the staining solution and cover the gel with 0.1 M Tris–H₃PO₄. Then, shake during 1–3 min. Discard the solution and cover the gel with 25 % (v/v) of methanol. Then shake during 1 min. Remove the methanol and wash the gel with 20 % (w/v) of ammonium sulfate for 24 h.
8. Images are digitized using a GS-800 Calibrated Densitometer (Bio-Rad, Hercules, USA), and then analyzed with PD-Quest software v8.1 (Bio-Rad, Hercules, USA).

9. Select a good image as a reference with a clear and representative spot pattern and with a minimum distortion, and align each of the images to the chosen reference. Select between three prominent spots to manually assignment, and use the automatic vector tool to add additional vectors.
10. After automatic spot detection and matching, check manually the spots with edition tools for correct detection.
11. Set gel groups according to the experimental design and normalize spot volume intensity ratios for each spot.
12. List all the spots together with their normalized volume.

3.6 Phylogenetic and Statistical Analyses

Prior to statistical and phylogenetic analyses, the volume of pixels for each band or spot is normalized according to the total volume of bands detected (SDS-PAGE) or to the total volume of valid spots in each gel (2-DE), respectively. Then, they are log-transformed, following the recommendations described by Valledor and Jorrin [34].

A multivariate analysis is carried out on two steps: firstly, hierarchical clustering is performed to check the entire dataset, and the results are indicated in dendograms using the cluster function of the software used; secondly, the dataset is analyzed by the use of Principal Component Analysis (PCA). The settings used for the PCA analysis are: co-variance matrix type, three principal components, onefold change, and 0.4 correlation threshold for clusters. PCA results are shown as a biplot.

Since the employed statistical methods tend to classify population together, this information is necessary for studying the possible correlation between distances and geographical and climate parameters.

3.7 Protein Identification

Spots are manually excised with a scalpel. Protein identification was carried out according to the Proteomics Service protocols of the University of Córdoba. Gel plugs are digested with modified porcine trypsin (Sequencing grade; Promega), by using an automatic ProGest digestion station (Genomics Solution). The conditions are two detailed steps for 30 min with 200 mM ammonium bicarbonate in 40 % (v/v) ACN at 37 °C; twice washed with 25 mM ammonium bicarbonate for 5 min and 25 mM ammonium bicarbonate in 50 % (v/v) ACN for 15 min respectively; dehydration with 100 % (v/v) ACN for 5 min and sample dried; hydration using 10 µL trypsin in a solution of 25 mM ammonium bicarbonate at a final concentration of 12.5 ng/µL for 10 min a room temperature, and the digestion is proceeded at 37 °C for 12 h. Subsequently, digestion is stopped by adding 10 µL of a solution of 0.5 % TFA in water.

Tryptic peptides are purified in an automatic ProMS station (Genomic Solutions) by using a resin C18 microcolumn (ZipTip, Millipore), and they are eluted directly with a matrix solution (α -cyano hydroxycinnamic acid at a concentration of 5 mg/mL in 70 % (v/v) ACN/0.1 % (v/v) TFA) on MALDI plaque in 1 μ L of final volume. After the cocrystallization on plaque, samples are analyzed by MALDI-TOF/TOF mass spectrometry to obtain the peptide mass fingerprinting (MS) in a 4800 Proteomics Analyzer (Applied Biosystems). The settings are: 800–4,000 m/z range, with an accelerating voltage of 20 kV, in reflection mode, with delayed extraction set to “on”, and an elapsed time of 120 ns. Spectra are internally calibrated with peptides from trypsin autolysis ($M + H^+ = 842.509$, $M + H^+ = 2,211.104$) with an m/z precision of ± 20 ppm. Most abundant peptide ions are subjected to MS/MS analysis, providing information that can be used to define the peptide sequence.

A combined search (PMF and MS/MS) is performed with GPS ExplorerTM software v3.5 (Applied Biosystems) over nonredundant NCBI databases using the MASCOT search engine (Matrix Science Ltd., London; <http://www.matrixscience.com>). The database search utilized the following parameters: taxonomy restrictions to *Viridiplantae*, one missed cleavage sites, 100 ppm mass tolerance in MS and 0.5 Da for MS/MS data, cysteine carbamidomethylation as a fixed modification, and methionine oxidation as a variable modification. The confidence in the peptide mass fingerprinting matches ($p < 0.05$) is based on the MOWSE score, and confirmed by the accurate overlapping of the matched peptides with the major peaks of the mass spectrum.

4 Final Remarks

Taking into account our own experience, in this chapter we have shown the usefulness of a basic proteomic approach—based on SDS-PAGE and 2-DE analyses of protein extracts from mature seeds and pollen—for variability studies in Holm oak. These analyses have allowed the separation and grouping of the populations according to its acorn morphometry, location (northern and southern), and climate conditions (xeric, mesic, and intermediate), confirming previous results of our research group by using acorn morphometry and NIRS chemical composition [15, 16, 35]. Therefore, gel-based Proteomics (SDS-PAGE and 2-DE) can be used to detect variability between different populations from different environments, and to correlate it to environmental conditions (Fig. 1).

5 Notes

1. Check the list of compatible chemicals and potential interfering chemicals typically found in the protein extraction buffer.
2. Do not freeze stock vial solutions more than once.
3. Prepare this solution on fresh each time.
4. TEMED accelerates the decomposition of APS molecules into sulfate free radicals and these, in turn, initiate the polymerization.
5. Prepare running buffer as 10× native buffer (0.25 M, Tris, 1.92 M glycine, and 1 % (w/v) SDS). Weigh 30.3 g of Tris, 144 g of glycine, and 10 g of SDS; mix and make it to 1 L with water. The pH of this solution should not be adjusted, and stored at room.
6. It is necessary to wear gloves and lab coat for all the procedures.
7. Be careful with liquid nitrogen due to its cool temperature (-195.8°C). It could cause severe frostbite. Pay attention to laboratory safety regulation.
8. Be particularly careful when handling the reagents TCA and phenol (consult safety data sheets) because they are corrosive. Use the fume hood with volatile reagents.
9. Be careful in not throwing out the pellet.
10. The volume of the solubilization solution added will depend on the quantity of precipitated proteins. It is advisable that samples be well concentrated.
11. There are three phases, namely: the upper phase (which is the phenolic phase where the proteins are), a white interphase, and a lower aqueous phase. Try to not to take part of the white interphase.

Acknowledgments

Jose Valero was recipient of an Alban Program fellowship (I06D00010MX). This work was supported by the Spanish Ministry of Science and Innovation cofinanced by the European Community FEDER funds: CGL2008-04503-C03-01/BOS, AGL2002-00530, and AGL2009-12243-C02-02; the Regional Government of Andalusia (Junta de Andalucía; the University of Córdoba (AGR-0164: Agricultural and Plant Biochemistry and Proteomics Research Group)); and the Autonomous University of Juarez City.

References

- Bahrman N, Petit RJ (1995) Genetic polymorphism in maritime pine (*Pinus pinaster* Ait.) assessed by two-dimensional gel electrophoresis of needle, bud, and pollen proteins. *J Mol Evol* 41:231–237
- Basha SM (1979) Identification of cultivar differences in seed polypeptide composition of peanuts (*Arachis hypogaea* L.) by two-dimensional polyacrylamide gel electrophoresis. *Plant Physiol* 63:301–306
- Chevalier F, Martin O, Rofidal V et al (2004) Proteomic investigation of natural variation between *Arabidopsis* ecotypes. *Proteomics* 4:1372–1381
- Emre I (2009) Electrophoretic analysis of some *Lathyrus* L. species based on seed storage proteins. *Genet Resour Crop Evol* 56:31–38
- Jacobsen S, Nesic L, Petersen M et al (2001) Classification of wheat varieties: use of two-dimensional gel electrophoresis for varieties that can not be classified by matrix assisted laser desorption/ionization-time of flight-mass spectrometry and an artificial neural network. *Electrophoresis* 22:1242–1245
- Ostergaard O, Melchior S, Roepstorff P et al (2002) Initial proteome analysis of mature barley seeds and malt. *Proteomics* 2:733–739
- Singh N, Matta N (2008) Variation studies on seed storage proteins and phylogenetics of the genus *Cucumis*. *Plant Syst Evol* 275:209–218
- Yüzbaşıoğlu E, Açık L, Özcan S (2008) Seed protein diversity among lentil cultivars. *Biol Plantarum* 52:126–128
- Cánovas FM, Dumas-Gaudot E, Recorbet G et al (2004) Plant proteome analysis. *Proteomics* 4:285–298
- Jorrín JV, Maldonado AM, Castillejo MA (2007) Plant proteome analysis: a 2006 update. *Proteomics* 7:2947–2962
- Jorrín-Novo JV, Maldonado AM, Echevarría-Zomeño S et al (2009) Plant proteomics update (2007–2008): second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. *J Proteomics* 72:285–314
- Hunt SM, Thomas MR, Sebastian LT et al (2005) Optimal replication and the importance of experimental design for gel-based quantitative proteomics. *J Proteome Res* 4:809–819
- Maldonado AM, Echevarría-Zomeño S, Jean-Baptiste S et al (2008) Evaluation of three different protocols of protein extraction for *Arabidopsis thaliana* leaf proteome analysis by two-dimensional electrophoresis. *J Proteomics* 71:461–472
- Damerval C, De Vienne D, Zivy M et al (1986) Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. *Electrophoresis* 7:52–54
- Valero GJ, Valledor L, Navarro CRM et al (2011) Studies of variability in Holm oak (*Quercus ilex* subsp. *ballota* [Desf.] Samp.) through acorn protein profile analysis. *J Proteomics* 74:1244–1255
- Valero GJ, Valledor L, Fernández GR et al (2012) Proteomic analysis of Holm oak (*Quercus ilex* subsp. *ballota* [Desf.] Samp.) pollen. *J Proteomics* 75:2736–2744
- Wang W, Vignani R, Scali M et al (2006) A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. *Electrophoresis* 27:2782–2786
- Jorrín JV, Maldonado AM, Castillejo MA (2007) Plant proteome analysis: a 2006 update. *Proteomics* 7:2947–2962
- Jorrín-Novo JV, Maldonado AM, Echevarría-Zomeño S et al (2009) Plant proteomics update (2007–2008): second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. *J Proteomics* 72:285–314
- Supek F, Peharec P, Krsnik-Rasol M et al (2008) Enhanced analytical power of SDS-PAGE using machine learning algorithms. *Proteomics* 8:28–31
- Petit RJ, Bahrman N, Baradat PH (1995) Comparison of genetic differentiation in maritime pine (*Pinus pinaster* Ait.) estimated using isozyme, total protein and terpenic loci. *Heredity* 75:382–389
- Vienne DD, Burstin J, Gerber S et al (1996) Two-dimensional electrophoresis of proteins as a source of monogenic and codominant markers for population genetics and mapping the expressed genome. *Heredity* 76:166–177
- Mathesius U, Keijzers G, Natera SH et al (2001) Establishment of a root proteome reference map for the model legume *Medicago truncatula* using the expressed sequence tag database for peptide mass fingerprinting. *Proteomics* 1:1424–1440

24. Matthias B, Michael MF, Kolbe M et al (2007) The state of the art in the analysis of two-dimensional gel electrophoresis images. *Appl Microbiol Biotechnol* 76:1223–1243
25. Chich JF, David O, Villers F (2007) Statistics for proteomics: experimental design and 2-DE differential analysis. *J Chromatogr B* 849:261–272
26. Rodríguez-Piñeiro AM, Rodríguez-Berrocal FJ, Páez de la Cadena M (2007) Improvements in the search for potential biomarkers by proteomics: application of principal component and discriminant analyses for two-dimensional maps evaluation. *J Chromatogr* 849:251–260
27. Grove H, Jørgensen BM, Jessen F (2008) Combination of statistical approaches for analysis of 2-DE data gives complementary results. *J Proteome Res* 7:5119–5124
28. Valledor L, Jorrín J (2011) Maximizing the information obtained by quantitative two dimensional gel electrophoresis analyses by an appropriate experimental design and statistical analyses. *J Proteomics* 74:1–18
29. Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. *Science* 312:212–217
30. Han X, Aslanian A, Yates IJR (2008) Mass spectrometry for proteomics. *Curr Opin Chem Biol* 12:483–490
31. Jorge I, Navarro RM, Lenz C et al (2005) The holm oak leaf proteome: analytical and biological variability in the protein expression level assessed by 2-DE and protein identification tandem mass spectrometry de novo sequencing and sequence similarity searching. *Proteomics* 5:222–234
32. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254
33. Gorg A, Postel W, Baumer M et al (1992) Two-dimensional polyacrylamide gel electrophoresis, with immobilized pH gradients in the first dimension, of barley seed proteins: discrimination of cultivars with different malting grades. *Electrophoresis* 13:192–203
34. Valledor L, Jorrín J (2010) Back to the basics: maximizing the information obtained by quantitative two dimensional gel electrophoresis analyses by an appropriate experimental design and statistical analyses. *J Proteomics* 74:1–18
35. Valero GJ, Jorrín-Novo J, Gómez CA et al (2011) Population variability based on the morphometry and chemical composition of the acorn in Holm oak (*Quercus ilex* subsp. *ballota* [Desf.] Samp.). *Eur J Forest Res* 131:893–904

Chapter 50

Immunoproteomics Analysis of Food Allergens

Rika Nakamura and Reiko Teshima

Abstract

Immunoblot-coupled proteomics based on two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), namely, immunoproteomics, has been used for comprehensive identification of food allergens, because it is a simple and inexpensive tool for rapid identification of several IgE-binding proteins. In this section, we describe our protocols for identification of food allergens using immunoproteomics and discuss a few technical points in detail.

Key words Food allergy, Allergen, IgE, Immunoproteomics, Identification, 2D-PAGE, MS/MS

1 Introduction

Some plants may trigger hypersensitivity reactions when patients with allergic diathesis eat parts of these plants or inhale their pollen. Most allergens in plants are proteinases, seed storage proteins, or enzyme (e. g., trypsin/amylase) inhibitor proteins. For identification of novel allergens, those allergens should have been extracted from food, purified and isolated by chromatography, such as ion-exchange and gel filtration, and their amino acid sequences decoded [1–3]. These methods involve complex steps, such as purifying the allergens from the crude extracts, and confirmation, at every step, by enzyme-linked immunosorbent assay (ELISA) or immunoblot analysis with the sera of the patients with food allergies, that the protein fraction has the capacity to bind IgE. Moreover, conventional methods to identify multiple allergens in foods are time-consuming, because only one allergen can be identified in a series of experiments. Proteomics has, therefore, been used for comprehensive identification of food allergens [4].

Immunoproteomics is a method coupling immunohistochemical analysis and proteomics based on 2-dimensional polyacrylamide gel electrophoresis (2D-PAGE). For identification of allergens, immunoglobulin E (IgE)-binding protein spots are detected on a 2D map of the food extract, the spots are excised from the 2D gel,

Table 1
Summary of the reports of 2D-immunoblot analysis of plant allergens

| Food plant | Identified allergen names | Authors | Year | Reference no. |
|------------|--|----------------|------|---------------|
| Celery | Api g 1 | Vallier et al. | 1992 | [5] |
| Hazelnuts | Cor a 9 (11S globulin) | Beyer et al. | 2002 | [6] |
| Sesame | Ses i 3 (7S vicilin-type globulin) Ses i 2 (2S globulin) Embryonic abundant protein Seed maturation protein | Beyer et al. | 2002 | [7] |
| Lupin | Conglutin γ 11S globulin basic subunits | Magni et al. | 2005 | [8] |
| Peanuts | Ara h 1 Ara h 3 Ara h 3/4 Ara h 4 Gly 1 iso-Ara h 3 | Boldt et al. | 2005 | [9] |
| Peanuts | Ara h 3 (basic subunit) | Restani et al. | 2005 | [10] |
| Apple | Mal d 1 Mal d 2 Mal d 3 Mal d 4 45 kDa basic protein | Herndl et al. | 2007 | [11] |
| Soybean | Late embryogenesis abundant protein | Batista et al. | 2007 | [12] |

followed by digestion with enzymes such as trypsin, and the allergens are identified by N-terminal amino acid sequencing or mass spectrometry/mass spectrometry (MS/MS) homology search. Immunoproteomics is simple and inexpensive method that allows rapid identification of several IgE-binding proteins at once. Since several IgE-binding proteins might be separated from crude extract using 2D-PAGE, minimum 5 short days would be needed to identify them in a food extract. In addition, protocols for 2D-PAGE have almost been optimized, whereas the protocols for allergen isolation need to be adjusted depending on the allergens. Therefore, the number of reports of identification of novel allergens and/or allergen location on 2D-map using immunoproteomics has been increasing recently [5–24] (Table 1).

In this section, we outline our protocols for the identification of food allergens using immunoproteomics and discuss a few technical points in detail. Of course, our protocols can be applied to 2D-western blotting to detect the target proteins using immunized animal antibodies [25, 26], and it is important to use optimal buffers for extracting allergens from food materials. The basic principle

is the same as that of 1D-immunoblotting, but double-staining of all protein spots and IgE-binding spots is adapted for precise localization of the IgE-binding spots. From our experience, 14 cm × 14 cm 2D-gels seems to be the optimal size for handling and obtaining good performance [11, 12, 18, 24].

2 Materials

Prepare all solutions using ultrapure water and analytical grade reagents.

Protect your eyes and hands when handling human serum, to prevent potential transmission of infections. Wearing gloves throughout the experiments is recommended for avoiding protein contamination.

2.1 Sample Preparation

1. Extraction buffer: 0.1 M phosphate-buffered saline (PBS, pH 7.2), 0.1 M Tris-HCl, pH 7.0, 1 M NaCl is frequently used (*see Note 1*).
2. 2-D Clean-Up Kit (GE Healthcare UK Ltd., Little Chalfont, UK): Store Wash buffer in the kit should be stored at -30 °C.
3. DeStreak Rehydration Solution (GE Healthcare): Before use, equilibrate at RT and dissolve urea crystals, and then add 0.5 % of IPG buffer and mix well.
4. Cellulose acetate 0.45 µm filter (Dismic-45, Advantec Toyo Kaisha Ltd., Tokyo, Japan).
5. Pierce BCA Assay Kit (Thermo Fisher Scientific K.K., Yokohama, Japan) or 2D-Quant Kit (GE Healthcare).

2.2 Isoelectro-phoresis and Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

1. Immobiline DryStrip pH3-10NL, 13 cm (GE Healthcare).
2. DryStrip holder, 13 cm (GE Healthcare).
3. IPG buffer pH3-10NL (GE Healthcare).
4. IPG phor III (GE Healthcare).
5. Cover oil fluid (GE Healthcare).
6. Equilibration buffer: 50 mM Tris-HCl, pH 8.8, 6 M urea, 30 % (v/v) glycerol, 2 % (w/v) SDS, and a small amount of bromophenol blue (BPB). Weigh 72.07 g of urea and 4 g of SDS and add 10 mL of 1 M Tris-HCl (pH 8.8), 69 mL of glycerol, and a small amount of BPB. Bring up the volume up to 200 mL and stir at room temperature (RT). Divide into aliquots and store at -20 °C.
7. Dithiothreitol: Just before use, add 100 mg (1 %) to 10 mL of equilibration buffer.
8. Iodoacetamide: Just before use, add 250 mg (2.5 %) to 10 mL of equilibration buffer.

9. Precast gels: Perfect NT gel S 10–20 %, 2-D well, 14 cm × 14 cm (D.R.C. Coop., Tokyo, Japan).
10. Electrophoresis system.
11. Agarose solution: 0.5 % low-melting agarose. Weigh 0.5 g of low-melting agarose and add 100 mL of SDS electrophoresis buffer. Warm to just before bubbling in a microwave. Divide 1 mL of agarose solution into microtubes and store at RT.
12. SDS electrophoresis buffer: 25 mM Tris, 192 mM glycine, 0.1 % (w/v) SDS. Dissolve 4.5 g of Tris base, 21.6 g of glycine, and 1.5 g of SDS in ultrapure water and top up the volume up to 1.5 L. In case of frequent use, make a tenfold concentrated stock solution and store at RT.
13. ECL Plex Fluorescent Rainbow Markers (GE Healthcare).

2.3 Double-Stained Immunoblotting

1. Tank transfer system.
2. Blotting buffer: 25 mM Tris, 192 mM glycine. Dissolve 15 g of Tris base and 72 g of glycine in ultrapure water and top up the volume to 5 L. In case of frequent use, make a tenfold concentrated stock solution and store at RT.
3. Container: The area of the bottom and volume should be more than 14 cm × 14 cm and 250 mL, respectively.
4. Immun-Blot PVDF (Bio-Rad Laboratories, Hercules, CA, USA): Cut a sheet into 14 cm × 14 cm. Incubate in methanol for 5 min just before use.
5. Filter paper: 20 cm × 20 cm, for transfer grade.
6. Cy5 Mono reactive dye (GE Healthcare): Dissolve 1 vial with 100 µL of dimethylformamide, vortex, and add to 100 mL of PBS just before use.
7. Wash buffer: 0.05 % Tween-20/PBS. Add 500 µL of Tween-20 to 1 L of PBS, mix gently.
8. Blocking buffer: 0.5 % casein/PBS. Dissolve 1 g of casein to 200 mL of PBS. Heat the solution at 60 °C for 1 h, because casein dissolves with difficulty. After cooling to RT, store at 4 °C.
9. Horseradish peroxidase (HRP)-linked anti-human IgE (Nordic Immunology, Tilburg, Netherlands): Dilute 1:1,000 with 0.1 % casein/PBS just before use.
10. ECL Plus (Plex) Western Blotting Detection Reagents (GE Healthcare): available for detection by a fluorescence imager.
11. Image scanner: We use Typhoon 9400 from GE Healthcare.
12. Low-fluorescence glass plate (two plates): Large enough to cover a 14 cm × 14 cm membrane.

2.4 In Gel Digestion of Protein Spots

1. Quick-CBB plus (Wako Pure Chemical Industries, Ltd., Osaka, Japan) or Silver Stain Kit for Mass Spectrometry (Thermo Fisher Scientific).
2. 1 M Ammonium bicarbonate: Weigh 67.9 mg and add 1 mL of ultrapure water. Prepare before every use.
3. Dehydration buffer: 50 % acetonitrile, 25 mM ammonium bicarbonate. Mix 1 mL of acetonitrile, 50 μ L of 1 M ammonium bicarbonate, and 950 μ L of ultrapure water.
4. Trypsin Gold (Promega): Dissolve in 100 μ L of ultrapure water. Divide aliquots and store at -80°C .
5. Protease Max (Promega): Divide 10 μ L into each tube and store at -80°C . Before use, add 990 μ L of 25 mM ammonium bicarbonate to make a 0.1 % solution.

2.5 Mass Spectrometry and Protein Identification

1. Zip-Tip C18 (Millipore Corp., MA, USA).
2. α -Cyano-4-hydroxycinnamic acid (α -CHCA, Sigma-Aldrich Corp., MO, USA): Add 50 % acetonitrile/0.1 % trifluoroacetic acid (TFA) to prepare a saturated solution (*see Note 2*). Dilute saturated α -CHCA to 1:4 with 50 % acetonitrile/0.1 % TFA just before use.
3. Opti-TOF 384 well Insert (123 mm \times 81 mm) (Life Technologies Corp., CA, USA).
4. 4800 MALDI TOF/TOF Analyzer (Life Technologies).
5. Mascot software (Matrix Science Inc., MA, USA).

3 Methods

3.1 Protein Sample Preparation

1. Extract allergens from the plant specimens. Homogenize the plant specimens and add an adequate volume of extraction buffer. Rotate for more than 3 h to overnight at 4°C and centrifuge at $8,000 \times g$ for 10 min. Filter the supernatants using a 0.45 μm filter.
2. Measure the protein concentration of the extracts with a Pierce BCA Assay Kit or 2-D Quant Kit, in accordance with the manufacturer's protocols. Divide into aliquots and store at -80°C .
3. Purify and desalt an amount of extract containing 50 μg proteins with a 2-D Clean-Up Kit (*see Note 3*).
4. Dissolve the protein pellet by vortexing in a 500 μL of DeStreak Rehydration Solution to obtain a protein concentration of 0.1 $\mu\text{g}/\mu\text{L}$. Divide the protein solutions into 25 μg aliquots and store at -80°C .

3.2 Isoelectric Focusing and SDS-PAGE (2D-PAGE)

1. Pipette 250 μ L of the sample solution on a 13-cm strip holder and place an Immobiline DryStrip (pH3-10NL, 13 cm) on it, taking care to avoid formation of large air bubbles (*see Note 4*).
2. Overlay the strip with 800 μ L of Cover Oil Fluid (*see Note 5*).
3. Set and run an isoelectric focusing protocol as follows: Dehydration, 12 h; Step 1 (step-n-hold), 500 V for 4 h; Step 2 (gradient), 1,000 V for 1 h; Step 3 (gradient), 8,000 V for 2.5 h; Step 4 (step-n-hold), 8,000 V for 1.5 h; current limit 50 μ A/strip, run under 20 $^{\circ}$ C.
4. When the protocol is completed, pick the holders out. The total value of V-hrs will be 22,000 or more (*see Note 6*).
5. Pick the strip up from the holder and gently remove the oil, pushing it to a paper towel, then place the strip into a 60-mm dish so as to have the gel side face the inside of the dish. If not performing the next SDS-PAGE immediately, keep the strips stored at -80 $^{\circ}$ C until use.
6. Heat an Agarose solution tube at 100 $^{\circ}$ C on a heat-block.
7. If the strips had been kept in a freezer, place them at room temperature until they return to RT (*see Note 7*).
8. Add 10 mL of SDS-equilibration buffer containing 1 % DTT to the dish and gently shake for 15 min at RT.
9. Discard the buffer and replace with 10 mL of SDS-equilibration buffer containing 2.5 % iodoacetamide, and gently shake for 15 min at RT.
10. Pick the strip up and place it on the slot of an acrylamide gel plate with the acidic ends to the left (*see Note 8*).
11. Pipette slowly 1 mL of agarose solution on the strip without creating any air bubbles between the strip and the gel. Pipette 10 μ L of fluorescence rainbow marker into the marker slot of the gel before the agarose is set, and allow the gels to stand for 1–2 min until the agarose sets.
12. Fill an electrophoresis tank with 1,200 mL of SDS electrophoresis buffer. Avoid formation of bubbles around the inner tank.
13. Run the SDS-PAGE protocol at a constant 220 V for 3 h.
14. Stop the electrophoresis when the BPB line reaches ~5 mm from the bottom. After the electrophoresis, wash the gel in a container filled with ultrapure water, followed by immersion in blotting buffer (*see Note 9*).

3.3 Double-Stained Immunoblotting

1. Incubate a 14 cm \times 14 cm PVDF membrane in methanol at RT for 5 min. Sandwich the gel between a filter paper and PVDF membrane and place another filter paper on the membrane. Roll a cylinder glass tube on the surface of the membrane placed on the gel to remove any air bubbles. Fix them in a cassette,

and place the cassette in a transfer tank. Pour 5 L of transfer buffer into the tank, and transfer at 180 mA overnight under cooling (*see Note 10*).

2. After blotting, the membrane is picked up and incubated with Cy5/PBS solution to label the transferred proteins with Cy5 (*see Note 11*). Gently shake for 1 h at RT, avoiding exposure to light.
3. Wash the membrane with 100 mL of methanol. Gently shake for 10 min at RT and repeat two times with fresh methanol.
4. Wash the membrane with 100 mL of 0.05 % Tween-20/PBS. Gently shake for 10 min at RT and repeat two times with fresh solution.
5. Block the membrane with 100 mL of 0.5 % casein/PBS. Gently shake for 2 h at RT.
6. Pipette diluted serum (*see Note 12*) on the transferred side of the membrane. Incubate the membrane for 1 h at RT and then overnight at 4 °C under gentle shaking.
7. Wash the membrane with 100 mL of 0.05 % Tween-20/PBS (*see Note 13*). Gently shake for 10 min at RT and repeat two times with fresh solution.
8. Incubate the membrane in HRP-linked anti-human IgE (1:1,000 diluted with 0.1 % casein/PBS) solution for 1.5 h at RT under gentle shaking.
9. Wash the membrane with 100 mL of 0.05 % Tween-20/PBS. Gently shake for 10 min at RT and repeat two times with fresh solution.
10. Pipette ECL plus (plex) reagent on to the membrane, and after incubation for an appropriate period of time, move the membrane into ultrapure water (*see Note 14*).
11. Place the membrane on a clean, low-fluorescence glass plate with the stained side facing down. Gently pour a few milliliters of ultrapure water on the surface of the membrane and cover it with another clean glass plate (Fig. 1, *see Note 15*).
12. Place the gel plate on the scanning table of the image scanner. Scan gel images of Cy2 (458 nm, band-pass 40) for ECL plus (plex), and Cy5 (670 nm, band-pass 30) (Fig. 2).

3.4 In Gel Digestion of Protein Spots

1. Prepare a gel separating 100 µg of proteins, stain with Coomassie brilliant blue or silver stain.
2. Excise gel pieces containing IgE-binding proteins from the 2D-gel and place them, one each in individual 1.5-mL centrifuge tubes.
3. The gel pieces are then decolorized with 50 µL of decolorizing buffer each. Replace fresh buffer and repeat washing the gel pieces until they are no longer stained.

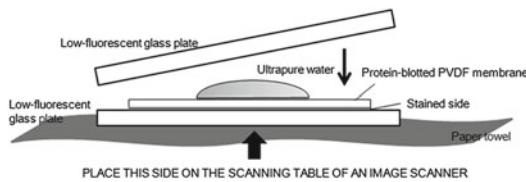


Fig. 1 Detection of fluorescence of Cy5 and ECL-double stained membrane. After incubation with ECL plus (plex) substrate, place the membrane on a low-fluorescent glass plate with the stained side on which the proteins are stuck facing down. Drip a few milliliters of water on the back of the membrane and place another glass plate on the membrane to allow the water to spread over the entire surface of the membrane to remove air bubbles

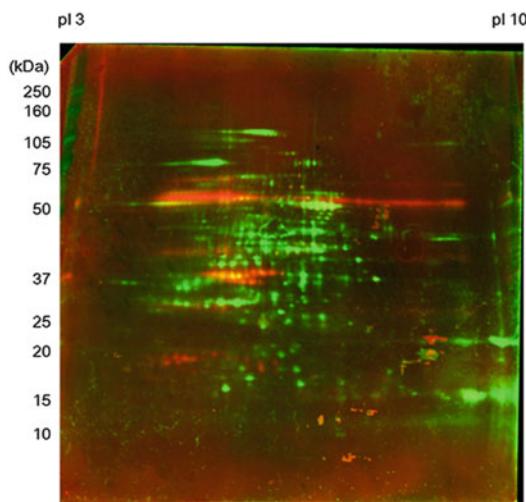


Fig. 2 A merge image of all soluble rice seed protein- and IgE-binding protein-patterns. All transferred rice seed proteins were labeled with Cy5, and the IgE-binding proteins were labeled with ECL after incubation with the serum of a patient with rice allergy (diluted 1:100) followed by that with HRP-linked anti human IgE (diluted 1:1,000). The merged scan image of the Cy5 (green) and ECL (red) images is shown

4. Dehydrate the gel pieces with dehydration buffer followed by acetonitrile, and air-dry.
 5. Add 10 μ L of trypsin digest solution (20 ng/ μ L of Trypsin Gold in 0.01 % Protease Max solution) to the completely dried gels. After keeping the tubes on ice for 15 min, incubate them for 2 h at 37 °C.
1. Desalt the tryptic digests with Zip-Tip C18.
 2. Mix 0.5 μ L of digest and 1 μ L of α -CHCA and spot onto an Opti-TOF 384 well Insert plate.

3.5 Mass Spectrometry and Protein Identification

3. Determine mass spectra and MS/MS fragment ion mass with a 4800 MALDI TOF/TOF analyzer.
4. Submit all product ions to a computer database search analysis with the Mascot MS/MS ion search.

4 Notes

1. Appropriate extraction buffers for plant specimens or allergens should be selected. In general, soluble buffers such as PBS, Tris-HCl or NaCl solution are used as the extraction buffer. Commercial protein extraction kits for plant specimens may also yield good results within a short time. Prior to the 2D-immunoblotting, it should be confirmed by ELISA or immunoblotting using the patients' sera, that the extracts contain allergens.
2. After addition, vortex well and allow standing until the resins are completely settled. Do not use if the powder color changes to yellow, or the saturated solution is not clear.
3. Careful attention should be paid to desalt samples, since a high-salt buffer may be used for extracting the allergens. Most failures in spot focusing may be caused by not paying sufficient attention to this step.
4. Move the strip up and down so that entire surface of the gel comes in contact with the sample solutions.
5. Slowly drop the oil from the tip of the holder. Excessively rapid dropping of the oil would cause the samples to float above the gels. Be careful to avoid flooding with oil from the side of the holder.
6. We set long-time runs at low voltage so that any salt contaminating the samples moves away to the tip of the holder. It would not be a problem even if the voltage did not reach up to 8,000 V at the end of the run. If the voltage does not rise at all during the first step, it could be assumed that the sample preparation or rehydration has not been satisfactory.
7. Addition of buffers to a frozen strip might cause incomplete reduction and/or alkylation. Put the side of the strip on the bottom of the dish so that the gels get completely flooded by the buffer.
8. Cut 5 mm off both sides of the plastic sheet on the gels as to obtain a good fit into the gel slot. Place the plastic side of the strip on the longer plate and put it on the acrylamide gel with the flat side of a spatula. Be careful not to put it into the gels. If air bubbles are present between the strip and gel, remove them with a spatula.

9. When the gel is taken off from the glass plates, take care not to break the gel. Cut the acidic corner of the gel to see its direction.
10. Air bubbles between the gel and membrane should be removed. Change the time and/or the electric current, depending on the target range of molecular weight. Semi-dried transfer systems may also be used for blotting proteins to the membrane, while we have obtained good results using tank transfer systems. After the electrical transfer, cut the acidic corner of the membrane to see its direction.
11. Incubate the membrane face down so as to prevent drying of the surface of the membrane. After the incubation, check whether the entire surface of the membrane is stained blue.
12. The membrane should be incubated with more than 20 mL of serum, sufficient to cover the entire surface of the membrane. Take care to avoid drying of the surface of the membrane. The serum dilution could be changed depending on the total amount of IgE. Usually, we use 100-fold diluted serum (Fig. 2).
13. After incubation, the serum solution must be sterilized with 1 % sodium hypochlorite before it is disposed of.
14. ECL plus (plex) reagent should cover the entire surface of the membrane. Use of excessive reagent or a long reaction time might result in excessive chemical reaction, resulting in void spots. Dilution of the reagent with ultrapure water or a short reaction time might yield better results.
15. Insufficient water may cause drying of the surface of the membrane. Slowly put the glass plate down on the water hill to allow the water to spread over the entire surface of the membrane. Spotty air bubbles between the membrane and the glass plate would yield unclear images.

Acknowledgments

This work was supported by a grant from the Ministry of Health, Labour and Welfare, Japan.

References

1. Burks AW, Williams LW, Helm RM et al (1991) Identification of a major peanut allergen, Ara h I, in patients with atopic dermatitis and positive peanut challenges. *J Allergy Clin Immunol* 88:172–179
2. Fränken J, Stephan U, Meyer HE et al (1994) Identification of alpha-amylase inhibitor as a major allergen of wheat flour. *Int Arch Allergy Immunol* 104:171–174
3. Daul CB, Slattery M, Reese G et al (1994) Identification of the major brown shrimp (*Penaeus aztecus*) allergen as the muscle protein tropomyosin. *Int Arch Allergy Immunol* 105:49–55

4. Picariello G, Mamone G, Addeo F et al (2011) The frontiers of mass spectrometry-based techniques in food allergenomics. *J Chromatogr A* 1218:7386–7398
5. Vallier P, DeChamp C, Valenta R et al (1992) Purification and characterization of an allergen from celery immunochemically related to an allergen present in several other plant species. Identification as a profilin. *Clin Exp Allergy* 22:774–782
6. Beyer K, Grishina G, Bardina L et al (2002) Identification of an 11S globulin as a major hazelnut food allergen in hazelnut-induced systemic reactions. *J Allergy Clin Immunol* 110: 517–523
7. Beyer K, Bardina L, Grishina G et al (2002) Identification of sesame seed allergens by 2-dimensional proteomics and Edman sequencing: seed storage proteins as common food allergens. *J Allergy Clin Immunol* 110:154–159
8. Magni C, Herndl A, Sironi E et al (2005) One- and two-dimensional electrophoretic identification of IgE-binding polypeptides of *Lupinus albus* and other legume seeds. *J Agric Food Chem* 53:4567–4571
9. Boldt A, Fortunato D, Conti A et al (2005) Analysis of the composition of an immunoglobulin E reactive high molecular weight protein complex of peanut extract containing Ara h 1 and Ara h 3/4. *Proteomics* 5:675–686
10. Restani P, Ballabio C, Corsini E et al (2005) Identification of the basic subunit of Ara h 3 as the major allergen in a group of children allergic to peanuts. *Ann Allergy Asthma Immunol* 94:262–266
11. Herndl A, Marzban G, Kolarich D et al (2007) Mapping of *Malus domestica* allergens by 2-D electrophoresis and IgE-reactivity. *Electrophoresis* 28:437–448
12. Batista R, Martins I, Jeno P et al (2007) A proteomic study to identify soya allergens—the human response to transgenic versus non-transgenic soya samples. *Int Arch Allergy Immunol* 144:29–38
13. Akagawa M, Handoyo T, Ishii T et al (2007) Proteomic analysis of wheat flour allergens. *J Agric Food Chem* 55:6863–6870
14. Sotkovský P, Hubálek M, Hernychová L et al (2008) Proteomic analysis of wheat proteins recognized by IgE antibodies of allergic patients. *Proteomics* 8:1677–1691
15. Fasoli E, Pastorello EA, Farioli L et al (2009) Searching for allergens in maize kernels via proteomic tools. *J Proteomics* 72:501–510
16. Postigo I, Guisantes JA, Negro JM et al (2009) Identification of 2 new allergens of *Phoenix dactylifera* using an immunoproteomics approach. *J Investig Allergol Clin Immunol* 19:504–7
17. Pignataro V, Canton C, Spadafora A et al (2010) Proteome from lemon fruit flavedo reveals that this tissue produces high amounts of the Cit s 1 germin-like isoforms. *J Agric Food Chem* 58:7239–7244
18. Nakamura R, Satoh R, Nakamura R et al (2010) Immunoproteomic and two-dimensional difference gel electrophoresis analysis of *Arabidopsis* dehydration response element-binding protein 1A (DREB1A)-transgenic potato. *Biol Pharm Bull* 33: 1418–1425
19. Palacín A, Tordesillas L, Gamboa P et al (2010) Characterization of peach thaumatin-like proteins and their identification as major peach allergens. *Clin Exp Allergy* 40:1422–1430
20. Satoh R, Nakamura R, Komatsu A et al (2011) Proteomic analysis of known and candidate rice allergens between non-transgenic and transgenic plants. *Regul Toxicol Pharmacol* 59:437–444
21. Satoh R, Nakamura R, Teshima R (2011) Proteomic identification of IgE-binding proteins in buckwheat. *Jpn J Food Chem Safety* 18:103–109
22. Luengo O, Mollá R, Gámez C et al (2008) Allergenicity and cross-reactivity of *Senecio* pollen: identification of novel allergens using the immunoproteomics approach. *Clin Exp Allergy* 38:1048–1060
23. Nakamura R, Satoh R, Nakajima Y et al (2009) Comparative study of GH-transgenic and non-transgenic amago salmon (*Oncorhynchus masou ishikawai*) allergenicity and proteomic analysis of amago salmon allergens. *Regul Toxicol Pharmacol* 55:300–308
24. Nakamura R, Nakamura R, Nakano M et al (2010) Allergenicity study of EGFP-transgenic chicken meat by serological and 2D-DIGE analysis. *Food Chem Toxicol* 48:1302–1310
25. Magni C, Ballabio C, Restani P et al (2005) Two-dimensional electrophoresis and western-blotting analyses with anti Ara h 3 basic subunit IgG evidence the cross-reacting polypeptides of *Arachis hypogaea*, *Glycine max*, and *Lupinus albus* seed proteomes. *J Agric Food Chem* 53:2275–2281
26. Chassaigne H, Trégoat V, Nørgaard JV et al (2009) Resolution and identification of major peanut allergens using a combination of fluorescence two-dimensional differential gel electrophoresis, Western blotting and Q-TOF mass spectrometry. *J Proteomics* 72:511–526

Chapter 51

Beer and Wort Proteomics

Takashi Iimure, Makoto Kihara, and Kazuhiro Sato

Abstract

Proteome analysis provides a way to identify proteins related to the quality traits of beer. A number of protein species in beer and wort have been identified by two-dimensional gel electrophoresis combined with enzyme digestion such as trypsin, followed by mass spectrometry analyses and/or liquid chromatography mass/mass spectrometry. In addition, low molecular weight polypeptides in beer have been identified by the combination of non-enzyme digestion and mass analyses. These data sets of various molecular weight polypeptides (i.e., proteomes) provide a platform for analyzing protein functions in beer. Several novel proteins related to beer quality traits such as foam stability and haze formation have been identified by analyzing these proteomes. Some of the proteins have been applied to the development of efficient protein or DNA markers for trait selection in malting barley breeding. In this chapter, recent proteome studies of beer and wort are reviewed, and the methods and protocols of beer and wort proteome analysis are described.

Key words Beer, Wort, Proteomics, Two-dimensional gel electrophoresis, Mass spectrometry

1 Introduction

Beer is one of the oldest fermented beverages, and is the most widely consumed alcoholic drink. The raw materials of beer are barley malt, hops, and sometimes adjuncts such as starch from corn and/or rice. Beer is produced by a number of processes including malting, mashing, wort boiling, fermentation, maturation, and filtration. The malting process consists of three steps: steeping, germination, and kilning. Various enzymes such as protease, amylase, and β -glucanase degrade proteins, starch, and β -glucan in the malting and mashing steps. Proteins and starch are degraded into amino acids and fermentable sugars as sources of nitrogen and carbon, respectively, to grow brewing yeast. In the wort boiling process, proteins and protein–polyphenol complexes coagulate and precipitate. Therefore, the resultant barley-derived proteins are both protease- and heat-stable. Several modifications such as glycation, acylation, and partial digestion also occur in these processes. In addition to these modified proteins, beer contains low molecular

weight polypeptides. Other than barley proteins, yeast proteins are eluted into beer in the fermentation process.

Beer proteins have been extensively analyzed and they are important components of beer quality traits such as foam, haze (clarity), gushing, and mouthfeel. Protein Z and lipid transfer protein 1 (LTP1) are abundant in beer. Therefore, a series of beer protein analyses have been conducted on these proteins. A negative effect on hordein or its derived polypeptides has been known to cause coeliac disease [1–3] with clinical symptoms of fatigue, diarrhea, abdominal distension, weight loss, and neurological disorders [4]. However, the molecular weight and amino acid sequences of hordein-derived polypeptides in beer are contentious.

These beer proteins have been analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), Western blotting, and enzyme-linked immunosorbent assay (ELISA) using specific antibodies. Since the 1980s, two-dimensional gel electrophoresis (2DE) has been applied to analyze beer protein profiles. However, proteins were not precisely identified in older studies [5–7]. In recent reports, beer proteins were analyzed by 2DE followed by mass spectrometry [8–11] and liquid chromatography mass/mass spectrometry (LC-MS/MS) followed by searches in DNA and/or protein sequence databases [11, 12]. In this chapter, these beer and wort proteome analyses are reviewed, and the methods of 2DE analysis of beer and wort samples are described.

2 Progress of Proteomics Analysis in Beer and Wort

The massive generation of nucleotide sequences in barley and yeast accelerated the studies to find the protein sequences responsible for specific traits in the target beer samples. In addition, the high sensitivity mass spectrometry systems improved the chances to find a number of expressed protein spots with the help of the database search engine. With this proteome approach, novel proteins have been identified in worts and beers as they are mapped on the gel. We can call these as proteome maps. The proteome map is the foundation stone for the quantification of trait-related proteins among the beer samples made from different materials and treatments.

2.1 Earlier Studies of Beer Proteins

A pioneering analysis of beer by 2DE was done by Marshall and Williams [5]. They analyzed commercial beer samples without concentration and simply compared their 2DE images. Dale and Young [6] analyzed beer samples by 2DE to compare the protein profile between the samples. They detected four specific protein spots in the beer from malt with 20 % wheat flour, which were absent in the all malt beer. Williams and Marshall [7] developed a high-resolution 2DE method for beer analysis comprising high-resolution 2DE and ultrasensitive silver staining for simultaneous

analysis of up to 300 polypeptide constituents in microliter volumes of unconcentrated beer. However, protein identification was not possible in these earlier studies. Therefore, the species of beer protein was unknown, and the information, whether it was related to beer quality, was not clear from these studies.

2.2 Proteome Maps of Beer and Wort

2.2.1 One Dimensional Gel Electrophoresis Analysis of Beer Proteins

A specific protein profile may be required in each beer type. Protein profiles are influenced by barley grain (cultivars) and their malting and brewing conditions. In addition, differences in beer and wort quality profiles can be detected by the comparison of protein concentrations. Beer protein analysis started with one-dimensional gel electrophoresis. The analysis of the protein profile of oat wort, as measured by using Lab-on-a-Chip analysis, revealed that there were no significant differences in the protein profiles between oat and barley worts [13, 14]. Bobálová et al. [15] analyzed barley grain and malt proteins of two cultivars by SDS-PAGE and mass spectrometry and found no significant difference in the two cultivars tested. Hao et al. [16] analyzed beer and beer foam proteins by SDS-PAGE and tandem mass spectrometry, and identified 21 and 24 proteins in beer and beer foam, respectively. However, these one-dimensional gel electrophoresis analyses did not have enough resolution to compare the protein profiles of the samples.

2.2.2 2DE and LC-MS/MS Analyses of Beer Proteins and Protein Identification

Two-dimensional gel electrophoresis (2DE) is more appropriate for comparing proteomes because it has improved resolution in protein detection compared with one-dimensional gel electrophoresis i.e., SDS-PAGE. In past decades, exponential improvement in mass spectrometry techniques such as matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) methods, and peptide searching systems in protein and DNA databases led to the identification of protein species on a large scale. Though the barley genome has not been completely sequenced to date, expressed sequence tag (EST) and full-length cDNA sequences strongly support the protein identification [17, 18]. Based on the advancement in these research tools, beer proteome analysis is progressing rapidly.

A proteome map, which displays identified protein spots on a 2DE image, is a platform for analyzing beer proteins. Perrocheau et al. [8] analyzed beer proteins by 2DE and mass spectrometry and identified 30 protein spots. Iimure et al. [9] identified 85 protein spots on 2DE gel out of 199 spots, and categorized them into 12 protein species. Based on these proteome maps, they compared the 2DE images of beer proteins, which were derived from malt samples of different barley cultivars and malt modifications, and found that the spot intensities of several beer-quality-related proteins were different among the beer samples. Fasoli et al. [19] identified more beer protein species, especially proteins derived from yeast using a unique technique as follows: they analyzed beer proteomes through prior capture with combinatorial peptide

ligand libraries (CPLL) (ProteoMiner as well as a homemade library of reduced polydispersity) at three different pH values. The CPLL technique offers a unique increment in sensitivity for low abundance protein species. Therefore, they identified 20 and 40 unique gene products from barley and *Saccharomyces cerevisiae*, respectively, by the mass spectrometry analysis of the recovered fractions. Avenin-like protein-a was identified in beer and the protein enriched using silica gel from beer foam by nano-ESI-MS/MS analyses [11, 12, 20]. Weber et al. [20] analyzed eight different beer products by LC-MS/MS after digestion with trypsin and identified B-, D-, and γ -hordeins and several albumins such as protein Z and LTP1. Colgrave et al. [12] analyzed wort, beer, and hordeins extracted from barley flour by LC-MS/MS and identified B-, C-, D-, and γ -hordeins and avenin-like protein-a. These results suggested the feasibility of alternative techniques to determine allergen and gluten markers in beer instead of ELISA-based methodologies. Table 1 summarizes a list of proteins identified in beer.

2.2.3 Proteome Analysis of Wort

2DE has been used to analyze barley grain and/or malt proteins for over 20 years and mass spectrometry has only recently been applied. Beer proteins consist predominantly of barley albumins and globulins, which are released into wort, and subsequently, appear in small amounts after processing (20–600 mg/L) [21]. Only approximately one-third of the total protein content remains in beer after degradation by mashing and precipitation during wort boiling. Therefore, quantitative and qualitative changes of proteins occur in malting and brewing processes. These changes are also significantly caused by malt modifications. Silva et al. [22] investigated the protein fractions of malt, wort, and beer from two barley cultivars by SDS-PAGE and reversed-phase HPLC (RP-HPLC), and showed that protein profiles from the wort samples between the two cultivars tested were different; however, those between beer samples were similar. Perrocheau et al. [8] analyzed barley grain, malt, and beer proteins by 2DE and mass spectrometry to investigate proteome changes during malting and brewing. In these processes, most of the heat-stable proteins are disulfide-rich. A comprehensive proteome map constructed for sweet wort by Iimure et al. [23] consisted of 63 identified proteins out of 202 protein spots on 2DE images, which were categorized into 20 protein species. Colgrave et al. [12] identified 27 proteins including hordeins in the wort sample by LC-MS/MS analysis.

2.3 Analysis of Low Molecular Weight Polypeptides in Beer and Wort

Barley proteins are digested in malting and mashing processes by malt proteases, and subsequently amino acids and low molecular weight polypeptides are present in the wort. 2DE cannot be used to analyze low molecular weight polypeptides, i.e., below approximately 8 kDa. To detect low molecular weight polypeptides, we should use alternative methods such as LC-MS/MS in addition to 2DE analysis.

Table 1
The list of identified proteins in beer by proteomics analysis

| Protein name | References | Protein name | References |
|---|----------------------|--|------------|
| <i>Hordeum vulgare</i> | | Calcium-dependent protein kinase | [10] |
| Alpha-amylase inhibitor BDAI-1 precursor | [8–10, 12], [16, 19] | Glucose and ribitol dehydrogenase homologue—barley | [10] |
| α-Amylase inhibitor, BMAl-1 precursor | [10, 12, 16, 19] | Grain softness protein | [10] |
| α-Amylase inhibitor | [10] | Pathogenesis-related protein 4 | [10] |
| α-Amylase inhibitor, CMa precursor ^a | [12, 16] | Triosephosphate isomerase, cytosolic | [10] |
| Chloroform/methanol-soluble protein CMb | [8–12, 16, 19] | 26 kDa endochitinase 2 | [12] |
| Barley trypsin inhibitor CMc | [12] | β-Amylase ^a | [12] |
| α-Amylase inhibitor, tetrameric, CMd | [10–12, 19] | Putative protease inhibitor | [12] |
| BTI-CMe ^a | [8–10, 12, 16, 19] | Thaumatin-like protein TLP6 ^b | [12] |
| Lipid transfer protein 1 ^a | [8–11, 16, 19] | Thaumatin-like protein TLP8 | [12] |
| LTP2 | [8, 11, 19, 45] | β-Glucosidase | [12] |
| Protein z-type serpin (Z4) ^a | [8–11, 16, 19] | Protein disulfide isomerase | [12] |
| Serpin-Z7 (HorvuZ7) (BSZ7) | [9, 10, 16, 19] | Elongation factor 1-alpha ^a | [12] |
| Serpin-Zx | [10, 19] | Reversed predicted protein | [12] |
| Subtilisin-chymotrypsin inhibitor CI-1A, 2A, 1B, 1C | [9, 10, 12, 16] | Oleosin | [12] |
| Chymotrypsin inhibitor 2 | [10] | Cystatin Hv-CP16 | [10] |
| Trypsin/amylase inhibitor pUP13 | [8–10] | Cystatin Hv-CP18 | [12] |
| Barwin ^a | [12, 16, 19] | Similar to PRF gamma hordein | [12] |
| Leaf specific thionin, Leaf specific thionin DB4 | [19] | Allene oxide synthase ^b | [12] |
| Glycine-rich RNA binding protein b1t801 | [19] | 26 kDa endochitinase 1 | [12] |

(continued)

Table 1
(continued)

| Protein name | References | Protein name | References |
|--|-----------------|---|----------------|
| Hordeindoline-A, B1, B2 ^a | [10, 12, 19] | Hordeum chilense | |
| Calmodulin | [19] | γ 3-Hordein ^b | [11] |
| 40S ribosomal protein S7 | [19] | S-adenosylmethionine decarboxylase proenzyme ^b | [11] |
| B1-hordein ^b | [11, 12] | Hordeum brevisubulatum | |
| B3-hordein ^a | [12] | B-hordein ^b | [11] |
| B-hordein ^a | [10-12] | Saccharomyces cerevisiae | |
| C-hordein ^b | [12] | Triosephosphate isomerase | [8, 9, 11, 19] |
| D-hordein ^a | [10, 12] | Yeast phosphorelay protein Ypd1 | [9] |
| Hordein γ -3 ^a | [8, 10, 12, 19] | Enolase 1, 2 | [8-11, 19] |
| Gamma-hordein-1 ^a | [12] | Thioredoxin; Trx2p | [9, 19] |
| Weakly similar to UP Gamma-ghadin ^b | [12] | 1, 3 β -Glucanosyltransferase GAS3 | [19] |
| Avenin-like protein a ^a | [11, 12] | Alcohol dehydrogenase 1,4 | [19] |
| Embryo globulin | [12] | Box C/D sno RNA protein 1 | [19] |
| Protein E13 ^b | [11] | Cell wall mannoprotein CIS3 | [19] |
| Putative glu-decarboxylase ^b | [11] | Cell wall protein ECM33 | [19] |
| P-type ATPase ^b | [11] | Cell wall protein, Scw10p | [10] |
| Late embryogenesis abundant protein | [12] | Cell wall protein, Scw4p | [10] |
| 17 kDa class I small heat shock protein | [12] | 2-Phosphoglycerate dehydratase | [10] |
| 18-kDa heat shock protein | [10] | 3-Phosphoglycerate kinase | [10] |
| Tetrameric phosphoglycerate mutase | [10] | Pyruvate kinase 1 | [19] |

| | | | |
|--|----------|---|------|
| Acyl-Co A-binding protein 2 | [10] | Saccharipepsin | [19] |
| Endo- β -1, 3-glucanase. | [10] | tRNA wybuto sine-synthesizing protein 1 | [19] |
| Major exo-1, 3- β -glucanase | [10] | Uncharacterized protein YGR237C, YOR020W-A | [19] |
| Glucan 1, 3- β -glucosidase | [10] | Oye2p | [10] |
| Thioredoxin peroxidase | [10] | Pep4p | [10] |
| Cytoplasmic thioredoxin isoenzyme | [10] | Pgl1p | [10] |
| Glycolipid-anchored surface protein | [10] | Pry1p | [10] |
| Heat shock protein of HSP70 family | [10] | Tos1p | [10] |
| Mannose-containing glycoprotein | [10] | UTH1 | [10] |
| Triose phosphate isomerase, abundant glycolytic enzyme | [10] | Vellp | [10] |
| Uncharacterized protein YER188W | [10] | Saccharomyces pastorianus | |
| Vacuolar proteinase B | [10] | α -Galactosidase | [19] |
| Coproporphyrinogen III oxidase | [10, 19] | Saccharomyces bayanus | |
| Glucan 1, 3 β -glucosidase, I/II | [19] | Glucan 1, 3 β -glucosidase | [19] |
| Glucose-6-phosphate isomerase | [19] | Others | |
| Glyceraldehyde-3-phosphate dehydrogenase, 2 ^a | [10, 19] | CM 17 protein precursor ^c | [10] |
| Hydroxymethylglutaryl-CoA synthase | [19] | Gliadin/avenin-like seed protein ^c | [10] |
| Invertase 1, 2 | [19] | Globulin 3B ^c | [10] |
| Long-chain fatty-acid-CoA ligase 4 | [19] | High molecular weight glutenin subunit ^c | [10] |
| Phosphoglycerate kinase | [19] | Os01g0915900 ^d | [10] |
| Phosphoglycerate mutase 1 | [19] | Os03g0393900 ^d | [10] |
| Pre-mRNA leakage protein 1 | [19] | Os06g0650100 ^d | [10] |

(continued)

Table 1
(continued)

| Protein name | References | Protein name | References |
|--|------------|--|------------|
| Probable family 17 glucosidase SCW4, SCW10 | [19] | Polyubiquitin containing seven ubiquitin monomers ^d | [10] |
| Probable glucosidase CRH1 | [19] | Cytochrome c ^c | [12] |
| Probable transporter SEO1 | [19] | HMV glutenin subunit x ^e , y ^f | [12] |
| Profilin | [10, 19] | Oxygen-evolving enhancer protein 3-2, chloroplastic ^g | [19] |
| Protein EGT2 | [19] | Protein terminal ear1 ^g | [19] |
| Protein NCA3, mitochondrial | [19] | Rp 1-like protein ^c | [12] |
| NCA3 | [10] | Wali3 protein ^f | [12] |
| Protein SIM1 | [19] | Peptidyl-propyl cis-trans isomerase ^{a,h} | [12] |
| Protein TBE1 | [19] | ATP synthase subunit alpha ^h | [12] |
| Protein TOS1 | [19] | Low molecular weight glutenin subunit ^h | [12] |
| Protein UTH1 | [19] | Serpi-Z2A ^{b,c} | [12] |
| Protein VEL1 | [19] | P0529E05_3 ^d | [11] |
| | | Glucoamylase G1 ⁱ | [10] |

^aIndicates low molecular weight (0.5–10 kDa) polypeptides as well as proteins with intact molecular weight
^bIndicates only low molecular weight (0.5–10 kDa) polypeptides identified

^c*Triticum aestivum*

^d*Oryza sativa*

^e*Secale cereale*

^f*Aegilops comosa*

^g*Zea mays*

^h*Triticum turgidum* subsp. *durum* × *Triticosecale* sp.

ⁱ*Aspergillus niger*

Low molecular weight polypeptides constituted over 75 % by weight of the total polypeptide content of beer, as estimated by peak area integration of size exclusion chromatography [11, 24]. Picariello et al. [11] analyzed beer polypeptides without enzyme digestion using nanoscale capillary liquid chromatography–electrospray tandem mass spectrometry (LC–ESI MS/MS) coupled with a hybrid quadrupole orthogonal acceleration time of flight mass spectrometer (Q-TOF MS). They identified several C-terminal fragments of protein Z ($M_w < 4,500$ Da), small size peptides derived from LTP1 ($M_w < 1,000$ Da), and various low molecular mass peptides ($M_w < 1,000$ Da) derived from γ 3-, B-, and B1-hordeins. In more recent research, Colgrave et al. [12] also analyzed the beer and wort proteins by LC–MS/MS without enzyme digestion and identified a number of proteins including γ -, C-, B1-, B3-, and D-hordeins and avenin-like protein-a. These two reports had a high impact in identifying hordein-derived polypeptides in beer. C-hordein degradation products were observed in wort but only trace amounts were found in beer, suggesting that the majority of these peptides do not survive the brewing and filtering processes [12].

2.4 Identification of Proteins Related to Beer Quality Traits

Foam and colloidal haze (clarity) are key factors for the visual impression of beer. Excellent beer foam gives an impression that the beer is fresh, pleasing to behold, and adequately carbonated. Also, foam stability and bubble size influence the feel on the lips and mouth. Beer foam quality is characterized by stability, quantity, lacing, whiteness, creaminess, density, viscosity, and strength [25]. Among these, foam stability is influenced by beer proteins. Several studies have shown that protein Z family (Z4 and Z7) [26–29] and LTP1 [30, 31] are involved in beer foam stability. Beer proteins also influence beer haze formation. Haze is a serious quality problem since stale beer displays haze. It has been suggested that hordein [32], which is the most abundant storage protein in barley grain is a negative factor for colloidal haze formation.

However, factors causing beer foam stability and haze formation are still controversial. One of the reasons for the poor understanding of protein factors controlling beer quality traits may come from the poor knowledge of protein species contained in beer samples. Thus, comprehensive analysis of beer protein by proteomics is essential to estimate the relationship between beer proteins and beer quality traits, e.g., foam stability and haze formation.

2.4.1 Beer Foam-Related Proteins

Protein Z and LTP1 are abundant proteins in beer, and these proteins have been suggested as foam-positive proteins. In particular, modification of protein Z and LTP1 was analyzed in detail using proteomics. van Nierop et al. [31] compared the 2DE images of LTP1 among normal, 96, and 105 °C heat treated samples, and found that heat treatment caused significant modification.

Reduction of disulfide bonds in barley grain, malt, and beer were surveyed by 2DE using the labeling of free thiol groups of water-soluble heat-stable proteins including LTP1 [33]. These results suggested the presence of free cysteine in malt and beer LTP1 but not in that of barley. In a very recent paper [34], differences in protein compositions of beer brewed from protein Z4 deficient, protein Z7 deficient, double deficient, and control barleys were detected by 2DE. The spots with different intensities among the beer samples were analyzed by matrix-assisted laser desorption ionization time-of-flight/time-of-flight mass spectrometry (MALDI TOF-TOF MS), resulted in the identification of lipid transfer protein 2 (LTP2). Iimure et al. [23] detected protein Z in sweet wort, boiled wort, and trub by 2DE, MALDI TOF-MS, and LC-MS/MS, and found that protein Z was precipitated during wort boiling by binding to comparatively small specific fragment(s) derived from the sweet wort protein, i.e., barwin.

Several novel foam-related proteins were identified by beer proteome analysis. In general, beer foam stability decreases by increasing malt modification. Okada et al. [35] found that foam stability of beer samples brewed from malts of two Japanese barley cultivars decreased as the level of malt modification increased, however, the foam stability of a Canadian barley cultivar tested stayed at a higher level. To identify the protein responsible for the foam stability of the Canadian cultivar, three fractions of beer samples namely beer whole proteins, salt-precipitated proteins, and the proteins concentrated from beer foam were analyzed by 2DE and mass spectrometry. As a result, barley dimeric alpha-amylase inhibitor-1 (BDAI-1) was identified as a foam-positive protein. They also found that beer protein composition depends much on the barley cultivar and malting condition such as the ex-steep moisture level. Iimure et al. [36] analyzed all-malt, standard (67 % of malt in total raw materials), and low malt (Happoshu: below 24 % of malt in total raw materials) beer samples with different foam stability using 2DE and mass spectrometry, and found that yeast thioredoxin was identified as a possible foam-negative protein. Proteinase A from yeast is well known as a foam negative factor [37–39]. Iimure et al. [36] first identified yeast thioredoxin as a foam-negative protein, other than proteinase A. Yeast thioredoxin is reported to be an intracellular protein [40, 41]; however, the identification of thioredoxin in beer suggested that yeast cells were damaged during brewing. In addition, several beer proteome analyses revealed that some yeast intracellular proteins such as enolase and triose-phosphate isomerase were contained in beer [8–11, 19]. Beer foam stability and flavor stability were also affected by the physiological condition of the yeast cells [42, 43]. These yeast proteins could be effective markers for controlling the physiological condition of brewing yeast in terms of stability in beer foam and flavor.

2.4.2 Beer Haze-Related Proteins

Haze-related proteins have been analyzed using the proteome approach. Haze-active proteins can be removed by physical treatment, i.e., application of silica gel [44, 45]. Therefore, brewers generally add silica gel during beer filtration to prevent beer colloidal haze. Robinson et al. [46] prepared an antibody that specifically bound to silica gel eluent proteins (SE proteins), and identified trypsin inhibitor CMe precursor (CMe) as a haze-active protein using 2DE and N-terminal sequence analysis of tryptic peptides. Iimure et al. [47] analyzed four haze samples and SE proteins by 2DE and mass spectrometry and suggested that BDAI-1, a component of tetrameric alpha-amylase inhibitor (CMb), and CMe were not predominant haze-active proteins, but growth factors for beer colloidal haze. In addition, Jin et al. [48] analyzed silica gel eluent proteins by 2DE and mass spectrometry and identified CMe, germin E, and protein Z as possible haze-active proteins. Jin et al. [49] suggested that hordein may be minor haze-active proteins but are crucial for haze formation according to a 2DE analysis of malt, wort, and haze samples. Colgrave et al. [12] quantified the contents of hordein family proteins such as avenin-like protein-a, B1-, B3-, D-, and γ -hordeins in commercial beer samples by Multiple Reaction Monitoring (MRM) mass spectrometry. The hordein family contents in beer depended on the samples, suggesting that barley cultivar, malting, and brewing conditions, type of beer, and adjuncts affected their contents. Polypeptides derived from the hordein family have been suggested to be foam-positive and haze-negative [32, 50]. However, the relation between hordein and foam quality has not been clarified [51, 52].

A number of beer proteins and polypeptides have been identified [8–12, 19], and these data form the basis of a strong platform for further study of beer proteomes. Although several proteins related to beer quality traits such as foam stability and haze formation have been identified, currently, the mechanism of the foam and haze formation is still controversial. The objectives of beer proteome study are identification of quality-related proteins, detection of proteome differences in malting and brewing conditions, and investigation of proteome differences among barley cultivars. From this perspective, however, beer proteome study is still at an early stage. As well as the advancement of mass spectrometry, development of fractionation methods of beer and wort proteins such as salt-precipitated proteins [35], foam-concentrated proteins [35], silica gel eluent proteins [46–48], and prior capture using CPLL technology [19] are important to detect objective protein species. In addition, collecting appropriate samples that differ in the target quality trait is essential to identifying the respective proteins. To apply the knowledge of beer and wort proteomes in malting barley breeding, the genetic analysis of objective proteins is necessary as demonstrated by Robinson et al. [46] and Iimure et al. [28]. From this perspective, proteomics must be further improved according to the advancement of barley genomics.

3 Protocol for Beer and Wort Proteomics Using Two-Dimensional Gel Electrophoresis

Desalting and concentration of proteins are keys to obtaining excellent gel images in 2DE of beer and wort proteins. To obtain clear protein spots on the 2DE gel, appropriate desalting is especially important. The 2DE step is applied as usual. As an example, the methods and protocol of sample preparation, 2DE, staining, and mass spectrometry are shown below, although other methods can also be applied. The strategy of beer and wort proteome analysis and typical 2DE images of beer and wort are shown in Fig. 1.

3.1 Materials

1. PD-10 column (GE Healthcare Biosciences, Tokyo, Japan).
2. Protein assay solution: Dilute (five-fold) Bio-Rad Protein Assay Kit II (Bio-Rad Laboratories, Tokyo, Japan) with ultrapure water, and filtered it through filter paper.
3. Filter paper: Advantec, #2, 90 mm (Toyo Roshi Kaisha, Ltd, Tokyo, Japan).
4. Bovine serum albumin (BSA): included in Bio-Rad Protein Assay Kit II (Bio-Rad Laboratories).
5. 96-well plate: assay plate, 96-well, flat bottom (well dia. 6.4 mm) (Iwaki, Tokyo, Japan).
6. VARIOSKAN microtiter plate reader (Thermo Electron Corporation, Yokohama, Japan).

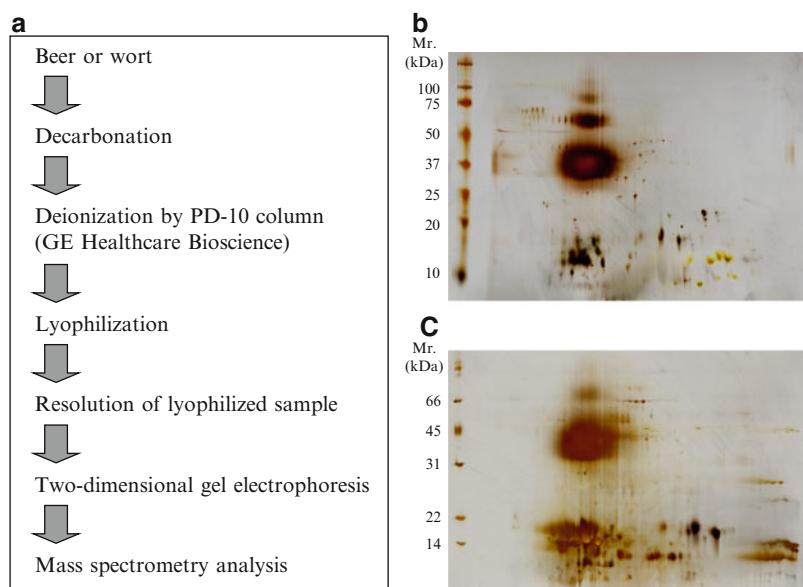


Fig. 1 Strategy of beer and wort proteome analysis. (a) Experimental flow of beer and wort proteome analysis; (b) Typical 2DE image of beer; (c) Typical 2DE image of wort

7. Dissolving buffer: 5 mL of 8 M urea (Wako, Japan) with 2 % 3-[(3-cholamidopropyl) dimethylammonio] propane-sulfonic acid (CHAPS) (Dojindo Laboratories, Kumamoto, Japan) and 0.28 % dithiothreitol (Wako, Tokyo, Japan) (*see Note 1*).
8. IPG buffer (GE Healthcare Bioscience) (*see Note 2*).
9. IPG dry strip (GE Healthcare Bioscience) (*see Note 2*).
10. Multiphor II system (GE Healthcare Bioscience).
11. Equilibration buffer: 50 mM Tris-HCl, pH 8.8, 6 M urea, 30 % v/v glycerol, 2 % sodium dodecyl sulfate (SDS) (*see Note 3*).
12. Precast XL 12–14 % gradient gel (GE Healthcare Bioscience).
13. Silver Staining Kit, Protein (GE Healthcare Bioscience). Silver Stain MS Kit (Wako).
14. Decolorizing buffer: 15 mM potassium ferricyanide and 50 mM sodium thiosulfate.
15. Zip-Tip (Nihon Millipore Ltd., Tokyo, Japan).
16. Voyager-DE STR (Applied Biosystems, Foster City, California USA).

3.2 Methods

3.2.1 Sample Preparation

1. From a case of beer, open a beer sample packed in a bottle, can, or any other packaging, and then pours it into a 300 mL conical flask (*see Note 4*). Agitate the flask by hand several times during foam formation to degas the beer. After no foam is generated by shaking, leave the flask overnight in a wrap (*see Note 5*). In the case of wort, this procedure is not required because wort is not carbonated.
2. Open the PD-10 column, and discard the buffer in the column. Load 5 mL of distilled deionized water into the column, and run the water through the column. To achieve equilibration of the column, repeat the water loading four times, with a total of 25 mL of the distilled deionized water. Then, load 2.5 mL of degassed beer or wort sample to the equilibrated column. After running through the column, load 3.5 mL of distilled deionized water in the column to elute desalted beer or wort proteins (*see Note 6*).
3. Determine the protein concentration of the desalted beer or wort protein solution using the Bradford method with bovine serum albumin as a standard [53]. For beer, no dilution is required. For wort, dilute the sample three- or five-fold with distilled deionized water if required (*see Note 7*). To determine the protein concentration, use 0.5 mL desalted samples. Dilute standard BSA to 0.5, 0.4, 0.3, 0.2, 0.1, and 0.05 mg/mL. Apply 10 μ L of the diluted standard BSA and triplicate samples to the 96-well plate. Then add 200 μ L of protein assay solution to each sample. After reacting for 10 min at room temperature, measure the absorbance at 595 nm using a

VARIOSKAN microtiter plate reader. Based on the standard curve of BSA, determine the protein concentration of the sample.

4. Stock the remaining desalted beer or wort sample in appropriate tubes, such as a 15 mL tube, and freeze dry it.
5. Add 0.5 mL of dissolving buffer to freeze-dried samples, and dissolve the sample completely using a vortex mixer.
1. Add an appropriate volume of sample solution containing 100 µg protein (see Note 8), 6 µL of IPG buffer and 10 µL of 0.1 % bromophenol blue solution to a 1.5 mL tube. Then, add dissolving buffer to obtain a total volume of 300 µL, and mix briefly using a vortex mixer and then completely spin down.
2. Apply a total of 300 µL of the sample solution to an IPG dry strip, and then rehydrate the IPG dry strip overnight.
3. Perform first dimensional gel electrophoresis, isoelectric focusing using a Multiphor II system. The electrophoresis condition depends on the IPG dry strip used. For details, see the manufacturer's instructions.
4. Equilibrate the IPG strip using equilibration buffer containing 10 mg/mL dithiothreitol for 15 min, followed by equilibration buffer containing 25 mg/mL iodoacetamide (Wako).
5. Perform 2D SDS-PAGE using a precast XL 12–14 % gradient gel using a Multiphor II system. For details, see the manufacturer's instructions.
6. Stain the gel with Silver Staining Kit, Protein. For mass spectrometry analysis, stain the gel with Silver Stain MS kit. For detailed procedures, see the manufacturer's instructions.

3.2.3 Mass-Spectrometry Analysis and Database Search

1. Excise the required protein spot on 2DE gel without contamination (see Note 9), and then stock the gel piece in a 1.5 mL tube.
2. Add 100 µL of decolorizing buffer to the gel and incubate for 5 min.
3. Remove decolorizing buffer, and then add 100 µL of new decolorizing buffer to the gel. Repeat this procedure three times.
4. Add 100 µL of ultrapure water to the gel and incubate for 10 min.
5. Remove the ultrapure water, and then add 100 µL of new ultrapure water to the gel. Repeat this procedure nine times.
6. Add 50 µL of 100 % acetonitrile and incubate for 15 min.
7. After removing the acetonitrile, add 50 µL of 50 mM ammonium hydrogen carbonate buffer to the gel and then incubate for 15 min.
8. After removing the ammonium hydrogen carbonate buffer, add 100 µL of 100 % acetonitrile.

9. Add 30 μ L of 10 mM dithiothreitol buffer to the gel and incubate for 10 min at 60 °C.
10. Remove the dithiothreitol buffer and then leave the gel for 20 min.
11. Add 30 μ L of 50 mM moniodoacetic acid to the gel and incubate 15 min.
12. Remove the moniodoacetic acid.
13. Add 40 μ L of 50 mM ammonium hydrogen carbonate buffer and incubate for 15 min.
14. After removing the ammonium hydrogen carbonate buffer, add 50 μ L of 100 % acetonitrile and incubate for 15 min. Repeat this step.
15. Volatilize acetonitrile using nitrogen gas.
16. Add trypsin solution to the gel and incubate for 2 h at 37 °C.
17. Add 10 μ L of ultrapure water and then incubate for 2 h at 37 °C again.
18. Add 10 μ L of 25 mM ammonium hydrogen carbonate buffer and incubate for 10 min.
19. Add 20 μ L of 100 % acetonitrile and incubate for 10 min.
20. Add 20 μ L of 10 % formic acid and incubate for 10 min.
21. Add 30 μ L of 100 % acetonitrile and incubate for 15 min.
22. Desalt the resultant sample solution using Zip-Tip and then elute with 80 % acetonitrile.
23. Analyze the sample using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI TOF-MS) using Voyager-DE STR.
24. Calibrate internal mass using trypsin autolysis products.
25. Identify the protein species by peptide mass fingerprinting (PMF) on the nonredundant amino acid database of the National Center for Biotechnology Information (NCBI-nr) using the MASCOT search engine [54] (see Notes 10 and 11).

3.3 Notes

1. Stock 5 mL of 8 M urea with 2 % CHAPS solution in 15 mL tube in -20 °C freezer. When used, add 14 mg of dithiothreitol (Wako) to the solution, and then completely dissolve.
2. IPG buffer and IPG dry strip depend on the objective pI of 2DE.
3. Stock 10 mL of the buffer in 15 mL tube in -20 °C freezer.
4. For one trial of 2DE analysis, 2.5 mL of beer sample is needed. Therefore, 50 mL of beer sample poured into the flask is enough.
5. Several proteins are specifically contained in the foam fraction. Therefore, a complete degas is needed to analyze total beer proteins
6. For more information, refer to the manufacturer's instructions.

7. Dilution rate depends on the sample used. Determine dilution rate of the sample if needed.
8. Calculate the volume of the solution needed based on the protein concentration of the sample analyzed by the Bradford method.
9. To avoid contamination, gloves and a mask are necessary.
10. The following parameters are used for searches in MASCOT. Database: NCBI nr; Taxonomy: all entries; enzyme: trypsin; number of missed cleavages: up to one missed cleavage; fixed modification: carbamidomethyl cysteine; variable modification: oxidation of methionine; peptide tolerance: below 50 ppm; mass values: MH⁺. The protein is considered a positive candidate when it satisfies all of the following three criteria: (1) significant MASCOT score ($P < 0.05$); (2) more than four peptides matched; and (3) sequence coverage of more than 15 %.
11. When candidate is not identified in NCBI-nr, the barley gene index (<http://compbio.dfci.harvard.edu/tgi/>) and HarvEST #31 (<http://harvest.ucr.edu/>), are appropriate using the same parameters to attempt protein identification.

References

1. Anderson RP, Degano P, Godkin AJ et al (2000) *In vivo* antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nat Med* 6:337–342
2. Shan L, Molberg Ø, Parrot I et al (2002) Structural basis for gluten intolerance in celiac sprue. *Science* 297:2275–2279
3. Tye-Din JA, Stewart JA, Dromey JA et al (2010) Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci Transl Med* 2:41–51
4. Green PHR, Jabri B (2006) Coeliac disease. *Annu Rev Med* 57:207–221
5. Marshall T, Williams KM (1987) High resolution two-dimensional electrophoresis of the proteins and macromolecular constituents of beer and wine. *Electrophoresis* 8:493–495
6. Dale CJ, Young TW (1988) Fractionation of high molecular weight polypeptides from beer using two dimensional gel electrophoresis. *J Inst Brew* 94:28–32
7. Williams KM, Marshall T (1995) Protein composition of beer as revealed by high-resolution two-dimensional electrophoresis. *Anal Proc* 32:25–28
8. Perrocheau L, Rogniaux H, Boivin P et al (2005) Probing heat-stable water-soluble proteins from barley to malt and beer. *Proteomics* 5:2849–2858
9. Iimure T, Nankaku N, Hirota N et al (2010) Construction of a novel beer proteome map and its use in beer quality control. *Food Chem* 118:566–574
10. Konečná H, Müller L, Dosoudilová H et al (2012) Exploration of beer proteome using OFFGEL prefractionation in combination with two-dimensional gel electrophoresis with narrow pH range gradients. *J Agric Food Chem* 60:2418–26
11. Picariello G, Bonomi F, Iametti S et al (2011) Proteomic and peptidomic characterization of beer: immunological and technological implications. *Food Chem* 124:1718–1726
12. Colgrave ML, Goswami H, Howitt CA et al (2012) What is in a beer? Proteomic characterization and relative quantification of hordein (gluten) in beer. *J Proteome Res* 11:386–396
13. Klose C, Thiele F, Arendt EK (2010) Changes in the protein profile of oats and barley during brewing and fermentation. *J Am Soc Brew Chem* 68:119–124
14. Klose C, Mauch A, Wunderlich S et al (2011) Brewing with 100% oat malt. *J Inst Brew* 117:411–421
15. Bobálová J, Salplachta J, Chmelík J (2008) Investigation of protein composition of barley by gel electrophoresis and MALDI mass spectrometry with regard to the malting and brewing process. *J Inst Brew* 114:22–26

16. Hao J, Li Q, Dong J et al (2006) Identification of the major proteins in beer foam by mass spectrometry following sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *J Am Soc Brew Chem* 64:166–174
17. Matsumoto T, Tanaka T, Sakai H et al (2011) Comprehensive sequence analysis of 24,783 barley full-length cDNAs derived from 12 clone libraries. *Plant Physiol* 156:20–28
18. Saisho D, Myoraku E, Kawasaki S et al (2007) Construction and characterization of a bacterial artificial chromosome (BAC) library from the Japanese malting barley variety 'Haruna Nijo'. *Breeding Sci* 57:29–38
19. Fasoli E, Aldini G, Regazzoni L et al (2010) Les Maîtres de l'orge: the proteome content of your beer mug. *J Proteome Res* 9:5262–5269
20. Weber D, Cléroux C, Godefroy SB (2009) Emerging analytical methods to determine gluten markers in processed foods – method development in support of standard setting. *Anal Bioanal Chem* 395:111–117
21. Hejgaard J, Kaersgaard P (1983) Purification and properties of the major antigenic beer protein (Antigen 1) of barley origin. *J Inst Brew* 89:402–410
22. Silva F, Nogueira LC, Gonçalves C et al (2008) Electrophoretic and HPLC methods for comparative study of the protein fractions of malts, worts, and beers produced from Scarlett and Prestige barley (*Hordeum vulgare* L.) varieties. *Food Chem* 106:820–829
23. Iimure T, Nankaku N, Kihara M et al (2012) Proteome analysis of the wort boiling process. *Food Res Intern* 45:262–271
24. Osman AM, Coverdale SM, Onley-Watson K et al (2003) The gel filtration chromatographic profiles of proteins and peptides of wort and beer: Effects of processing -malting, mashing, kettle boiling, fermentation and filtering. *J Inst Brew* 109:41–50
25. Bamforth CW (1985) The foaming properties of beer. *J Inst Brew* 91:370–383
26. Evans DE, Sheehan MC (2002) Don't be fobbed off: the substance of beer foam-A review. *J Am Soc Brew Chem* 60:47–57
27. Evans DE, Hejgaard J (1999) The impact of malt derived proteins on beer foam quality. Part I. The effect of germination and kilning on the level of protein Z4, protein Z7 and LTP1. *J Inst Brew* 105:159–169
28. Iimure T, Kihara M, Ichikawa S et al (2011) Development of DNA markers associated with beer foam stability for barley breeding. *Theor Appl Genet* 122:199–210
29. Maeda K, Yokoi S, Kamada K et al (1991) Foam stability and physicochemical properties of beer. *J Am Soc Brew Chem* 49:14–18
30. Sørensen SB, Bech LM, Muldbjerg M et al (1993) Barley lipid transfer protein 1 is involved in beer foam formation. *Tech Q Master Brew Assoc Am* 30:136–145
31. van Nierop SNE, Evans DE, Axcell BC et al (2004) Impact of different wort boiling temperatures on the beer foam stabilizing properties of lipid transfer protein 1. *J Agric Food Chem* 52:3120–3129
32. Asano K, Shinagawa K, Hashimoto N (1982) Characterization of haze-forming proteins of beer and their roles in chill haze formation. *J Am Soc Brew Chem* 40:147–154
33. Perrocheau L, Bakan B, Bovin P et al (2006) Stability of barley and malt lipid transfer protein 1 (LTP1) toward heating and reducing agents: Relationships with the brewing process. *J Agric Food Chem* 54:3108–3113
34. Iimure T, Kimura T, Araki S et al (2012) Mutation analysis of barley malt protein Z4 and protein Z7 on beer foam stability. *J Agric Food Chem* 60:1548–1554
35. Okada Y, Iimure T, Takoi K et al (2008) The influence of barley malt protein modification on beer foam stability and their relationship to the barley dimeric α -amylase inhibitor-I (BDAI-I) as a possible foam-promoting protein. *J Agric Food Chem* 56:1458–1464
36. Iimure T, Takoi K, Kaneko T et al (2008) Novel prediction method of beer foam stability using protein Z, barley dimeric α -amylase inhibitor-1 (BDAI-1) and yeast thioredoxin. *J Agric Food Chem* 56:8664–8671
37. Leisegang R, Stahl U (2005) Degradation of a foam-promoting barley protein by a proteinase from brewing yeast. *J Inst Brew* 111:112–117
38. Kogin A, Fukui H, Furukubo S et al (1999) Regulation of protease activity in beer. *Tech Q Master Brew Assoc Am* 36:67–70
39. Kondo H, Yomo H, Furukubo S et al (1999) Advanced method for measuring Proteinase A in beer and application to brewing. *J Inst Brew* 105:293–300
40. Gan ZR (1991) Yeast thioredoxin genes. *J Biol Chem* 266:1692–1696
41. Pedrajas JR, Kosmidou E, Miranda-Vizuete A et al (1999) Identification and functional characterization of a novel mitochondrial thioredoxin system in *Saccharomyces cerevisiae*. *J Biol Chem* 274:6366–6373
42. Dreyer T, Biedermann K, Ottesen M (1983) Yeast proteinase in beer. *Carlsberg Res Commun* 48:249–253
43. Guido LF, Rodrigues PG, Rodrigues JA et al (2004) The impact of the physiological condition of the pitching yeast on beer flavor stability: an industrial approach. *Food Chem* 87:187–193
44. Leiper KA, Stewart GG et al (2003) Beer polypeptides and silica gel part I. Polypeptides involved in haze formation. *J Inst Brew* 109:57–72

45. Siebert KJ, Lynn PY (1997) Mechanisms of beer colloidal stabilization. *J Am Soc Brew Chem* 55:73–78
46. Robinson LH, Juttner J, Milligan A et al (2007) The identification of a barley haze active protein that influences beer haze stability: cloning and characterization of the barley SE protein as a barley trypsin inhibitor of the chloroform/methanol type. *J Cereal Sci* 45:343–352
47. Iimure T, Nankaku N, Watanabe-Sugimoto M et al (2009) Identification of novel haze-active beer proteins by proteome analysis. *J Cereal Sci* 49:141–147
48. Jin B, Li L, Li B et al (2009) Proteomics study of silica eluent proteins in beer. *J Am Soc Brew Chem* 67:183–188
49. Jin B, Li L, Feng ZC et al (2011) Investigation of hordeins during brewing and their influence on beer haze by proteome analysis. *J Food Biochem* 35:1522–1527
50. Sheehan MC, Skerritt JH (1997) Identification and characterisation of beer polypeptides derived from barley hordeins. *J Inst Brew* 103:297–306
51. Leiper KA, Stewart GG, McKeown IP (2003) Beer polypeptides and silica gel. Part II. Polypeptides involved in foam formation. *J Inst Brew* 109:73–79
52. Evans DE, Robinson LH, Sheehan MC et al (2003) Application of immunological methods to differentiate between foam-positive and haze-active proteins originating from malt. *J Am Soc Brew Chem* 61:55–62
53. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254
54. Perkins DN, Pappin DJC, Creasy DM et al (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20:3551–3567

Chapter 52

Precipitation of Champagne Base Wine Proteins Prior to 2D Electrophoresis

Clara Cilindre

Abstract

Numerous methods have been employed to depict the protein content of wines. Among them, two-dimensional electrophoresis (2D-E) presents a powerful resolution, but has been poorly applied to wine. Furthermore, 2D-E was coupled with various extraction methods of proteins without any reference method for wine. Here, we describe a rapid method to extract proteins from a champagne base wine through ultrafiltration followed by precipitation with ethanol and trichloroacetic acid. More than 50 spots were visualized on 2D-gels (7 cm, pH 3–6) by colloidal Coomassie Brilliant Blue staining.

Key words Wine proteomics, Champagne, Grape, Chardonnay, Ultrafiltration, Precipitation, 2D-electrophoresis

1 Introduction

Proteins are not the major components of wine, although they are essential compounds, contributing to many organoleptic characteristics. Indeed, wine proteins are implied in the foaming properties of sparkling wines [1, 2], the interaction with wine volatile compounds [3], the stabilization of tartaric salts [4], the decrease in wine astringency [5] and, unfortunately, the formation of haze in white wines [6, 7]. Various factors are likely to modify the protein content of a wine, as for example, the grape variety [8], the infection of grape berries by a widespread phytopathogen *Botrytis cinerea* [9, 10], the yeast strain employed for alcoholic fermentation [11], the aging on lees [12], and also various fining treatments [1, 13–16].

Since the end of the 90s, a broad range of methods for the separation of specific or total proteins have been developed and applied to wine [17], including chromatographic techniques [6, 13, 16, 18–20], capillary gel electrophoresis [18, 21], combinatorial peptide ligand libraries [22] or one-dimensional polyacrylamide gel electrophoresis [1, 6, 14, 15, 23, 24]. 2D-E has been

first introduced to wine in 1987 by Marshall and Williams [25] and has also been widely used in food proteomics [26]. However, this highly resolutive method has been poorly applied for wine protein analysis to date [7, 9, 10, 15, 27, 28].

Among these studies, different methods were employed to extract soluble proteins from wine before 2D-E analysis. In 2006, Okuda et al. [28] have undergone successive steps to extract soluble proteins from a Chardonnay wine. Briefly, the wine was concentrated by rotary evaporation at 40 °C, precipitated by ammonium sulfate 80 %, ultracentrifugated, dialyzed, and finally the lyophilized proteins were fractionated by Sephadex G-100 chromatography before 2D-E (16 cm, pH 3–6). The main disadvantage of this method might be the numerous steps that are time consuming, though more than 300 spots were visualized by Coomassie Brilliant Blue (CBB-R250) staining. More recently, Batista et al. (2009) [7] purified the proteins from an Arinto white wine by FPLC cation exchange chromatography, subsequently desalted and lyophilized before 2D-E (13 cm, pH 3–10). A few number of spots was retrieved on 2D gels, likely due to the low sensitivity of the staining method employed, i.e., CBB-R250 or an insufficient amount of protein loaded (not specified). More recently, Sauvage et al. (2010) [15] employed 2D electrophoresis (18 cm, pH 3–10) to monitor the impact of enological treatments on the protein content of a Chardonnay wine. In this latter study, a partial purification was undergone through adsorption of polyphenols on Fractogel, removal of polysaccharides (>150 kDa) and then, concentration of wine proteins (between 5 and 150 kDa) by successive ultrafiltrations. Only ten spots were visualized by colloidal CBB (G-250) staining, despite the high amount of protein loaded on the IPG strips (i.e., 300 µg).

Here, we present an alternative method to extract proteins from a Champagne base wine with minor modifications as compared to the protocol published in our previous studies [9, 10]. Briefly, a Champagne base wine (Chardonnay) was subjected to an ultrafiltration in order to concentrate ten times the wine protein content. The retentate was precipitated in ethanol with 15 % (w/v) of TCA (ET/TCA) and finally, the protein pellet was solubilized in an appropriate rehydration buffer. The wine protein sample was separated and analyzed by 2D-E, using two IPG strips (7 cm) with different pH gradients, pH 3–10 and pH 3–6. Thus, we were able to visualize around 30 and 50 spots on pH 3–10 and pH 3–6 gradients, respectively, by using a sensitive colloidal CBB-G250 staining [29]. Our precipitation method is easy to conduct, well adapted to wine and has been applied successfully to must and champagne (*data not shown*). Another advantage is the use of 7 cm IPG strips, instead of 18 cm [9, 10] that allows a lower protein load, without disturbing the subsequent MS analysis. It is an important fact to take into account, since Champagne wines generally contain a low

amount of proteins (<20 mg/L). On the other hand, 2D-E is still a method of choice to have a detailed overview of the proteins present in a wine and, is therefore better suited for a comparative analysis than 1D-electrophoresis.

2 Materials

The following solutions are prepared from high purity chemicals (i.e., electrophoresis grade) and ultrapure deionized water (e.g., 18.2 MΩ). Disposable gloves are worn during each step to avoid contaminations.

2.1 Wine Samples

A Chardonnay base wine elaborated by the Cooperative Nogent l'Abbesse according to the traditional method was tested. The protein content was estimated around 10 mg/L (Bovine Serum Albumin was used as a protein standard) (*see Note 1*).

2.2 Isoelectro-focusing Components

1. Standard rehydration solution (1×): 7 M Urea, 2 M Thiourea, 4 % (w/v) CHAPS, 0.5 % (v/v) IPG Buffer 3–10, traces of Bromophenol blue, 60 mM DTT (*see Note 2*).
2. Glycerol (Sigma Chemical Company, St. Louis, MO, USA). Store at room temperature (*see Note 3*).
3. Mineral oil (Sigma Chemical Company, St. Louis, MO, USA). Store at room temperature.
4. Electrode wicks (Bio-Rad).
5. Readystrips™ IPG strips, pH 3–10 and pH 3–6, 7 cm (Bio-Rad).
6. IPGphor Isoelectric Focusing System (Amersham Pharmacia Biotech).

2.3 SDS Polyacrylamide Gel Components

1. 1 M Tris–HCl, pH 8.8, buffer: Weigh 121.1 g of Tris and transfer to a glass beaker. Add water to a volume of 900 mL. Mix and adjust to pH 8.8 with HCl 6 N. Transfer the solution to a graduated flask and make up to 1 L with water. Store at 4 °C.
2. Acrylamide–Bis-acrylamide 37.5:1, 40 % (Bio-Rad). Store at 4 °C.
3. SDS 10 % (w/v) solution in water. Store at room temperature.
4. Ammonium persulfate (APS) 10 % (w/v) solution in water (*see Note 4*).
5. TEMED, store at room temperature.
6. 1.2 M Bis-Tris/ 0.8 M HCl: Dissolve 12.55 g of Bis-Tris in 25 mL of water. Add 3.94 g of HCl 37 % (12 N) and adjust to 50 mL with water. Store at 4 °C.

7. Overlay agarose solution: 1 % (w/v) agarose low-melting, 0.2 % SDS, 0.15 M Bis-Tris/0.1 M HCl, 1 μ g/ μ L bromophenol blue. Store at room temperature.
8. SDS Running buffer (1 \times): 25 mM Tris-HCl, pH 8.3, 192 mM glycine, 0.1 % (w/v) SDS. The buffer is prepared from a 10 \times stock commercial solution (Bio-Rad) and is refrigerated at 4 °C before use.
9. Mini-Protean 3 system, Spacer plates with 1.0 mm spacers (Bio-Rad).

2.4 Equilibration

1. Reduction solution: 50 mM Tris-HCl, pH 8.8, 6 M urea, 30 % (v/v) glycerol, 2 % (w/v) SDS. Add 1 % (w/v) DTT prior to use (see Note 5).
2. Alkylation solution: 50 mM Tris-HCl, pH 8.8, 6 M urea, 30 % (v/v) glycerol, 2 % (w/v) SDS, traces of bromophenol blue. Add 2.5 % (w/v) iodoacetamide prior to use (see Note 5).

2.5 Wine Protein Extraction

1. Ultrafiltration device with a 10 kDa membrane molecular weight cut-off (MWCO) (see Note 6).
2. Ethanol/TCA 15 % (w/v) solution: Prepare a fresh solution before use by adding 15 g of trichloroacetic acid to 100 mL of glacial ethanol (see Note 7).

3 Methods

3.1 Wine Ultrafiltration

1. Using an ultrafiltration device (see Note 8), the wine sample is concentrated ten times.
2. The ultrafiltrate is discarded, and deionized water is added up to the initial wine volume to dialyze the wine retentate.
3. Dialysis against deionized water is repeated one more time (step 2).
4. The final wine retentate is immediately precipitated or stored at -80 °C.

3.2 Wine Protein Precipitation

1. Mix 5 mL of wine retentate with 40 mL of ethanol/TCA 15 % (w/v) freshly prepared.
2. Store the sample at -20 °C during 1 h (see Note 9).
3. Centrifuge 10 min at 9,500 \times g and 4 °C, then discard the supernatant.
4. Wash the pellet with 20 mL of ethanol, to remove TCA (see Note 10) and centrifuge as described above. Repeat this step again.
5. Dry the resulting pellet with nitrogen gas to remove ethanol and solubilize the wine protein extract with 1 mL of standard rehydration solution (see Note 11).

6. Determine the protein concentration of the extract (*see Note 12*).
7. Prepare aliquots of the wine protein extract and store them at -80°C .

3.3 IEF Parameters

The IPG strips are rehydrated and focused in an IPGphor IsoElectric Focusing System (GE Healthcare).

1. For in gel active rehydration, load approximately 20 μg of proteins. Mix the corresponding volume of wine protein extract, with 12.5 μL of glycerol and the remaining volume of rehydration solution for a final volume of 125 μl (*see Note 13*).
2. Pipet slowly the sample solution in the strip holder.
3. Place carefully the gel of the IPG strip in contact with the sample solution and proceed to the active rehydration of the IPG strip with the IPGphor, at 30 V for 15 h.
4. At the end of the active rehydration step, place electrode paper wicks between the IPG strips and the wire electrodes. Wet one paper with ultrapure deionized water and insert the paper between the IPG strip and the cathode. Wet another paper with the rehydration solution and insert it between the IPG strip and the anode (identified by +). Cover the IPG strip with mineral oil and initiate the focusing steps.
5. The following running conditions are applied according to the pH gradient:
 - pH 3–10, linear, 7 cm: 300 V for 15 min, 4,000 V linear ramp for 2 h, 4,000 V for 10,000 Vh, 300 V as an additional step.
 - pH 3–6, linear, 7 cm: 50 V for 15 min, 250 V for 15 min, 4,000 V linear ramp for 2 h, 4,000 V for 10,000 Vh, 300 V as an additional step.
6. At the end of IEF, remove the focused IPG strips from the strip holder and place it, gel-side up, on a sheet of paper to remove the excess of oil.
7. IPG strips can be used for the following equilibration steps or stored at -80°C until use.

3.4 Equilibration of Focused Proteins

1. Place the IPG strip in a 10 mL screw-cap tube containing 5 mL of reduction solution. Incubate 15 min at room temperature, with gentle shaking.
2. Remove the IPG strip from the reduction solution and place it in a tube containing 5 mL of alkylation solution. Incubate 15 min at room temperature with gentle shaking.
3. Blot the IPG strip, gel-side up, on a sheet of paper to remove the excess of liquid and apply immediately the IPG strip on the top of the SDS polyacrylamide gel (*see Note 14*).

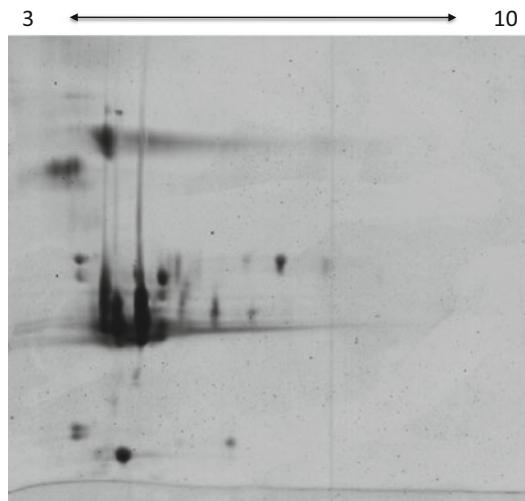


Fig. 1 2D gel electrophoresis of champagne base wine proteins stained with colloidal Coomassie Brilliant Blue. Wine proteins were separated by IEF in 7 cm long pH 3–10 IPG strips, followed by SDS-PAGE in vertical 12 % gels. Champagne base wine proteins are mainly located between pH 3 and 5

3.5 Second Dimension (SDS-PAGE)

1. Prepare the 12 % polyacrylamide gels (composition of two gels, 1 mm width, for Mini-Protean 3 system, Bio-Rad) by mixing slowly 4.875 mL of deionized water, 5.4 mL of 1 M Tris-HCl, pH 8.8 buffer, 4.5 mL Acrylamide–Bisacrylamide 40 % ratio 37.5:1 (Bio-Rad), 150 μ L of SDS 10 % (w/v) in a 25 mL glass beaker. Add 75 μ L of APS 10 % (w/v) and 7.5 μ L of TEMED. Cast gel in a 7.25 cm \times 10 cm \times 1 mm gel cassette. Allow space (\approx 1 cm) for the IPG strip and cover the gel with a few milliliters of water-saturated *n*-butanol or deionized water.
2. After polymerization (\approx 1 h at room temperature), discard water-saturated *n*-butanol and wash thoroughly with deionized water.
3. Pipet the overlay agarose solution, previously warmed, on the top of the vertical gel and place immediately the IPG strip in close contact with the SDS polyacrylamide gel (*see Note 15*).
4. Let the overlay agarose solution to set at room temperature for at least 1 min.
5. Run the electrophoresis at 30 V during 20 min, then at 175 V until the front dye reached 5 mm from the bottom of the gel.
6. The gels were stained with colloidal Coomassie Brilliant Blue according to Candiano et al. [29]. We were able to visualize around 30 and 50 spots on pH 3–10 (Fig. 1) and pH 3–6 (Fig. 2) gradients, respectively.

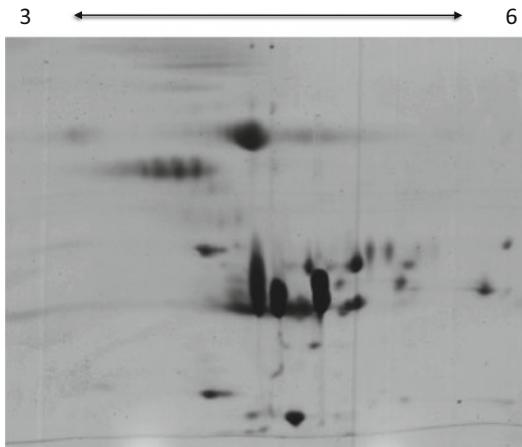


Fig. 2 2D gel electrophoresis of champagne base wine proteins stained with colloidal Coomassie Brilliant Blue. Wine proteins were separated by IEF in 7 cm long pH 3–6 IPG strips, followed by SDS-PAGE in vertical 12 % gels

3.6 Gel image Analysis

Images of the gels were acquired at 63.5 μm resolution using the GS-800 scanner and Quantity One v 4.62 software (Bio-Rad, Hercules, CA). Spot detection and quantification were performed using the PDQuest Basic 8.0.1 software (Bio-Rad, Hercules, CA). Reproducibility of 2D gels was assayed by running three replicates. Selected protein features were modelled as Gaussians, and the relative optical densities (OD), i.e., the feature OD divided by the total OD over the whole image, were computed. Means \pm standard deviation ($n=3$) were calculated.

4 Notes

1. A commercial wine or a wine elaborated in a laboratory can also be used. Wines made with red grape varieties, and thus containing an elevated content of phenolic compounds have not been tested by this method. To eliminate these interfering compounds an additional step might be undergone by adding PVPP.
2. Add DTT prior to use. This solution is stored in aliquots at -80°C .
3. Aliquots of glycerol (12.5 μL correspond to 10 % (v/v) of the final rehydration volume. 125 μL is the final rehydration volume needed for a 7 cm length strip) can be prepared and stored at -20°C .
4. Prepare the solution freshly.
5. 10 mL aliquots of the solution, without DTT or iodoacetamide, are stored at -20°C . Dissolution of DTT and iodoacetamide can be difficult and may be achieved by a thorough shaking.

6. Use a centrifugal filter device to ultrafiltrate 15 mL of wine or, if feasible, higher volume ultrafiltration systems, in order to obtain a sufficient volume of retentate.
7. Ethanol can be stored at -20 °C. TCA must be added prior use since the solution is unstable at -20 °C and might become yellow after storage. TCA is extremely hazardous, thus it is important to avoid inhalation, skin and eyes contacts. Furthermore, TCA is highly hygroscopic, thus it is important to store the product in an appropriate place in order to weigh the correct amount of TCA.
8. A hydrophilic polysulfone membrane (30 UFIB, Setric Genie Industriel, SGI, France) with a 10 kDa molecular weight cut-off was used. Wine ultrafiltration was carried out at 12 °C using a crossflow filtration module coupled with the SGI Hi-Flow system (pumping system plus glass tank). Alternative ultrafiltration device (with an equivalent molecular weight cut-off) might be used, as for example centrifugal filter devices.
9. Alternatively, the solution can be left overnight at -20 °C. In this case, TCA might be more difficult to eliminate and an additional washing step might be needed (step 4, Subheading 3.2). The mentioned volumes can be modified, even if it is better to keep these proportions: one volume of wine concentrate diluted with eight volumes of ethanol/TCA.
10. The volume of ethanol must be at least 20 mL, to maximize the removal of TCA. Precipitation with TCA has both advantage and disadvantage. It allows proteases inactivation, reduces salt concentration but interferes with IEF leading to horizontal streaking.
11. In this experiment, the protein pellet was mixed with 1 mL of rehydration buffer. A lower volume, as for example 500 µL, might be used if all the proteins are well solubilized. It also depends on the initial protein concentration of the wine sample.
12. The protein content was estimated with the Bradford method as modified by Fey et al. [30], using Bovine Serum Albumin as protein standard.
13. The protein load is adapted according to the length of the strip and the staining method used.
14. The second dimension gel can be prepared 24 h in advance and stored in the casting gel cassette with a layer of deionized water to exclude the drying of the gel.
15. Use forceps and a spatula to place correctly the IPG strip. Avoid trapping air bubbles between the IPG strip and the second dimension gel. The gel-side of the IPG strip must be in contact with the little glass plate.

Acknowledgment

We express our gratitude to the Association Recherche Oenologie Champagne et Université (Reims, France) for supporting grant. The authors are also indebted to the Région Champagne-Ardenne, the Ville de Reims, the Union des Oenologues de Champagne, and the Conseil Général de la Marne for supporting our research.

References

1. Dambrouck T, Marchal R, Cilindre C et al (2005) Determination of the grape invertase content (using PTA-ELISA) following various fining treatments *vs* changes in the total protein content of wine. relationships with wine foamability. *J Agric Food Chem* 53: 8782–8789
2. Moreno-Arribas V, Pueyo E, Nieto FJ et al (2000) Influence of the polysaccharides and the nitrogen compounds on foaming properties of sparkling wines. *Food Chem* 70: 309–317
3. Guichard E (2006) Flavour retention and release from protein solutions. *Biotechnol Adv* 24:226–229
4. Moine-Ledoux V, Perrin A, Paladin I et al (1997) First result of tartaric stabilization by adding mannoproteins (Mannostab). *J Int Sci Vigne Vin* 31:23–31
5. Escot S, Feuillat M, Dulau L et al (2001) Release of polysaccharides by yeasts and the influence of released polysaccharides on colour stability and wine astringency. *Aust J Grape Wine Res* 7:153–159
6. Esteruelas M, Poinsaut P, Sieczkowski N et al (2009) Characterization of natural haze protein in Sauvignon white wine. *Food Chem* 113:28–35
7. Batista L, Monteiro S, Loureiro VB (2009) The complexity of protein haze formation in wines. *Food Chem* 112:169–177
8. Monteiro S, Picarra-Pereira MA, Teixeira AR et al (2003) Environmental conditions during vegetative growth determine the major proteins that accumulate in mature grapes. *J Agric Food Chem* 51:4046–4053
9. Cilindre C, Castro AJ, Clement C et al (2007) Influence of *Botrytis cinerea* infection on Champagne wine proteins (characterized by two-dimensional electrophoresis - immunodetection) and wine foaming properties. *Food Chem* 103:139–149
10. Cilindre C, Jégou S, Hovasse A et al (2008) Proteomic approach to identify Champagne wine proteins as modified by *Botrytis cinerea* infection. *J Proteome Res* 7:1199–1208
11. Younes B, Cilindre C, Villaume S et al (2011) Evidence for an extracellular acid proteolytic activity secreted by living cells of *Saccharomyces cerevisiae* PIR1: Impact on grape proteins. *J Agric Food Chem* 59:6239–6246
12. Alcaide-Hidalgo JM, Martinez-Rodriguez AJ, Martin-Alvarez PJ et al (2008) Influence of the elaboration process on the peptide fraction with angiotensin I converting enzyme inhibitor activity in sparkling wines and red wines aged on lees. *Food Chem* 111:965–969
13. Vanrell G, Canals R, Esteruelas M et al (2007) Influence of the use of bentonite as a riddling agent on foam quality and protein fraction of sparkling wines (Cava). *Food Chem* 104: 148–155
14. Weber P, Steinhart H, Paschke A (2009) Determination of the bovine food allergen casein in white wines by quantitative indirect ELISA, SDS-PAGE, Western Blot and immunostaining. *J Agric Food Chem* 57:8399–8405
15. Sauvage FX, Bach B, Moutounet M et al (2010) Proteins in white wines: thermosensitivity and differential adsorption by bentonite. *Food Chem* 118:26–34
16. Monaci L, Losito I, Palmisano F et al (2011) Towards the quantification of residual milk allergens in caseinate-fined white wines using HPLC coupled with single-stage Orbitrap mass spectrometry. *Food Addit Contam A* 28: 1304–1314
17. Giribaldi M, Giuffrida MG (2010) Heard it through the grapevine: Proteomic perspective on grape and wine. *Proteomics* 73:1647–1655
18. Luguera C, Moreno-Arribas V, Pueyo E et al (1998) Fractionation and partial characterization of protein fractions present at different stages of the production of sparkling wines. *Food Chem* 63:465–471
19. Goncalves F, Heyraud A, de Pinho MN et al (2002) Characterization of white wine mannoproteins. *J Agric Food Chem* 50:6097–6101

20. Marangon M, Van Sluyter SC, Haynes PA et al (2009) Grape and wine proteins: their fractionation by hydrophobic interaction chromatography and identification by chromatographic and proteomic analysis. *J Agric Food Chem* 57:4415–4425
21. Rodriguez-Delgado MA, Malovana S, Montelongo FJ et al (2002) Fast analysis of proteins in wines by capillary gel electrophoresis. *Eur Food Res Technol* 214:536–540
22. D'Amato A, Fasoli E, Kravchuk AV, Righetti PG (2011) *Mehercules, adhuc Bacchus!* The debate on wine proteomics continues. *J Prot Res* 10:3789–3801
23. Kwon SW (2004) Profiling of soluble proteins in wine by nano-high-performance liquid chromatography/tandem mass spectrometry. *J Agric Food Chem* 52:7258–7263
24. Wigand P, Tenzer S, Schild H et al (2009) Analysis of protein composition of red wine in comparison with rosé and white wines by electrophoresis and high-pressure liquid chromatography-mass spectrometry (HPLC-MS). *J Agric Food Chem* 57:4328–4333
25. Marshall T, Williams KM (1987) High resolution two-dimensional electrophoresis of the proteins and macromolecular constituents of beer and wine. *Electrophoresis* 8:493–495
26. Pedreschi R, Maarten H, Lilley KS et al (2010) Proteomics for the food industry: opportunities and challenges. *Crit Rev Food Sci* 50: 680–692
27. Hsu JC, Heatherbell DA (1987) Heat-unstable proteins in wine. I. Characterization and removal by bentonite fining and heat treatment. *Am J Enol Viticul* 38:11–16
28. Okuda T, Fukui M, Takayanagi T et al (2006) Characterization of major stable proteins in chardonnay wine. *Food Sci Technol Res* 12:131–136
29. Candiano G, Bruschi M, Musante L et al (2004) Blue silver: a very sensitive colloidal coomassie G-250 staining for proteome analysis. *Electrophoresis* 25:1327–1333
30. Fey SJ, Nawrocki A, Larsen MR et al (1997) Proteome analysis of *Saccharomyces cerevisiae*: a methodological outline. *Electrophoresis* 18: 1361–1372

Chapter 53

The Minimal Information About a Proteomics Experiment (MIAPE) from the Proteomics Standards Initiative

Salvador Martínez-Bartolomé, Pierre-Alain Binz, and Juan P. Albar

Abstract

During the last 10 years, the Proteomics Standards Initiative from the Human Proteome Organization (HUPO-PSI) has worked on defining standards for proteomics data representation as well as guidelines that state the minimum information that should be included when reporting a proteomics experiment (MIAPE). Such minimum information must describe the complete experiment, including both experimental protocols and data processing methods, allowing a critical evaluation of the whole process and the potential recreation of the work. In this chapter we describe the standardization work performed by the HUPO-PSI, and then we concentrate on the MIAPE guidelines, highlighting its importance when publishing proteomics experiments particularly in specialized proteomics journals. Finally, we describe existing bioinformatics resources that generate MIAPE compliant reports or that check proteomics data files for MIAPE compliance.

Key words Proteomics, HUPO-PSI, MIAPE, Standards, Reporting, Guidelines, Semantic validator

1 Introduction

With the growing number of variations of high-throughput proteomics techniques in the last 10 years, the scientific community soon detected the need to define reporting guidelines that ensure some minimal data and meta-data quality and consistency as a requirement prior to publish protein and peptide identification data. Following this, a group of experts proposed the so-called PARIS guidelines in 2004 [1]—later revised in 2005—and then implemented in the instructions for authors by the journal Molecular and Cellular Proteomics. Then, during 2006 and 2007, the Proteomics Standards Initiative (PSI) [2] from the Human Proteome Organization (HUPO), developed the concept of the set of Minimal Information about a Proteomics Experiment (MIAPE) guidelines, based on the experience of the micro-array community with their MIAME guidelines. A MIAPE parent document [3, 4] which defines the principles and objectives of MIAPE

guidelines laying the foundations for a set of MIAPE guideline modules would be published during the following years.

The MIAPE guidelines, as defined by HUPO-PSI, aim at listing the information that should be provided while describing a proteomics experiment. As a complete proteomics experiment can be divided into smaller experimental and data analysis steps, the various HUPO-PSI working groups, composed of experts from different proteomics fields, has defined one or more MIAPE modules each covering one of these steps. The different MIAPE modules are the result of discussions between working group experts, software developers, hardware developers and end users. The documents have further followed the formal review process internally defined by PSI [3] and are available on the HUPO-PSI Web site (www.psidev.info). In most of the cases, these guidelines have also passed the review process of *Nature Biotechnology* as they have been published there as well.

Subsequently, several joint sessions between HUPO-PSI and publication committees were held [4, 5], in which journals fed back their opinions about how to ensure proteomics data quality before publishing it; how to allow the reprocessing of the data; whether to require raw data deposition in a public repository or not; and how to tackle the adoption of MIAPE guidelines in their respective instructions for authors. Most importantly, proteomics journals have been taking part in these meetings, such as *Nature Biotechnology*, *Molecular and Cellular Proteomics*, *Proteomics*, *Journal of Proteomics* and *Journal of Proteome Research*. As a result, several of these journals have been encouraging MIAPE compliance [1, 8, 10] to submit manuscripts for several years.

It is quite important for the scientific community that MIAPE guidelines become a standard that regulates the minimum information compiled by every proteomics scientific publication, since they will turn into a quality stamp to ensure a critical review and a potential repetition of the results. In addition, it is also crucial that tools are available to help authors of proteomics manuscripts as well as proteomics data producers to reach compliance with minimal technical burden.

2 The Proteomics Standards Initiative

Initiated on 2002, the Proteomics Standards Initiative (PSI) is an open community and one of the 13 scientific initiatives of HUPO that involves researchers, database providers, vendors, software developers and publishers. Its mission statement is: HUPO-PSI defines community standards for data representation in proteomics to facilitate data comparison, exchange and verification. Five main

working groups divide the vertical topics: Molecular Interactions (MI), Proteomics Informatics (PI), Mass Spectrometry (MS), Protein Separations (PS) and Protein Modifications (MOD). Additionally, a proposal for the creation of a new working group dedicated to proteomics experimental protocol standardization was presented in the last PSI meeting in San Diego (USA) on March 2012. It also has a steering committee that takes care of horizontal administrative and logistical tasks. Every year, a workshop [6–16] is organized in which the participants report on the status of the projects and split in different tracks to address specific outstanding issues. Developments and definitions are also carried out by individuals during the year, followed up and coordinated via regular teleconferences and in some cases, by face-to-face workshops on a specific topic. Additionally, progresses of HUPO-PSI standard developments are presented every year in the HUPO World Congress where a satellite meeting or a specific session [17–20] is dedicated; this also provides a wider panel to make the deliverables visible and to allow for additional discussions with the proteomics community.

HUPO-PSI is producing the following types of standards:

- *Standard formats for proteomics data representation.* These data formats are primarily XML formats defined by XML schemas (xsd) (Extensible Markup Language). In addition to the schema, a semantic validator [21] is provided in most of the cases to check the schema but, more importantly, also to ensure the proper usage of controlled vocabulary terms as well as the compliance with specific constraints (see Subheading 5.2.3). These standard data formats are meant to allow seamless exchange of data all the way along a proteomics workflow. They also, with the help of converter tools, aim to solve the problem of the difficulty of handling different proprietary data formats from various vendors while proposing a common way to represent data and meta-data. Current released formats are mzML [22], mzIdentML [23, 24], traML [25], gelML [26], and PSI-MI XML [27].
- *Controlled vocabularies.* A number of controlled vocabularies are developed to support the implementation of the data formats. They cover all concepts, procedures, materials, equipment, bioinformatics tools, etc. that are necessary to encode information into the data formats. They are integrated in *The Open Biological and Biomedical Ontologies (OBO) foundry* [28], a collaborative project for the establishment of a set of principles for ontology development in the biomedical domain. Proteomics information is coded into the standard data files using these ontology terms, which enormously facilitates the data interpretation by third-party bioinformatics tools.

- *MIAPE guidelines*. They describe the minimum information needed to appropriately report information and data about a proteomics experiment. Objectives and principles of MIAPE guidelines are defined in a parent document [4], which describes the minimum information that describes the experimental context, allows the understanding of the results and their interpretation sufficiently to permit a critical evaluation and, in principle, a potential recreation of the work. However, it is important to stress the fact that MIAPE guidelines intend neither to make any data quality judgement, nor to fix any data format for its representation or to try to establish the way to run an experiment. MIAPE guidelines have been defined in a modular way, and each MIAPE module defines the minimum information related to a certain part of the proteomics data flow. HUPO-PSI has currently produced nine different MIAPE documents/modules that will be described later in this chapter.

Data Formats, Controlled Vocabularies and MIAPEs are strongly interrelated since Controlled Vocabularies provide the way to represent proteomics data in the standard data files, and MIAPE modules define which data the file should contain to be MIAPE-compliant (Table 1).

Standards produced by HUPO-PSI are reported to the scientific community following a set of formal requirements defined by the initiative itself. Accordingly, documents can be one of the following types:

- (a) *Community practice documents*, which inform and influence the community regarding an approach or process that is considered to be widely accepted by consensus and practice in the Proteomics community.
- (b) *Informational documents*, which inform the community of an interesting and useful proteomics-related technology, architecture, framework or concept
- (c) *MIAPE (minimum information about a proteomics experiment) documents*, which inform the community as to the minimal information that should be captured about an experiment to enable its results to be clearly interpreted and validated.
- (d) *Recommendation documents*, which describe a particular technical specification or a particular set of guidelines for the application of a technical specification. Recommendations are intended to guide interoperability and promote standard approaches.

Table 1

The MIAPE module defined by each PSI working group is showed in each row, together with its corresponding standard data format (if any) and ontology

| Working group | Reporting guide lines (MIAPE) | Data exchange format | Controlled vocabulary | |
|-----------------------------|--|----------------------|------------------------------|--------------------------|
| Protein separations (PS) | Gel Electrophoresis (MIAPEGE) | gelML (vl.1) | Separation methods CV | |
| | Gel Informatics (MIAPE GI) | – | | |
| | Capillary Electrophoresis (MIAPE CE) | spML (milestone 1) | | |
| | Column Chromatography (MIAPE CC) | – | | |
| Mass spectrometry (MS) | Mass Spectrometry (MIAPE MS) | mzML (vl.1) | PSI mass spectrometry CV | Protein modifications CV |
| | – | traML (vl.1) | PSI mass spectrometry CV | Protein modifications CV |
| Proteomics informatics (PI) | Mass Spectrometry Informatics (MIAPE MSI) | mzIdentML (vl.1) | PSI mass spectrometry CV | Protein modifications CV |
| | Mass Spectrometry Quantification (MIAPE Quant) | mzQuantML (v1.0) | PSI mass spectrometry CV | Protein modifications CV |
| Molecular interactions (MI) | Molecular Interactions (MIMIx) | PSI-MI XML (v2.5) | Molecular interaction CV | |
| | Protein Affinity Reagent (MIAPAR) | PSI-PAR | Protein affinity reagents CV | |
| | Bioactive Entity (MIABE) | – | – | |

Additionally, standards are usually published in specialized journals (see references) for a greater spreading in the community.

Before being published, each one of the previously described document types must pass a formal review process also defined by the initiative itself [5]. Such a review process defines different review phases. The first review phase (Fig. 1a), is performed by the steering committee, and the second by reviewers selected by the PSI editor and, in parallel, by all the scientific community since the document is made public to receive any other comment or critical review. In the event of informational or community practice documents, the review simply consists of a public phase (Fig. 1b).

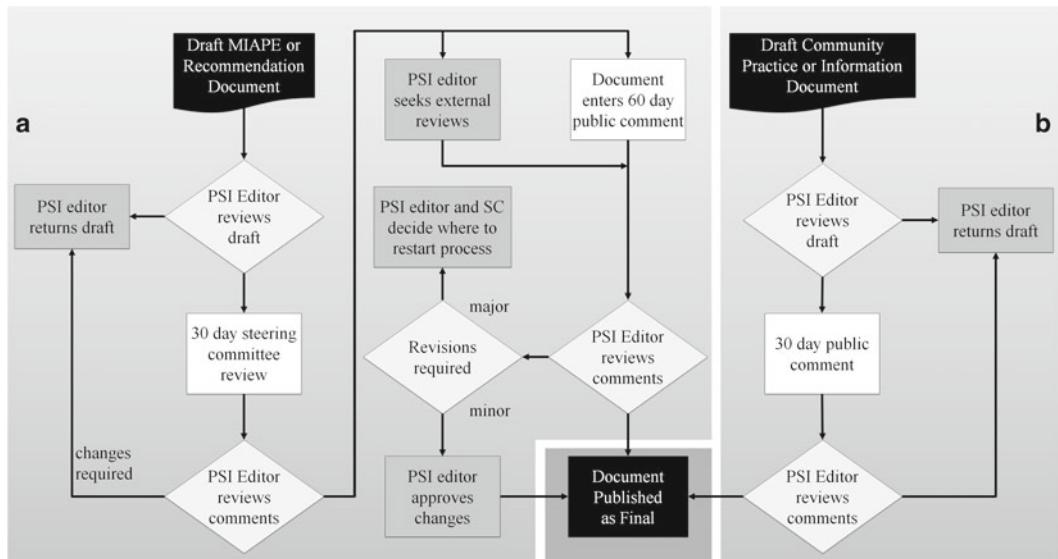


Fig. 1 Phases of the PSI formal document process by which standards are reviewed before to be published. MIAPE and recommendation documents (a); Community practice and information documents (b)

3 MIAPE Modules

As it has been already commented, MIAPE guidelines have been defined in a modular way, so each module contains a checklist of information to include when reporting a certain part of the workflow in a proteomics experiment.

In a typical proteomics experiment, we can differentiate different phases or steps: a sample separation in one or more dimensions in order to reduce its complexity; the acquisition of mass spectra by analyzing the sample in a mass spectrometer; and the bioinformatics analysis of the data generated in the previous step in order to identify or quantify proteins and/or peptides from the sample. Accordingly, with these proteomics workflow differentiations, HUPO-PSI defined the following MIAPE modules:

(a) MIAPE modules on complexity reduction of samples:

- MIAPE *gel electrophoresis* (MIAPE-GE) [29]: it describes the experimental protocol by which a sample is submitted to electrophoretic separation in a one or two dimensional gel matrix. This includes the gel matrix preparation and manufacturing, the run electrophoresis conditions, visualization techniques such as gel staining, as well as the scanning method performed to obtain the digitalized images of the gel matrixes.

- MIAPE *capillary electrophoresis* (MIAPE-CE) [30]: it describes the experimental protocol by which a sample is submitted to capillary electrophoresis, including different steps such as: pre-conditioning, injection, separation, and post-conditioning, and compiling different parameters such as temperature, detection length, pressure, voltages.
- MIAPE *column chromatography* (MIAPE-CC) [31]: it describes the experimental protocol by which a sample is submitted to a column chromatography separation step, including information such as column configuration, the appropriate mobile phase selection, the performed gradients for the chromatographic run, the sample fractioning collection.

(b) MIAPE module on mass spectrometry experiment:

- MIAPE *mass spectrometry* (MIAPE-MS) [32]: it describes the process in which a sample is analyzed by a mass spectrometer to generate the raw data files as well as the process that generates the processed spectra or peak lists. It also describes the employed equipment, including its configuration and mass spectra acquisition parameters.

(c) MIAPE modules about the bioinformatics analysis of data:

- MIAPE *gel informatics* (MIAPE-GI) [33]: it describes all the processes in which the digitalized images of the gels are analyzed by software to detect and quantify the spots or bands. This includes the experimental design description, including the description of selected group and/or replicates, the software description and parameters, and the detailed description of methods for the image alignment, detection, matching and quantification of spots, or the statistical analysis performed to determine the confidence of differential expression results.
- MIAPE *mass spectrometry informatics* (MIAPE-MSI) [34]: it describes the process by which acquired mass spectra are analyzed to identify proteins and peptides that exist in the sample. This includes all the database search engine processes, describing the software and parameters, any performed de-novo sequencing analysis, or any statistical analysis or post-processing of the identification data.
- MIAPE *mass spectrometry quantification* (MIAPE-Quant): it describes all the performed data analysis processes by way of a wide range of quantification techniques by mass spectrometry, such as chemical (iTRAQ, TMT, ICPL, ...) or metabolic (SILAC) labelling-based techniques, label-free techniques (spectral counting, chromatogram alignment, etc.) or targeted quantification techniques such as

Multiple/Selected Reaction Monitoring (SRM/MRM). While writing this chapter, the guidelines for quantification were being reviewed under the PSI formal document process.

(d) MIAPE modules on molecular interaction experiments:

- *Minimum information about a molecular interaction experiment* (MIMIx) [35]: it describes the minimum information to report about molecular interaction experiments. It includes data, such as the host organism where the interaction has been detected, the interaction detection method, the list of participants in the interaction together with their identifiers in a public database, their biological and experimental roles.
- *Minimum information about a protein affinity reagent* (MIAPAR) [36]: it describes the protein affinity reagent characterization, such as antibodies used as protein identification tools. This includes the description of molecules participating in the interaction—i.e., the reagent molecule and the target molecule—features about the interaction, such as sensitivity, selectivity, the recognized epitope, binding constants, applications; and the characterization of the interaction by information such as kinetic constants, affinity measures, or the description of the characterization method of the reagent.
- *Minimum information about a bioactive entity* (MIABE) [37]: it describes drug-target data, including information such as molecule properties, molecule production, physicochemical properties, in vitro cell-free assays, whole organism studies and pharmacokinetic studies.

4 MIAPE Guidelines Evolution

Proteomics is a rapidly evolving field. New technologies and modifications of methods are regularly appearing in the literature and on the market that widen the range of tools available to proteomics scientists. Better understanding of the strength and limitations of these processes involves also the need for more precise requirements and SOPs that target optimized quality of the obtained results. As a consequence the MIAPE guidelines need to stay aligned with this evolution. Also, as already mentioned, scientific journals have actively participated in dedicated round tables to better align the generic MIAPE technical reporting guidelines with their own submission guidelines that include quality judgment aspects in addition to descriptive data and meta-data requirements [4, 5]. Representatives from each journal have shown statistics and data about how scientists

are able to comply with these guidelines. This has highlighted that in practice some fine tuning related to required information should be addressed before that MIAPE guidelines are fully adopted by journals as integral part of their rules. For example, the percentage of publication in which raw data has been made available by submitting it to a public repository is still low, often due to technical problems (data storage and/or bandwidth of transferring large amount of data) or a lack of bioinformatics resources to do it. Other explicitly required details in MIAPE guidelines were also difficult to compile by authors, due to ignorance or to lack of clarity on some definitions in the MIAPE modules.

All that reasons have been resulted in a further review of some MIAPE modules to facilitate their adoption as much as possible to the real life. So, the PSI document process actually allows the review of documents describing standards even if they were already published previously, provided that changes are well documented and justified [3]. It is the case of MIAPE MS and MSI modules, that were published on 2008 [32, 34], and they are currently under the PSI document process, after having changed some of their content. The reasons for these modifications have been the following:

- Some sections about quantification processes were present in 2008 MIAPE MS and MSI modules. These sections were deleted when a new MIAPE module (MIAPE-Quant) about quantification data was decided to be defined.
- A number of terms were removed and replaced by more generic terms to reduce over-specification.
- Some parts of the guidelines were restructured and definitions in the appendix were modified to improve the clarity of the checklist and make it easier to comply with journals requirements and PSI data standards.

So, hopefully, MIAPE MS and MSI documents will again be released after passing by a new PSI formal document process.

But, MIAPE modules are not the only standards that have been redefined. For example the standards for proteomics data representation mzML and mzIdentML were also slightly modified (from v1.0 to 1.1) after the detection of some technical issues by some developers that were implementing these formats in their tools.

MIAPE guidelines are not sufficient to describe all kinds of publishable data and results. They cover a large part of proteomics related experiments, but do not cover other fields such as micro array experiments, RNAi experiment, a T Cell Assay, a genotyping experiment. In that respect, MIAPE is registered to the MIBBI foundry [38], where over 30 similar projects covering a wide range of fields and technologies have deposited minimal information guidelines.

5 MIAPE Tools and Implementations

While reading through the documents, a proteomics scientist might feel that complying with MIAPE guidelines—i.e., providing all required experimental and analysis parameters—is not a trivial task. This is particularly true when some part of the experiment is done as a service, for instance when samples are analyzed by Mass Spectrometry by a core facility.

There is a real need for bioinformatics resources that help to search, extract, compile, check and store such minimal information in the most automated way. Below we will describe available resources related to the management of MIAPE information.

5.1 Semi-automatic Generation and Storage of MIAPE Reports

A MIAPE-compliant document describing a proteomics experiment can be included as a materials and methods section or as supplementary material in a manuscript for publication. To compile MIAPE information and generate a legible MIAPE document, storing them in a repository is just the purpose of these two bioinformatics tools: *MIAPEGelDB* [42] developed by the Swiss Institute of Bioinformatics (SIB) and the MIAPE Generator Tool [43] developed by the bioinformatics working group of ProteoRed-ISCIII [44, 45].

Both tools provide similar functionalities, being based on a set of Web forms dedicated to compile MIAPE information. In order to create a MIAPE-compliant report the user has to go through all requested fields typing in the data manually while it is stored in a database. MIAPE documents can be later viewed in HTML or plain text formats, exported to XML, and can be accessed by a permanent URL from remote sites.

MIAPEGelDB allows for the creation of MIAPE GE documents that describe gel-based proteomics experiments. The MIAPE Generator Tool allows for the generation of several MIAPE documents such as GE, GI, MS and MSI. Moreover, all the information stored in the repository can be easily reused as templates to generate new MIAPE-compliant reports, since the same protocols and/or parameters are commonly repeated through several experiments in a given laboratory, and solely the results or the data generated by those protocols are really specific for each experiment. Additionally, it provides a permission management system that allows users to share their documents with others, assigning them different levels of operational capabilities.

Regarding MIAPE GE documents that both tools can generate, it is important to stress that both tools allow exporting that information to its corresponding standard data schema, that is, the gelML [26] data format.

5.2 Automatic Extraction and Validation of MIAPE Information

If MIAPE documents describe what information is to be provided for a given experiment, the task of the Proteomics data representation formats as defined by HUPO-PSI is to provide a mean to store this information in a standardized manner. So for example, mzML data schema [22] presents all the information related to the processing of a sample as performed by a mass spectrometer, including instrument configuration, acquisition parameters, as well as the generated data itself, such as mass spectra and chromatograms. The mzIdentML data schema [23, 24] represents all the information related to the bioinformatics analysis of mass spectra that leads to the identification of peptides and proteins, which means the description of the process in which mass spectra are assigned to amino acid sequences belonging to protein databases. This includes all parameters used when submitting spectra to a search engine and possibly validation tools, as well as the results themselves, that is, the obtaining peptide and protein lists with attributed confidence levels. In addition to PSI standards, other data formats are aiming at storing relevant information from proteomics experiments. For example, the PRIDE data schema [39, 40] was developed by the European Bioinformatics Institute (EBI) in parallel to mzML and mzIdentML for the purpose of converting data and meta-data from proteomics experiments not expressed in PSI standard formats. PRIDE has become one of the most important repositories of protein identification data by mass spectrometry, currently containing almost 300 million of spectra, more than 50 million peptides and almost 9 million proteins in its database. In a near future, the PRIDE team will natively support mzML and mzIdentML data in their importing workflows.

More and more tools and software are capable of reading and/or writing these standard data formats (see <http://psidev.info/mzml> for mzML and <http://psidev.info/mzidentml> for mzIdentML), including free-tools, open-source tools and commercial tools. Proteomics data deposition in public repositories is enormously facilitated by these standards, which is required or strongly recommended to authors when publishing MS data and results in most of the proteomics journals.

An alignment between data contained by standard files and the information required by MIAPE guidelines is therefore of crucial importance. In other words, before to store standard files in a repository, it becomes necessary to check if the encoded information is MIAPE-compliant or not, and in case of no, it is also necessary to detect which information is missing and to provide the way to include it. The *ProteoRed MIAPE Web Toolkit* (PMWTK) [41] is able to do it, providing a MIAPE quality stamp to a given file indicating that the experiment is perfectly described and that could be evaluated and potentially reproduced by other scientists.

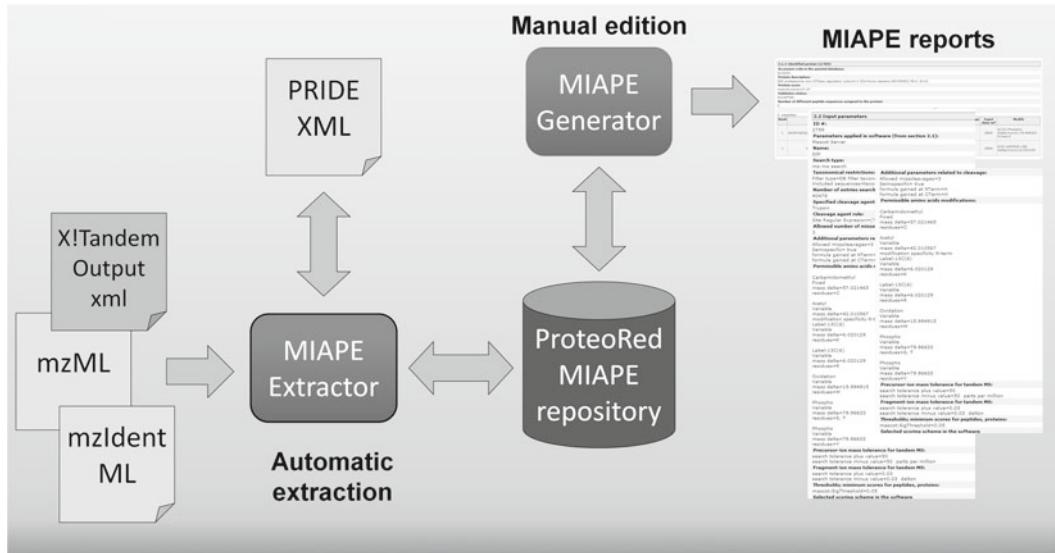


Fig. 2 Schematic workflow of the MIAPE Extractor and MIAPE Generator tools. MIAPE information is automatically extracted from mzML, mzIdentML, X!Tandem xml output file and PRIDE xml files, and it is stored in the ProteoRed MIAPE repository. This MIAPE information can then be manually checked and edited using the online MIAPE Generator tool which can generate MIAPE compliant reports that can serve as materials and methods section or as supplementary material in a manuscript for publication

5.2.1 The MIAPE Extractor

The *MIAPE extractor* tool provides automation to the PMWTK. It is a Java® stand-alone application that allows, in a fully automatic way, to extract MIAPE information from files such as PRIDE XML, mzML, or mzIdentML, and stores it in the ProteoRed MIAPE repository. So, the user can automatically generate MIAPE MS and MSI compliant reports from his proteomics data and then he can easily add missing MIAPE information by completing the corresponding Web forms from the MIAPE Generator tool. Additionally, the MIAPE Extractor tool is able to export MIAPE MS and/or MSI documents from the repository to a PRIDE XML file that will be MIAPE-compliant in case of these MIAPE reports were complete, and that can be sent to the PRIDE public repository for publication (Fig. 2).

5.2.2 The ProteoRed Java MIAPE-API

The MIAPE Extractor tool is developed using the ProteoRed Java MIAPE API, an Application Programming Interface (API) coded in Java that provides a data model for representing MIAPE GE, GI, MS and MSI documents; it includes a set of functionalities to extract MIAPE information from standard proteomics files; it also offers several programmatic ways to manage and store MIAPE information. This API licensed under an open-source license, is the unique programmatic resource that allows the development of tools managing MIAPE information.

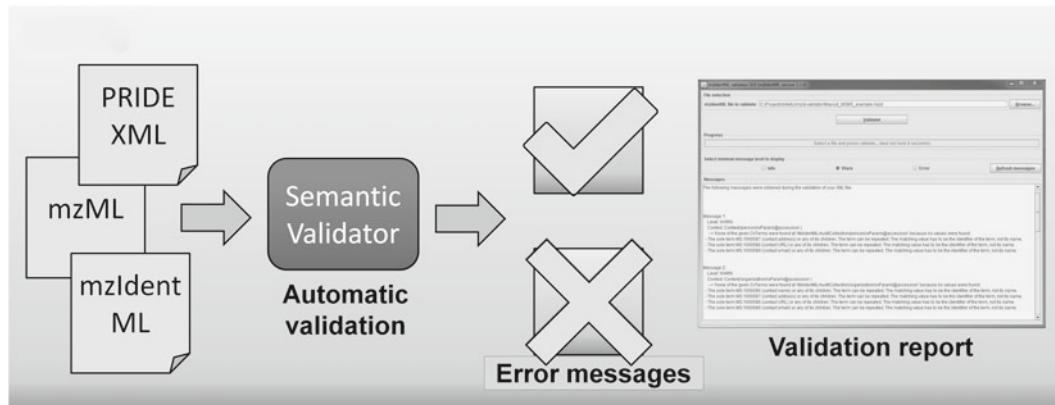


Fig. 3 Schematic workflow of the semantic validators. Standard data files are submitted to the validator software. If the file is MIAPE compliant, that is, it is semantically valid and contains all the data required by MIAPE guidelines, the validator will show a validation report confirming the MIAPE compliance of the file. However, if the file is not valid, the validation report will contain some error messages explaining which information is missing or misspelled in the file

Additionally, the bioinformatics working group of ProteoRed is also providing a set of Web-services that offers interaction with the ProteoRed MIAPE repository. So, it provides a remote and secure access to stored MIAPE documents, as well as a way to remotely store new MIAPE documents. The MIAPE Extractor tool uses these Web-services to interact with the ProteoRed MIAPE repository.

5.2.3 Semantic Validators

Together with the proteomics data representation schemas, a system for validating standard data files has been developed by HUPO-PSI. The called PSI semantic validator framework [21] not only checks the XML syntax but it also enforces rules regarding the use of an ontology class or Controlled Vocabulary (CV) terms by checking that the terms exist in the resource and that they are used in the correct location of a document. Moreover, this framework is extremely fast, even on sizable data files, and flexible, as it can be adapted to any standard by customizing the parameters it requires: an XML Schema Definition, one or more CVs or ontologies, and a mapping file describing in a formal way how the semantic resources and the format are interrelated. As such, the validator provides a general solution to the common problem in data exchange: how to validate the correct usage of a data standard beyond simple XML schema definition validation. It also tackles the issue of checking that experimental data reported using a specific format, CVs and public bio-ontologies (e.g., Gene Ontology, NCBI taxonomy) are compliant with the MIAPE recommendations. So, implementing the PSI semantic validator framework for each standard data format, HUPO-PSI provides a tool able to validate all the standard files (Fig. 3).

6 Discussion

The HUPO Proteomics Standards Initiative's developments are nowadays crucial for proteomics data interpretation, sharing and reprocessing. The definition of standards formats and the provision of tools that read and write these formats are particularly beneficial for highly collaborative scientific projects in which different proteomics laboratories with different equipment are exchanging data. This is also true for centralizing data and meta-data from variable sources in a repository such as PRIDE and therefore facilitates compliance with journals' requirements on data deposition. However, a minimal "quality control" (i.e., in the sense of MIAPE compliance) on proteomics data is desired in order to ensure that any reported, shared or published proteomics work has the minimum information that will allow a clear review of the performed procedures and produced data, as well as to allow a potential recreation of the work. We have described in this chapter a number of different bioinformatics resources that are available to help users to comply with required guidelines.

With this chapter, we present PSI standardization developments to the scientific community, in particular the MIAPE guidelines and related resources. These guidelines and associated tools are useful to help the community in producing, exchanging, interpreting and finally publishing proteomics data using a common language. In this era where datasets can live and exist by themselves, where they can obtain a DOI and become citable as proposed by the proteomeXchange consortium (www.proteomeXchange.org), it is of interest for the whole community to appropriately generate well annotated datasets that can be reused and reinterpreted if needed.

Acknowledgments

This publication was supported by the Spanish National Research Council (CSIC), the Spanish National Proteomics Institute (ProteoRed-ISC III) (grant number 2005X747_3) and by the EU FP7 "ProteomeXchange" (EU FP7 grant number 260558) grants. The Swiss-Prot group is part of the SIB Swiss Institute of Bioinformatics. This activity is supported by the Swiss Federal Government through the The State Secretariat for Education, Research and Innovation SERI and by the European Union project "ProteomeXchange" (EU FP7 grant number 260558).

References

1. Carr S, Aebersold R, Baldwin M et al (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
2. Orchard S, Hermjakob H, Apweiler R (2003) The proteomics standards initiative. *Proteomics* 3:1374–1376
3. Vizcaino JA, Martens L, Hermjakob H et al (2007) The PSI formal document process and its implementation on the PSI website. *Proteomics* 7:2355–2357
4. Orchard S, Binz PA, Hermjakob H (2009) Second joint HUPO publication and proteomics standards initiative workshop. *Proteomics* 9:4426–4428
5. Orchard S, Ping P (2009) HUPO world congress publication committee meeting Aug 2008, Amsterdam, The Netherlands. *Proteomics* 9:502–503
6. Orchard S, Kersey P, Zhu W et al (2003) Progress in establishing common standards for exchanging proteomics data: the second meeting of the HUPO proteomics standards initiative. *Comp Funct Genom* 4:203–206
7. Orchard S, Kersey P, Hermjakob H et al (2003) The HUPO proteomics standards initiative meeting: towards common standards for exchanging proteomics data. *Comp Funct Genom* 4:16–19
8. Orchard S, Hermjakob H, Taylor CF et al (2005) Further steps in standardisation. Report of the second annual proteomics standards initiative spring workshop (Siena, Italy 17–20th Apr 2005). *Proteomics* 5:3552–3555
9. Orchard S, Apweiler R, Barkovich R et al (2006) Proteomics and beyond: a report on the 3rd annual spring workshop of the HUPO-PSI 21–23 Apr 2006, San Francisco, CA, USA. *Proteomics* 6:4439–4443
10. Orchard S, Hermjakob H, Taylor C et al (2006) Autumn 2005 workshop of the human proteome organisation proteomics standards initiative (HUPO-PSI) Geneva, 4–6 Sept 2005. *Proteomics* 6:738–741
11. Orchard S, Taylor CF, Jones P et al (2007) Entering the implementation era: a report on the HUPO-PSI fall workshop 25–27 Sept 2006, Washington DC, USA. *Proteomics* 7:337–339
12. Orchard S, Montecchi-Palazzi L, Deutsch EW et al (2007) Five years of progress in the standardization of proteomics data 4th annual spring workshop of the HUPO-proteomics standards initiative Apr 23–25, 2007 Ecole Nationale Supérieure (ENS), Lyon, France. *Proteomics* 7:3436–3440
13. Orchard S, Albar JP, Deutsch EW et al (2008) Annual spring meeting of the proteomics standards initiative 23–25 Apr 2008, Toledo, Spain. *Proteomics* 8:4168–4172
14. Orchard S, Deutsch EW, Binz PA et al (2009) Annual spring meeting of the proteomics standards initiative. *Proteomics* 9:4429–4432
15. Orchard S, Jones A, Albar JP et al (2010) Tackling quantitation: a report on the annual spring workshop of the HUPO-PSI 28–30 Mar 2010, Seoul, South Korea. *Proteomics* 10:3062–3066
16. Orchard S, Albar JP, Deutsch EW et al (2011) Enabling BioSharing—a report on the annual spring workshop of the HUPO-PSI Apr 11–13, 2011, EMBL-Heidelberg, Germany. *Proteomics* 11:4284–4290
17. Orchard S, Jones AR, Stephan C et al (2007) The HUPO pre-congress proteomics standards initiative workshop. HUPO 5th annual world congress. Long Beach, CA, USA 28 Oct–1 Nov 2006. *Proteomics* 7:1006–1008
18. Orchard S, Martens L, Tasman J et al (2008) 6th HUPO annual world congress—proteomics standards initiative workshop 6–10 Oct 2007, Seoul, Korea. *Proteomics* 8:1331–1333
19. Orchard S, Albar JP, Deutsch EW et al (2010) Implementing data standards: a report on the HUPOPSI workshop Sept 2009, Toronto, Canada. *Proteomics* 10:1895–1898
20. Orchard S, Albar JP, Deutsch EW et al (2012) From proteomics data representation to public data flow: a report on the HUPO-PSI workshop Sept 2011, Geneva, Switzerland. *Proteomics* 12:351–355
21. Montecchi-Palazzi L, Kerrien S, Reisinger F et al (2009) The PSI semantic validator: a framework to check MIAPE compliance of proteomics data. *Proteomics* 9:5112–5119
22. Martens L, Chambers M, Sturm M et al (2011) mzML—a community standard for mass spectrometry data. *Mol Cell Proteomics* 10(1):R110.000133. doi:[10.1074/mcp.R110.000133-2](https://doi.org/10.1074/mcp.R110.000133)
23. Eisenacher M (2011) mzIdentML: an open community-built standard format for the results of proteomics spectrum identification algorithms. *Methods Mol Biol* 696:161–177
24. Jones AR, Eisenacher M, Mayer G et al (2012) The mzIdentML data standard for mass spectrometry-based proteomics results. *Mol Cell Proteomics* 11:M111.014381
25. Deutsch EW, Chambers M, Neumann S et al (2011) TraML: a standard format for exchange

- of selected reaction monitoring transition lists. *Mol Cell Proteomics* 10:O111.015446
26. Gibson F, Hoogland C, Martínez-Bartolomé S et al (2010) The gel electrophoresis markup language (GelML) from the proteomics standards initiative. *Proteomics* 10: 3073–3081
27. Hermjakob H, Montecchi-Palazzi L, Bader G et al (2004) The HUPO PSI's molecular interaction format – a community standard for the representation of protein interaction data. *Nat Biotechnol* 22:177–183
28. Smith B, Ashburner M, Rosse C et al (2007) The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol* 25:1251–1255
29. Gibson F, Anderson L, Babnigg G et al (2008) Guidelines for reporting the use of gel electrophoresis in proteomics. *Nat Biotechnol* 26:863–864
30. Domann PJ, Akashi S, Barbas C et al (2010) Guidelines for reporting the use of capillary electrophoresis in proteomics. *Nat Biotechnol* 28:654–655
31. Jones AR, Carroll K, Knight D et al (2010) Guidelines for reporting the use of column chromatography in proteomics. *Nat Biotechnol* 28:654
32. Taylor CF, Binz PA, Aebersold R et al (2008) Guidelines for reporting the use of mass spectrometry in proteomics. *Nat Biotechnol* 26:860–861
33. Hoogland C, O'Gorman M, Bogard P et al (2010) Guidelines for reporting the use of gel image informatics in proteomics. *Nat Biotechnol* 28:655–656
34. Binz PA, Barkovich R, Beavis RC et al (2008) Guidelines for reporting the use of mass spectrometry informatics in proteomics. *Nat Biotechnol* 26:862
35. Orchard S, Salwinski L, Kerrien S et al (2007) The minimum information required for reporting a molecular interaction experiment (MIMIx). *Nat Biotechnol* 25:894–898
36. Bourbeillon J, Orchard S, Benhar I et al (2010) Minimum information about a protein affinity reagent (MIAPAR). *Nat Biotechnol* 28:650–653
37. Orchard S, Al-Lazikani B, Bryant S et al (2011) Minimum information about a bioactive entity (MIABE). *Nat Rev Drug Discov* 10:661–669
38. Taylor CF, Field D, Sansone SA et al (2008) Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. *Nat Biotechnol* 26:889–896
39. Martens L, Hermjakob H, Jones P et al (2005) PRIDE: the proteomics identifications database. *Proteomics* 5:3537–3545
40. Jones P, Cote RG, Martens L et al (2006) PRIDE: a public repository of protein and peptide identifications for the proteomics community. *Nucleic Acids Res* 34:D659–D663
41. Medina-Aunon JA, Martínez-Bartolomé S, Lopez-Garcia MA et al (2011) The ProteoRed MIAPE web toolkit: a user-friendly framework to connect and share proteomics standards. *Mol Cell Proteomics* 10:M111.008334
42. Robin X, Hoogland C, Appel RD et al (2010) MIAPEGelDB, a web-based submission tool and public repository for MIAPE gel electrophoresis documents. *Journal of proteomics* 71:249–251
43. Martínez-Bartolomé S, Medina-Aunon JA, Jones AR, et al (2010) Semi-automatic tool to describe, store and compare proteomics experiments based on MIAPE compliant reports. *Proteomics* 10:1256–1260
44. Martínez-Bartolomé S, Blanco F, Albar JP (2010) Relevance of proteomics standards for the ProteoRed Spanish organization. *Journal of proteomics* 73:1061–1066
45. Paradela A, Escuredo PR, Albar JP (2006) Geographical focus. Proteomics initiatives in Spain: ProteoRed. *Proteomics* 6 Suppl 2:73–76

INDEX

A

- Agrobacterium* 657–659, 662
Algorithm 5, 17, 19, 23, 36, 54, 59, 98, 104, 176, 180, 181, 213, 282, 283, 297, 305, 355, 385
Allergen 6, 12, 123, 125, 380, 435, 437, 443, 446, 447, 725–734, 740
Aluminium hydroxide metal-oxide affinity chromatography (Al(OH)₃-MOAC) 622
Amyloplasts 470
Antibodies 11, 18, 19, 22, 34, 46, 47, 113, 200, 243, 249, 371, 372, 374, 456, 457, 459, 462, 513, 598, 656–658, 660–662, 726, 738, 747, 772
Arabidopsis 6, 11, 31–39, 111–114, 126, 187–190, 247, 260, 275–287, 333–335, 367, 394, 410, 436–447, 453–464, 495, 496, 500, 509, 528–536, 541–550, 597, 605, 610–619, 623, 655–662, 667–672, 689
Avidin affinity chromatography 113, 678, 679, 682

B

- Barley 125, 224, 233, 234, 236, 237, 368, 500, 683, 684, 737–740, 746, 747, 752
Beer 125, 737–752
Biofuel 6
Bioinformatics 4, 5, 11, 25, 38, 94, 165, 171–182, 373, 380, 394, 398, 399, 401, 535, 542, 709, 767, 770–775, 777, 778
Biotin switch 609–620
Blotting 46, 416, 420, 459, 513, 657–661, 726, 728, 730, 731, 734, 738
Blue Native PAGE (BN-PAGE) 10, 317, 328, 668, 690
Botrytis cinerea 96–100, 104–106, 755
Bottom-up 9, 10, 75, 98, 172, 374
Brachypodium distachyon 332, 482
Brassica oleracea 392–395

C

- Callus culture 407, 408, 414–415
Carnation (*Dianthus caryophyllus*) 573–590

- Cell-free protein expression system 260
Cell wall 85, 90, 94, 96, 112, 116, 317, 324, 327–328, 335, 348, 355, 386, 392, 409, 410, 448, 454, 461–463, 535, 563, 664, 695, 742
Chernobyl 555–560
Chlamydomonas reinhardtii 16, 19–21, 23, 25, 35, 352
Chloroplasts 31, 35, 123, 145, 284, 318, 454–456, 460–465, 469, 470, 472, 473, 476–479, 502, 504–505, 511–513, 516, 547, 548, 550, 551, 671, 677, 744
Chromoplasts 470
Chymotrypsin 166, 185–197, 741
CID. *See* Collision induced dissociation (CID)
Clear Native PAGE (CN-PAGE) 10, 668, 694
Clustering comparative 340
Collision induced dissociation (CID) 66, 71, 75, 77, 157, 163, 164, 211, 281, 354, 419, 476, 617, 629
Combinatorial peptide ligand library (CPLL) 11, 112, 114–127, 739–740, 747, 755
Cytosol 453–455, 458, 459, 461, 462, 550
- D**
- Dataset 25, 30, 31, 33, 35, 40, 52–56, 58, 164, 168, 180–182, 277, 282–284, 304, 322, 424, 520, 719, 778
Data transformation 55
1-DE. *See* One-dimensional electrophoresis (1-DE)
2-DE. *See* Two-dimensional electrophoresis (2-DE)
Density gradient centrifugation 329, 472–473, 510–511, 544, 545, 689
Descriptive proteomics 57, 94
Detergent 86, 90, 96, 103, 116, 121, 138, 152, 160, 216, 260, 266, 267, 347–357, 387, 447, 483, 488–489, 650, 668–672, 675, 676, 689, 690, 700, 711
Difference gel electrophoresis (DIGE) 9, 10, 142, 150, 153, 159, 320, 325, 365, 367, 501, 505–507, 514–517, 523

- 3D Ion Trap (3D-IT) 63–65, 70, 72, 73, 77
 Drift-time ion mobility mass spectrometry (DT-IMMS) 74, 75
 Dynamic range 6, 11, 73, 76–78, 112, 114, 115, 117–118, 156, 159, 182, 205–207, 209, 219, 306, 347, 362, 369–373, 380, 500
- E**
- EDMAN sequencing 8, 45
 Electrospray ionization (ESI) 10, 68, 69, 72, 195, 219, 323, 343, 354, 369, 424, 428, 683, 691, 739
 Electrospray ionization-tandem mass spectrometry (ESI-MS/MS) 219, 317, 320, 323, 337, 500, 507, 516, 517, 740
 Embryo 131–134, 136–138, 362–364, 366, 382, 428, 503, 583, 584, 714, 726
 Endoplasmic reticulum (ER) 368, 369, 444, 458, 459, 462, 527–538, 547, 646, 689, 696, 697
 Environmental stress 290, 327, 469
 ESI. *See* Electrospray ionization (ESI)
 ESI-LTQ-Orbitrap 74
 ESI-MS/MS. *See* Electrospray ionization-tandem mass spectrometry (ESI-MS/MS)
 EST. *See* Expressed sequence tag (EST)
 Experimental design 9, 11, 52, 53, 94, 146, 158, 159, 286, 301, 373, 421, 577, 588, 629, 709, 719, 771
 Expressed sequence tag (EST) 31, 98, 131, 337, 340, 345, 373, 386, 393, 503, 739
 Extracellular proteome 407–431

F

- False discovery rate (FDR) 57, 165, 177, 181, 211, 277, 282, 285, 298, 355, 419, 431, 476, 617, 620, 630
 FFE. *See* Free-flow electrophoresis (FFE)
 Fluorescent label 141–153
 Fluorochromes 141
 Forest tree 141–153
 Fourier transform ion cyclotron resonance (FT-ICR) 66, 69, 71, 73, 77
 Free-flow electrophoresis (FFE) 444, 500–502, 513, 518, 527–538
 FT-ICR. *See* Fourier transform ion cyclotron resonance (FT-ICR)
 Full-scan selective peptide extraction (Selpex) 303–312
 Fungi 93–106, 152, 380
Fusarium 394
Fusarium oxysporum f. sp. *dianthi* 574, 588

G

- GAL4 242–244, 247–250, 254, 255, 257
 Gel-fractionation 347–357
 Genome annotation 18, 25, 31, 290
 Glycoblotting 646–650, 652
 Glycome 646, 650, 651
 Glycoprotein 94, 112, 145, 325, 327, 372, 409, 633–636, 638–641, 645–652
 Golgi apparatus 527–538
 Gradient purification 499
 Grain 30, 123, 187, 190, 196, 224, 233, 234, 236, 362, 435–438, 440–444, 446–448, 550, 715, 739–741, 745, 746
Grape vine (Vitis vinifera) 160, 428
 Guidelines 12, 38, 53, 112, 765, 766, 768, 770, 772–775, 777, 778

H

- HAMMOC. *See* Hydroxy acid-modified metal oxide chromatography (HAMMOC)
 Heat-mapping 58
 Heavy peptide internal standard 202, 204
 High-performance liquid chromatography (HPLC) 69, 104, 120, 159, 160, 163, 166, 190, 195, 279, 281, 287, 348, 357, 385, 414, 418, 419, 464, 500, 507, 517, 519, 571, 613, 617, 619, 622, 650, 740
 HILIC. *See* Hydrophilic interaction chromatography (HILIC)
*Holm oak (Quercus ilex subsp. *ballota*)* 709–721
 HPLC. *See* High-performance liquid chromatography (HPLC)
 Human Proteome Organization (HUPO) 765–767, 778
 Hybrid mass spectrometers 163
 Hydrophilic interaction chromatography (HILIC) 647, 649, 652
 Hydrophobic proteins 43, 96, 105, 185–197, 337, 427, 481, 482, 500, 711
 Hydroxy acid-modified metal oxide chromatography (HAMMOC) 595–605

I

- ICAT. *See* Isotope-coded affinity tag (ICAT)
 IEF. *See* Isoelectric focusing (IEF)
 Immobilized metal-ion affinity chromatography (IMAC) 46, 597, 598, 622
 Immobilized pH gradients (IPG) 44, 100, 103, 104, 133, 135, 143, 147–150, 153, 315, 317, 318, 320, 321, 336, 371, 412, 419–421, 429, 430, 506, 514, 515, 557, 558, 567, 575, 576, 580, 581, 584–587, 589, 590, 712, 718, 727, 749–751, 756, 757, 759–762
 Immunoproteomics 725–734

- In-gel digestion 98, 189–190, 194–195, 291, 292, 294–296, 348, 418, 423, 493, 729, 731–732
- In-solution digestion 493, 494
- Interactome 374
- Interactomics 5, 11
- Ion trap (IT) 63–65, 69–75, 77, 211, 276, 277, 427, 476
- Isobaric tags for relative and absolute quantitation (iTRAQ) 9, 10, 155–168, 171, 172, 470, 471, 473–477, 678, 771
- Isoelectric focusing (IEF) 44, 85, 87, 96, 97, 103, 106, 121, 132, 134, 149, 151, 153, 212, 276, 317, 318, 320, 321, 365, 371, 420, 429, 500, 506, 514–515, 522, 557, 558, 560, 570, 575, 580, 581, 585, 587, 590, 689, 693, 697, 698, 700, 701, 714, 718, 730, 750, 759–762
- Isotope-coded affinity tag (ICAT) 9, 10, 171, 471, 474, 475, 609–620, 677–684
- iTRAQ. *See* Isobaric tags for relative and absolute quantitation (iTRAQ)
- L**
- Label-free 5, 9, 10, 12, 59, 104, 105, 171–175, 299, 300, 304, 373
- Label-free proteomics (GeLC) 173, 176, 299, 300, 336
- LC. *See* Liquid chromatography (LC)
- LC-MS/MS. *See* Liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS)
- LDS-PAGE. *See* Lithium dodecylsulfate-PAGE (LDS-PAGE)
- Leaf 113, 115, 123, 124, 276, 285, 293, 297, 304, 305, 317–319, 322, 392, 476, 478, 479, 541–551, 563, 565, 566, 568–570, 597, 659, 660, 741
- Lectin 46, 325, 371, 372, 402, 403, 633–641
- Lilium longiflorum* 436, 437, 440, 441
- Linear ion trap (LIT) 70–73, 77, 163, 164, 476
- Linear trap quadrupole (LTQ) 72, 74, 276, 277, 571, 629
- Lipid 11, 85–90, 96, 116, 121, 145, 160, 186, 235, 238, 260, 261, 323, 363, 374, 386, 387, 443, 444, 469, 481, 514, 518, 689, 738, 741, 746
- Lipid removal 86, 116
- Liposome 260, 261, 263–266, 269–271
- Liquid chromatography (LC) 8–10, 39, 69, 94, 132, 134–139, 199, 217, 281, 291, 322, 336, 456, 458, 482, 530, 630
- Liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) 31, 97, 121, 124, 125, 141, 157, 158, 160, 162, 174, 190, 192, 193, 195–196, 212, 213, 296, 304, 306, 321, 336, 347, 348, 354–355, 392, 393, 400, 414, 418–419, 423, 424, 458, 464, 474, 475, 479, 481, 488, 492, 500, 534, 598, 604, 678–680, 683, 738–740, 745, 746
- LIT. *See* Linear ion trap (LIT)
- Lithium dodecylsulfate-PAGE (LDS-PAGE) 667–676
- Low-abundant proteins 206, 403
- LTQ. *See* Linear trap quadrupole (LTQ)
- Lysyl endopeptidase (LysC) 305, 599, 601, 604, 605, 626
- M**
- Magnaporthe oryzae* 564
- Maize (*Zea mays*) 30, 35, 368, 689
- MALDI. *See* Matrix-assisted laser desorption/ionization (MALDI)
- MALDI-LTQ-Orbitrap 74
- MAPA. *See* Mass accuracy precursor alignment (MAPA)
- Marker enzyme 445–446, 448, 501, 504–505, 512–513, 547, 548, 550, 689
- MASCOT 104, 282, 385, 414, 419, 423, 517, 519, 530, 534, 535, 587, 596, 617, 682, 714, 720, 729, 733, 751, 752
- Mass accuracy precursor alignment (MAPA) 15–26, 306, 355, 631
- Mass analyzers 10, 22, 61–78, 212, 629
- Mass spectrometric imaging (MSI) 223–239, 769, 771, 773, 774, 776
- Mass western 15–26, 199–207
- Matrix-assisted laser desorption/ionization (MALDI) 68, 322, 400, 691, 739, 746, 751
- MAXQUANT 629–631
- Medicago truncatula* 6, 34, 35, 303–312, 334, 442
- Membrane
- detergent-resistant 482
 - microdomain 481–497
- Metabolites 8, 16, 36, 37, 76, 77, 85, 96, 116, 207, 223–239, 316, 329, 408, 409, 428, 453, 564
- Metabolomics 7, 16, 29, 37, 141, 325, 328, 564
- Metal-oxide affinity chromatography (MOAC/MOC) 16, 23, 621–631
- Microbe 35, 368, 409
- The Minimal Information About a Proteomics Experiment (MIAPE) 12, 31, 32, 38–39, 765–778
- Missing value 52, 54, 55, 180, 312
- Mitochondria 145, 260, 304, 328, 329, 444, 445, 448, 454, 458, 459, 462, 499–523, 527, 528, 538, 541, 547, 550, 696, 744

- MOAC/MOC. *See* Metal-oxide affinity chromatography (MOAC/MOC)
- MS/MS. *See* Tandem mass spectrometry (MS/MS)
- Multidimensional protein identification technology (MudPIT).....9, 10, 98, 172, 173, 290, 301, 336, 424, 458, 469, 571
- Multiple reaction monitoring (MRM).....5, 38, 77, 207, 747, 772
- Multiple-stage MS (MSn).....71
- Multiplexing.....141–153
- Multivariate statistics.....16, 18, 23, 24, 53, 55, 57–58, 355, 631
- Myristylation.....260
- N**
- Nano-electrospray ionisation tandem mass spectrometry (nano-ESI-MS/MS).....740
- Nitrosothiols.....610, 611, 614, 618–620
- Nitrosylation.....609–620
- Non parametric tests.....52, 56–57
- O**
- Oil bodies.....85, 86, 185–197
- One-dimensional electrophoresis (1-DE).....9, 39, 94, 96, 97, 100, 103, 105, 106, 393, 416–417, 429, 482, 574, 579–580, 582, 584, 587, 589
- Open Reading Frames (ORF).....25
- Orbitrap.....35, 59, 71, 73, 74, 77, 163, 164, 206, 212, 305, 476, 629
- Organelle.....4, 9, 36, 85, 94, 112, 159, 186, 269, 276, 284, 317, 318, 327, 329, 330, 362, 443–445, 448, 453–459, 461, 462, 464, 465, 469, 470, 478, 479, 499–501, 510, 513, 516, 521, 527, 528, 534, 535, 538, 541–551, 695
isolation.....276, 470
- Orphan plant species.....333–345, 386
- Osborne.....379–387
- P**
- PAGE. *See* Polyacrylamide gel electrophoresis (PAGE)
- Partial least squares, discriminant analysis (PLS-DA).....58
- PCA. *See* Principal component analysis (PCA)
- Pea (*Pisum sativum*).....371, 470, 500, 542
- Peptide
based analysis.....336
desalting.....348, 352–354, 623, 626–627
proteotypic.....17, 22
quantification.....306
- Peroxidases.....324, 392, 410, 418, 426, 687–704
- Peroxisomes.....329, 459, 462, 511, 512, 527, 528, 541–551, 696
- Phenolics.....11, 85, 105, 145, 316, 318, 341, 356, 428, 502, 521, 568, 578, 624, 637, 662, 690, 703, 721, 761
- Phosphopeptide.....23, 24, 34, 595–605, 622, 623, 627–631
- Phosphoproteomics.....16, 24, 25, 595–605, 622
- Phosphorylation.....17, 24, 34, 35, 38, 76, 159, 164, 179, 374, 409, 596, 621, 622, 630, 631
- Plant defence.....113, 409, 574, 662
- Plant pathogenic fungi.....93–106
- Plant-pathogen interaction.....587, 588
- Plastids.....30, 35, 276, 283, 284, 304, 458, 459, 462, 469–479, 536, 541, 646, 696
- PLS-DA. *See* Partial least squares, discriminant analysis (PLS-DA)
- Pollen
grain.....435–438, 440–444, 446–448, 715
tube.....24, 435, 440–442, 445
- Polyacrylamide gel electrophoresis (PAGE).....668, 689, 717
- Polyploidy.....334
- Polysaccharides.....11, 85, 94, 96, 112, 115, 121, 126, 387, 415, 429, 574, 756
- Post-translational modification (PTM).....5, 8, 10–12, 201, 202, 213, 214, 355, 366, 374, 596
- Precipitation.....86, 87, 89, 90, 95, 96, 101, 102, 112, 113, 116, 117, 125, 127, 143, 145–148, 161, 316, 318, 319, 321, 335, 338, 349, 370, 371, 381, 384, 402, 412, 415, 428, 429, 464, 479, 505, 514, 518, 549, 560, 565, 577, 578, 580, 614, 688, 695, 716, 740, 755–762
- Precursor fragmentation.....74
- Prefractionation.....47, 159–163, 174, 176, 212, 348, 369, 570, 596, 598
- Principal component analysis (PCA).....20, 57–58, 423, 719
- Proteasome.....123, 255, 327, 454, 655–663
- Protein
coverage.....118, 186
digestion.....9, 118–120, 161, 166, 171, 186, 336, 339, 342, 349–350, 352–354, 357, 413, 418, 419, 423, 557, 626, 639–640
extraction.....5, 8, 9, 11, 39, 85–90, 94, 96, 98, 101–103, 112, 113, 123, 126, 133–134, 138, 145, 147, 166, 217, 246–247, 279–280, 291–294, 300, 305, 316–322, 335, 338–342, 348–351, 379–387, 412, 415, 427, 428, 462, 471, 473, 478, 555–558, 564–567, 573–590, 596, 599, 601, 622–624, 636–638, 657, 659, 683, 709–712, 714–716, 721, 733, 758
protein interaction.....30, 32–34, 241–257, 542
quantitation.....156, 351, 356, 416
- Proteomics.....5
- Proteomics Identifications (PRIDE)
database.....31–33, 35, 38, 775, 776, 778

- Proteotyping 6, 10, 12, 709–721
- Protoplasts 94, 392, 454, 457–465, 528–529, 531, 536
- PTM. *See* Post-translational modification (PTM)
- Q**
- Quadrupole 10, 22, 63–66, 69–77, 209, 211, 323, 745
- Quadrupole-time of flight (Q-TOF) 71, 163, 164, 343, 559, 617, 619
- Quantitative proteomics 10, 16–18, 25, 38, 45, 51–53, 77, 98, 156, 159, 171–182, 275–288, 292, 299, 300, 373, 469–479, 500, 678, 711
- R**
- Radioactivity 266, 555, 556
- Repository 38–40, 766, 773–778
- Reproducibility 44–45, 97, 139, 312, 716, 761
- Resolution 23, 53, 67, 71, 73, 74, 78, 86, 96–98, 132, 134, 139, 143, 150, 163, 172, 177, 186, 196, 205–207, 223, 228, 231–236, 238, 239, 276, 306, 315, 316, 318, 322, 336, 347, 354, 356, 374, 419, 421, 430, 476, 482, 528, 570, 574, 580, 582–584, 586, 589, 590, 628–630, 668, 674, 688–690, 694, 697, 703, 738, 739, 748, 761
- Reverse-phase high-performance liquid chromatography (RP-HPLC) 69, 501, 518, 519, 740
- Rhizobium* 326, 327, 657
- Rice (*Oryza sativa*) 33, 34, 111, 293, 368
- Root 36, 55, 73, 145, 224, 226, 227, 229, 236, 276, 278, 285, 292, 295, 304, 317–320, 323–328, 392, 394, 428, 470, 502, 546, 556, 574, 577, 580, 582, 584, 587, 588, 689, 694, 695
- RP-HPLC. *See* Reverse-phase high-performance liquid chromatography (RP-HPLC)
- S**
- SDS-PAGE. *See* Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)
- Secretome 94, 96, 410, 414, 563–571
- Seed
- coat 362, 363, 366
 - storage protein 379
- Selected reaction monitoring (SRM) 38, 209–220, 455, 456, 459, 772
- Selpex. *See* Full-scan selective peptide extraction (Selpex)
- Sensitivity 5, 11, 39, 67, 69–77, 98, 114, 195, 202, 204–207, 209, 223, 236, 242, 337, 347, 362, 366, 374, 430, 447, 455, 482, 528, 574, 583, 590, 738, 740, 756, 772
- Sequence tag 164, 424
- SEQUEST 59, 104, 176, 177, 179, 305, 355, 385, 419
- Shotgun 5, 17, 18, 25, 34, 39, 43–46, 98, 119, 201, 205, 209–212, 214, 217, 218, 220, 289–301, 303–312, 371, 380, 470, 474, 476, 481–497, 595–605, 630
- SILAC. *See* Stable isotope labeling by/with amino acids in cell culture (SILAC)
- Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) 18, 22, 31, 85, 87, 97, 103, 115, 119, 121, 124, 125, 134, 135, 149–151, 153, 173, 174, 186–188, 192–194, 265, 268–271, 278, 280, 287, 289–301, 317, 318, 320, 326, 336, 348, 349, 351–352, 365, 382, 384, 385, 394, 400–401, 403, 412–413, 416–421, 448, 458, 465, 500, 506, 515, 557, 575, 579–580, 582–585, 638, 639, 657–660, 662, 693, 697–702, 711, 713, 715, 719, 720, 727–728, 730, 738–740, 760–761
- Solid phase extraction (SPEC) 338, 518, 626, 640, 682
- Solubilization 88–90, 95, 96, 99, 102, 103, 105, 113, 116, 126, 152, 166, 305, 318–320, 348, 350, 356, 382–384, 387, 481, 482, 552, 557, 560, 576–579, 668–669, 672, 689, 699–707, 711, 712, 715, 716, 721
- Soybean (*Glycine max*) 36, 334, 362
- Spectral counting 9, 171–182, 276, 283, 287, 298–300, 304, 373, 501, 520, 535, 771
- Spectrum/Spectra 8, 17, 35–36, 38, 39, 59, 72, 73, 76, 78, 85, 98, 151–153, 156, 163–165, 167, 172, 173, 176–177, 181, 196, 197, 210, 212–214, 220, 223, 228, 231, 232, 236–239, 276, 277, 282–284, 287, 292, 297–299, 336–338, 343, 419, 423, 424, 428, 476, 517, 519, 535, 613, 629, 649, 652, 679, 683, 720, 733, 770, 771, 775
- SRM. *See* Selected reaction monitoring (SRM)
- Stable isotope labeling by/with amino acids in cell culture (SILAC) 9, 10, 172, 771
- Stress 12, 36, 47, 142, 290, 291, 299, 304, 316, 319, 323–330, 380, 387, 394, 402, 409, 410, 453, 454, 469, 499, 522, 536, 564, 610, 617, 618, 656, 688, 689, 768, 774
- Strong cation exchange (SCX) chromatography 157–160, 162, 172, 173
- Subcellular proteomics 10, 11, 316, 327–329, 403, 527
- Suspension cultured cells (SCC) 407, 414, 536
- Systems biology 4, 15–26, 300, 303–312
- T**
- Tandem mass spectrometry (MS/MS) 38, 66, 70, 76, 290
- Targeted analysis 18, 19, 74, 213, 228
- Target(ed) proteomics 4, 15–19, 38, 71–77, 209, 216

- Target peptide 304, 306, 310
 Thioredoxin 677–684, 746
 Thylakoid 472, 478, 502, 510, 547, 548, 668–672, 675
 Time-of-flight (TOF) 10, 66, 69–72, 75, 77
 TiO₂-MOAC. *See* Titanium dioxide metal-oxide affinity chromatography (TiO₂-MOAC)
 Tip growth 435
 Titanium dioxide metal-oxide affinity chromatography (TiO₂-MOAC) 24, 622
 Tobacco (*Nicotiana tabacum*) 437
 Tomato (*Solanum lycopersicum*) 334
 Top-down 9–11, 75, 98
 Transcriptomics 7, 16, 29, 37, 141, 324, 564
 Translational proteomics 12
 Triple quadrupole (TQ) 22, 66, 71, 72, 76, 77, 199, 205–206, 209, 211, 212, 215, 219, 456
 Trypsin 9, 98, 104, 115, 119–121, 144, 151, 157, 159, 161, 164, 166, 167, 174, 185–197, 212, 214, 230, 231, 238, 239, 279, 281, 283, 291, 292, 296, 300, 305, 323, 324, 344, 345, 349, 350, 352, 353, 357, 385, 413, 418, 419, 423, 424, 430, 431, 458, 474, 476, 481, 484, 485, 492, 501, 506, 516, 517, 519, 521, 530, 533, 557, 559, 571, 587, 599, 601, 604, 605, 613, 615, 626, 627, 640, 669, 680, 681, 713, 719, 720, 725, 726, 729, 732, 740, 741, 747, 751, 752
 Two-dimensional electrophoresis
 (2-DE) 5, 9, 10, 30, 34, 36, 37, 39, 43–48, 52–55, 59, 94, 96–98, 100, 103–106, 113, 141–153, 290, 291, 315, 316, 318–323, 325, 326, 328, 329, 337, 393, 410, 412–413, 419–422, 429, 481, 482, 555, 560, 564, 567–571, 573, 574, 576, 579–582, 584–587, 589, 590, 709, 711–713, 715, 718–720, 738–740, 748–752, 755–762
- U**
 Ubiquitin 454, 655, 656, 659, 660, 662
 Ultrafiltration 756, 758, 762
 Univariate statistics 52, 55, 57, 711
- V**
 Value mining 16, 23, 201
 Value imputation 54–55
- W**
 Wheat (*Triticum durum* Desf.) 138, 635, 689
 Wheat germ cell-free system 259–271
 Wide genomics 8, 9
 Wine 114, 117, 125, 334, 755–762
 Workflow 7–12, 16, 39, 52, 53, 57, 78, 94, 95, 156, 157, 171–182, 200, 201, 210, 224, 225, 238, 275–288, 292, 293, 338, 347–357, 369, 380, 501, 569, 570, 574, 576, 577, 587–588, 596, 611, 622, 633, 634, 679, 709, 710, 767, 770, 775–777
 Wort 737–752
- X**
Xanthomonas oryzae 564, 566
 Xylem 326, 391, 392
 Xylem sap 391–403
- Y**
 Yeast 34, 44, 125, 241–246, 248–251, 253–256, 380, 633, 657, 737–739, 746, 755
 Yeast two-hybrid system 241–257