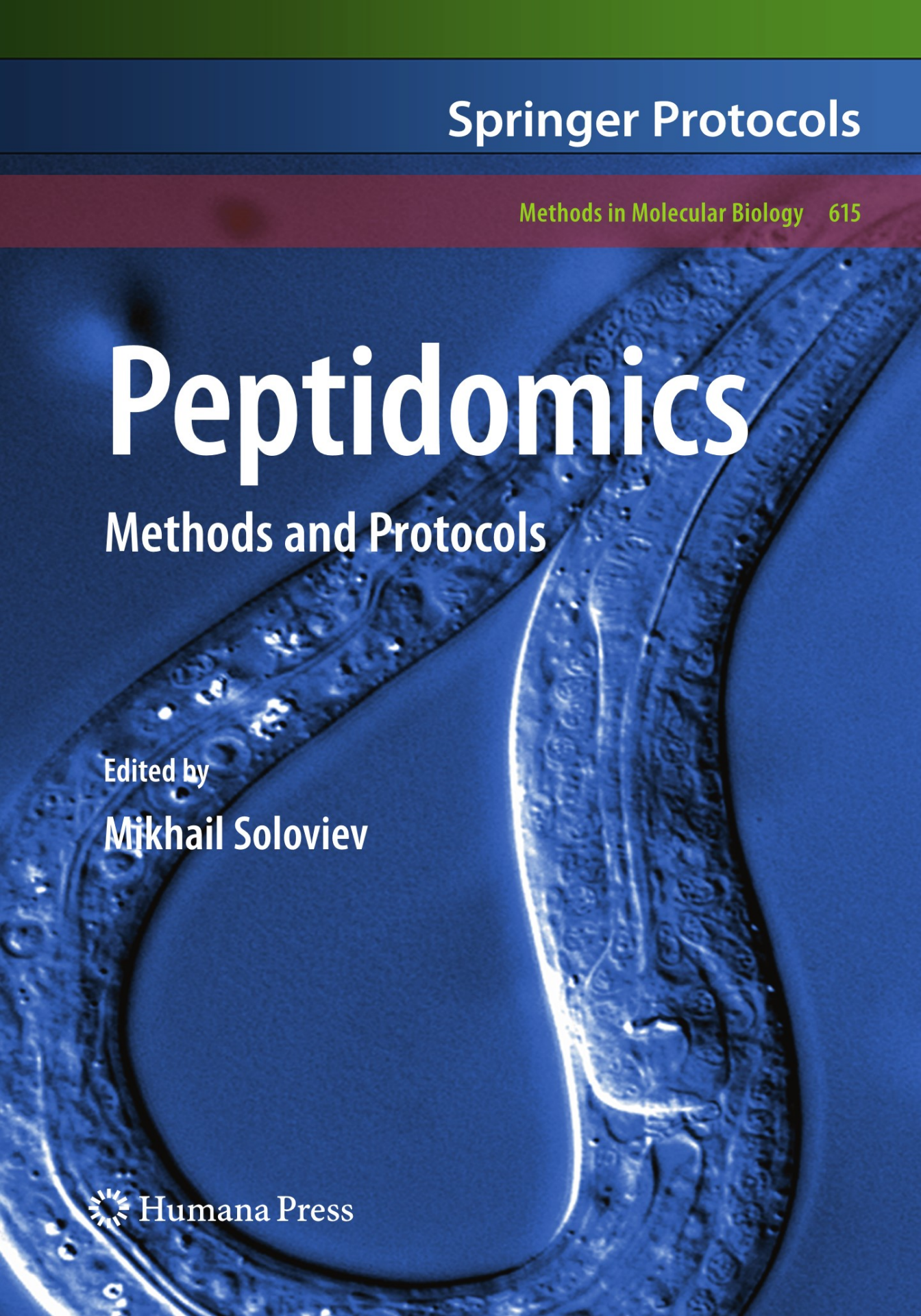
Springer Protocols

Methods in Molecular Biology 615



Peptidomics

Methods and Protocols

Edited by

Mikhail Soloviev

Humana Press

METHODS IN MOLECULAR BIOLOGY™

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

For other titles published in this series, go to
www.springer.com/series/7651

Peptidomics

Methods and Protocols

Edited by

Mikhail Soloviev

Royal Holloway, University of London, Egham, UK

Humana Press

Editor
Mikhail Soloviev
Royal Holloway University of
London
School of Biological Sciences
Egham Hill
Egham, Surrey
United Kingdom TW20 0EX
mikhail.soloviev@rhul.ac.uk

ISSN 1064-3745 e-ISSN 1940-6029
ISBN 978-1-60761-534-7 e-ISBN 978-1-60761-535-4
DOI 10.1007/978-1-60761-535-4

Library of Congress Control Number: 2009940608

Preface

Despite being known and studied for years, peptides have never before attracted enough attention to necessitate the invention of the term "Peptidomics" in order to specify the study of the complement of peptides from a cell, organelle, tissue or organism. This volume presents a comprehensive range of analytical techniques for analysis of the peptide contents of complex biological samples. The emphasis is often on higher throughput techniques, suitable for the analysis of large numbers of peptides typically present in the *peptidome*s or other complex biological samples. A wide range of methods is presented, covering all stages of peptidomic research including, where applicable, organism handling, tissue and organ dissection, cellular and subcellular fractionation, peptide extraction, fractionation and purification, structural characterisation, molecular cloning and sequence analysis. In addition to this, a selection of methods suitable for quantification, display, immunochemical and functional analysis of peptides and proteins are presented. The methods and techniques covered in this volume encompass a number of species ranging from bacteria to man and include model organisms such as *Caenorhabditis elegans*, *Drosophila melanogaster* and *Mus musculus*. Strong emphasis is placed on data analysis, including mass spectra interpretation and in silico peptide prediction algorithms. Where relevant, the *peptidomic* approaches are compared to the *proteomic* methods. Here is a snapshot of the practical information, peptidomic methods and other related protocols included in this volume:

Target organisms and samples covered: Bacteria (Chapter 2), hydra (Chapter 21), nematode (Chapter 3), mollusc (Chapter 4), crab (Chapter 5), spider venoms (Chapters 6 and 7), insects (Chapters 8, 9, 10, 11, 25), amphibians (Chapters 12, 13, 14), rodents (Chapters 15, 16, 17, 18), samples of human origin (Chapters 19, 20, 22, 23) and plants (Chapter 26).

Peptide extraction and Liquid chromatography fractionation methods (mostly size exclusion, ion exchange, reverse-phase modes or their combinations) can be found in Chapters 2, 3, 4, 5, 6, 7, 8, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22). These include OFF-line and ON-line techniques. The former are often used with MALDI-MS detection (e.g. Chapters 3, 4, 5, 6, 7, 8, 12, 15, 16, 19) whilst the latter more generally with single or multidimensional hyphenated LC^N-MS^N techniques (e.g. Chapters 2, 3, 15, 17, 18, 21).

Other separation and fractionation methods covered include microdialysis of live animals (Chapter 5), SDS-PAGE (Chapters 6, 18), magnetic bead based purification (Chapter 20) and solid-phase extraction (Chapters 2, 6, 12, 19, 22).

Affinity peptide detection including anti-peptide antibody development and characterisation, Affinity peptidomics, ELISA and microarray affinity assays are covered in Chapters 22, 23 and 24.

Mass spectrometry techniques include MALDI-TOF MS (e.g. Chapters 3, 6, 7, 8, 12, 16, 19, 20), MALDI-TOF with PSD (Chapter 8), MALDI-TOF MS/MS (e.g. Chapters 4, 6, 15, 21); ESI-MS/MS techniques (Chapters 3, 6, 16, 17, 18) or high-resolution FTMS (Chapter 2). Direct MALDI-MS peptide profiling from cells and tissues is described in Chapters 9, 10 and 11.

The description of *functional assays* can be found in Chapters 7, 14 and 21. Of particular interest in this respect is Chapter 21, where functional activity of the peptides is assessed through the analysis of mRNA transcription levels changes in response to the peptide application. That chapter contains a selection of protocols for peptide extraction, fractionation and functional testing using a combination of molecular biology techniques, cellular and morphological assays.

Molecular cloning of peptide cDNAs and the associated techniques are described in Chapters 13 and 14.

Issues related to peptide *sequence analysis* are addressed in many chapters dealing with MS spectra interpretation, but of special interest in this respect are Chapters 25 and 26, dealing with *in silico peptide prediction* techniques and Chapter 20 which includes a section on bioinformatics analysis of *peptide expression profiling data*. *Differential peptide expression* issues are also covered in Chapter 2.

Peptidomics is 10-years old. My congratulations go to all scientists who have created and developed the science of Peptidomics through their research and especially those who found time to contribute their invaluable know-how in the form of methods and protocols for inclusion in this volume. *Peptidomics: Methods and Protocols* is designed to complement previously published titles in the *Methods in Molecular Biology*™ series, which focused on *protein* analysis. This volume will help the beginner to become familiar with this fascinating field of research and will provide scientists at all levels of expertise with easy-to-follow practical advice needed to set up and carry out analysis of the peptide contents of complex biological samples.

Royal Holloway University of London Mikhail Soloviev
December 2009

Contents

Contributors

Noura Abul-Husn • *Department of Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, NY, USA*

Joshua N. Adkins • *Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA*

Suresh P. Annagudi • *Department of Chemistry and the Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL, USA*

Charles Ansong • *Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA*

Neil Audsley • *The Food and Environment Research Agency, Sand Hutton, York, UK*

Geert Baggerman • *ProMeta, Interfacultary Center for Proteomics and Metabolomics, K.U. Leuven, Leuven, Belgium*

Julian Bailes • *School of Biological Sciences, Royal Holloway University of London, Egham, Surrey, UK*

Heidi L. Behrens • *Department of Chemistry, University of Wisconsin-Madison, Madison, WI, USA*

Kurt Boonen • *Functional Genomics and Proteomics Research Unit, Department of Biology, K.U. Leuven, Leuven, Belgium*

Philippe Bulet • *TIMC-IMAG, UMR 5525, Domaine de Chosal, Archamps, France*

Ittai Bushlin • *Department of Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, NY, USA*

Elke Clynen • *Functional Genomics and Proteomics, Department of Biology, K.U. Leuven, Leuven, Belgium*

J. Michael Conlon • *Department of Biochemistry, Faculty of Medicine and Health Sciences, United Arab Emirates University, Al-Ain, UAE*

Lakshmi A. Devi • *Department of Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, NY, USA*

James A. Dowell • *Department of Chemistry, School of Pharmacy, University of Wisconsin-Madison, Madison, WI, USA*

Anna Dulneva • *School of Biological Sciences, Royal Holloway University of London, Egham, Surrey, UK*

Tsezi A. Egorov • *Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia*

Anne Fogli • *GreD UMR INSERM 931 CNRS 6142, Faculté de Médecine, Clermont-Ferrand, France*

Toshitaka Fujisawa • *Institute of Zoology, University of Heidelberg, Heidelberg, Germany*

Eugene V. Grishin • *Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia*

Eisuke Hayakawa • *Functional Genomics and Proteomics Research Unit, Department of Biology, K.U. Leuven, Leuven, Belgium*

Mingyue He • *The Babraham Institute, Cambridge, UK*

Fred Heffron • *Department of Molecular Microbiology and Immunology, Oregon Health and Sciences University, Portland, OR, USA*

Lianghai Hu • *Key Laboratory of Separation Science for Analytical Chemistry, National Chromatographic R&A Centre, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, China*

Steven J. Husson • *Functional Genomics and Proteomics, Department of Biology, K.U. Leuven, Leuven, Belgium*

Shawichi Iwamuro • *Department of Biology, Faculty of Science, Toho University, Funabashi, Chiba, Japan*

Tom Janssen • *Functional Genomics and Proteomics, Department of Biology, K.U. Leuven, Leuven, Belgium*

Mie Kanke • *Clinical Proteomics and Molecular Medicine, St. Marianna University Graduate School of Medicine, Kawasaki, Japan*

Tomohiro Kato • *Clinical Proteomics and Molecular Medicine, St. Marianna University Graduate School of Medicine, Kawasaki, Japan*

Tetsuya Kobayashi • *Department of Regulation Biology, Faculty of Science, Saitama University, Saitama, Japan*

Sergey A. Kozlov • *Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia*

Manae S. Kurokawa • *Clinical Proteomics and Molecular Medicine, St. Marianna University Graduate School of Medicine, Kawasaki, Japan*

Ren Lai • *Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China*

Bart Landuyt • *Functional Genomics and Proteomics Research Unit, Department of Biology, K.U. Leuven, Leuven, Belgium*

Kevin A. Lease • *Division of Biological Sciences and Bond Life Sciences Centre, University of Missouri, Columbia, MO, USA*

Jérôme Leprince • *European Institute for Peptide Research (IFRMP 23), INSERM U-413, UA CNRS, University of Rouen, Mont-Saint-Aignan, France*

Ka Wan Li • *Department of Molecular and Cellular Neurobiology, Centre for Neurogenomics and Cognitive Research, Faculty of Earth and Life Sciences, VU University Amsterdam, Amsterdam, The Netherlands*

Lingjun Li • *Department of Chemistry, School of Pharmacy, University of Wisconsin-Madison, Madison, WI, USA*

Marleen Lindemans • *Functional Genomics and Proteomics, Department of Biology, K.U. Leuven, Leuven, Belgium*

Feng Liu • *Data Analysis and Modeling Group, Transportation Research Institute, Hasselt University, Diepenbeek, Belgium*

Walter H.M.L. Luyten • *Department Woman and Child, Faculty of Medicine, K.U. Leuven, Leuven, Belgium*

Nathan P. Manes • *Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA*

Thomas O. Metz • *Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA*

Heather Mottaz • *Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA*

Susanne Neupert • *Institute of General Zoology and Animal Physiology, Friedrich-Schiller-University, Jena, Germany*

Nigel M. Page • *School of Life Sciences, Kingston University London, Kingston-upon-Thames, Surrey, UK*

Reinhard Predel • *Institute of General Zoology and Animal Physiology, Friedrich-Schiller-University, Jena, Germany*

Dinah Lee Ramos-Ortolaza • *Department of Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, NY, USA*

Joachim Schachtner • *Department of Biology, Animal Physiology, Philipps-University, Marburg, Germany*

Liliane Schoofs • *Functional Genomics and Proteomics Research Unit, Department of Biology, K.U. Leuven, Leuven, Belgium*

August B. Smit • *Department of Molecular and Cellular Neurobiology, Centre for Neurogenomics and Cognitive Research, Faculty of Earth and Life Sciences, VU University Amsterdam, Amsterdam, The Netherlands*

Richard D. Smith • *Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA*

Mikhail Soloviev • *School of Biological Sciences, Royal Holloway University of London, Egham, Surrey, UK*

Liang Songping • *College of Life Sciences, Hunan Normal University, Changsha, China*

Oda Stoevesandt • *Babraham Bioscience Technologies, Cambridge, UK*

Robert M. Sturm • *Department of Chemistry, School of Pharmacy, University of Wisconsin-Madison, Madison, WI, USA*

Jonathan Sweedler • *Department of Chemistry and the Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL, USA*

Toshio Takahashi • *Suntory Institute for Bioorganic Research, Osaka, Japan*

Yukiko Takakuwa • *Clinical Proteomics and Molecular Medicine, St. Marianna University Graduate School of Medicine, Kawasaki, Japan*

Alexander A. Vassilevski • *Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia*

John C. Walker • *Division of Biological Sciences and Bond Life Sciences Center, University of Missouri, Columbia, MO, USA*

Robert J. Weaver • *The Food and Environment Research Agency, Sand Hutton, York, UK*

Christian Wegener • *Emmy Noether Neuropeptide Group, Animal Physiology, Philipps-University, Marburg, Germany*

Nicola J. Weston-Bell • *Genetic Vaccine Group, Cancer Sciences Division, Southampton General Hospital, University of Southampton School of Medicine, Southampton, Hampshire, UK*

Geert Wets • *Data Analysis and Modeling Group, Transportation Research Institute, Hasselt University, Diepenbeek, Belgium*

Yang Xiang • *Clinical Proteomics and Molecular Medicine, St. Marianna University Graduate School of Medicine, Kawasaki, Japan*

Mingliang Ye • *Key Laboratory of Separation Science for Analytical Chemistry, National Chromatographic R&A Centre, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, China*

Fan Zhang • *School of Biological Sciences, Royal Holloway University of London, Egham, Surrey, UK*

Hanfa Zou • *Key Laboratory of Separation Science for Analytical Chemistry, National Chromatographic R&A Centre, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, China*

Section I

Introduction

Chapter 1

Peptidomics: Divide et Impera

Mikhail Soloviev

Abstract

The term "peptidomics" can be defined as the systematic analysis of the peptide content within a cell, organelle, tissue or organism. The science of peptidomics usually refers to the studies of naturally occurring peptides. Another meaning refers to the peptidomics approach to protein analysis. An ancient Roman strategy *divide et impera* (divide and conquer) reflects the essence of peptidomics. Most effort in this field is spent purifying and dividing the peptidomes, which consist of tens, hundreds or sometimes thousands of functional peptides, followed by their structural and functional characterisation. This chapter introduces the concept of peptidomics, outlines the range of methodologies employed and describes key targets – the peptide groups which are often sought after in such studies.

Key words: Peptidomic, peptidome, peptide, functional peptide, methods.

1. Introduction

Polypeptides, being short stretches of amino acids or small proteins, occupy a strategic position between proteins and amino acids and play, for the most part, fundamental roles by regulating the vast majority of biological processes in the animal kingdom. Whilst perhaps sometimes overlooked, the importance of the regulatory role of peptides is truly great and hard to overestimate. Peptides have in fact been the focus of much research for decades with the first successful attempts to analyse peptide content of various biological samples (including from urine, blood and brain tissues) having been reported over 60 years ago. These relied mostly on chromatography, including two-dimensional liquid chromatography (LC), or mass spectrometry (MS) techniques. Recent advances in MS and further developments in liquid chromatography, including nano-LC (1) and associated "omics"

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_1,

techniques, resulted in dramatic improvements in the sensitivity and high throughput of protein and peptide analyses (2) and generated unprecedented growth in the number of relevant publications (**Fig. 1.1**).

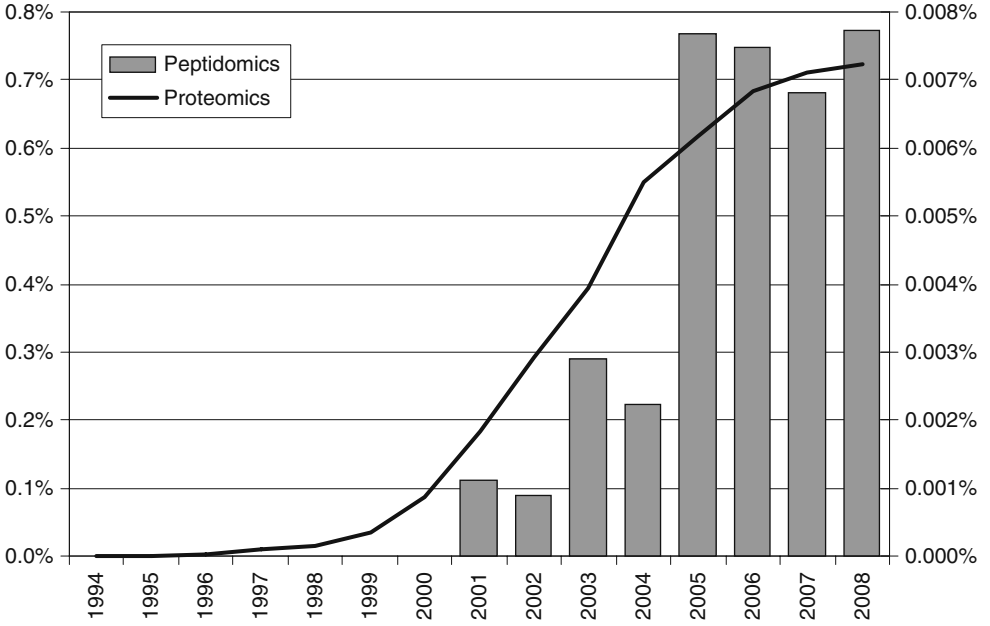


Fig. 1.1. Peptidomics publications since 1999. The *bars* represent the number of publications in PubMed containing "peptidomics OR peptidomic OR peptidome" normalised to the total number of publications added to PubMed each year. Proteomics publications in PubMed (found similarly) follow the same trend (*solid line*). *Vertical axes* show normalised data (in %) for proteomics papers (*left*) and peptidomics papers (*right*). Total number of publications on 1 January 2009 was 28,273 (proteomics) and 246 (peptidomics).

2. Peptidomics of Naturally Occurring Peptides and Peptide Pools

Similarly to "proteomics", the term "peptidomics" can be defined as the systematic analysis of the peptide content within an organism, tissue or cell (3) in order to determine peptides' identity, quantity, structure and function. Such interpretation made its public debut at the 2nd International Seminar on the Enabling Role of MS in 1999 (4), before finally appearing in press in 2001 in research papers by Peter D.E.M. Verhaert et al. (5), Peter Schulz-Knappe et al. (6) and Elke Clynen et al. (7). The discipline of peptidomics focuses on peptides that often display biological activity such as hormones, cytokines, toxins, neuropeptides and alike, which are generated from larger precursors, as well as biomarker-type peptides that may not have any bioactivity but are indicative of a particular pathology, for example the up/down-regulation of many serum peptides that result from proteolysis.

Peptidomics is in its infancy relative to other "omics" (28,273 papers on PubMed for a "proteomics" search compared to just 246 for "peptidomics" as of January 1st, 2009) but is expanding rapidly (**Fig. 1.1**).

3. Peptidomics Approach to Proteomics

In parallel, and independently of the peptidomics definition given in (5–7), another meaning was introduced by Barry et al. (8) in relation to the analysis of peptide pools (of biological fluids, tissues or cells) obtained by means of proteolytic digestion of these samples and in particular using affinity-based analysis (hence "Affinity peptidomics") e.g. in the form of protein arrays (9, 10). Since biologically occurring peptides (whether biologically active or not) are strictly speaking also the products of proteolysis (e.g. insulin pre-pro-insulin, or biologically active peptides obtained through "non-specific" proteolysis of e.g. haemoglobin), both definitions of "peptidomics" are therefore very similar in that they refer to the analysis of partially or fully proteolytically digested proteins, i.e. peptides. And finally, to acknowledge everyone involved in the birth of the "peptidomics" as a separate field of chemical biology, we should mention a Germany-based company "BioVisioN AG" which filed a trademark "peptidomics" in 1999 to cover "Chemicals used in science, in particular for analysis, other than for medical or veterinary purposes"; "Medical, veterinary and pharmaceutical products" as well as the "Scientific and industrial research; conducting medical and non-medical analyses; services in the field of diagnostics; development of pharmaceutical active substances; purchasing, licensing and exploitation of intellectual property" (11).

4. Peptidomics: The Methodologies

In the past, the majority of separation techniques used in protein and peptide analysis relied on their physical properties, such as protein or peptide size, shape, polarity, pI, the distribution of ionisable, polar and non-polar groups on the molecule surface, and their affinity towards specific or non-specific affinity capture reagents. Modern separation techniques rely on a combination of isoelectric focusing, electrophoretic separation and a great variety of liquid chromatography techniques, often linked together to yield two- or three-dimensional separation approaches, and

frequently backed up by serious automation. Highly parallel analysis is often attempted through miniaturisation (12) and the use of chip-based techniques (13–16) or the Agilent 2100 Bioanalyzer (www.chem.agilent.com). The inherent heterogeneity of the proteins' and to a lesser degree peptides' physical properties which underlies all of the above separation options is, at the same time, the inherent problem of any highly parallel protein analysis. A single universal system suitable for extraction and separation (let alone functional analysis) of all classes of proteins is yet to be reported. Unlike proteins, the peptides are often less heterogeneous in their physico-chemical properties and therefore the complete peptidomic analysis of samples, tissues and in some cases whole organisms is more straightforward than proteomic analysis.

In addition to physical methods of analysis and separation, chemical biology offers a number of other approaches, which rely directly or indirectly on chemical modifications and separation principles based on chemical properties of proteins and peptides. Chemical modification of the side chains of proteins and peptides was first reported many decades ago (see 17 for a review) and has been used widely since for protein modifications, labelling and cross-linking, but not so widely for protein separations – the latter because of the issues related to the availability and surface exposure of the reactive groups. Unlike proteins, peptides offer a unique chance to apply chemical selection techniques because of the lack of complex secondary structure and virtually complete exposure to solvent of all of the reactive groups. A number of reports utilising chemical biology approach to peptide separation and analysis have been published more recently. In most cases these describe various group-specific labelling procedures, often linked to peptide quantification (18, 19) as well as chemical depletion approaches (17, 20). Among the other "omics" technologies and approaches, "peptidomics" is the most comparable to "Proteomics" and although the terms are not synonymous, the underlying techniques and approaches are almost identical. For example, MS, a cornerstone of modern proteomics, in most cases actually analyses peptides (obtained through proteolytic digestion of proteins) or their fragments (obtained through e.g. CID), not proteins. The difference between the terms "peptidomics" and the "proteomics" is therefore blurred, especially if "methods" are being considered.

5. Peptidomics: The Targets

Target-wise, the peptidomics research is often focused, although not always, on studying peptides formed in vivo by proteolysis of specialised or non-specialised precursor proteins (often bioactive

peptides), rather than "artificially" or in vitro-produced peptides. The range of biological activities displayed by naturally occurring peptides is truly remarkable; it ranges from toxins that can paralyse or kill to peptides that have the ability to heal. The venoms of arthropods such as spiders and scorpions, as well as other species such as cone snails, comprise a vast number of neuromodulatory peptides that are capable of serious harm, but also serve as a highly useful point to discover new drugs such as painkillers (21). The identification and functional characterisation of peptides from all species including humans is crucial in the discovery of novel biomarkers and drug targets, and may yield novel therapeutic agents such as peptide-based vaccine Glatiramer acetate (GA) (Copaxone) used for the treatment of relapsing and remitting cases of multiple sclerosis (22). The suitability of peptides as biomarkers stems from the fact that they are present in all body fluids, cells and tissues (23), and many approaches focus on identifying them from such samples (24, 25). Peptides also play crucial roles in innate and adaptive immune responses by forming complexes with MHC-I, MHC-II and T cells where they stimulate defensive immune responses (26–28). The importance of peptides in cell-to-cell communication underpins the importance of peptidomics in understanding multiple pathologies that result from these communication processes going wrong. The peptide content of biological fluids, such as urine for example, can be used to produce a complete peptidomic fingerprint of an individual's health (29). The evolutionary evidence to support the importance of peptides in such widespread biological roles is evident when one examines the conservation of peptide families across species. The Tachykinin peptides for example, the largest known neuropeptide family, are found in vertebrates, protochordates and invertebrates (30). On the other end of the scale, even primitive microorganisms rely on peptide signalling, such as for example bacterial quorum sensing (31, 32) and yeast mating factors (33).

The following chapters provide a comprehensive guide to peptidomics methods and applications, spanning a range of species from bacteria to man and covering a wide range of relevant methods from basic biochemistry techniques to in silico tools and protocols.

References

1. Chervet, J.P., Ursem, M., and Salzmann, J.B. (1996) Instrumental requirements for nanoscale liquid chromatography. *Anal. Chem.* **68**, 1507–1512.
2. Quadroni, M. and James, P. (1999) Proteomics and automation. *Electrophoresis* **20**, 664–677.
3. Schrader, M. and Schulz-Knappe, P. (2001) Peptidomics technologies for human body fluids. *Trends Biotechnol.* **19**, S55–S60.
4. Verhaert, P., Vandesande, F., and De Loof, A. (1999) Automated analysis of the peptidome. No longer science fiction. In: 2nd

International Seminar on the Enabling Role of MS in Manchester.
5. Verhaert, P., Uttenweiler-Joseph, S., de Vries, M., Loboda, A., Ens, W., and Standing, K.G. (2001) Matrix-assisted laser desorption/ionization quadrupole time-of-flight mass spectrometry: an elegant tool for peptidomics. *Proteomics* **1**, 118–131.
6. Schulz-Knappe, P., Zucht, H.D., Heine, G., Jürgens, M., Hess, R., and Schrader, M. (2001) Peptidomics: the comprehensive analysis of peptides in complex biological mixtures. *Comb. Chem. High Throughput Screen* **4**, 207–217.
7. Clynen, E., Baggerman, G., Veelaert, D., Cerstiaens, A., Van der Horst, D., Harthoorn, L., Derua, R., Waelkens, E., De Loof, A., and Schoofs, L. (2001) Peptidomics of the pars intercerebralis–corpus cardiacum complex of the migratory locust, *Locusta migratoria*. *Eur. J. Biochem.* **268**, 1929–1939.
8. Scrivener, E., Barry, R., Platt, A., Calvert, R., Masih, G., Hextall, P., Soloviev, M., and Terrett, J. (2003) Peptidomics: a new approach to affinity protein microarrays. *Proteomics* **3**, 122–128.
9. Barry, R., Diggle, T., Terrett, J., and Soloviev, M. (2003) Competitive assay formats for high-throughput affinity arrays. *J. Biomol. Screen.* **8**, 257–263.
10. Barry, R. and Soloviev, M. (2004) Quantitative protein profiling using antibody arrays. *Proteomics* **4**, 3717–3726.
11. Community Trade Mark No. 001274646; http://oami.europa.eu
12. Marko-Varga, G., Nilsson, J., and Laurell, T. (2003) New directions of miniaturization within the proteomics research area. *Electrophoresis* **24**, 3521–3532.
13. Hoa, X.D., Kirk, A.G., and Tabrizian, M. (2007) Towards integrated and sensitive surface plasmon resonance biosensors: a review of recent progress. *Biosens. Bioelectron.* **23**, 151–160.
14. Kurosawa, S., Aizawa, H., Tozuka, M., Nakamura, M., and Park, J.W. (2003) Immunosensors using a quartz crystal microbalance. *Meas. Sci. Technol.* **14**, 1882–1887.
15. Lion, N., Rohner, T.C., Dayon, L., Arnaud, I.L., Damoc, E., Youhnovski, N., Wu, Z.Y., Roussel, C., Josserand, J., Jensen, H., Rossier, J.S., Przybylski, M., and Girault, H.H. (2003) Microfluidic systems in proteomics. *Electrophoresis* **24**, 3533–3562.
16. Lion, N., Reymond, F., Girault, H.H., and Rossier, J.S. (2004) Why the move to microfluidics for protein analysis?. *Curr. Opin. Biotechnol.* **15**, 31–37.
17. Soloviev, M. and Finch, P. (2005) Peptidomics, current status. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **815**, 11–24.
18. Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999.
19. DeSouza, L., Diehl, G., Rodrigues, M.J., Guo, J., Romaschin, A.D., Colgan, T.J., and Siu, K.W. (2005) Search for cancer markers from endometrial tissues using differentially labeled tags iTRAQ and cICAT with multidimensional liquid chromatography and tandem mass spectrometry. *J. Proteome Res.* **4**, 377–386.
20. Soloviev, M., Barry, R., Scrivener, E., and Terrett, J. (2003) Combinatorial peptidomics: a generic approach for protein expression profiling. *J. Nanobiotechnol.* **1**, 4.
21. Rash, L.D. and Hodgson, W.C. (2002) Pharmacology and biochemistry of spider venoms. *Toxicon* **40**, 225–254.
22. Perumal, J., Filippi, M., Ford, C., Johnson, K., Lisak, R., Metz, L., Tselis, A., Tullman, M., and Khan, O. (2006) Glatiramer acetate therapy for multiple sclerosis: a review. *Expert Opin. Drug Metab. Toxicol.* **2**, 1019–1029.
23. Adermann, K., John, H., Ständker, L., and Forssmann, W.G. (2004) Exploiting natural peptide diversity: novel research tools and drug leads. *Curr. Opin. Biotechnol.* **15**, 599–606.
24. Zimmerman, L.J., Wernke, G.R., Caprioli, R.M., and Liebler, D.C. (2005) Identification of protein fragments as pattern features in MALDI-MS analyses of serum. *J. Proteome Res.* **4**, 1672–1680.
25. Vidal, B.C., Bonventre, J.V., and I-Hong Hsu, S. (2005) Towards the application of proteomics in renal disease diagnosis. *Clin. Sci. (Lond).* **109**, 421–430.
26. Desjardins, M., Houde, M., and Gagnon, E. (2005) Phagocytosis: the convoluted way from nutrition to adaptive immunity. *Immunol. Rev.* **207**, 158–165.
27. Cresswell, P., Ackerman, A.L., Giodini, A., Peaper, D.R., and Wearsch, P.A. (2005) Mechanisms of MHC class I-restricted antigen processing and cross-presentation. *Immunol. Rev.* **207**, 145–157.
28. Van der Merwe, P.A. and Davis, S.J. (2003) Molecular interactions mediating T cell antigen recognition. *Annu. Rev. Immunol.* **21**, 659–684.

29. Metzger, J., Schanstra, J.P., and Mischak, H. (2009) Capillary electrophoresis-mass spectrometry in urinary proteome analysis: current applications and future developments. *Anal. Bioanal. Chem.* **393**, 1431–1442.
30. Severini, C., Improta, G., Falconieri-Erspamer, G., Salvadori, S., and Erspamer, V. (2002) The tachykinin peptide family. *Pharmacol. Rev.* **54**, 285–322.
31. Miller, M.B. and Bassler, B.L. (2001) Quorum sensing in bacteria. *Annu. Rev. Microbiol.* **55**, 165–199.
32. Gibbs, R.A. (2005) Trp modification signals a quorum. *Nat. Chem. Biol.* **1**, 7–8.
33. Kalkum, M., Lyon, G.J., and Chait, B.T. (2003) Detection of secreted peptides by using hypothesis-driven multistage mass spectrometry. *Proc. Natl. Acad. Sci. USA.* **100**, 2795–2800.

Section II

From Bacteria to Men

Chapter 2

Performing Comparative Peptidomics Analyses of *Salmonella* from Different Growth Conditions

Joshua N. Adkins, Heather Mottaz, Thomas O. Metz, Charles Ansong, Nathan P. Manes, Richard D. Smith, and Fred Heffron

Abstract

Host–pathogen interactions are complex competitions during which both the host and the pathogen adapt rapidly to each other in order for one or the other to survive. *Salmonella enterica* serovar Typhimurium is a pathogen with a broad host range that causes a typhoid fever-like disease in mice and severe food poisoning in humans. The murine typhoid fever is a systemic infection in which *S.* typhimurium evades part of the immune system by replicating inside macrophages and other cells. The transition from a foodborne contaminant to an intracellular pathogen must occur rapidly in multiple, ordered steps in order for *S.* typhimurium to thrive within its host environment. Using *S.* typhimurium isolated from rich culture conditions and from conditions that mimic the hostile intracellular environment of the host cell, a native low molecular weight protein fraction, or peptidome, was enriched from cell lysates by precipitation of intact proteins with organic solvents. The enriched peptidome was analyzed by both LC–MS/MS and LC–MS-based methods, although several other methods are possible. Pre-fractionation of peptides allowed identification of small proteins and protein degradation products that would normally be overlooked. Comparison of peptides present in lysates prepared from *Salmonella* grown under different conditions provided a unique insight into cellular degradation processes as well as identification of novel peptides encoded in the genome but not annotated. The overall approach is detailed here as applied to *Salmonella* and is adaptable to a broad range of biological systems.

Key words: Comparative proteomics, *Salmonella*, mass spectrometry, peptide extraction, native proteases, accurate mass.

1. Introduction

Controlled and coordinated protein degradation is critical for biological systems to function properly. The processes of protein degradation have roles in the cell-cycle (e.g., cyclins), signaling cascades (e.g., receptor shedding), protein maturation (e.g.,

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_2,

plasminogen), and nutrient cycling. Despite the critical roles of protein degradation in biological processes, there have been surprisingly few systematic global analyses of protein degradation; the majority of studies that have been performed focus on eukaryotic systems. Specific protein degradation processes are very highly regulated in bacteria and determined by environmental conditions. Selective degradation of proteins followed by cannibalization of the released amino acids is the most efficient process for bacterial adaptation to changing metabolic requirements (1, 2). Indeed, the ability of a pathogen to survive in the host and exploit new resources is an essential virulence trait.

The development of novel antibiotics against bacterial pathogens represents just a single discipline that can benefit from the elucidation of selective protein degradation processes. Recently our group developed an LC–MS/MS-based approach to globally profile a sub-set of peptides in a biological sample. Peptides, defined here, are short chains of amino acids linked via peptide bonds and are typically composed of fewer than 100 amino acids. The source of peptides in a biological system may result from short genes or through targeted degradation of proteins. Most of the peptides observed in this recent study were found to be the products of protein degradation (3); regardless of source we refer here to this naturally occurring peptide fraction as the "peptidome".

Interestingly, nearly 2% of the 4550 predicted proteins in *S.* typhimurium are annotated as being involved in protein degradation. Importantly, nearly all of these proteolytic proteins were identified in an early analysis of the *S.* typhimurium proteome, indicating that there is an upregulation of these functions under some of the growth conditions studied (4). The following is a step-by-step description of the sample preparation and analytical procedures that were used in determining the *Salmonella* peptidome. In addition, a discussion of the data analysis concerns that are unique to analyzing peptidomics samples is included.

2. Materials

Unless stated otherwise, Materials were obtained from Sigma Aldrich, St. Louis, MO.

2.1. Cell Growth and Isolation

Cellgro Dulbecco's Phosphate Buffered Saline (Mediatech, Mannasas, VA).

2.2. Lysis/Peptide Extraction Reagents

1. Water purified using a NANOpure® or equivalent system ($\geq$ 18 M$\Omega\times$cm, Barnstead International, Dubuque, Iowa).

2. Ammonium bicarbonate, isopropanol, and methanol (Sigma Aldrich).
3. Protease inhibitor cocktail formulated for use with bacterial cell extracts (Cat. No. P8465, Sigma)
4. 0.1 mm zirconia/silica beads (BioSpec Products, Bartlesville, OK) were used for cell lysis.
5. 10–20% Tris-Tricine Ready Gels® (Bio-Rad, Hercules, CA) and GelCode® Blue reagent from Pierce for SDS-PAGE analyses.
6. OMIX® C-18 tips (100 μL) (Varian, Inc, Palo Alto, CA) for sample solid-phase extraction (SPE) clean-up prior to MS analysis.
7. SpeedVac (Thermo Fisher Scientific, Waltham, MA) to concentrate samples.

2.3. Liquid Chromatography–Mass Spectrometry/Mass Spectrometry

1. Ion trap mass spectrometers (LTQ, Thermo Fisher Scientific, San Jose, CA)
2. Water purified using a NANOpure® or equivalent system ($\geq$ 18 MΩ×cm)
3. Mobile phase A: Degassed 0.2% acetic acid, 0.05% trifluoroacetic acid in water (Sigma Aldrich)
4. Mobile phase B: Degassed 0.1% trifluoroacetic acid in 90% acetonitrile (ACN), 10% water (Sigma Aldrich)
5. 5-μm Jupiter C_{18} stationary phase (Phenomenex, Torrence, CA) packed into 60-cm (360 μm o.d. X 150 μm i.d.) fused silica capillary tubing (Polymicro Technologies Inc., Phoenix, AZ)
6. Liquid chromatography system is described elsewhere by Livesay et al. (15)

2.4. Liquid Chromatography–High-Resolution Mass Spectrometry

1. Fourier transform ion cyclotron resonance (FTICR) mass spectrometer, either a custom-built 11 T instrument or 9.4 T instrument (Bruker Daltonics, Billerica, MA).
2. *See* **Section 2.3** for details on mobile and stationary-phase materials.

2.5. LC–MS/MS Data Analyses

1. SEQUEST® version [TurboSEQUEST® (cluster) v.27 (rev. 12), Thermo Fisher Corp.]

.

2.6. Proteomics

1. RapiGest™ (Waters, Milford, MA) is a surfactant to aid in the solubilization and trypsin digestion of proteins.
2. Trypsin for protein digestion (Promega, Madison, WI)

3. Bicinchoninic Acid (BCA) Protein Assay kit (Pierce, Rockford, IL) for quantitation of peptides

2.7. Data Visualization and Cluster Analysis

1. *DAnTE*, freely available software for comparative analysis of proteomics data available at http://omics.pnl.gov/software/
2. *MultiExperiment Viewer* (MEV) is also freely available and designed for use in microarray experiments, but can be particularly useful for proteomics data visualization and clustering and is available at http://www.tm4.org/mev.html.

3. Methods

3.1. Culturing Conditions

The culturing conditions of the bacteria are not the focus of this review but are summarized here. The primary difference between the culture conditions used is that various forms of stresses, some relevant to pathogenesis, were compared relative to a rich growth medium at middle logarithmic growth phase. Wild-type *S.* typhimurium strains 14028 and LT2 were grown to mid-logarithmic (Log) and stationary (Stat) phases in Luria-Bertani (LB) broth and harvested for analysis. Two other cell growth conditions were used that differed only in the pre-growth conditions. In one, the bacteria were grown to stationary phase in LB, the bacteria were isolated, washed, and then grown in magnesium-minimal acidic medium (Shock); in the other, the bacteria were diluted 1:100 and grown in acidic minimal media overnight (Dilu). All cultures were harvested following standard batch culture techniques as outlined (*see* references (3–5) for more detail of culture methods). Aliquots of cell cultures (corresponding to 0.15 g cell pellets) were pelleted, washed in PBS, flash frozen with liquid N_2, and used as needed to prepare samples.

3.2. Sampling Preparation and Peptide Extraction

The procedures outlined here are specific to samples that require Biosafety Level 2 (BSL2) containment and treatment. Many of the precautions (e.g., O-ring sealed cryovials, cooling following vortexing) are to prevent aerosolization of unlysed pathogenic organisms. When developing these protocols, we lysed cells in the presence of a protease inhibitor cocktail formulated for use with bacterial cell extracts. However, we did not evaluate corresponding analyses without the protease inhibitor cocktail (*see* **Note 1**). The listing of class-specific chemical inhibitors of proteases found in the excellent review by Overall and Blobel (6) may be consulted if protease inhibition is desired. In our previous work, we

performed tests to mimic poor sample handling (incubation at 22°C for 20 min without inhibitors) and compared these results to those obtained when the samples were prepared at ~7°C with a cooling block (normal handling temperatures with inhibitors) (*see* **Note 2**). We found no significant variation in the peptides identified. The procedure below is based on an isopropanol extraction that causes larger proteins to precipitate while endogenous peptides are maintained in solution. Different concentrations of isopropanol were tested and it was determined that a ratio of 3:2 resulted in the best recovery of endogenous peptides from *S.* typhimurium. This may not hold true for all biological samples.

1. Lysis of bacterial cells is accomplished by first resuspending the cell pellet in an equivalent volume of 100 mM NH_4HCO_3, followed by transfer of the sample to a 2.0-mL O-ring sealed cryovial. Next, 0.1-mm zirconia/silica beads are added to half of the volume in the tube, and the tube is then vortexed for 30 s, followed by cooling for 1 min in a cold-block. Six cycles of vortexing and cooling are performed. The lysate is then removed from the top of the settled beads, and the beads are rinsed five times with buffer. The lysates and rinses are then pooled separately in microcentrifuge tubes (*see* **Note 3**).
2. The pooled lysate is centrifuged at 16,000×*g* for 10 min at room temperature to pellet insoluble and precipitated proteins. Transfer the supernatant to a new microcentrifuge tube, and ensure that the entire pellet is left behind. The supernatant is now considered a cleared lysate. An aliquot of the cleared lysate can be saved for SDS-PAGE as a reference.
3. Isopropanol is then added to the cleared lysate in an appropriate ratio (we used 1:1, 3:2, 2:1, or 5:2 (v/v, isopropanol:lysate)), and the samples were mixed by vortexing. Pre-cooling the isopropanol to 4°C before adding to the lysate can assist with precipitation of proteins. The samples are then incubated at 4°C for 15 min, then microcentrifuged at 16,000×*g* for 10 min at 4°C to remove precipitated proteins. The resulting supernatants are transferred to new microcentrifuge tubes and concentrated in a SpeedVac to ~75 μL. Ten-microliter aliquots can be removed at this time for SDS-PAGE.
4. Peptide concentrations are determined by BCA protein assay, and SDS-PAGE analyses are performed using 10–20% Tris-Tricine Ready Gels®. The Tris-Tricine gels are used because they are specific for the separation of extremely small proteins and peptides. Gels are fixed for 30 min in 40% methanol/10% acetic acid and then stained for 60 min using GelCode® Blue reagent.

5. Prior to MS analysis, the concentrated isopropanol extracts are cleaned via solid-phase extraction using OMIX C18 pipette tips. These tips are monolithic, rather than particulate, and are therefore much easier to use without clogging, while providing better recovery and reproducibility. Thirty micrograms of peptide mass from each sample is applied to a 100-μL tip. The directions provided by the manufacturer are used to condition, wash, and load the samples. Peptides are eluted from the tips with 80:20 ACN:H_2O containing 0.1% TFA. Eluted peptides are concentrated to ~15 μL in a SpeedVac.
6. Alternatively, samples can be fractionated using strong cation exchange (SCX) HPLC to minimize sample complexity prior to each LC–MS/MS analysis, as described previously (7). Each fractionation is performed using approximately 150 μg (peptide mass) of concentrated isopropanol extract, resulting in 25 fractions that are concentrated in a SpeedVac to dryness. The samples are then reconstituted in 25 mM NH_4HCO_3 to a volume appropriate for LC–MS/MS analysis.

3.3. Liquid Chromatography–Mass Spectrometry/Mass Spectrometry

Our analytical instrumentation consists of commercially available platforms [e.g., ion traps (LTQ from ThermoFisher) and FTICR–MS (BrukerDaltonics)] that are in-house modified to increase the sensitivity and throughput of the analyses. However, the below LC–MS(/MS) approaches can be applied at a reasonable level of quality with more generally available off-the-shelf instrumentation. LC–MS/MS analyses are useful for making identifications and for semi-quantitation based on "spectrum counting" techniques (4, 8–10). These analyses are also used to build a database of identified peptides annotated with determined reversed-phase elution times (11) and calculated masses. This database (also referred to as a mass and time tag lookup table) is used with results from the high-resolution MS analyses (**Section 3.4**) to increase throughput, perform label-free quantitation, and improve peptide-sampling methods in the MS experiment. This is a simplified description of the accurate mass and time (AMT)-tag process developed in our laboratory, which has been extensively discussed elsewhere (12–14).

1. The concentrated C18 SPE eluents from the peptide clean-up procedure and the SCX fractions are then analyzed by reversed-phase microcapillary HPLC (15) interfaced through nanoelectrospray ionization (nanoESI) to an ion trap mass spectrometer, as described previously (4). Briefly, the technique used in our laboratory entails gradient elution of peptides over 100 min using a 360 μm OD × 150 μm ID × 65 cm long capillary column packed with 5 μm Jupiter C18 particles.

2. For typical "bottom-up" proteomics experiments, in which the proteins are digested with trypsin, the charge states of peptides detected during LC–MS/MS are typically +2 and +3. Detected peptides are then fragmented using collision-induced dissociation. It should be noted that the peptides detected from the *S.* typhimurium endogenous peptidome include more +4 and +5 charge states than typically observed for other sample types. Due to the larger number of higher charged species, electron transfer dissociation may be considered for future analyses of the endogenous peptidome.

3.4. LC–MS Analyses

1. Concentrated C18 SPE eluents are also analyzed in our laboratory by reversed-phase microcapillary HPLC–nanoESI–FTICR–MS (11.5 T) (16). The same chromatographic platforms are used for LC–MS/MS analyses as is used with the FTICR–MS, and during analysis of multiple samples to be compared, the same chromatography column and electrospray emitter is preferred. This reduces the number of confounding variables during an experiment for downstream data analysis.
2. The analysis order for an experiment such as this needs to be addressed to minimize the effects of analysis time and possibility of carryover from highly abundant peptides. This is referred to as "randomized block design" and is meant to remove experimental nuisance factors that can obscure true differences between samples (*see* **Note 4**). These blocks typically contain one replicate for each experiment and the order of the analyses within a block is randomized.
3. Peptides from the LC–MS spectra are identified using the AMT tag approach (14), including any peptides with +4 and +5 charge states. The necessary software tools are publicly available (http://omics.pnl.gov). This approach uses the calculated mass and the observed normalized elution time (NET) of each filter-passing peptide identification (*see* **Section 3.5**) from the previous LC–MS/MS analyses to construct a reference database of AMT tags. Features from LC–MS analyses (i.e., m/z peaks deconvoluted of isotopic and charge state effects and then annotated by mass and NET) are matched (13) to AMT tags to identify peptides in a manner that results in roughly 5% false-positive identifications. For each protein, the sum of its peptide peak areas (NET vs. peak height) is used as a measure of the abundance of its fragments within the peptidome.

3.5. LC–MS/MS Data Analyses

Peptides can be identified using a number of different publicly available software packages.

1. In this example, we utilize *SEQUEST®* to search the resulting MS/MS spectra against the annotated *S.* typhimurium FASTA data file of proteins translated from genetic code provided by the J. Craig Venter Institute – formerly TIGR (4550 protein sequences, http://www.jcvi.org/) (17). These analyses used a standard parameter file with a peptide mass tolerance = 3, fragment ion tolerance = 0, and no amino acid modifications. Also, these analyses search for all possible peptide termini (i.e., not limited to only tryptic termini). Separate *SEQUEST®* searches that use the above FASTA data file but with scrambled amino acid sequences are performed in parallel to estimate the false discovery rate.
2. *SEQUEST®* generally returns multiple peptide identifications for each MS/MS spectrum and for each parent ion charge state. Therefore, for each MS/MS spectrum and for each parent ion charge state, only the peptide identification with the highest XCorr value (i.e., the "top ranked hit") is retained here.
3. Limiting false identification of peptides is an especially challenging issue for natively produced peptides because cleavage state (i.e., trypsin cleavage sites) is often used in making confident identifications. PeptideProphet (18) values are also not applicable because of a strong bias for "tryptic" peptides. The estimated percentage of false-positive peptide identifications can be defined as $\%FP_{est.} = 100\% \times$ (number of scrambled peptide identifications) / (number of normal peptide identifications) (19). $\%FP_{est.}$ should be calculated for each charge state, XCorr_Cutoff value (the minimum XCorr value requirement, which ranged from 1.5 to 5 in units of 0.02), and ΔCn_Cutoff value (i.e., the minimum ΔCn value requirement, which ranged from 0 to 0.4 in units of 0.005). In an effort to maximize identifications, a two-dimensional analysis of the XCorr_Cutoff and ΔCn_Cutoff is used for each parent ion charge state. This method is different from typical proteomics analyses in that it does not use a single ΔCn_Cutoff value.
4. The optimal XCorr_Cutoff and ΔCn_Cutoff values for each parent ion charge state (+1 to +5) was determined in our previous work to be 1.84 and 0.21 (+1), 2.1 and 0.21 (+2), 2.8 and 0.23 (+3), 3.56 and 0.265 (+4), and 4.16 and 0.22 (+5), respectively.
5. A rough measure of the abundance of each parent protein and its fragments within the peptidome can be attained using a spectrum counting (i.e., tallying of filter-passing peptide identifications) approach (20).

3.6. Comparison to Proteomics

Peptidomics data (samples acquired without digestion) should ideally be compared to proteomics data (samples acquired using typical bottom-up proteomics approaches including the use of trypsin) from the same source material. This comparison ensures that the peptidomics results are interpreted and can be compared with peptides resulting from abundant proteins being nonspecifically degraded. We performed a proteomics analysis with the same starting sample material to that used in the peptidomics experiment (4). Briefly, proteins are isolated and digested as described in the protocol provided by Waters with the modification of 2.0% TFA rather than using concentrated HCL to adjust to a pH of 3.0. Acid incubation occurred at 37°C for 1 h to fully precipitate the RapiGest™ surfactant. The samples are centrifuged in a microcentrifuge at full speed to pellet the RapiGest™ and the supernatant is returned to neutral pH with NH_4OH to allow for digested peptide concentration determination by BCA protein assay.

The resultant peptides are then fractionated using strong SCX HPLC (7) into 25 fractions. A single unfractionated sample and the full set of 25 SCX fractions are then analyzed by reversed-phase LC–MS/MS. MS/MS spectra are searched using SEQUEST® and filtered to reduce false-positive peptide identifications (3, 4, 20).

3.7. Data Visualization and Cluster Analysis

The comparative interpretation of the identified proteins and peptides can present unique challenges. In the case of comparing environmentally induced changes in the *S.* typhimurium proteome and peptidome, one challenge is that many proteins are not commonly observed across all conditions. If one generates a matrix of protein/peptides (rows) by experimental conditions (column) populated with values of spectral observations or peak abundance measurements, the unobserved proteins/peptides are sometimes referred to as "missing data". The source of an unobserved species can be the result of either of the following: (1) its actual absence in a sample, (2) it is present, but below the detection limit of the mass spectrometer, or (3) the identification did not pass various quality thresholds used for confident peptide identifications. This results in a less than ideal direct application of statistical methods typically used for comparisons of high-throughput data (microarrays) such as an analysis of variance (ANOVA). For this reason, we typically try to combine the abundance values for all peptides from a source protein into a single representative protein abundance for comparison across conditions. This collapsing of peptide abundance to protein abundance is often referred to by us as "protein roll-up" (*see* **Note 5**). These protein values are then grouped by similar abundance profile

changes using methods such as a hierarchical clustering, which are common for microarray analysis comparisons. The comparative analyses of the peptide and protein abundances are enabled with the use of data mining tools that offer clustering and heatmap visualization of the matrix form of the experimental results, e.g., *DAnTE* (21), *OmniViz*® (22), or *MeV* (23). Some considerations that must be made when analyzing the data are listed below:

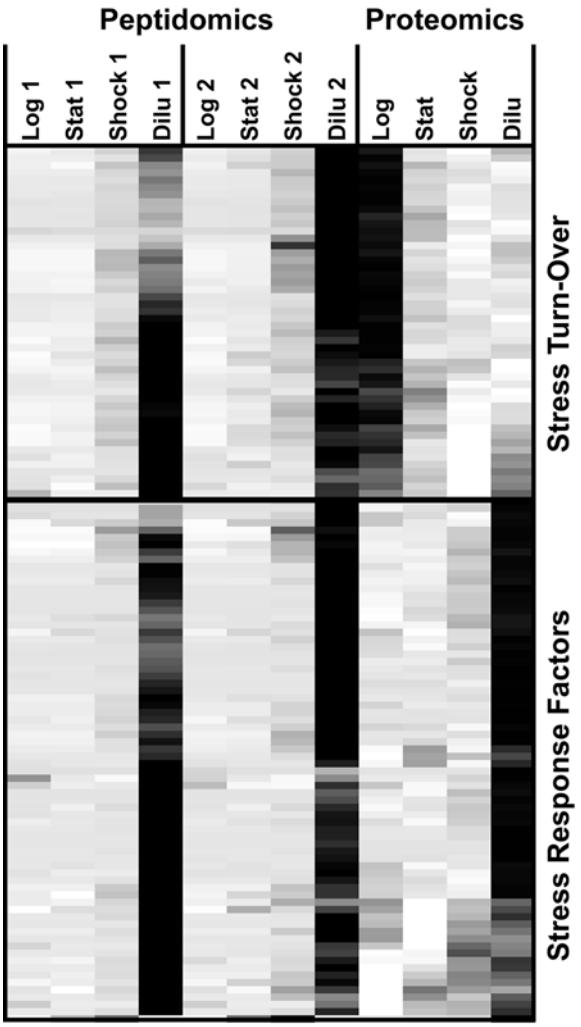


Fig. 2.1. Example heat map showing endogenous peptidomics results compared to global proteomics results. Observations across conditions were scaled using the Z-score across protein (with black representing a Z-score of 2.5 and white a Z-score of –1.0). Two selected regions were taken from data found elsewhere. "Stress response factors", in this case endogenously occurring peptides, correspond well with the abundance of the proteins in the proteomics experiments. The "stress turn-over" peptides appear to be scavenged in the "Dilu" stress condition, and these proteins appear to only be overly abundant in the rich logarithmic growth condition.

1. One of the first decisions is whether to fill arbitrary values into the unobserved peptide/protein abundances to make the analysis more amenable to various downstream data analysis methods typically applied in transcriptional microarray data analysis, such as ANOVA, principle component analysis, and/or clustering methods. If the number of spectra observed in a protein are used as a surrogate for an abundance measurement, filling might include applying the

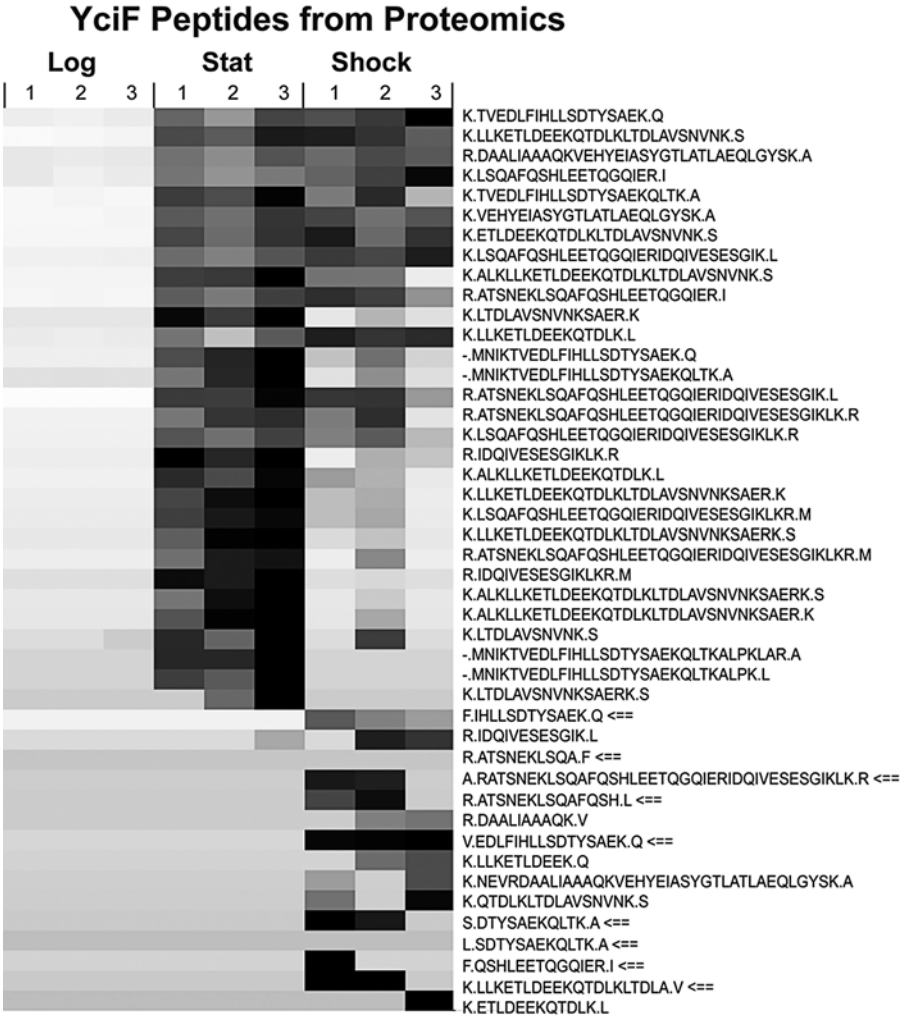


Fig. 2.2. A demonstration of proteomics results in the context of endogenous peptidomics. Although tryptic digestion was use in this example, the disappearance of a number of peptides between the stationary (stat) and shock conditions indicates that the protein is being differentially acted upon by proteases in the cell between the two conditions. This is especially true when this protein YciF was observed to be particularly abundant in the peptidome in the shock condition previously (3). As a secondary confirmation, new peptides that were not observed in the stationary condition appear in the shock condition. New "partially tryptic peptides" also appear and are highlighted with '<==' in the figure. The numbers under the conditions represent the biological replicate of that growth condition.

minimum number of required peptides for protein identification (*see* **Note 6**).

2. In both the peptide-centric and protein-centric (using a single abundance value for the protein) analysis, the difference in abundance between the most abundant peptide/protein versus the least abundant species may range several orders of magnitude. This large dynamic range of measurements may lead to difficulty comparing proteins with similar trends in a set of experiments. To use clustering tools, this dynamic range must be compensated for by scaling to similar magnitudes for comparison (i.e., a trend that is varied across 2 orders of magnitude should be grouped with other similar trends varying across 2 orders of magnitude even if the most abundant value to least abundant value between protein is across 6 orders of magnitude). Depending on the nature of quantitation (spectrum count versus peak area) and the number of experiments being compared (fewer than six versus thirty or more), different scaling approaches are preferred (*see* **Note 7**).
3. Once these steps are performed, comparisons between the experimental samples (both from the undigested native peptidome and the digested proteome) can be performed using heat maps of the clustered results (**Fig. 2.1**).
4. Once an endogenous peptidome analysis has been performed, and knowledge of proteins that are subject to native proteolysis is obtained, it is then possible to extract some additional information utilizing only a proteomics (i.e., trypsin was used) analysis by looking for non-tryptic cleavage sites (for an example **Fig. 2.2**).

4. Notes

1. It is reasonable to consider the goals of the experiment, there may be a specific desire to leave a class of proteases active to amplify the abundance of the cleaved products.
2. Set cooling block between 6 and 8°C, leaving the cooling blocks in the refrigerator 1 day prior to the experiment. Be sure to confirm that freezing will not occur by using microcentrifuge tubes of ~100 μL of water in the block during cooling.
3. All microcentrifuge tubes from this point forward should be siliconized (Fisher 02-681-332) to prevent polymer contamination, which is detrimental to downstream LC–MS(/MS) analyses.

4. The USA National Institute of Standards and Technology maintains an electronic *Engineering Statistics Handbook* (http://www.itl.nist.gov/div898/handbook) with a useful discussion of "Randomized block designs" for experiments.
5. "Protein roll-up" refers to methods that attempt to give a single value for each protein for quantitative purposes, even though each protein identification in a bottom-up proteomics experiment typically is based on more than one peptide identification. As of this writing, *DAnTE* offers multiple methods for protein roll-up (21).
6. Typically, an identification of a specific protein based on its tryptic cleavage products requires identification of three separate tryptic peptides. For native peptidomics this is not realistic because there is a high likelihood that only a single species will be present. Biological conclusions based on single peptide identifications should be based on methods with better relative abundance measurements such as the spectral peak abundance.
7. For large experiments, a Z-score (24) analysis can be helpful to visualize significant trends that are further than expected by a normal distribution. This is also better suited for peak area-based quantitation where the values are non-integers. For smaller experiments, dividing each value in a peptide or protein row by the associated sum, mean, or median of that entire row can be a useful method to scale the results.

Acknowledgments

This work was supported by the National Institute of Allergy and Infectious Diseases (NIH/DHHS through interagency agreement Y1-AI-4894-01 and Y1-AI-8401-01). The authors also acknowledge the US Department of Energy Office of Biological and Environmental Research and National Center for Research Resources (RR18522) for the development of the instrumental capabilities used for the research. Significant portions of this research were performed in the Environmental Molecular Sciences Laboratory, a US Department of Energy (DOE) national scientific user facility located at the Pacific Northwest National Laboratory (PNNL) in Richland, Washington. PNNL is a multiprogram national laboratory operated by Battelle Memorial Institute for the DOE under Contract No. DE-AC05-76RLO-1830.

References

1. El-Sharoud, W.M. and Niven, G.W. (2007) The influence of ribosome modulation factor on the survival of stationary-phase *Escherichia coli* during acid stress. *Microbiology* **153**, 247–253.
2. Wada, A., Mikkola, R., Kurland, C.G., and Ishihama, A. (2000) Growth phase-coupled changes of the ribosome profile in natural isolates and laboratory strains of *Escherichia coli*. *J. Bacteriol.* **182**, 2893–2899.
3. Manes, N.P., Gustin, J.K., Rue, J., Mottaz, H.M., Purvine, S.O., Norbeck, A.D., Monroe, M.E., Zimmer, J.S., Metz, T.O., Adkins, J.N., Smith, R.D., and Heffron, F. (2007) Targeted protein degradation by *Salmonella* under phagosome-mimicking culture conditions investigated using comparative peptidomics. *Mol. Cell Proteomics* **6**, 717–727.
4. Adkins, J.N., Mottaz, H.M., Norbeck, A.D., Gustin, J.K., Rue, J., Clauss, T.R., Purvine, S.O., Rodland, K., Heffron, F., and Smith, R.D. (2006) Analysis of the *Salmonella typhimurium* proteome through environmental response toward infectious conditions. *Mol. Cell Proteomics* **5**, 1450–1461.
5. Beuzon, C.R., Meresse, S., Unsworth, K.E., Ruiz-Albert, J., Garvis, S., Waterman, S.R., Ryder, T.A., Boucrot, E., and Holden, D.W. (2000) Salmonella maintains the integrity of its intracellular vacuole through the action of SifA. *EMBO J.* **19**, 3235–3249.
6. Overall, C.M. and Blobel, C.P. (2007) In search of partners: linking extracellular proteases to substrates. *Nat. Rev. Mol. Cell Biol.* **8**, 245–257.
7. Adkins, J.N., Varnum, S.M., Auberry, K.J., Moore, R.J., Angell, N.H., Smith, R.D., Springer, D.L., and Pounds, J.G. (2002) Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. *Mol. Cell Proteomics* **1**, 947–955.
8. Gao, J., Friedrichs, M.S., Dongre, A.R., and Opiteck, G.J. (2005) Guidelines for the routine application of the peptide hits technique. *J. Am. Soc. Mass Spectrom.* **16**, 1231–1238.
9. Gao, J., Opiteck, G.J., Friedrichs, M.S., Dongre, A.R., and Hefta, S.A. (2003) Changes in the protein expression of yeast as a function of carbon source. *J. Proteome Res.* **2**, 643–649.
10. Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., and Mann, M. (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell Proteomics* **4**, 1265–1272.
11. Monroe, M.E., Shaw, J.L., Daly, D.S., Adkins, J.N., and Smith, R.D. (2008) MASIC: A software program for fast quantitation and flexible visualization of chromatographic profiles from detected LC–MS(/MS) features. *Comput. Biol. Chem.* **32**(3), 215–217.
12. Norbeck, A.D., Monroe, M.E., Adkins, J.N., Anderson, K.K., Daly, D.S., and Smith, R.D. (2005) The utility of accurate mass and LC elution time information in the analysis of complex proteomes. *J. Am. Soc. Mass Spectrom.* **16**, 1239–1249.
13. Monroe, M.E., Tolic, N., Jaitly, N., Shaw, J.L., Adkins, J.N., and Smith, R.D. (2007) VIPER: an advanced software package to support high-throughput LC–MS peptide identification. *Bioinformatics* **23**, 2021–2023.
14. Zimmer, J.S., Monroe, M.E., Qian, W.J., and Smith, R.D. (2006) Advances in proteomics data analysis and display using an accurate mass and time tag approach. *Mass Spectrom. Rev.* **25**, 450–482.
15. Livesay, E.A., Tang, K., Taylor, B.K., Buschbach, M.A., Hopkins, D.F., LaMarche, B.L., Zhao, R., Shen, Y., Orton, D.J., Moore, R.J., Kelly, R.T., Udseth, H.R., and Smith, R.D. (2008) Fully automated four-column capillary LC–MS system for maximizing throughput in proteomic analyses. *Anal. Chem.* **80**, 294–302.
16. Gorshkov, M.V., Pasa Tolic, L., Udseth, H.R., Anderson, G.A., Huang, B.M., Bruce, J.E., Prior, D.C., Hofstadler, S.A., Tang, L., Chen, L.Z., Willett, J.A., Rockwood, A.L., Sherman, M.S., and Smith, R.D. (1998) Electrospray ionization-Fourier transform ion cyclotron resonance mass spectrometry at 11.5 tesla: instrument design and initial results. *J. Am. Soc. Mass Spectrom.* **9**, 692–700.
17. Peterson, J.D., Umayam, L.A., Dickinson, T., Hickey, E.K., and White, O. (2001) The comprehensive microbial resource. *Nucleic Acids Res.* **29**, 123–125.
18. Keller, A., Nesvizhskii, A.I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392.
19. Qian, W.J., Liu, T., Monroe, M.E., Strittmatter, E.F., Jacobs, J.M., Kangas, L.J., Petritis, K., Camp, D.G., 2nd, and Smith,

R.D. (2005) Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: the human proteome. *J. Proteome Res.* **4**, 53–62.

20. Qian, W.J., Jacobs, J.M., Camp, D.G., 2nd, Monroe, M.E., Moore, R.J., Gritsenko, M.A., Calvano, S.E., Lowry, S.F., Xiao, W., Moldawer, L.L., Davis, R.W., Tompkins, R.G., and Smith, R.D. (2005) Comparative proteome analyses of human plasma following in vivo lipopolysaccharide administration using multidimensional separations coupled with tandem mass spectrometry. *Proteomics* **5**, 572–584.
21. Polpitiya, A.D., Qian, W.J., Jaitly, N., Petyuk, V.A., Adkins, J.N., Camp, D.G., 2nd, Anderson, G.A., and Smith, R.D. (2008) DAnTE: a statistical tool for quantitative analysis of -omics data. *Bioinformatics* **24**, 1556–1558.
22. Saffer, J.D., Burnett, V.L., Chen, G., and van der Spek, P. (2004) Visual analytics in the pharmaceutical industry. *IEEE Comput. Graph Appl.* **24**, 10–15.
23. Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., and Quackenbush, J. (2003) TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34**, 374–378.
24. Adkins, J.N., Monroe, M.E., Auberry, K.J., Shen, Y., Jacobs, J.M., Camp, D.G., 2nd, Vitzthum, F., Rodland, K.D., Zangar, R.C., Smith, R.D., and Pounds, J.G. (2005) A proteomic study of the HUPO Plasma Proteome Project's pilot samples using an accurate mass and time tag strategy. *Proteomics* **5**, 3454–3466.

Chapter 3

Approaches to Identify Endogenous Peptides in the Soil Nematode *Caenorhabditis elegans*

Steven J. Husson, Elke Clynen, Kurt Boonen, Tom Janssen, Marleen Lindemans, Geert Baggerman, and Liliane Schoofs

Abstract

The transparent soil nematode *Caenorhabditis elegans* can be considered an important model organism due to its ease of cultivation, suitability for high-throughput genetic screens, and extremely well-defined anatomy. *C. elegans* contains exactly 959 cells that are ordered in defined differentiated tissues. Although *C. elegans* only possesses 302 neurons, a large number of similarities among the neuropeptidergic signaling pathways can be observed with other metazoans. Neuropeptides are important messenger molecules that regulate a wide variety of physiological processes. These peptidergic signaling molecules can therefore be considered important drug targets or biomarkers. Neuropeptide signaling is in the nanomolar range, and biochemical elucidation of individual peptide sequences in the past without the genomic information was challenging. Since the rise of many genome-sequencing projects and the significant boost of mass spectrometry instrumentation, many hyphenated techniques can be used to explore the "peptidome" of individual species, organs, or even cell cultures. The peptidomic approach aims to identify endogenously present (neuro)peptides by using liquid chromatography and mass spectrometry in a high-throughput way. Here we outline the basic procedures for the maintenance of *C. elegans* nematodes and describe in detail the peptide extraction procedures. Two peptidomics strategies (off-line HPLC–MALDI-TOF MS and on-line 2D-nanoLC–Q-TOF MS/MS) and the necessary instrumentation are described.

Key words: Nematode, *Caenorhabditis elegans* , neuropeptide, insulin, FMRFamide-like peptide, *flp* , neuropeptide-like protein, G-protein-coupled receptor, mass spectrometry.

1. Introduction

1.1. Caenorhabditis elegans Is an Ideal Model Organism

The transparent, free-living, non-parasitic soil nematode *Caenorhabditis elegans* (*Caeno,* recent; *rhabditis,* rod; *elegans,* nice) of only 1 mm in length can be safely handled and is easy to

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_3,

grow and maintain. Since its introduction as a model system in the 1960s, *C. elegans* was used widely in many research laboratories due to the ease of handling and the well-defined anatomy. *C. elegans* contains exactly 959 cells that are ordered in sets of fully differentiated tissues. There are two sexes. Hermaphrodites can self-fertilize or mate with males in order to produce over 300 offspring. Although hermaphrodites are the most common sex in nature, mating with males will yield a 50% male progeny. In the laboratory, self-fertilization of the hermaphrodites or crossing with males can easily be manipulated for genetic studies. In addition, *C. elegans* has a short life cycle. It takes about 3–4 days from egg to egg and it goes through four larval stages (L1–L4) until reaching adulthood. A developmentally arrested "dauer" larva can be formed under conditions of starvation or overcrowding. These thinner dauers have a relative impermeable cuticle, are non-feeding, and can survive for months, in contrast to the average life span of around 2–3 weeks under standard conditions.

A sophisticated knowledge infrastructure has been developed, with many research methods and protocols that are widely shared in the "worm-community." Most information can be found in the easily accessible database "WormBase" at (http://www.wormbase.org). The "WormBook" (http://www.wormbook.org) can be considered as the open-access collection of peer-reviewed chapters that covers all kinds of different topics and protocols related to *C. elegans.* This nematode is also perfectly suited for light microscopy due to its transparency. For high-end visual analysis of *C. elegans,* the microscope has to be equipped with differential interference contrast (DIC; Nomarski) optics for obtaining 3D-like view of the tissues. This way, individual neurons can be observed and recognized. As an example, a DIC image of an L1 larva is shown in **Fig. 3.1**. Detailed DIC and electron microscopic images are available on "WormAtlas" (http://www.wormatlas.org), together with a plethora of detailed schematic representations. *C. elegans* was the first multicellular organism to have its genome fully sequenced (1). Its genome (about 100 Mb) encodes for over 20,000 proteins and its size is about 1/30th of that of a human. The awarding of the Nobel Prize to the three "worm-pioneers" Sydney Brenner, Robert Horvitz, and John Sulston in 2002 for their discoveries concerning genetic regulation of organ development and programmed cell death, to Andrew Fire and Craig Mello in 2006 for their discovery of RNA interference in this nematode, and to Martin Chalfie in 2008 for the discovery and development of the green fluorescent protein (GFP), emphasizes the great potential of this tiny nematode of being a model organism, just like the fruit fly *Drosophila melanogaster,* the marine snail *Aplysia californica*, and the mouse *Mus musculus.*

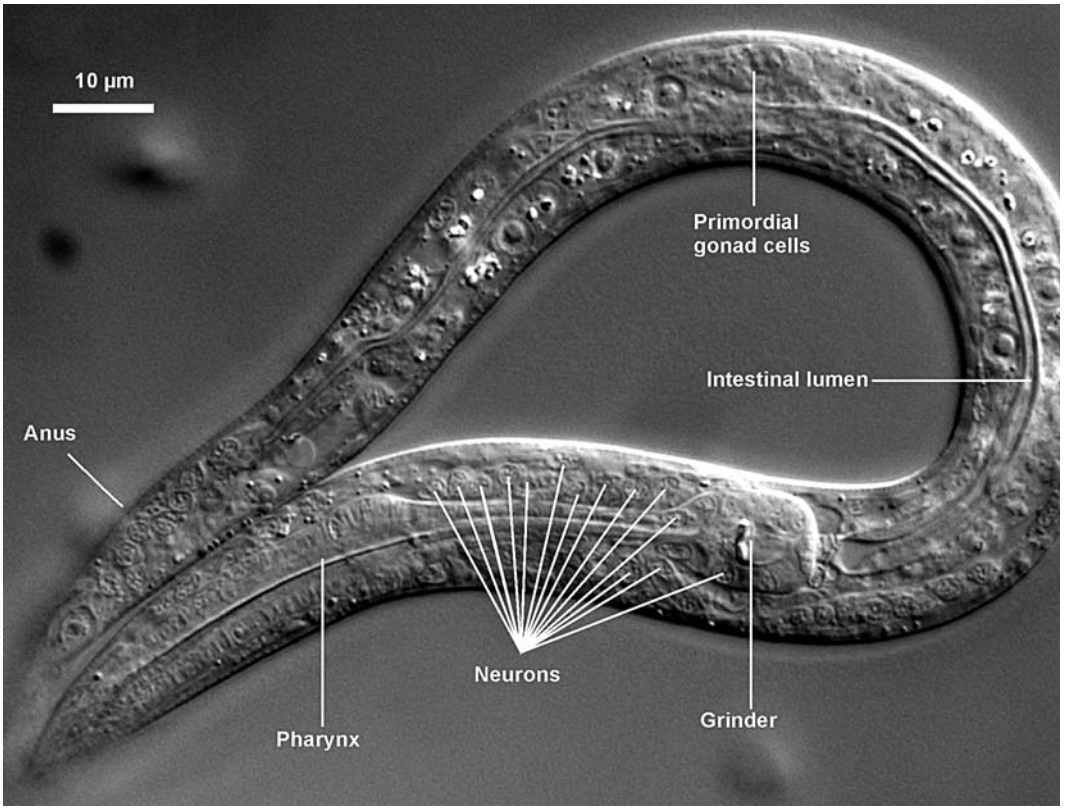


Fig. 3.1. Differential interference contrast image of a *C. elegans* L1 larva. The first larval stage of the nematode *C. elegans* is shown. This picture was taken using an Axio Observer Z1 instrument (Zeiss) equipped with differential interference contrast (DIC) or Nomarski optics to allow a clear 3D-like structure of individual neurons.

1.2. Peptidomics of C. elegans

(Neuro)peptides are small messenger molecules that are derived from larger precursor proteins by the highly controlled action of processing enzymes. These biologically active peptides can be found in all metazoan species where they orchestrate a wide variety of physiological processes. The knowledge of the primary amino acid sequence of the neuropeptidergic signaling molecules is absolutely necessary to understand their function and interactions with G-protein-coupled receptors. Three classes of neuropeptide-encoding genes have been predicted from the genomic data of *C. elegans*. Initially, 24 FMRFamide-like peptide (*flp*) genes have been found by searching cDNA libraries and genomic sequences (2–4); more *flp* genes were identified by mining the EST data (5) (*see* **Table 3.1**). By searching the *C. elegans* genome for predicted proteins with the structural hallmarks of neuropeptide precursors, 32 so-called neuropeptide-like protein (*nlp*) genes have been identified (6) (*see* **Table 3.2**). These neuropeptide preproproteins all contain peptides without the RFamide motif, but display sequence homology with other

Table 3.1
FLP neuropeptides of *C. elegans*

Gene		Peptide sequence[a]	Gene		Peptide sequence[a]
-LRFa family			*-MRFa family*		
flp-1		**SADPNFLRFa**	*flp-3*		SPLGTMRFa
		SQPNFLRFa			**TPLGTMRFa**
		ASGDPNFLRFa			**EAEEPLGTMRFa**
		SDPNFLRFa			**NPLGTMRFa**
		AAADPNFLRFa			**ASEDALFGTMRFa**
		(**K**)**PNFLRFa**			EDGNAPFGTMRFa
		AGSDPNFLRFa			**SAEPFGTMRFa**
flp-14	4×	**KHEYLRFa**			**SADDSAPFGTMRFa**
flp-15		**GGPQGPLRFa**			**NPENDTPFGTMRFa**
		RGPSGPLRFa	*flp-6*	6×	**KSAYMRFa**
flp-18		(**DFD**)**GAMPGV LRFa**	*flp-20*	2×	AMMRFa
		EMPGVLRFa	*flp-22*	3×	**SPSAKWMRFa**
	3×	(**SYFDEKK**)**SVP GVLRFa**	*flp-27*		(**EASAFGDIIGELKGK**) **GLGGRMRFa**
		EIPGVLRFa	*flp-28*		**APNRVLMRFa**
		SEVPGVLRFa	*-VRFa family*		
		DVPGVLRFa	*flp-7*	3×	**SPMQRSSMVRFa**
flp-21		GLGPRPLRFa		2×	**TPMQRSSMVRFa**
flp-23		TKFQDFLRFa			**SPMERSAMVRFa**
flp-26		(**E**)**FNADDLTLRFa**			**SPMDRSKMVRFa**
		GGAGEPLAFSPD MLSLRFa	*flp-9*	2×	**KPSFVRFa**
-IRFa family			*flp-11*		**AMRNALVRFa**
flp-2		**SPREPIRFa**			**ASGGMRNALVRFa**
		LRGEPIRFa			**NGAPQPFVRFa**
flp-4		(**GLRSSNGK**) **PTFIRFa**	*flp-16*	2×	**AQTFVRFa**
		ASPSFIRFa			**GQTFVRFa**
flp-5		**GAKFIRFa**	*flp-17*	2×	KSAFVRFa
		AGAKFIRFa	*flp-19*		**WANQVRFa**
		APKPKFIRFa			**ASWASSVRFa**
flp-8	3×	**KNEFIRFa**	*flp-24*		**VPSAGDM**(ox)**M**(ox)**VRFa**
flp-10		pQPKARSGYIRFa	*flp-25*		**DYDFVRFa**
flp-12		RNKFEFIRFa	*flp-32*		**AMRNSLVRFa**
flp-13		(**SDRPTR**)**AMD SPFIRFa**	*-PRFa family*		

(continued)

Table 3.1 (continued)

Gene	Peptide sequence[a]	Gene	Peptide sequence[a]
	AADGAPFIRFa	*flp-33*	**APLEGFEDMSGFLRTIDGI QKPRFa**
	APEASPFIRFa		
	ASPSAPFIRFa		
	SPSAVPFIRFa		
	ASSAPFIRFa		
	SAAAPLIRFa		
flp-17	KSQYIRFa		
flp-25	**ASYDYIRFa**		

[a]Sequences shown in bold have been confirmed by Edman degradation, MALDI-TOF MS, or Q-TOF mass spectrometry.

Table 3.2
NLP neuropeptides of *C. elegans*

Gene		Peptide sequence[a]	Gene	Peptide sequence[a]
nlp-1	×3	**MDANAFRMSFa**	*nlp-21*	GGARAMLH
		MDPNAFRMSFa		GGARAFSADVGDDY
		VNLDPNSFRMSFa		**GGARAFYDE**
nlp-2		SIALGRSGFRPa		**GGARAFLTEM**
		SMAMGRLGLRPa		**GGARVFQGFEDE**
	×3	SMAYGRQGFRPa		GGARAFMMD
nlp-3		AINPFLDSMa		GGGRAFGDMM
		AVNPFLDSIa		GGARAFVENS
		YFDSLAGQSLa		GGGRSFPVKP GRLDD
nlp-4		SLILFVILLVAFA AARPVSEEVDRV		**pQYTSELEEDE**
		DYDPRTEAPRRLPA DDDEVDGEDRV	*nlp-22*	SIAIGRAGFRPa
		DYDPRTDAPIRVPV DPEAEGEDRV	*nlp-23*	LYISRQGFRPA
nlp-5		**SVSQLNQYAGFD TLGGMGLa**		SMAIGRAGMRPa
		ALSTFDSLGGMGLa		AFAAGWNRa
		ALQHFSSLDTL GGMGFa	*nlp-24*	pQWGGGPYGGYGP
nlp-6		(MA)**APKQMVFGFa**		GYGGGYGGa
		YKPRSFAMGFa		YGGYGa
		AAMRSFNMGFa		FTGPYGGYGa
		LIMGLa		GPYGYGa
nlp-7		pQADFDDPRMFTSSFa		GPYGGGGLVGALLa

(continued)

Table 3.2 (continued)

Gene		Peptide sequence[a]	Gene		Peptide sequence[a]
		SMDDLDDPRL MTMSFa	*nlp-25*		IGTEVAEGVLVA EEVSEAIa
		MILPSLADLH RYTMYD			GGGYGGGYGGGFGA QQAYNVQNAA
		LYLKQADFDDP RMFTSSFa	*nlp-26*		pQFGFGGQQSFGa
nlp-8		**AFDRFDNSGV FSFGA**			GGQFGGMQ
		AFDRMDNSDFFGA			GGFNGN
		SFDRMGGT EFGLM			**GGFGQQSQFGa**
		YPYLIFPASPSS GDSRRLV		×2	GGNQFGa
nlp-9		GGARAFYGF YNAGNS			GGSQFNa
		GGGRAFNHN ANLFRFD			GGFGFa
		GGGRAFAGSWSPYLE	*nlp-27*		pQWGYGGMPYGGYGGM GGYGMGGYGMGY
		TPIAEAQGAPE DVDDRRELE			MWGSPYGGYGGY GGYGGWa
nlp-10		AIPFNGGMYa	*nlp-28*		GYGGYa
		STMPFSGGMYa			GYGGYGGYa
		AAIPFSGGMYa		×2	GYGGYGGYa
		GAMPFSGGMYa			GMYGGWa
nlp-11		**HISPSYDVEIDAG NMRNLLDIa**	*nlp-29*		pQWGYGGYa
		SAPMASDYGN QFQMYNRLIDAa			GYGGYGGYa
		SPAISPAYQFENA FGLSEALERAa		×3	GMYGGYa
nlp-12	×2	**DYRPLQFa**			GMYGGWa
		DGYRPLQFa	*nlp-30*		pQWGYGGYa
nlp-13		**NDFSRDIMSFa**			GYGGYGGYa
		SGNTADLYDR RIMAFa			GYGGYa
		pQPSYDRDIMSFa			GMWa
		SAPSDFSRD IMSFa			PYGGYGWa
		SSSMYDRDIMSFa	*nlp-31*		pQWGYGGYa
		SPVDYDR PIMAFa		×2	GYGGYGGYa
		AEDYERQIMAFa			GYGGYa
nlp-14	×2	ALDGLDGSGFGFD			GMYGGYa
	×5	**ALNSLDGAGFGFE**			PYGGYGWa

(continued)

Table 3.2 (continued)

Gene		Peptide sequence[a]	Gene		Peptide sequence[a]
	×3	ALDGLDGAGFGFD	*nlp-32*		YGGWGa
		ALNSLDGQGFGFE			GGWa
	×3	ALNSLDGNGFGFD			GGa
nlp-15		**AFDSLAGSGFDNGFN**			GYGa
	×2	AFDSLAGSGFGAFN			GGGWGa
		AFDSLAGSGFSGFD			GGGWa
		AFDSLAGQGFTGFE			GGGa
		AFDTVSTSGFDDFKL			FGYGGa
nlp-16		STEHHRV			GWa
		SEGHPHE	*nlp-33*		pQWGYGGPYGGYG GGYGGGPWGYGGGW
		ATHSPEGHIVA KDDHHGHE			HWGGYGGGPWGG YGGGPWGGYY
		SSDSHHGHQ	*nlp-34*		PYGYGGYGGW
		SVDEHHGHQ			PYGYGWa
		NAEDHHEHQ	*nlp-35*		**AVVSGYDNIYQVLAPRF**
		SEHVEHQAEM HEHQ	*nlp-36*		DDDVTALERWGY
		STQEVSGHP EHHLV			NIDMKLGPH
nlp-17		**GSLSNMMRIa**			**SMVARQIPQT VVADH**
		pQQEYVQFPNEGVV PCESCNLGTLMRIa	*nlp-37*		**NNAEVVNHILK NFGALDRLGDVa**
nlp-18		**SPYRAFAFA**	*nlp-38*		(**ASDDR**)**VLGWNKAHGLWa**
		ARYGFA			**TPQNWNKLNSLWa**
		SPYRTFAFA			**SPAQWQRANGLWa**
		ASPYGFAFA	*nlp-39*		**EVPNFQADNV PEAGGRV**
		SDEENLDFLE	*nlp-40*		**APSAPAGLEEKL**(**R**)
nlp-19		IGLRLPNFLRF			MVAWQPM
		IGLRLPNML	*nlp-41*		**APGLFELPSRSV**(**RLI**)
		MGMRLPNIFLRNE	*nlp-42*		SALLQPENNPEWNQLGWAWa
nlp-20		FAFAFA			NPDWQDLGFAWa
		SGPQAHEGA GMRFAFA	*nlp-43*	×2	KQFYAWAa
		APKEFARFARASFA	*nlp-44*		APHPSSALLVPYPRVa
					LYMARVa
					AFFYTPRIa
			nlp-45		RNLLVGRYGFRIa
			nlp-46		NIAIGRGDGLRPa
			nlp-47		PQMTFTDQWT

[a]Sequences shown in bold have been confirmed by Edman degradation, MALDI-TOF MS, or Q-TOF mass spectrometry.

invertebrate neuropeptides. Finally, a systematic search for genes encoding members of the insulin superfamily revealed the presence of 40 insulin-like genes (7). Neuropeptidergic signaling in the nematode *C. elegans* has recently been reviewed (8). Based on such sequence information alone, one cannot deduce whether all the predicted peptides are actually expressed and properly processed. Therefore, each such neuropeptide needs to be purified and characterized biochemically. In the past, biochemical purification and elucidation of neuropeptide sequences required multiple chromatographic separation steps to purify an individual biologically active peptide. This approach appeared to be problematic, especially for small-sized animals, such as *C. elegans*. Previously, only 12 neuropeptides of *C. elegans* could be biochemically isolated and identified using Edman degradation analysis or gas-phase sequencing (9–14). Recently we set out to systematically search for and characterize neuropeptides of *C. elegans* using high-throughput peptidomics techniques. A peptidomics approach aims to identify endogenous (neuro)peptides using liquid chromatography and mass spectrometry. We aimed to elucidate which peptides were actually present in the nematode and to identify any post-translational modifications, which are often required for the peptide's bioactivity. We successfully analyzed the peptidome of *C. elegans* (15, 16), and *C. briggsae* (17), while the *Ascaris suum* peptidome has been explored by others (18, 19). Differential peptidomics techniques allowed us to characterize the neuropeptide precursor processing enzymes EGL-3 (20, 21) and EGL-21 (22) and the neuroendocrine chaperone protein 7B2 (23). In this chapter we mainly focus on the basic techniques and methods required to culture the nematodes and to perform the sample preparation. Then, different technologies that can be used in peptidomic research are described and a short overview is provided of the instrumentation needed.

2. Materials

2.1. C. elegans Culture

1. *C. elegans* strains can be ordered from the *Caenorhabditis* Genetics Center (CGC, http://www.cbs.umn.edu/CGC/), which is supported by the National Institutes of Health–National Center for Research Resources. This center collects, maintains, and distributes all kinds of *C. elegans* strains at a $7 fee per strain in the case of academic/nonprofit organizations or a $100 fee per strain for commercial organizations, in addition to the annual fee of $25. *C. elegans* N2 (Bristol) is referred to as the wild-type

reference strain. The nematodes are sent by regular post as starved cultures on small Petri dishes.

2. *Escherichia coli* OP50 bacteria are also available at the CGC.
3. Nematode Growth Medium (NGM): Dissolve 3 g NaCl, 17 g agar, and 2.5 g peptone in 1 L H_2O. Sterilize by autoclaving, add 1 mL of 1 M $CaCl_2$, 1 mM of 5 mg/mL cholesterol in ethanol, 1 mL of 1 M $MgSO_4$, and 25 mL of 1 M KPO_4. Pour NGM medium in Petri dishes under sterile conditions (*see* **Note 1**).
4. Incubators (15–22°C) (*see* **Note 2**).
5. Drigalski spatula.

2.2. Sample Preparation

1. 60% sucrose solution; sugar can also be used. This solution can be stored at 4°C for a couple of weeks.
2. 0.1 M NaCl solution. Make this solution fresh each time.
3. Extraction solvent: methanol:water:acetic acid (90:9:1), used ice-cold.
4. 50% acetonitrile containing 0.1% trifluoroacetic acid (TFA).
5. Sample reconstitution buffers: 2–5% acetonitrile and 0.1% TFA (for HPLC analysis); 2–5% acetonitrile and 0.1% formic acid (FA) (for nano LC-ESI-Q-TOF MS).
6. *n*-hexane, ethyl acetate.
7. Solid-phase extraction cartridges, such as SepPak C18 cartridge (Waters, Milford, MA).
8. Glass homogenizator, sonicator (Sanyo MSE Soniprep 150 ultrasonic disintegrator or Branson 5510 ultrasonic cleaner).
9. SpeedVac vacuum centrifuge (Savant and Flexi-Dry MP, FTS systems).
10. 22-μm spin filter (Ultrafree-MC, Millipore Corporation, Bedford, MA).

2.3. Peptidomics Analyses

1. High-performance liquid chromatograph (Beckmann, Fullerton, CA) equipped with a programmable solvent module 126 and a Diode Array Detector Module 168 (Gold System).
2. Symmetry C18 column (5 μm, 4.6 × 250 mm, Waters) for use with solvent flow rates of ~1 mL/min. Symmetry C18 column (2.1 × 150 mm, 3.5 μm, Waters) for use with flow rates of ~300 μL/min.
3. Matrix-assisted laser desorption ionization mass spectrometer (MALDI-TOF MS) Reflex IV (Bruker Daltonic

GmbH, Germany); UltraflexII MALDI-TOF MS (Bruker Daltonic GmbH, Germany). Both mass spectrometers are operated using *FlexControl* software. The *FlexAnalysis* program is used to process mass readouts.

4. Standard calibration peptide mixture: Angiotensin 2 (1045.54 Da), angiotensin 1 (1295.68 Da), substance P (1346.73 Da), bombesin (1618.82 Da), ACTH clip 1–17 (2092.08 Da), and ACTH clip 19–39 (2464.19 Da) (Bruker Daltonic GmbH, Germany).
5. *Mascot* search engine (http://www.matrixscience.com).
6. Miniaturized LC system (nanoLC) comprising *Ultimate* HPLC pump, a *Switchos* column-switching device, and a *Famos* autosampler (LC Packings, Amsterdam, the Netherlands).
7. Electrospray quadrupole time-of-flight mass spectrometer (ESI-Q-TOF MS) (Waters-Micromass, Manchester, UK) (*see* **Note 3**).
8. Stainless steel emitter (Proxeon, Odense, Denmark).
9. C18 pre-column (μ-guard column MGU-30 C18, LC-Packings).
10. Strong cation exchange column (Bio-SCX, 500 μm × 15 mm, LC-Packings).
11. Symmetry C18 column (3.5 μm, 75 μm × 100 mm, Waters); PepMap C18 column (3 μm, 75 μm × 150 mm, LC Packings).
12. *ProteinLynx* software (Waters-Micromass).
13. Solvents: Water, CH_3CN, TFA, H_2O. All solvents have to be HPLC grade.
14. Saturated α-cyano-4-hydroxycinnamic acid in acetone.
15. Pre-spotted anchorchip targets (Bruker Daltonics GmbH, Germany).

3. Methods

3.1. Maintenance of C. elegans Cultures

Here, we shortly describe how to get the nematode culture started.

1. *C. elegans* is normally grown using the *E. coli* OP50 as a food source (*see* **Note 4**).
2. OP50 bacteria can be grown using conventional microbiological methods and LB broth at 37°C.

3. Apply 50–100 μL of an overnight grown culture of bacteria on medium-sized NGM plates.
4. Spread bacteria, let them dry, and allow them to grow overnight on the bench (20°C) or in a 37°C incubator to form a nice OP50 lawn (*see* **Notes 5–7**).
5. Equilibrate plate at 20°C before using them for culturing the nematodes.
6. Several methods can be used to transfer the worms from an old plate to a new one in order to expand the mass of nematodes for a peptidomics analysis, or to keep the nematodes in culture. We cut out a small piece of agar from the old plate, containing the worms, and transfer it to a new NGM plate using a sterile scalpel or spatula (*see* **Note 8**). Alternatively, individual animals can also be picked up using a home-made "worm-picker", which is a small platinum wire with a flattened end that is melted into a glass Pasteur pipette.
7. To maintain the worm lines, the worms should be transferred to new plates weekly (*see* **Notes 9 and 10**).

3.2. Sample Preparation

1. Collect the mixed-stage worms from 10–15 fully grown Petri dishes by rinsing the plates with a 0.1 M NaCl solution (*see* **Note 11**).
2. Living animals shall be separated from the *E. coli* bacteria and dead animals by flotation on 30% sucrose or sugar. Add an equal volume of a 60% sucrose or sugar solution to the 0.1 M NaCl solution containing the worms. Centrifuge for 4 min at 500×*g*; the living animals will float on top of the sugar gradient. Harvest the nematodes and wash four times with 0.1 M NaCl (*see* **Note 12**).
3. Transfer the nematodes to 15 mL of an ice-cold extraction solvent (*see* **Note 13**).
4. Homogenize the worms using a glass stick homogenizator and sonicate the solution prior to centrifugation.
5. Discard the pellet, evaporate the methanol using a SpeedVac concentrator.
6. The remaining aqueous solution, containing the peptides, has to be delipidated by re-extraction with ethyl acetate or *n*-hexane (*see* **Note 14**). Add equal volume of organic solvent to the aqueous solution that contains the peptides. Mix by vigorous inversion of the sample, and centrifuge briefly (1 min at 13,000 rpm using a benchtop centrifuge) to separate the phases. Carefully remove and discard the top (organic) layer.
7. Desalt the aqueous solution using solid-phase extraction with a SepPak C18 cartridge (*see* **Note 15**). Activate the

cartridge using 50–100% of CH_3CN, rinse the column using water containing 0.1% TFA, add the aqueous peptide sample. Wash the cartridge with 0.1% TFA in water. Elute the peptides with 50% (or higher) acetonitrile containing 0.1% TFA.

8. The desalted peptide sample shall be stored at 4°C prior to analysis. Alternatively, samples can be lyophilized by using a SpeedVac concentrator and stored at –20°C.
9. Immediately prior to the analysis by HPLC and MALDI-TOF MS, reconstitute the samples in water containing 2–5% acetonitrile and 0.1% TFA and filter them using 22-μm spin filters. For the analysis by nano LC-ESI-Q-TOF MS, samples should be reconstituted in water containing 2–5% acetonitrile and 0.1% FA.

3.3. Peptidomics Analyses

Here we describe two general strategies for the peptidomics analysis of *C. elegans*. The first method is an off-line strategy, in which the generated HPLC fractions are characterized using a MALDI-TOF instrument (summarized in **Fig. 3.2**). This strategy allows an easy comparison of different fractions from various mutant strains and is therefore preferred for differential peptidomics analysis. Peptides of interest can be sequenced later using, for example, MALDI-TOF/TOF MS. The other approach relies on a high-throughput two-dimensional separation of the peptide extract and the automated MS and MS/MS measurements using an ESI-Q-TOF instrument (summarized in **Fig. 3.3**). Using that on-line approach, the peptidomes of the fruitfly *D. melanogaster* (24) and the nematode *C. elegans* (15) have been successfully characterized in our lab.

3.3.1. Off-Line HPLC–MALDI-TOF MS

1. Inject the peptide extract and wash the column for 10 min using 4% acetonitrile in 0.1% TFA (*see* **Note 16**).
2. Start a linear gradient of 4% acetonitrile in 0.1% TFA to 50% CH_3CN in 0.1% TFA (60 min). Endogenous peptides tend to elute between 22 and 37% of acetonitrile (*see* **Note 17**).
3. Collect fractions eluted from HPLC once every minute (*see* **Note 18**).

3.3.1.1. High-Performance Liquid Chromatography (HPLC) (*see* **Note 16**)

3.3.1.2. Matrix-Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-TOF MS)

This off-line HPLC–MALDI-TOF MS approach allows fast screening of the peptide content of different *C. elegans* strains as the mass readouts can be compared easily. We found 75 peptides using this robust peptidomics protocol (17, 20–23).

1. Vacuum dry one-fifth to one-half of each of the generated HPLC fractions and reconstitute each in 1 μL of 50% acetonitrile in 0.1% TFA prior to applying them to the ground steel target MALDI plate.

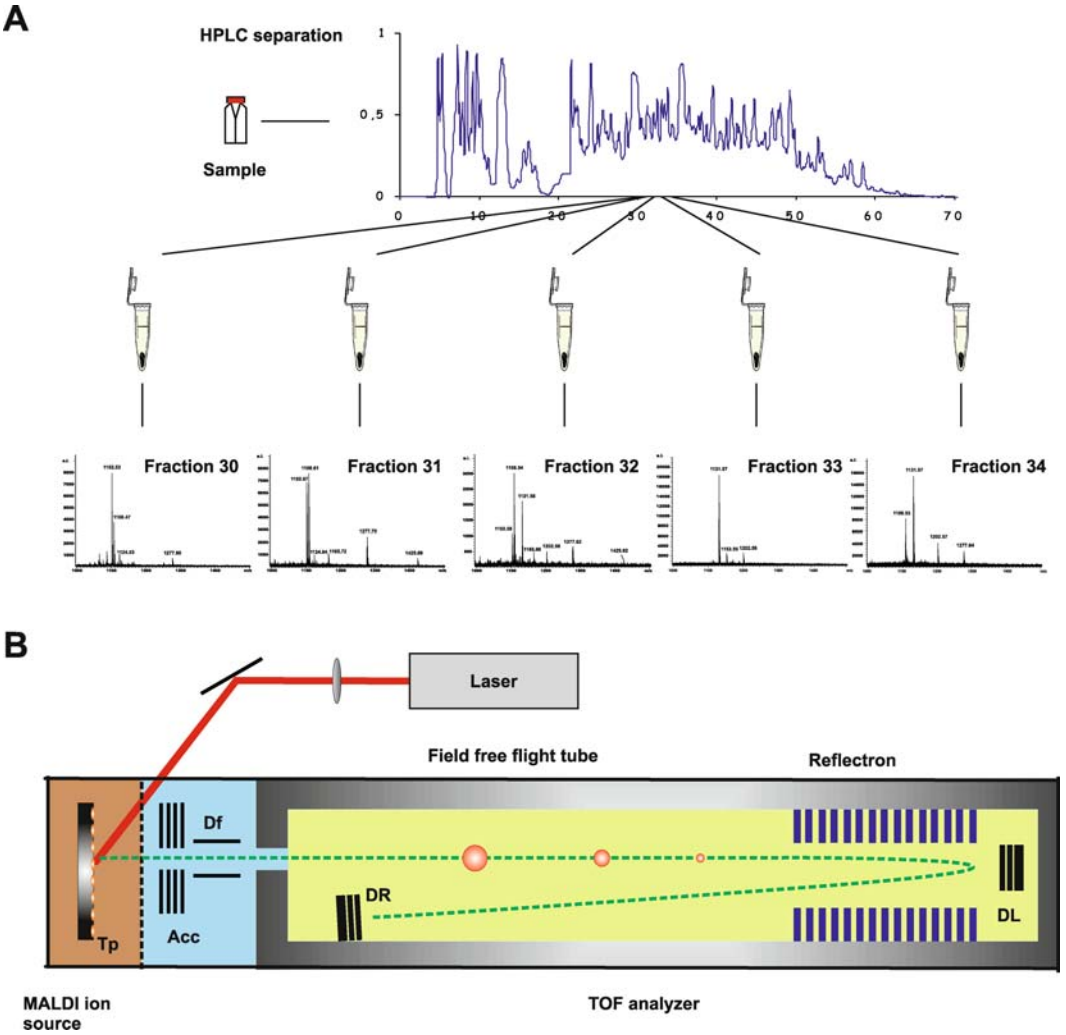


Fig. 3.2. Overview of the off-line HPLC–MALDI-TOF MS workflow. (**a**) The peptide extract was separated using a reversed-phase C18 column to generate a chromatogram as shown. Absorbance was monitored at 214 nm. Each HPLC fraction was then analyzed by MALDI-TOF mass spectrometry to generate a peptide profile. Only fractions 30–34 are shown. (**b**) Schematic representation of a typical MALDI-TOF instrument. All samples are deposited on a stainless steel target plate, together with an UV-absorbing matrix like α-cyano-4-hydroxycinnamic acid. When a pulsed laser beam hits the target plate, an ion plume is generated. Next, the ions are accelerated by an electrostatic field that is applied on the acceleration plates (Acc), and guided through the deflectors (Df) before entering the field-free flight tube. This time-of-flight (TOF) analyzer measures the time an accelerated ion needs to reach the detector at the end of the flight tube. These data can be converted into *m*/*z* units as the kinetic energies of all ions in the flight tube are equal. When measuring in "reflectron mode", an electrostatic mirror lengthens the flight path to increase the resolution and mass accuracy.

2. Mix the droplets with the saturated solution of α-cyano-4-hydroxycinnamic acid in acetone (*see* **Note 19**). Dry the target plate and insert it into the mass spectrometer.

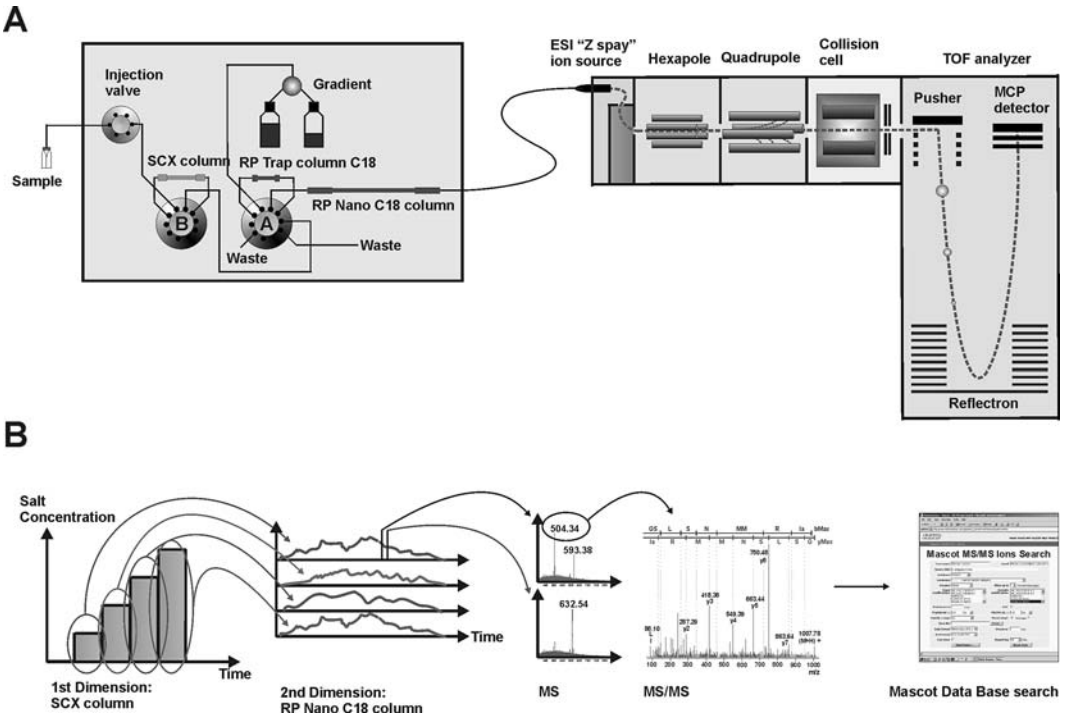


Fig. 3.3. Overview of the on-line 2D-nanoLC–Q-TOF MS/MS workflow. (**a**) Schematic representation of the hardware used: an Ultimate high pressure LC pump, a Switchos column-switching device, a Famos autosampler (all LC Packings) and a quadrupole – time-of-flight mass spectrometer (Q-TOF) (Micromass-Waters). Two nanoscale columns (a strong cation exchange (SCX) column and a reversed-phase C18 column) are placed in line. Each fraction that elutes from the first SCX column will undergo a subsequent separation on the second reversed-phase column. This way, ten successive separations are performed. The eluent is directly connected to the Q-TOF mass spectrometer. Individual ions are formed in the electrospray source (Z-spray ESI source), which are guided through the hexapole (six parallel rods) to enter the quadrupole (four parallel rods) mass filter. This Q-TOF instrument allows a selection of particular ions in the first quadrupole (narrow bandpass mode), while the other non-resonant ions get lost. After fragmentation of the selected ion by collision with an inert gas in the collision cell, the generated fragments are measured in the time-of-flight (TOF) analyzer to generate the fragmentation or MS/MS spectrum. This TOF analyzer is equipped with a reflectron (to lengthen the flight path) and a multi-channel plate (MCP) detector. (**b**) Visualization of the data obtained. All spectra are converted into typical peak list files which can be submitted to a bioinformatics program that matches the experimental data against any protein database. For our work, we used a home-made database containing the predicted neuropeptide precursors of *C. elegans*.

3. Calibrate the instrument using a standard peptide mixture containing angiotensin 2, angiotensin 1, substance P, bombesin, ACTH clip 1–17, and ACTH clip 19–39.
4. Record spectra using the reflectron mode within a mass range of 500–3000 Da. Adjust the laser intensity to obtain optimal resolution and sensitivity.
5. Mass readouts can automatically be processed in the *FlexAnalysis* program to obtain peak list files. Experimental *m/z* values can then be compared with the theoretical masses of the predicted peptides (*see* **Note 20**).

3.3.2. On-Line 2D-NanoLC ESI Q-TOF MS/MS

The main advantage of this approach is that the peptides are automatically sequenced in a high-throughput manner. Using this method we sequenced ~60 endogenous peptides (15, 16); these peptides are indicated in bold font in **Tables 3.1 and 3.2**.

1. Load 20 μL of the peptide sample (corresponding to two fully grown NGM plates) onto a strong cation exchange column (Bio-SCX, 500 μm × 15 mm) using 2% acetonitrile in 0.1% FA and the flow rate of 30 μL/min. This cation exchange column was placed on-line with a C18 pre-column or trapping column (μ-guard column MGU-30 C18, LC-Packings).
2. After loading the sample, the SCX column should be switched off-line, and the reversed-phase pre-column should be rinsed for 5 min.
3. Switch the reversed-phase trapping column on-line with the nanoscale Symmetry C18 column (3.5 μm, 75 μm × 100 mm) or a PepMap C18 column (3 μm, 75 μm × 150 mm). Separate the peptides using a linear gradient from 2% to 50% acetonitrile containing 0.1% FA at a flow rate of 200 nL/min for 50 min.
4. Elute the second fraction of peptides from the SCX column by injecting 20 μL of a 20 mM ammonium acetate solution. Concentrate and desalt these peptides again on the C18 pre-column prior to the nanoscale HPLC and MS analysis.
5. Repeat this elution procedure ten times using different salt plugs of ammonium acetate (0, 20, 50, 100, 200, 400, 600, 800, 1000, and 2000 mM).
6. The 2D-LC system should be connected directly to the electrospray interface of the Q-TOF mass spectrometer through a stainless steel emitter.
7. The mass spectrometer should be set to automatic data-dependent MS to MS/MS switching when the intensity of the doubly and triply charged parent ions increases above 15 counts/s. The applied collision energy of the argon gas should be chosen automatically (between 25 and 40 eV) depending on the number of charges and the mass range of the selected parention.
8. Transform the MS/MS data of all ten SCX fractions into *pkl* (peak list) files using the *ProteinLynx* software.
9. Submit these text files to a *Mascot* search to identify the peptides (*see* **Note 21**).

4. Notes

1. We prefer to use Petri dishes that allow air to float under the lid as the nematodes need oxygen to survive. Depending on the amount of plates needed, a peristaltic pump can be used to pour the NGM.
2. *C. elegans* is normally cultured at 20°C. Depending on the planning of the extractions, temperature can be lowered or increased to slow down or speed up the growth. Nematode cultures can also be stored on the bench when a constant room temperature of about 20°C is maintained.
3. The nanoLC column was directly coupled to the ESI-Q-TOF MS.
4. This bacterial strain is uracil auxotroph and thus has a limited growth on NGM plates.
5. It is very important not to damage the NGM surface as the worm will tend to crawl into the agar. Also, when spreading the bacteria, try not to cover the total surface of the plate as the nematodes will crawl up the sides of the plate and die when the bacterial lawn reaches the edges of the Petri dish.
6. Depending on the experiments planned, the bacteria can be grown for longer or shorter periods. In order to get more nematodes, we prefer to extend the incubation time to produce a thicker bacterial lawn. Also, conventional LB agar (35 g/L) can be used instead of 3 g NaCl, 17 g agar, and 2.5 g peptone in 1 L H_2O as described in **Section 2.1**.
7. (Seeded) plates may be stored at 4°C for a couple of weeks, although it is better to use fresh plates.
8. This technique is referred to as "chunking." Worms will crawl out of the chunk and a typical sinusoidal "footprint" is generated by the worms. The worms can easily be visualized using a dissecting microscope or a stereo microscope. This method is preferred when a large numbers of nematodes are required, e.g., when starting a new peptidomics experiment.
9. This frequency will depend on the size of the chunks, the dimensions of the Petri dishes, and the growth temperature.
10. For a typical off-line HPLC–MALDI-TOF MS experiment, we use 10–20 fully grown Petri dishes (90 mm diameter) of *C. elegans*. Two plates of the starting material should be sufficient for an on-line 2D-nanoLC ESI Q-TOF MS/MS setup.

11. Be careful not to damage the NGM surface when collecting the nematodes.
12. Ten to fifteen fully grown Petri dishes will yield a pellet containing ~500 μL of living nematodes.
13. This extraction solvent is specially designed to extract small endogenous peptides, while larger proteins precipitate. When interested in larger peptides (5–15 kDa) such as the insulin-like peptides, diluted acids might be a better extraction solvent. All steps have to be performed on ice to avoid degradation of the proteins. Active peptidases result in degradation of proteins and might result in shortened and/or fragmented peptides, which are obviously not of interest.
14. Both solvents for re-extraction of the peptide extract perform equally well in our hands, but may have ramifications with other peptidomics experiments. If lots of lipids appear to be present, extraction with both organic solvents can be performed.
15. Other solid-phase cartridges may be used, e.g., Oasis HLB extraction cartridges (10 mg, Waters, Milford, MA). These are a good alternative to the SepPak C18 solid-phase extraction cartridges. The HLB column is equilibrated with methanol and then with water. After loading the aqueous solution of peptides, the cartridge is washed with water containing 5% methanol. Finally, peptides are eluted with 100% methanol.
16. Many different HPLC columns are available, we prefer a Symmetry C18 (5 μm, 4.6 × 250 mm) column that operates at a solvent flow-rate of 1 mL/min. Depending on the amount of starting material, a smaller Symmetry C18 column (2.1 × 150 mm, 3.5 μm) with a flow rate of 300 μL/min might be used (15, 17, 20, 22, 23).
17. Three-step gradient may be used at this step. For example: from 2% to 22% acetonitrile (in 0.1% TFA) for 20 min, followed by 22–37% acetonitrile (in 0.1% TFA) for 30 min, followed by 37–50% acetonitrile (in 0.1% TFA) for 10 min.
18. We prefer to collect the generated HPLC fractions automatically from the beginning of the (three-step) gradient.
19. We prefer to use α-cyano-4-hydroxycinnamic acid as matrix, because it is ideally suited for use with small peptides. If higher sensitivity is required, pre-spotted anchorchip targets can be used.
20. When using a LIFT/TOF or TOF/TOF instrument (like the Ultraflex II), fragmentation of ion peaks of interest can yield sequence information. MS/MS spectra can be

analyzed by de novo sequencing. However, because a good protein database of *C. elegans* is available, we prefer to use search engines such as "Mascot."

21. Our in-house *Mascot* server matches the fragmentation data from the peak list files against our home-made database containing all known FLP and NLP precursors. Individual ions with *Probability Based Mowse Scores* above the threshold ($P<0.05$) are further analyzed and annotated to gain sequence information.

Acknowledgments

Research in the authors' lab was sponsored by the Research Foundation Flanders (FWO-Vlaanderen grant G.0434.07 and 1.5.137.06). The authors strongly acknowledge the Interfacultary Centre for Proteomics and Metabolomics "Prometa", K.U. Leuven, and wish to thank the *Caenorhabditis* Genetics Center for providing all the *C. elegans* strains. S.J. Husson, T. Janssen, M. Lindemans and E. Clynen are postdoctoral fellows of the Research Foundation Flanders (FWO-Vlaanderen).

References

1. The *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2018.
2. Kim, K. and Li, C. (2004) Expression and regulation of an FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. *J. Comp. Biol.* **475**, 540–550.
3. Li, C., Nelson, L.S., Kim, K., Nathoo, A., and Hart, A.C. (1999) Neuropeptide gene families in the nematode *Caenorhabditis elegans*. *Ann. N.Y. Acad. Sci.* **897**, 239–252.
4. Li, C., Kim, K., and Nelson, L.S. (1999) FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. *Brain Res.* **848**, 26–34.
5. McVeigh, P., Leech, S., Mair, G.R., Marks, N.J., Geary, T.G., and Maule, A.G. (2005) Analysis of FMRFamide-like peptide (FLP) diversity in phylum Nematoda. *Int. J. Parasitol.* **35**, 1043–1060.
6. Nathoo, A.N., Moeller, R.A., Westlund, B.A., and Hart, A.C. (2001) Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species. *Proc. Natl. Acad. Sci. USA.* **98**, 14000–14005.
7. Pierce, S.B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S.A., Buchman, A.R., Ferguson, K.C., Heller, J., Platt, D.M., Pasquinelli, A.A., Liu, L.X., Doberstein, S.K., and Ruvkun, G. (2001) Regulation of DAF-2 receptor signaling by human insulin and *ins-1*, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes Dev.* **15**, 672–686.
8. Husson, S.J., Mertens, I., Janssen, T., Lindemans, M., and Schoofs, L. (2007) Neuropeptidergic signaling in the nematode *Caenorhabditis elegans*. *Prog. Neurobiol.* **82**, 33–55.
9. Marks, N.J., Shaw, C., Maule, A.G., Davis, J.P., Halton, D.W., Verhaert, P., Geary, T.G., and Thompson, D.P. (1995) Isolation of AF2 (KHEYLRFamide) from *Caenorhabditis elegans*: evidence for the presence of more than one FMRFamide-related peptide-encoding gene. *Biochem. Biophys. Res. Commun.* **217**, 845–851.
10. Marks, N.J., Maule, A.G., Geary, T.G., Thompson, D.P., Davis, J.P., Halton, D.W., Verhaert, P., and Shaw, C. (1997) APEASPFIRFamide, a novel FMRFamide-related decapeptide from *Caenorhabditis elegans*:

structure and myoactivity. *Biochem. Biophys. Res. Commun.* **231**, 591–595.
11. Marks, N.J., Maule, A.G., Geary, T.G., Thompson, D.P., Li, C., Halton, D.W., and Shaw, C. (1998) KSAYMRFamide (PF3/AF8) is present in the free-living nematode, *Caenorhabditis elegans*. *Biochem. Biophys. Res. Commun.* **248**, 422–425.
12. Marks, N.J., Maule, A.G., Li, C., Nelson, L.S., Thompson, D.P., Alexander-Bowman, S., Geary, T.G., Halton, D.W., Verhaert, P., and Shaw, C. (1999) Isolation, pharmacology and gene organization of KPSFVR-Famide: a neuropeptide from *Caenorhabditis elegans*. *Biochem. Biophys. Res. Commun.* **254**, 222–230.
13. Marks, N.J., Shaw, C., Halton, D.W., Thompson, D.P., Geary, T.G., Li, C., and Maule, A.G. (2001) Isolation and preliminary biological assessment of AADGAPLIR-Famide and SVPGVLRFamide from *Caenorhabditis elegans*. *Biochem. Biophys. Res. Commun.* **286**, 1170–1176.
14. Rosoff, M.L., Doble, K.E., Price, D.A., and Li, C. (1993) The *flp-1* propeptide is processed into multiple, highly similar FMRFamide-like peptides in *Caenorhabditis elegans*. *Peptides* **14**, 331–338.
15. Husson, S.J., Clynen, E., Baggerman, G., De Loof, A., and Schoofs, L. (2005) Discovering neuropeptides in *Caenorhabditis elegans* by two dimensional liquid chromatography and mass spectrometry. *Biochem. Biophys. Res. Commun.* **335**, 76–86.
16. Husson, S.J., Clynen, E., Baggerman, G., De Loof, A., and Schoofs, L. (2005) Peptidomics of *Caenorhabditis elegans*: in search of neuropeptides. *Commun. Agric. Appl. Biol. Sci.* **70**, 153–156.
17. Husson, S.J., Landuyt, B., Thomas, N., Baggerman, G., Boonen, K., Clynen, E., Lindemans, M., Janssen, T., and Schoofs, L. (2008) Comparative peptidomics of *Caenorhabditis elegans* versus *C. briggsae* by LC–MALDI-TOF MS. *Peptides*, **30**, 449–457.
18. Yew, J.Y., Dikler, S., and Stretton, A.O. (2003) De novo sequencing of novel neuropeptides directly from *Ascaris suum* tissue using matrix-assisted laser desorption/ionization time-of-flight/time-of-flight. *Rapid Commun. Mass Spectrom.* **17**, 2693–2698.
19. Yew, J.Y., Kutz, K.K., Dikler, S., Messinger, L., Li, L., and Stretton, A.O. (2005) Mass spectrometric map of neuropeptide expression in *Ascaris suum*. *J. Comp. Neurol.* **488**, 396–413.
20. Husson, S.J., Clynen, E., Baggerman, G., Janssen, T., and Schoofs, L. (2006) Defective processing of neuropeptide precursors in *Caenorhabditis elegans* lacking proprotein convertase 2 (KPC-2/EGL-3): mutant analysis by mass spectrometry. *J. Neurochem.* **98**, 1999–2012.
21. Husson, S.J. and Schoofs, L. (2006) Characterization of a key neuropeptide processing enzyme in *C. elegans* by mass spectrometry. *Commun. Agric. Appl. Biol. Sci.* **71**, 171–174.
22. Husson, S.J., Janssen, T., Baggerman, G., Bogert, B., Kahn-Kirby, A.H., Ashrafi, K., and Schoofs, L. (2007) Impaired processing of FLP and NLP peptides in carboxypeptidase E (EGL-21)-deficient *Caenorhabditis elegans* as analysed by mass spectrometry. *J. Neurochem.* **102**, 246–260.
23. Husson, S.J. and Schoofs, L. (2007) Altered neuropeptide profile of *Caenorhabditis elegans* lacking the chaperone protein 7B2 as analyzed by mass spectrometry. *FEBS Lett.* **581**, 4288–4292.
24. Baggerman, G., Boonen, K., Verleyen, P., De Loof, A., and Schoofs, L. (2005) Peptidomic analysis of the larval *Drosophila melanogaster* central nervous system by two-dimensional capillary liquid chromatography quadrupole time-of-flight mass spectrometry. *J. Mass Spectrom.* **40**, 250–260.

Chapter 4

Mass Spectrometric Analysis of Molluscan Neuropeptides

Ka Wan Li and August B. Smit

Abstract

The central nervous systems of molluscan species contain high levels of structurally diverse peptides that function as neurotransmitters, neuromodulators or neurohormones. Peptide diversity is believed to be a way to increase the information handling capacity of neurons in the context of a brain with low cell numbers and neuronal connectivity. Accordingly, much effort has been made to identify peptides from single neurons and tissues of interest. In the past decade a mass spectrometry-based approach has been applied to detect and characterize peptides from single neurons, nerves and tissues of the molluscan brain. Peptides from single neurons are often analysed directly by mass spectrometry without prior sample preparation. Single neurons from the molluscan brain may be identified based on their position, cell morphology and colour. Neurons that cannot be readily identified can be tagged functionally or chemically. For the analysis of peptides from tissues, special extraction methods in conjunction with peptide separation by liquid chromatography coupled to mass spectrometry have been developed. Tens to hundreds of peptides from the tissue extract can be detected and characterized in a single analysis.

Key words: Neuropeptides, MALDI mass spectrometry, single-cell analysis, tissue extraction, retrograde labelling.

1. Introduction

Peptidergic neurons constitute the major class of nerve cells in the molluscan brain. Some of the neuropeptides are released from neurohemal areas as hormones. More often, neuropeptides are released from axon endings that closely appose the target cells and function as neuromodulators and/or neurotransmitters and are involved in fast cell–cell communication. The diversity of molluscan peptides is large (1) and estimated to be in the order of several hundreds within a given species. In contrast, peptide diversity

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_4,

in the vertebrate nervous system is low and classical transmitters, such as glutamate and GABA, are preferentially used for neurotransmission.

Peptides are differentially expressed in distinct populations of neurons. While some neurons express a single peptide, other neurons may express a large number of structurally diverse peptides (2). Recent studies revealed the molecular mechanisms that are used to generate peptide diversity in single cells, including differential expression of multiple peptide precursors, alternative splicing of a single precursor, differential processing of peptide domains from a precursor, and different types of posttranslational modifications of peptides such as phosphorylation, glycosylation and hydroxylation (2). Peptide diversity is believed to be a way to increase the information handling capacity of the cells (3). This might be the evolutionary outcome of selecting complex behaviours, such as feeding or reproduction, whilst using a brain with low number of cells and limited neuronal connectivity. Much effort has been made to identify peptides from single neurons and tissues of interest.

Mass spectrometry-based techniques play a major role in the detection and characterization of peptides in molluscan nervous systems (1, 4–6). In cases where peptide diversity is low, for example in a single cell that contains few tens of peptides, sample can be analysed directly by mass spectrometry without prior sample preparation (2). When nerves, organelles or the released peptides from nervous system need to be analysed (1, 6–8), special extraction methods in conjunction with peptide separation by liquid chromatography should be used. This (pre-) fractionation step is aimed to remove the interfering molecules and reduce sample complexity. This should increase the sensitivity and capacity of the mass spectrometric analysis of the peptides.

Peptides may be analysed from different sources ranging from single neurons to nerves and whole tissues such as the reproductive organs (1, 2, 9). As these samples differ greatly in complexity and the ease of extraction, optimized methodologies for each sample type have been developed. In molluscs many giant neurons can be individually identified based on their position, colour and size. These neurons may be picked individually and analysed by mass spectrometry without extra treatment (2). Neurons that cannot be readily identified visually should be tagged functionally or chemically. To our advantage is the fact that neurons, functionally connected to the same target, often share a common nerve, and therefore they can be retrograde-labelled from this nerve (4). These back-filled cells can then be isolated for subsequent analysis. When extracting neuropeptides from tissues, one should avoid conventional homogenization-based extraction methods that extract large number of proteins in addition to neuropeptides. This would complicate the subsequent peptide

fractionation and mass spectrometric analysis. We routinely use acetone extraction (10, 11). Acetone causes partial dehydration of the tissue resulting in the extraction of the small molecules including neuropeptides into the acetone solvent. The majority of proteins remain within the organ.

2. Materials

1. Solvents: 0.1% TFA in water; 60% acetonitrile in 0.1% TFA; acetone/HCl/H_2O solvent (40:1:6).
2. Reversed–phase solid-phase extraction column Supeclean (Supelco).
3. Saline buffer: 4 mM $CaCl_2$, 1.7 mM KCl, 1.5 mM $MgCl_2$, 30 mM NaCl, 5 mM $NaHCO_3$, 10 mM $NaCH_3SO_4$, and buffered with 10 mM HEPES to pH 7.8. All the chemicals are reagent-grade.
4. Saturated dithiooxamide: Add dithiooxamide (Sigma-Aldrich) to ethanol until it is saturated. Keep the supernatant.
5. Nickel-lysine solution: Add 1.7 g $NiCl_2 \times 6H_2O$ and 3.5 g L-lysine to 20 mL H_2O.
6. Sylgard dish (Dow Corning)
7. Vaseline
8. Matrix solution (except for single-cell analysis): Dissolve 7 mg of α-cyano-4-hydroxycinnamic acid (ultra-pure grade, Sigma-Aldrich) in 1 mL of 50% acetonitrile/50% 10 mM ammonium monobasic phosphate (*see* **Notes 1** and **2**).
9. Matrix solution for single-cell analysis: 10 mg/mL 2,5-dihydroxybenzoic acid (Sigma-Aldrich) in acetonitrile (HPLC grade)/water/TFA (50%/50%/0.1%).
10. Solvents for high-performance liquid chromatography (HPLC). Solvent A: 5% acetonitrile in 0.05% TFA; Solvent B: 80% acetonitrile in 0.04% TFA.
11. Nano HPLC system complete with Ultimate LC system (LC-Packing)
12. MALDI target spotter Probot (Dionex)
13. Pipette capable of handling ~0.5 μL volumes.
14. MALDI mass spectrometry: Large variety of MALDI mass spectrometers are available from different vendors. Our protocols have been optimized for use with the 4800 Proteomics Analyzer (Applied Biosystems).

3. Methods

3.1. Single-Cell Analysis

1. Dissect the brain of the mollusc and pin it down on a Sylgard disc containing saline buffer.
2. Carefully remove the connective tissue with a pair of forceps under a stereo microscope.
3. Loosen the neuron of interest from the brain with tiny hooks.
4. Use a glass pipette to pierce through the neuron. Aspirate the cell content into the pipette and transfer to mix with 1 μL drop of matrix solution on a MALDI-metal plate (*see* **Notes 1** and **3**).
5. Let the matrix to dry at room temperature for a few minutes before inserting the sample plate into the mass spectrometer for analysis.

3.2. Retrograde Labelling of Neurons

1. Carefully cut open the skin of the head region to expose the brain.
2. Cut off all the nerves from the brain except the nerve of interest, which should be cut as far away as possible from the brain.
3. Transfer the brain to a dry Sylgard dish.
4. Use several pins to pierce through the connective tissues of the brain onto the Sylgard disc to fix the brain in position.
5. Cut the nerve of interest tens of centimetres away from the brain.
6. Apply a ring of Vaseline around the cut end of the nerve, and within a minute add 1–2 drop of nickel-lysine solution to the ring of Vaseline. The nickel-lysine solution should completely immerse the cut end of the nerve.
7. Seal the Vaseline ring with additional layer of Vaseline to cover the nickel-lysine solution.
8. Add enough saline buffer to the Sylgard disk to submerge the brain and leave it at room temperature overnight.
9. Transfer the brain to another Sylgard dish containing saline buffer.
10. Wash the brain once in fresh saline buffer.
11. Add the saturated dithiooxamide solution to the Sylgard dish containing the brain. Use 1 drop of dithiooxamide solution per 1 mL of saline buffer.
12. When the retrograde-labelled neurons appear brownish black in colour, they can be removed as shown in **Section 3.1** for analysis.

3.3. Analysis of Neuropeptides from a Single Nerve

1. Dissect the nerve of interest under stereo microscope and transfer the nerve into a 2-μL drop of matrix solution spotted on the MALDI-metal plate (*see* **Note 3**).
2. Break and tear apart the nerve with a pair of forceps while maintaining the nerve in the matrix solution. Neuropeptides diffuse from the damaged nerve into the matrix solution.
3. Remove the nerve debris from the matrix solution with a pair of forceps. The whole procedure should be finished within 2–3 min after applying the matrix on the stainless steel target.
4. Let the matrix to dry at room temperature for a few minutes before inserting the sample plate into the mass spectrometer for analysis.

3.4. Analysis of Neuropeptides from Molluscan Tissues and Organs

1. Dissect tissue and store at –80°C until used. We usually collect 20–50 samples for a single experiment, depending on the weight of the tissue collected.
2. Add 5–10 volumes of extraction solvent, acetone/HCl/H_2O solvent (40:1:6), in a glass beaker and stir overnight at 4°C (*see* **Note 4**).
3. Dilute the solvent containing the neuropeptide extract tenfold with water.
4. Prepare the C18 solid-phase extraction column. Condition the column with 2 volumes of 100% methanol and then wash with 5 volumes of 0.1% TFA.
5. Aspirate the diluted solvent containing the neuropeptides into a 10-mL plastic syringe and slowly inject into the conditioned C18 solid-phase extraction column. The flow rate should not be too high; we usually add the solvent to the column at around 2–4 mL per min. Do not dry the column.
6. Apply the rest of the solvent to the column in 10-mL aliquots.
7. Wash the column with 5 volumes of 0.1% TFA.
8. Elute the neuropeptides from the column with 2–3 volumes of 60% acetonitrile in 0.1% TFA, and collect in a 1.5-mL eppendorf tube (*see* **Note 5**).
9. Dry the neuropeptides in a SpeedVac.
10. Re-suspend the dried sample in 0.1% TFA and fractionate the neuropeptides with HPLC using a nano-C18 column.
11. Collect the fractions on a MALDI-metal plate for mass spectrometric analysis.

3.5. HPLC Separation of Neuropeptides

1. Redissolve the peptides in 20 μL 0.1% TFA.
2. Inject the peptides into a 3-μm nano-C18 LC column.
3. Separate the peptides using a linearly increasing concentration of acetonitrile from 5 to 50% in 30 min and to 100% in 5 min. Set the flow rate to 400 nL/min.
4. Mix the eluent from LC column with matrix (α-cyano-4-hydroxycinnamic acid) delivered at a flow rate of 1.5 μL/min, and deposit off-line to the MALDI-metal plate every 15 s for a total of 192 spots, using an automatic robot, such as Probot or similar others.
5. Analyse the peptides with MALDI MS/MS.

3.6. Maldi Ms/Ms

1. Analyse peptides on an MALDI MS/MS, such as ABI 4800 or similar proteomics analyzer.
2. Perform MS analysis. We usually acquire 1250 MS spectra per fraction/sample (*see* **Note 6**).
3. Select peptides with signal-to-noise ratio above 50 at the MS mode for MS/MS experiment; a maximum of 25 MS/MS is allowed per spot. Set the precursor mass window to 200 (*see* **Note 7**).
4. Perform collision-induced dissociation on each of the peptides at 1 kV with air as the collision gas.
5. Collect MS/MS spectra from 2500 laser shots per peptide.

3.7. Recrystallization of MALDI Matrix

1. Add α-cyano-4-hydroxycinnamic acid to 100 mL ethanol; heat in a boiling water bath until saturation (*see* **Note 8**).
2. Pour the solution into a beaker, and store it at –20°C for 2 days.
3. Matrix appears as yellow precipitate in the solution. Collect the matrix and air dry on a Whatman paper.
4. Break the matrix on the Whatman paper, and transfer to a Buchner funnel.
5. Wash the matrix with a few volumes of ethanol.
6. Weigh the air-dried matrix and put 7 mg into a 1.5-mL Eppendorf tube. Store the matrix at –20°C. The matrix is stable for years.
7. Dissolve matrix in 1 mL solvent for off-line LC analysis.

4. Notes

1. We use two types of matrix for different applications. The preferred matrix for the off-line LC analysis is α-cyano-4-hydroxycinnamic acid. It forms a homogeneous layer of fine

crystals on the MALDI-metal plate, which facilitates automatic MALDI MS and MS/MS analysis.

2. Alternatively, matrix can be purified from analytical grade to high purity by recrystallization (as described in **Section 3.7**). A considerable amount of matrix will be used for LC analysis. It is more economical to re-crystallize reagent-grade matrix rather than to purchase the expensive ultra-pure-grade matrix. This option is especially attractive when larger quantity is required, for example in the case of LC-MALDI MS.
3. For the direct single-cell or nerve MS analysis both 2,5-dihydroxybenzoic acid and α-cyano-4-hydroxycinnamic acid can be used. The advantage of using the 2,5-dihydroxybenzoic acid is that it does not crystallize as fast as α-cyano-4-hydroxycinnamic acid. Therefore, more time will be available to break and mix the cell content in the matrix. However, 2,5-dihydroxybenzoic acid forms inhomogeneous spear-shaped crystals with sweet spots. These sweet spots represent the site where neuropeptides are co-crystallized with the matrix. They are usually formed around the rim of the matrix crystals. The analysis requires manual searching of sweet spots in the spear-shaped crystal. The laser beam is targeted on the crystal and the peptide peak intensity is continuously monitored. If a sweet spot is found, multiple mass spectra can be generated with higher sensitivity.
4. Neuropeptides and other small molecules will be preferentially extracted into the solvent.
5. The amount of neuropeptides is generally low. The low peptide concentration increases the risk of their loss during the sample handling and storage. We use Eppendorf safe-lock 1.5-mL microcentrifuge tubes because they have low peptide absorption. Furthermore, these tubes do not contain low molecular weight contaminants that may interfere with subsequent MALDI MS analysis.
6. A typical MALDI spot can withstand thousands of laser shoots before it is depleted of material. So multiple analyses can be performed on a single spot. To get high MS1 sensitivity it is possible to increase the laser energy until the gain of peak intensity reaches the plateau. A higher number of MS1 spectra can also be summed together to reduce background noise. Whereas a routine MS1 analysis is about 1250 shoots, we occasionally use up to 7000 laser shoots per analysis.
7. It is possible to select peptides with signal-to-noise ratio below 50 at the MS mode for MS/ MS experiment. However, the signal intensity of the MS/MS spectra may be low, and often only a few fragment ions are detected.

Nevertheless, the low number of fragment ions detected may still be useful for some analyses – for example, for the confirmation of peptide identities in samples from, e.g. a single cell, often containing less than ten different peptide sequences (12)

8. Keep adding α-cyano-4-hydroxycinnamic acid until no more can be dissolved. We use about 10 g α-cyano-4-hydroxycinnamic acid per 100 mL ethanol.

References

1. El Filali, Z., Van Minnen, J., Liu, W.K., Smit, A.B., and Li, K.W. (2006) Peptidomics analysis of neuropeptides involved in copulatory behavior of the mollusk *Lymnaea stagnalis*. *J. Proteome Res.* **5**, 1611–1617.
2. Jimenez, C.R., Spijker, S., de Schipper, S., Lodder, J.C., Janse, C.K., Geraerts, W.P., van Minnen, J., Syed, N.I., Burlingame, A.L., Smit, A.B., and Li, K.W. (2006) Peptidomics of a single identified neuron reveals diversity of multiple neuropeptides with convergent actions on cellular excitability. *J. Neurosci.* **26**, 518–529.
3. Brezina, V., Orekhova, I.V., and Weiss, K.R. (1996) Functional uncoupling of linked neurotransmitter effects by combinatorial convergence. *Science* **273**, 806–810.
4. El Filali, Z., Hornshaw, M., Smit, A.B., and Li, K.W. (2003) Retrograde labeling of single neurons in conjunction with MALDI high-energy collision-induced dissociation MS/MS analysis for peptide profiling and structural characterization. *Anal. Chem.* **75**, 2996–3000.
5. Hummon, A.B., Amare, A., and Sweedler, J.V. (2006) Discovering new invertebrate neuropeptides using mass spectrometry. *Mass Spectrom. Rev.* **25**, 77–98.
6. Jimenez, C.R., Li, K.W., Smit, A.B., and Janse, C. (2006) Auto-inhibitory control of peptidergic molluscan neurons and reproductive senescence. *Neurobiol. Aging* **27**, 763–769.
7. Jakubowski, J.A., Hatcher, N.G., and Sweedler, J.V. (2005) Online microdialysis–dynamic nanoelectrospray ionization–mass spectrometry for monitoring neuropeptide secretion. *J. Mass Spectrom.* **40**, 924–931.
8. Jimenez, C.R., ter Maat, A., Pieneman, A., Burlingame, A.L., Smit, A.B., and Li, K.W. (2004) Spatio-temporal dynamics of the egg-laying-inducing peptides during an egg-laying cycle: a semiquantitative matrix-assisted laser desorption/ionization mass spectrometry approach. *J. Neurochem.* **89**, 865–875.
9. Smit, A.B., van Kesteren, R.E., Spijker, S., Van Minnen, J., van Golen, F.A., Jimenez, C.R., and Li, K.W. (2003) Peptidergic modulation of male sexual behavior in *Lymnaea stagnalis*: structural and functional characterization of -FVamide neuropeptides. *J. Neurochem.* **87**, 1245–1254.
10. Van Golen, F.A., Li, K.W., Chen, S., Jimenez, C.R., and Geraerts, W.P. (1996) Various isoforms of myomodulin identified from the male copulatory organ of *Lymnaea* show overlapping yet distinct modulatory effects on the penis muscle. *J. Neurochem.* **66**, 321–329.
11. Van Golen, F.A., Li, K.W., De Lange, R.P., Van Kesteren, R.E., Van Der Schors, R.C., and Geraerts, W.P. (1995) Co-localized neuropeptides conopressin and ALA-PRO-GLY-TRP-NH2 have antagonistic effects on the vas deferens of *Lymnaea*. *Neuroscience* **69**, 1275–1287.
12. Li, K.W., Miller, S., Klychnikov, O., Loos, M., Stahl-Zeng, J., Spijker, S., Mayford, M., and Smit, A.B. (2007) Quantitative proteomics and protein network analysis of hippocampal synapses of CaMKIIalpha mutant mice. *J. Proteome Res.* **6**, 3127–3133.

Chapter 5

Monitoring Neuropeptides In Vivo via Microdialysis and Mass Spectrometry

Heidi L. Behrens and Lingjun Li

Abstract

Neuropeptides are important signaling molecules that regulate many essential physiological processes. Microdialysis offers a way to sample neuropeptides in vivo. When combined with liquid chromatography–mass spectrometry detection, many known and unknown neuropeptides can be identified from a live organism. This chapter describes sample preparation techniques and general strategies for the mass spectral analysis of neuropeptides collected via microdialysis sampling. Methods for the in vitro microdialysis of a neuropeptide standard as well as the in vivo microdialysis sampling of neuropeptides from a live crab are described.

Key words: Neuropeptides, mass spectrometry, microdialysis, crustacean, hemolymph.

1. Introduction

Microdialysis is an in vivo sampling technique that allows the collection of molecules in real time with minimal disturbance to the organism; it produces relatively clean samples that do not require extensive preparation before analysis. Microdialysis is a well-established technique for sampling low molecular weight molecules from the brain, blood, and peripheral tissues (1). More recently, microdialysis has been applied to study larger molecules, such as neuropeptides. Neuropeptides are endogenous signaling molecules that are known to regulate many physiological processes. Most of the microdialysis studies of neuropeptides use radioimmunoassay (RIA) detection. While it has superb sensitivity, RIA cannot distinguish among the members of a neuropeptide family due to the similarity of their amino acid sequences.

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_5,

To overcome this lack of specificity, many researchers employ mass spectrometry (MS) to detect neuropeptides because it can sequence both known and unknown neuropeptides. This chapter describes a method to sample neuropeptides from the hemolymph (blood) of the crab, *Cancer borealis*, and detect them using liquid chromatography–mass spectrometry (LC–MS). Decapod crustaceans are important model organisms for studying the neuromodulatory and hormonal control of physiological processes (2, 3) and numerous studies have been published identifying the peptides present in the hemolymph and various tissues of these animals (4–9).

This chapter contains protocols for performing in vitro microdialysis with LC–MS quantification of recovery as well as in vivo microdialysis, dialysate sample preparation, and the MS analysis of neuropeptides from microdialysates. In the in vitro microdialysis experiment a microdialysis probe is used to recover a standard peptide from artificial crab saline. This kind of experiment is often done to test the viability of the probe and the microdialysis setup. While it does not completely resemble in vivo conditions, the in vitro microdialysis experiment can provide a rough idea of the recovery of the particular analyte in vivo. If a higher recovery is desired, one can increase the surface area or pore size of the probe membrane, decrease the flow rate, or add molecules to the perfusion fluid that bind the analyte (10). Once the experiment is complete, the probe can be stored for later use. The sample preparation of in vitro and in vivo dialysates is fairly straightforward and the quantification of in vitro dialysates is accomplished via integration of the chromatographic peak using LC–MS software. In vivo microdialysis of the hemolymph in the pericardial sinus of a live crab is described, with details on probe preparation, implantation, and post-experiment probe visualization. Finally, we include some general suggestions for the MS analysis of neuropeptides from microdialysates.

While this chapter focuses on a microdialysis and LC–MS method for the crab nervous system, many of the techniques described here can be readily applied to other systems. The principles of the in vitro microdialysis recovery experiment are the same regardless of the analyte or the organism. The section on quantification using LC–MS can also be used quite generally. Fewer desalting steps will be required for the LC–MS detection of neuropeptides from non-marine organisms.

2. Materials

2.1. In Vitro Microdialysis of 2 μM Arg Vasopressin in Saline

1. Crab saline: 440 mM NaCl, 11 mM KCl, 26 mM $MgCl_2$, 13 mM $CaCl_2$, 11 mM Trizma base, 5 mM maleic acid, pH 7.45. Store at 4°C.

2. Arg vasopressin (AVP, CYFQNCPRGa with Cys_1–Cys_6 forming disulfide bridge, MW 1083.45, American Peptide Company, Sunnyvale, CA).
3. 2.5 mL glass syringe and attached needle (CMA Microdialysis, North Chelmsford, MA).
4. 50% ethanol in water.
5. Water, double distilled by filtration system (Millipore, Bedford, MA).
6. Syringe pump, CMA/102.
7. Fluorinated ethylene propylene (FEP) tubing connectors (CMA), stored in ethanol at room temperature.
8. Microdialysis probe: CMA/20 Elite with 20 kDa molecular weight cut-off and 4 mm polyarylethersulfone membrane.
9. 20 gauge, 1″ long PrecisionGlide hypodermic needles (BD, Franklin Lakes, NJ).
10. Total recovery collection vials (Waters, Milford, MA).

2.2. Sample Preparation of In Vitro Microdialysates

1. 0.1% formic acid (FA) in water
2. ZipTipsC18 (Millipore)
3. Solvent A, aqueous solvent for LC gradient: 95% water, 5% (v/v) acetonitrile (ACN), 0.1% (v/v) FA. Store at room temperature.

2.3. Analysis of In Vitro Microdialysates Using LC–MS to Determine AVP Recovery

1. LC–MS software with peak integration capabilities, such as Mass Lynx, version 4.0 (Waters)

2.4. In Vivo Microdialysis of Live Crab

In addition to the supplies listed in **Section 2.1** , the following items will also be required:

1. FEP tubing, 0.12 mm inner diameter (CMA)
2. Hot glue and glue gun
3. 1–2 sheets of plexiglass, 3/8″ thick, sized to fit your saltwater tank
4. Plumber's epoxy (Poxy Plus, Inc., Sussex, WI)
5. Rotary tool, such as Dremel 7.2 V MultiPro Cordless (Dremel, Racine, WI)
6. Rotary tool drill bit, 1/32″ (0.8 mm) (Dremel)
7. Super epoxy (Poxy Plus, Inc.)
8. Green food dye
9. Dissecting tools: Side cutters or rongeurs, spatula, small scissors

2.5. Sample Preparation of In Vivo Microdialysates

In addition to the supplies listed in **Section 2.2** , the following items will also be required:

1. Arg vasopressin (AVP, CYFQNCPRGa with Cys_1–Cys_6 forming disulfide bridge, MW 1083.45, American Peptide Company, Sunnyvale, CA).
2. Vivapure C_{18} micro spin columns (Vivaproducts, Inc., Littleton, MA).

3. Methods

3.1. In Vitro Microdialysis of 2 μM AVP in Saline (*see* Note 1)

1. Degas about 15 mL of crab saline to use as perfusate (*see* **Note 2**).
2. Prepare about 6–8 mL of peptide solution by dissolving AVP in crab saline (final concentration 2 μM AVP). Vortex to mix. Transfer to a 10 mL beaker before the in vitro experiment. Keep cold (*see* **Note 3**).
3. Wash syringe by rinsing with 50% ethanol three times, with water three times, and with the perfusate three times. Fill syringe with perfusate and push all air bubbles out of the syringe needle.
4. Set filled syringe in syringe pump. Attach three-prong clamp in clamp holder to burette stand. Attach hosecock clamp to three-prong clamp and hang probe from hosecock clamp by tightening on the plastic tab of the probe where the inlet and outlet tubing come together (*see* **Note 4**).
5. Fill a 10 mL beaker with about 6 mL of crab saline. Lower the probe into the beaker so that the membrane is completely immersed but does not touch the walls of the beaker.
6. Remove the FEP tubing connectors from ethanol and use them to make connections between the syringe, inlet tubing, probe, and outlet tubing. Blot dry with a kimwipe or a similar cleaning tissue. Let air dry for 10 min (*see* **Note 5**).
7. Set the outlet in a waste vial and start the pump at 10 μL/min for 5 min to push a few internal volumes of perfusate through the system (*see* **Notes 6** and **7**).
8. Use a ×10 magnifier to check that there are no bubbles in the membrane or shaft of the probe (*see* **Note 8**).
9. Stop the pump and slowly remove the probe from the crab saline solution and lower it into a 10 mL beaker filled with about 6 mL of the AVP solution. Remove 50 μL of

the AVP solution now and save for LC–MS analysis (*see* **Notes 9** and **10**).

10. Push the perfusate out of the outlet tubing by starting the pump at 5 μL/min for 2 min (*see* **Note 11**).
11. Stop the pump and set the outlet tubing in a collection vial. Start the pump at 0.5 μL/min for the experiment and collect for 30 min. Transfer the outlet tubing to a new vial and collect fractions every 60 min for 3 h (three fractions of 30 μL each). Collect the fractions on ice and store at –20°C immediately upon collection (*see* **Notes 12–15**).
12. To reuse the probe and/or tubing, stop the pump and suspend the probe in a beaker of water. Replace the perfusate in the syringe with water (rinse with water three times). Reapply the tubing connectors and let air dry. Flush the system with several internal volumes of water (10 μL/min, 5 min).
13. To store the probe for later reuse, fill a 50 mL centrifuge tube with enough water to immerse the membrane (about 10–15 mL). Use a 20 G needle to poke two holes in the cap and feed the inlet and outlet tubes through these holes until the probe can be suspended in the water without touching the bottom or sides of the tube. Screw the cap tight. Place a square of parafilm around each tubing end to prevent dust from getting inside the tubing. Store in the refrigerator (*see* **Notes 16** and **17**).

3.2. Sample Preparation of In Vitro Microdialysates

1. Prepare the fractions for analysis by vacuum drying and redissolving in 10 μL of 0.1% FA. Centrifuge for 5 min at 10,000×*g* (*see* **Notes 18–20**).
2. Desalt with ZipTips. In the elution step, elute in 2–3 μL of elution solution, then dilute up to 30 μL with Solvent A. Elute into the LC–MS total recovery vial (*see* **Note 21**).
3. Analyze samples using LC–MS, preferably injecting full loop to increase injection reproducibility. Run all the fractions and the medium (*see* **Notes 22** and **23**).

3.3. Analysis of In Vitro Microdialysates Using LC–MS to Determine AVP Recovery

1. Determine the experimental m/z value (± 0.1) of each charge state of AVP. AVP displays $(M+H)^+$ at around m/z 1084.5 and $(M+2H)^{2+}$ at around m/z 542.7 (*see* **Notes 24** and **25**).
2. Create an extracted ion chromatogram (EIC) for the experimental m/z values of singly and doubly charged AVP. For each peak in the EIC, check the MS to determine if AVP is present. From this, determine the retention time of AVP for the specific LC–MS gradient used.

3. Smooth and integrate the AVP peak for each charge state (*see* **Notes 26** and **27**).
4. Repeat Steps 1–3 for each LC–MS file.
5. Add together the peak areas for both AVP charge states for each microdialysis fraction you collected as well as for the 2 μM AVP medium. To determine the recovery for each fraction, divide the total AVP peak area for the fraction by the total AVP peak area for the medium. This percentage is the recovery of AVP in that fraction. To determine the recovery of AVP for the microdialysis experiment, average together the recoveries from the fractions after the 30 min equilibration period (average the recoveries of the 90, 150, and 210 min fractions). This averaged value is the final recovery value (*see* **Notes 28** and **29**).

3.4. In Vivo Microdialysis of a Live Crab *(see* Note 1*)*

1. Degas about 15 mL of crab saline to use as perfusate (*see* **Notes 2** and **30**).
2. Cut desired lengths of FEP tubing to extend the inlet and outlet tubing of the probe (*see* **Notes 31** and **32**).
3. Clean syringe by rinsing with 50% ethanol three times, with water three times, and then perfusate three times. Fill syringe with perfusate and push all air bubbles out of the syringe needle.
4. Set filled syringe in syringe pump. Attach three-prong clamp in clamp holder to burette stand. Attach hosecock clamp to three-prong clamp and hang probe from hosecock clamp by tightening on the plastic tab of the probe where the inlet and outlet tubing come together (*see* **Note 4**).
5. With the hot glue gun, apply a small ball of hot glue to the probe shaft about 1 cm from the probe tip, being careful not to get any glue on the membrane. Let dry.
6. Fill a 10 mL beaker with about 6 mL of crab saline. Lower the probe into the beaker so that the membrane is completely immersed but does not touch the walls of the beaker.
7. Remove the FEP tubing connectors from ethanol and use them to make connections between the syringe, FEP tubing, inlet tubing, probe, and outlet tubing. Blot dry with a kimwipe. Leave to air dry for 10 min (*see* **Note 5**).
8. Set the outlet in a waste vial and start the pump at 10 μL/min for 10 min to push a few internal volumes of perfusate through the system (*see* **Notes 7** and **33**).
9. Use a ×10 magnifier to check that there are no bubbles in the membrane or shaft of the probe (*see* **Note 8**).

10. Set up a couple of plexiglass sheets in the tank to section off one side/corner in which to confine the crab during the microdialysis experiment (*see* **Note 34**).
11. Remove the crab from the tank and place it in a bucket filled with ice; note the time when the crab was first placed on ice. Dry the shell above the pericardial sinus and use a marker to mark where you will drill a hole later. Mix up the plumber's epoxy on a plate and roll into a long cylinder. Press this in a circle around the pericardial region, so that the drilling site is in the center (**Fig. 5.1**). Move the ice up around the crab to cover the shell, but leave the area around the plumber's epoxy dry. Leave the crab on ice about 20–25 min (*see* **Notes 35** and **36**).
12. While the crab is on ice, set up the surgery area in a cold room, preferably in the same room as the crab tank. Set out a dissection pan and fill it with a thin layer of ice. Set out the rotary tool with drill bit attached, have the super epoxy ready to mix on a plate with a plastic knife, and make sure the hot glue gun is heating. Set out wipes to blot dry hemolymph during surgery.
13. The probe should be rinsed by now (*see* Step 8 above), the syringe pump should be switched off, and both the probe and the pump should be transferred close to the surgery setup. Just before surgery, remove the probe from the beaker and unscrew the hosecock clamp slightly so the probe is sitting loosely in the clamp (*see* **Notes 37** and **38**).
14. Once the plumber's epoxy is set and the crab is anesthetized, remove the crab from the bucket and place it dorsal side up in the dissection pan for surgery. Have someone else hold the crab while you dry the area inside the plumber's epoxy with a kimwipe and mix the super epoxy on a plate.
15. Use the rotary tool to drill a hole in the shell over your mark. Quickly grab the probe and place it inside the hole, pushing down until the glue ball meets the shell (the probe tip should now be about 1 cm deep inside the crab). Hold the probe in the crab (*see* **Note 39**).
16. While holding the probe, scoop the super epoxy around the probe in the well created by the plumber's epoxy. Fill this well with super epoxy. Hold the probe at an angle so that its tip faces the heart and keep it steady for about 10 min while you wait for the super epoxy to dry (*see* **Notes 40–42**).
17. Once the super epoxy has formed a gel, add hot glue around the base of the probe shaft to glue the probe shaft

to the super epoxy. Let dry for a few minutes. Check that the probe is secure and let go off the probe (*see* **Note 43**).

18. Pick up the crab in the pan and the syringe pump (unplugged) and move it over to the tank. Carefully place the crab in the tank and expel gas from the stomach. Place the lid on the tank and place the syringe pump and a small bucket filled with ice on the lid. Set the outlet tubing in a collection vial in the ice bucket (*see* **Note 44**).
19. Plug in the syringe pump and begin the flow at 10 μL/min. Watch for fluid to come out of the outlet tubing (*see* **Note 45**).
20. Once you see fluid flowing out, change the flow to 0.5 μL/min for the microdialysis experiment. Set up the outlet tubing in the collection vial and tape the outlet tubing to the ice bucket so that it does not fall out of the collection vial. Begin collecting individual fractions. Immediately store collected fractions at –20°C (*see* **Notes 46–49**).
21. Once the experiment is completed, stop the syringe pump and remove the FEP tubing connector between the syringe needle and inlet tubing. Fill the syringe with 0.5 mL of green dye. Re-attach the inlet tubing to the syringe needle with a new FEP tubing connector, blot dry with a kimwipe, and let air dry for 10 min.
22. Start the pump at 5 μL/min. You should see green dye flow out the outlet tubing after a few minutes. Place the outlet tubing in a vial, wait a few minutes more, and then stop the pump (*see* **Note 50**).
23. Disconnect the inlet tubing from the syringe pump and remove the syringe pump and ice bucket from the lid of the tank. Remove the crab from the tank and place in a large bucket of ice for 20 min to anesthetize. Be careful not to touch the probe (*see* **Note 51**).
24. While the crab is on ice, set out the dissection pan, spatula, side cutters, and small scissors.
25. Once the crab is anesthetized, place the crab in the dissecting pan and begin by using the side cutters to remove the claws and legs at the base where they meet the body.
26. Use the side cutters to crunch around the outer rim of the crab shell. Then use the spatula to reach in between the top and bottom shells and separate the hypodermis from the upper shell.
27. Use the side cutters to remove the upper shell back to the pericardial region, saving the area behind the pericardial ridges. Use the scissors to cut the connective tissue between

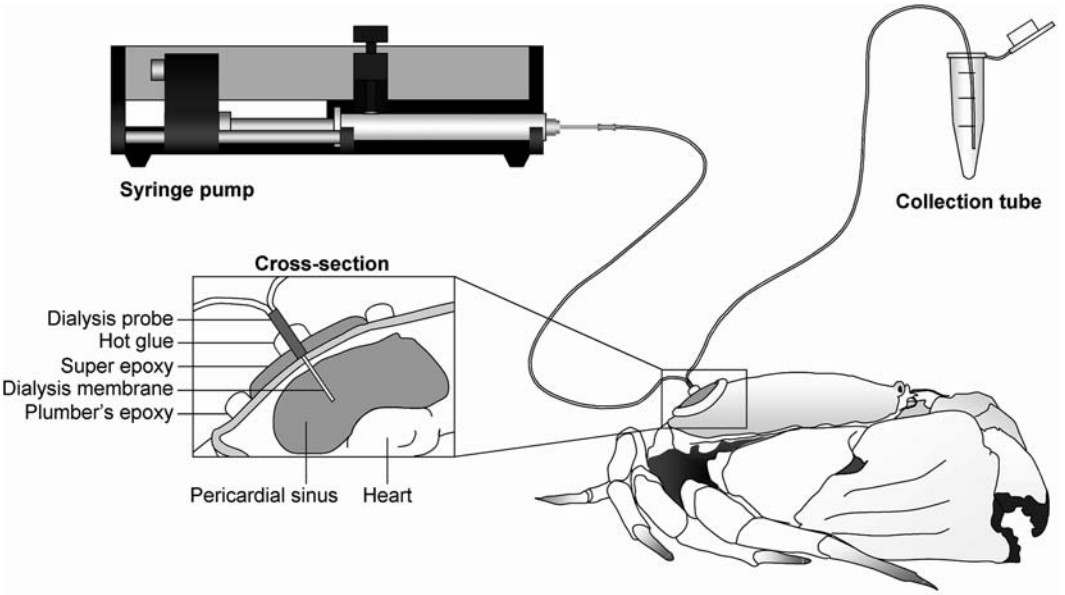


Fig. 5.1. Schematic representation of the microdialysis setup, the microdialysis probe placement in the pericardial sinus of a crab, and of the glues used to stabilize the probe.

the upper shell and the pericardial region so that you can pull the shell off.

28. Pull the pericardial shell off, leaving the part next to the tail attached if possible. Upon removing the shell, you should be able to see where the probe tip was located. Look for a small amount of green dye under the probe tip, probably on the surface of the heart.
29. Finish dissecting the crab and dispose it off according to the rules and procedures.

3.5. Sample Preparation of In Vivo Microdialysates

1. Defrost fractions and spike them with AVP to 1 nM. Vacuum dry the spiked fractions and redissolve in 10–200 μL of 0.1% FA, depending on the initial fraction volume and the desired desalting method. Centrifuge for 5 min at 10,000×*g*; a small white pellet may be visible (*see* **Notes 19, 20, 52,** and **53**).
2. Desalt the supernatant of all fractions using ZipTips for smaller volumes (less than 50 μL) or micro spin columns for larger volumes (50–200 μL). For LC–MS analysis, elute in a small volume (2–10 μL) then dilute up to the desired volume with Solvent A. For MALDI MS analysis, elute in 2–3 μL elution solution. If necessary, store samples at –20°C (*see* **Notes 20, 21,** and **54**).
3. Analyze desalted microdialysis fractions by LC–MS or MALDI MS (*see* **Notes 55** and **56**).

4. Notes

1. All solutions could be chilled and the AVP solution kept on ice or, alternatively, perform all steps in a cold room.
2. Should be done on the day of the experiment.
3. Note that micromolar concentrations of AVP will degrade within a few months.
4. Many microdialysis suppliers offer probe clips that may simplify the suspension of the probe in the analyte solution.
5. Some probes (including the CMA/20) come with tubing connected to them and you can just use this or you can extend the attached tubing with FEP tubing; extending the tubing length is usually necessary for in vivo experiments. Other probes do not have tubing connected and you will have to attach your own FEP tubing to the probe.
6. The total internal volume of this setup is about 12 μL, so about four internal volumes were rinsed through this system. For other setups, one will need to measure the internal volume of the probe-tubing system. This can be calculated from the probe and tubing internal volumes (usually given in the accompanying product manuals). Typical probe internal volumes are 3–5 μL; 0.12 mm ID FEP tubing is 1.2 μL/100 mm. A good estimate for rinsing is 5 min at 10 μL/min.
7. Once the pump has started, check the system for leaks by looking for small beads of fluid, usually at the tubing connectors. If you find a leak, stop the pump and replace the tubing connector with one that has soaked in ethanol for 5–10 min. Blot dry with kimwipe and let air dry for 10–15 min. Gently test that the connection is secure and start the pump again.
8. If there are any bubbles, these need to be removed before continuing because they will decrease the recovery of your probe and alter the results. The easiest way to remove bubbles is to try a higher flow rate (20 μL/min or more) for a few minutes. If this does not work, remove the probe and clamp and carefully tap the clamp against a metal surface. You can also try forcefully swinging the probe and clamp in the air. If none of these options work, use a new probe.
9. The analyte solution can be stirred with a magnetic stirrer on a stirring plate if desired; this may increase recovery (11).

10. The volume of the analyte solution should be no less than 5–10 mL; this will depend on the flow rate, expected recovery of the specific analyte, and the total time of the experiment.
11. To determine the void time, one needs to know the internal volume from the probe tip out through the outlet tubing to the vial. Typically, this is 5–10 μL which can be flushed out by running at 5 μL/min for 1–2 min.
12. The duration of the fractions will vary depending on the volume requirement and sensitivity of the analysis method as well as the number of analysis replicates desired.
13. For LC–MS, fractions could be injected without further sample preparation. In such a case the samples should be collected in the vials used for sample injection. This will reduce sample handling and decrease sample losses.
14. You may want to start by collecting fractions of larger volumes than needed to aid sample preparation and handling and allow for extra MS replicates.
15. The first 30 min are required to equilibrate the analyte solution and the perfusate in the membrane; this fraction will have a lower recovery than later fractions. Collect at least 2–3 fractions after the equilibration period.
16. Used probes can be stored for several months. When storing a probe, it is important to keep the membrane wet and free from contact with the storage container. Tubing should be free of salts to avoid clogging.
17. Alternatively, the same procedure can be done with all-purpose contact lens solution instead of water. The probe can be stored at room temperature, but the contact solution will need to be replaced every 2 weeks to minimize bacterial growth.
18. If water was used as the perfusate, it is not necessary to vacuum dry or desalt; the microdialyzed fractions can be analyzed by LC–MS directly.
19. Fractions with larger initial volumes will need more liquid to dissolve all the salt that precipitates out upon vacuum drying. Use just enough FA to dissolve the precipitates.
20. For MALDI MS analysis, 0.1% trifluoroacetic acid (TFA) in water can be substituted with the 0.1% FA.
21. Final ACN concentration in the samples for LC–MS analysis is 5–25%. Whilst using ZipTips, elute samples in small volumes of 50% ACN, then dilute the samples by adding aqueous LC–MS solvent. Alternatively, the samples can be

vacuum-dried after ZipTip purification and redissolved in the required LC–MS solvent.

22. The volume of the fractions may need to be adjusted based on the requirements of the specific LC–MS system. For a 5 μL sample loop, full-loop mode with an overfill factor of 2 would require 10 μL, whilst only 5 μL would be injected. This means that 2–3 LC–MS analyses can be performed on each of the 30 μL fractions.
23. Run multiple replicates of fractions and of 2 μM AVP solution. Run each set of replicates on the same day, starting with the lowest analyte concentration. Wash the injectors with water to avoid carryover of analyte between runs.
24. The lower charge states of the analyte would require less signal integration.
25. Failure to determine the correct experimental m/z values can affect the extraction ion chromatogram and thus the integration, which could change the final recovery value.
26. Smoothing is not usually necessary, but it can help with noisy chromatograms.
27. Make sure that the integrated area includes the AVP charge state desired but not much of anything else. For instance, if there is an intense contaminant eluting shortly after and overlapping with AVP, one can modify the tail of the integrated peak to stop before the MS signal from the contaminant overwhelms that of AVP. Regardless of the integration method chosen, it is important to be consistent for all of the LC–MS files.
28. Expected recovery for AVP at this flow rate and with this probe is approximately 15–20%.
29. If you inject 5 μL of 2 μM AVP, you would be loading roughly 1–10 ng of AVP onto the instrument, depending on the microdialysis recovery. Make sure that the instrument can accurately quantify AVP over the specific concentration range chosen.
30. If the desired neuropeptide is particularly hydrophobic, it may adsorb to the microdialysis probe and the walls of the tubing, making its detection difficult. This sticking can be minimized by adding a small amount (0.5%, w/v) of bovine serum albumin (BSA) to the perfusate (12, 13). The BSA can be precipitated out of the dialysed samples with the addition of methanol or acetonitrile followed by centrifugation.
31. Use a commercial tube cutter to achieve straight edges and prevent leaks.

32. Minimize total tubing length because peptides can stick to the walls of tubing. For our tank setup, we use about 45 cm of FEP to extend the probe inlet tubing and about 65 cm of FEP to extend the probe outlet tubing.
33. The total internal volume of this setup with the FEP tubing extensions is about 25 μL, so about four internal volumes were rinsed through this system. For other setups, one will need to know the internal volume of the probe-tubing system. This can be calculated from the probe and tubing internal volumes (usually given in the accompanying product manuals). Typical probe internal volumes are 3–5 μL; 0.12 mm ID FEP tubing is 1.2 μL/100 mm. A good estimate for rinsing is 10 min at 10 μL/min.
34. You should allow the crab to move, but excessive movements would require longer tube lengths. Typically we confine crabs to an area of about 11 in^2.
35. Leaving the crab on ice for too long (45 min or more) may kill the animal.
36. The plumber's epoxy needs to be placed on the crab before transferring the animal to ice because it takes about 20–25 min to harden. Building the plumber's epoxy circle a little higher on the side next to the tail will help to prevent the super epoxy from flowing down the tail during surgery.
37. Do not remove the probe from the beaker far in advance of surgery to prevent drying out of the membrane (it will become unusable).
38. The probe is still attached to the syringe through the inlet tubing, so the pump should be close enough that it can be connected to the crab without straining the tubing.
39. Keep drilling the shell until hemolymph flows out, indicating that you penetrated the pericardial cavity. This will happen shortly after the drill penetrates the shell (felt through the reduced mechanical resistance to drilling).
40. The hemolymph will keep flowing out of the crab during this step, but the super epoxy will gradually begin to harden, slowing the flow of hemolymph.
41. Try not to move the probe drastically, as you could run into some tissue on the sides of the pericardial cavity and clog the probe membrane.
42. After 8–10 min, you can tap the super epoxy to see if it is dry. The super epoxy will not be completely hard, but it will resemble a gel, which would provide sufficient adhesion. The probe may still be slightly moveable; the hot glue (added next) will set quickly and fix the probe in place.

43. The hot glue will set very quickly, especially if the surgery is performed in a cold room. If the probe is not fixed after the first applicaiton, more hot glue could be added around the probe.
44. When placing the crab in the tank, be careful not to touch the walls of the tank with the probe. Try to keep the end of the outlet tubing out of the water to keep it clean and prevent saltwater from getting inside. This tube can be taped to the top of the tank to keep it out of the way whilst the crab is transferred into the tank. Try to keep the tubing away from the crab and tangle-free.
45. If no fluid is flowing out of the outlet tubing after a few minutes, the flow rate could be increased and check for leaks. If there are no leaks, it is likely that the system is blocked. Try cutting the outlet tubing by a few inches in case the clog is near the outlet. If this does not work, the clog is probably near the probe and a new probe would have to be used and the experiment re-started.
46. Using a tube holder prevents the collection tube from moving as the ice melts. A simple holder can be made by cutting a hole in a 1-inch-wide strip of cardboard and taping this across the ice bucket.
47. The first 30 min is an equilibration period and is not representative of basal conditions. The first 10–12 h may contain stress-induced neuropeptides.
48. Once basal samples have been collected, the in vivo analyte concentration can be estimated using the zero net flux microdialysis method (14).
49. Change the ice and collection tubes at least 2–3 times a day to minimize temperature-induced neuropeptide degradation.
50. Alternatively, you could use a disposable plastic syringe and needle for the green dye, as it can be quite messy; the injection rate is not critical for this step. The FEP tubing connectors fit well over a 21 G needle.
51. The tubing may be cut near the probe to simplify disassembly.
52. The spiked standard serves as a way to monitor sample loss and neuropeptide concentration as well as being an internal standard for MALDI MS. More standards can be added to cover the mass range of analysis.
53. Centrifugation helps to prevent large molecules or precipitated salts from clogging the membrane of the desalting devices.

54. Neuropeptides exist at very low concentrations in the hemolymph, so increase the concentration factor until you can see neuropeptides with your instrument. Concentrating dialysed samples 6- to 10-fold should allow observation of neuropeptides at basal levels in sensitive MS instruments. The samples may need to be concentrated further (100- to 200-fold), depending on the concentration of the desired neuropeptide and the physiological state of the organism.
55. For MALDI, alpha-cyano-4-hydroxycinnamic acid (CHCA) works better for microdialysates, although some neuropeptides are detected easier with different matrices such as 2,5-dihydroxybenzoic acid (DHB).
56. When analyzing neuropeptides from in vivo microdialysates, it is important to adequately desalt the dialysates and employ a sensitive and selective detection scheme. Biological fluids contain high concentrations of various salts, which can be problematic for MS detection. Desalting of dialysates can be done offline using solid support (15), ZipTips, or reversed-phase capillary columns (16). Online desalting methods such as reversed-phase trap columns (17) and microdialysis-based devices (18) decrease sample handling and increase automation.

In addition to desalting, it is essential to maximize the sensitivity of the MS detection. Neuropeptides exist at micromolar to picomolar levels in vivo so the MS instrument must be able to detect attomoles of neuropeptides in microliter samples. Triple quadrupole (16, 19), quadrupole ion trap (20, 21), time-of-flight (21–23), and quadrupole time-of-flight mass spectrometers have all been successfully employed to detect neuropeptides from microdialysates. For ion trap and quadrupole-based instruments, the sensitivity can be improved by detecting a narrow m/z range (24) or by performing single reaction monitoring experiments (16).

Front-end separation of dialysates by reversed-phase LC or capillary electrophoresis (16, 25–27) can further enhance MS sensitivity by simplifying the sample that enters the mass spectrometer. In LC, a small inner diameter (ID) column ($\leq$ 75 μm) is essential to neuropeptide detection, with very small ID columns providing the best sensitivity (24). Several microliters of dialysate should be injected and subjected to a slow elution gradient, ideally increasing by less than 1% organic solvent per minute.

Finally, the enormous variety in neuropeptide sequences requires the selectivity of tandem MS detection to discern the correct amino acid sequence. Using MS/MS, one can sequence numerous neuropeptides from microdialysates. However, if the instrument is capable, MS^3 can give better signal-to-noise ratios, resulting in more confident peptide identifications (20). Chemical derivatization is another strategy to improve peptide identification

from tandem MS data and can be particularly helpful in determining the identity of the C- and N-terminal amino acids as well as the presence of certain internal amino acids (28, 29).

Acknowledgments

The authors wish to thank Professor Craig Berridge (Department of Psychology, University of Wisconsin-Madison) for helpful discussions about microdialysis. This work was supported in part by the School of Pharmacy and the Wisconsin Alumni Research Foundation at the University of Wisconsin-Madison, a National Science Foundation CAREER Award (CHE-0449991), and the National Institutes of Health through Grant 1R01DK071801. L.L. acknowledges an Alfred P. Sloan Research Fellowship. H.L.B. acknowledges the National Institutes of Health Biotechnology Training Grant 5T32 GM08349.

References

1. Robinson, T.E. and Justice, J.B. (eds.) (1991) Microdialysis in the Neurosciences. Elsevier Science Publishing Company, London.
2. Skiebe, P. (2001) Neuropeptides are ubiquitous chemical mediators: using the stomatogastric nervous system as a model system. *J. Exp. Biol.* **204**, 2035–2048.
3. Nusbaum, M.P. and Beenhakker, M.P. (2002) A small-systems approach to motor pattern generation. *Nature* **417**, 343–350.
4. Turrigiano, G.G. and Selverston, A.I. (1990) A cholecystokinin-like hormone activates a feeding-related neural circuit in lobster. *Nature* **344**, 866–868.
5. Li, L., Kelley, W.P., Billimoria, C.P., Christie, A.E., Pulver, S.R., Sweedler, J.V., and Marder, E. (2003) Mass spectrometric investigation of the neuropeptide complement and release in the pericardial organs of the crab, *Cancer borealis*. *J. Neurochem.* **87**, 642–656.
6. Messinger, D.I., Kutz, K.K., Le, T., Verley, D.R., Hsu, Y.W., Ngo, C.T., Cain, S.D., Birmingham, J.T., Li, L., and Christie, A.E. (2005) Identification and characterization of a tachykinin-containing neuroendocrine organ in the commissural ganglion of the crab *Cancer productus*. *J. Exp. Biol.* **208**, 3303–3319.
7. DeKeyser, S.S., Kutz-Naber, K.K., Schmidt, J.J., Barrett-Wilt, G.A., and Li, L. (2007) Imaging mass spectrometry of neuropeptides in decapod crustacean neuronal tissues. *J. Proteome Res.* **6**, 1782–1791.
8. Stemmler, E.A., Gardner, N.P., Guiney, M.E., Bruns, E.A., and Dickinson, P.S. (2006) The detection of red pigment-concentrating hormone (RPCH) in crustacean eyestalk tissues using matrix-assisted laser desorption/ionization-Fourier transform mass spectrometry: [M + Na]+ ion formation in dried droplet tissue preparations. *J. Mass Spectrom.* **41**, 295–311.
9. Saideman, S.R., Ma, M., Kutz-Naber, K.K., Cook, A., Torfs, P., Schoofs, L., Li, L., and Nusbaum, M.P. (2007) Modulation of rhythmic motor activity by pyrokinin peptides. *J. Neurophysiol.* **97**, 579–595.
10. Duo, J., Fletcher, H., and Stenken, J.A. (2006) Natural and synthetic affinity agents as microdialysis sampling mass transport enhancers: current progress and future perspectives. *Biosens. Bioelectron.* **22**, 449–457.
11. Rojas, C., Nagaraja, N.V., and Derendorf, H. (2000) In vitro recovery of triamcinolone acetonide in microdialysis. *Pharmazie* **55**, 659–662.
12. Lanckmans, K., Sarre, S., Smolders, I., and Michotte, Y. (2008) Quantitative liquid chromatography/mass spectrometry for the analysis of microdialysates. *Talanta* **74**, 458–469.

13. Trickler, W.J. and Miller, D.W. (2003) Use of osmotic agents in microdialysis studies to improve the recovery of macromolecules. *J. Pharm. Sci.* **92**, 1419–1427.
14. Guiard, B.P., David, D.J., Deltheil, T., Chenu, F., Le Maitre, E., Renoir, T., Leroux-Nicollet, I., Sokoloff, P., Lanfumey, L., Hamon, M., Andrews, A.M., Hen, R., and Gardier, A.M. (2008) Brain-derived neurotrophic factor-deficient mice exhibit a hippocampal hyperserotonergic phenotype. *Int. J. Neuropsychopharmacol.* **11**, 79–92.
15. Pettersson, A., Amirkhani, A., Arvidsson, B., Markides, K., and Bergquist, J. (2004) A feasibility study of solid supported enhanced microdialysis. *Anal. Chem.* **76**, 1678–1682.
16. Andren, P.E. and Caprioli, R.M. (1999) Determination of extracellular release of neurotensin in discrete rat brain regions utilizing in vivo microdialysis/electrospray mass spectrometry. *Brain Res.* **845**, 123–129.
17. Bengtsson, J., Jansson, B., and Hammarlund-Udenaes, M. (2005) On-line desalting and determination of morphine, morphine-3-glucuronide and morphine-6-glucuronide in microdialysis and plasma samples using column switching and liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **19**, 2116–2122.
18. Jakubowski, J.A., Hatcher, N.G., and Sweedler, J.V. (2005) Online microdialysis-dynamic nanoelectrospray ionization-mass spectrometry for monitoring neuropeptide secretion. *J. Mass Spectrom.* **40**, 924–931.
19. Lanckmans, K., Stragier, B., Sarre, S., Smolders, I., and Michotte, Y. (2007) Nano-LC–MS/MS for the monitoring of angiotensin IV in rat brain microdialysates: limitations and possibilities. *J. Sep. Sci.* **30**, 2217–2224.
20. Baseski, H.M., Watson, C.J., Cellar, N.A., Shackman, J.G., and Kennedy, R.T. (2005) Capillary liquid chromatography with MS3 for the determination of enkephalins in microdialysis samples from the striatum of anesthetized and freely-moving rats. *J. Mass Spectrom.* **40**, 146–153.
21. Reed, B., Zhang, Y., Chait, B.T., and Kreek, M.J. (2003) Dynorphin A(1–17) biotransformation in striatum of freely moving rats using microdialysis and matrix-assisted laser desorption/ionization mass spectrometry. *J. Neurochem.* **86**, 815–823.
22. Zhang, H., Stoeckli, M., Andren, P.E., and Caprioli, R.M. (1999) Combining solid-phase preconcentration, capillary electrophoresis and off-line matrix-assisted laser desorption/ionization mass spectrometry: intracerebral metabolic processing of peptide E in vivo. *J. Mass Spectrom.* **34**, 377–383.
23. Wilson, S.R., Boix, F., Holm, A., Molander, P., Lundanes, E., and Greibrokk, T. (2005) Determination of bradykinin and arg-bradykinin in rat muscle tissue by microdialysis and capillary column-switching liquid chromatography with mass spectrometric detection. *J. Sep. Sci.* **28**, 1751–1758.
24. Haskins, W.E., Wang, Z., Watson, C.J., Rostand, R.R., Witowski, S.R., Powell, D.H., and Kennedy, R.T. (2001) Capillary LC–MS2 at the attomole level for monitoring and discovering endogenous peptides in microdialysis samples collected in vivo. *Anal. Chem.* **73**, 5005–5014.
25. Davies, M.I., Cooper, J.D., Desmond, S.S., Lunte, C.E., and Lunte, S.M. (2000) Analytical considerations for microdialysis sampling. *Adv. Drug Deliv. Rev.* **45**, 169–188.
26. Shackman, H.M., Shou, M., Cellar, N.A., Watson, C.J., and Kennedy, R.T. (2007) Microdialysis coupled on-line to capillary liquid chromatography with tandem mass spectrometry for monitoring acetylcholine in vivo. *J. Neurosci. Methods* **159**, 86–92.
27. Myasein, K.T., Pulido, J.S., Hatfield, R.M., McCannel, C.A., Dundervill, R.F., 3rd, and Shippy, S.A. (2007) Sub-microlitre dialysis system to enable trace level peptide detection from volume-limited biological samples using MALDI-TOF-MS. *Analyst* **132**, 1046–1052.
28. Cruz-Bermudez, N.D., Fu, Q., Kutz-Naber, K.K., Christie, A.E., Li, L., and Marder, E. (2006) Mass spectrometric characterization and physiological actions of GAHKNYLR-Famide, a novel FMRFamide-like peptide from crabs of the genus *Cancer*. *J. Neurochem.* **97**, 784–799.
29. Ma, M., Kutz-Naber, K.K., and Li, L. (2007) Methyl esterification assisted MALDI FTMS characterization of the orcokinin neuropeptide family. *Anal. Chem.* **79**, 673–681.

Chapter 6

Protocols for Peptidomic Analysis of Spider Venoms

Liang Songping

Abstract

Spider venom contains a complex mixture of components with a large range of molecular masses (0.1–60 kDa) exhibiting a diverse array of actions. Most of these components are proteinaceous molecules – biologically active proteins and peptides. Proteomics profiling of spider venoms (the components with MW >10 kDa) could be achieved through conventional 2-DE-based proteomics methods combined with MS or MS/MS detection. Peptidomic profiling (of the components with MW below ~10 kDa) is usually achieved through off-line separation by a combination of ion-exchange and reverse-phase chromatography, and it relies more heavily on de novo sequencing by Edman degradation or MS/MS for peptide identification.

Key words: Spider venom, peptidomics, multidimensional separations, mass spectrometer.

1. Introduction

Spider venom is a complex mixture of components with a large range of molecular masses (0.1–60 kDa), exhibiting a diverse array of functional activities. Spider venoms are chemically diverse and include proteins, peptides and small organic molecules such as acylpolyamines. It has been estimated that spider venoms may contain of the order of ~500 different proteinaceous components of varying weight, pI, hydrophobicity and of highly variable abundance (1–3). Extracting proteins from spider venoms presents essentially the same challenges as does the protein and peptide extraction from cells, tissues and body fluids (the latter typically available in abundance, compared to venom samples). In addition to the variability of their physical and chemical properties, the abundance of individual components of spider toxins varies

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_6,

significantly. Estimates indicate up to 6 orders of magnitude differences in protein and peptide expression level within individual venoms. In classical proteomics most of the proteins have molecular weights above 10 kDa, whilst in peptidomics these are typically below 10 kDa. Spider venoms contain a complex mixture of peptides *and* proteins and therefore a special strategy is required for their efficient isolation.

The rapid progress of proteomics and mass spectrometry technologies makes it possible to access the full peptide and protein complement of spider venoms. Proteomic and Peptidomic analyses have been successfully applied to the studies of venoms from several spider species including *Ornithoctonus huwena*, *Chilobrachys jingzhao*, *Atrax robustus* and *Hadronyche versuta* (1, 3–5). **Figure 6.1** shows a typical strategy for the combined proteomics and Peptidomics analyses of spider venom by using the

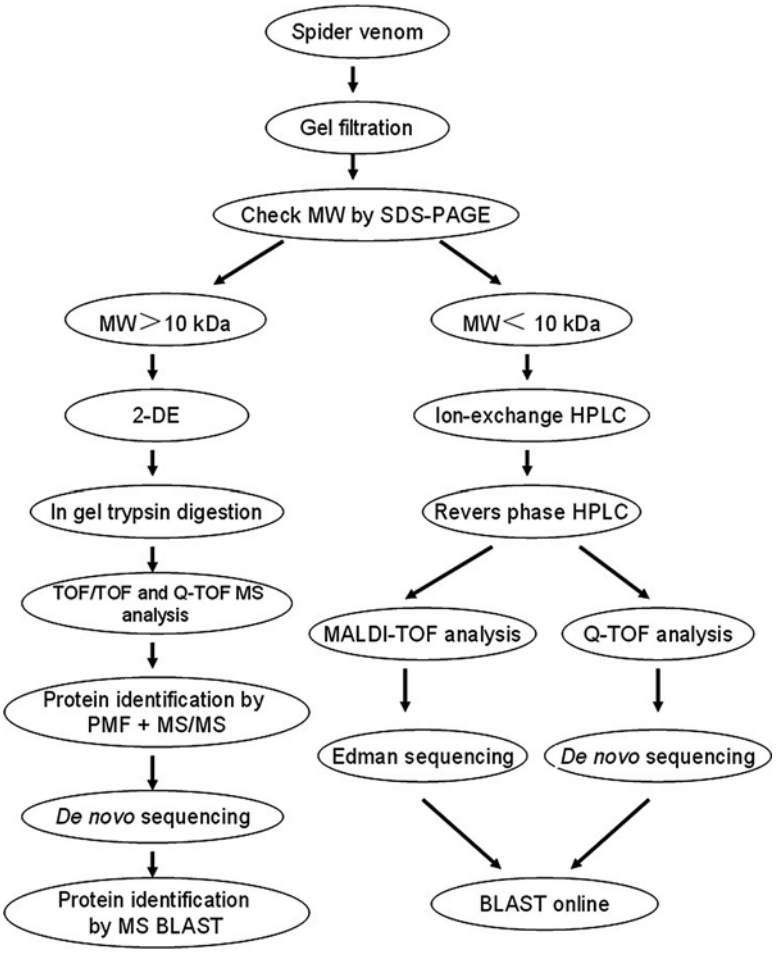


Fig. 6.1. Schematic overview of the strategy for the proteomic and peptidomic analyses of spider venoms.

combination 2-DE and mass spectrometry or a multidimensional liquid chromatography (MDLC) in combination with MALDI-TOF or Q-TOF analyses and peptide sequencing. The protein components of the spider venoms having molecular weight over ~10,000 Da could be analysed using conventional 2-DE-based proteomic approach. Peptidomic profiling is however more difficult because of the huge diversity of venom peptides and the lack of genomic data to support peptide mass matching analysis and peptide identification by mass spectrometry. Therefore, de novo sequencing becomes the main approach for peptide identification from venoms. After an off-line separation using a combination of ion-exchange and reverse-phase HPLC, de novo peptide sequencing can be attempted using either automatic Edman degradation or tandem mass spectrometry.

2. Materials

1. Sephadex G-75, IPG Buffer pH 3–10, DryStrips (180×30×0.5 mm), Cover fluid, Agarose and colloidal Coomassie blue (GE Healthcare, formerly Amersham Biosciences).
2. Deionized water was prepared with a tandem Milli-Q system and used for the preparation of all buffers.
3. Rehydration solution: 8 M Urea, 2 M Thiourea, 4% CHAPS, 20 mM Tris-base, 0.5% (v/v) IPG buffer, 18 mM DTT, bromophenol blue (trace amount to facilitate visualization of the samples), pH3.
4. Reduction solution: 50 mM Tris-HCl, 6 M urea, 30% glycerol, 2% SDS, 125 mM DTT, pH 6.8.
5. Alkylation solution: 50 mM Tris-HCl, 6 M urea, 30% glycerol, 2% SDS, 125 mM iodoacetamide, pH 6.8.
6. LC solvents. Buffer A: 0.1% formic acid, 4.9% ACN, 95% H_2O (v/v/v); Buffer B: 0.1% formic acid, 4.9% H_2O, 95% ACN (v/v/v); Buffer C: 0.1% v/v TFA in water; Buffer D: 0.1% v/v TFA in acetonitrile.
7. Gel staining with Coomassie Brilliant Blue G250: Dissolve 100 mg of Brilliant Blue G250 in 50 mL of 95% ethanol. Mix with 100 mL of 85% phosphoric acid and made up to 1 L with distilled water.
8. Matrix solution: CHCA, saturated in 97% Acetone/0.1% TFA solution Recrystallization solution: 100 mg CHCA dissolved in 10 mL of the solution ethanol/acetone/0.1% TFA (6:3:1).

9. IPGphor IEF system (Amersham Pharmacia Biotech).
10. Protean II Electrophoresis system (Bio-Rad).
11. ProXPESS 2D Proteomic Imaging System (Perkin Elmer).
12. *PDQuest* spot detection software software Version 6.1 (Bio-Rad).
13. Q-TOF mass spectrometer with a nanoelectrospray ionization source (Micromass).
14. MassLynx software for MS/MS data analysis (Micromass).
15. MALDI-TOF-TOF mass spectrometer (UltraFlex I), AnchorChip^{TM} (Bruker Daltonics).
16. BioTools v2.2 software for the analysis of LIFT-MS/MS spectra (Bruker).
17. Accell Plus Sep-Pak CM cation exchange cartridges (10 mm ×100 mm, Waters).
18. Vydac C18 reversed-phase HPLC column (218TP54, 300 Å, 4.6 mm × 250 mm, Grace Davison Discovery Sciences).
19. HPLC capillary column CapLC-MS/MS (75 μm×150 mm, Waters) for protein identification with CapLC-MS/MS (Waters).
20. Hiprep^{TM} 16/10 CM FF pre-packed column (Pharmacia).
21. Precise 491A sequencer (Applied Biosystems).
22. The venoms were obtained by stimulating the cheliceral claw of spiders using electro-pulse stimulator. The expressed venom was collected from the fang tips into a glass vial and freeze-dried.
23. Protein Assay Kit for protein concentration determination (Bio-Rad).

3. Methods

3.1. Gel Filtration and SDS-PAGE

1. Pool a few venom samples from several spiders of the same species and of the same sex (*see* **Note 1**).
2. Load pooled venom samples (10 mg) onto a 10 × 450 mm Sephadex G-75 column pre-equilibrated with 20 mM NH_4HCO_3 at pH 6.8.
3. Elute the venom using the equilibration buffer with a flow rate of 1.0 mL/min at room temperature (25°C). Monitor the elution at 215 and 280 nm. Collect 500 μL fractions (*see* **Note 2**).

4. Check molecular weight of the collected peptides by SDS-PAGE, using 15% separation gel and 4.8% stacking gel.
5. Pool fractions containing venomous proteins with MW above 10 kDa for 2-DE analysis. Pool the remaining fractions (MW below 10 kDa) for HPLC separation.
6. Determine protein concentration using Bio-Rad Protein Assay Kit or a similar method.

3.2. Separation of Venom Proteins by 2D Electrophoresis

Run 2D electrophoresis using IPGphor IEF system or a similar system.

1. Combine 300 μg of the pooled venom protein samples after the gel filtration separation with 50 μL of the rehydration solution; apply to IPG dry strips.
2. Rehydrate for 14 h, run IEF using step-n-hold protocol: 500 V for 1 h; 1000 V for 1 h; and 8000 V for 6 h at 50 μA/strip.
3. After focusing, soak the strips for 20 min in the reduction solution followed by 20 min incubation in alkylation solution.
4. Carry out the second-dimensional run on incontinuity SDS-polyacrylamide vertical slab gels 1 mm thick, with 4.8% stacking gels and 12.5% separation, in a Bio-Rad Protein II electrophoresis apparatus. Run the gel at a 12.5 mA/gel constant current, use water cooling to maintain the temperature at 10°C.
5. Stain the gel with Coomassie Brilliant Blue. Scan the gel with ProXPESS 2D Imaging System. For spot detection use *PDQuest* software.

3.3. Protein In-Gel Digestion

1. Manually excise the Coomassie blue-stained protein spots from the 2-DE gel using a puncher and place them into 500-μL microcentrifuge tubes. Store excised samples at –20°C prior to the digestion.
2. To perform in-gel digestion, first destain each spot with 50 μL of 50% ACN in 25 mM NH_4HCO_3. Incubate at 37°C for 30 min, change the destaining buffer once and repeat the incubation.
3. Reduce proteins with 10 mM DTT solution in 25 mM NH_4HCO_3 at 56°C for 1 h and then alkylate proteins with 55 mM iodoacetamide solution in 25 mM NH_4HCO_3) in the dark at room temperature for 45 min in situ.
4. Wash the gel slices or spots with 25 mM NH_4HCO_3 in water/acetonitrile (1:1, v/v) solution and dry completely in a SpeedVac. Then digest the protein in-gel using 25 μL of trypsin solution (10 ng/μL in 25 mM NH_4HCO_3) by

incubation overnight at 37°C. Extract the peptides in 50 μL of 50% acetonitrile containing 2.5% TFA. Concentrate the sample to about 4 μL for MALDI-TOF/TOF analysis.

3.4. MALDI-TOF/TOF and Q-TOF Analysis of Tryptic Peptides (see Note 3)

Protein digests obtained in **Section 3.3** above are analysed with MALDI-TOF/TOF MS, followed by protein identification with peptide mass fingerprinting (PMF) and LIFT-MS/MS (*see* **Note 4**) data searching.

1. Touch the surface of the AnchorChip™ MALDI plate with a pipette tip containing 1 μL of cyano-4-hydroxycinnamic acid (CHCA) matrix; aspirate the excess of the matrix with the same tip. The CHCA thin layer is formed within seconds.
2. Apply 2 μL of the extracted peptides directly onto the AnchorChip™ plate preloaded with CHCA matrix and left to dry for 3 min. Add 2 μL 0.1%TFA to the sample to wash out contaminants, and 4 s later remove the remaining solution with a pipette. Subsequently, add 1 μL of the recrystallization solution to the sample; ensure that sample is concentrated in the centre of the anchors.
3. For the calibration use a mixture of peptide standards, such as for example Bruker Daltonics Pepmix containing Angiotensin II, $[M+H]^+ = 1046.5420$; Angiotensin I, $[M+H]^+ = 1296.6853$; Substance P, $[M+H]^+ = 1347.7361$; Bombesin, $[M+H]^+ = 1619.8230$; ACTH clip 1–17, $[M+H]^+ = 2093.0868$; ACTH clip 18–39, $[M+H]^+ = 2465.1990$; Somatostatin 28, $[M+H]^+ = 3147.4714$.
4. Set up the Ultraflex™ TOF/TOF mass spectrometer using the *FlexControl* (TM) software; choose the reflectron mode and the accelerating voltage of 25 kV.
5. For the MS/MS analysis, choose a maximum of four precursor ions per sample. In the TOF1 stage, accelerate all ions to 8 kV to promote metastable fragmentation conditions. Select the jointly migrating parent and fragment ions in a timed ion gate; lift them to high potential energy in the LIFT cell (19 kV). Their masses could be analysed simultaneously and with high accuracy in the reflectron mode.
6. For the Q-TOF MS analysis of the peptide mixtures from the in-gel digestions use nanoelectrospray ionization source coupled with the HPLC capillary column (CapLC-MS/MS).
7. Reconstitute peptides in an aqueous solution of 5% ACN before injection.
8. For the on-line LC separation use a gradient elution (Buffer A/Buffer B) as follows: (i) (95/5–50/50%) for 65 min,

followed by (35/65–5/95%) for 10 min, followed by (5%/95%) elution for 10 min. Set flow rate to 3 μL/min. Use nanoelectrospray to inject the eluted peptides into the coupled Q-TOF MS.

9. Use the data-dependent MS/MS mode to analyse peptides over the *m/z* range of 400–2000.
10. Acquired and process MS/MS data automatically by using the *MassLynx* software.

3.5. Processing of the Mass Spectrometric Data and Database Searching (see Note 5)

1. Use *BioTools* software to analyse PMF and LIFT-MS/MS spectra with Mascot (www.matrixscience.com). Set mass tolerance in PMF to 100 ppm and MS/MS tolerance to 1.0 Da; one missed cleavage site; choose Cysteine modification by carbamidomethylation and oxidized methionine. For the purposes of protein identification, no other post-translational modifications should be taken into account and no restrictions should be imposed on species of origin of the analysed proteins. The probability score calculated by the software should be used as the only criterion of correct identification.
2. If no identification was achieved with Mascot, or if the proposed hits were not statistically significant (having scores below the threshold score suggested by Mascot), MS/MS spectra obtained from Q-TOF should be sequenced de novo and analysed with the assistance of *MassLynx* software.
3. Merge all the candidate sequences, which were interpreted from MS/MS spectra by de novo sequencing, into a single search string for MS BLAST search. Perform the search against a nonredundant protein database *nrdb95* (genetics.bwh.harvard.edu/msblast) using the default settings. Only consider the hits which yield statistically significant MS BLAST scores. If there are several hits, select the top hit (*see* **Note 6**).

3.6. Separation of Venom Peptides by HPLC

1. Equilibrate the Hiprep^{TM} 16/10 CM FF column with 0.02 M sodium phosphate buffer (pH 6.25).
2. Dilute peptide fractions (MW below 10 kDa) eluted from the gel filtration column (*see* **Section 3.1**) in distilled water until final concentration of 5∼10 mg/mL. Load 1 mL of the sample onto the equilibrated Hiprep^{TM} 16/10 CM FF column through a syringe loading sample injector equipped with 2 mL loop. Inject the sample into the column by pumping 20 mM sodium phosphate buffer through the loop.
3. Elute the peptides from the column with a linear gradient of 0–60% of 1 M sodium chloride (pH 6.25) over 60 min at

a flow rate of 1 mL/min. Peptide elution should be monitored at 215 nm.

4. Collect the eluted peptide peaks and load them onto the analytical Vydac C18 reverse-phase HPLC column. The eluted samples will have high salt concentration and large volume; these can be loaded onto the column through the Pump.
5. Elute peptides from the C18 column with a flow rate of 1 mL/min using a gradient of (Buffer C/Buffer D) 100/0–80/20% over 5 min, followed by a gradient of 80/20–55/45% over 50 min, followed by a gradient 55/45–40/60% over 10 min. Monitor peptide elution at 215 nm. Collect the eluted fractions and characterize the molecular mass with MALDI-TOF mass spectrometry. Freeze-dry the fractions and store them at –20°C until further analysis
6. Reduce the disulphide bonds of the peptide with 0.4 μmol of DTT in 200 μl 0.25 M Tris-HCl, containing 1 mM EDTA and 8 M guanidine HCl (pH 8.5, 37°C), incubate for 2 h), perform reaction under nitrogen.
7. Alkylate the peptide by adding 8 mg iodoacetamide to the solution of reduced peptides. Incubate at 37°C in the dark overnight under nitrogen. Terminate the reaction by injecting the sample into RP-HPLC.
8. Elute with a linear gradient of 0–60% buffer D in 80 min.
9. For automated gas-phase sequencing use 100 μg of peptide (90% purity) (*see* **Note 7**).
10. If only small amount of the peptide is available use tandem mass spectrometry for de novo peptide sequencing (*see* **Note 8**).

3.7. Peptide Bioinformatics

1. The theoretical p*I* and MW of the fully sequenced peptides can be calculated online (www.expasy.org/tools/pi_tool).
2. Determine sequence homologies by using sequence obtained from literature data and searching the nonredundant protein databases, via the *BLAST* server (www.ncbi.nlm.nih.gov/BLAST).
3. Edit the fully sequenced peptides using the Bioedit Sequence Alignment Editor software. Align and refine manually the sequences using *ClustalW 1.83* software (www.genebee.msu.su/clustal).
4. Construct phylogenetic trees using the *MEGA 3.1* and neighbour-joining method. Open the *MEGA* and click "Click me to activate a data file". Browse to and select the file containing previous aligned peptide sequences. Click OK for protein sequence. To infer a phylogenetic tree,

select Neighbor-Joining (NJ) from the Phylogeny menu. To obtain a tree with bootstrap values, repeat selecting Neighbor-Joining (NJ) from the Phylogeny menu and click the Test of Phylogeny tab and select Bootstrap of Replications (making a large number (1000) of random samples). The tree will be reconstructed and bootstrap values will be automatically added onto the original tree.

4. Notes

1. Spider venoms are heterogeneous, not only between species but also within species. The quantity and/or quality of venom have been shown to vary according to the sex, size, age and geographic origin of the spider even within the same species. Sex is a substantial factor in the intra-specific variation of spider venoms. It is better to not mix together the venom samples from both sexes but collect and analyse them separately. The heterogeneity of spider venoms within a single species due to age and geographical origin appears relatively lower.
2. There are two reasons for doing subfractionation of spider venom by gel filtration before conducting the proteome and Peptidome analysis. One is to reduce sample complexity and also to enrich the low-abundance components. The other reason is to divide the venom into two major groups: proteins with high molecular weight of above ~10 kDa for the analysis by 2-DE analysis, and peptides with MW below ~10 kDa for the analysis by LC-MS.
3. Many different types of hybrid mass spectrometers have been used for the proteome profiling of spiders venoms, including MALDI-TOF-MS, MALDI-TOF/TOF-MS/MS, ESI-Q-TOF-MS/MS, ESI-iTrap-MS/MS and ESI-Fourier-transform ion cyclotron resonance (FT-ICR) MS/MS. These differ in their sensitivity, resolving power and measurement accuracy and have been extensively reviewed (6, 7). From our own experience and that of other Labs, good results are achievable with any of these instruments, but the instruments and the protocols require careful optimizations and skillful interpretation of data.
4. LIFT-MS/MS is an ion fragmentation method by acceleration fragments for a given precursor ion in a “LIFT box” so as to overcome the disadvantages of PSD, for example, the broad distribution of fragments, the low efficiency of fragment ions detection, etc. The potential lift is the heart of the LIFT technology. It consists of three stages between four

grids: the first stage is the actual potential lift, the second stage is a focusing cell to modulate the speed of the ions, and the third stage is a post-acceleration cell. The fragment ions of different kinetic energies are accelerated when they leave the lift cell and focused onto the detector with the narrow energy distribution. These properties are essential for the high sensitivity and high signal-to-noise ratio achieved by utilizing this design.

5. Conventional methods of database searching rely heavily on matching masses of intact peptides (peptide mass fingerprinting) or their fragments(as in tandem mass spectrometry) to the predicted masses of peptides and/or peptide fragments obtained by *in silico* processing of protein sequences from database entries. But this processing is restricted to the species, for which either a complete genome and/or a substantial number of cDNA sequences are available. Unfortunately no complete genome sequence is available for even a single spider species and the number of nucleic acid and protein database entries are also relatively low. Therefore, a special consideration is needed for doing the processing of mass spectrometric data and database searching. For example, tryptic PMF alone without any tandem MS/MS data is not sufficient for the spider protein and peptide identification. Examples from publications reporting proteomic analyses of the venoms from *O. huwena* and *C. jingzhao* (3, 4) indicate that many spots from 2D-PAGE possess high-quality PMF data or tandem MS/MS spectra, but despite that have low statistical significance of the retrieved protein hits and result in ambiguous identification. So such spots should be subjected to a further de novo sequencing and the mass spectrometry-driven BLAST (MS BLAST) should be used to degenerate and interpret the tandem mass spectra. Alternatively, if the found hits are not statistically confident (their Mascot scores fall below the threshold level), their MS/MS spectra obtained from Q-TOF should be analysed and interpreted manually, e.g. with the assistance of MassLynx software (3).
6. The MS BLAST searching is performed following the procedure described by Shevchenko et al. (8).
7. The advantages of the automated gas-phase sequencing is to produce an accurate and reliable result and the ability to get the full-length sequences of the peptide toxins with about 40 residues by a single run. But this method is limited by slow speed, high cost and it requires rather large amounts of material.
8. The state-of-the-art method for de novo sequencing is mass spectrometry. And there are two kinds of approaches for

doing this. The first one is termed "bottom-up", in which the toxin is first digested by endopeptidase and then analysed and sequenced using hyphenated MS/MS techniques. The second approach is termed "top-down" due to the fact that the entire toxin is fragmented in the gas-phase within the instrument without enzymatic digestion (9). Although each approach has some shortcomings, peptide sequencing by mass spectrometry is fast and extraordinarily sensitive.

References

1. Escoubas, P., Sollod, B., and King, G.F. (2006) Venom landscapes: mining the complexity of spider venoms via a combined cDNA and mass spectrometric approach. *Toxicon* **47**, 650–663.
2. Escoubas, P. (2006) Mass spectrometry in toxicology: a 21st-century technology for the study of biopolymers from venoms. *Toxicon* **47**, 609–613.
3. Yuan, C., Jin, Q., Tang, X., Hu, W., Cao, R., Yang, S., Xiong, J., Xie, C., Xie, J., and Liang, S. (2007) Proteomic and peptidomic characterization of the venom from the Chinese bird spider, *Ornithoctonus huwena* Wang. *J. Proteome Res.* **6**, 2792–2801.
4. Liao, Z., Cao, J., Li, S., Yan, X., Hu, W., He, Q., Chen, J., Tang, J., Xie, J., and Liang, S. (2007) Proteomic and peptidomic analysis of the venom from Chinese tarantula *Chilobrachys jingzhao*. *Proteomics* **7**, 1892–1907.
5. Huang, L., Li, B., Luo, C., Xie, J., Chen, P., and Liang, S. (2004) Proteome comparative analysis of gynogenetic haploid and diploid embryos of goldfish (*Carassius auratus*). *Proteomics* **4**, 235–243.
6. Reinders, J., Lewandrowski, U., Moebius, J., Wagner, Y., and Sickmann, A. (2004) Challenges in mass spectrometry-based proteomics. *Proteomics* **4**, 3686–3703.
7. Glish, G.L. and Burinsky, D.J. (2008) Hybrid mass spectrometers for tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* **19**, 161–172.
8. Shevchenko, A., Sunyaev, S., Loboda, A., Shevchenko, A., Bork, P., Ens, W., and Standing, K.G. (2001) Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. *Anal. Chem.* **73**, 1917–1926.
9. Escoubas, P., Quinton, L., and Nicholson, G.M. (2008) Venomics: unravelling the complexity of animal venoms with mass spectrometry. *J. Mass Spectrom.* **43**, 279–295.

Chapter 7

Purification and Characterization of Biologically Active Peptides from Spider Venoms

Alexander A. Vassilevski, Sergey A. Kozlov, Tsezi A. Egorov, and Eugene V. Grishin

Abstract

Spider venoms represent invaluable sources of biologically active compounds suitable for use in life science research and also having a significant potential for biotechnology and therapeutic applications. The methods reported herewith are based on our long experience of spider venom fractionation and peptides purification. We routinely screen new peptides for antimicrobial and insecticidal activities and our detailed protocols are also reported here. So far these have been tested on species of Central Asian and European spiders from the families Agelenidae, Eresidae, Gnaphosidae, Lycosidae, Miturgidae, Oxyopidae, Philodromidae, Pisauridae, Segestriidae, Theridiidae, Thomisidae, and Zodariidae. The reported protocols should be easily adaptable for use with other arthropod species.

Key words: Spider venom, spider toxin, peptidomics, venomics, antimicrobial peptide, cytolytic peptide, insecticidal peptide, chromatography, mass spectrometry, protein sequencing, specific proteolysis.

1. Introduction

Natural venoms and spider venoms in particular represent invaluable sources of biologically active compounds of interest to basic and applied research. For instance, the classical ω-agatoxin IVA selectively identifies voltage-sensitive P-type calcium channels (1), whereas a more recently described toxin GsMTx-4 will probably become a marker of mechanosensitive (stretch-activated) ion channels (2). A number of spider venom constituents are believed to exhibit high therapeutic potential and have been proposed as possible drug leads (3). For example, psalmotoxin 1 that blocks

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_7,

a subtype of acid-sensing ion channels presents potent analgesic properties and has been suggested as a new-generation painkiller (4). Other venom-derived peptides possess antimicrobial properties and may represent a novel class of antibiotics (5, 6). Given that insects are the main biological target of spider venom, spider toxins are regarded as highly specific insecticides that could ensure effective pest control in agriculture (7, 8). Natural venoms represent complex mixtures of chemically diverse substances with spider venoms being probably the most complex. It is estimated that a few hundreds of biologically active peptides may be present in any single venom, forming the so-called evolutionary edited combinatorial libraries (9–11). This remarkable level of venoms' complexity renders the task of purification and functional assignment of individual venom peptides very difficult. Here we describe a comprehensive set of protocols used in our laboratory for the isolation and characterization of biologically active peptides from spider venoms. We focus on two biological properties of spider venom peptides – their antimicrobial and insecticidal activities (12, 13). These represent two key properties of significant practical interest, but our peptidomic protocols can be easily adapted for use with other functional tests (14).

2. Materials

1. Deionized water (resistivity of 18.2 MΩ-cm) should be used in all experiments.
2. All chemicals and solvents should be of analytical grade.

2.1. Venom Fractionation and Peptide Purification

1. Spider venoms: The protocols described here are adapted for use with crude spider venoms from the families Agelenidae, Eresidae, Gnaphosidae, Lycosidae, Miturgidae, Oxyopidae, Philodromidae, Pisauridae, Segestriidae, Theridiidae, Thomisidae, and Zodariidae (Fauna Laboratories, Almaty, Kazakhstan). These reach us in the form of lyophilized powder. Crude lyophilized venoms should be stored at –80°C (*see* **Note 1**).
2. System Gold(R) high-performance liquid chromatography (HPLC) instrument (Beckman Coulter Inc., Fullerton, CA) (*see* **Note 2**).
3. TSK 2000SW size-exclusion chromatography (SEC) column (7.5 × 600 mm, 12.5 nm pore size, 10 μm particle size; Toyo Soda Manufacturing Co., Tokyo, Japan) (*see* **Note 2**).
4. Elution solvent 1: 10% (v/v) acetonitrile, 0.1% (v/v) trifluoroacetic acid (TFA) (*see* **Note 3**).

5. Kontron HPLC instrument (Kontron Instruments, Milan, Italy) (*see* **Note 2**).
6. Vydac 218TP54 C_{18} reversed-phase (RP) column (4.6 × 250 mm, 30 nm pore size, 5 μm particle size; Separations Group Inc., Hesperia, CA) (s*ee* **Note 2**).
7. Elution solvent 2: 0.1% (v/v) TFA in water.
8. Elution solvent 3: 0.1% (v/v) TFA in acetonitrile.
9. "Ascentis" RP-Amide column (4.6 × 150 mm, 10 nm pore size, 3 μm particle size; Sigma-Aldrich Corp., St. Louis, MO) (*see* **Note 2**).

2.2. Structural Characterization of Venom Peptides

1. ZipTip pipette tips with 0.6 μL C_{18} resin (Millipore, Billerica, MA).
2. A vacuum concentrator Savant SpeedVac (GMI Inc., Ramsey, MN) or a similar instrument.
3. FreeZone freeze dry system (Labconco Corp., Kansas City, MO) or a similar instrument.
4. Reaction solution 1: 0.2 M Tris–HCl, 6 M Gu-HCl, 2 mM EDTA, pH 8.
5. Dithiothreitol solution: 1 M DTT in acetonitrile (*see* **Note 4**).
6. 4-Vinylpyridine solution: 50% (v/v) 4-VP in isopropanol.
7. Peptide sequencing by Edman degradation: Procise Model 492 protein sequencer and all reagents (Applied Biosystems Inc., Foster City, CA) (*see* **Note 5**).
8. Reaction solution 2: 50 mM sodium phosphate buffer, 10 mM DTT, 1 mM EDTA, pH 7.
9. Pyroglutamate aminopeptidase (pyroglutamyl-peptidase I) – a cysteine peptidase from the archeon *Pyrococcus furiosus* (EC: 3.4.19.3; Sigma-Aldrich). To make the enzyme stock solution 1, add 1 milliunit of pyroglutamate aminopeptidase per 10 μL of the reaction solution 2 (*see* **Note 6**).
10. Reaction solution 3: 80% TFA (*see* **Note 7**).
11. Cyanogen bromide solution 1: 5 M CNBr in acetonitrile.
12. Cyanogen bromide solution 2: 50 mM CNBr in acetonitrile.
13. Proteolytic enzymes for protein sequencing: Trypsin from bovine pancreas (EC: 3.4.21.4; F. Hoffmann-La Roche Ltd., Basel, Switzerland), Endoproteinase Arg-C (tissue kallikrein) – a serine protease from mouse submaxillary gland (EC: 3.4.21.35; Sigma-Aldrich), Endoproteinase Lys-C (lysyl endopeptidase) – a serine protease from the bacterium *Lysobacter enzymogenes* (EC: 3.4.21.50; Sigma-Aldrich), Endoproteinase Asp-N (peptidyl-Asp met-

alloendopeptidase) – a metalloprotease produced by the bacterium *Pseudomonas fragi* (EC: 3.4.24.33; F. Hoffmann-La Roche), Endoproteinase Glu-C (V8 protease; glutamyl endopeptidase) – a serine protease produced by the bacterium *Staphylococcus aureus* strain V8 (EC: 3.4.21.19; F. Hoffmann-La Roche) (*see* **Note 8**).

14. Reaction solutions, enzyme stock solutions – *see* **Table 7.1** (*see* **Note 6**).

Table 7.1
Proteolytic enzymes for protein sequencing

Enzyme	Stock solution (kept at –80°C)	Reaction solution	Enzyme–substrate ratio (w/w)
Trypsin	1 μg/μL, 1 mM HCl	100 mM Tris–HCl, pH 8.5	1:50
Arg-C	0.1 μg/μL, 1 mM HCl	100 mM NH_4HCO_3, pH 8.5	1:50
Lys-C	0.1 μg/μL, 50 mM Tricine, 10 mM EDTA, pH 8	100 mM NH_4HCO_3, pH 8.5	1:20
Asp-N	0.04 μg/μL, 10 mM Tris–HCl, pH 7.5	50 mM Tris–HCl, pH 8	1:20
Glu-C	1 μg/μL in water	50 mM CH_3COONH_4, pH 4	1:50

2.3. Functional Characterization Of the Isolated Venom Peptides

1. Sterile flat-bottom polypropylene 96-well microtiter plates, for example, BD Falcon microplates (BD Biosciences, San Jose, CA) (*see* **Note 9**).

2.3.1. Antimicrobial Activity

2. Biological material: Frozen bacterial cultures (*Escherichia coli* DH5α, *Staphylococcus aureus* 209P) (*see* **Note 10**).
3. Instruments: A laminar airflow cabinet, such as HLAF or VLAF (Gelaire, Sydney, Australia); a shaking incubator, for instance from LabScientific Inc. (Livingston, NJ) or similar; iMark microplate absorbance reader (Bio-Rad Laboratories Inc., Hercules, CA) or equivalent plate reader.

2.3.2. Insecticidal Activity

1. Biological material: Flesh fly *Sarcophaga carnaria* maggots (weight ~50–60 mg); these are usually available in abundance from good fishing tackle shops (*see* **Note 11**).
2. Instruments: Microliter syringes (10 μL size; Hamilton Company, Bonaduz, Switzerland) or similar 10 μL syringes.

3. Methods

3.1. Fractionation of Spider Venoms

Traditional approaches to venom fractionation rely on a combination of RP-HPLC and ion-exchange chromatography (15–17). In our experience the best fractionation strategy capable of tackling spider venom complexity and yielding homogenous preparations of individual venom peptides includes SEC followed by two rounds of RP-HPLC (*see* **Fig. 7.1**). Detailed description of the

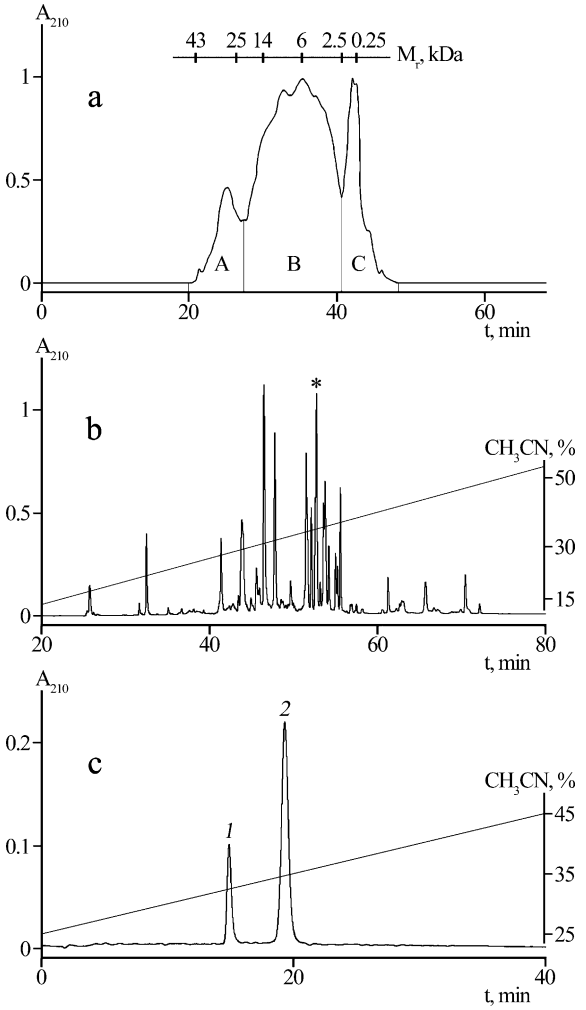


Fig. 7.1. Purification of biologically active peptides from spider venom. (**a**) *Lachesana tarabaevi* crude venom (10 μL) separation using SEC on a TSK 2000SW column; isolated fractions A, B, and C are indicated. Molecular standard retention times and the corresponding molecular masses are shown at the top. (**b**) Second step chromatography of peptidic fraction B using RP-HPLC on a Vydac 218TP54 C_{18} column in a linear gradient of acetonitrile (shown with a *line*). Antimicrobial and insecticidal fraction containing latarcin 2a (6) and cyto-insectotoxin 1a (12) is indicated with a *star*. (**c**) Final step purification of latarcin 2a (*1*) and cyto-insectotoxin 1a (*2*) using RP-HPLC on an Ascentis RP-Amide column in a linear gradient of acetonitrile (shown with a *line*).

purification methods may be found elsewhere (18). This approach is fully compatible with different mass spectrometry (MS) techniques and can be automated easily (16).

3.1.1. Size-Exclusion Chromatography (SEC)

1. Set up SEC system: Wash the tubes and the injector, pre-equilibrate the column with elution solvent 1, set the flow rate to 0.5 mL/min, and check the pressure (*see* **Note 12**).
2. Run a blank experiment: Inject 50–100 μL of pure elution solvent 1 instead of your sample; monitor the baseline for approximately 2 h (*see* **Note 13**).
3. Prepare sample: Dissolve lyophilized spider venoms in elution solvent 1 (the equivalent of 10 μL of crude venom should be dissolved in ~100 μL of the solvent) (*see* **Notes 14** and **15**); spin samples in a microcentrifuge at ~15,000×*g* for 15 min and carefully transfer clear supernatant into a fresh tube.
4. Inject venom sample in the SEC column; monitor effluent absorbance at 210 and/or 280 nm (*see* **Note 16**); collect 0.5 mL fractions. An example of separation is shown in **Fig. 7.1a**.
5. Store collected fractions at 4°C (*see* **Note 1**).
6. After use, the SEC column should be washed with the running solvent followed by ethanol and sealed; columns can be stored with ethanol at room temperature for over a year.

3.1.2. Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC), First Round

1. Set up RP-HPLC system: Wash the tubes and the injector, pre-equilibrate the column with elution solvent 2, set the flow rate to 0.7 mL/min, and check the system pressure (*see* **Note 12**).
2. Run a blank experiment: Inject 1 mL of pure elution solvent 2 (no venom components should be injected at this stage), perform blank separation with a quick sharp gradient of elution solvent 3 concentration (0–80% in 20 min), and monitor the baseline (*see* **Note 13**).
3. Prepare sample: Concentrate the eluted fraction (0.5 mL fractions obtained in **Section 3.1.1**, Step 4) using vacuum concentrator to reduce sample volume to ~0.1 mL; add water to the final volume of 1 mL (*see* **Note 17**).
4. Inject sample in the RP column, perform separation with a long shallow gradient of elution solvent 3 concentration (0–60% in 90 min); monitor absorbance at 210 and/or 280 nm (*see* **Note 16**). At this stage it might be best to collect fractions manually. An example of separation is shown in **Fig. 7.1b**.
5. Store collected fractions at 4°C (*see* **Notes 1** and **18**).

6. After use, the RP column should be washed with the running solvent followed by ethanol and sealed; columns can be stored with ethanol at room temperature for over a year.

3.1.3. Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC), Second Round (see **Note 19***)*

1. Set up RP-HPLC system: Wash the tubes and the injector, pre-equilibrate the column with elution solvent 2, set the flow rate to 1 mL/min, and check the system pressure (*see* **Note 12**).
2. Run a blank experiment: Inject 1 mL of pure elution solvent 2 (no venom components should be injected at this stage), perform blank separation with a quick sharp gradient of elution solvent 3 concentration (0–80% in 20 min), and monitor the baseline (*see* **Note 13**).
3. Prepare sample: Dilute the eluted fractions (obtained in **Section 3.1.2**, Step 4) in water 2–4-fold (s*ee* **Note 17**).
4. Inject sample in the RP column, perform separation with a long shallow gradient of elution solvent 3 concentration (0–60% in 120 min); monitor absorbance at 210 and/or 280 nm (*see* **Note 16**). At this stage it might be best to collect fractions manually. An example of separation is shown in **Fig. 7.1**c.
5. Store collected fractions at 4°C (*see* **Notes 1** and **18**).
6. After use, the RP column should be washed with the running solvent followed by ethanol and sealed; columns can be stored with ethanol at room temperature for over a year.

3.2. Structural Characterization of Venom Peptides

In our work we have chosen to follow a simple scheme of polypeptide structure analysis described in 19. First, the presence or absence of disulfide bonds should be assessed. This is followed by Edman degradation, used to determine the peptide sequence. If polypeptide lengths exceed ~40–50 residues, selective proteolysis should be performed (*see* **Fig. 7.2**). We routinely use a number

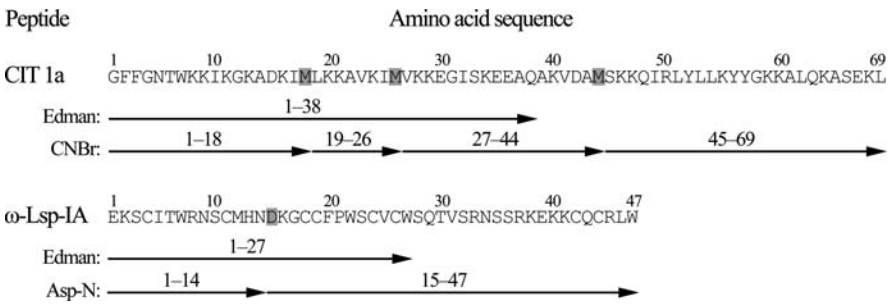


Fig. 7.2. Peptide sequencing techniques. To establish full amino acid sequences of long peptides, selective proteolysis should supplement Edman degradation. Two examples are listed including cyto-insectotoxin 1a (CIT 1a) (12) and ω-Lsp-IA (14). Selective cleavage sites are indicated by shading, the corresponding peptide fragments are represented by arrows.

of cleavage reagents for protein sequencing that have also been reviewed by others (20). MS-based analysis is fully compatible with this approach and could be used to complement or replace Edman sequencing.

3.2.1. Disulfide Bond Reduction and Thiol Group Alkylation

1. Freeze-dry 0.1–1 nmol of the peptide obtained in **Section 3.1.3** (*see* **Note 20**). For MS analysis only, use smaller aliquots (10 pmol of the peptide should suffice).
2. Dissolve the sample in 50 μL of reaction solution 1.
3. Fill the tube with nitrogen to remove atmospheric oxygen and incubate at 60°C for 1 h.
4. Add 2 μL of DTT solution.
5. Incubate at 40°C for 18 h.
6. Add 4 μL of 4-VP solution.
7. Incubate sample in the dark at room temperature for 15 min.
8. For MS analysis only, desalt your sample using ZipTip.
9. For Edman sequencing, run RP-HPLC separation immediately, *see* **Section 3.1.2** for details. Elute the excess reagents and side-products isocratically in 5% of elution solvent 3, and then run a gradient elution as described in **Section 3.1.2**, Step 4 (*see* **Notes 1, 18,** and **21**).

3.2.2. N-Terminal Sequencing by Edman Degradation and Pyroglutamate Removal

We rely on the automated stepwise Edman degradation using Applied Biosystems Model 492 protein sequencer and would not recommend any departures from the manufacturer's protocol. Cysteine residues are determined as *S*-pyridylethylated derivatives. A pyroglutamic acid residue (originating from a glutamine) sometimes N-terminally blocks peptides from spider venom and should be removed to allow sequencing.

1. Freeze-dry 0.1–1 nmol of the peptide obtained in **Section 3.2.1** (or **Section 3.1.3** for peptides that do not contain cysteine residues) (*see* **Note 20**).
2. Dissolve the sample in 50 μL of reaction solution 2.
3. Add 10 μL of enzyme stock solution 1.
4. Incubate at 75°C for 2 h (the enzyme is heat-stable).
5. Separate proteolytic fragments by RP-HPLC; follow the same steps as in **Section 3.1.2** (*see* **Notes 1** and **18**).

3.2.3. Selective Proteolysis with Cyanogen Bromide (CNBr)

1. Freeze-dry 0.1–1 nmol of the peptide obtained in **Section 3.2.1** (or **Section 3.1.3** for peptides that do not contain cysteine residues) (*see* **Note 20**). For MS analysis only, use smaller aliquots (10 pmol of the peptide should suffice).
2. Dissolve the sample in 50 μL of reaction solution 3.
3. Add 1 μL of CNBr solution 1. For MS analysis only, add 1 μL of CNBr solution 2.

4. Incubate the reaction in the dark at room temperature overnight.
5. Terminate the reaction by adding 0.5 mL of water.
6. For MS analysis only, evaporate the sample on a vacuum concentrator to the volume of ~10 μL. Peptide mixture may be analysed by matrix-assisted laser desorption/ionization (MALDI) MS without further purification (*see* **Note 18**).
7. For Edman sequencing, dry the sample on a vacuum concentrator to the volume of 10–50 μL, then dilute the sample in 1 mL of elution solvent 2, and run RP-HPLC separation as in **Section 3.1.2**. Elute the excess reagents and side-products isocratically in elution solvent 2 (s*ee* **Notes 1, 18, 22,** and **23**). An example of peptide sequencing using CNBr is listed in **Fig. 7.2**.

3.2.4. Enzymatic Cleavage

1. Freeze-dry 0.1–1 nmol of the peptide obtained in **Section 3.2.1** (or **Section 3.1.3** for peptides that do not contain cysteine residues) (*see* **Note 20**). For MS analysis only, use smaller aliquots (10 pmol of the peptide should suffice).
2. Dissolve the sample in 50 μL of the required reaction solution (*see* **Table 7.1**).
3. Add the enzyme stock solution (*see* **Table 7.1** for details).
4. Incubate at 40°C for 4 h.
5. For MS analysis only, desalt the sample using ZipTip. Otherwise, run RP-HPLC separation as described in **Section 3.1.2** (*see* **Notes 23** and **24**). An example of peptide sequencing using enzymatic cleavage is listed in **Fig. 7.2**.

3.3. Functional Characterization of the Isolated Venom Peptides

3.3.1. Antimicrobial Activity (see **Note 25***)*

1. Prepare bacterial night culture: Transfer a piece of stock frozen bacterial culture stored at –80°C into 5 mL of LB medium in 50 mL tubes with a sterile pipette tip (keep frozen bacteria on dry ice during the procedure); grow bacteria overnight at 37°C with a vigorous shaking (220 rpm).
2. Dilute the resulting culture 200-fold using LB medium.
3. Continue incubation for another 3–5 h till the optical density of the culture at 600 nm (OD_{600}) reaches ~0.6.
4. Dilute the fresh culture 100,000-fold (*see* **Note 26**) using two serial dilutions: First transfer 10 μL of culture into 1 mL of LB in a 1.5 mL sterile tube, then transfer 10 μL of the resulting dilution into 10 mL of LB in a 15 mL sterile tube. Use this final bacterial suspension for the experiment.
5. Freeze-dry 0.2–1 nmol of peptides obtained in **Section 3.1.2** or **3.1.3** (*see* **Notes 20** and **27**); dissolve the lyophilized samples in 22 μL of water.
6. Using sterile flat-bottom 96-well plate, mix 90 μL aliquots of bacterial suspension with 10 μL peptide aliquots; use

10 μL of water as a negative control (uninhibited bacterial growth). Use wells filled with 100 μL of LB medium to control for medium sterility. Each peptide or control well should be tested at least in duplicate. Fill all of the unused wells with LB medium or water to reduce evaporation of the test wells and seal the plate. All the procedures should be performed under sterile conditions, e.g. in a laminar airflow cabinet.

7. Incubate the plate at 37°C in a shaking incubator (220 rpm) overnight.
8. Measure OD_{600} in each well using a plate reader (*see* **Note 28**). Sterility control wells should remain clear. Lower absorbance values in the test wells or clear test wells would indicate the antimicrobial activity.

3.3.2. Insecticidal Activity

1. Freeze-dry 0.2–1 nmol of peptides obtained in **Section 3.1.2** or **3.1.3** (*see* **Notes 20** and **29**); dissolve the lyophilized samples in 22 μL of water (*see* **Note 30**).
2. Inject up to 10 μL of samples into the third to fifth segments of larvae using a microsyringe. Test at least two individual larvae for each peptide. Inject water as negative control.
3. Monitor toxic (paralytic or lethal) effects for 24 h following the injection.

4. Notes

1. Venoms should best be kept as lyophilized powder. When dissolved, aliquots (1–100 μL) should also be kept at –80°C until use. Polypeptide fractions containing acetonitrile could be kept at –20°C or 4°C; lyophilized peptides should be stored at –20°C.
2. There is an abundance of HPLC instruments and separation media on the market and individual laboratory setups are likely to be different from the ones described here. Our protocols can be easily adapted for use with different HPLC hardware.
3. The mobile phase used for SEC has been adjusted to minimize non-specific sorption of the venom components onto the stationary phase, to improve resolution and the accuracy of estimating the molecular weights of the eluted components. Other elution solvent compositions could be used if necessary, e.g. phosphate-buffered saline or Tris-based buffers, but the separation profiles may be substantially different.

4. Use a freshly prepared solution. Water, acetonitrile, ethanol, isopropanol can all be used to dissolve DTT. In case of dithioerythritol (DET), however, water cannot be used.
5. Other hardware or commercial peptide sequencing services can be used instead.
6. Store enzyme solutions at –20°C or –80°C.
7. Alternatively, 70% formic acid or 0.1 M HCl can be used.
8. We routinely use these enzymes and also CNBr; together these should be sufficient for sequencing. The enzymes should be of protein sequencing grade. This is usually specified by the manufacturer.
9. Standard bacterial media, e.g. Muller-Hinton (MHB) or Luria-Bertani (LB) broth, and procedures can be used, but importantly, polystyrene plates should not be used, as they adsorb peptides. The choice of plates will affect the result strongly. It is therefore advisable to use the same type of plates in all experiments to reduce the variability.
10. We recommend to use *E. coli* ATCC 25922, *Enterococcus faecalis* ATCC 29212, *Pseudomonas aeruginosa* ATCC 27853, *S. aureus* ATCC 27660 or ATCC 29213 (21, 22). However, for the purpose of novel antimicrobials discovery this is not always necessary, and strains conventionally found in laboratories may be used. We also recommend checking activity on at least one Gram-positive and one Gram-negative bacterium.
11. It is possible to use other insects, which might be easier to obtain, such as housefly *Musca domestica* and the oriental cockroach *Blatta orientalis* or the speckled cockroach *Nauphoeta cinerea*. However, the susceptibility of different species to toxins may vary greatly and more material would be required for testing larger insects. *Drosophila melanogaster* would be an ideal test insect, but more sophisticated instrumentation might be required due to small size of the insects.
12. The flow rates and pressure limits are specified by the SEC/RP-HPLC column and chromatograph producers. In SEC, column equilibration is a time-consuming process, because it normally takes 5–10 column volumes of elution solvent to equilibrate the column prior to separation. We recommend running the procedure overnight at a low rate of ~0.1 mL/min. In RP-HPLC, column equilibration is achieved within a few minutes at normal speed rate.
13. If any peaks are detected, new solvents should be made and the system be washed again. If the peaks persist, the blank profile may be subtracted from the experimental traces.

14. To estimate the amount of the crude venom required, we assume an average total polypeptide concentration in spider venom of ~0.2 mg/μL, although actual values will differ between species.
15. Smaller sample volumes should be used to improve resolution in SEC. We recommend that no more than 1/100 of the column void volume is loaded.
16. For protein characterization, both absorbance values may be useful (peptide bond absorbs at ~210 nm; aromatic residues at ~280 nm).
17. Solvent evaporation and addition of water are needed to lower acetonitrile concentration and thereby allow sample absorption on the reversed phase. We do not recommend drying the samples using a vacuum concentrator due to the risk of polypeptide adsorption on the tube surface; alternatively, lyophilize the samples. Before running RP-HPLC, make sure the concentration of acetonitrile in the sample is at least 15% lower than the corresponding elution conditions.
18. Small fractions may be set aside at this stage for MS analysis (off-line LC-MS configuration). We use M@LDI-LR (Micromass, Manchester, UK) for routine fraction characterization or Ultraflex TOF-TOF (Bruker Daltonik GmbH, Bremen, Germany) for tandem MS/MS analysis and follow the manufacturer's protocols. Normally 0.5–10 μL aliquots of the RP-HPLC eluates are sufficient for MALDI-MS analysis.
19. The third step of separation is usually required to purify a selected component to homogeneity. We prefer running a second round of RP-HPLC. The conditions will vary depending on the peptide being purified but usually these will differ from the conditions used for the first round of RP-HPLC. One may either change stationary phase (e.g. as in the example described), mobile phase (e.g. use alcohol instead of acetonitrile, pentafluoropropionic acid instead of TFA) or both. Alternatively, the temperature can be changed.
20. To estimate peptide concentration measure UV absorbance at 280 (A_{280}) and 260 (A_{260}) nm. Peptide concentration (C) may be calculated as follows: $C\ [mg/mL] = 1.55 \times A_{280} - 0.76 \times A_{260}$. For peptides with known amino acid sequences their molar extinction coefficients should be used to precisely determine concentration. For those peptides that do not contain aromatic residues, concentration may be measured during the first step of Edman degradation.

21. The retention time of the modified peptide may change significantly and several products may be recovered from a multi-chain peptide. The alkylated peptides absorb better at 280 nm. Possible side-products include hypo- and hyper-alkylated peptides and 4-VP self-polymerization products (because of the exposure to light at Step 7). These could be identified by comparing the absorbance at 210 and 280 nm (typical absorbance ratio 210/280 for peptides is usually ~10 or greater).
22. A small fraction of eluted products (usually <5%) will contain a C-terminal homoserine instead of a homoserine lactone.
23. Here we often use a smaller column (diameter 1–2 mm), such as for example, a Luna C_{18} column (2 × 150 mm, 10 nm pore size, 3 μm particle size; Phenomenex Inc., Torrance, CA) and a different flow rate of 50–300 μL/min.
24. The reported protocol usually works well with the presented set of enzymes. Occasionally we have incomplete and/or non-specific cleavage; these products may be easily identified using MS.
25. Determination of peptide minimal inhibitory concentrations is usually performed using a microtiter broth dilution assay (21, 22). Described here is a simplified version which is adequate for screening purposes.
26. This will usually result in a final concentration of 10^4–10^5 colony-forming units/mL.
27. Antimicrobial peptides are active in the micromolar range. To reach the final test concentration of 1 μM in a volume of 100 μL, one needs 0.1 nmol of peptide.
28. The results usually come out in a "yes–no" fashion; the active fractions producing clear wells can be detected without the use of a plate reader.
29. Lethal doses of neurotoxins for fly larvae usually lie in the 0.1–10 mg/kg range. Similar to **Note 27**, 0.1 nmol of peptide will give an average dose of 10 mg/kg in a 50 mg-larva.
30. Physiological saline is recommended, but water may be used during screening.

Acknowledgments

This work was funded by the Russian Foundation for Basic Research (Grant Nos. 08-04-00454 and 08-04-90444), the Fed-

eral Agency for Education of the Russian Federation (State Contact No. P 1388) and the Program of Cell and Molecular Biology of the Russian Academy of Sciences.

References

1. Mintz, I.M., Venema, V.J., Swiderek, K.M., Lee, T.D., Bean, B.P., and Adams, M.E. (1992) P-type calcium channels blocked by the spider toxin omega-Aga-IVA. *Nature* **355**, 827–829.
2. Suchyna, T.M., Johnson, J.H., Hamer, K. et al. (2000) Identification of a peptide toxin from *Grammostola spatulata* spider venom that blocks cation-selective stretch-activated channels. *J. Gen. Physiol.* **115**, 583–598.
3. Estrada, G., Villegas, E., and Corzo, G. (2007) Spider venoms: a rich source of acylpolyamines and peptides as new leads for CNS drugs. *Nat. Prod. Rep.* **24**, 145–161.
4. Mazzuca, M., Heurteaux, C., Alloui, A. et al. (2007) A tarantula peptide against pain via ASIC1a channels and opioid mechanisms. *Nat. Neurosci.* **10**, 943–945.
5. Kuhn-Nentwig, L. (2003) Antimicrobial and cytolytic peptides of venomous arthropods. *Cell. Mol. Life Sci.* **60**, 2651–2668.
6. Kozlov, S.A., Vassilevski, A.A., Feofanov, A.V., Surovoy, A.Y., Karpunin, D.V., and Grishin, E.V. (2006) Latarcins, antimicrobial and cytolytic peptides from the venom of the spider *Lachesana tarabaevi* (Zodariidae) that exemplify biomolecular diversity. *J. Biol. Chem.* **281**, 20983–20992.
7. Nicholson, G.M. (2007) Insect-selective spider toxins targeting voltage-gated sodium channels. *Toxicon* **49**, 490–512.
8. King, G.F. (2007) Modulation of insect Ca(v) channels by peptidic spider toxins. *Toxicon* **49**, 513–530.
9. Sollod, B.L., Wilson, D., Zhaxybayeva, O., Gogarten, J.P., Drinkwater, R., and King, G.F. (2005) Were arachnids the first to use combinatorial peptide libraries? *Peptides* **26**, 131–139.
10. Escoubas, P. and Rash, L. (2004) Tarantulas: eight-legged pharmacists and combinatorial chemists. *Toxicon* **43**, 555–574.
11. Escoubas, P., Sollod, B., and King, G.F. (2006) Venom landscapes: mining the complexity of spider venoms via a combined cDNA and mass spectrometric approach. *Toxicon* **47**, 650–663.
12. Vassilevski, A.A., Kozlov, S.A., Samsonova, O.V. et al. (2008) Cyto-insectotoxins, a novel class of cytolytic and insecticidal peptides from spider venom. *Biochem. J.* **411**, 687–696.
13. Lipkin, A., Kozlov, S., Nosyreva, E., Blake, A., Windass, J.D., and Grishin, E. (2002) Novel insecticidal toxins from the venom of the spider *Segestria florentina*. *Toxicon* **40**, 125–130.
14. Pluzhnikov, K., Vassilevski, A., Korolkova, Y. et al. (2007) Omega-Lsp-IA, a novel modulator of P-type Ca2+ channels. *Toxicon* **50**, 993–1004.
15. Escoubas, P., Quinton, L., and Nicholson, G.M. (2008) Venomics: unravelling the complexity of animal venoms with mass spectrometry. *J. Mass Spectrom.* **43**, 279–295.
16. Liang, S. (2008) Proteome and peptidome profiling of spider venoms. *Expert Rev. Proteomics* **5**, 731–746.
17. Kozlov, S.A., Vassilevski, A.A., and Grishin, E.V. (2008) in "Peptidomics: Methods and Applications" (Soloviev, M., Shaw, C. and Andrén, P., Eds.), John Wiley & Sons, Hoboken, NJ, pp. 55–70.
18. Cutler, P. (2004) Protein purification protocols, Humana Press, Totowa, NJ.
19. Darbre, A. (1986) Practical protein chemistry: a handbook, Wiley, Chichester, NY.
20. Smith, B.J. (2003) Protein sequencing protocols, Humana Press, Totowa, NJ.
21. Otvos, L. and Cudic, M. (2007) Broth microdilution antibacterial assay of peptides. *Methods Mol. Biol.* **386**, 309–320.
22. Wiegand, I., Hilpert, K., and Hancock, R.E. (2008) Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* **3**, 163–175.

Chapter 8

MALDI-TOF Mass Spectrometry Approaches to the Characterisation of Insect Neuropeptides

Robert J. Weaver and Neil Audsley

Abstract

The diversity of insect neuropeptides coupled with the limitations from the small size of the insects themselves combine to make positive identification through peptide sequencing a highly challenging task. The advent of the "soft-ionisation" techniques of MALDI-TOF and electrospray (ESI)-Q-TOF mass spectrometry, coupled with the additional information from insect genome projects have revolutionised the characterisation of insect neuropeptides, such that sequences can now be obtained from just a few cells, where before thousands of insects had to be laboriously dissected, extracted and purified. Some of the procedures that are now used to identify these peptides are described here. Once the neuropeptides have been identified, it then becomes possible to use this knowledge to define physiological functionality.

Key words: MALDI-TOF, insect neuropeptide, direct tissue analysis, perfusion extraction, reflectron mode analysis, post-source decay, "De novo" sequencing.

1. Introduction

The sequencing and identification of small quantities of neuropeptides using mass spectrometric techniques has become an increasingly important component of the study of physiological processes in insects and other invertebrates. This is largely because of the sensitivity that can be achieved in relation to the source material, which is typically very small, such that peptides may be identified and sequenced from extracts of relatively few tissues (1), and sometimes even single glands (2), nerves (3) and individual cells (4). Whereas it was necessary previously to collect laboriously thousands of individual tissues and conduct several HPLC steps, e.g. for Edman degradation, it is now possible to obtain sequences from just a few femtomoles of material (just a few tissues) and with very little purification. The two

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_8,

main approaches that are used are the so-called "soft-ionisation" techniques (5) of matrix-assisted laser desorption-ionisation time-of-flight (MALDI-TOF) and electrospray ionisation quadrupole time-of-flight (ESI-Q-TOF), depending on the instruments that are available and the particular requirements. Both techniques may be used to achieve the same aims, the identification and sequencing of neuropeptides through the mass measurement of the ions from both the intact and fragmented peptide, although only the details for MALDI-TOF are dealt with here. The MALDI-TOF process involves the use of a specialised instrument, in which a small quantity of sample is mixed and co-crystallised with an organic matrix (simply by allowing it to dry) and is irradiated by a pulsed nitrogen laser generating ions that are measured using a mass spectrometer. The time required for ions to reach a detector at the opposite end of a flight tube is measured. This is converted into a mass/charge ratio (m/z) which can be related to the mass of the intact peptide or fragment. The number of ions reaching the detector at any given time is also measured, which is reported as the signal intensity, and is an indication of the abundance of the ions at any particular m/z. Various procedures, which are instrument- or manufacturer-dependent, may be used to improve resolution of the ions such as "Delayed Extraction™" and "velocity focussing", and there are differences in the way that the ions are recorded and results visualised. For peptide analysis, the parent ions and fragments are measured separately. An "ion gate" is used to "select" individual parent peaks, which are then increasingly fragmented by upwards adjustment of the laser energy. In a reflectron time-of-flight mass spectrometer, the peptides are fragmented by a process known as "post-source decay" (PSD) (6), whilst in a tandem mass spectrometer (MS-MS, TOF-TOF or MS^2) the process is collision-induced dissociation (CID) (7). In either case, the spectrum resulting from the fragmentation of a singly charged peptide molecule will contain mass signals coming from immonium ions, N-terminal fragment ions (*a*-, *b*-, *c*- and *d*-type ions), C-terminal fragment ions (*x*-, *y*- and *z*-type ions) and internal (double cleavage) fragment ions. In addition, many of these ions may yield satellite peaks due to loss of ammonia from lysine or arginine (–17 mass units), or loss of water from serine or threonine (–18 mass units) (8). All of this makes the interpretation of PSD and MS2 data a potentially complex procedure, although confirmation of the sequence data by matching experimentally attained fragmentation spectra to "theoretical" spectra is often more readily achievable than "de novo' interpretation. There are several published and web-based descriptions and tutorials on the interpretation of PSD (8, 9) and MS^2 data (10, and references), which should be consulted for further guidance. The procedures outlined here describe the application of some of these techniques for the identification of insect neuropeptides.

2. Materials

2.1. Tissue Dissection and Peptide Extraction

1. Insect saline (e.g. 150 mM NaCl, 10 mM KCl, 4 mM CaCl2, 2 mM MgCl2, 10 mM HEPES, pH 7.0, sterile filtered) or Phosphate-buffered saline (e.g. Dulbecco's PBS, sterile filtered). Store at 4°C.
2. Distilled or purified water (De-ionised, 0.22-μm filtered)
3. Methanol (HPLC grade)
4. Trifluoroacetic acid (TFA, 0.1% v/v)

2.2. High-Performance Liquid Chromatography (HPLC) of Peptide Extracts

1. HPLC solvents: Acetonitrile (HPLC-grade)/0.1% TFA (v/v) and water (HPLC grade)/0.1% TFA (*see* **Note 1**)
2. TFA (0.1%) for dilution of extracts
3. Acetonitrile/water (70:30) for redissolving dried extracts
4. Narrow-bore column LC column: e.g. Jupiter C_{18} 10 μm 300 Å (250 mm × 2.1 mm i.d.; Phenomenex)
5. Guard column: similar packing material (as above).

2.3. MALDI-TOF Mass Spectrometry

1. MALDI-TOF matrix (α-CHCA): For routine MALDI-TOF matrix, α-cyano-hydroxy-4-cinnamic acid (Sigma C-2020) is dissolved at approximately 10 mg/mL in 50% acetonitrile, 50% 0.1% TFA (ultra-pure) (*see* **Notes 2** and **3**)
2. Direct tissue analysis matrix: Routine matrix (as above) diluted 50:50 with methanol.
3. Standard peptides (for calibration): e.g. Sequazyme Peptide Mass Standards Kit (P2-3143-00). Alternatively, individual components, e.g. angiotensin I (Bachem H-1680), *des*-Arg^1-bradykinin (Bachem H-2200), Glu^1-fibrinopeptide B (Bachem H-2950) are dissolved at 1–2 μM (1 or 2 pmol/μL) in 30% acetonitrile, 0.01% TFA. The standard peptide solutions should be stored in small aliquots (e.g. 20 μL) at –20°C.

3. Methods

Neuroendocrine tissues are dissected from insects, under saline, using fine dissecting scissors, watchmaker's forceps and a dissecting microscope (*see* **Note 4**). Individual tissues should be briefly rinsed in sterile saline or water before transfer either to extraction/perfusion tubes (microfuge tubes) or directly onto a stainless steel target plate (*see* below). Tissues may be analysed directly,

although individual peptides may not always be present in sufficient amounts (or ionised sufficiently well) to provide good sequence information. For most PSD analyses, several tissues are collected together for perfusion or extraction, followed either by direct MALDI-TOF analysis (e.g. using a fraction of the perfusate) or by HPLC separation and then analysis of individual fractions.

3.1. Tissue Dissection and Peptide Extraction

3.1.1. Direct Tissue Analysis

1. Individual tissues or pieces of tissue are transferred from the dissection using a dissecting pin or loop and placed directly on a stainless steel MALDI-TOF target plate together with a small drop (< 0.5 μL) of ice-cold distilled water, which is blotted off within 30 s using a small piece of cellulose filter paper.
2. The tissue is immediately covered with small droplet (ca. 0.2 μL) of α-CHCA matrix diluted 50:50 with methanol and allowed to dry (the so-called "dried-droplet" technique). The preparation is now ready for MALDI-TOF analysis without further processing and without removing the tissue sample from the target plate (*see* **Note 5**).
3. Standard peptides, mixed in similar matrix, are spotted adjacent to the tissue samples. The samples and standards are allowed to air dry for a few minutes and the target plate can then be inserted into the mass spectrometer.
4. For direct tissue analysis, the laser shots are directed at the matrix area immediately surrounding the dried tissue, into which the methanol-soluble peptides will have been extracted. For all other samples, the laser can be targeted in any region of the sample spot (*see* **Note 6**).

3.1.2. Perfusion Extraction of Neuropeptides

1. To increase the concentration of peptides, either for direct analysis or HPLC separation, several tissues are collected into the same tube and perfused with 100% methanol on ice for 15–30 min. Before transfer into the methanol, the tissues are first rinsed in saline or water which is then blotted from the forceps or transfer pin using a small piece of cellulose paper, taking care not to touch the tissue itself (*see* **Note 7**).
2. After perfusing the tissues in methanol, the supernatant is removed to a fresh microfuge tube into which several such perfusates may be combined.
3. The tissues may then be separately extracted using 0.1% TFA, which encourages the elution of some peptides that do not appear to be readily eluted using methanol alone (*see* **Note 8**). The methanol extracts can be either partially concentrated by vacuum evaporation prior to direct analysis

(*see* **Note 9**) or may be diluted for HPLC separation. For direct analysis, an aliquot (e.g. 0.5–1 μL) of concentrated methanol extract is mixed with an equal volume of α-CHCA matrix and then applied to the MALDI-TOF target plate using a micro-pipettor (1–2 μl). This can be done by adding a droplet of the sample to a droplet of matrix on the plate, or by pre-mixing droplets of the matrix and sample on a small piece of parafilm M®, before transfer to the plate (*see* **Note 10**).

3.2. High-Performance Liquid Chromatography (HPLC) of Peptide Extracts

1. Methanol extracts, or vacuum-concentrated methanol extracts remaining after direct analysis of perfusate, are diluted (at least 10-fold) with 0.1% TFA for HPLC separation (*see* **Note 11**). Tissue extracts can be suitably diluted by drawing into a polyethylene syringe pre-loaded with 0.1% TFA and by back-flushing into the sample tube.
2. The entire syringe contents (e.g. 5 mL) are then loaded onto the HPLC column via an equivalent-sized loop. Reversed-phase gradient HPLC may be conducted using any suitable system and column combination, although for best results, particularly when peptide amounts are minimal, a narrow bore or microbore column and detector with small volume flow cell should be used. For many peptides, it is important that the column packing material be 300 Å (*see* **Note 12**).
3. The column is eluted with a linear gradient of 5–60% acetonitrile/0.1% TFA over 55 min at a flow rate of 0.2 mL/min, with elution being monitored at 214 nm (*see* **Note 13**). Fractions (0.2 mL) are collected automatically at 1 min intervals or manually if collection of specific UV-absorbing peaks is required (*see* **Note 14**).
4. The HPLC fractions are concentrated to a small volume (~ 10–20 μL) using a vacuum centrifuge (*see* **Note 9**) and aliquots (e.g. 0.5–1 μL) are taken for MALDI-TOF analysis.

3.3. MALDI-TOF Mass Spectrometry

3.3.1. Positive Ion, Reflectron Mode Analysis

1. After the target plate has been inserted into the MALDI-TOF instrument, there is a short period of equilibration (5–10 min), whilst the vacuum is re-adjusted and the high tension (voltage) is turned on and stabilised. The samples are now ready for analysis.
2. The machine is calibrated and adjusted in positive ion/reflectron mode by targeting the laser at a calibration spot containing a range of pre-selected peptide standards. The peptide standards kits will contain details of the exact monoisotopic masses ($[M+H]^+$) for each standard peptide, and these details will be used to provide accurate calibration. Depending on the instrument, it will be necessary to adjust

e.g. laser intensity (usually arbitrary units), delay time (ns), accelerating voltage, guide wire and grid voltages in order to optimise peak shape (i.e. for peptide ions of interest) and resolution and signal-to-noise ratio. For accurate analysis, it is essential to achieve isotopic resolution for both standard and target peptides. Peptides larger than ca. 5000 may not be isotopically resolved, in which case average mass data will need to be used, although the majority of insect peptides are somewhat smaller in size.

3. Typically instrument settings for peptide analysis using the Voyager DE-STR would be accelerating voltage, 20 kV; grid voltage, 68%; guide wire, 0.001%; extraction delay, 150 ns; and acquisition range (effectively mass range, *see* **Note 15**) set to record between 500 and 5000 (max.). The laser fluence (= power) is manually set to a level that produces ions of sufficient intensity, with relatively low background. If the laser intensity is too high, the signal intensity may be saturated, in which case decreasing the laser intensity should optimise the signal.
4. A full spectrum can now be recorded for each intact tissue, tissue extract (i.e. perfusate), or HPLC fraction by moving the laser to the appropriate sample position and recording the summated ion signal(s) from several laser shots. There will be variation in the ion signal intensity at different positions over each target spot and it is best to summate the signals, both at each laser position (e.g. number of shots = 50) and at several positions over the sample spot (e.g. 5 positions × 50 shots). Typical spectra for tissue and tissue extracts from two insect species are shown in **Figs. 8.1** and **8.2**.
5. The ion spectra for each sample are examined for masses (m/z) that may correspond to known or suspected peptides (*see* **Note 16**).
6. For each sequence, the monoisotopic protonated mass ($[M+H]^+$) should be recorded, which can then be matched against the observed experimental masses (m/z) (*see* **Note 17**). For the whole-tissue extracts, there may be relatively few peptides that yield ions of sufficient intensity for PSD or MS^2 sequencing, although this will depend upon the nature and size of the tissue being examined as well as the degree of peptide concentration that can be achieved (actually minimisation of dilution) by keeping matrix/spot volumes as low as possible.

3.3.2. Post-source Decay and MS^2 Analysis

1. The MALDI-TOF or tandem-TOF instruments are generally pre-programmed to conduct PSD or CID MS^2 analysis,

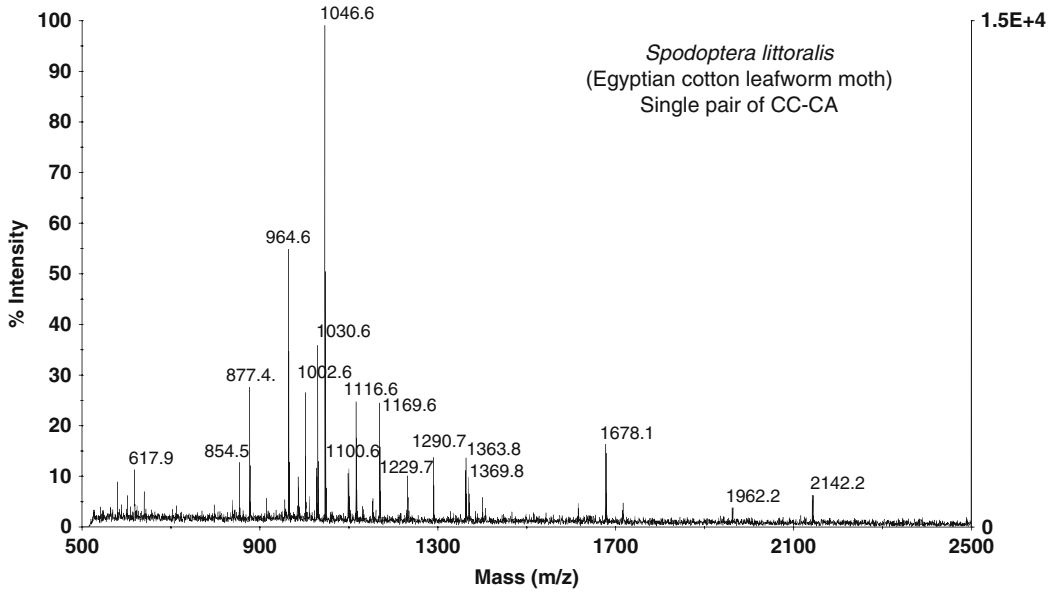


Fig. 8.1. MALDI-TOF mass spectrum in positive-ion reflector mode (Voyager DE-STR) of a methanol/matrix infusion (α-CHCA) of a single pair of retrocerebral glands (corpora allata–corpora cardiaca, CC-CA) extracted from behind the brain of adult Egyptian cotton leafworm moth, *Spodoptera littoralis*. Such extracts of single endocrine tissues will often provide ample signal intensity, such that peptide sequences can be either confirmed or identified de novo, using subsequent PSD analysis on the same sample spot.

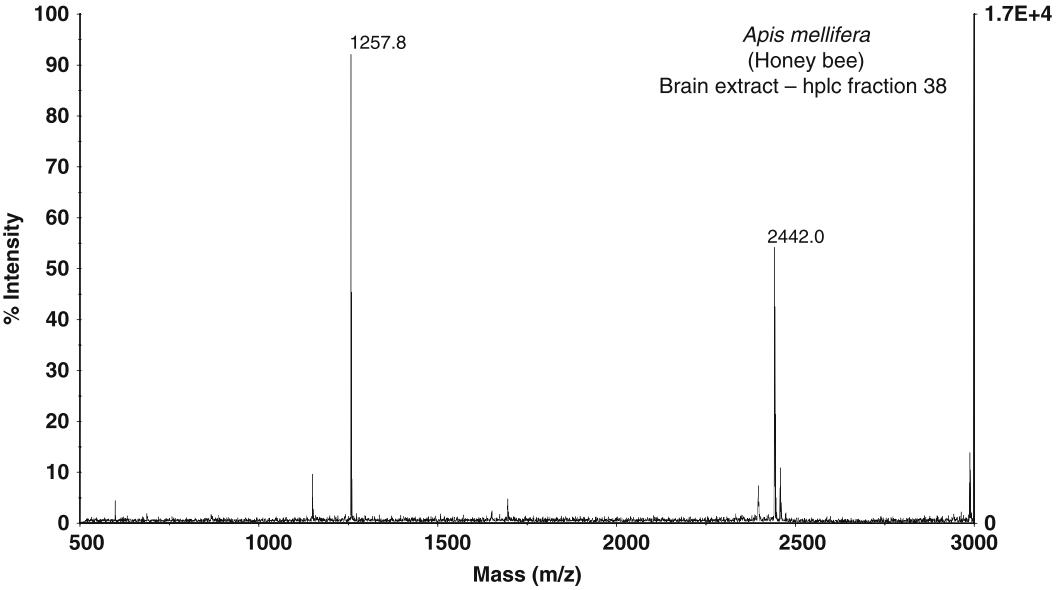


Fig. 8.2. MALDI-TOF mass spectrum of fraction 38 from a reversed-phase HPLC separation of a methanol brain extract of adult Honey bee (*Apis mellifera*). The fraction contains two different peptides, as shown by their mass/charge ratio (*m*/*z*), were subsequently fragmented and sequenced using post-source decay. The peptide at 1257.7 was found to be identical to a previously identified peptide (Leucomyosuppressin, pyroGlu-Asp-Val-Asp-His-Val-Phe-Leu-Arg-Leu-amide) found in cockroaches. The crude starting material contained at least 55 peptides, of which, after a single HPLC clean-up step, at least 14 could be assigned using PSD analysis and reference to genomic sequences.

which is done on the same sample as the reflectron analysis. The machine is first re-calibrated by opening an appropriate method file (pre-supplied) with which to accurately measure selected fragments of a known peptide. For the Applied Biosystems Voyager DE-STR, this is done using a PSD standard peptide, angiotensin I (*m/z* 1296.6853), the sequence and fragments of which are shown in **Fig. 8.3**.

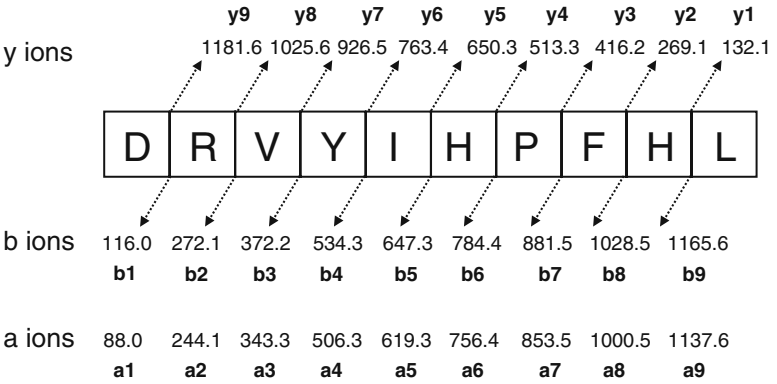


Fig. 8.3. Nomenclature and *m/z* values for theoretical PSD fragment ions that may be obtained from the vertebrate peptide Angiotensin I (Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu-OH, *m/z* = 1296.68 monoisotopic). The **b** and **y** ion pairs derive from fragmentation between the CO–NH bond, whereas the **a** and **x** ions (the latter not shown) derive from fragmentation between the CH–CO bond. The **b**, **y** and **a** ions, together with some internal fragment ions and immonium ions, will tend to predominate in the PSD spectra. Note that the sum of each pair of **y** and **b** ions is always equal to $[M+H]^+$ plus 1. This can be used as a starting point to identify the possible **y** and **b** ion pairings in an unknown sequence.

2. Once the instrument has been calibrated in the appropriate mode, the selected *m/z* peak value from the reflectron analysis for the peptide for which the information is required is entered (precursor ion selection), and in the case of the Voyager DE-STR a series of measurements (segments) are made in which stepwise adjustments to laser intensity and mirror (= reflectron) ratio enable the capture of a successive series of fragment ions. These "segments" are then electronically "stitched together" by the software, to produce a composite spectrum of all fragments, and consisting of a mixture of predominantly a-, b- and y-series ions, plus ammonia and water losses, internal fragments and immonium ions. A typical PSD spectrum for a known insect peptide is shown in **Fig. 8.4**. The challenge now is to identify the unknown peptide that has generated such a fragmentation spectrum.

3.3.3. Interpretation of MALDI-TOF Spectra

The number and variety of different peptides and neuropeptides, both within any individual insect or other invertebrate species and between different species, even from the same genera, is likely to

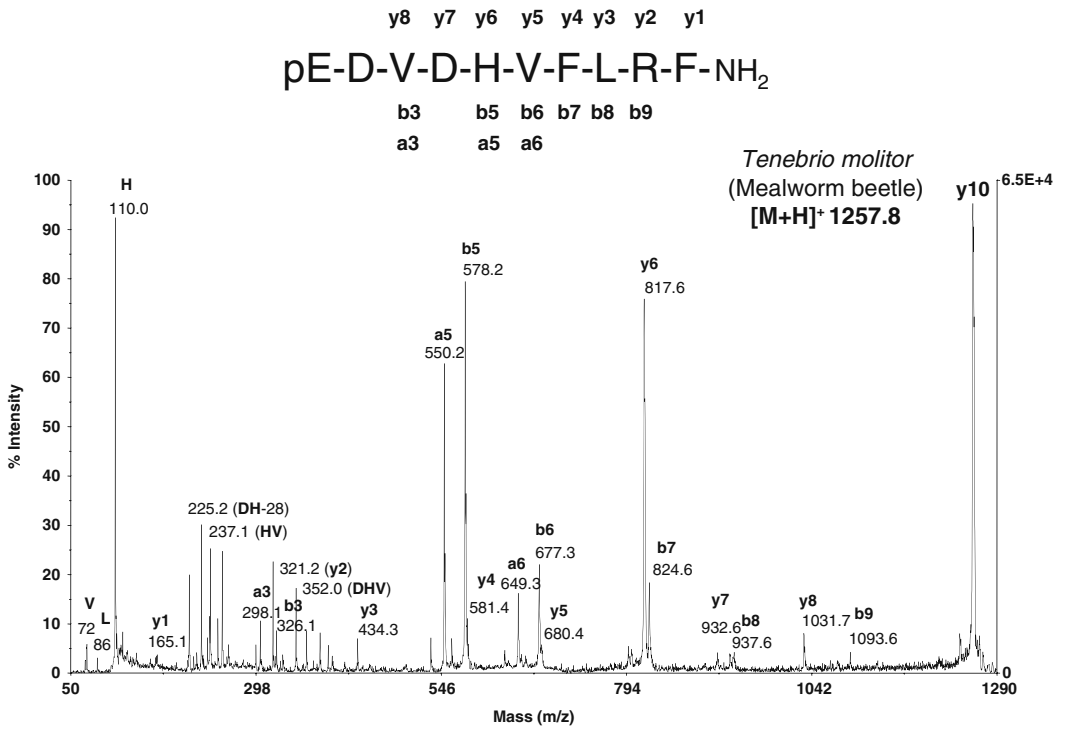


Fig. 8.4. Post-source decay MALDI-TOF mass spectrum of the ion at 1257.7 from a methanol infusion of adult male brains of the mealworm beetle, *Tenebrio molitor*. An almost complete series of y and b ions is observed, together with several a ions, internal sequences and immonium ions. The mass and the derived sequence of the peptide are consistent with the cockroach peptide leucomyosuppressin. (Re-drawn from (13) with permission from Elsevier Science).

be very large. For example, there are potentially at least 50 different "structural classes" or "types" of neuropeptides in any given species, based on known peptides, biological activities, and in several cases, association with known receptors, cell types and target tissues. Moreover, whilst some peptides appear to be highly conserved (just one, or very few types occurring across many species), others may exist in multiple types or homologous series even within the same species, and may vary considerably from species to species and within different genera, families and insect orders (e.g. the FGL-amide allatostatins (11)). All of this makes the identification of insect peptides an even greater challenge.

1. The first step is to attempt to provisionally identify any "known" peptides (either from the species under investigation or from other insect species) which may be giving rise to the observed precursor ion and associated fragment ion spectra. This is often quite difficult, and at the moment there is no comprehensive database of insect peptides other than may be gleaned from web-based sources, such as NCBI, Swiss-Prot etc., and from numerous individual publications.

2. A second step, therefore, is to compile a list of all known invertebrate peptide sequences, from as many sources as possible, and to assign mass and monoisotopic ion values to each sequence. This mass assignment can be done, either using the instrument manufacturer's software, if available, or by making use of web-based programmes, such as the Protein Prospector MS-Product module, developed by the University of California, San Francisco (http://prospector.ucsf.edu/). This programme is also used for generating the theoretical fragments of a peptide sequence, making it invaluable for PSD and CID interpretation.
3. In compiling the list of peptides and monoisotopic ion masses, it may also be useful to list the equivalent m/z values for sodiated and potassiated adducts, particularly for adipokinetic hormones.
4. The next task would be to identify potential sequences from genomic and cDNA (e.g. EST databases) sources. For this, a variety of publicly available web-based tools (e.g. the BLAST resources at NCBI) may be used to generate information on possible neuropeptide precursor genes. It will then be necessary to identify likely neuropeptide cleavage sites, from which the sequences of putative peptides can be deduced.
5. After this, it may be necessary to take account of any potential post-translational modifications (e.g. C-terminal pyroglutamyl formation, N-terminal amidation, cysteine cross-linking or cyclisation, tyrosine sulphation etc.), as well as partial processing, and even partial degradation.
6. Once any sequence has been putatively "identified", on the basis of monoisotopic mass alone, it is then necessary to generate a theoretical fragmentation fingerprint (e.g. using Protein Prospector MS-Product, or Manufacturer's software) and assessing whether a sufficient number of fragment ions and immonium ions are present in the spectrum of the unknown sample to make it a probable match. The more fragments that can be matched, the greater the likelihood that the identification is sound, although once a provisional identification has been made, additional reassurance can be gained by comparing the spectrum of the sample against that for a synthetic peptide of identical sequence.
8. Where no match is made to the initial precursor peak, it will then be necessary to rely on "de novo" sequencing, for which the quality of the PSD or CID spectrum is extremely important (*see* **Notes 18** and **19**).

4. Notes

1. Secondary solvent systems (e.g. acetonitrile/0.1% heptafluorobutyric acid) may be required to resolve peptides that are poorly resolved using only an acetonitrile/0.1% TFA gradient.
2. It is not necessary to weigh out the exact amount of matrix. The matrix should be prepared on the same day as a saturated solution and diluted, if necessary (with the acetonitrile/0.1% TFA, 50:50), to provide even crystals when dried down after mixing with sample and spotted on the target plate.
3. For best results, α-CHCA should be re-crystallised as follows: Dissolve 100 mg α-CHCA in 10 mL water and add conc. Ammonium hydroxide (ca. 150 μL) dropwise until almost all matrix has dissolved. Slowly add conc. HCl to the solution until the majority has precipitated (ca. pH 2). Collect precipitant by filtration or centrifugation, discard solution, and wash precipitant several times in 0.1 M HCl. Dry overnight in a vacuum dessicator. Dispense into aliquots (e.g. 3–10 mg) in microfuge tubes. Store in dark at –20°C.
4. The incisions and direction of entry will depend upon the particular tissue under investigation, the stage and species concerned and, to a certain extent, upon individual preference. Micro-dissecting pins are used to expose the target tissues to best advantage. Insect haemolymph may be viscous and semi-opaque in some species, and will obscure the intended target. This is best flushed away with fresh saline using a drawn-out glass pasteur bulb pipette. The composition of saline is not generally important, because rapid dissection will produce best results. The insects should be surface sterilised (e.g. with 80% alcohol) and rinsed with water (sterile) prior to dissection. It will also help to immobilise them by placing on ice for several minutes.
5. For optimum sensitivity, the volume of matrix droplet should be kept as low as possible and preferably applied using a nanolitre injector (3).
6. MALDI-TOF instruments are equipped with a means to visualise the laser beam hitting within the target spot. For example, the Applied Biosystems Voyager DE-STR has a video camera and monitor, whereas the Bruker Daltronics Ultraflex displays via a computer screen image. In either case, the laser beam can be targeted to different areas of the

sample spot, either manually or automatically, e.g. according to a predetermined pattern.

7. It is important to minimise the amount of saline that is transferred along with the tissue. The quantity of tissues required will depend upon the species, stage, tissue type and peptides under investigation; however, 5–50 tissues will generally prove adequate for the analysis of many of the more abundant peptides. The numbers of tissues required may need to be greater when using HPLC separation. The volume of methanol used is not important, although 20–50 μL would be quite adequate for 10–20 corpora cardiaca–corpora allata from a medium-sized insect.
8. There are several alternatives to the perfusion technique, if progressing to HPLC separation, which have been used by others for extraction and identification of insect peptides. For example, extraction in methanol/water/acetic acid (90:9:1), together with homogenisation and/or sonication, will extract most peptides and some small proteins, but will also co-extract many other small molecules and salts, which will need to be separated away by at least one HPLC step. Another alternative is to use distilled water (to burst cells) rather than organic solvent, although this may lead to indiscriminate proteolysis of extracted peptides.
9. Peptide losses can occur if samples are concentrated to dryness, it is much better to concentrate samples down to a small volume (e.g. 5–10 μL). Some peptides are also not easily solubilised after drying, in which case a few microlitres of 80% acetonitrile should be used to aid resolubilisation. It may be necessary to dry HPLC fractions completely, as each fraction will centrifugally vacuum evaporate at a slightly different rate because of the different organic solvent content.
10. Using parafilm M® helps prevent the droplets from spreading. It should make little difference whether the sample droplet is added to a matrix droplet or vice versa. This should be done one sample at a time, to avoid evaporation of the target droplet. Samples that contain a higher proportion of organic solvent may "spread" excessively on the parafilm or target plate, in which case adding sample to matrix may be the preferred option. Care should be taken to avoid touching the target plate with the pipette tip when spotting final sample/matrix mixture.
11. It is preferable to dilute samples directly to reduce the organic content to less than 10%, rather than by taking to dryness and re-dissolving. This helps to minimise peptide losses.

12. A set-up that has been used extensively in our laboratory for insect neuropeptide identification comprises a dual-pump programmable solvent module, coupled to a variable wavelength UV detector (0.5 mm flow cell), and fraction collector. The samples are loaded via a Rheodyne loop injector onto a Jupiter C18 10 μM 300 Å narrow-bore column (250 mm × 2.1 mm i.d.; Phenomenex, Macclesfield, UK) fitted with guard column (30 mm × 2.1 mm i.d.) of similar packing material.
13. HPLC-grade acetonitrile, TFA and ultra-pure water are essential when monitoring for UV absorption at 214 nm. Collection tubes and transfer pipettes should also be kept scrupulously clean and not handled without disposable gloves.
14. The connections between pumps, injector, column, detector and fraction collector should be kept as small as practicable (length and diameter), consistent with optimal flow rates and pressures and wherever possible zero-dead volume HPLC unions should be used. Columns should not be used for, or calibrated with, any synthetic peptides that are likely to be encountered in the sample. Instead, a non-invertebrate, synthetic peptide should be chosen instead.
15. MALDI-TOF mass spectrometers may be set to run in either positive- or negative-ion mode. For most peptide analysis it is better to use the positive-ion mode, in which case the ions will be predominantly singly protonated ($[M+H]^+$), although at higher laser fluences doubly charged ions may be observed ($[M+H]^{2+}$). In the first instance, the m/z ratio will be equal to the peptide mass + 1; in the second case, it will be half that value. If there is salt in the sample, one may also observe catonated ions $[M+Na]^+$ and $[M+K]^+$ that will be, respectively, + 22 and +38 m/z units larger than the protonated ions. For a few peptides, most notably the insect adipokinetic hormones, the predominant ions observed are typically sodiated $[M+Na]^+$ and potassiated $[M+K]^+$ under MALDI-TOF conditions, with protonated ions generally not being observed. A mass difference of 16 (38 minus 22) between two adjacent peaks is generally indicative of a sodiated/potassiated pair.
16. This will be done by reference to previously characterised peptides from the species being studied (1), by comparison with closely (and sometimes distantly) related species, and by reference to genomic and EST databases (both for same or similar species) (12–14). In each case, a list of all potential peptides (e.g. all known lepidopteran peptides)

should be drawn up, including details of any known or suspected post-translational modifications and together with sequences of any predicted or inferred precursor or degradation products.

17. Ion peaks in MALDI-TOF instruments are visualised and plotted as intensity, on the *y*-axis, versus mass/charge ratio (*m/z*) on the *x*-axis. Intensity is normally recorded in arbitrary units, and is dependent upon the type of digitiser used to convert ions detected into electronic signal. The *m/z* range can be set anywhere between 0 and 300,000 (depending on the machine), but for peptide analysis (in reflectron mode) this is optimally restricted to between 500 and 5000 *m/z*. Depending on the matrix used, there may be considerable interference from matrix ions and other non-peptide small molecules in the range below ca. 500 *m/z*, whilst peptides/proteins producing ions greater than 5000 *m/z* will not be isotopically resolved and peaks will be broader.
18. It will be necessary to follow the steps outlined in the tutorials and guidelines referred to earlier (8–10) and even then there may be too many ambiguities that will need to be resolved by further processing (e.g. derivatisation, deuterisation, selected cleavage).
19. A limitation in MALDI-TOF PSD sequencing is that the internal energies of the $[M+H]^+$ ions are generally insufficient to yield side-chain cleavages that would allow the discrimination of leucine from iso-leucine.

References

1. Audsley, N. and Weaver, R.J. (2003) A comparison of the neuropeptides from the retrocerebral complex of adult male and female *Manduca sexta* using MALDI-TOF mass spectrometry. *Regul. Pept.* **116**, 127–137.
2. Baggerman, G., Clynen, E., Huybrechts, J., Verleyen, P., Clerens, S., De Loof, A., and Schoofs, L. (2003) Peptide profiling of a single *Locusta migratoria* corpus cardiacum by nano-LC tandem mass spectrometry. *Peptides* **24**, 1476–1485.
3. Predel, R. (2001) Peptidergic neurohemal system of an insect: Mass spectrometric morphology. *J. Comp. Neurol.* **436**, 363–375.
4. Neupert, S. and Predel, R. (2005) Mass spectrometric analysis of single identified neurons of an insect. *Biochem. Biophys. Res. Commun.* **327**, 640–645.
5. Hillenkamp, F. and Karas, M. (2007) The MALDI process and method. in, MALDI MS, A Practical Guide to Instrumentation, Methods and Applications (Hillenkamp, F., Peter-Katalinić, J., eds.), Wiley-VCH, Weinheim, pp. 1–28.
6. Spengler, B., Kirsch, D., Kaufmann, R., and Jaeger, E. (1992) Peptide sequencing by matrix-assisted laser-desorption mass-spectrometry. *Rapid Commun. Mass Spectrom.* **6**, 105–108.
7. O'Connor, P.B. and Hillenkamp, F. (2007) MALDI mass spectrometry instrumentation. in MALDI MS, A Practical Guide to Instrumentation, Methods and Applications (Hillenkamp, F., Peter-Katalinić, J., eds.) Wiley-VCH, Weinheim, pp. 29–82.
8. Spengler, B. (1997) Post-source decay analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules. *J. Mass Spectrom.* **32**, 1019–1036.

9. Charaund, P., Luetzenkirchen, F., and Spengler, B. (1999) Peptide and protein identification by matrix-assisted laser desorption ionization (MALDI) and MALDI-post-source decay time-of-flight mass spectrometry. *J. Am. Soc. Mass Spectrom.* **10**, 91–103.
10. http://www.ionsource/tutorial/DeNovo/DeNovoTOC.htm
11. Hult, E.F., Weadick, C.J., Chang, B.S.W., and Tobe, S.S. (2008) Reconstruction of ancestral FGLamide-type insect allatostatins: A novel approach to the study of allatostatin function and evolution. *J. Insect Physiol.* **54**, 959–968.
12. Audsley, N. and Weaver, R.J. (2006) Analysis of the peptides in the brain and copora cardiaca–corpora allata of the honey bee, *Apis mellifera* using MALDI-TOF mass spectrometry. *Peptides* **27**, 512–520.
13. Weaver, R.J. and Audsley, N. (2008) Neuropeptides of the beetle, *Tenebrio molitor* identified using MALDI-TOF mass spectrometry and deduced sequences from the *Tribolium castaneum* genome. *Peptides* **29**, 168–178.
14. Gäde, G., Marco, H., Šimek, P., Audsley, N., Clark, K.D., and Weaver, R.J. (2008) Predicted versus expressed adipokinetic hormones, and other small peptides from the corpus cardiacum–corpus allatum: a case study with beetles and moths. *Peptides* **29**, 1124–1139.

Chapter 9

Direct MALDI-TOF Mass Spectrometric Peptide Profiling of Neuroendocrine Tissue of *Drosophila*

Christian Wegener, Susanne Neupert, and Reinhard Predel

Abstract

Direct MALDI-TOF mass spectrometric peptide profiling is increasingly used to analyze the peptide complement in the nervous system of a variety of invertebrate animals, from leech to *Aplysia* and many arthropod species, especially insects and crustaceans. Proper sample preparation is often the most crucial step to obtain the necessary data. Here, we describe protocols for the use of MALDI-TOF mass spectrometry to directly analyze the peptidome of neuroendocrine tissues of insects, particularly *Drosophila melanogaster*, by MALDI-TOF MS.

Key words: Peptidomics, neuropeptides, post-translational processing, *Drosophila*, perisympathetic organs (PSOs), corpora cardiaca, ring gland, MALDI-TOF tandem mass spectrometry, insects.

1. Introduction

Direct peptide profiling is related to MALDI-TOF imaging, but is carried out on whole organs or tissues, rather than on tissue sections. Compared to LC/MS analyses, it holds several advantages: (i) it is quick, (ii) it leads to a selective extraction of neuropeptides and peptide hormones, (iii) it does not need expensive equipment besides the mass spectrometer, (iv) it can be performed on tissue from single animals, and (v) it also gives an idea about the peptide location in tissues.

The first comprehensive peptidomic analysis by direct peptide profiling was performed in the American cockroach, *P*eriplaneta

All authors contributed equally.

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_9,

americana (1). This study demonstrated the advantage of profiling neuroendocrine tissue of insects, particularly neurohemal organs. Neurohemal organs such as the corpora cardiaca (CC), thoracic perisympathetic organs (tPSOs), and abdominal perisympathetic organs (aPSOs) store large amounts of peptide hormones. The profiling of these tissues is definitely the easiest and fastest approach to analyze the species-specific composition of many insect neuropeptide families such as CAPA-peptides (accumulated in aPSOs; (2)), extended FMRFamides (accumulated in tPSOs; (3)), FXPRLamides (projection area of the nervi corporis allati-2; (1–5)), and brain peptides such as corazonin, allatostatin-A, allatostatin-C, myosuppressin (accumulated in the CC; *see* (1, 6)). This approach was also used to analyze the neuroendocrine system of adults (7) and larvae (8) of the fruitfly *Drosophila melanogaster*.

Due to its genetic amenability, *Drosophila* is a perfect organism to investigate molecular and cellular aspects of neuropeptides. A potential problem, however, is its small size: the whole brain of an adult fruit fly is about the size of a single bag cell neuron of the sea slug, *Aplysia*. In the last few years, we have optimized direct peptide profiling protocols for neurohemal organs of *Drosophila* to study the distribution and post-translational processing of neuropeptides. De novo peptide sequencing by direct MALDI-TOF-TOF peptide profiling of neuroendocrine tissue, which has been achieved for larger flies (9), is usually not necessary for *Drosophila* since the genomes of a number of *Drosophila* species are fully sequenced. Thus, a partial fragmentation is sufficient to unequivocally identify fruitfly neuropeptides including those with unpredicted cleavage sites, or to assign mass-identical neuropeptides such as HUGIN-pyrokinin (pyrokinin-2) and DTK-2 (*Drosophila* tachykinin 2). In this section, we describe our preferred *Drosophila* protocols which are generally suitable for other insects as well.

2. Materials

2.1. Sample Preparation

1. Dissecting saline: 128 mM NaCl, 2 mM KCl, 1.8 mM $CaCl_2$, 4 mM $MgCl_2$, 36 mM sucrose, 5 mM HEPES, pH 7.1; or 80 mM NaCl, 5 mM KCl, 1.5 mM $CaCl_2$, 20 mM $MgCl_2$, 10 mM $NaHCO_3$, 5 mM trehalose, 115 mM sucrose, 5 mM HEPES, pH 7.2 (*see* **Note 1).**
2. A pair of fine forceps (e.g., sharpened Dumont No. 5)
3. Ultra-fine spring or clipper scissors (Fine Science Tools GmbH, Heidelberg, Germany)

4. Tungsten micro-needles (custom-made from electrolytically sharpened tungsten wire)
5. Pulled uncoated glass capillaries (e.g., Hilgenberg GmbH) fitted to a tube with mouthpiece (e.g., a sterile pipette tip)
6. Sylgard-coated preparation dish
7. Dissecting microscope with high magnification

2.2. MALDI-TOF Matrix Application

1. Re-crystallized α-cyano-4-hydroxycinnamic acid (CHCA, Sigma-Aldrich)
2. Methanol (MeOH), Trifluoroacetic acid (TFA), Acetonitrile (ACN), all HPLC grade
3. Water, double-distilled or HPLC-grade
4. Nanoliter applicator (we use e.g. a manual oocyte injector Drummond Digital, Broomall, PA, USA, or the nanoliter injector from World Precision Instruments, Berlin, Germany)

2.3. MALDI-TOF Mass Spectrometry

1. MALDI target plates (we typically use simple stainless steel target plates)
2. MALDI-TOF mass spectrometer, e.g., Voyager-DE STR biospectrometry workstation (Applied Biosystem), or UltraFlex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics), or 4700/4800 MALDI TOF/TOF ™ Analyzer (Applied Biosystem)
3. *FlexAnalysis* software (Bruker Daltonics), Voyager *DataExplorer*™ 2.4 (Applied Biosystem) or similar.
4. *ProteinProspector* – Proteomics tools for mining sequence databases in conjunction with mass spectrometry experiments (http://prospector.ucsf.edu)
5. NeuroPred (http://neuroproteomics.scs.uiuc.edu/neuropred.html) (*see* **Note 2**).
6. NCBI/BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (*see* **Note 3**).

3. Methods (*see* Note 4)

3.1. Sample Preparation

3.1.1. Preparation of Larval Ring Glands

In larvae of cyclorraphous flies, the corpora cardiaca (CC), corpora allata (CA), and prothoracic gland are fused into the so-called ring gland. To dissect the ring gland, first dissect the central nervous system in saline (*see* **Note 5**). A very useful online demonstration of the preparation is given by Kei Ito (http://jfly.iam.u-tokyo.ac.jp/html/movie/index.html); the ring gland should be visible after that procedure.

Method A

1. Carefully remove all imaginal disks attached to the brain. Punch out the ring gland with a pulled-out glass capillary, the sharp tip of which is broken to a convenient diameter.
2. Suck the isolated ring gland into the glass capillary and transfer to a MALDI target.
3. Blow out the gland onto the MALDI target. Remove all carried-over saline with a non-abrasive tissue (e.g., KimWipes) or aspirate off with a glass capillary; let dry.

Method B

1. Fix the ring gland with a fine forceps using one of the two nerves leaving the ring gland into the periphery. Cut off the connection between the ring gland and the brain with a clipper scissor. Imaginal disks need not to be removed.
2. Use a stainless steel insect pin (size 0 or 00) mounted on a pin holder to transfer the gland to a MALDI target (*see* **Note 6**).
3. Dip the ring gland briefly in a small drop of distilled water to remove salt contamination and pull the gland out of the drop; disrupt the tissue, allow to dry. Alternatively, the drop of water can be aspirated away with a pulled-out glass capillary, while the gland is fixed on the target plate with the insect pin.

3.1.2. Preparation of the Corpora Cardiaca of Adults

In adult cyclorrhaphous flies, the ring gland has transformed into a well-separated CA and the CC which are fused with the hypocerebral ganglion; the prothoracic gland has disappeared (10).

Method A

1. Cut off the legs, wings, and abdomen.
2. Fix the fly with a fine insect needle through the head.
3. Open the thorax from the dorsal side, and carefully remove all muscle tissue until the esophagus/gut is visible.
4. Carefully remove all tissue and the sternites below the gut/esophagus (*see* **Note 7**).
5. Punch out the CC/hypocerebral ganglion with a glass capillary and transfer as in **Section 3.1.1**.

Method B

1. Cut off the legs and wings.
2. Fix the fly in lateral position with a fine insect needle through the head.
3. Use two forceps for pulling the thorax apart from the head to disrupt the cervical (neck) membrane (*see* **Note 8**).

4. Fix the thorax with an insect pin and separate, from anterior to posterior, the nervus corporis cardiaci and the CC/hypocerebral ganglion with attached nerves from the gut/aorta by using micro-needles.
5. Transfer the isolated tissue by one of the techniques described in **Section 3.1.1**.

3.1.3. Preparation of Trachea with Attached Peritracheal Cells

1. Fix a larva with fine insect needles through the mouthparts and the posterior end.
2. Hold the dorsal cuticle with fine forceps above the heart and lift gently.
3. Cut a hole into the dorsal cuticle directly below the forceps.
4. Widen this hole by carefully cutting along the dorsal midline anterior to the mouthparts and posterior to the spiracles.
5. Remove gut and fat body.
6. Hold the main trachea at the posterior spiracle and gently pull it off from the body wall.
7. Transfer the isolated trachea by one of the techniques described in **Section 3.1.1** (*see* **Note 9**).

3.1.4. Preparation of Perisympathetic Organs (PSOs) of Larvae

1. Dissect the larval CNS in saline as described in **Section 3.1.1**. Fix the CNS in a lateral position in a preparation dish with black background.
2. In the lateral position, the abdominal median/transverse nerves are well visible; the first three nerves contain neurosecretions from the ventral median neurosecretory cells (V_a-neurons) of the abdominal ventral nerve cord (*capa*-neurons). Fix one of these nerves with a forceps and use a clipper scissor to cut the nerve off proximally.
3. Transfer the isolated nerve(s) (i.e., the abdominal PSO) by one of the techniques described in **Section 3.1.1** (*see* **Note 10**).

3.1.5. Preparation of Perisympathetic Organs of Adults

Adult cyclorraphous flies do not possess PSOs located outside the central nervous system. Instead, release sites of neurosecretions from the ventral nerve cord branch out and intermingle directly beneath the dorsal ganglionic sheath (11, 12).

1. Cut off the legs and wings.
2. Fix the fly with insect pins through head and end of abdomen.
3. Open the thorax/abdomen with a dorso-median incision; move the thoracic muscles sideways until the fused ventral nerve cord is visible.

4. Use a very fine clipper scissor to make a posterior cut across the dorsal ganglionic sheath.
5. Fix the incised ganglionic sheath with fine forceps and dissect the complete dorsal ganglionic sheath using the clipper scissor or (preferably) micro-needles.
6. Before cutting off the anterior attachment, remove pieces of the nervous system which are attached to the ventral surface of the dorsal ganglionic sheath, using a micro-needle (*see* **Note 11**).
7. Transfer the isolated ganglionic sheath by one of the techniques described in **Section 3.1.1** (*see* **Note 12**).

3.2. MALDI-TOF Matrix Application

Method A

1. If the transfer of the sample was made with a capillary, add a small drop of ice-cold water containing 0.1–0.5% TFA onto the dried tissue, and immediately remove it with a glass capillary or a non-abrasive tissue (e.g., KimWipes). This step is important to remove salts that might interfere with the crystallization of the matrix and the ionization process (*see* **Note 13**).

Method B

1. If the transfer of the sample was made with an insect pin into a drop of water, just proceed with the next step.
2. Apply a small amount of matrix solution (50–100 nl) on top of the tissue, let dry (**Fig. 9.1**) (*see* **Note 14**).
3. Apply a drop of Aqua bidest or 0.1% TFA onto the dried sample which has to be removed after few seconds, let dry (*see* **Note 15**).

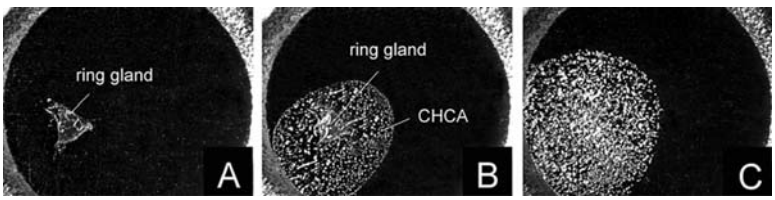


Fig. 9.1. Application of matrix crystals (CHCA) onto a *Drosophila* ring gland preparation. The black circle represents a single sample spot on a stainless steel MALDI target. (**A**) Ring gland after washing off the insect saline as described in **Section 3.1.2**, (**B**) Matrix crystallization after application of a suitable concentration of CHCA. The ring gland is still visible. (**C**) Matrix crystallization after application of an overdose of CHCA. The ring gland is hidden in the matrix crystals.

3.3. MALDI-TOF Mass Spectrometry

1. Analyze the sample in the reflector mode as recommended for peptide samples. Limit the amount of laser shots for each spectrum acquisition to 20–50, accumulate the spectrum, and use the same spot again. If the ion intensity has

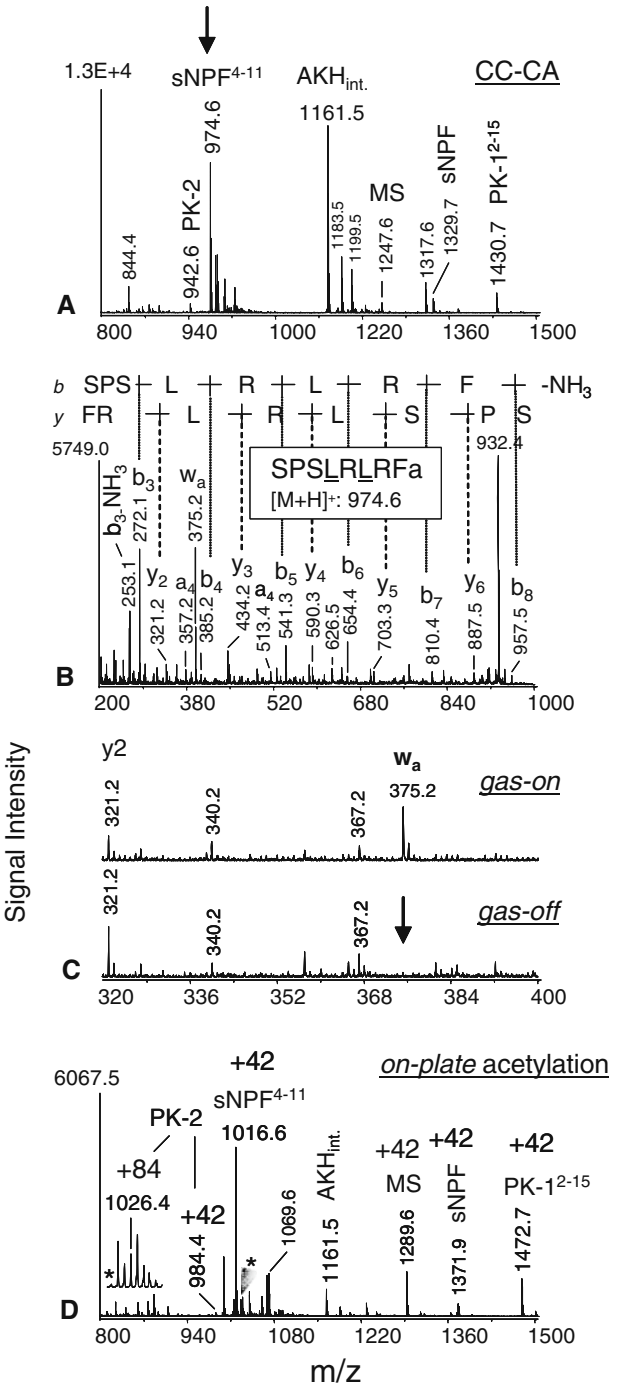


Fig. 9.2. MALDI-TOF mass spectra obtained from a single corpora cardiaca (CC) preparation of *D. melanogaster.* (**A**) Mass fingerprint, representing adipokinetic hormone (AKH int. = intermediate product), pyrokinins (PK), myosuppressin (MS), and short neuropeptide F (sNPF). (**B**) MALDI-TOF/TOF tandem mass spectrum of the peptide at $[M+H]^+$: 974.6 Da. (**C**) MALDI-TOF/TOF tandem mass spectrum of the peptide at $[M+H]^+$: 974.6 Da under conditions of high collision energy (gas on) showing side-chain fragments typical of leucine. (**D**) Mass spectrum of the same preparation after on-plate acetylation.

markedly decreased, discard the last spectrum and move on to a new spot. Move the laser to different spots on and directly around the tissue.

2. For fragmentation, find a spot with reasonable ion intensity. For that, use a low number of shots for each spectrum acquisition to minimize peptide loss prior to fragmentation. Select the parent ion for fragmentation and begin fragmentation (*see* **Notes 16** and **17**)
3. Samples with peptides that contain Lys/Gln ambiguities can be analyzed again after dissolving the respective preparations in acetic anhydride (2:1 methanol/acetic anhydride) which results in acetylation of the ε-amino group of Lys (*see* **Fig. 9.2D**).

3.4. Data Analysis

1. Analyze the data with the appropriate software (*see* **Section 2.3**). Compare the masses of observed peaks with the theoretical masses of known or predicted *Drosophila* peptides (*see* **Note 18** and **19**). Post-translational peptide processing can be predicted by the web-based program *NeuroPred*.
2. To obtain an idea about the relative abundances of the peptides in native samples, the MALDI-behavior of the peptides can be studied by using synthetic peptides under identical conditions (*see* **Note 20**).
3. If short peptide sequence stretches have been obtained by fragmentation, a BLAST search should be performed. We recommend using high expectation (>1) and the PAM30 matrix.

4. Notes

1. More information on the choice of buffers can be found in earlier publications (13, 14).
2. NeuroPred is a tool for predicting cleavage sites in neuropeptide precursors and provides the masses of the resulting peptides (15).
3. **B**asic **L**ocal **A**lignment **S**earch **T**ool (BLAST) is a tool for searching and comparing primary biological sequence information, such as amino acid sequences.
4. In some cases, different approaches of sample preparation for MALDI-TOF mass spectrometry are suggested (*see* e.g. **Section 3.1.1**, A, B). The preferred technique depends on the specific preparation skills available. In cases where one of these procedures does not work, try the other approach.

5. For the preparation steps, it is crucial to work as quickly as possible. Substantial peptide loss can occur once the tissue is separated from the nervous system.
6. The transfer must be quick once the pin with the attached ring gland is out of the saline. In the drop of water, the ring gland comes off the pin easily if it is not completely dry.
7. You should end up with an intact head to which only the gut and ventral ganglion is attached. Just before the proventriculus, the fused CC/hypocerebral ganglion can be recognized by its bluish tinge (Tyndall effect) dorsal to the esophagus (for morphology *see* (10)).
8. The gut and aorta (still beating!) should be visible.
9. The preparation contains the peritracheal cells which produce neuropeptides of the *eth* (eclosion triggering hormone, ETH) gene. The cells are located at the obtuse angle of the primary tracheal branches.
10. The bulb-like thoracic PSOs of larvae, which contain neurosecretions from the T_v-neurons of the thoracic ventral nerve cord (extended FMRFamide neurons), are too small to be dissected manually. Information regarding the processing of the FMRFamide prepropeptide in larvae can be obtained by profiling pieces of the CNS, or by profiling of the appropriate portion of the dorsal neural sheath as in **Section 3.1.5**.
11. We suggest to first test this procedure using larger flies with similar anatomy, e.g., blowflies.
12. During or following the dissection, the dorsal ganglionic sheath can be divided into an anterior and posterior part to partially separate the neurosecretions from the thorax and abdomen.
13. In preparations containing a high amount of salts, peptides can be seen to form alkali ion adducts ($[M+Na]^+$ or $[M+K]^+$). Sometimes, these alkali adducts can be useful to separate peptide from non-peptide mass peaks, and may increase ionization of peptides that are otherwise hard to protonate. The adipokinetic hormone (AKH), for example, does not typically occur as a $[M+H]^+$ adduct, but as alkali adducts in MALDI-TOF mass spectra.
14. We use CHCA dissolved in 30% MeOH/30%EtOH/40% Aqua bidest./0.1%TFA, or 60% ACN/40%Aqua bidest. /0.1%TFA. Always make fresh matrix and briefly spin down non-dissolved matrix crystals in a centrifuge if saturated CHCA is used. The use of a nanoliter injector or similar equipment ensures the desired long application time combined with a low amount of applied matrix solution. For

larger peptides (>3 kDa), DHB in 60%MeOH/40% Aqua bidest usually works better. Mass spectrometers with a laser operating at a wavelength of 355 nm (e.g., Applied Biosystems ABI 4800 TOF/TOF mass spectrometer), however, seem to be less sensitive when using DHB as matrix.

15. Keep the sample dark and dry. In that way, it is usually sufficiently stable for up to a week or more.
16. With low amounts of material, best results are usually obtained when the collision gas source is turned on.
17. Some MALDI-TOF mass spectrometers (e.g., Applied Biosystems ABI 4700/4800 TOF/TOF mass spectrometer) allow the unambiguous assignment of isomeric leucine and isoleucine amino acids even when profiling small insect samples (16). For that, retake the spectra under conditions of high gas pressure (*see* **Fig. 9.2**).
18. On request, a mass list can be obtained from the authors for following species: *D. melanogaster*, *D. virilis*, *D. pseudoobscura*, *D. sechellia*, *D. mojavensis*.
19. Data bank searches based on peptide mass peaks (such as Mascot) do not normally result in peptide identification due to the low scores obtainable with short peptide sequences or single masses. If used, make sure that the peptide modifications known to occur in insects are recognized (*see* **Table 9.1**).

Table 9.1
Mass changes due to post-translational modifications (PTMs) of insect neuropeptides

PTM	Mass difference (Da)
C-terminal amidation	–1
Disulfide bridge	–2
Methylation	+14
Oxidation	+16
Pyroglutamic acid formation	–17
Sulfation[1]/phoshorylation	+80

[1]In MALDI-TOF mass spectrometry, sulfation is only detectable in the negative mode.

20. Relative ion intensities of neuropeptides in mass spectra from neuroendocrine tissue are usually quite reproducible, but the ion intensities will depend on the peptide sequences as well as on the amounts.

Acknowledgments

Original work was supported by the Deutsche Forschungsgemeinschaft (DFG, PR 595/1-1. . .7, 6-1. . .4; WE 2652/2), Peter und Traudl Engelhorn Stiftung (to SN), and the Fonds der Chemischen Industrie (FCI, to CW). We like to thank William K. Russell (College Station, TX), Jörg Kahnt, RK Thauer (Max-Planck-Institute of Terrestrial Microbiology, Marburg, Germany) and Anton Gorbashov (Marburg, Germany) for supporting this study.

References

1. Predel, R. (2001) Peptidergic neurohemal system of an insect: mass spectrometric morphology. *J. Comp. Neurol.* **436**, 363–375.
2. Predel, R. and Wegener, C. (2006) Biology of the CAPA peptides in insects. *Cell. Mol. Life Sci.* **63**, 2477–2490.
3. Predel, R., Neupert, S., Wicher, D., Gundel, M., Roth, S., and Derst, C. (2004) Unique accumulation of neuropeptides in an insect: FMRFamide related peptides in the cockroach, *Periplaneta americana*. *Eur. J. Neurosci.* **20**, 1499–1513.
4. Predel, R., Eckert, M., Pollák, E., Molnár, L., Scheibner, O., and Neupert, S. (2007) Peptidomics of identified neurons demonstrates a highly differentiated expression pattern of FXPRLamides in the neuroendocrine system of an insect. *J. Comp. Neurol.* **500**, 498–512.
5. Clynen, E., Baggerman, G., Huybrechts, J., Vanden Bosch, L., De Loof, A., and Schoofs, L. (2003) Peptidomics of the locust *corpora allata*: identification of novel pyrokinins (-FXPRLamides). *Peptides* **24**, 1493–1500.
6. Baggerman, G., Clynen, E., Huybrechts, J., Verleyen, P., Clerens, S., De Loof, A., and Schoofs, L. (2003) Peptide profiling of a single *Locusta migratoria* corpus cardiacum by nano-LC tandem mass spectrometry. *Peptides* **24**, 1475–1485.
7. Predel, R., Wegener, C., Russell, W.K., Tichy, S.E., Russell, D.H., and Nachman, R.J. (2004) Peptidomics of CNS-associated neurohemal systems of adult *Drosophila melanogaster*: a mass spectrometric survey of peptides from individual flies. *J. Comp. Neurol.* **474**, 379–392.
8. Wegener, C., Reinl, T., Jänsch, L., and Predel, R. (2006) Direct mass spectrometric peptide profiling and fragmentation of larval peptide hormone release sites in *Drosophila melanogaster* reveals tagma-specific peptide expression and differential processing. *J. Neurochem.* **96**, 1362–1374.
9. Nachman, R.J., Russell, W.K., Coast, G.M., Russell, D.H., Miller, J.A., and Predel, R. (2006) Identification of PVK/CAP2b neuropeptides from single neurohemal organs of the stable fly and horn fly via MALDI-TOF/TOF tandem mass spectrometry. *Peptides* **27**, 521–526.
10. Shiga, S. (2003) Anatomy and functions of brain neurosecretory neurons in Diptera. *Microsc. Res. Tech.* **2003**, 114–131.
11. Truman, J.W. (1990) Metamorphosis of the central nervous system of *Drosophila*. *J. Neurobiol.* **21**, 1072–1084.
12. Santos, J.G., Pollák, E., Rexer, K.H., Molnár, L., and Wegener, C. (2006) Morphology and metamorphosis of the peptidergic Va neurons and the median nerve system of the fruit fly, *Drosophila melanogaster*. *Cell Tissue Res.* **326**, 187–199.
13. Jan, L.Y. and Jan, Y.N. (1976) Properties of the larval neuromuscular junction in *Drosophila melanogaster*. *J. Physiol.* **262**, 189–214.
14. Stewart, B.A., Atwood, H.L., Renger, J.J., Wang, J., and Wu, C.F. (1994) Improved stability of *Drosophila* larval neuromuscular preparations in haemolymph-like physiological solutions. *J. Comp. Physiol. A* **175**, 179–191.
15. Southey, B.R., Amare, A., Zimmerman, T.A., Rodriguez-Zas, S.L., and Sweedler, J.V. (2006) NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. *Nucleic Acids Res.* **34**, 267–272.
16. Nachman, R.J., Russell, W.K., Coast, G.M., Russell, D.H., and Predel, R. (2005) Mass spectrometric assignment of Leu/Ile in neuropeptides from single neurohemal organ preparations ofnsects. *Peptides* **26**, 2151–2156.

Chapter 10

Direct Peptide Profiling of Brain Tissue by MALDI-TOF Mass Spectrometry

Joachim Schachtner, Christian Wegener, Susanne Neupert, and Reinhard Predel

Abstract

Direct MALDI-TOF mass spectrometric peptide profiling is increasingly used to analyze the peptide complement in the nervous system of a variety of invertebrate animals from leech to *Aplysia* and many arthropod species, especially insects and crustaceans. Here, we describe a protocol for direct peptide profiling of defined areas of the central nervous system of insects. With this method, one can routinely and reliably obtain neuropeptide signatures of selected brain areas from various insects.

Key words: Peptidomics, neuropeptides, brain tissue profiling, antennal lobes, neuropil regions, MALDI-TOF mass spectrometry, insects.

1. Introduction

This chapter describes approaches to analyze the peptidome of the CNS by direct profiling of pieces of the brain or other parts of the CNS. The rule "the smaller (and less complex) the sample, the better the mass spectrum" particularly applies to this method of peptide profiling; i.e., large brain tissue samples usually yield low-quality spectra. Another drawback of direct peptide profiling of tissues of the CNS is the somewhat stochastically obtained ion intensities that depend on the quality of the preparation as well as the location of the laser beam. This makes (semi)quantifications even within single preparations difficult. In a number of cases, however, neither single-cell profiling nor analysis of the peptide complement of neurohemal organs gives the necessary

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_10,

information. Single-cell dissection is only possible if the respective neurons can be properly identified. The peptidome of interneurons which cannot be traced by backfilling with fluorescent dyes or expression of fluorescent marker proteins can only be obtained by dissecting a tissue area containing the target neuron. In most of these cases, the putative peptide-expressing neurons have been localized by immunocytochemistry in earlier experiments. A number of neuropeptides do not occur or are not accumulated in neurohemal organs. To study processing of the respective prepropeptides, profiling of CNS tissue may become necessary, as e.g. in the case of tachykinin-related peptides (TKRPs) (1).

The focus of this chapter is on the occurrence and functional significance of neuropeptides in specific neuropil areas such as the antennal lobes, which are the primary integration centers for odor information in the insect brain (2). The approaches described in this chapter refer mainly to methods developed for the analysis of this defined brain area (3), of which the neuroarchitecture is well described for several insect species (2). The function and identity of neuropeptides is poorly known in the insect antennal lobes and in their vertebrate counterpart, the olfactory bulbs. For this reason, a direct profiling protocol was developed which allows a fast and reliable detection of neuropeptides from not only antennal lobes but also other brain areas.

2. Materials

2.1. Sample Preparation

1. Dissecting saline: 128 mM NaCl 128, 2.7 mM KCl, 2 mM $CaCl_2$, 1.2 mM $NaHCO_3$, pH 7.25 (*see* **Note 1**).
2. A pair of fine forceps (e.g., sharpened Dumont No. 5), ultrafine spring or clipper scissors (Fine Science Tools GmbH, Heidelberg, Germany), tungsten micro-needles (custom-made from electrolytically sharpened tungsten wire).
3. Pulled uncoated glass capillaries (e.g., Hilgenberg GmbH) fitted to a tube with mouthpiece (e.g., a sterile pipette tip).
4. Sylgard-coated preparation dish.
5. Dissecting microscope with high magnification.

2.2. MALDI-TOF Matrix Application

1. Re-crystallized α-cyano-4-hydroxycinnamic acid (CHCA, Sigma-Aldrich).
2. Methanol (MeOH), ethanol (EtOH), trifluoroacetic acid (TFA), acetonitrile (ACN), all HPLC grade.
3. Water, double-distilled or HPLC-grade.

4. Nanoliter applicator (manual oocyte injector, Drummond Digital, Broomall, PA, USA) or a nanoliter injector (World Precision Instruments, Berlin, Germany) or equivalent.

2.3. MALDI-TOF Mass Spectrometry

1. MALDI target plates (we typically use simple stainless steel target plates).
2. MALDI-TOF mass spectrometer, e.g., Voyager-DE STR biospectrometry workstation (Applied Biosystem), or UltraFlex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics), or 4700/ 4800 MALDI TOF/TOF™ Analyzer (Applied Biosystem).
3. ***FlexAnalysis*** software (Bruker Daltonics), Voyager ***DataExplorer***™ 2.4 (Applied Biosystem) or similar.
4. ***ProteinProspector*** – Proteomics tools for mining sequence databases in conjunction with mass spectrometry experiments (http://prospector.ucsf.edu).
5. NeuroPred (http://neuroproteomics.scs.uiuc.edu/neuropred.html) (*see* **Note 2**).
6. NCBI/BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (*see* **Note 3**).

3. Methods

3.1. Sample Preparation

3.1.1. Dissection of Brain Tissue

1. Dissect the brain in cold insect saline.
2. Remove all attached tissues (e.g., muscles, fat body, trachea) and disrupt the ganglionic sheath. In case of larger insect brains (e.g., *Manduca sexta*, *Locusta migratoria*, *Periplaneta americana*) remove the complete ganglionic sheath.
3. Cut the brain into small pieces with fine insect pins.
4. Transfer the separated tissues directly into small drops of water on the MALDI sample plate using an insect pin mounted on a pin holder or a glass capillary connected to a mouthpiece (*see* **Note 4**).
5. After a few seconds, pull the tissue out of the drop of water, disrupt the tissue with the insect pin, and let dry.

3.1.2. Dissection of Defined and Morphologically Distinct Brain Regions

Some neuropil regions of the brain are easily identifiable and dissectable due to their compact appearance. Although they typically do not contain cell bodies, they contain a large amount of certain neuropeptides. This was shown for TKRPs in the tritocerebral glomeruli (1) and for the dorso-caudal neuropil in the terminal ganglion of many insects **(Fig. 10.1**). In other cases,

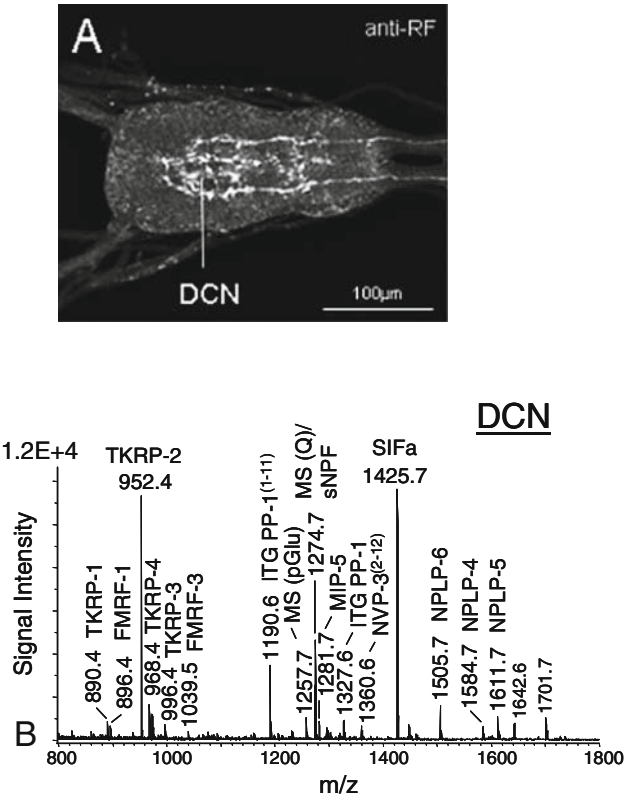


Fig. 10.1. MALDI-TOF mass spectrum obtained from a preparation of the dorso-caudal neuropil of the terminal ganglion of *Tribolium castaneum*. (**A**) Anti-RF immunostaining in the dorso-caudal neuropil of the terminal ganglion that exemplifies this part of the CNS. (**B**) The mass spectrum illustrates the high number of ion signals obtained, all of which represent neuropeptides (TKRP, tachykinin-related peptides; NPLP, neuropeptide-like precursor peptides; MIP, myoinhibitory peptides; MS, myosuppressin; sNPF, short neuropeptide F; FMRFamides; ITP, ion transport peptides; SIFamides, NVP containing peptides. Preparations of other regions of the terminal ganglion yield considerably less complex spectra.

such as the antennal lobes, the majority of the somata of the peptidergic neurons innervating the olfactory glomeruli (mostly local interneurons) are part of the antennal lobe compartments. However, the cell bodies of some large peptidergic centrifugal neurons are located in other brain areas outside the antennal lobes (2, 4) and project from there into the antennal lobe neuropil. The following dissection protocol refers to the antennal lobes.

1. Dissect the brains out of the head capsule and disassemble it into defined parts such as the antennal lobes (ALs) (*see* **Note 5**).
2. Suck the isolated piece of tissue into the tip of a pulled-out glass capillary.

3. Transfer the tissue to the MALDI target as explained in **Section 3.1.1**.

3.2. MALDI-TOF Matrix Application

Apply a small amount of matrix solution (saturated CHCA) on top of the dried tissue and allow to air-dry (*see* **Notes 6** and **7**). The matrix solvent should contain a higher methanol/ethanol/acetonitrile concentration than the matrix solutions which are described for profiling of neurohemal organs or single cells (*see* **Notes 8** to **10**).

3.3. MALDI-TOF Mass Spectrometry

1. Analyze the sample in the reflector mode as recommended for peptide samples. Limit the amount of laser shots for each spectrum acquisition to 20–50, accumulate the spectrum, and use the same spot again. If the ion intensity has markedly decreased, discard the last spectrum and move on to a new spot. Move the laser to different spots on and directly around the tissue. Start the analysis with relatively low laser energy and scan the sample for regions with good ion signals.
2. For fragment analyses, use these spots without further depleting the sample prior to the fragmentation. For that, use a low number of shots for each spectrum acquisition to minimize peptide loss prior to fragmentation. Select the parent ion for fragmentation and begin fragmentation. With low amounts of material, best results are usually obtained when the collision gas source is turned on.

4. Notes

1. Alternatively, use the saline which is used routinely, e.g., Weevers saline for *M. sexta* (5). Alternatively, phosphate-buffered saline may work as well.
2. NeuroPred is a tool for predicting cleavage sites in neuropeptide precursors and provides the masses of the resulting peptides (6)
3. **B**asic **L**ocal **A**lignment **S**earch **T**ool (BLAST) is a tool for searching and comparing primary biological sequence information, such as amino acid sequences.
4. Smaller and less complex tissue samples yield better mass spectra.
5. In small insects, including, e.g., *D. melanogaster*, *T. castaneum*, and *Aedes aegypti*, single ALs can be directly transferred to the target plate (**Fig. 10.2**). In larger insects like *M. sexta*, *Heliothis virescens*, or *A. mellifera*, whole neuropil areas are often too large to obtain good ion signals. Thus,

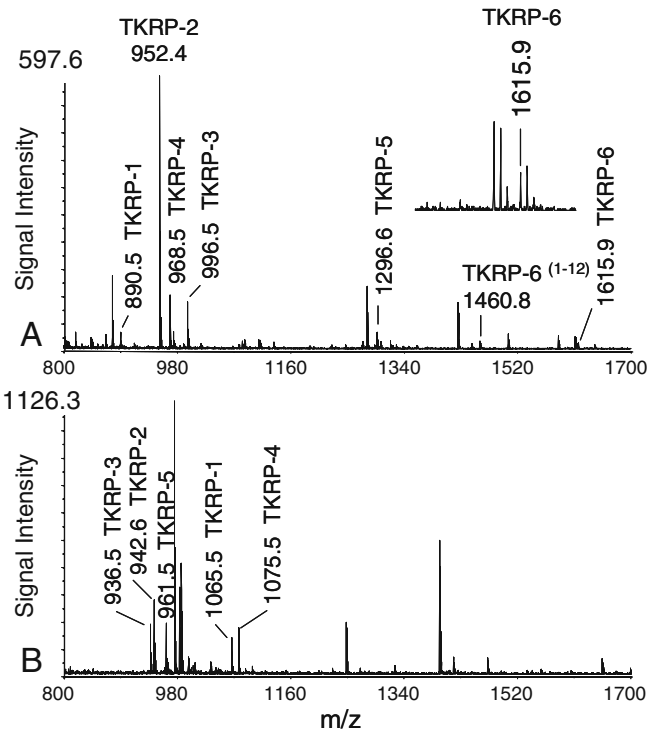


Fig. 10.2. MALDI-TOF mass spectra obtained from preparations of single antennal lobes. Tachykinin-related peptides (TKRPs) are labeled, respectively. (**A**) *Tribolium castaneum* and (**B**) *Drosophila melanogaster*. For information on *tkrp*-gene products in *T. castaneum* and *D. melanogaster see* 8, 9 and 10.

they have to be further broken down into smaller pieces. For example in *M. sexta*, a larger lateral and a smaller medial cell group can easily be distinguished on the surface of the isolated ALs. The cell groups can be selectively peeled off from the underlying central neuropil by using ultra-fine scissors and micro-needles (3). In *H. virescens* and *A. mellifera* whole ALs usually give reasonable signals (7).

6. The amount of the matrix solution necessary for optimal ion signals depends on the size of the tissue sample.
7. Always cover the tissue completely with matrix solution.
8. The high methanol/ethanol/acetonitrile concentration results in a better elution of analyte molecules out of the non-uniform tissues; it particularly improves the detection of neuropeptides from regions deep inside the tissue. Since the high methanol/ethanol/acetonitrile concentration causes a fast evaporation, it is best to use an injector for a constant supply with matrix solution over a period of 10–20 s.

9. The peptide concentration in brain tissue samples is usually much lower than the peptide concentration in neurohemal organs. For that reason, dilute the matrix solution if the mass spectra are disappointing (low signal-to-noise ratio).
10. In *M. sexta*, different solvent mixtures were used to yield good spectra from ALs of different developmental stages (3).

Acknowledgments

Original work was supported by the Deutsche Forschungsgemeinschaft (DFG, SCHA 678/3-3; PR 766/9-1 to R. Predel, J. Schachtner, and C. Wegener; WE 2652/2) and Peter und Traudl Engelhorn Stiftung (to SN). We like to thank William K. Russell (College Station, TX), Jörg Kahnt, and Lotte Søgaard-Andersen (Max-Planck-Institute of Terrestrial Microbiology, Marburg) for the use of the mass spectrometer and Sandra Utz (Marburg, Germany) for supporting this study.

References

1. Predel, R., Neupert, S., Roth, S., Derst, C., and Nässel, D. (2005) Tachykinin-related peptide precursors in two cockroach species: molecular cloning and peptide expression. *FEBS J.* **272**, 3365–3375.
2. Schachtner, J., Schmidt, M., and Homberg, U. (2005) Organization and evolutionary trends of primary olfactory brain centers in Tetraconata (Crustacea + Hexapoda). *Arthropod Struct. Dev.* **34**, 257–299.
3. Utz, S., Huetteroth, W., Predel, R., Wegener, C., Kahnt, J., and Schachtner, J. (2007) Direct peptide profiling of lateral cell groups of the antennal lobes of *Manduca sexta* reveals specific composition and changes in neuropeptide expression during development. *Dev. Neurobiol.* **67**, 764–777.
4. Utz, S., Huetteroth, W., Vömel, M., and Schachtner, J. (2008) Mas-allatotropin in the developing antennal lobe of the sphinx moth *Manduca sexta*: Distribution, time course, developmental regulation and colocalization with other neuropeptides. *Dev. Neurobiol.* **68**, 123–142.
5. Weevers, R.D. (1966) A lepidopteran saline: the effects of inorganic cation concentrations on sensory reflex and motor responses in a herbivorous insect. *J. Exp. Biol.* **44**, 163–176.
6. Southey, B.R., Amare, A., Zimmerman, T.A., Rodriguez-Zas, S.L., and Sweedler, J.V. (2006) NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. *Nucleic Acids Res.* **34**, 267–272.
7. Berg, B.G., Schachtner, J., Utz, S., and Homberg, U. (2007) Distribution of neuropeptides in the primary olfactory centre of the heliothine moth *Heliothis virescens*. *Cell Tissue Res.* **327**, 385–398.
8. Li, B., Predel, R., Neupert, S., Hauser, F., Tanaka, Y., Verleyen, P., Cazzamali, G., Williamson, M., Schoofs, L., Schachtner, J., Grimmelikhuijzen, C., and Park, Y. (2008) Genomics, transcriptomics, peptidomics, and evolution of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. *Genome Res.* **18**, 113–122.
9. Siviter, R.J., Coast, G.M., Winther, Å.M.E., Nachman, R.J., Taylor, C.A.M., Shirras, A.D., Coates, D., Isaac, R.E., and Nässel, D.R. (2000) Expression and functional characterization of a *Drosophila* neuropeptide precursor with homology to mammalian preprotachykinin A. *J. Biol. Chem.* **275**, 23273–23280.
10. Winther Å.M.E., Siviter R.J., Isaac R.E., Predel R., Nässel D.R. (2003). Neuronal expression of tachykinin-related peptides and gene transcript during postembryonic development of Drosophila. *J. Comp. Neurol.* 464:180–196.

Chapter 11

Peptidomic Analysis of Single Identified Neurons

Susanne Neupert and Reinhard Predel

Abstract

Today, commercially available mass spectrometers increasingly meet all the demands of the proteomics community including high throughput, high sensitivity, and significant fragmentation capability for sequence determinations. Therefore, proper sample preparation is often the most crucial step to obtain the necessary data, particularly when working with biological samples. Depending on the size, sample preparation techniques differ and have to be optimized empirically. This is particularly apparent at the single cell level. In this chapter, we describe protocols for the use of MALDI-TOF mass spectrometry to directly analyse the peptidome of single insect neurons.

Key words: Peptidomics, single-cell analysis, neuropeptides, retrograde filling, dye injection, green fluorescence protein (GFP), MALDI-TOF mass spectrometry, insects.

1. Introduction

Studying the function of the CNS or behavioural patterns in general without exact structural knowledge of the neuropeptides involved can only give an incomplete view of the physiological processes in an organism. Clearly, information about cell-specific expression or cell-specific post-translational modifications of gene products and relative abundances of products from different genes cannot be deduced from genome information and are also not detectable by proteomic analyses of tissue extracts. Thus, single-cell analysis can contribute essentially to a better understanding of the complex functions of neuronal circuits. MALDI-TOF mass spectrometry is the method of choice to study peptidergic intercellular communication capabilities of neurons (1, 2).

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_11,

In contrast to proteomic analysis which usually has to deal with highly complex mixtures of proteins and peptides, the challenge in single-cell peptidomics is the development of feasible cell preparation protocols that ensure optimal signal intensity in subsequent mass spectrometric analyses. In this context, it is important to correctly identify specific cells and to avoid contaminations during cell dissection. This is much more critical than any subsequent step during the acquisition of mass spectra from single-cell samples. Reducing the sample complexity from tissues to single cells also reduces the complexity of the peptidome and may lead to the detection of peptides that are otherwise obscured by abundant signals in tissue samples. In addition, single-cell analysis offers unparalleled information about co-localized neuropeptides and therefore complements or verifies immunocytochemical findings. In this chapter, we present an overview of the methods that are used for cell identification, dissection, and subsequent mass spectrometric analysis of peptidergic neurons in insects. Such approaches were successfully used for the analysis of cockroach neurons (3), moth neurons (4), and *Drosophila* neurons with a size of 10 μm and revealed novel insights about prohormone processing (5).

2. Materials

2.1. Sample Preparation

1. Dissecting saline: 128 mM NaCl, 2.7 mM KCl, 2 mM $CaCl_2$, and 1.2 mM $NaHCO_3$, pH 7.25.
2. Dextran-tetramethylrhodamine (Molecular Probes).
3. Stereo fluorescence microscope, e.g. SteREO Lumar.V12 (Carl Zeiss AG, Germany) equipped with an EX BP 450-490 (FITC) and an EX BP 550/25 (Cy3) filter or similar equipment.
4. Inverse fluorescence microscope with digital camera, e.g. Nikon Eclipse TE 2000 U (Nikon GmbH, Germany).
5. A pair of fine forceps (e.g. sharpened Dumont No. 5), ultrafine spring or clipper scissors (Fine Science Tools GmbH, Heidelberg, Germany), tungsten micro-needles (custom-made from electrolytically sharpened tungsten wire).
6. Home-made uncoated glass capillaries (e.g. Hilgenberg GmbH) fitted to a tube with mouthpiece (e.g. a sterile pipette tip).
7. Sylgard-coated preparation dish.

2.2. MALDI-TOF Matrix Application

1. Re-crystallized α-cyano-4-hydroxycinnamic acid (CHCA, Sigma-Aldrich).
2. 2,5-Dihydroxybenzoic acid (DHB, Fluka).
3. Methanol (MeOH), Trifluoroacetic acid (TFA), Acetonitrile (ACN), all HPLC grade.
4. Water, double-distilled or HPLC-grade.
5. Nanoliter applicator (World Precision Instruments, Berlin, Germany) or equivalent.

2.3. MALDI-TOF Mass Spectrometry

1. MALDI target plates (we typically use simple stainless steel target plates).
2. MALDI-TOF mass spectrometer, e.g. LBMS Voyager-DE STR biospectrometry workstation (Applied Biosystem), or UltraFlex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics), or 4700/4800 MALDI TOF/TOF™ Analyzer (Applied Biosystem).
3. ***FlexAnalysis*** software (Bruker Daltonics), Voyager ***DataExplorer***™ 2.4 (Applied Biosystem), or similar.
4. ***ProteinProspector*** – Proteomics tools for mining sequence databases in conjunction with mass spectrometry experiments (http://prospector.ucsf.edu).

3. Methods

3.1. Sample Preparation

3.1.1. Tissue Dissection

Dissect insect ganglia, containing the cells of interest, in insect saline. Remove the connective tissue (muscles or fat body). If retrograde labelling has to be performed for cell identification, attached nerves should be severed as distal as possible (*see* **Note 1**).

3.1.2. Cell Identification

3.1.2.1. Retrograde Labelling of Neurons with Peripheral Projection

Transfer the ganglion into a chamber containing insect saline (4°C) and place the attached nerve that contains the neurites of the cells of interest in a drop of 10% dextran-tetramethylrhodamine. The fluorescence dye should be separated from the main chamber by vaseline. Staining of cell bodies via passive diffusion into the ganglion starts after 20–48 h (**Fig. 11.1a**).

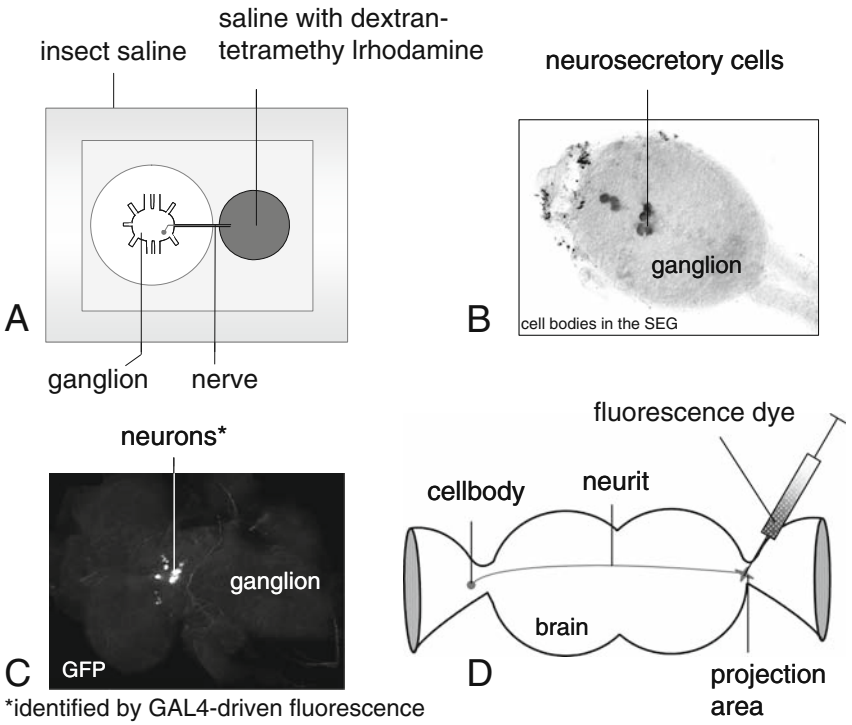


Fig. 11.1. Methods for cell identification. (**a**) Retrograde filling of neurons via nerves, which contain the respective neurites and are placed in a separate chamber containing fluorescence dye. (**b**) Cell identification in the subesophageal ganglion (SEG) via the bluish colour which is typical of some neurosecretory neurons (Tyndall effect). (**c**) In a number of insects, neurons can be visualized by GAL4-lines which promote the expression of marker proteins such GFP or YFP. (**d**) Neurons (e.g. interneurons) which cannot be visualized by retrograde filling of external nerves or GAL4-lines can be stained by injection of fluorescence dye (dextran-tetramethylrhodamine) into the projection area of these neurons.

3.1.2.2. Dye Injection

Open the integument and remove all tissues (e.g. fat body, muscles, trachea) which cover the ganglion. Disrupt the ganglionic sheath in the vicinity of the projection area of the neurons of interest using fine forceps or ultra-fine scissors. A small volume (< 1 nl) of dye solution has to be injected directly into the projection area of the neuron by a glass capillary fitted to a tube with mouthpiece (**Fig. 11.1d**). Subsequently, seal the integument with superglue. Staining of neurons via passive diffusion starts after 12 h (*see* **Note 2**).

3.1.2.3. GAL4-Driven Expression of Fluorescent Proteins in Specific Neuron Populations

A variety of Gal4-lines are available for insects such as *Drosophila melanogaster* and *Tribolium castaneum*. These are suitable for the identification of different neuron populations (*see* **Fig. 11.1c**) (*see* **Note 3**).

3.1.2.4. Tyndall Effect

Some neurosecretory neurons can easily be identified due to their slightly bluish colour, this phenomenon is known as the Tyndall effect. In transmitted light the neurons appear blue because of the light scattering due to peptide vesicles that they contain.

The intensity of the colour depends on the concentration of these dense-core vesicles (**Fig. 11.1b**) and varies between insect species (*see* **Note 4**).

3.1.3. Cell Dissection

1. Fix the ganglia with micro-needles, search for the labelled neurons with a stereo fluorescence microscope (cell identification **Sections 3.1.2.1, 3.1.2.2, and 3.1.2.3**) or a stereo microscope without fluorescence option (cell identification **Section 3.1.2.4**), and disrupt the dorsal ganglionic sheath in the vicinity of the labelled cells using ultra-fine scissors.
2. Without prior enzyme treatment, remove the labelled cells step by step using an uncoated glass capillary fitted to a tube with mouthpiece and transfer the cells to a stainless steel sample plate for MALDI-TOF mass spectrometry. This is a purely mechanical isolation (**Fig. 11.2**) (*see* **Note 5**).

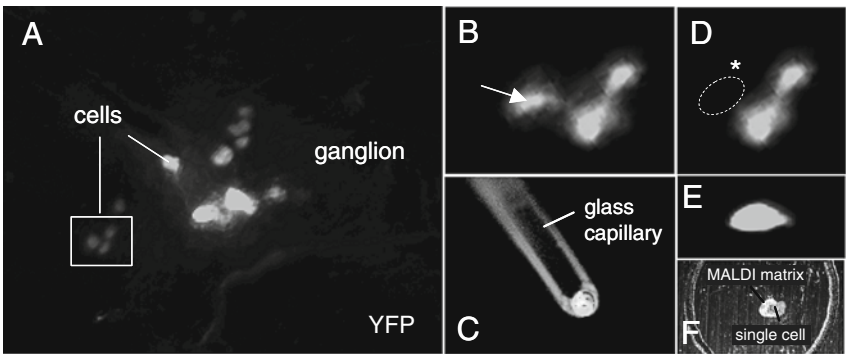


Fig. 11.2. Cell dissection of identified neurons for MALDI-TOF MS. The figure shows *Drosophila* neurons which have been visualized by expression of YFP (9). (**a**) The ganglionic sheath of the isolated CNS is disrupted to approach and separate the labelled cells. (**b**, **d**) Detail of the nervous tissue before and after the isolation of a single cell, which was removed by means of a glass capillary (as in panel **c**). (**e**) Microphotograph of the intact cell body on the MALDI sample plate. To check the success of the single-cell dissection, the preparation was always examined and documented using an inverted fluorescence microscope (panels **a**–**d**). (**f**) Microphotograph of the single-cell preparation after matrix application as described in the text. The photo was taken using a digital camera mounted on the dissecting stereo microscope.

3. For documentation, microphotographs of the ganglia before and after cell dissections should be taken using the stereo microscope or (preferably) an inverse microscope each time a single cell is removed.

3.2. MALDI-TOF Matrix Application

1. Remove the insect saline, containing the isolated cell, from the MALDI target by using the same glass capillary which was used for the cell transfer. The neuron should stick on the surface of the sample plate.
2. Rinse the rim around the dried cell, not the cell itself, with water for a few seconds to decrease salt contamination; the water can be removed with a glass capillary.

3. Apply 5–20 nl of matrix solution (depending on the cell size) onto the dried cell over a period of about 2–5 s using a Nanoliter injector. To analyse peptides with a mass <3 kDa, we usually try saturated α-cyano-4-hydroxycinnamic acid first, dissolved in 50% methanol (1:1 or 1:2). For peptides with an ion mass >3 kDa, we dilute 2,5-dihydroxybenzoic acid (50 mg/ml) in 30% acetonitrile and 0.1% trifluoroacetic acid (*see* **Fig. 11.3**). The optimal ratio of matrix molecules and analyte molecules should be determined empirically for each cell type (*see* **Note 6**).

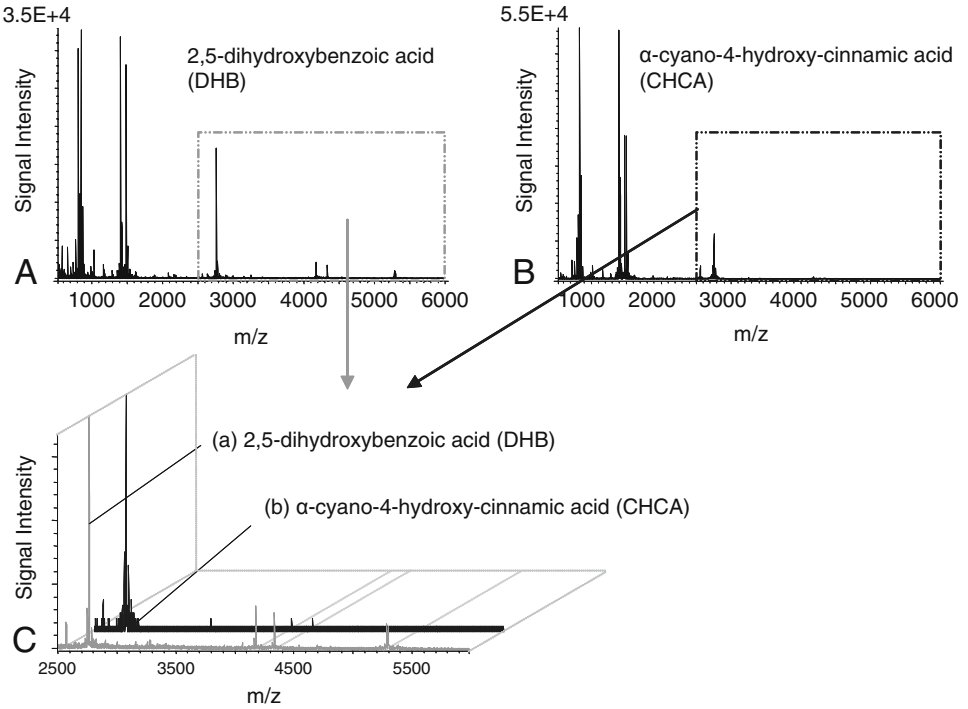


Fig. 11.3. Comparison of mass spectra from single identified neurons prepared with different matrix solutions. (**a**) 2,5-dihydroxybenzoic acid (DHB) and (**b**) α-cyano-4-hydroxy-cinnamic acid (CHCA). DHB enhances the intensity of the ion signals above 3 kDa but results in less reproducible mass spectra when analysing small insect neurons.

4. Cover the dried preparations (only CHCA-preparations) with pure water or 0.1% TFA, which should be removed after a few seconds by cellulose paper. Rinsing with water is commonly used to reduce high salt content in biological samples as well as to remove free matrix crystals.
5. Place an appropriate peptide standard close to the sample spots on the sample plate. This standard peptide mixture is important not only for an optimal calibration of the mass spectrometer but also for the tuning of the settings such as grid voltage, guide wire voltage, delay time, and laser intensity (*see* **Note 7**).

3.3. MALDI-TOF Mass Spectrometry

1. Load the sample plate into the mass spectrometer and use the synthetic peptide standard for calibration and tuning of the mass spectrometer.
2. Analyse the sample in the reflector mode as recommended for peptide samples. Limit the amount of laser shots for each spectrum acquisition to 20–50, accumulate the spectrum, and use the same spot again. If the ion intensity has markedly decreased, discard the last spectrum and move on to a new spot. Move the laser to different spots on and directly around the tissue (*see* **Note 8**).
3. For fragmentation, find a spot with reasonable ion intensity. For that, use a low number of shots for each spectrum acquisition to minimize peptide loss prior to fragmentation. Select the parent ion for fragmentation and start fragmentation. With low amounts of material, best results are usually obtained when the collision gas source is turned on (*see* **Note 9**).

4. Notes

1. Be careful not to damage the nerve attached to the ganglion; otherwise, backfilling will not work effectively.
2. Interneurons without peripheral projections cannot be visualized by retrograde labelling of external nerves but can often be visualized in larger insects after injection of dextran-tetramethylrhodamine in the putative projection area of these neurons within the ganglia (6).
3. Retrograde filling or dye injections are not suitable techniques to identify specific neurons in insects as small as *D. melanogaster* and *T. castaneum*. Instead, Gal4-lines allow the identification of different neuron populations by expression of fluorescent marker proteins such as green fluorescence protein (GFP) under upstream activating sequence (UAS) control (7). Fluorescence protein in neurons does not suppress ion signals in mass spectra.
4. Examples of neurons with distinct Tyndall effect are PBAN-expressing neurons in the subesophageal ganglion of moths (4) and neurons of the *pars intercerebralis* in many fly species.
5. If cell bodies become damaged during the dissection, the fluorescence signal of the labelled cells disappears and subsequent mass spectra are usually poor even if the cell is properly placed on the sample plate.
6. To avoid contamination, use a new or thoroughly rinsed capillary for each cell sample.

7. The synthetic peptides should be analysed using the same settings as expected for the cell samples (use very low peptide concentrations!). The single-cell preparations do not contain enough material for testing and selecting the optimal conditions.
8. Start analyses with relatively low laser energy and few laser shots. The cell samples deplete quickly since modern mass spectrometers operate with a high frequency of laser shots; this is not advantageous for single-cell analyses.
9. Some MALDI-TOF mass spectrometers (e.g. Applied Biosystems ABI 4700/4800 TOF/TOF mass spectrometer) allow the unambiguous assignment of isomeric leucine and isoleucine amino acids even when profiling small insect samples (8). For that, retake the spectra under conditions of high gas pressure.

Acknowledgments

Original work was supported by the Deutsche Forschungsgemeinschaft (DFG, PR 595/1-1...7, 6-1...4; and PR 766/9-1 to R Predel, J Schachtner and C Wegener) and by the Peter und Traudl Engelhorn Stiftung (to SN).

References

1. Hummon, A.B., Amare, A., and Sweedler, J.V. (2006) Discovering new invertebrate neuropeptides using mass spectrometry. *Mass Spectrom. Rev.* **25**, 77–98.
2. Rubakhin, S.S. and Sweedler, J.V. (2007) Characterizing peptides in individual mammalian cells using mass spectrometry. *Nat. Protoc.* **2**, 1987–1997.
3. Neupert, S. and Predel, R. (2005) Mass spectrometric analysis of single identified neurons of an insect. *Biochem. Biophys. Res. Commun.* **327**, 640–645.
4. Ma, P.W., Garden, R.W., Niermann, J.T., O'Connor, M., Sweedler, J.V., and Roelofs, W.L. (2000) Characterizing the Hez-PBAN gene products in neuronal clusters with immunocytochemistry and MALDI MS. *J. Insect Physiol.* **46**, 221–230.
5. Neupert, S., Johard, H.A., Nassel, D.R., and Predel, R. (2007) Single-cell peptidomics of *Drosophila melanogaster* neurons identified by Gal4-driven fluorescence. *Anal. Chem.* **79**, 3690–3694.
6. Reischig, T. and Stengl, M. (2002) Optic lobe commissures in a three-dimensional brain model of the cockroach *Leucophaea maderae*: a search for the circadian coupling pathways. *J. Comp. Neurol.* **443**, 388–400.
7. Brand, A. (1999) GFP as a cell and developmental marker in the *Drosophila* nervous system. *Methods Cell Biol.* **58**, 165–181.
8. Nachman, R.J., Russell, W.K., Coast, G.M., Russell, D.H., and Predel, R. (2005) Mass spectrometric assignment of Leu/Ile in neuropeptides from single neurohemal organ preparations of insects. *Peptides* **26**, 2151–2156.
9. Melcher, C. and Pankratz, M.J. (2005) Candidate gustatory interneurons modulating feeding behavior in the *Drosophila* brain. *PLoS. Biol.* **3**, e305.

Chapter 12

Identification and Analysis of Bioactive Peptides in Amphibian Skin Secretions

J. Michael Conlon and Jérôme Leprince

Abstract

Skin secretions from anurans (frogs and toads), particularly those species belonging to the Hylidae and Ranidae families, are a rich source of biologically active peptides. Cytolytic peptides with broad-spectrum antimicrobial activities and highly variable amino acid sequences are often released into these secretions in high concentrations. Identification and characterization of these components can prove to be valuable in species identification, elucidation of evolutionary histories and phylogenetic relationships between species, and may lead to development of agents with potential for therapeutic application. This chapter describes the use of norepinephrine (injection or immersion) to stimulate peptide release in a procedure that does not appear to cause distress to the animals. The peptide components in the secretions are separated by reversed-phase HPLC on octadecylsilyl silica (C_{18}) columns under standard conditions after partial purification on Sep-Pak cartridges. Individual peptides are identified by determination of their molecular masses by MALDI-TOF mass spectrometry and from their retention times. The use of mixtures of synthetic peptides of appropriate molecular mass as calibration standards enables mass determination to a high degree of precision.

Key words: Frog skin secretions, antimicrobial peptide, reversed-phase HPLC, MALDI-TOF mass spectrometry.

1. Introduction

Analysis of skin secretions and/or skin extracts from different species of Anura (frogs and toads) has led to the characterization of a wide range of peptides with biological activity that may have potential for development into therapeutically valuable agents. Examples include antimicrobial peptides with broad-spectrum activity against bacteria and fungi; myotropic peptides

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_12,

such as bradykinin- and tachykinin-related peptides, caeruleins and bombesins; mast cell degranulating peptides; neuroendocrine peptides such as TRH, angiotensin-II and opioids; inhibitors of proteolytic enzymes and nitric oxide synthase; and pheromones [reviewed in (1)]. Skin secretions from frogs belonging to the subfamilies Phyllomedusinae (South American tree frogs) (2) and Pelodryadinae (Australian tree frogs) (3) in the family Hylidae, and North American and Eurasian frogs in the family Ranidae (4) have proved to be particularly rich sources of bioactive peptides.

Cytolytic peptides synthesized in the skins of many, although by no means all, frog species play a role defending the animal against invasion by pathogenic microorganisms, and may also be important in deterring ingestion by predators (5). On the basis of limited similarities in amino acid sequence, the frog skin antimicrobial peptides may be grouped together in families that share a common evolutionary origin (4, 6) but the variation in primary structure among individual family members is considerable. It is rare that a peptide from one species is found with an identical amino acid sequence in another even when those species are closely related phylogenetically. Consequently, determination of the primary structures of these peptides can be used to complement morphological and other types of molecular analysis, such as comparisons of nucleotide sequences of orthologous genes, to provide valuable insight into taxonomy and phylogenetic relationships (4). In addition, morphological differences between species belonging to the same genus are often slight so that taxonomic classification of specimens can be difficult. Unambiguous identification of individuals is especially challenging in regions where several species coexist and produce hybrids so that peptidomic analysis of skin secretions is a useful molecular technique that can be used to aid the taxonomic classification. Quantitative and qualitative peptidomics data, however, must be interpreted with caution, because the expression of the genes encoding skin peptides is seasonally and hormonally dependent (7). Individuals that are nominally assigned to the same species group but are from different geographical regions may synthesize dermal peptides with different amino acid sequences because of the extreme hypermutability of the genes, particularly those encoding antimicrobial peptides.

The dermal peptides are synthesized and stored in granular glands present in the skin and are released into skin secretions, often in very high concentrations, in a holocrine manner upon stress or injury as a result of contraction of myocytes surrounding the glands. In the laboratory, electrical stimulation is an effective and non-invasive method of inducing secretion of skin peptides but may cause the animals some distress and is difficult to perform in the field (8). Intradermal injection of norepinephrine is

an equally effective method of releasing peptides into skin secretions that is well tolerated by animals (9) and is the procedure that will be described in this chapter. A variation on the method involves immersion of the animal in a solution of norepinephrine. This is a less effective stimulus but completely non-invasive and so may be more appropriate for rare and endangered species. The method of peptidomic analysis of the secretions involves separation of the peptides by chromatography on an octadecysilyl silica (C_{18}) column under standard conditions and determination of the molecular masses of individual components by high-precision Matrix-Assisted Laser Desorption Time-of-Flight (MALDI-TOF) mass spectrometry. In comparison to other classes of vertebrates, relatively little work has been done to characterize the genomes of amphibian species so that it is necessary to characterize structurally the peptides present in skin secretions by determination of their amino acid sequences using Edman degradation or tandem MS/MS mass spectrometry. In this way, researchers can create a database of dermal peptides for each species that is of particular value in species identification and taxonomic classification. In view of the fact that >5000 species of frogs are currently included in the Amphibian Species of the World on-line database (10) and skin secretions of <100 species have been examined in detail, it must be pointed out that this work is in its infancy. At this time, data cannot be analysed in the unambiguous way that is possible with species for which there is complete or partial structural characterization of the genome. Identification of peptides in skin secretions from a particular species that has been studied previously is made on the basis of molecular mass and observed retention time on reversed-phase HPLC.

2. Materials

2.1. Collection and Partial Purification of Skin Secretions

1. Norepinephrine bitartrate salt (Sigma-Aldrich).
2. Collection buffer: 25 mM sodium chloride, 25 mM ammonium acetate pH 7.0.
3. Concentrated hydrochloric acid (36%).
4. Sep-Pak C-18 cartridges (Waters Associates).
5. Peristaltic pump (Gilson Minipuls 3).
6. Speed-Vac concentrator with multi-tube rotor (Savant).

2.2. Reagents and Equipment for HPLC

1. Acetonitrile (HPLC Spectro grade, Pierce).
2. Water (Milli-Q purified; 18.2 mohm.cm^{-1}) or HPLC Spectro grade (Pierce).

3. Trifluoroacetic acid (99.8% purity; Sequenal grade, Pierce).
4. An HPLC system capable of generating a binary gradient using pumps operating in the flow rate range 0.1–10 mL/min, a Rheodyne 7125 injection system equipped with a 2 mL loop, a detector capable of simultaneously monitoring at two wavelengths (typically 214 and 280 nm) (*see* **Note 1**).
5. HPLC column: The size of C_{18} column chosen for the separation step is dependent upon the amount of peptide to be fractionated after recovery from Sep-Pak cartridges. For samples containing <1 mg of material, a (0.46 × 25-cm) Vydac 218TP54 analytical C_{18} column (Separations Group) should be used at a flow rate of 1.5 mL/min. For material in the 1–10 mg range, a (1.0 × 25-cm) Vydac 218TP510 semi-preparative C_{18} column should be used at a flow rate of 2.0 mL/min. In the case of samples containing large amounts of peptide material (> 10 mg), a (2.2 × 25-cm) Vydac 218TP1022 preparative C_{18} column should be used and the flow rate increased to 6 mL/min. The approximate amount of peptide material may be estimated using a BCA (bicinchoninic acid) protein assay reagent kit following the manufacturer's recommended procedure.
6. Helium cylinder for degassing of solvents (*see* **Note 2**).
7. 2.5 mL Gastight model 1002 injection syringe (Hamilton).
8. Polypropylene tubes 12 mm × 75 mm and 15 mm × 100 mm (Nunc) (*see* **Note 3**).
9. Microcentrifuge (Eppendorf model 5415D).

2.3. MALDI-TOF Mass Spectrometry

1. Sample diluent: 50% (v/v) acetonitrile-water containing 0.1% (v/v) trifluoroacetic acid (*see* **Section 2.2**).
2. MALDI matrix: α-cyano-4-hydroxycinnamic acid (α-CHCA) recrystallized and cation-depleted (LaserBio Labs).
3. Peptide calibration samples. Mixture 1: [des-Arg^1] bradykinin 2.3 μg; angiotensin-I 4.2 μg; [Glu^1] fibrinopeptide B 5.1 μg; neurotensin 0.2 μg. Mixture 2: Angiotensin-I 6.5 μg; $ACTH_{(1–17)}$ 10.5 μg; $ACTH_{(18–39)}$ 9.3 μg; $ACTH_{(7–38)}$ 27.5 μg; bovine insulin 50.2 μg. These calibration mixtures are supplied in lyophilized form by Applera.
4. RBS35 detergent for cleaning sample plate (Fisher Scientific).
5. Ultrasonic bath (LEO-50, Fischer Scientific).

3. Methods

3.1. Collection of Skin Secretions

1. Animals are individually injected at two sites within the dorsal sac with a freshly prepared solution of norepinephrine in water (2 nmol/g body weight in a volume of 200 μL) (*see* **Note 4**).
2. Immediately after injection, the animal is placed for a period of 15 min in the collection buffer contained in a covered glass beaker (typically 100 mL with a greater volume for particularly large specimens) (*see* **Note 5**).
3. The solution containing the secretions is acidified immediately after collection with concentrated hydrochloric acid (final concentration 1% v/v) and stored at –20°C until time of analysis (*see* **Note 6**).

3.2. Sample Preparation Using Sep-Pak Cartridges

1. Preparation of HPLC solvent A: 1.2 mL Trifluoroacetic acid is added to 1000 mL water (*see* **Note 7**).
2. Preparation of HPLC solvent B: 1.0 mL Trifluoroacetic acid is added to 700 mL acetonitrile/300 mL water.
3. Sep-Pak cartridges are activated by pumping acetonitrile (2 mL per cartridge) at a flow rate of 2 mL/min using a peristaltic pump or manually using a 20-mL plastic syringe, followed by solvent A (2 mL/cartridge). Up to ten Sep-Pak cartridges may be connected in series depending on the volume and amount of peptide material in the extract to be processed.
4. The combined skin secretions and washings are centrifuged (5000×*g* for 30 min) (*see* **Note 8**).
5. The supernatant is pumped through the Sep-Pak cartridges at a flow rate of 2 mL/min. If a peristaltic pump is not available, a polypropylene syringe (50 mL) can be used to apply the solution manually.
6. Solvent A (4 mL/cartridge) is pumped at a flow rate of 4 mL/min and the eluate discarded.
7. Solvent B (2 mL/cartridge) is pumped at a flow rate of 1 mL/min and the eluate collected into polypropylene tubes.
8. The volume of eluate is reduced to approximately 1.5 mL under reduced pressure in a Speed-Vac concentrator (*see* **Note 9**). With an efficient vacuum pump, this step can be accomplished in approximately 60 min at room temperature or in 30 min with external heating. Sample may be stored at 4°C overnight or frozen at –20°C for several days.

3.3. Peptide Separation by Reversed-Phase HPLC

1. HPLC solvent A and solvent B (*see* **Section 3.2**) are degassed with helium for 1 min.
2. Column preparation. Before injecting the sample, it is necessary to "condition" the column in order to improve resolution. The column is irrigated at an appropriate flow rate (*see* Step 5 in **Section 2.2**) with solvent B for 20 min. The concentration of solvent B is decreased to 0% over 10 min using a linear gradient. The concentration of solvent B is increased to 100% over 10 min. The concentration of solvent B is decreased to 0% over 10 min. The column is equilibrated with solvent A for 20 min.
3. The sample is centrifuged for 5 min at 13,000×*g* in a 1.5-mL polypropylene Eppendorf tube to ensure clarity of solution (*see* **Note 10**).
4. The HPLC system is programmed to perform chromatography under the following conditions using linear gradients for elution: (a) increase concentration of solvent B from 0 to 30% over 10 min, (b) increase concentration of solvent B from 30 to 90% over 60 min, (c) increase concentration of solvent B from 90 to 100% over 1 min and hold at 100% until UV-absorbance returns to baseline value (*see* **Note 11**).
5. For the example shown in **Fig. 12.1**, the sample is injected onto a (1.0 × 25-cm) Vydac 218TP510 semi-preparative

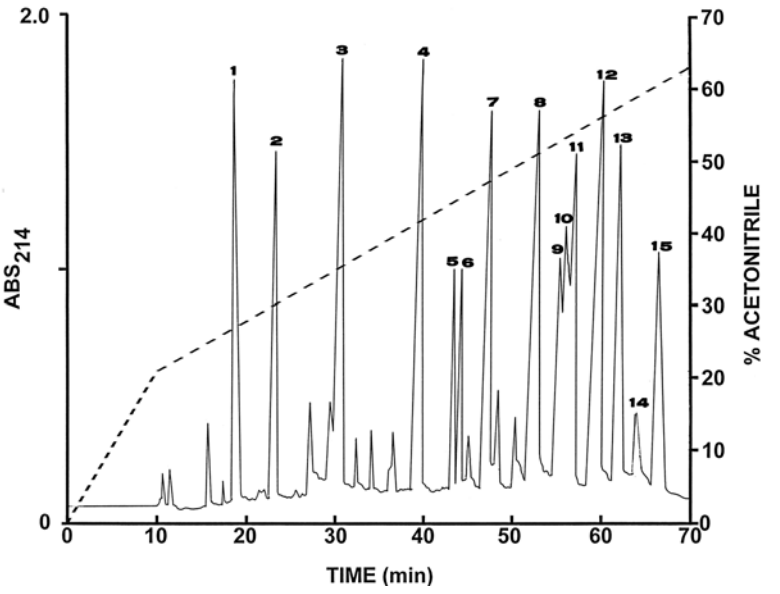


Fig. 12.1. Elution profile on a semi-preparative Vydac C_{18} column of skin secretions from the Colorado spotted frog *Rana luteiventris*, after partial purification on Sep-Pak cartridges. Aliquots of the peaks designated 1–15 were subjected to MALDI-TOF mass spectrometry. The *dashed line* shows the concentration of acetonitrile in the eluting solvent. The retention times of the major peaks are compared with those obtained during chromatography of an extract of *R. luteiventris* skin under the same conditions (13).

C_{18} column equilibrated with solvent A at a flow rate of 2 mL/min for 20 min. Up to 1.5 mL may be injected into a 2 mL loop.

6. The concentration of solvent B is increased according to the elution program listed in Step 5 in **Section 3.3** . Fractions corresponding to each UV-absorbing peak are collected by hand into polypropylene tubes (*see* **Note 12**).
7. Aliquots (20–200 μL depending on the size of the UV-absorbing peak) of each fraction are taken for analysis by mass spectrometry and dried under reduced pressure in a Speed-Vac concentrator (*see* **Note 13**). The elution profile on a semi-preparative Vydac C_{18} column of skin secretion obtained from the North American frog *Rana luteiventris* Thompson, 1913 is shown in **Fig. 12.1**. Aliquots of the major peaks designated 1–15 were subjected to analysis by MALDI-TOF mass spectrometry (*see* **Note 14**).

3.4. Preparation of the MALDI Matrix Solution

1. α-Cyano-4-hydroxycinnamic acid is dissolved in sample diluent to give a final concentration of 10 mg/mL.
2. The solution is vortexed for 15 s at low speed and placed in an ultrasonic bath for 1 min.
3. The solution is centrifuged (5000×*g* for 1 min) to remove any undissolved matrix and the supernatant allowed to stand for 10 min (*see* **Note 15**).

3.5. Preparation of Analyte and Peptide Calibration Mixtures

1. Each lyophilized HPLC fraction is reconstituted in 10 μL of sample diluent.
2. Samples are vortexed for 15 s at low speed and centrifuged (5000×*g* for 1 min).
3. The peptide calibration mixtures are reconstituted in 100 μL of sample diluent (*see* **Note 16)**. Working solutions are prepared by dilution of 1 μL of the stock solution in 24 μL of matrix solution.
4. Preparation of the sample plate: Sample plate is scrubbed clean with detergent and rinsed extensively with ethanol, 50% (v/v) ethanol-water, and deionised water. The plate is allowed to dry in air.
5. 1 μL of each reconstituted HPLC fraction is loaded onto the sample plate in a defined position.
6. 1 μL of matrix solution is loaded onto each sample drop and mixed by aspiration into the pipette tip (*see* **Note 17**).
7. The analyte/matrix mixture is allowed to dry in air at room temperature.
8. The sample plate is loaded into the mass spectrometer.

3.6. Determination of the Peptide Mass Range in Each Sample (*see* Note 18)

1. To obtain a preliminary spectrum, each sample is analysed using the following parameters: linear mode (*see* **Note 19**), positive polarity (detection of MH^+ ions), wide mass range (500–15,000 Da), and default instrument settings (accelerating voltage 20,000 V; grid 95%; guide wire 0.05%; delay time 450 ns; 500 shots/spectrum).
2. The laser intensity is adjusted manually to improve signal-to-noise ratio (approximately 50:1). If the laser intensity is too high, the signal may be saturated.
3. A second spectrum of each sample is obtained using the following parameters: reflector mode (*see* **Note 19**), positive polarity, narrower mass range (500–6000 Da) and default instrument settings (accelerating voltage 20,000 V; grid 76%; guide wire 0.002%; delay time 255 ns; 500 shots/spectrum). The laser intensity is again adjusted manually to improve signal-to-noise ratio.
4. The sample plate is ejected from the instrument.

3.7. Determination of the Accurate Mass of the Peptides in Each Sample (*see* Note 20)

1. A peptide calibration mixture appropriate to the mass range is assigned to each sample.
2. 0.5–1 μL of working solution of the appropriate peptide calibration mixture is applied to the sample plate as close as possible to the sample.
3. The mixture is allowed to dry in air at room temperature.
4. The sample plate is loaded into the instrument.
5. In general, peptides from amphibian skin secretions do not exceed 6000 Da (50 amino acid residues) and are readily detectable in reflector mode and positive polarity. Consequently, two methods of spectrum acquisition covering two mass ranges are used. The first method covers the mass range from 500 to 2500 Da and employs peptide calibration mixture 1. The default instrument settings are as follows: accelerating voltage 20,000 V; grid 76%; guide wire 0.002%; delay time 100 ns; 500 shots/spectrum. The second method covers the mass range from 500 to 6000 Da and uses peptide calibration mixture 2. The default instrument settings are as follows: accelerating voltage 20,000 V; grid 76%; guide wire 0.002%; delay time 150 nc; 500 shots/spectrum.
6. The calibration peptide spot assigned to the first sample is selected.
7. This spot is analysed using a default calibration and the appropriate method. The laser intensity is adjusted manually to improve signal-to-noise ratio (approximately 50:1). Resolution is optimized by setting the grid and guide wire

voltages and delay time. Acceptable resolution is determined by the mass range. In the range 500–2500 Da, 6000 or greater is acceptable; for the range 500–6000 Da, 7000 is required.

8. The best spectrum is saved and the values of the observed masses of the peptides are adjusted to correspond to the reference masses. Calibration standards appropriate to the mass of the frog skin peptide that will be analysed are selected. The following monoisotopic reference masses (Da) are used: [des-Arg^1]bradykinin 904.468; angiotensin-I 1296.685; [Glu^1]fibrinopeptide B 1570.677; neurotensin 1672.917; ACTH$_{(1–17)}$ 2093.087; ACTH$_{(18–39)}$ 2465.199; ACTH$_{(7–38)}$, 3657.929; bovine insulin, 5730.609 ($n = +1$) and 2865.808 ($n = +2$). The calibrated spectrum of the standard peptide mixture is

Table 12.1
Identification by MALDI-TOF mass spectrometry of the major peptide components in skin secretions from *Rana luteiventris*

Peak	M_r	Peptide	Amino acid sequence
1	1849.0	Unknown	
2	1059.7	Bradykinin	RPPGFSPFR
3	8523.7	Unknown	
4	3286.9	Ranatuerin-2La	GILDSFKGVAKGVAKDLA GKLLDKLKCKITGC
5	2990.7 2611.5	Unknown	
6	3231.9	Ranatuerin-2Lc	GILSSFKGVAKGVAKDLA GKLLDTLKCKITGC
7	3196.9	Ranatuerin-2Lb	GILSSIKGVAKGVAKNVAA QLLDTLKCKITGC
8	3747.4	Esculentin-2La	GILSLFTGGIKALGKTLFKMAG KAGAEHLACKATNQC
9	1366.0	Temporin-La	VLPLISMALGKLL.NH_2
10	2841.5	Ranatuerin-2Ld	GILSSIKGVAKNVAAQ LLDTLKCKITGC
11	3741.8	Esculentin-2Lb	SIFSLLTAGAKVLGKTLLKM AGKAGAEHLACKATNQC
12	2578.4	Brevinin-1Lb	FLPMLAGLAASMVPK FVCLITKKC
13	1367.7	Temporin-Ld	FLPILGNLLSGLL.NH_2
14	1545.0	Unknown	
15	1573.9	Temporin-1Lb	NFLGTLINLAKKIM.NH_2

The peak numbers refer to those indicated in **Fig. 12.1**. M_r refers to monoisotopic relative molecular mass.

generated using the software provided by the manufacturer of the instrument and saved.

9. The sample spot adjacent to the calibration peptide spot is selected.
10. The sample is analysed using the calibrated spectrum of the standard peptide mixture as external calibration file. The laser intensity and instrument settings are adjusted to obtain a spectrum with acceptable signal-to-noise ratio and resolution. The spectrum provides an accurate mass of each peptide component with a precision of 0.002% according to the specifications of the instrument.
11. The procedure is repeated for each sample/peptide calibration mixture pair. The monoisotopic masses of the peptides present in peaks 1–15 (**Fig. 12.1**) from *R. luteiventris* skin secretions are shown in **Table 12.1** (*see* **Note 21**).

4. Notes

1. A dual-pen flatbed chart recorder (Kipp and Zonen) to supplement any on-screen computer recording system is useful especially when there is a significant delay in the response of the absorbance detector and appearance of the peak on the computer screen.
2. Degassing of solvents may be achieved using an ultrasonic bath but degassing under reduced pressure is not recommended as it may lead to relative loss of volatile components.
3. Polypropylene tubes should be used throughout, not glass or polystyrene, in order to minimize irreversible binding of peptides to the tubes.
4. Animals that appear to be agitated or highly mobile can be partially anaesthetized by immersion for 5 min in crushed ice. There is no indication that this procedure affects the concentration and distribution of peptides in the norepinephrine-stimulated secretions.
5. Norepinephrine injection is generally well tolerated by the animals without signs of discomfort and fatalities are extremely rare. However, in the case of protected or endangered species, it may be impossible to obtain a permit to carry out an invasive procedure. In this case, the animals may be immersed for 15 min in collection buffer containing 200 μM norepinephrine. The agent is absorbed through the frog's skin and stimulates peptide release, although with

reduced effectiveness compared with injection (11). The collection buffer is processed as in Step 2 in **Section 3.1**.

6. If secretions are collected in the field, it may not be possible to freeze the sample immediately. In this case, secretions can be kept at 0°C in an ice bath for several hours until access to a freezer is obtained.
7. A slightly greater concentration of trifluoroacetic acid in solvent A than in solvent B produces a flat baseline under HPLC gradient elution conditions.
8. Frog skin secretions contain lipid components but it is generally unnecessary to remove them by extraction with an organic solvent prior to partial purification on Sep-Pak cartridges.
9. The sample should not be dried completely except when this is unavoidable (e.g. for shipment to a different laboratory) as this can lead to formation of insoluble material.
10. Filtering the sample is not recommended unless absolutely necessary as it can lead to appreciable loss of peptide by irreversible binding to the filter material.
11. At the end of the experiment, the column is washed with acetonitrile (100 mL) and stored in this solvent.
12. Fractions may be stored in stoppered tubes at –20°C for up to several months. However, peptides containing methionine and tryptophan residues may oxidize on prolonged storage, even at low temperature.
13. When mass spectrometry is to be carried out in the chromatographer's own laboratory or institute, it is preferable not to dry the fractions completely but reduce the volume to approx. 10–20 μl in order to minimize losses of peptide due to irreversible binding to the plastic tubes. When samples are to be analysed by an external core facility, lyophilization to dryness cannot be avoided.
14. Peak collection by hand is preferable to use of a fraction collector and often results in peptide fractions that are sufficiently pure for amino acid sequence analysis. When mass spectrometry reveals that the fraction contains multiple components, peptides may be separated by further chromatography on a (0.46 × 25-cm) Vydac 214TP54 analytical butylsilyl silica (C_4) column and a (0.46 × 25-cm) Vydac 219TP510 analytical phenyldimethylsilyl silica column. The use of these columns and the general strategy for purification of peptides to near homogeneity by reversed-phase HPLC are discussed in detail in a recent article (12).
15. The matrix solution is stored at 4°C and may be used for up to 1 week.

16. The peptide calibration mixtures are stored in single-use aliquots (1 μL) at –20°C.
17. It is important not to touch the surface of the plate with the pipette tip to avoid uneven crystallization.
18. The procedure described here refers to the use of a Voyager DE-PRO mass spectrometer equipped with delayed extraction reflector (Applied Biosystems) but is readily adaptable to other instruments.
19. The linear mode of operation is the most sensitive due to shorter flight path whereas reflector mode provides higher resolution and greater mass accuracy due to longer flight path and focusing action at the detector.
20. The first round of spectral acquisitions has allowed determination of the mass range of the constituents in each fraction and the best mode of analysis for each sample.
21. On the basis of these masses (*see* **Table 12.1**) it was possible to identify the components present in peaks 4, 7, 8–10, 12, 13 and 15 with previously characterized antimicrobial peptides that were isolated from an extract of R. *luteiventris* skin (13). Comparison of their observed retention times on reversed-phase HPLC (**Fig. 12.1**) with those reported in by Goraya et al. (13) provides confirmation that the identifications are correct. Peaks 6, 11 and 13 contained peptides whose masses had not been reported previously. Determination of their amino acid sequences by automated Edman degradation demonstrated that the peptides belonged to the esculentin-2, ranatuerin-2 and temporin families of antimicrobial peptide. Peak 2 contained bradykinin, confirmed by measurement of its retention time on HPLC.

References

1. Lazarus, L.H. and Attila, M. (1993) The toad, ugly and venomous, wears yet a precious jewel in his skin. *Prog. Neurobiol.* **41**, 473–507.
2. Nicolas, P. and Amiche, M. (2006) The dermaseptins, in *Handbook of Biologically Active Peptides* (Kastin, A.J., ed.), Elsevier, San Diego, CA, pp. 295–304.
3. Apponyi, M.A., Pukala, T.L., Brinkworth, C.S., Maselli, V.M., Bowie, J.H., Tyler, M.J., et al. (2004) Host-defence peptides of Australian anurans: structure, mechanism of action and evolutionary significance. *Peptides* **25**, 1035–1054.
4. Conlon, J.M., Kolodziejek, J., Nowotny, N. (2004) Antimicrobial peptides from ranid frogs: taxonomic and phylogenetic markers and a potential source of new therapeutic agents. *Biochim. Biophys. Acta* **1696**, 1–14.
5. Nascimento, A.C.,Fontes, W., Sebben A. and Castro, M.S. (2003) Antimicrobial peptides from anurans skin secretions. *Protein Pept. Lett.* **10**, 227–238.
6. Vanhoye, D., Bruston, F., Nicolas, P. and Amiche, M. (2003) Antimicrobial peptides from hylid and ranin frogs originated from a 150-million-year-old ancestral precursor with a conserved signal peptide but a hypermutable antimicrobial domain. *Eur. J. Biochem.* **270**, 2068–2081.
7. Ohnuma, A., Conlon, J.M., Kawasaki, H. and Iwamuro S. (2006) Developmental and

triiodothyronine-induced expression of genes encoding preprotemporins in the skin of Tago's brown frog *Rana tagoi*. *Gen. Comp. Endocrinol.* **146**, 242–250.
8. Tyler, M.J., Stone, D.J. and Bowie, J.H. (1992) A novel method for the release and collection of dermal, glandular secretions from the skin of frogs. *J. Pharmacol. Toxicol. Methods* **28**, 199–200.
9. Nutkins, J.C. and Williams, D.H. (1989) Identification of highly acidic peptides from processing of the skin prepropeptides of *Xenopus laevis*. *Eur. J. Biochem.* **181**, 97–102.
10. Frost, D.R. (2008) Amphibian species of the world: an online reference. Version 5.2. Electronic database accessible at http://research.amnh.org/herpetology/amphibia/index.php American Museum of Natural History, New York.
11. Davidson, C., Benard, M.F., Shaffer, H.B., Parker, J.M., O'Leary, C., Conlon, J.M., et al. (2007) Effects of chytrid and carbaryl exposure on survival, growth and skin peptide defenses in foothill yellow-legged frogs. *Environ. Sci. Technol.* **41**, 1771–1776.
12. Conlon, J.M. (2007) Purification of naturally occurring peptides by reversed-phase HPLC. *Nat. Protoc.* **2**, 191–197.
13. Goraya, J., Wang, Y., Li, Z., O'Flaherty, M., Knoop, F.C., Platz, J.E., et al. (2000) Peptides with antimicrobial activity from four different families isolated from the skins of the North American frogs *Rana luteiventris*, *Rana berlandieri* and *Rana pipiens*. *Eur. J. Biochem.* **267**, 894–900.

Chapter 13

An Efficient Protocol for DNA Amplification of Multiple Amphibian Skin Antimicrobial Peptide cDNAs

Shawichi Iwamuro and Tetsuya Kobayashi

Abstract

Antimicrobial peptides (AMPs) play an important role in the host's innate defence system in many organisms. Amphibian skin is expected to be a particularly rich source of novel AMPs. In amphibians, AMPs are produced from precursor proteins via specific cleavage by processing enzymes. While the nucleotide sequences of the AMP coding region in precursors are hypervariable, those of other regions, including the 5′- and 3′-untranslated regions (UTRs), are highly or relatively conserved in different precursors. Such nucleotide sequence conservation suggests an efficient strategy for molecular cloning of the antimicrobial peptide genes by 3′-rapid amplification of cDNA ends (3′-RACE) and reverse transcriptase polymerase chain reaction (RT-PCR) methods using specific primers. With this strategy in mind we have established an efficient protocol suitable for amplification of multiple cDNAs encoding amphibian AMP precursor proteins.

Key words: Amphibian skin, antimicrobial peptides, molecular cloning, 3′-RACE, RT-PCR.

1. Introduction

Antimicrobial peptides (AMPs) are an evolutionarily well-conserved component of the host innate defence system in a wide range of organisms, from bacteria to mammals. AMPs display a broad spectrum of antimicrobial activities against pathogenic microorganisms, including bacteria, viruses, and fungi (1–5), and might therefore provide useful therapeutic agents against antibiotic-resistant environmental pathogens. In contrast to antibiotics, the antimicrobial activities of AMPs are dependent on their primary and secondary structures. Although there

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_13,

is no common consensus amino acid sequence that is associated with antimicrobial activity, almost without exception, typical AMPs are cationic and hydrophobic and have a propensity to form an amphipathic helical conformation, which might interact with the negatively charged phospholipid head groups of the external leaflet of the microbial cytoplasmic membrane.

Amphibian skin is known to be a rich source of AMPs. To date, a large number of AMP sequences have been reported from various frog species, and this number is increasing quickly. Although there are many amino acid sequence variations among amphibian AMPs, their precursor proteins show a high degree of sequence identity, suggesting that they have arisen from a common ancestor (6). Typically an amphibian AMP precursor protein consists of three domains – a signal peptide region, an intervening sequence region, and an AMP region; the AMPs are produced via specific cleavage by processing enzymes (**Fig. 13.1**). The nucleotide sequences of the signal peptide and the intervening sequence coding regions of their mRNAs, especially in the signal peptide region, are highly conserved among different precursors. In addition, the nucleotide sequences of 5′- and 3′-untranslated regions (UTRs) of these precursor mRNAs are also relatively well conserved. Thus, the hypervariable AMP coding regions in the genes for amphibian AMP precursors are surrounded by the relatively well-conserved nucleotide sequences.

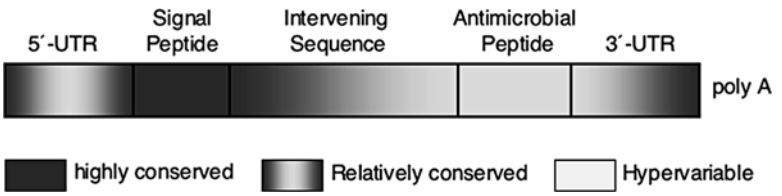


Fig. 13.1. A schematic drawing of the amphibian typical antimicrobial peptide precursor cDNAs. While the nucleotide sequences of the antimicrobial peptide region are hypervariable, those of the signal peptide region are highly conserved and the 5′-UTR, intervening sequence regions, and the 3′-UTR are relatively well conserved among different precursors.

These nucleotide sequence features of amphibian AMP precursor mRNAs are useful for molecular cloning of cDNAs encoding AMPs by 3′-rapid amplification of cDNA ends (3′-RACE) and reverse transcription-polymerase chain reaction (RT-PCR) methods using specific primers. In this section, an efficient protocol for molecular cloning of multiple genes encoding a probable AMP sequence from amphibian skin total RNA specimens will be introduced as part of a "peptidomics" strategy for exploration of bioactive peptides.

2. Materials

Unless otherwise stated, all experimental materials should be of biochemistry or molecular biology grade, which can be obtained easily from various commercial suppliers. In preparations, autoclaved "Milli-Q" (Millipore, Billercica, MA) grade ($\geq$18.0 MΩ·cm) distilled water (dH_2O) should be used (*see* **Note 1**).

2.1. RNA Extraction by Acid Guanidinium Isothiocyanate-Phenol-Chloroform (AGPC) Protocol

1. Modified Chomczynski and Sacchi (7) denaturing solution: 4 M guanidinium isothiocyanate (GTC), 0.1 M Tris-HCl, pH 7.5, 1% β-mercaptoethanol (β-ME) (*see* **Note 2**). This solution is stable for at least 1 month at 4°C.
2. Water-saturated phenol solution containing 0.1% (w/v) hydroxy quinoline. Store at 4°C in a dark bottle. Use the yellow (organic) phase but do not use if the colour has turned brown (*see* **Note 3**).
3. Tris-saturated phenol/chloroform/isoamyl alcohol (PCIA): A mixture of 0.1 M Tris (pH 8.0)-saturated phenol, containing 0.1% (w/v) hydroxyquinoline, with chloroform and isoamyl alcohol (25:24:1, v/v/v). Use the lower phase. Store at 4°C in a dark bottle.
4. Sodium acetate buffers: 2 M Sodium acetate, pH 4.0 and 3 M Sodium acetate, pH 5.2. Sterilize by autoclaving. Store at room temperature (*see* **Note 4**).
5. 8 M LiCl in water. Sterilize with 0.22-μm pore size filter. Store at 4°C.
6. CIA: Chloroform/isoamyl alcohol (49:1, v/v). Store at room temperature.
7. TE: 10 mM Tris-HCl, 1 mM EDTA, pH 8.0. Store at room temperature.
8. Blade edge homogenizer: e.g., Physcotron NS series (Microtec, Funabashi, Japan) or Polytron PT series (Kinematica, Lucerne, Switzerland) (*see* **Note 5**).

2.2. Synthetic DNA Oligonucleotides for PCR

1. Primers for 3′-RACE and RT-PCR: Designed to correspond to the nucleotide sequence of commonly conserved regions among amphibian AMP precursor cDNAs. These are available from a variety of commercial suppliers, e.g., Sigma Genosys (St. Louis, MO) or The Midland Certified Reagent Company (Midland, TX). For all primers a cartridge column or better purification should be selected. Examples of nucleotide sequences and combinations of primers that have been used successfully for cDNA amplification are shown in **Table 13.1**. Dissolve the oligonucleotides with dH_2O

Table 13.1
Combinations and nucleotide sequences of PCR primers that have been used successfully for amplification of frog skin AMP precursor cDNAs

Primers combination	AMP encoded	Frog species	Acc. No.	References
3′-RACE				
(F) 5′-ATGTTCA CCTTGAAGAA TC-3′ (ATG primer)	Temporin-1SKa	*Rana sakuraii*	AB275357	(11)
	Peptide VR23	*Rana sakuraii*	AB325526	(14)
(R) Oligo dT-3 sites adaptor primer (Takara)				
(F) 5′-GAWYYA YYHRAGCCYA AADATG-3′ (Degenerated primer)	Brevinin-1P	*Pelophylax plancyifukienensis*	AJ971789	(9)
	Brevinin-1S	*Odorrana schmackeri*	AJ971790	(9)
(R) 3′-RACE cDNA synthesis primers (Takara)	Brevinin-1 V	*Odorrana versabilis*	AJ971791	(9)
	Palustirn-1c	*Rana (Odorrana) versabilis*	AM113507	(10)
	Brevinin-1VEb	*Rana (Odorrana) versabilis*	AM113508	(10)
	Ranatuerin-2VEa	*Rana (Odorrana) versabilis*	AM113509	(10)
	Temporin-1VE	*Rana (Odorrana) versabilis*	AM113510	(10)
	Ranatuerin-2VEb	*Rana (Odorrana) versabilis*	AM113511	(10)
	Esculentin-2VEb	*Rana (Odorrana) versabilis*	AM113512	(10)
	Palustrin-3b	*Rana (Odorrana) versabilis*	AM113513	(10)
	Esculentin-1VEb	*Rana (Odorrana) versabilis*	AM113514	(10)
RT-PCR				
(F) 5′-ATGTTCAC CTTGAAGAATC-3′ (ATG)	Temporin-1TGa	*Rana tagoi*	AB219400	(12)
(R) 5′-AGATGATTT CCAATTCCAT-3′	Temporin-1TGb	*Rana tagoi*	AB219401	(12)
	Temporin-1Oa1	*Rana ornativentris*	AB274920	(13)
	Temporin-1Ob1	*Rana ornativentris*	AB274921	(13)
	Temporin-1Oc1	*Rana ornativentris*	AB274922	(13)
	Temporin-1Oe1	*Rana ornativentris*	AB274923	(13)
	Temporin-1SKc	*Rana sakuraii*	AB275358	(11)
	Temporin-1SKd	*Rana sakuraii*	AB275359	(11)
	Brevinin-1Ja	*Rana japonica*	AB373713	(15)

(F), forward primer; (R), reverse primer.

or TE buffer at appropriate concentrations (generally 50–100 μM). Dispense into aliquots, and store at –20°C.

2. Primers for nucleotide sequencing: T7 primer (5′-TAATACGACTCACTATAGGG-3′), SP6 primer (5′-ATTTAGGTGACACTATAG-3′).

2.3. cDNA Amplification by 3′-RACE

2.3.1. Reverse Transcription Reaction

1. Avian myeloblastosis virus (AMV)-derived reverse transcriptase XL (5 U/μL; Life Sciences Inc., St. Petersburg, FL). Store at –20°C.
2. RNase inhibitor (40 U/μL; Takara, Ohtsu, Japan). Store at –20°C (*see* **Note 6**).
3. Oligo dT-3 sites Adaptor Primer (Takara). This primer has oligo-dT region and three restriction sites of *Bam*HI, *Kpn*I, and *Xba*I. Store at –20°C.
4. The "3 sites" Adaptor Primer: 5′-CTGATCTAGAGGTACCGGATCC-3′ (20 μM) (Takara). Store at –20°C.
5. RNA PCR buffer (10×, Takara): 100 mM Tris-HCl, 500 mM KCl, pH 8.3. Store at –20°C.
6. Deoxyribonucleotides (dNTP) mixture (10×, Takara): 10 mM each of dNTP, ultrapure quality. Store at –20°C.

2.3.2. Polymerase Chain Reaction (PCR)

1. DNA polymerase: *Ex Taq* (5 U/μL, Takara). Store at –20°C.
2. *Ex Taq* buffer (Mg^{2+} plus) (10×). Store at –20°C.
3. 2.5 mM dNTP mixture. Store at –20°C.
4. Forward primer (20 μM). Store at –20°C.
5. "3 sites" Adaptor Primer (20 μM). Store at –20°C.

2.4. cDNA Amplification by RT-PCR

1. OneStep RT-PCR Enzyme Mix (Qiagen, Valencia, CA): Omniscript reverse transcriptase, Sensiscript reverse transcriptase, HotStart *Taq* DNA polymerase. Store at –20°C.
2. OneStep RT-PCR Buffer (5×) (Qiagen). Store at –20°C.
3. dNTP Mixture: 10 mM each of dNTP, ultrapure quality (Qiagen). Store at –20°C.
4. RNase inhibitor
5. Forward and reverse primers: Dilute to 30 μM with RNase-free H_2O as stock solutions. Store at –20°C.

2.5. Purification of the Amplified cDNA by Agarose Gel Electrophoresis

1. 1.5–2.0% agarose gel: Use Type II medium EEO agarose (Sigma) for conventional experiments or SeaKem GTG agarose (Takara) for gel purification.

2. TBE buffer (10×): 0.9 M Tris-borate, 20 mM EDTA. Or TAE buffer (10×): 0.4 M Tris-acetate, 10 mM EDTA. Sterilize by autoclaving. Store at room temperature.
3. Loading buffer (6×): 30% (v/v) glycerol, 0.25% (w/v) bromophenol blue (BPB) in dH_2O. Store at 4°C.
4. Molecular weight markers: 1 kb and 100 bp DNA ladder molecular weight markers (New England Biolabs, Ipswich, MA). Stable for at least 3 months at 4°C. A temperature of –20°C is recommended for long-term storage.
5. Gel staining solution: Dilute 10 mg/mL ethidium bromide (Merck, Darmstadt, Germany) to 0.5 μg/mL in TBE (*see* **Note 7**).
6. Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI): Membrane Binding Solution, Membrane Wash Solution, Wizard SV Minicolumns, Collection Tubes.
7. Pellet Paint Co-Precipitant (Novagen, Darmstadt, Germany): A visible dye-labelled carrier for alcohol precipitation of nucleic acids. Dispense into small aliquots (~50 μL) and store at –20°C.
8. A mini-gel electrophoresis system, such as Mupid series (Advance, Tokyo, Japan).

2.6. TA-Cloning

1. pSTBlue-1 AccepTor Vector (Novagen): 50 ng/μL, store at –20°C (*see* **Note 8**).
2. Clonables 2× Ligation Premix (Novagen). Store at –20°C.
3. Competent cells: *E. coli* JM109 competent cells ($>1\sim2\times10^9$ bacteria/mL; Takara). Store at –80°C.
4. Antibiotic (1000×): 100 mg/ml (w/v) Ampicillin (Wako, Osaka, Japan) in 50% ethanol. Store at –20°C.
5. Luria-Bertani (LB) broth: 1% Bacto-tryptone (Becton-Dickinson, Franklin Lakes, NJ), 0.5% Bacto-yeast extract (Becton-Dickinson), 1% NaCl (w/v) in dH_2O, pH 7.0–7.5. Sterilize by autoclaving. Store at room temperature. Add the antibiotic just before use.
6. LB plates: 1.5% (w/v) agar in LB containing the antibiotic. Inverse the plates and store at 4°C.
7. SOC medium: Add 2 mL of filter-sterilized 1 M glucose to 100 mL of SOB. No antibiotics. Dispense into aliquots (1–2 mL) and use once only. Store at –20°C. SOB medium: 2% Bacto-tryptone, 0.5% Bacto-yeast extract (Becton-Dickinson), 0.05% NaCl (w/v), pH 7.0. Sterilize by autoclaving. Add 5 mL of autoclaved 2 M $MgCl_2$ just before use.
8. PCR primers: T7 and SP6 primers (5 μM).

9. Ribonuclease A (RNase A) (optional for DNA purification): 1 mg/mL calf pancreatic RNase A (Sigma, Saint Louis, MO) in TE. Store in aliquots at –20°C.

2.7. Purification of the Plasmid DNA

1. LB broth containing the antibiotic (LB/amp$^+$).
2. Quantum Prep Plasmid Miniprep Kit (BioRad, Hercules, CA): Resuspension Solution, Lysis Solution, Neutralization Solution, Quantum Prep Matrix, Wash Buffer, Spin Filters. Store at room temperature.
3. Restriction enzyme: *Eco*RI (8∼12 U/μL, Takara). Store at –20°C (*see* **Note 9**).
4. 10× H buffer (Takara): 500 mM Tris-HCl, $MgCl_2$, 10 mM dithiothreitol (DTT), 1 M NaCl, pH 7.5. Store at –20°C.

2.8. Preparation of Cycle-Sequencing Samples (*see* Note 10)

1. BigDye Terminator v1.1 or v3.1 Cycle Sequencing Kits (Applied Biosystems): Ready Reaction Premix (2.5×), BigDye Sequencing Buffer (5×)
2. Sequence primer: T7 and SP6 primers (3.2 pmol)
3. Template DNA: 150–300 ng

2.9. Nucleotide and Amino Acid Sequences Analysis

1. Computer software for nucleotide and amino acid sequence analysis, such as *Genetyx* (Software Development Corporation, Osaka, Japan)

3. Methods

3.1. Approaches to the Design of PCR Primers

The design of nucleotide sequences for use as PCR primers is a key point for the success of cDNA amplification of amphibian AMP precursors. It is common, but not universal, in many frog genera that the N-terminal signal peptide region of amphibian AMP precursors is comprised of 22 amino acid residues and terminated by a Cys residue (6, 8). Therefore, a forward primer sequence based on this region might be suitable for the amplification of AMP precursor cDNAs. An example of a nucleotide sequence for the forward primer is 5′-ATGTTCACCTTGAAGAAATC-3′, designated here as the "ATG primer." A combination of the ATG primer and oligo-(dT) primer in 3′-RACE has a high probability of successful amplification of AMP precursor cDNAs (**Table 13.1**). A "degenerate" primer, 5′-GAWYYAYYHRAGCCYAAADATG-3′ (W = A+T, Y = C+T, H = A+T+C, R = A+G, D = A+T+G), designed according to a conserved region at the 5′-UTR in front of the signal peptide coding region of frog AMPs works well as a forward primer in 3′-RACE (**Table 13.1**) (9, 10). Specific reverse

primers designed according to the nucleotide sequence within the 3′-UTR of amphibian AMP precursor cDNAs might be used to avoid amplification of nonspecific cDNAs (**Table 13.1**), but the possibility of finding cDNAs encoding novel AMPs will be reduced. Although the nucleotide sequences of the 3′-UTR of amphibian AMP precursors are relatively well conserved among several AMP groups, they are not very long and tend to possess inverted repeats (11–15). Thus, the reverse primer should be designed such that its nucleotide sequence will not form primer dimers or hairpin or stem-loop structures.

3.2. Total RNA Extraction from Frog Skin Samples by AGPC Protocol

1. Immerse adult frog(s) in ice-cold water until anaesthetized, then sacrifice by decapitation (*see* **Note 11**). Remove the skin and mince on ice immediately, or freeze on dry ice or in liquid nitrogen and store at –80°C until use. Experiments should be carried out according to the guidelines of the appropriate animal ethics committees.
2. Place 0.5–1 g of the skin sample into a 50-mL Falcon tube and add 10 mL of denaturing solution per 1 g of sample. Agitate by inversion on a rotator at 4°C overnight (*see* **Note 12**).
3. Homogenize the samples with a blade edge homogenizer on ice. If multiple samples are processed, rinse the probe of the homogenizer with dH_2O prior to each use.
4. Spin the homogenate at 10,000×*g* for 20 min at 4°C and transfer the supernatant to a new tube. Add 1 volume of water-saturated phenol, 0.1 volume of 2 M sodium acetate, pH 4.0, and 0.25 volumes of CIA to the supernatant, with vigorous mixing after the addition of each reagent. Leave the sample on ice for 15 min.
5. Spin the sample at 10,000×*g* for 15 min at 4°C. The suspension will be separated into the aqueous phase, interphase, and phenol phases. Transfer the upper aqueous phase to a new tube, taking care not to aspirate the interphase. Add 1 volume of isopropanol and mix vigorously, then place at –20°C for at least 1 h to precipitate RNA (*see* **Note 13**).
6. Centrifuge the suspension at 10,000×*g* for 30 min at 4°C to precipitate RNA, the colour of which may be dark brown or black due to contamination by dyes from pigment cells, which are difficult to remove. Dissolve the resulting crude RNA pellet in 400 μL of dH_2O, transfer to a pre-chilled 1.5-mL microcentrifuge tube, add 400 μL of 8 M LiCl, and place at 4°C overnight (*see* **Note 14**).
7. Spin at 16,000×*g* for 30 min at 4°C. Add 200 μL dH_2O to the pellet, and then dissolve completely by pipetting. Add 200 μL of PCIA, vortex vigorously, and spin at 16,000×*g*

for 5 min at 4°C. Transfer the aqueous phase to a 1.5-mL microcentrifuge tube, and precipitate the RNA sample by a conventional ethanol precipitation procedure (*see* **Note 15**).

8. Dissolve the RNA pellet in the appropriate volume (50–200 μL) of RNase-free dH_2O or TE. Prepare serially diluted (50–400 fold) RNA solutions and measure the absorbance at 260 and 280 nm for the quantification and quality check of the specimens. Prepare 1 μg/μL and 100 ng/μL RNA dilutions in dH_2O for subsequent cDNA synthesis and amplification, dispense into aliquots, and store at –80°C (*see* **Note 16**).

3.3. cDNA Amplification by 3′-RACE

3.3.1. Reverse Transcription

1. Prepare a reaction mixture in a pre-chilled 0.2-mL PCR tube by combining the following reagents: 2 μL of 10× RNA PCR Buffer, 4 μL of 25 mM $MgCl_2$, 2 μL of 10× dNTP mixture, 1 μL of AMV Reverse Transcriptase XL, 0.5 μL of RNase Inhibitor, 1 μL of Oligo dT-3 sites Adaptor Primer, 1 μL of skin total RNA (1 μg), and 8.5 μL of RNase-free dH_2O (total volume: 20 μL).
2. Place the tube in a Thermal Cycler and set the parameters under the following conditions: 30°C for 10 min, 50°C for 30 min, 95°C for 5 min, and 5°C for 5 min, then store at 4°C (*see* **Note 17**).

3.3.2. PCR

1. Spin the reverse transcription tube briefly, then add the following reagents directly to the tube: 10 μL of 10× *Ex Taq* buffer, 16 μL of 2.5 mM dNTP mixture, 0.5 μL of *Ex Taq* DNA polymerase, 1 μL of forward primer, 1 μL of 3 sites Adaptor Primer, 20 μL of the obtained reverse transcription reaction product, and 51.5 μL of sterilized dH_2O (total volume: 100 μL). A negative control (without template RNA) should be included in each amplification.
2. Place the tube in a Thermal Cycler and perform DNA amplification under the following conditions: pre-heating for 15 min at 72°C for hot start, followed by 30 cycles of 30 s at 94°C, 30 s at 50°C, 2 min at 72°C, with a final extension step of 7 min at 72°C, and then store at 4°C. The reaction products may be stored at 4°C for up to 1 week in the tubes. For long-term storage, precipitation by a conventional ethanol precipitation protocol is recommended.

3.4. cDNA Amplification by RT-PCR

1. Prepare a reaction mixture in a pre-chilled 0.2-mL PCR tube by combining the following reagents: 1 μL of the total RNA template (100 ng/μL), 10 μL of 5× OneStep PCR buffer, 2 μL of 10 mM dNTP mixture, 1 μL of forward primer (30 μM), 1 μL of reverse primer (30 μM), 2 μL of OneStep RT-PCR Enzyme Mix, 0.5 μL of RNase inhibitor, and

32.5 μL of RNase-free dH_2O (total volume: 50 μL). A negative control (without template RNA) should be included in each amplification.

2. Place the tubes in a Thermal Cycler and perform cDNA amplification under the following conditions: 30 min at 50°C for reverse transcription, 15 min at 95°C for denaturation of the reverse transcriptase, 5 min at 94°C for denaturation of the DNA, followed by 35 cycles of 30 s at 94°C, 30 s at 50°C, 2 min at 72°C, with a final extension step of 7 min at 72°C.

3.5. Purification of the Amplified cDNA by Agarose Gel Electrophoresis

1. Prepare a 1.5–2.0% agarose mini gel using SeaKem GTG agarose and 1× TBE (or 1× TAE) running buffer. Take 10-μL aliquots of the 3′-RACE or RT-PCR product and add 2 μL of sample loading buffer. Load samples and the molecular weight marker(s) onto the gel; leave empty lanes between the samples to avoid cross-contamination. Run electrophoresis at a constant voltage of 100 V.
2. At the end of the run, soak the gel in ethidium bromide solution (0.5 μg/mL) for 30 min with gentle shaking, then visualize and photograph the DNA on a UV-transilluminator (*see* **Note 18**).
3. Excise the DNA fragment of interest in a minimum amount of agarose using a clean scalpel or razor blade and place the gel slice in a 1.5-mL microcentrifuge tube. The gel slice may be stored at 4°C or –20°C for up to 1 week in a tightly closed tube under nuclease-free conditions.
4. Add 10 μL of Membrane Binding Solution per 10 mg of gel slice, vortex, and incubate at 55°C until the gel slice has melted completely. Vortex the tube every 2–3 min during the incubation. Transfer the dissolved gel mixture to an SV Minicolumn assembly inserted into a Collection Tube and incubate at room temperature for 1 min. Spin at 16,000×*g* for 1 min, discard the flow-through fraction, and reassemble the column.
5. Add 700 μL of Membrane Wash Solution to the column assembly, spin at 16,000×*g* for 1 min, and discard the flow-through fraction. Repeat the step with 500 μL of Membrane Wash Solution and spin for 5 min. Leave the column assembly for a few minutes to allow residual ethanol in the Membrane Wash Solution to evaporate.
6. Place the Minicolumn in a clean 1.5-mL microcentrifuge tube, add 100 μL of sterile dH_2O, incubate for 1 min at room temperature, and spin at 16,000×*g* for 1 min. Repeat this step once more. Add 1–2 μL of Pellet Paint to the eluate, 0.1 volume of 3 M sodium acetate, and 2 volumes of

ethanol, vortex vigorously, and spin at 16,000×*g* for 5 min at room temperature. Discard the supernatant and rinse the pink-coloured precipitate with 100 μL of 70% ethanol. Spin briefly, discard the supernatant, and vacuum dry the DNA. Resuspend the DNA in 10–15 μL of TE and store at –20°C (*see* **Note 19**).

3.6. TA-Cloning

3.6.1. Ligation Reaction

1. Prepare a reaction mixture in a pre-chilled 0.2-mL PCR tube by combining the following reagents: 1 μL of AccepTor Vector (50 ng/μL), 2.5 μL of Minicolumn-purified DNA, and 3 μL of Clonables 2× Ligation Premix (total volume: 6 μL). Prepare a negative control (without the insert DNA). Incubate the reaction at 16°C for at least 30 min (preferably 2 h). Store the reaction products at –20°C (*see* **Notes 20 and 21**).
2. Add 1 μL of the reaction products directly to 30 μL of fresh competent JM109 *E. coli* cells in a sterile pre-chilled 1.5-mL microcentrifuge tube and stir by gentle tapping. Place the tube on ice for 30 min, heat the tube for 30–45 s in a water bath at 42°C, and chill on ice immediately (*see* **Note 22**).
3. After 3 min, add 300 μL of SOC medium and incubate at 37°C for 1 h in a water bath. Plate the transformation mixture on a well-dried LB-agar/amp$^+$ plate, spread gently, and allow to be adsorbed. Invert the plate and incubate overnight at 37°C (15–18 h). SOC medium and LB-agar/amp+ plates should be warmed at room temperature prior to use.

3.6.2. Rapid Screening by Colony PCR

1. Choose colonies that are at least 1 mm in diameter. Pick a single colony from the agar plate using a sterile toothpick. The colonies should be marked and numbered in ink on the bottom of the plate to allow correspondence with the results of screening, and re-incubated at 37°C until visible growth can be seen.
2. Transfer the bacteria to a 0.6-mL sterile microcentrifuge tube containing 50 μL of sterile dH_2O and vortex well. Heat at 100°C for 5 min to lyse the cells and denature DNAs. Spin at 16,000×*g* for 1 min to remove cell debris and transfer 1 μL of the supernatant to a pre-chilled 0.2-mL PCR tube.
3. Add the following reagents directly to the PCR tube: 5 μL of 10× *Ex Taq* buffer, 10 μL of 2.5 mM dNTP mixture, 1 μL of T7 primer, 1 μL of SP6 primer, 31.5 μL of sterilized dH_2O, and 0.5 μL of *Ex Taq* DNA polymerase (total volume: 50 μL) (*see* **Note 23**).
4. Place the tube in a Thermal Cycler and perform DNA amplification under the following conditions: pre-heat for 5 min

at 94°C, followed by 30 cycles of 30 s at 94°C, 30 s at 55°C, 2 min at 72°C, with a final extension step of 7 min at 72°C (*see* **Note 24**).

5. Transfer 5-μL aliquots of the PCR product to microcentrifuge tube, add 1 μL of 10× H or equivalent buffer, 1 μL of *Eco*RI, and 3 μL of sterile dH_2O. Incubate at 37°C for at least 30 min.
6. Add 2 μL of loading buffer and separate the products by electrophoresis on 1.5–2.0% agarose gels. Visualize and photograph the DNA as described in **Section 3.5** . Compare the sizes of DNA on the gel and select colonies if they have DNA of different sizes. The corresponding colonies on the agar plate will be subjected to plasmid purification (*see* **Note 25**).

3.7. Purification of Plasmid DNA

1. Pick the corresponding colonies subjected to rapid screening from the agar plate and inoculate in 2 mL of LB/amp+ broth in a 15-mL conical tube with a loosened cap. Shake at 200 rpm at 37°C for at least 6 h (preferably overnight but use 4–5 mL LB broth in this case). Prepare multiple samples.
2. Transfer the culture (~1.4 mL) to a 1.5-mL microcentrifuge tube, pellet the cells by centrifugation for 30 s at 14,000×*g*, and remove all the supernatant. Add 200 μL of the Cell Resuspension Solution and vortex vigorously until the pellet is completely resuspended. Add 250 μL of the Cell Lysis Solution and mix by gently inverting the tube ten times (do not vortex). Add 250 μL of Neutralization Solution and mix in the same way. Spin at 14,000×*g* for 5 min at room temperature. A compact pellet of white debris will form along the sides and bottom of the tube. Transfer the supernatant to a Spin Filter inserted into a 2-mL wash tube taking care not to aspirate the debris.
3. Add 200 μL of thoroughly suspended Quantum Prep Matrix, mix completely by gentle pipetting, and spin at 14,000×*g* for 30 s to pull fluid through the column. Remove the Spin Filter, discard the filtrate, and replace the column in the wash tube. Add 500 μL of Wash Buffer and spin at 14,000×*g* for 30 s.
4. Repeat the previous step once again but increase the time for centrifugation to 2 min. Remove the Spin Filter into a clean 1.5-mL microtube, and add 100 μL of sterile dH_2O. Heat at 70°C for a few minutes and elute the DNA by centrifugation at 14,000×*g* for 1 min at room temperature. Store at –20°C.

3.8. Preparation of Cycle-Sequencing Samples

1. Prepare two reaction mixtures in pre-chilled 0.2-mL PCR tubes by combining the following reagents: 6 μL of the purified plasmid DNA from **Section 3.6** above, 8 μL of

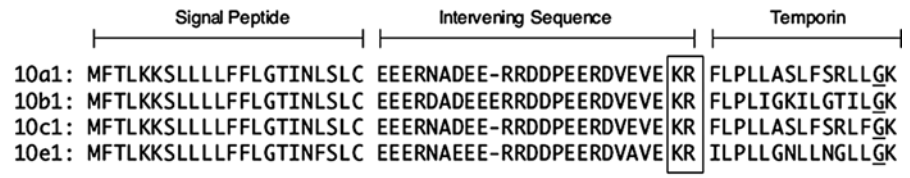
Terminator Ready Reaction Mix, 3.2 pmol of T7 primer for one tube or 3.2 pmol of SP6 primer for the remaining tube, and sterile dH_2O to a total volume of 20 μL. Mix well and spin briefly (*see* **Note 26**).

2. Place the tubes in a Thermal Cycler and run the reaction under the following conditions: 25 cycles of 10 s at 96°C, 5 s at 50°C, and 4 min at 60°C, and then 4°C for 10 min. Samples can be stored for a few days at 4°C or for several days at –20°C (*see* **Note 27**).

3.9. Sequence Analysis to Find Probable AMPs

1. Translate the obtained nucleotide sequences to amino acid sequences using a computer program such as *Genetyx*, then perform queries with these sequences against GenBank using BLAST (Basic Local Alignment Search Tool; http://blast.ncbi.nlm.nih.gov/Blast.cgi). If the query sequences show high similarity to the amphibian AMP precursors, in particular within the regions of signal peptide and intervening sequence, try to find di-basic amino residue

A)



B)

Temporin	Net Charge	Antimicrobial activities		Reference
		E. coli	*S. aureus*	
10a: FLPLLASLFSRLL.NH2	(+2)	weak	strong	17
10b: FLPLIGKILGTIL.NH2	(+2)	weak	strong	17
10c: FLPLLASLFSRLF.NH2	(+2)	NA	strong	17
10e: ILPLLGNLLNGLL.NH2	(+1)	NA	NA	13

Fig. 13.2. Comparisons of the deduced amino acid sequences of *R. ornativentris* preprotemporin −10a1, −10b1, −10c1, and −10e1 (**A**) and the primary structures and antimicrobial activities of the mature temporins (**B**). In panel **A**, the *box* shows the proposed processing enzyme cleavage site consisting of di-basic amino acid residues. The Gly residues at the C-terminus of temporin domains, which may function as nitrogen donors for C-terminal amidation, are underlined. In panel **B**, the positions of the functional groups providing a net positive charge at physiological pH are underlined. The peptides were incubated in Mueller-Hinton broth (100 μL) with an inoculum (10 μL of 5×10^5 colony forming units/mL) from a log-phase culture of either *Staphylococcus aureus* (*S. aureus*) or *Escherichia coli* (*E. coli*) in 1% BSA-coated 96-well microtiter cell culture plates for 18 h at 37°C in air. At the end of the incubation, the absorbance at 630 nm of each well was determined using a microtiter plate reader. Minimum inhibitory concentrations (MICs) of the peptide against these microorganisms were determined by the standard microdilution method (18). NA, not active at a concentration below 150 μM.

sites (KK, KR, RK, or RR) within the sequence. These sites may be recognized by the processing enzymes and the subsequent C-terminal sequence in the precursors may be an AMP.

2. Check the C-terminal amino acid sequence of the peptides. If a Gly residue is present at the C-terminus, it may function as a nitrogen donor (**Fig. 13.2**) and the C-terminus of the peptide will be an amide. The C-terminal amide provides +1 net positive charge as well as a basic amino acid residue, Lys or Arg. Calculate the net positive charge at physiological pH. An acidic amino acid residue (Asp or Glu) within the peptide neutralizes a net positive charge. If the total value of the net charge is higher than +2, the peptide is more likely to be "antimicrobial" (**Fig. 13.2**).
3. Predict probable secondary structures of the peptides and draw a helical wheel diagram using Genetyx software (*see* **Note 28**). If the peptides are cationic and have an amphipathic helical conformation, these represent good agreement with typical features of AMPs and it may be worthwhile preparing synthetic replicates for antimicrobial assays.

4. Notes

1. Water treated with diethyl pyrocarbonate (DEPC; note it is harmful) to remove nucleases is often used in molecular biology experiments. However, fresh high-performance liquid chromatography (HPLC)-grade H_2O can also be used as nuclease-free water.
2. GTC is hazardous. Laboratory gloves, mask, coat, and safety goggles should be worn when handling the denaturing solution.
3. Phenol is highly corrosive and causes severe burns. Laboratory gloves, coat, and safety goggles should be worn when handling phenol. Rinse with a large amount of water and wash with soap if the solution comes into contact with skin.
4. A relatively large amount of acetic acid is required to adjust the pH of the solution to 4.0. Care should be taken to avoid adding too much acetic acid at once; otherwise, the pH of the solution will drop below pH 4.0.
5. A conventional sonicator (e.g., Sonifier, Branson, Danbury, CT) can also be used. Sonication should be performed on ice otherwise the temperature of the suspension will become very high. Check the colour of the extraction

buffer every 2–3 min and stop sonication when the colour turns dark brown or black. Avoid long-term sonication to avoid fragmenting the DNA and RNA in the tissues. Ear protectors should be used when operating a sonicator.

6. Enzymes and RNase inhibitors for molecular biology experiments should be used immediately. A portable bio-cooler that can keep such reagents at –15°C on a laboratory bench for about 1–1.5 h might be useful.
7. Ethidium bromide is a nucleic acid intercalating agent and it is a strong mutagen. Laboratory gloves and goggles should be worn when handling the agent. To reduce hazardous waste in the laboratory, a cartridge system filtering the material, such as Fluor/Away(TM) (Triangle Biomedical Sciences, Durham, NC), is recommended.
8. PCR products generated by DNA polymerases such as KOD XL polymerase and native and recombinant *Taq* polymerases leave single 3′-dA overhangs. Thus, linearized plasmid vectors containing single 3′-dT DNA ends are compatible with direct ligation of these 3′-dA overhanging products. The vector posseses with duel opposed T7 and SP6 promoters available for nucleotide sequencing from both directions.
9. There are two *Eco*RI sites in the pSTBlue-1 plasmid vector, one in front and another at the rear of the inserted DNA. Thus, the inserted DNA can be obtained by *Eco*RI digestion unless *Eco*RI recognition sites are present within the inserted DNA.
10. A DNA sequencer is an expensive piece of equipment. Reagents for sequencing reactions are also expensive. Many commercial suppliers provide various types of custom DNA sequencing service at a reasonable cost. In most cases it is more cost-effective to use external sequencing services, which accept purified DNA samples after cycle-sequence reactions.
11. Adult frog skin should be used. Larval amphibian skin is known to very rarely produce AMPs (12, 14, 16).
12. Frog skin is very tough and so the process will contribute to homogenization. A sufficient amount of total RNA for the PCR experiments may be extracted from 1 g of skin samples after two overnight incubations with agitation in the denaturing solution.
13. RNA is present in the aqueous phase, whereas proteins and DNA are present in the interphase. If the sample does not separate into aqueous and phenol phases, add a little more CIA to the suspension, vortex vigorously, incubate on ice, and spin again.

14. DNA is soluble in 4 M LiCl, whereas RNA is precipitated in the solution.
15. Conventional ethanol precipitation procedure: Add 0.1 volume of 3 M sodium acetate (pH 5.2) and 2 volumes of ethanol to the aqueous phase of the PCIA extract, vortex vigorously, and place at –80°C for 30 min. Spin at 16,000×*g* for 15 min at 4°C. Discard the supernatant, rinse the pellet with 400 μL of 70% ethanol, sediment again by centrifugation at 16,000×*g* for 5 min at 4°C, and vacuum dry the pellet.
16. A_{260}/A_{280} ratio for pure RNA is generally expected to range from 1.8 to 2.0. However, in the case of frog skin total RNA, the values will be relatively lower (~1.6–1.8). This may be due to contamination by dyes from pigment cells in their skin, but this does not adversely affect subsequent experiments.
17. Due to the small volume of PCR mixtures, use a mineral oil-free type Thermal Cycler throughout this protocol to save time and effort of the oil removal.
18. Use a long-wavelength UV lamp to reduce nicking. A face protector should be worn to avoid irradiation. Amplified fragments of approximately 200–700 bp can be observed. Check the DNA fragments carefully. In a few cases, particularly in RT-PCR products, they appear as doublet bands. In this case, carefully separate the bands. If DNA bands are not detected, decrease the annealing temperature to around 40–45°C.
19. Pellet Paint is known to contribute to the A_{260}. If the DNA is quantified from the A_{260}, a water blank containing the same ratio of Pellet Paint should also be prepared.
20. The composition of the reagents in this protocol is optimized for the cloning of amphibian AMP cDNA. For a standard reaction and according to the manufacturer's protocol, 1 μL of AccepTor Vector is ligated with 0.15 pmol of amplified product (50 ng of a 500-bp fragment) in a total volume of 10 μL.
21. To perform the reaction at 16°C, place a conventional water bath in a cold room or a chromatochamber and set the temperature to 16°C. Alternatively, a Thermal Cycler can also be used for incubating samples at 16°C.
22. Thaw the cells just before use and keep on ice whenever possible. The cells can be refrozen in dry ice-ethanol and reused; however, this may reduce the transformation efficiencies by more than tenfold.

23. In TA-cloning, the direction of the inserted DNA fragment in the vector is unknown. If two sets of primers consisting of one insert-derived sequence, such as the ATG primer, and one sequence primer (T7 or SP6 primers) are used for screening, the direction can be determined from the presence or absence of the amplified band on agarose gel electrophoresis.
24. If checking for the DNA amplification only, take a 5-μL aliquot of the PCR products to a microcentrifuge tube, add 1 μL of the sample loading buffer and load onto a 1.5–2.0% agarose gel for electrophoresis. To save time, a gel containing ethidium bromide is recommended.
25. The rapid screening step can be skipped. A standard miniprep procedure is also useful for screening. In most cases, a fragment of the insert DNA may be obtained by *Eco*RI digestion but use at least 40% overnight culture broth of the inoculated colonies to obtain the amounts of small DNA fragments sufficient for visualization on a gel. RNase A solution should be added to the miniprep samples (0.5 μL per sample) to degrade ribosomal RNA.
26. The samples may be shipped for sequencing to commercial suppliers at 4°C. The purified plasmids are often accepted by such services without the need for Step 2 (in **Section 3.7**) which may be omitted.
27. Before nucleotide sequencing analysis, the remaining components, such as unincorporated dyes, salt ions, and dNTPs remaining after the cycle-sequencing reactions, should be removed. These are likely to interfere with DNA sequencing. Standard ethanol precipitation is the simplest protocol for this purpose. A microspin column filled with a matrix for gel filtration chromatography (e.g., Sephadex G-50; GE Healthcare) may also be used. BigDye(R) XTerminator(TM) Purification (Applied Biosystems) is strongly recommended by the manufacturer.
28. The *PredictProtein* program (http://www.predictprotein.org/) is also useful for predicting secondary structures of the peptides.

Acknowledgments

The authors wish to thank the members in the laboratory of Regulatory Biology in Toho University for their useful suggestions for the preparation of the manuscript. This work was supported in part by a Grant-in-Aid for Scientific Research (19570063) from the Japan Society for the Promotion of Science to S.I. and T.K.

References

1. Boman, H. G. (1995) Peptide antibiotics and their role in innate immunity. *Annu. Rev. Immunol.* **13**, 61–92.
2. Ganz, T. and Lehrer, R. I. (1999) Antibiotic peptides from higher eukaryotes: biology and applications. *Mol. Med. Today* **5**, 292–297.
3. Hancock, R. E. and Diamond, G. (2000) The role of cationic antimicrobial peptides in innate host defences. *Trends Microbiol.* **8**, 402–410.
4. Zasloff, M. (2002) Antimicrobial peptides of multicellular organisms. *Nature* **415**, 389–395.
5. Radek, K. and Gallo, R. (2007) Antimicrobial peptides: natural effectors of the innate immune system. *Semin. Immunopathol.* **29**, 27–43.
6. Nicolas, P., Vanhoye, D. and Amiche, M. (2003) Molecular strategies in biological evolution of antimicrobial peptides. *Peptides* **24**, 1669–1680.
7. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanated-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156–159.
8. Duda, T. F., Jr., Vanhoye, D. and Nicoras, P. (2002) Roles of diversifying selection and coordinated evolution in the evolution of amphibian antimicrobial peptides. *Mol. Biol. Evol.* **19**, 858–864.
9. Chen, T., Li, L., Zhou, M., Rao, P., Walker, B. and Shaw, C. (2006) Amphibian skin peptides and their corresponding cDNAs from single lyophilized secretion samples: identification of novel brevinins from three species of Chinese frogs. *Peptides* **27**, 42–48.
10. Chen, T., Zhou, M., Rao, P., Walker, B. and Shaw, C. (2006) The Chinese bamboo leaf odorous frog (*Rana (Odorana) versabilis*) and North American *Rana* frogs share the same families of skin antimicrobial peptides. *Peptides* **27**, 1738–1744.
11. Suzuki, H., Iwamuro, S., Ohnuma, A., Coquet, L., Leprince, J., Jouenne, T., Vaudry, H., Taylor, C. K., Abel, P. W. and Conlon, J. M. (2007) Expression of genes encoding antimicrobial and bradykinin-related peptides in skin of the stream brown frog *Rana sakuraii*. *Peptides* **28**, 505–514.
12. Ohnuma, A., Conlon, J. M., Kawasaki, H. and Iwamuro, S. (2006) Developmental and triiodothyronine-induced expression of genes encoding preprotemporins in the skin of Tago's brown frog *Rana tagoi*. *Gen. Comp. Endocrinol.* **146**, 242–250.
13. Ohnuma, A., Conlon, J. M., Yamaguchi, K., Kawasaki, H., Coquet, L., Leprince, J., Jouenne, T., Vaudry, H. and Iwamuro, S. (2007) Antimicrobial peptides from the skin of the Japanese mountain brown frog *Rana ornativentris*: evidence for polymorphism among preprotemporin mRNAs. *Peptides* **28**, 524–532.
14. Suzuki, H., Conlon, J. M. and Iwamuro, S. (2007) Evidence that the genes encoding the melittin-related peptides in the skins of the Japanese frogs *Rana sakuraii* and *Rana tagoi* are not orthologous to bee venom melittin genes: developmental- and tissue-dependent gene expression. *Peptides* **28**, 2061–2068.
15. Koyama, T. and Iwamuro, S. (2008) Molecular cloning of a cDNA encoding atypical antimicrobial and cytotoxic brevinin-1Ja from the skin of the Japanese brown frog, *Rana japonica*. *Zool. Sci.* **25**, 487–491.
16. Reilly, D. S., Tomassini, N. and Zasloff, M. (1994) Expression of magainin antimicrobial peptide genes in the developing granular glands of *Xenopus* skin and induction by thyroid hormone. *Dev. Biol.* **162**, 123–133.
17. Kim, J. B., Iwamuro, S., Knoop, F. C. and Conlon, J. M. (2001) Antimicrobial peptides from the skin of the Japanese mountain brown frog, *Rana ornativentris*. *J. Peptide Res.* **58**, 349–356.
18. Barchiesi, F., Colombo, A. L., McGough, D. A. and Rinaldi, M. G. (1994) Comparative study of broth macrodilution and microdilution techniques for in vitro antifungal susceptibility testing of yeasts by using the National Committee for Clinical Laboratory Standards' proposed standards. *J. Clin. Microbiol.* **32**, 2494–2500.

Chapter 14

Combined Peptidomics and Genomics Approach to the Isolation of Amphibian Antimicrobial Peptides

Ren Lai

Abstract

A large number of diverse antimicrobial peptides have been found in amphibian skins, and many more remain to be identified. It is sufficiently easy to obtain amounts of gland secretions sufficient for both identification and functional testing of the bioactive peptides. We describe here a systematic peptidomics approach which we combined with genomics and functional testing. This has proven to be an effective way to identify amphibian antimicrobial peptides, including novel peptide families. Protocols are exemplified for *Bombina maxima* and *Odorrana grahami* and can be easily adapted for use with other amphibian species.

Key words: Amphibian, antimicrobial peptides, peptidomics, genomics, skin secretion.

1. Introduction

1.1. Antimicrobial Peptides Are Attractive Candidates as New Anti-infective Agents

A large amount of drug-resistant microorganisms are emerging because of abused "conventional" anti-infective agents and this leads to a serious problem of resistance to several of the therapeutically used antibiotics (1). The growing problem of microorganism resistance to conventional antibiotics has urged development of new human therapeutics. The gene-encoded antimicrobial peptides (AMPs) play an important role in innate immunity against noxious microorganisms. Endogenous antimicrobial peptides have become recognized as important, ubiquitous and ancient contributors to the innate mechanisms that permit animals and plants to resist infection (2). Since they cause much less drug resistance of microbes than conventional antibiotics, AMPs nowadays attract considerable interest for the development of new anti-infective agents (1–3).

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_14,

1.2. Amphibian Skins Are Rich Resources for Antimicrobial Peptides

Much attention has been paid to amphibian peptides for their wide-ranging pharmacological properties, clinical potential and gene-encoded origin (3, 4). Some amphibian peptides have been extensively studied, such as antimicrobial peptides bombesins and bradykinins. Granular glands in the skin of anuran amphibians, particularly those belonging to the families *Pipidae*, *Hylidae*, *Hyperoliidae*, *Pseudidae* and *Ranidae*, synthesize and secrete a remarkably diverse array of antimicrobial peptides. These are typically 10–50 residues long and are released onto the outer layer of the skin to provide an effective and fast-acting defence against harmful microorganisms (5–11). More than 300 antimicrobial peptides have been identified from amphibians and many more are expected to be identified in the future. Antimicrobial peptides from ranid frogs are suggested as taxonomic and phylogenetic markers and potential sources of new therapeutic agents (5–11).

1.3. Major Problems in Working on Amphibian Antimicrobial Peptides

Most of bioactive compounds including antimicrobial peptides exist in amphibian granular gland secretions. It is not always possible to obtain the amounts of gland secretions sufficient for the identification and functional testing of bioactive compounds. In some cases, many hundreds, even thousands of amphibians have been sacrificed for extracting sufficient quantities of gland secretions for biochemical or pharmacological prospecting (4, 12–14). New, more humane methods have to be developed to avoid killing and skinning amphibians.

1.4. Genomics Approaches to the Identification of Antimicrobial Peptide-Like Peptides

Many antimicrobial peptides cannot be easily isolated from amphibian skin secretions and characterized because of their low abundance, occasionally unusual anionic character and often only weak antimicrobial activity which makes their functional characterization more difficult. On the other hand, known amphibian antimicrobial peptide precursors have highly conserved signal peptide sequences (**Tables 14.1** and **14.2**). For example, the ranid antimicrobial peptides have a common N-terminal precursor sequence, which is highly conserved even between different species. The conserved precursor sequence comprises a hydrophobic signal peptide of 22 residues followed by a 16–25-residue acidic fragment which is terminated by a typical processing signal Lys-Arg (5–11, 15). The existence of highly conserved signal peptide sequences provides an alternative approach to the discovery of novel antimicrobial peptides through molecular biology rather than proteomic approaches. It is relatively easy to find sequence regions suitable for the design of DNA primers for screening cDNA libraries.

Table 14.1
Sequence comparison of antimicrobial precursors from amphibians of Ranid family

Precursors	Sequences
Gaegurin–4	MFTMKKSLLFLFFLGTISLSLCEEERSADEDDGGEMTEEEVKRG ILDTLKQFAKGVGKDLVKGAAQGVLSTVSCKLAKTC
Brevinin–1E	MFTLKKSMLLLFFLGTINLSLCEEERDADEEERRDNPD ESEVEVEKRFLPLLAGLAANFLPKIFCKITRKC
Palustrin 1C	MFTMKKSLLLLFFLGTISLSLCEEERGADEEEGDGEKL TKRALSILRGLEKLAKMGIALTNCKATKKC
Brevinin–2GHC	MFTMKKSLLLLFFLGMISLSLCEQERGADEDEGEVEEQIKRSI WEGIKNAGKGFLVSILDKVRCKVAGGCNP
Ranalexin	MFTLKKSLLLLFFLGTINLSLCEEERNAEEERRDNPDERDV EVEKRFLGGLIKIVPAMICAVTKKC
Esculentin–1B	MFTLKKPLLLIVLLGMISLSLCEQERNADEEEGSEIKRGIFS KLAGKKLKNLLISGLKNVGKEVGMDVVRTGIDIAGCKIKGEC
Esculentin–2S	MFTLKKSLLVLFFLGTISLSLCEQERAADDEDNGEVEEVKRGL FTLIKGAVKMIGKTVAKEAGKTGLELMACKVTNQC
Ranatuerin–2VA	MFTLKKSFLLLFFLGTITLSLCEQERGADEDDGVEMTEEEVKRG LLDTIKNTAKNLAVGLLDKIKCKMTGC
Palustrin–OG1	MFTMKKSPLVLFFLGIVSLSLCQEERSADDEEGEVIEEEVKRGF WDTIKQAGKKFFLNVLDKIRCKVAGGCRT
Temporin B	MFTLKKSLLLLFFLGTINLSLCEEERNAEEERRDEPDER DVQVEKRLLPIVGNLLKSLLGK
Nigrocin–2	MFTLKKSLLLLFFLGMVSLALCEQERDANEEERRDELDER DVEAIKRGLLSKVLGVGKKVLCGVSGLC
Odorranain–A1	MFTMKKSLLLLFFLGTISLSLC–EQERDADEE–––– EGSENGAEDIKLNRVVKCSYRLGSPDSRCN
Odorranain–B1	MFTLKKSLLLLFFLGIISLSFR–EQERDADED–––– DGGEVTGEEVKRAALKGCWTKSIPPKPCFGKR
Odorranain–C6	MFTMKKSLLLLLFLGTISLSLC–EEERDADEE–––– EG–EMTEEEVKRGVLGTVKNLLIGAGKSAAQSVLKTLSCKLSNDC
Odorranain–F1	MFTMKKSLLVLFFLGIVSLSLC–QEERSADDE–––– EG–EVIEEEVKRGFMDTAKNVAKNVAVTLLDNLKCKITKAC
Odorranain–G1	MFTMKKSLLLLFFLATINLSLCEEERNAEEERRDDP DEMNAEVEKRFMPILSCSRFKRC
Odorranain–H3	MFTLKKSLLLLFFLGTINLSLC–QDETNAEEE–– RRDEEVAKMEEIKRGLFGKILGVGKKVLCGLSGMC
Odorranain–I1	MFTMKKSLLVLFFLGIVSLSLCQEERTAEEEDNGEVEEEKRGFF TLIKAANKLINKTVNKEAGKGGLEIMA
Odorranain–J1	MFTLKKPLLVLFFLGTISLSLC–EQERAADEE–––– DNGEIEEVNIGLFTLIKCAYQLIAPTVACN
Odorranain–K1	MFTMKKSLLVLFFLGTISLSLC–QEERAADEE–––– DNGEVEEVKRGLFTLIKGAAKLIGKTVPKKQARLGMNLWLVKLPTNVKT
Odorranain–L1	MFTLKKSLLLLFFLGTISLPLCEQERDADEEGNEENRVEVQV RDKGKGIYGLSPLRQPAP

(continued)

Table 14.1 (continued)

Precursors	Sequences
Odorranain–M1	MFTLKKFLLLLFFLGIVSSSPC–LRKRDADEEGNEENGGEAKMEDIKRAT AWDFGPHGLLPIRPIRIRPLCGKDKS
Odorranain–N1	MFTLKKSLLLIVLLGIISLSLC–EQERAADED– – – – ETNAEEERRDEKGPKWKR
Odorranain–O1	MFTLKKSLLLLFFLGTISLS– – – – – – – – ADDE– – – – DGGEAKLEDIKRAVPLIYNRPGIYVTKRPKGK
Odorranain–P1–2	MFTLKKSLLLLFFLGTINLFFCQEEERNADEEERRDERDVEVEKRVIPF VASVAAETMQHVYCAASKKMLKLNWKSSDVENHLAKC
Odorranain–P2–1	MFTLKKSLLLLFFLGTINLSLC–RDETNAEEE–RRDEEVAKMEEIKRG LLSGILGAGKHIVCGLSGPCQSLNRKSSDVEYHLAKC
Odorranain–Q2	MFTLKKSLLLLFFLGTISLSLC–EEERDADEERRDDEVEETRRAPFRMWY MYHKLKDMEPKPMA
Odorranain–R1	MFTLKKSLLLIVLLGIISLSLC–EQERAADED– – – – EGNEIKRGFSPNLPGKGLRIS
Odorranain–S1	MFTMKKSLLLLFFLGAISLSLC–EQERDADEE– EENGGEAKVEEIKRFLPPSPWKETFRTS
Odorranain–T1	MFTLKKSLLLLFFLGTISLSLC–EQERDADEESNEENGVEAKVKELKRTS RCYIGYRRKVVCS
Odorranain–U1	MFTLKKSLLLLFFLGTISLSLC–EEERDADEEGNEENGGEAKLEVVKRGC SRWIIGIHGQICRD
Odorranain–V1	MFPLKKSLLLLFFLGTINLSFCQDETNAEEERRDEE VAKMEEIKRGLLSGTSVRGSI
Odorranain–W1	MFTLKKSLLLLFFLGTINLSLCQDETNAEEERRDEE VAKMEEIKRGLFGKSSVWGRKYYVDLAGCAKA

Antimicrobial peptides: Gaegurin-4 (26), Brevinin-1E (27), palustrin 1c (28), brevinin-2GHc (29), Ranalexin (30), esculentin 1b (27), esculentin-2S (31), Ranatuerin-2Va (28), palustrin-OG1 (27), temporin B (32), Nigrocin-2 (GenBank Accession CAL25905); other antimicrobial peptides are from reference (11).

Table 14.2 Sequence comparison of antimicrobial precursors from amphibians of *Bombina* genus

Precursors	Sequences
Bombina variegata Bombinin-like peptide	MNFKYIVAVSILIASAYARSEENDIQSLSQRDVLEEESLREIRG IGGALLSAAKVGLKGLAKGLAEHFANGKRTAEER. . .
Bombina orientalis Bombinin-like peptide 1	MNFKYIVAVSILIASAYARSEENDIQSLSQRDVLEEESLR EIRGIGASILSAGKSALKGLAKGLAEHFANGKRTAEDH. . .
Bombina maxima Maximin 3	MNFKYIVAVSFLIASAYARSVQNDEQSLSQRDVLEEESLREIRGI GGKILSGLKTALKGAAKELASTYLHRRRTAEEH. . .

Antimicrobial peptides: *Bombina variegata* Bombinin-like peptide (33), *Bombina orientalis* Bombinin-like peptide 1 (34), *Bombina maxima* Maximin 3 (7).

2. Materials

1. Protocols described here were adapted for use with the following species: *Bombina maxima*, *Bombina microdeladigitora*, *Odorrana grahami*, *Rana pleuraden*, *Rana nigrovittata*, *Amolops loloensis*. Animals were collected from the Yunnan province of China. The sea frog, *Rana cancrivora*, was from the Hainan province of China (*see* **Notes 1 and 2**).
2. Size-exclusion chromatography medium: Sephadex G-50 (Superfine).
3. Ion exchange chromatography media: DEAE-Sephadex A-50 and CM-Sephadex C-25 (Amersham Biosciences).
4. Reversed-phase high-performance liquid chromatography (RP-HPLC) pre-packed C_{18} column (Hypersil BDS 300A, 30 × 0.46 cm, Dalian Elite Analytical Instruments Co, Da Lian, China).
5. Protease inhibitor cocktail for general use (Sigma).
6. Protein sequencing: LC 491 Protein Sequencing System (Applied Biosystems).
7. Peptide synthesis: ABI 433A Peptide Synthesizer (Applied Biosystems). Peptides should be analysed by HPLC and MALDI-TOF mass spectrometry to confirm that the purity is higher than 95%. All peptides should be dissolved in water.
8. DNA sequencing: ABI PRISM 377 (Applied Biosystems).
9. RNA isolation: TRIzol(R) (Life (Pacific) Technologies, China).
10. DNA amplification and cloning: PCR thermal cycler (Bio-Rad); SMART(TM) PCR cDNA synthesis kit and DNA Advantage polymerase (Clontech, USA); pGEM®-T Easy vector (Promega, WI).
11. Microorganism strains: Gram-positive bacterial strains *Staphylococcus aureus* (ATCC2592), *Bacillus megatherium*, Gram-negative bacterial strains *Escherichia coli* (ATCC25922), *Bacillus pyocyaneus* (CMCCB10104), *Bacillus dysenteriae*, *Klebsiella pneumoniae*, and fungal strains *Candida albicans* (ATCC2002), *Aspergillus flavus* (IFFI4015) and *Penicillium uticale* (IFFI2001) were obtained from Kunming Medical College.
12. Fungi culture medium: Dissolve 10 g BactoYeast extract, 20 g BactoPeptone, 20 g Dextrose and 20 g agar in 500 mL water, add water to 1000 mL and autoclave.

13. DNA synthesis: Model 381A (Applied Biosystems).
14. Matrix-Assisted Laser Desorption Ionization Time-Of-Flight Mass Spectrometry: DETM MALDI-TOF-MS (Voyager DE Pro), operated in positive ion and linear mode. The ion acceleration voltage 20 kV; polypeptide mass standard (Applied Biosystems).
15. Fast atom bombardment (FAB) MS: Autospec-3000 (VG, Manchester, UK). Matrix mix: glycerol/3-nitrobenzyl alcohol/dimethyl sulphoxide = 1/1/1 (v:v:v). The ion gun was routinely operated at 25 kV with a current of 1 μA, using Cs+ as the bombarding gas.

3. Methods

3.1. Collecting Skin Secretions

This is one of the oldest methods suitable for the extraction of skin secretions from all kinds of amphibians including frogs, toads and salamanders. An obvious disadvantage is the need to sacrifice the animals for their skins. This approach was widely used in the past but is no longer popular (*see* **Note 3**).

3.1.1. Extraction of Amphibian Skin Secretions from Homogenized Amphibian Skins

3.1.2. Stimulated Extraction of Amphibian Skin Secretions Using Adrenaline or Noradrenalin

Adrenaline and noradrenalin occur naturally in amphibians and small amounts are present in skin glands. Adrenergic receptor activation induces the contractions of the glandular myoepithelium and the outflow of skin secretion. Injection of adrenaline or noradrenalin can activate adrenergic receptors of amphibians and stimulate section. Skin secretions of most of the amphibians could be extracted by this method.

1. Dissolve adrenaline or noradrenalin in 0.5 mL of saline solution and inject the dorsal subepidermal tissue of amphibian at the base of the spine. The final dose should be 50 μg/kg.
2. Several minutes after the initial injection, skin secretions could be seen on the skins of amphibians (*see* **Note 4**).
3. Collect the skin secretions from the amphibian by scraping gently with toothpicks or similar tools and transfer them to a suitable buffer solution (*see* **Note 5**).

3.1.3. Stimulated Extraction of Amphibian Skin Secretions Using Physical Stimuli

Mild electrical stimulation has been proved to be an effective method to stimulate the release of amphibian skin secretions even in species lacking enlarged compact glands (12).

1. Use platinum electrodes and 5–10 V AC current (50 Hz, 4 ms pulse width) to stimulate the amphibian; gently rub the electrodes over the moistened dorsal skin surface for 5–20 s (*see* **Note 6**).

2. Another simple method to get amphibian skin secretions is to press the granular glands gently so as to milk skin secretions. This latter method is most suitable for amphibian species with enlarged compact glands.

3.1.4. Stimulated Extraction of Amphibian Skin Secretions Using Ether (see **Note 7***)*

1. Transfer several frogs or toads (5–10, depending on the size of the animals) into a glass beaker (1000 mL).
2. Moisten a piece of absorbent cloth with anhydrous ether (~1 mL) and put it on top of the container with animals.
3. Cover the container with a lid and wait for 1–2 min (*see* **Note 8**).
4. Collect the skin secretions of the animals with the solution of protease inhibitors in 0.1 M NaCl.
5. Transfer the anaesthetized amphibians back into a clean moist pond (*see* **Note 9**).
6. Without any further delay centrifuge the eluted peptide secretion for 20 min at 5000×*g* to remove any cell debris. Collect the supernatant and freeze it immediately in liquid nitrogen. Freeze-dry the samples and store the lyophilized secretions at –20°C (*see* **Note 10**).

3.2. Isolation of Antimicrobial Peptides from Amphibian Skin Secretions

Most of the antimicrobial peptides isolated from amphibians are cationic in nature and contain some ~10–50 amino acids. The best purification strategy would therefore include a combination of cation exchange (e.g. CM-Sephadex C-25, see **Fig. 14.1A**) to enrich for basic peptides and/or size-exclusion chromatography (e.g. with Sephadex G-50 or Superdex Peptide gel filtration, see **Fig. 14.1B**) to remove larger molecular weight proteins. The final step is usually reversed-phase high performance liquid chromatography (RP-HPLC). In cases when limited amounts of crude amphibian skin secretions are available, the samples may be fractionated with RP-HPLC directly (see **Fig. 14.1C**) (*see* **Note 11**).

3.2.1. Isolation of Antimicrobial Peptides from the Skin Secretions of Bombina maxima (see **Note 12***)*

1. Dissolve ~1.8 g of lyophilized skin secretions of *B. maxima* (total absorbance at 280 nm is 300) in 10 mL of 50 mM Tris-HCl, 5 mM EDTA, pH 7.3.
2. Dialyze the sample against the same buffer (2 L) at 4°C for 8 h.
3. Load the dialyzed sample onto DEAE-Sephadex A-50 ion exchange chromatography column (26 × 400 mm) equilibrated with the same buffer.
4. Elute the peptides with two column volumes of the same buffer; collect 2.5 mL fractions. Monitor the elution by measuring absorbance at 280 nm.

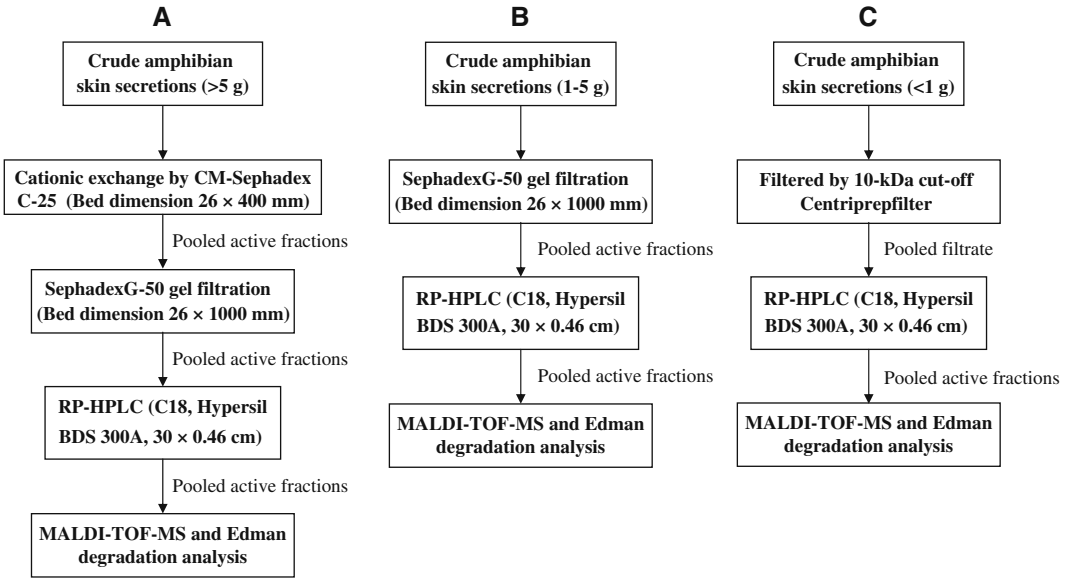


Fig. 14.1. Different isolation procedures of amphibian antimicrobial peptides based on different amounts of crude amphibian skin secretions.

5. Equilibrate Sephadex G-50 column (100 × 2.6 cm) with 0.15 M phosphate buffer, pH 7.8, and apply peptide fractions eluted from DEAE-Sephadex A-50 column (*see* **Note 13**).
6. Collect fractions (2.5 mL) and analyse them for antimicrobial activity (as described in **Section 3.4**).
7. Fractions possessing functional activity can now be characterized by FAB mass spectrometry (*see* **Note 14**).
8. To further fractionate the peptides, pool any functionally active fractions (from Step 5 above), freeze-dry, re-dissolve in smaller volume of water (~10 mL) and dialyze for 12 h against 0.1 M phosphate buffer solution, pH 7.8.
9. Apply the dialyzed sample to CM-Sephadex C-25 ion exchange column. Elute with a linear (0–0.8 M) gradient of NaCl (500 mL PBS: 500 mL PBS containing 0.8 M NaCL) and set flow rate to 15 mL/h (*see* **Note 15**).
10. Collect eluted peaks and analyse for antimicrobial activity (as described in **Section 3.4**) (*see* **Note 15**).
11. Apply the eluted peptide fractions onto a Hypersil BDS C18 RP-HPLC column (30 × 0.46 cm) equilibrated with 0.1% (v/v) trifluoroacetic acid/water. Elute with acetonitrile gradient (0–60%) over 30 min at a flow rate of 0.7 mL/min. (*see* **Note 16**).

3.2.2. Isolation of Antimicrobial Peptides from the Skin Secretions of Odorrana Grahami (see **Note 17***)*

1. Dissolve ~3.5 g of lyophilized skin secretions of *O. grahami* (total absorbance at 280 nm is 1000) in 10 mL of 0.1 M phosphate buffer, 5 mM EDTA, pH 6.
2. Equilibrate Sephadex G-50 column (2.6 × 100 cm) with 0.1 M phosphate buffer, pH 6.0 and apply the dissolved skin secretions.
3. Collect 3 mL fractions. Monitor the elution by measuring absorbance at 280 nm.
4. Analyse collected fraction for antimicrobial activity (as described in **Section 3.4**).
5. Pool any functionally active fractions (from Step 4 above), freeze-dry and re-dissolve in 2 mL of 0.1 M phosphate buffer, pH 6.0.
6. Apply the peptide samples on a Hypersil BDS C18 RP-HPLC column (30 × 0.5 cm) equilibrated with 0.1% (v/v) trifluoroacetic acid in water. The elution was performed with the gradients of acetonitrile (from 0 to 100% over 60 min) in 0.1% (v/v) trifluoroacetic acid in water at a flow rate of 0.7 mL/min (*see* **Note 18**).
7. Collect fractions and analyse for functional activity (as described in **Section 3.4**) (*see* **Note 19**).
8. Analyse functionally active fractions by Edman degradation (*see* **Note 20**) and MALDI-MS or FAB-MS (*see* **Note 21**).

3.3. cDNA Synthesis and Amplification

3.3.1. SMART cDNA Synthesis

1. Extract total RNA from the skin of a single sample of *O. grahami* using TRIzol reagent.
2. Synthesize cDNA, e.g. using SMART(TM) PCR cDNA synthesis kit and the SMART primers supplied with the kit.
3. Amplify cDNA using Advantage polymerase and 5′ PCR primer II A, 5′-AAGCAGTGGTATCAACGCAGAGT-3′.

3.3.2. Screening of cDNA Encoding Antimicrobial Peptides

cDNAs encoding antimicrobial peptides can be amplified using PCR and primers designed to match highly conserved sequence encoding the signal peptide.

1. For ranid frogs use the following forward PCR primer: S1 (5′-CCAAA(G/C)ATGTTCACC(T/A)TGAAGAAA (T/C)-3′.
2. If, as here, the cDNA was made using SMART cDNA synthesis kit, use the reverse primer II A supplied with the kit.
3. Typical PCR amplification should be conducted as follows: 2 min at 94°C, followed by 30 cycles of 10 s at 92°C, 30 s at 50°C, 40 s at 72°C.
4. Clone the PCR products using pGEM®-T Easy vector or similar "TA" cloning system (*see* **Note 22**).

3.4. Testing Antimicrobial and Antifungal Activity of Peptides

1. Grow bacteria in LB broth to an OD600 = 0.8.
2. Add a 10-μL aliquot of the bacterial culture to 8 mL of fresh LB broth containing 0.7% agar and pour into a 90-mm Petri dish containing sterile 1.5% agar in LB broth (~25 mL).
3. Wait till the top agar sets. The plates are now ready for screening the antimicrobial samples.
4. A 20-μL aliquot of the fractionated peptide sample (from **Section 3.2**) filtered through a 0.22-μm Millipore filter should be added onto the surface of the top agar.
5. Wait until the sample dries. Incubate the plate overnight at 37°C.
6. Clear zones formed on the surface of the top agar will indicate the presence of antimicrobial activity.
7. Determine the minimal inhibitory concentration (MIC) by incubating the bacteria in LB broth with variable concentrations of the sample peptide. Record MIC at which no visible growth occurs.
8. To assay anti-fungal activity of samples use strains *A. flavus* and *P. uticale*, and count the fungal spore concentration under a microscope. Culture the fungi (initial concentration of 10^5 spores/mL) in yeast extract-peptone-dextrose broth; use different amounts of the test samples.

4. Notes

1. Further information on the species used is available from literature (7, 11, 16–22). Our methods should be easy to adapt for use with other species.
2. It is better to catch toads or frogs at night because most amphibians are accustomed to stay in an unshielded open place at night during summer and autumn periods. Most of the amphibians can be kept alive in the laboratory for a long time, and attention should be paid to their diet. A suitable diet includes live insects, such as the larvae of *Tenebrio molitor* or crickets, but the choice might be different for other species.
3. One other problem is that many other compounds from skin and surrounding tissues may co-extract with the secretions. However, should the skins become available, these should be cut into small pieces and grinded in suitable solvents such as water, methanol or buffer solutions (4, 12–14) prior to peptide extraction.

4. In the case of *Xenopus laevis*, droplets of a white viscous solution could be seen on the back and hind legs (23, 24).
5. This may be 0.9% sodium chloride or 0.1 M ammonium acetate or 50 mM Tris-HCl (pH 7.8) or 0.1 M phosphate buffered saline (pH 6.0).
6. Optimal volt strength and the duration time would vary between different amphibian species. In the case of the frog *Rana palustris*, the volt strength and the duration time are 5 V and 10 s, respectively (25).
7. We developed this new extraction method and used it successfully with many amphibians except *Bufo* amphibians (7, 11, 16–22).
8. Some stress-like behaviour may be observed (e.g. amphibians start jumping), but they will become anaesthetized quickly and their skins will exude copious secretions.
9. Importantly, there should be no water in the pond until the anaesthetized amphibians are awake fully.
10. This method causes less discomfort to the amphibians compared to injection-based methods, electrical or mechanical stimulation. No foreign compounds are introduced into the samples either, due to the highly volatile nature of ether.
11. No single universal method has yet been reported for the isolation of antimicrobial peptides from amphibian skin secretions. The choice of purification strategy would depend on the type of secretion and of the anticipated amount of the peptides. A summary of purification strategies used in our work is shown in **Fig. 14.1**.
12. *Bombina maxima* secretions contain many anionic peptides and a large amount of high molecular weight proteins. Therefore, we used DEAE-Sephadex A-50 ion exchange chromatography to remove anionic peptides, and Sephadex G-50 size-exclusion chromatography to remove high molecular weight proteins. Furthermore, because several serine protease inhibitors have been found in skin secretions of *Bombina* amphibians, we did not use protease inhibitor cocktail in this particular case.
13. Size-exclusion separation yielded five peaks; fraction IV displayed antimicrobial activity.
14. In this particular experiment the active fraction contained a mixture of peptides with molecular masses ranging from ~2500 to 2900 Da. Peak IV was pooled, lyophilized.
15. In our experiments three peaks were eluted; the antimicrobial activity was found mostly in peaks I and II.

16. We obtained five different antimicrobial peptides and named them *maximins*. Starting with ~1.8 g of the lyophilized skin secretions of *B. maxima* we obtained 4.5, 1.6, 5.2, 2.6, 1 mg of *maximins* 1, 2, 3, 4 and 5, respectively (7).
17. This section exemplifies a protocol for use with smaller samples of crude skin secretions (~1–5 g). We normally omit the ion exchange chromatography step when working with these amounts of starting material.
18. In our experiments six peaks were eluted from the Sephadex G-50 column; the antimicrobial activity was found mostly in peaks V and VI. Following RP-HPLC these yielded 58 and 69 individual fractions, respectively.
19. We identified 21 functionally active peptide fractions after RP-HPLC separation.
20. All 21 functionally active fractions have been analysed by Edman degradation. We identified 17 different antimicrobial peptides families in *O. grahami* of which 13 were novel groups.
21. We chose to analyse all 127 RP-HPLC fractions by MALDI-TOF. We found that 56 fractions contained masses matching ones predicted based on cDNA sequences of *O. grahami* (we predicted 47 antimicrobial peptides, *see* **Section 3.2.2**). Most of the masses were consistent with the proposed structures and demonstrated the presence of a cysteine bridge in the *Brevinin-1E-OG1*, *Brevinin-1E-OG6, Brevinin-2E-OG3*, *Esculentin-1-OG3*, *Esculentin-2-OG6*, *Nigrocin-OG1*, *Nigrocin-OG20*, *Odorranain-A1, -B1, -C6, -G1, -H1, -H2, -J1* and *-T1* peptides (11).
22. The technique described allowed us to identify 372 sequences of antimicrobial peptide-like sequences from *O. grahami* (GenBank accession DQ672724-DQ673095). These sequences can be classified into 30 divergent groups containing 107 non-identical antimicrobial peptides. Six of the thirty groups match antimicrobial peptide families found in other amphibians. These have named *Brevinin-1E-OG*, *Brevinin-2E-OG*, *Esculentin-1-OG*, *Esculentin-2-OG*, *Nigrocin-OG* and *Palustrin-OG*, respectively. Other 24 families appear to represent novel families.

References

1. Zasloff, M. (2002) Antimicrobial peptides of multicellular organisms. *Nature* **415**, 389–395.
2. Epand, R.M. and Vogel, H.J. (1999) Diversity of antimicrobial peptides and their mechanisms of action. *Biochim. Biophys. Acta* **1462**, 11–28.
3. Borgden, K.A. (2005) Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat. Rev. Microbiol.* **3**, 238–250.

4. Clarke, B.T. (1997) The natural history of amphibian skin secretions, their normal functioning and potential medical applications. *Biol. Rev. Camb. Philos. Soc.* **72**, 365–379.
5. Duda, T.F., Jr, Vanhoye, D. and Nicolas, P. (2002) Roles of diversifying selection and coordinated evolution in the evolution of amphibian antimicrobial peptides. *Mol. Biol. Evol.* **19**, 858–864.
6. Conlon, J.M., Kolodziejek, J. and Nowotny, N. (2004) Antimicrobial peptides from ranid frogs: taxonomic and phylogenetic markers and a potential source of new therapeutic agents. *Biochim. Biophys. Acta* **1696**, 1–14.
7. Lai, R., Zheng, Y.T., Shen, J.H., Liu, G.J., Liu, H., Lee, W.H., et al. (2002) Antimicrobial peptides from skin secretions of Chinese red belly toad *Bombina maxima*. *Peptides* **23**, 427–435.
8. Conlon, J.M., Al-Ghaferi, N., Abraham, B., Jiansheng, H., Cosette, P., Leprince, J., et al. (2006) Antimicrobial peptides from diverse families isolated from the skin of the Asian frog, *Rana grahami*. *Peptide* **27**, 2111–2117.
9. Barra, D., Simmaco, M. and Boman, H.G. (1998) Gene-encoded peptide antibiotics and innate immunity. Do 'animalcules' have defence budgets? *FEBS Lett.* **430**, 130–134.
10. Matutte, B., Storey, K.B., Knoop, F.C. and Conlon, J.M. (2000) Induction of synthesis of an antimicrobial peptide in the skin of the freeze-tolerant frog, *Rana sylvatica*, in response to environmental stimuli. *FEBS Lett.* **483**, 135–138.
11. Li, J., Xu, X., Xu, C., Zhou, W., Zhang, K., Yu, H., et al. (2007) Anti-infection peptidomics of amphibian skin. *Mol. Cell. Proteomics* **6**, 882–894.
12. Tyler, M.J., Stone, D.J. and Bowie, J.H. (1992) A novel method for the release and collection of dermal, glandular secretions from the skin of frogs. *J. Pharmacol. Toxicol. Methods* **28**, 199–200.
13. Erspamer, V. (1971) Biogenic amines and active polypeptides of amphibian skin. *Ann. Rev. Pharmacol.* **2**, 327–350.
14. Daly, J.W., Brown, G.B., Mensah-Dwumah, M. and Myers, C.W. (1978) Classification of skin alkaloids from neotropical poison-dart frogs (Dendrobatidae). *Toxicon* **16**, 163–188.
15. Zhou, M., Wang, L., Owens, D.E., Chen, T., Walker, B. and Shaw, C. (2007) Rapid identification of precursor cDNAs encoding five structural classes of antimicrobial peptides from pickerel frog (*Rana palustris*) skin secretion by single step "shotgun" cloning. *Peptides* **28**, 1605–1610.
16. Lu, X., Ma, Y., Wu, J. and Lai, R. (2008) Two serine protease inhibitors from the skin secretions of the toad, *Bombina microdeladigitora*. *Comp. Biochem. Physiol. B* **149**, 608–612.
17. Liu, X., You, D., Chen, L., Wang, X., Zhang, K. and Lai, R. (2008) A novel bradykinin-like peptide from skin secretions of the frog, *Rana nigrovittata*. *J. Pept. Sci.* **14**, 626–630.
18. Wang, X., Song, Y., Li, J., Liu, H., Xu, X., Lai, R., et al. (2007) A new family of antimicrobial peptides from skin secretions of *Rana pleuraden*. *Peptides* **28**, 2069–2074.
19. Lu, Y., Ma, Y., Wang, X., Liang, J., Zhang, C., Zhang, K., et al. (2008) The first antimicrobial peptide from sea amphibian. *Mol. Immunol.* **45**, 678–681.
20. Li, J., Zhang, C., Xu, X., Wang, J., Yu, H., Lai, R., et al. (2007) Trypsin inhibitory loop is an excellent lead structure to design serine protease inhibitors and antimicrobial peptides. *FASEB J.* **21**, 2466–2473.
21. Lu, Y., Li, J., Yu, H., Xu, X., Liang, J., Tian, Y., et al. (2006) Two families of antimicrobial peptides with multiple functions from skin of rufous-spotted torrent frog, *Amolops loloensis*. *Peptides* **27**, 3085–3091.
22. Che, Q., Zhou, Y., au>Yang, H., Li, J., Xu, X. and Lai, R. (2008) A novel antimicrobial peptide from amphibian skin secretions of *Odorrana grahami*. *Peptides* **29**, 529–535.
23. Nutkins, J.C. and Williams, D.H. (1989) Identification of highly acidic peptides from processing of the skin prepropeptides of *Xenopus laevis*. *Eur. J. Biochem.* **181**, 97–102.
24. Giovannini, M.G., Poulter, L., Gibson, B.W. and Williams, D.H. (1987) Biosynthesis and degradation of peptides derived from *Xenopus laevis* prohormones. *Biochem. J.* **243**, 113–120.
25. Marenah, L., Flatt, P.R., Orr, D.F., McClean, S., Shaw, C. and Abdel-Wahab, Y.H. (2004) Brevinin-1 and multiple insulin-releasing peptides in the skin of the frog *Rana palustris*. *J. Endocrinol.* **181**, 347–354.
26. Kwon, S.Y., Carlson, B.A., Park, J.M. and Lee, B.J. (2000) Structural organization and expression of the gaegurin 4 gene of *Rana rugosa*. *Biochim. Biophys. Acta* **1492**, 185–190.
27. Simmaco, M., Mignogna, G., Barra, D. and Bossa, F. (1994) Antimicrobial peptides from skin secretions of *Rana esculenta*. Molecular cloning of cDNAs encoding esculentin and brevinins and isolation of new active peptides. *J. Biol. Chem.* **269**, 11956–11961.
28. Chen, T., Zhou, M., Rao, P., Walker, B. and Shaw, C. (2006) The Chinese bamboo leaf odorous frog (*Rana (Odorrana) versabilis*)

and North American *Rana* frogs share the same families of skin antimicrobial peptides. *Peptides* **27**, 1738–1744.
29. Zhou, J., McClean, S., Thompson, A., Zhang, Y., Shaw, C., Rao, P., et al. (2006) Purification and characterization of novel antimicrobial peptides from the skin secretion of *Hylarana guentheri*. *Peptides* **27**, 3077–3084.
30. Clark, D.P., Durell, S., Maloy, W.L. and Zasloff, M. (1994) Ranalexin, A novel antimicrobial peptide from bullfrog (*Rana catesbeiana*) skin, structurally related to the bacterial antibiotic, polymyxin. *J. Biol. Chem.* **269**, 10849–10855.
31. Chen, T., Zhou, M., Chen, W., Lorimer, J., Rao, P., Walker, B., et al. (2006) Cloning from tissue surrogates: antimicrobial peptide (esculentin) cDNAs from the defensive skin secretions of Chinese ranid frogs. *Genomics* **87**, 638–844.
32. Simmaco, M., Mignogna, G., Canofeni, S., Barra, D., Simmaco, M. and Rivas, L. (1996) Temporins, antimicrobial peptides from the European red frog *Rana temporaria*. *Eur. J. Biochem.* **242**, 788–792.
33. Mignogna, G., Simmaco, M., Kreil, G. and Barra, D. (1993) Antibacterial and haemolytic peptides containing D-alloisoleucine from the skin of *Bombina variegate*. *EMBO J.* **12**, 4829–4832.
34. Gibson, B.W., Tang, D.Z., Mandrell, R., et al. (1991) Bombinin-like peptides with antimicrobial activity from skin secretions of the Asian toad, *Bombina orientalis*. *J. Biol. Chem.* **1266**, 23103–23111.

Chapter 15

Identification and Relative Quantification of Neuropeptides from the Endocrine Tissues

Kurt Boonen, Steven J. Husson, Bart Landuyt, Geert Baggerman, Eisuke Hayakawa, Walter H.M.L. Luyten, and Liliane Schoofs

Abstract

Endocrine tissues like the pituitary, hypothalamus and islets of Langerhans are rich in bioactive peptides. These are used for intercellular signalling and are involved in regulation of almost all physiological processes. Peptidomics is the comprehensive analysis of peptides in tissues, fluids and cells. Peptidomics applied to (neuro-)endocrine tissues aims therefore to identify as many bioactive peptides as possible. Peptidomics of (neuro-)endocrine tissues requires an integrated approach that consists of careful sample handling, peptide separation techniques, mass spectrometry and bioinformatics. Here we describe the methods for isolation and dissection of endocrine tissues, the extraction of bioactive peptides and further sample handling and identification of peptides by mass spectrometry and hyphenated techniques. We also present a straightforward method for the comparison of relative levels of bioactive peptides in these endocrine tissues under varying physiological conditions. The latter helps to elucidate functions of the bioactive peptides.

Key words: Neuropeptide, *Mus musculus* , peptidomics, pituitary, hypothalamus, islets of Langerhans, endocrine pancreas, MALDI-TOF/TOF MS, Q-TOF MS, two-dimensional HPLC, nanoLC.

1. Introduction

Bioactive peptides are peptides that are used for intercellular signalling. These are present in neural as (neuropeptides) well as non-neural tissues and perform key roles in the regulation of almost every aspect of physiology (1). These peptides may act as neurotransmitters, neuromodulators or neurohormones in the nervous system and have important paracrine and hormonal effects in non-neural endocrine tissues like the gut or pancreas. Bioactive peptides and their corresponding receptors and processing enzymes are important drug targets (2). Bioactive

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_15,

peptide identification and quantification are therefore of vital importance in the understanding of many diseases. Neuropeptides are cleaved out of their precursors in the endoplasmic reticulum by proprotein convertases (PCs). The PCs typically cleave peptides at mono- or di-basic cleavage sites. The exposed C-terminal basic amino acids are further removed by exoproteases like carboxypeptidase E. Many peptides need to be modified to improve their stability or to become bioactive (1, 3).

Mass spectrometry (MS) is widely used in proteomics research along with traditional methods based on SDS-PAGE electrophoresis and is routinely capable of attomolar detection levels and femtomolar quantities can be sufficient for peptide sequencing. Peptides fall outside the range of masses suitable for routine SDS-PAGE analysis and this hindered their analysis. However, the advent of MS and hyphenated techniques (like nanoLC-MS) speeded up peptide analysis significantly and helped to define the peptidomics as an extension of proteomics that specialised in the characterisation of peptides. Peptidomics can be defined as the comprehensive analysis of all peptides in a biological sample like body fluids, tissues, organs and even cells (4–7). LC, MS and bioinformatics form the core of the peptidomic approaches. In this chapter we present the analysis of bioactive peptides from a few selected (neuro-)endocrine organs. The endocrine pancreas, the pituitary and the hypothalamus are chosen because of their physiological importance and their high concentrations of neuropeptides. The MS analysis of neuropeptides from brain tissues dates back to the 1990s (8). Invertebrate peptidomics studies were initially very successful and the peptidomes of nervous tissues of *Drosophila melanogaster* (9), *Locusta migratoria* (10) and *Periplaneta americana* (11) have been characterised with these new techniques. Vertebrate studies pose more problems because excessive sample handling may induce proteolysis. The proteolytic peptides can be very abundant making the analyses of bioactive almost impossible because these cannot be discriminated in the mass spectrometer. This was the case, for example, in porcine brain (12) and some murine brain studies (13). One of the solutions to circumvent this problem was the use of Cpe^{fat}/Cpe^{fat} mice by the research group of Fricker (14). These mice lack carboxypeptidase E and all the neuropeptides therefore have a C-terminal extension of one or two basic amino acids. An anhydrotrypsin affinity resin could therefore be used for enrichment of neuropeptides in the sample. Another early study of the laboratory of Andrén proved the usefulness of microwave irradiation (15). A 5 kW microwave was used to heat the murine brain quickly and thus to inactivate proteases and peptidases. Microwave irradiation can be done using a specialised small-animal microwave or a regular microwave. A small-animal microwave has the advantage of higher radiation power that is

able to bring the temperature of the brain to more than 90°C in 1.4 s (animals can be sacrificed this way) (15). However, small-animal microwaves are costly. Regular microwaves may be used for heating the brain rapidly after decapitation (16), but their use requires extensive optimisation. Various brain tissues (like pituitary, striatum or hypothalamus) can also be snap-frozen using liquid nitrogen (17). If smaller brain regions are wanted (like hypothalamic nuclei), cryostat slices can be used to dissect the region of interest. The brain is first dissected and quickly cooled in dry ice chilled isobutanol (at a temperature of –40°C) (18). In this chapter we present a method which relies on snap-freezing of the brain tissues with liquid nitrogen. We also describe here a method for the isolation of the islets of Langerhans from the murine pancreas, sample handling and the subsequent peptide extraction for the MS analysis (19). Since the living islets are used for the extraction of neuropeptides, proteolytic degradation has almost no influence on the outcome of peptidomic experiments. The peptides are extracted in acidified methanol at low temperatures. Other methods rely on boiling the tissues in water or acidified water. Combinations of various methods of peptide extraction are also used, e.g. a combined extraction with boiling water followed by cold acid extraction, or the combination of boiling water and acidified organic solvents or cold acidified organic solvent extraction combined with cold acid extraction (20, 21).

In our work we rely on a combination of the two complementary techniques of MALDI-TOF/TOF and ESI Q-TOF. MALDI-TOF/TOF is combined with the off-line 1D LC. The inherent complexity of (in)vertebrate neuropeptide extracts makes off-line 2D LC the most appropriate separation procedure for use with ESI Q-TOF. Nanospray ESI has the advantage that ion suppression effects do not occur, making the ion intensities reproducible between different MS runs. This allows for relative quantification of neuropeptides without the need for labelling methods, which makes this approach straightforward (22). Standardisation of the resulting intensities is of utmost importance, since ion intensities depend on the performance of the LC-MS system. Standardisation can be done with by adding internal standards or with normalisation with the total ion current (22). Ion intensities are divided by the value of the ion intensity of the internal standard or with the sum of all ion intensities of the peptides detected in the LC-MS run. This can be done manually or with recently developed software like e.g. *DeCyder MS* (23), which helps to visualise 1D LC-MS data. The detection of differentially expressed peptides is backed up by the statistics analysis; the normalisation of data is possible with the two methods mentioned above. Quantitative analysis of neuropeptides can provide valuable information of their possible functions in various disease states.

2. Materials

2.1. Dissection of Brain Tissues, Islets Isolation and Neuropeptide Extraction

1. Mice (c57bl) are purchased from Harlan Nederland (Horst, the Netherlands). Two-week-old mice are used for islets isolation from the pancreas. Pituitary and hypothalamus isolation are performed on 3-month-old mice or older.
2. Hank's medium containing 25 mM HEPES.
3. Freshly made collagenase P (SERVA NB8), 1 mg/ml in Hank's medium.
4. A dextran stock solution of 0.28 g/ml. Dextran gradient: Layer I (0.26 g/ml); Layer II (0.22 g/ml); Layer 3 (0.12 g/ml). Make fresh solutions for each experiment (only useable for a week).
5. Stereomicroscope.
6. Extraction medium: Methanol:Water:Formic Acid (FA) (90:9:1; v:v:v), freshly made, cool to 0°C before use.
7. Tissue pulveriser (Bessman) and sonicator (Branson 5510 ultrasonic cleaner or MSE Soniprep 150 ultrasonic disintegrator).
8. Liquid nitrogen.

2.2. Sample Preparation

1. Centrifugal filter units (0.22-μm pore size, Millipore).
2. Vacuum concentrator (SpeedVac concentrator and refrigerated condensation trap, Savant).
3. Solid-phase extraction cartridges: Ziptip C18 (15 μl, Millipore), SepPak C18 (Waters) or Oasis HLB (Waters).
4. Solvents: Deionised H_2O, acetonitrile, methanol, FA and trifluoroacetic acid (TFA) (*see* **Note 1**).

2.3. Liquid Chromatograph and Mass Spectrometry

1. Miniaturised LC system (Ultimate^{TM} 3000 Nano and Capillary LC system, Dionex).
2. HPLC system (Beckman LC system with the programmable solvent module 126) and diode array detector module 165 (Gold System, Beckman Coulter, The Netherlands).
3. Reverse-phase (RP) chromatography: Symmetry C18 column (5 μm particle size, 4.6 × 250 mm, Waters)
4. Strong cation exchange (SCX) chromatography: Spherisorb column (5 μm, 4.6 × 150 mm, Waters); BioSCX column (5 μm particle size, 1000 μm × 15 cm, Dionex)

5. NanoLC columns: Precolumn (μ-guard column MGU-30 C18, LC Packings) and analytical column (Pepmap C18, 3 μm particle size, 75 × 150 mm, LC Packings)
6. Deionised H_2O, acetonitrile, FA and TFA.
7. MALDI-TOF/TOF (Ultraflex II, Bruker Daltoniks, Germany), equipped with a N_2 laser and pulsed ion extraction accessory.
8. Matrix: α-cyano-4-hydroxycinnamic acid.
9. Q-TOF (MicroTOF-Q, Bruker Daltonics, Germany).
10. Stainless steel emitter (Proteon).
11. Solvent A: Deionised water containing 0.1% FA; Solvent B: acetonitrile containing 0.1% FA; Solvent C: 100 mM FA, 5% acetonitrile, pH 2.7; Solvent D: 32 mM FA, 5% acetonitrile, 600 mM ammonium acetate, pH 5.9.
12. Standard calibration peptide mixture: Angiotensin 2 (1045.54 Da), angiotensin 1 (1295.68 Da), substance P (1346.73 Da), bombesin (1618.82 Da), ACTH clip 1–17 (2092.08 Da) and ACTH clip 19–39 (2464.19 Da) (Bruker Daltonic GmbH, Germany).

2.4. Data Processing and Peptide Identification

1. DataAnalysis software package (Bruker Daltonics)
2. *Biotools*(TM) (Bruker Daltonic GmbH, Germany)
3. *Mascot* search engine for rapid protein identification using MS data (http://www.matrixscience.com)
4. Differential analysis software (*DeCyder MS* 2.0, GE Healthcare, Sweden)
5. *CompassXport* software (Bruker Daltonic GmbH, Germany)

3. Methods

3.1. Dissection of Brain Tissues, Islets Isolation and Neuropeptide Extraction

3.1.1. Peptide Extraction from the Pituitary and Hypothalamus

1. The mice are sacrificed by spinal dislocation. The mouse is then decapitated and the brain is dissected as quickly as possible. A spatula is used to remove the brain from the base of the skull.
2. The pituitary is visible as a lob between the optic nerves, protected by the sella turcica. Place the pituitary in liquid nitrogen. The hypothalamus can be dissected from the brain with a scalpel and pincers before snap-freezing. The pituitary can be minced in an Eppendorf tube with the tip of a pair of pincers (*see* **Note 2**). The hypothalamus is best grinded

with pestel and mortar and then transferred to an Eppendorf tube.

3. Add the chilled Extraction medium and keep the sample on ice.
4. Homogenise the sample further with a bar sonicator three times for 5 s, with 15 s intervals using the MSE Soniprep 150 ultrasonic disintegrator (or a sonication bath (Branson 5510) for 2 min). Centrifuge the sample for 20 min at 4°C. The supernatant is filtered using the centrifugal filter units.

3.1.2. Extraction of Peptides from the Endocrine Pancreas

The number of mice needed depends on the experimental setup. Typically, an average of 100 islets of Langerhans are obtained from one 2-week-old mouse. Extracts from 100–200 islets of Langerhans are sufficient for one LC-MALDI run. An off-line 2D LC-MS/MS experiment using Q-TOF requires 500 islets of Langerhans (*see* **Note 3**).

1. Dissect the pancreata and place them immediately in Hank's medium.
2. Remove the fat of the pancreata.
3. Digest the pancreata enzymatically with the collagenase P solution. Add 5 ml of collagenase P solution to three 50-ml flasks. Place pancreata in flask No. 1 and shake for 3 min by hand (37°C). Transfer the non-digested pancreatic tissue (residue) to flask No. 2 and shake for 3 min. Transfer the remaining residue to flask No. 3. Shake flasks No. 1 and 3 together for 3 min. Add 40 ml of cooled Hank's medium to flask No. 2 to inactivate the collagenase. Transfer the residue from flask No. 3 to flask No. 1 and continue shaking both flasks for 3 min. Remove the residue from both flasks and add 40 ml of cooled Hank's medium to each flask. Incubate for 2 min and change the Hank's medium in all three bottles. Repeat the medium change twice (*see* **Note 4**).
4. Decant the Hank's medium and transfer the residues from all the three bottles to a fresh 50-ml tube and centrifuge at 2000×*g* for 2 min.
5. Discard the supernatant and resuspend the residue in 10 ml dextran solution (0.28 g/ml). A three-layered dextran gradient is added: 4 ml Layer I, 4.5 ml of Layer II and 4 ml of Layer III. Centrifuge at 2000×*g* for 20 min.
6. Transfer the islets to a Petri dish with a Pasteur pipette (*see* **Note 5**).
7. Transfer the islets to a centrifugal filter device (*see* **Note 6**). Centrifuge at 10,000×*g* for 1 min at 4°C to remove the Hank's medium. Transfer the filter containing the islets

to another tube (from which the original filter is removed) and add the chilled Extraction medium (*see* **Note 7**). Centrifuge the extract at 10,000×*g* for 2 min to remove the islet debris.

3.2. Sample Preparation

1. Vacuum-dry the extracts.
2. For 1D nanoLC-MS analysis only dissolve the samples in 10 μl of deionised water containing 5% acetonitrile and 0.1% FA (*see* **Note 8**).
3. For the RP-LC only dissolve the sample in 500 μl of deionised water containing 5% acetonitrile and 0.1% FA.
4. For the SCX separation only dissolve the sample in deionised water containing 10–15% acetonitrile and 0.1% FA.
5. For the solid-phase extraction dissolve the sample in deionised water containing 5% acetonitrile and 0.1% FA. SepPak C18 (Waters) and Oasis HLB (Waters) are well suited for larger sample volumes. Activate the SepPak C18 cartridge with 50–100% acetonitrile. Rinse the column with 0.1% TFA (3 times). Apply the dissolved peptide extract onto the cartridge and wash with 0.1% TFA (3 times). Elute the bound peptides with 70% acetonitrile containing 0.1% FA. To activate the Oasis HLB cartridges add acetonitrile and rinse with deionised water. Load the peptide extract, wash with 5% methanol. Elute the peptides with methanol or acetonitrile.

3.3. Liquid Chromatograph and Mass Spectrometry

3.3.1. NanoLC

1. Dissolve the sample in 10 μl of deionised water containing 5% acetonitrile and 0.1% FA.
2. Load the sample onto the precolumn with deionised water containing 2% acetonitrile and 0.1% FA (flow rate 30 μl/min). Switch the column-switching valve to connect the precolumn online with the analytical column.
3. Separate the peptides using a linear gradient from 95%/5% (Solvent A/Solvent B) to 50%/50% (Solvent A/Solvent B). Use the flow rate of 150 nl/min. The 1D nanoLC setup should be coupled directly to the Bruker microTOF-Q.

3.3.2. RP LC

1. Inject 500 μl of a peptide extract into the HPLC system.
2. Wash for 10 min using 2% of acetonitrile in 0.1% TFA.
3. The RP separation is done on a Symmetry C18 column at a flow rate of 1 ml per minute. A linear gradient over 60 min to a final concentration of 50% acetonitrile in 0.1% TFA separates the peptides. Collect 1-ml fractions (every minute) starting from the beginning of the gradient.

3.3.3. Off-Line 2D LC (see Note 9)

1. Reconstitute the sample in deionised water containing 10–15% acetonitrile and 0.1% FA. Inject 500 μl into the BioSCX column.
2. Elute the peptides with a 40 min gradient from 100% Solvent C to 100% Solvent D, followed by 10 min of isocratic elution with 100% solvent D.
3. Collect and dry the fractions.
4. Re-dissolve the peptide samples in 5% acetonitrile containing 0.1% FA and continue peptide separation with 1D RP nanoLC as described in **Section 3.3.1**.

3.3.4. Q-TOF MS

1. Connect the nanoLC system in series with the electrospray interface of the mass spectrometer (*see* **Note 10).** A voltage of 2 kV should be applied between the stainless steel emitter and the cone.
2. Ions of sufficient abundance and preferably doubly charged (with exclusion of singly charged ions) are selected for fragmentation by the automatic charge state recognition software. Each depicted spectrum is typically a summation of 10^4 spectra. 3. We use argon as the collision gas; the collision energy is automatically selected depending on the mass and the charge of the selected parent ion.
4. The detection window in the survey scan is *m/z* 400–1500. Fragmentation spectra are usually acquired from *m/z* 40–1400.
5. The instrument should be calibrated (quadratic calibration) on a daily basis with the Agilent Tune Mix (*see* **Note 11).**

3.3.5. MALDI-TOF/TOF MS (see Note 12)

1. Resuspend the dried peptide samples in water containing 2% acetonitrile and 0.1% FA (*see* **Note 13).**
2. Spot the HPLC fractions onto the ground steel MALDI target.
3. Mix the droplets with the saturated solution of α-cyano-4-hydroxycinnamic acid in ethanol:acetonitrile (50:50) and dry under a gentle flow of argon to speed up the evaporation of the solvents (*see* **Note 14).**
4. In MS mode, the following voltage settings are applied: Ion source 1: 25.02 kV; Ion source 2: 21.67 kV; Lens: 9.61 kV; Reflectron 1: 26.31 kV and Reflectron 2: 13.81 kV.
5. Pulsed ion extraction should be used. The laser (N_2) frequency is 100 Hz and the laser intensity is adapted to maximum sensitivity and accuracy. Ions are detected between *m/z* 500 and 4000.

6. Calibrate the instrument using a standard peptide mixture containing angiotensin 2, angiotensin 1, substance P, bombesin, ACTH clip 1–17i and ACTH clip 19–39.
7. Record spectra using the reflectron mode.
8. The voltage settings in MS/MS mode have the following values: Ion source 1: 8.00 kV; Ion source 2: 7.15 kV; Lens: 3.50 kV; Reflectron 1: 29.50 kV; Reflectron 2: 13.80 kV; Lift 1: 19.00 kV and Lift 2: 3 kV.
9. Mass readouts can be processed in the *FlexAnalysis* program to obtain peak list files.

3.4. Data Processing and Peptide Identification (see Note 15)

1. For analysing the Q-TOF data, use the Find Compounds MS^n algorithm from the DataAnalysis package. This program selects all the fragmentation spectra with an intensity threshold above 1000. Use a filter (fragments qualified by amino acids) to select only fragmentation spectra of putative peptides. Deconvolute the charges with the maximum number of charges being 4 for MS and 3 for MS/MS. Spectra range from *m/z* 250 to 4000 for MS and from *m/z* 50 to 3000 for MS/MS. The data are transported as Mascot generic files (MGF) to the search engine.
2. The Bruker Ultraflex II data are processed with the *FlexAnalysis* software. In MS mode, peaks are detected with the PMF-SNAP method. This method creates a mass list using the peak detection algorithm SNAP. A signal-to-noise threshold of 2 should be applied. Set the maximum number of peaks in a single spectrum to 200. Set the quality factor threshold to 50. Use averaging as the SNAP average composition. Set the baseline subtraction parameter to median (flatness 0.8; median level 0.8). Smooth the spectra using the Savitzky Golay algorithm with a width of *m/*z 0.15 and for 4 cycles. Select the high-accuracy MS masses for MS/MS (**Fig. 15.1**).
3. In MS/MS, the LIFT-SNAP method is used for the formation of mass lists. The peak detection algorithm is SNAP. Apply signal-to-noise threshold of 1.5. Set the maximum number of peaks in a single spectrum to 200. Set the quality factor threshold to 30. Use averaging as the SNAP average composition. Set the baseline subtraction parameter to TopHat. Smoothen the spectra using the Savitzky Golay algorithm with a width of *m/*z 0.15 and for 4 cycles. Use *FlexAnalysis* to create the mass lists, these can be further analysed with *Biotools* (TM). This software has a direct link to the *Mascot* search engine.
4. *Mascot* is a bioinformatics resource which matches the fragmentation data against any sequence databases (*see* **Note**

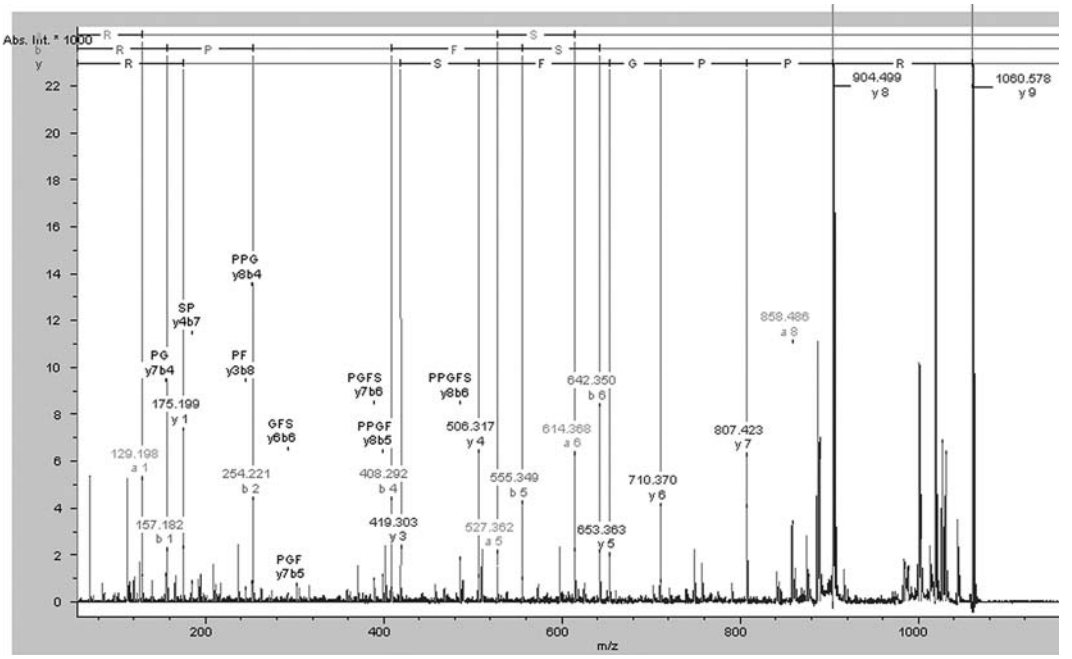


Fig. 15.1. MALDI-TOF/TOF fragmentation spectrum of bradykinin (RPPGFSPFR, monoisotopic mass of 1060.578 Da). An extract of 400 islets of Langerhans was separated on a Waters symmetry C18 column with a Beckman LC system using a 1 h gradient (see text). Fractions (1 ml) were dried and re-dissolved in 2 μl 50% acetonitrile containing 1% FA and subsequently spotted on a stainless steel MALDI target plate. α-cyano-4-hydroxycinnamic acid in ethanol/:acetonitrile (50:50) was used as matrix. Bradykinin eluted after 30 min. Sequencing was done on an Ultraflex II MALDI-TOF/TOF (Bruker Daltonics). The spectrum was analysed with the *FlexAnalysis* and *Biotools* software (Bruker Daltonics) and submitted to *Mascot* for searching an in-house neuropeptide precursor database.

16). Select the following as possible post-translational modifications: pyroglutamic acid, carboxy-terminal amidation, acetylating, sulfatation and methionine-oxidation. Leave the fields of enzyme and fixed modifications empty. Peptide mass tolerance and MS/MS tolerance depends on the mass spectrometer used and the calibration and should be set accordingly.

5. *Decyder MS* 2.0 is a differential analysis software tool that also allows for easy visualisation of LC-MS data by creating 2D maps (*m/z* values vs. retention times). The data from the MicroTOF-Q can be converted to the mzXML format using *CompassXport*. The mzXML data format is compatible with *Decyder MS*. In the import module of the Decyder MS software, set the resampling values to proportional and set maximal resampling error to 20 ppm. The retention time is reduced, starting from minute 20 instead of minute 0 because most of the peptides elute in this time frame (*see* **Note 17**). The PepDetect module offers the visualisation and detection tool (**Fig. 15.2**). Peptide detection

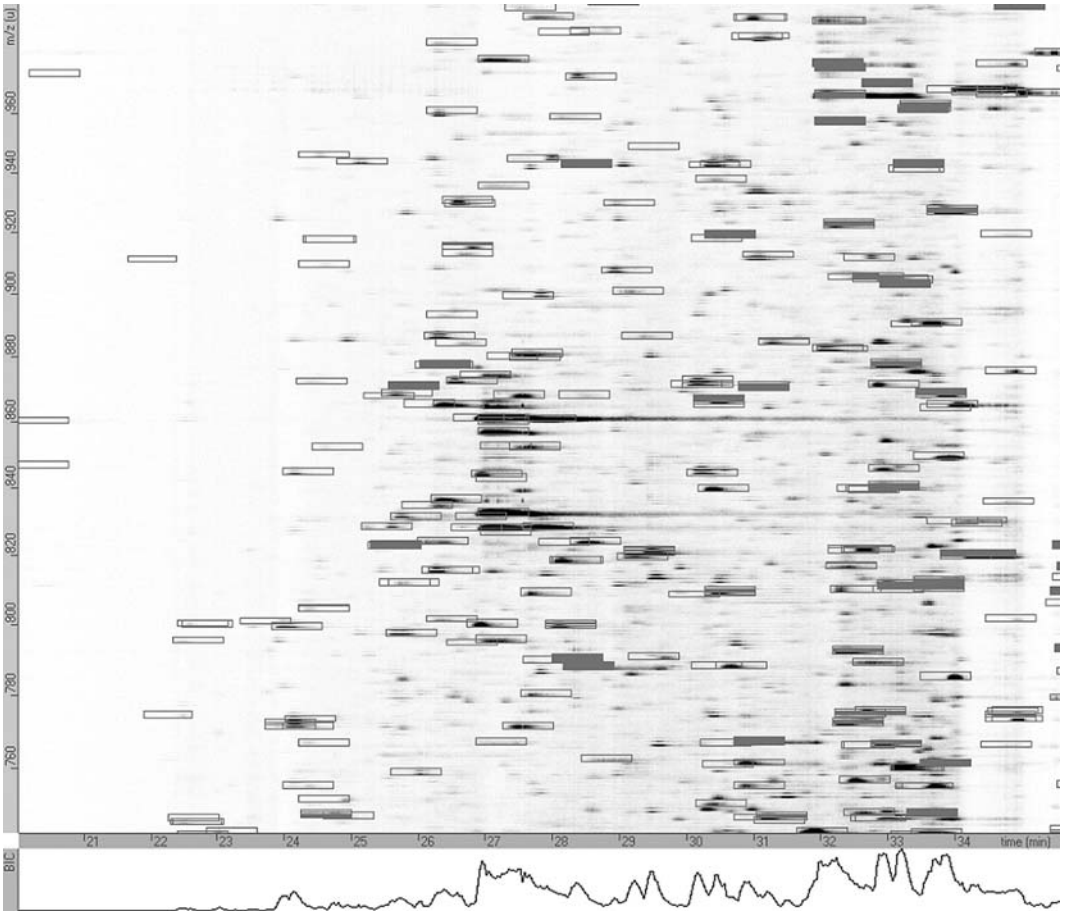


Fig. 15.2. Visualisation of 1D nanoLC MS data of the islets of Langerhans with *DeCyder MS* (GE Healthcare). An extract of 200 islets of Langerhans was separated on a Pepmap C18 column, using nanoLC system (Ultimate™ 3000 Nano and Capillary LC system). The nanoLC was directly coupled to the MicroTOF-Q for a MS survey run. The MS data were converted with the *CompassXport* software (Bruker Daltonics) to the mzXML format that is compatible with *DeCyder MS*. The *PepDetect* module was used to visualise the MS data. The *x*-axis represents the time and the *y*-axis the *m/z* values. Signal intensities are represented by grayscale (with black being the most intense). The map shown is from the time interval of 20 to 36 min and from *m/z* 740 to 1000. The total ion current (TIC) is depicted at the bottom. Over 600 peptides were detected using this approach.

parameters have to be entered. The typical elution peak width and MS resolution should be calculated for each run separately. Peptides with two to six charges are taken into account. Only peptides with a signal-to-noise threshold over 4 are considered. The detected peaks are evaluated subsequently. Editing can be done by altering the charge state, the mass or elution time. Redundant peptides are excluded. A two-dimensional visualisation and a working table of the peptides are presented after confirmation of the detection. The masses presented in these tables can be used to search sequence databases, e.g. *SwePep*

(www.swepep.org) or our home-made database (available from http://www.peptides.be). For the differential peptidomics analysis, the *PepDetect* files should be loaded into the *PepMatch* module. First, the experimental groups are defined. Then, the intensity maps are aligned and peptides are subsequently matched between the different runs. A maximum tolerance of 1 min for elution time and maximum *m/z* shift of 0.15 should be used. There are two different methods for data normalisation. One method normalises individual ion intensities relevant to the intensities of the whole peptide population. The other method relies on internal standards. The *PepMatch* software uses the internal standards intensities to normalise the data from the runs (*see* **Note 18**). Following the normalisation, the Student's *t*-test can be used chosen for a group-to-group comparison; alternatively, ANOVA can be used for multi-group comparisons (**Fig. 15.3**).

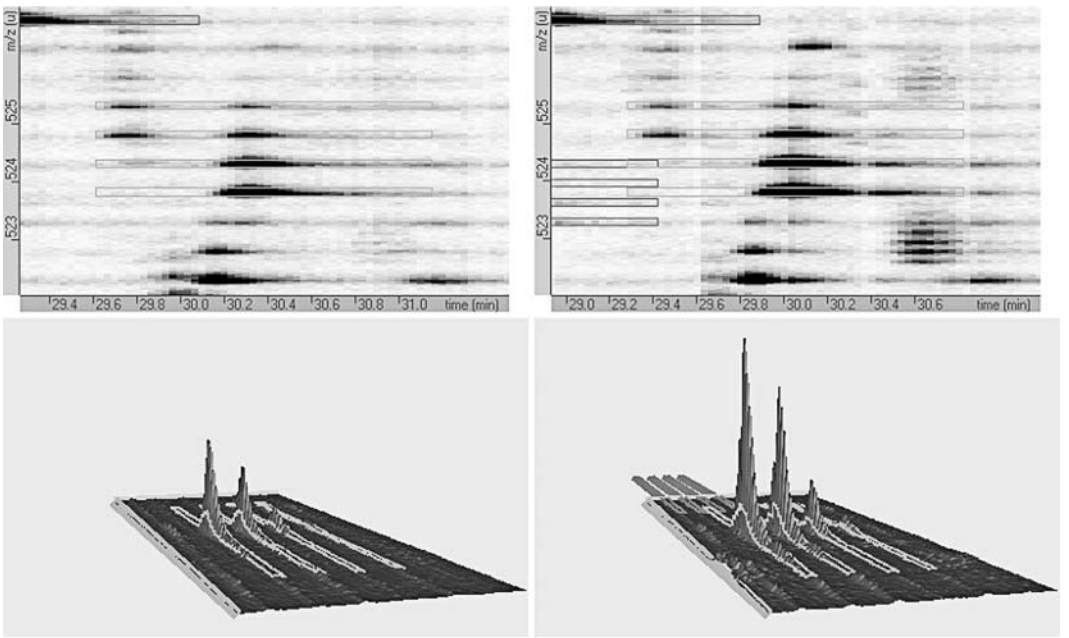


Fig. 15.3. Relative quantification of neuropeptides using *DeCyder MS* (GE Healthcare). One large extract of 800 islets of Langerhans was dissolved in 40 μl of 2% acetonitrile containing 1% FA and four 1-μl aliquots were taken. Samples A contained 50 fmol of the Bruker Peptide standard mixture and 100 fmol of a peptide mixture containing kinetensin, neurotensin, neuromedin and oxytocin. Samples B contained 100 fmol of the Bruker Peptide standard mixture and 50 fmol of the peptide mixture. The four samples (with experimental groups A and B) were analysed with 1D nanoLC Q-TOF MS and the data were loaded into *DeCyder MS*. A statistical analysis was performed with the *PepMatch* Module. The data were normalised using measured intensity distribution and the peptides of the four runs were matched. A *t*-test was chosen for the statistical analysis of the two experimental groups. Angiotensin 2 is depicted in the figure. The average linear ratio of the normalised intensity of angiotensin 2 is 0.62 (A compared to B) with a *t*-test *p*-value of 0.012.

4. Notes

1. Make sure that all the reagents are of the highest purity. Low-quality solvents and Eppendorf tubes may contain polymers that elevate background in MS analyses, interfering with the detection and quantification of peptide ions. Regular Eppendorf tubes are suitable, siliconised tubes should be avoided.
2. Use cooled forceps to handle frozen tissues.
3. The resolution of the peptide separation deteriorates if the nanoLC column is overloaded. Insulin is a highly abundant polypeptide and wash steps between subsequent runs are recommended. An alternative solution is to use 3 kDa cut-off filter (e.g. the 3 kDa Centriplus filter (Amicon, Danvers, USA)). These will also eliminate a number of other regulatory peptides from the sample, like e.g. glucagon and peptide YY.
4. Islets of Langerhans adhere to transfer pipettes. This is especially the case when they are suspended in the dextran stock solution. The contact between the transfer pipettes and the islets in solution should therefore be minimal.
5. The islets are mostly found between layers II and III, but the presence of islets between layers I and II is also possible. Islets of Langerhans are easily visible against a dark background as they are white and not transparent (as the exocrine tissue is). Islets can be picked up with a 10-μl pipette. Diluting the dextran solution after transferring them to the Petri dish makes picking easier. The addition of foetal calf serum also reduces sticking of the islets to the Petri dish.
6. Alternatively, the islets can be kept in RPMI 1640 medium with 10 mmol/L HEPES, 10% foetal calf serum and 100 μg/ml streptomycin for 24 h at 37°C. This way, the islets recover and peptide concentrations increase. Make sure that the islets are well rinsed since the foetal calf serum contains peptide growth factors.
7. Sonication enhances the extraction.
8. This is sometimes difficult due to impurities in the sample. In such a case the samples should be dissolved in 50 μl and cleaned by solid-phase extraction with Ziptip C18 (15 μl, Millipore). The eluted sample can be concentrated under vacuum, but should not be dried completely. Typically, 30–60 s of vacuum concentration is sufficient to remove

organic solvents, thereby leaving the peptides in the aqueous solvents.

9. The first separation dimension is a strong cation exchange. This chromatography step is performed on the Beckman HPLC system. This has several advantages, including the increased capacity of the column (so it is possible to use larger amounts of starting material) and salt gradient elution (improves the resolution compared to elution with salt plugs).
10. The eluent is emitted through a stainless steel emitter (Proteon) and nitrogen is used as nebulising and drying gas.
11. Further details on Q-TOF analysis can be found in **Section 3.3.2** (**Chapter 3**, Husson et al).
12. We prefer to use Bruker Ultraflex II: MALDI-TOF/TOF over the Reflex IV MALDI-TOF instrument.
13. Alternatively, the dried peptide sample may be dissolved in 50% acetonitrile and 0.1% FA.
14. It is possible to use other matrices such as sinapinic acid and 2,5,dihydroxybenzoic acid; these would suite longer peptides. The use of (pre-spotted) AnchorChip targets will result in a significant increase in sensitivity. Several protocols for optimising mass spectrometric analysis of proteins and peptides are available, e.g. from 24.
15. 1D or 2D nanoLC MS/MS experiments create a large amount of mass spectrometric data. MS instrument manufacturers provide software packages for the automatic processing of the MS data. The peak lists obtained with these tools are usually compatible and can be submitted to search engines like *Mascot*, *Sequest* or *MS-Tag* (Protein prospector, http://prospector.ucsf.edu/).
16. A local *Mascot* server has the additional advantage of allowing to use own sequence databases. We use several homemade neuropeptide and neuropeptide precursor databases: a database containing all mouse peptides found by the approach of Feng (25), a neuropeptide precursor database and a neuropeptide database containing the in silico spliced precursors are available (*See* **Chapter 25** by Clynen et al.)
17. The elution time will depend on the experimental setup. Choosing a narrow time frame speeds up the processing time. Too narrow time frames can cause inaccuracies when using the total ion current normalisation.
18. The type of normalisation depends on the experimental setup. Total ion current normalisation is applicable only if the total concentration of peptides is approximately the same for the experimental groups. This can for example be

altered if proteolytic peptides are present in only one condition. Total ion current normalisation takes many sources of error into account (like sample handling, tissue weight, LC-MS performance etc.) and is therefore more accurate than internal standard normalisation. Internal standard should be added as soon as possible in the peptide extract in order to avoid sample-handling errors.

Acknowledgments

K. Boonen and B. Landuyt are supported by grants of Institute for the Promotion of Innovation through Science and Technology (I.W.T.)-Flanders. S. J. Husson is a postdoctoral fellow of the Research Foundation Flanders (F.W.O.-Vlaanderen). The authors also acknowledge the Interfacultary Centre for Proteomics and Metabolomics "Prometa", K.U. Leuven.

References

1. Strand, F.L. (1999) Neuropeptides. Cambridge, The MITT Press.
2. Hökfelt, T., Bartfai, T. and Bloom, F. (2003) Neuropeptides: opportunities for drug discovery. *Lancet Neurol.* **2**, 463–472.
3. Fricker, L.D. (2005) Neuropeptide-processing enzymes: applications for drug discovery. *AAPS J.* **7**, E449–E455.
4. Boonen, K., Landuyt, B., Baggerman, G., Husson, S.J., Huybrechts, J. and Schoofs, L. (2008) Peptidomics: the integrated approach of MS, hyphenated techniques and bioinformatics for neuropeptide analysis. *J. Sep. Sci.* **31**, 427–445.
5. Soloviev, M. and Finch, P. (2005) Peptidomics, current status. 2. *J. Chromatogr B. Analyt. Technol. Biomed. Life Sci.* **815**, 11–24.
6. Ivanov, V.T. and Yatskin, O.N. (2005) Peptidomics: a logical sequel to proteomics. *Expert Rev. Proteomics.* **2**, 463–473.
7. Svensson, M., Sköld, K., Nilsson, A., Fälth, M., Svenningsson, P. and Andrén, P.E. (2007) Neuropeptidomics: expanding proteomics downwards. *Biochem. Soc. Trans.* **35**, 588–593.
8. Desiderio, D.M. (1996) Mass spectrometry, high performance liquid chromatography, and brain peptides. *Biopolymers.* **40**, 257–264.
9. Baggerman, G., Cerstiaens, A., De Loof, A. and Schoofs, L. (2002) Peptidomics of the larval *Drosophila melanogaster* central nervous system. *J. Biol. Chem.* **277**, 40368–40374.
10. Clynen, E., Baggerman, G., Veelaert, D., Cerstiaens, A., Van der Horst, D., Harthoorn, L., *et al.* (2001) Peptidomics of the pars intercerebralis–corpus cardiacum complex of the migratory locust, *Locusta migratoria*. *Eur. J. Biochem.* **268**, 1929–1939.
11. Predel, R. and Gäde, G. (2002) Identification of the abundant neuropeptide from abdominal perisympathetic organs of locusts. *Peptides* **23**, 621–627.
12. Minamino, N., Tanaka, J., Kuwahara, H., Kihara, T., Satomi, Y., Matsubae, M., *et al.* (2003) Determination of endogenous peptides in the porcine brain: possible construction of peptidome, a fact database for endogenous peptides. *J. Chromatogr B. Analyt. Technol. Biomed. Life Sci.* **792**, 33–48.
13. Sköld, K., Svensson, M., Kaplan, A., Björkesten, L., Aström, J. and Andren, P.E. (2002) A neuroproteomic approach to targeting neuropeptides in the brain. *Proteomics* **2**, 447–454.
14. Che, F.Y., Yan, L., Li, H., Mzhavia, N., Devi, L.A. and Fricker, L.D. (2001) Identification of peptides from brain and pituitary

of Cpe(fat)/Cpe(fat) mice. *Proc. Natl. Acad. Sci. USA* **98**, 9971–9976.
15. Svensson, M., Sköld, K., Svenningsson, P. and Andren, P.E. (2003) Peptidomics-based discovery of novel neuropeptides. *J. Proteome Res.* **2**, 213–219.
16. Che, F.Y., Lim, J., Pan, H., Biswas, R. and Fricker, L.D. (2005) Quantitative neuropeptidomics of microwave-irradiated mouse brain and pituitary. *Mol. Cell. Proteomics.* **4**, 1391–1405.
17. Boonen, K., Husson, S.J., Baggerman, G., Cerstiaens, A., Luyten, W. and Schoofs, L. (2008) Peptidomics in neuroendocrine research: A *Caenorhabditis elegans* and *Mus musculus* study. In: Peptidomics: Methods and Applications, pp 355–386. Wiley, Hoboken, New Jersey.
18. Dowell, J.A., Heyden, W.V. and Li, L. (2006) Rat neuropeptidomics by LC-MS/MS and MALDI-FTMS: Enhanced dissection and extraction techniques coupled with 2D RP-RP HPLC. *J. Proteome Res.* **5**, 3368–3375.
19. Boonen, K., Baggerman, G., D'Hertog, W., Husson, S.J., Overbergh, L., Mathieu, C., et al. (2007) Neuropeptides of the islets of Langerhans: a peptidomics study. *Gen. Comp. Endocrinol.* **152**, 231–241.
20. Conlon, J.M. (1997) Preparation of neuropeptide-containing fractions from biological materials. In: Neuropeptide Protocols, pp 1–8. Humana Press, Totowa, New Jersey.
21. Che, F.Y., Zhang, X., Berezniuk, I., Callaway, M., Lim, J. and Fricker, L.D. (2007) Optimization of neuropeptide extraction from the mouse hypothalamus. *J. Proteome Res.* **6**, 4667–4676.
22. Cutillas, P.R. (2008) Quantification of polypeptides by mass spectrometry. In: Peptidomics: Methods and Applications, pp 291–316. Wiley, Hoboken, New Jersey.
23. Johansson, C., Samskog, J., Sundström, L., Wadensten, H., Björkesten, L. and Flensburg, J. (2006) Differential expression analysis of *Escherichia coli* proteins using a novel software for relative quantitation of LC-MS/MS data. *Proteomics* **6**, 4475–4485.
24. Chapman, J.R. (ed.) (2000) Methods in molecular biology: Mass spectrometry of proteins and peptides. Humana Press, New Jersey.
25. Liu, F., Baggerman, G., Schoofs, L. and Wets, G. (2008) Construction of a database of signalling peptides in Metazoa. *J. Proteome Res.* 7, 4119–4131.

Chapter 16

Peptidome Analysis of Mouse Liver Tissue by Size Exclusion Chromatography Prefractionation

Lianghai Hu, Mingliang Ye, and Hanfa Zou

Abstract

Here we report our approach to the peptidomic analysis of mouse liver. We use ultrafiltration for peptide prefractionation, which is followed by size exclusion chromatography. The low molecular weight peptides (MW below ~3 kDa) are analysed directly by nanoLC-MS/MS, and the higher molecular weight peptides (MW above ~3 kDa) are characterized with MALDI-TOF MS first and then proteolytically digested prior to nanoLC-MS/MS analyses.

Key words: Mouse liver, peptidomics, ultrafiltration, size exclusion chromatography, multidimensional separation, mass spectrometry.

1. Introduction

The unprecedented growth and development of separation and detection technologies in recent years has led to rapid progress in proteomic research (1, 2). However, only a few proteins have been validated as disease biomarkers because of the low abundance of the potential biomarkers and the complexity of biological samples (3). The low molecular weight (LMW) fraction of proteome (termed "peptidome"), considered previously as biological debris (4), has attracted increasing attention recently (5). The area of research involved in comprehensive study of peptides or LMW proteins expressed by a cell, tissue and organs of an organism has become known as Peptidomics (6). The peptidome analysis mainly focuses on the quantitative and qualitative analysis of peptides, which can be divided into two classes: (I) the

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_16,

endogenous peptides that exert vital functions in biological process such as hormones, cytokines, growth factors, MHC class I peptides and alike (6, 7); (II) the degraded fragments of proteins which reflect the proteolytic enzyme species and biological state of individual (8, 9). The endogenous peptides play crucial roles in the respiratory, cardiovascular, endocrine, inflammatory and nervous systems (7, 10). Discovering novel neuropeptides has been extensively studied and some databases have been established for the endogenous peptides (11, 12). The degraded fragments of proteins are generated by the proteolytic enzymes and can be considered as the metabolic products of proteins. Circulating protein fragments generated in the body fluid or tissues may reflect the biological events and provide a rich bank for diagnostic biomarkers (13). It is believed that peptide concentration in tissues should be higher than that in the blood and thus screening for peptide biomarkers in tissues may be another way for speeding up the biomarker discovery (14). However, unlike the abundance of publications on the peptidomic analysis of body fluids, few peptidomic analyses from other tissues were reported (15). Liver is a vital organ, which is considered as the main "chemical factory" and "energy plant" for the body (16) and can therefore provide a rich source of peptides indicative of body metabolism and hepatic function. NanoLC-MS/MS (nano-liquid chromatography tandem mass spectrometry) has high detection sensitivity and is capable of high throughput, which makes this technique suitable for the peptidomic analyses. However, few peptides over 3000 Da can be characterized directly. To fully characterize the peptidome over the whole range of molecular weights, we developed a comprehensive method which relies on a simple but highly reproducible ultrafiltration step to extract the liver peptidome, followed by prefractionation of the peptidome using size exclusion chromatography and nanoLC-MS/MS (17). The flow chart summarizing the procedure is shown in **Fig. 16.1**.

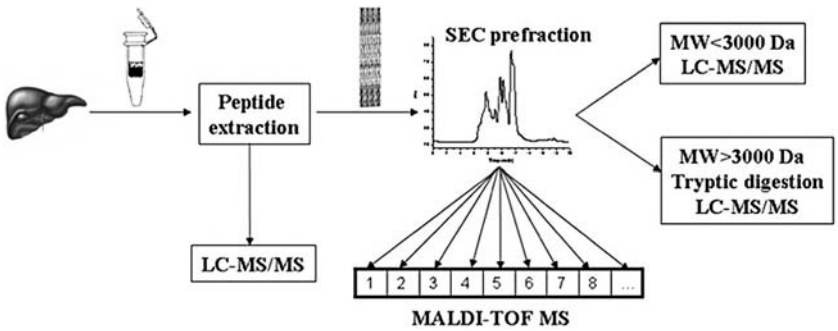


Fig. 16.1. Outline of the experimental approach to the comprehensive peptidomic analysis of mouse liver tissue.

2. Materials

1. HPLC pump: LC-10ADvp pump (Shimadzu, Kyoto, Japan). Large variety of instruments exist and any instrument which can operate at low flow rates (<1 mL/min) and tolerate pressure of 2000 psi can be used.
2. UV detector: SPD-M10Avp UV-vis detector (Shimadzu, Kyoto, Japan). Other suitable detectors with adjustable wavelength of 200–400 nm may be used.
3. Size exclusion column: TSK SuperSW 2000 (4 μm, 125 Å, 4.6 mm i.d. × 300 mm; TOSOH, Tokyo, Japan) (*see* **Note 1**).
4. Capillary separation column: Fused-silica AQ C18 packed capillary (5 μm, 120 Å, 75 μm i.d. × 120 mm; Michrom BioResources, CA, USA).
5. Mobile phases. Solvent 1: 45% Acetonitrile (ACN) in 0.1% trifluoroacetic acid (TFA). Solvent 2: 0.1% Formic acid in water. Solvent 3: 0.1% Formic acid in acetonitrile. All reagents should be chromatographic grade. High-purity deionized water should be used in all experiments, e.g. purified with Milli-Q system (Milford, MA, USA) (*see* **Note 2**).
6. Peptide extraction. Extraction buffer: 0.25% Acetic acid. Adult female C57 mice. High-purity deionized water should be used in all experiments, e.g. purified with Milli-Q system (Milford, MA, USA). Centrifugal filter: Amicon Ultra-15, (Millipore, Milford, MA, USA) or equivalent centrifugal filter with a nominal molecular mass limit of 10 kDa.
7. Standard proteins: 10 mg/mL cytochrome *c* in Solvent 1; 10 mg/mL insulin in Solvent 1; 10 mg/mL insulin chain B in Solvent 1.
8. Peptide digestion: 1 μg/μL TPCK-treated trypsin; 1 M dithiothreitol (DTT); 1 M iodoacetamide (IAA); 50 mM ammonium bicarbonate.
9. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS): Bruker Autoflex(TM) (Bruker, Bremen, Germany). This instrument is equipped with a nitrogen laser ($\lambda = 337$ nm) and the accelerating potential is in the range of +20/–20 kV. MALDI stainless-steel sample target (MTP 384). All mass spectra should be obtained in the positive-ion detection mode. MALDI-TOF MS is used for determining the molecular

weight of peptides in fractions obtained after size exclusion chromatography (SEC).

10. MALDI-TOF MS reagents. Matrix solvent: 33% ACN/0.1% TFA (33%/67%, v/v). Matrix: 20 mg/mL 2,5-dihydroxybenzoic acid (DHB) in matrix solvent.
11. High-performance liquid chromatography (HPLC) MS/MS: Finnigan surveyor MS pump (ThermoFinnigan, San Jose, CA) and LTQ linear ion trap mass spectrometer ThermoFinnigan (San Jose, CA, USA) with a nanospray source. HPLC MS/MS is used for separation and sequencing of peptides.
12. *SEQUEST* database search and data analysis software (Thermo Electron, San Jose, CA).

3. Methods

3.1. Sample Preparation (see Note 3*)*

1. Sacrifice the mouse and remove the liver, wash it in extraction buffer to remove traces of blood.
2. Mince the liver with scissors on ice and homogenize with extraction buffer at 4°C. Use 5 ml of extraction buffer per gram of liver (equivalent concentration 0.2 g of tissue per mL of buffer).
3. Sonicate the suspension for 90 s at 450 W at 4°C (*see* **Note 4**).
4. Centrifuge the suspension at 25,000×*g* at 4°C for 1 h.
5. Transfer the supernatant to Amicon Ultra-15USA and centrifuge at 5000×*g* for 30 min at 4°C (*see* **Note 5**).
6. Collect and lyophilize the filtrate.
7. Redissolve the lyophilized peptides sample in Solvent 2 (use 1 ml per 6 g of tissue) and store at –20°C until use.

3.2. SEC Prefractionation of the Peptide Sample (see Note 6*)*

1. Equilibrate HPLC system and the TSK SuperSW 2000 column with Solvent 1.
2. Set flow rate to 0.35 ml/min, use isocratic elution (*see* **Note 7**).
3. UV detection should be set to 214 nm.
4. Use standard proteins (cytochrome *c*, insulin and insulin chain B) for evaluating the separation efficiency. An example of the chromatogram is shown in **Fig. 16.2a**, where the standard proteins are well separated from each other.

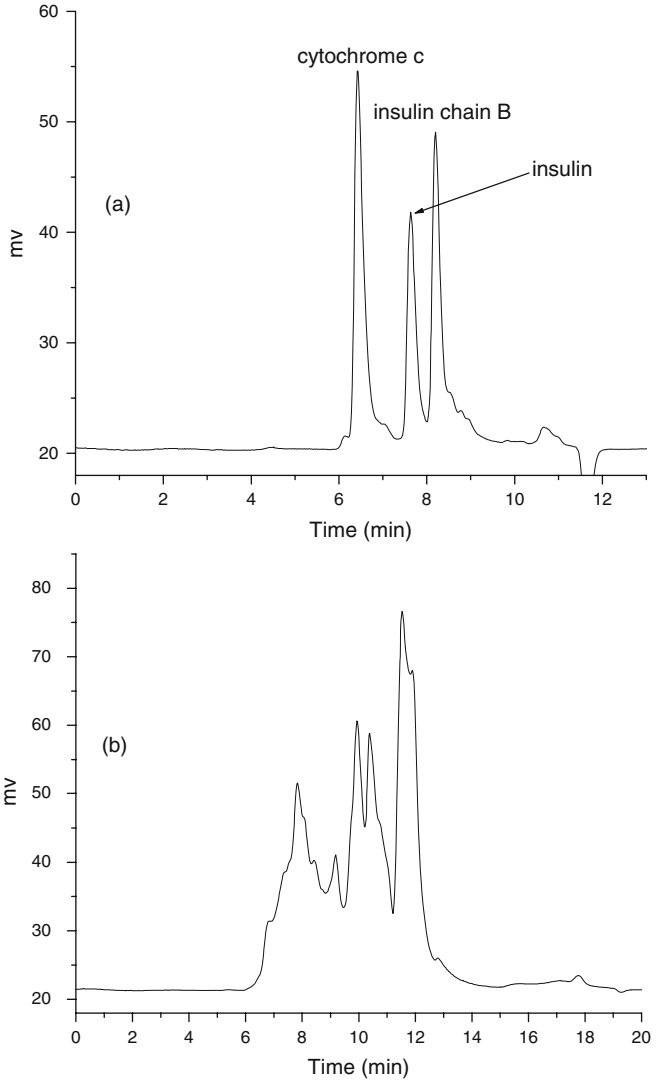


Fig. 16.2. Size exclusion chromatography separation of (**a**) standard sample containing cytochrome *c* (12327 Da), insulin (5734 Da) and insulin chain B (3496 Da) and (**b**) the extracted peptides from mouse liver tissue. Mobile phase: 45% ACN in 0.1% TFA (isocratic elution); flow rate 0.35 ml/min; UV detection at 214 nm.

5. Separate the extracted peptides (from Step 7, **Section 3.1**) on the SEC column.
6. Collect the eluted peptides (~200 μl fractions or smaller, approximately 30 s per sample). An example of the chromatogram is shown in **Fig. 16.2b**.
7. Freeze-dry the collected samples.
8. Redissolve peptides in Solvent 2 and store at –20°C until use.

3.3. MALDI-TOF MS

Prepare matrix solution:

1. Add 3 μL of Matrix (20 mg/ml DHB) to 1 μL of the peptide sample.
2. Deposit 0.5 μL of the mixture on the MALDI target plate, air dry.
3. Perform MALDI analysis (*see* **Note 8**). An example of mass spectra obtained for different fractions is shown in **Fig. 16.3** (*see* **Note 9**).

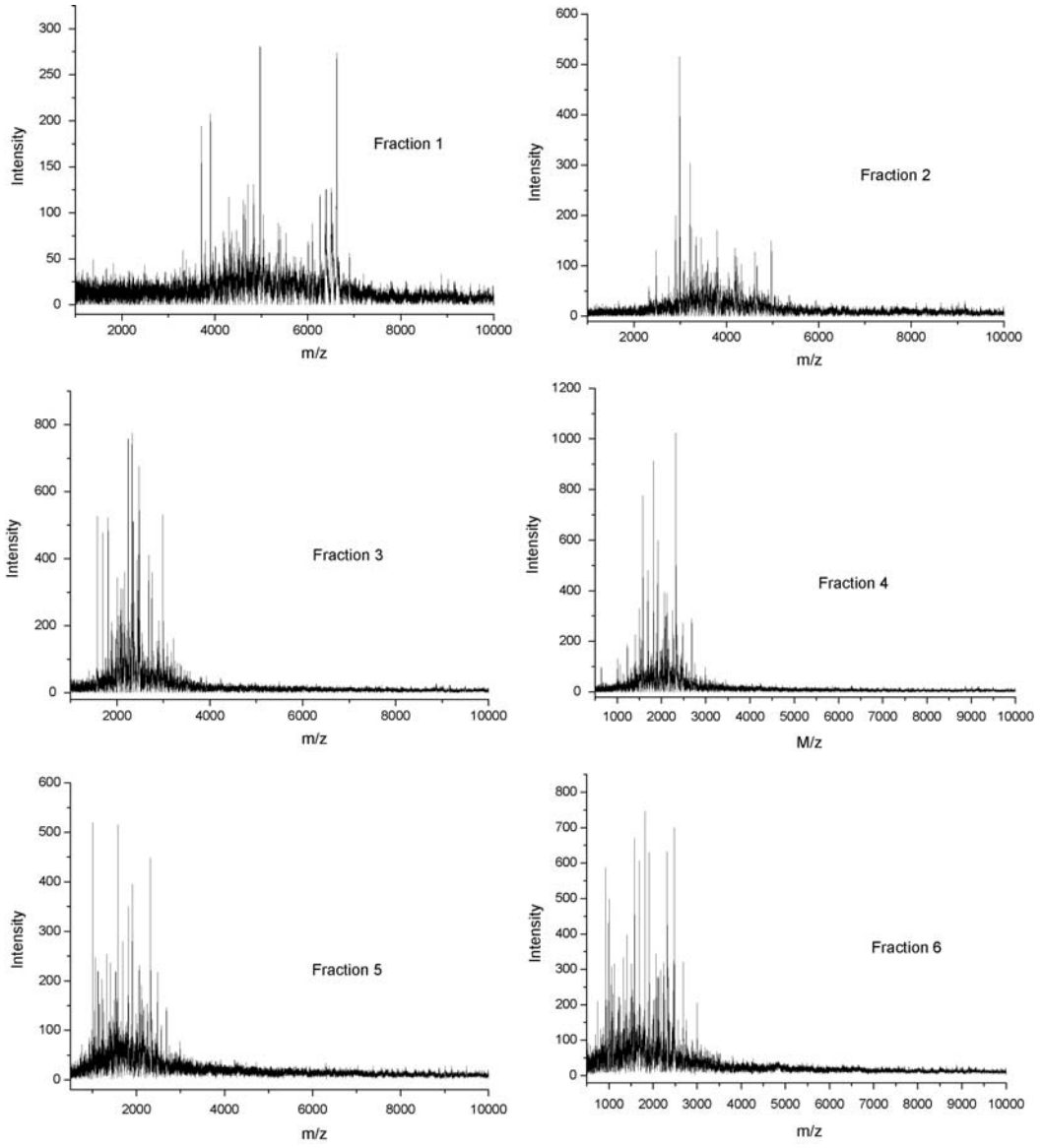


Fig. 16.3. MALDI-TOF MS analysis of the peptide fractions eluted from the size exclusion chromatography column (six consecutive fractions).

3.4. Digestion of the HMW Peptides (see Note 10)

1. Redissolve high molecular weight fractions (MW above ~ 3000 Da) collected from SEC column in 150 μL solution of 50 mM ammonium bicarbonate.
2. Add 1 μL of 1 M DTT and incubate the mixture at 37°C for at least 2 h.
3. Add 1 μL of 1 M IAA to the mixture and incubate the mixture at room temperature for 30 min in the dark.
4. Add 1 μL of 1 μg/μL trypsin and incubate the mixture at 37°C overnight (*see* **Note 11**).
5. Freeze-dry the digests; redissolve in 5 μL of Solvent 2 for LC-MS/MS analysis.

3.5. NanoLC-MS/MS

1. Set the column flow rate to ~200 nL/min (*see* **Note 12**).
2. The μRPLC column should be coupled directly to a LTQ linear ion trap mass spectrometer with a nanospray source.
3. The LTQ instrument should be operated at positive-ion detection mode. A voltage of 1.8 kV is applied to the cross. Set the capillary temperature to 170°C. Set the normalized collision energy to 35.0.
4. The mass spectrometer should be set at one full MS scan followed by ten MS/MS scans on the ten most intense ions from the MS spectrum.
5. Inject 1 μL of the fractions eluted from the SEC columns or the fractions digested with trypsin into C18 nanoLC column for nanoLC-MS/MS analysis.
6. Use gradient elution (Solvent 3/Solvent 2) as follows: (2%/98% to 10%/90%) for 3 min, followed by (10%/90% to 35%/65%) for 33 min, followed by (35%/65% to 80%/20%) for 2 min and maintain the flow at (Solvent 3/Solvent 2 = 80%/20%) for 10 min. Re-equilibrate the column as follows: use a fast gradient (80%/20% to 2%/98%) for 3 min, then maintain flow at (Solvent 3/Solvent 2 = 2%/98%) for 9 min.

3.6. Data Processing and Analysis

1. Peptides can be identified by searching against sequence databases. We use ***SEQUEST*** database search and data analysis software and download sequence data from ftp://ftp.ebi.ac.uk/pub/databases/IPI/old/MOUSE/ipi.MOUSE.v3.08.fasta.gz.
2. Search parameters should be set as follows: no enzyme, set variable modification to oxidation of Met. If the samples were digested with trypsin, set fixed modification to carbamidomethylation of Cys and specify the enzyme as partially tryptic (*see* **Note 13**).
3. Specify the mass as monoisotopic.

4. Combine output data and remove keratins and any redundant data.
5. Filter the search results by setting lowest X_{corr} as 1.9, 2.2 and 3.75 (corresponding to 1+, 2+ and 3+ charge states, respectively).
6. A minimum delta correlation (ΔC_n) of 0.2 is required for the identification to be considered positive. Determine the false discovery rate (FDR) of the peptide identification (*see* **Note 14**).

4. Notes

1. There are different SEC columns for different samples with different ranges of molecular weight of proteins. For example, in the TSK series there are TSK 2000 (MW range 500–100,000 Da), TSK 3000 (2,000–500,000 Da), TSK 4000 (20,000–7,000,000 Da), etc.
2. All the mobile phases for chromatography should be filtered through a 0.22-μm or 0.45-μm membrane to avoid clogging the columns.
3. Extraction is the first step in the peptidomic analysis, and, therefore, the reproducibility of extraction is crucial in determining the outcome of the whole process. All procedures should be carried at temperatures below 4°C to reduce protein degradation. Delays and unnecessary sample storage steps should be avoided.
4. The tube should be immersed in ice water and the liquid level of the suspension should be lower than the ice water to keep the suspension at 4°C (7, 10, 18, 19).
5. The speed limit is different for different type of filters. The speed limit should be observed to avoid damaging the filter membrane.
6. LMW peptides (MW below ~3 kDa) should be separated from HMW components (MW above ~3 kDa) prior to the proteolysis or direct MS analysis.
7. If the SEC column has been used with other solvents such as the phosphate buffer, the column should be first equilibrated with water at a flow rate of 0.2 mL/min and then with the mobile phase (45% ACN in 0.1% TFA). For storage (over a week), the solvent should be replaced with 0.05% NaN_3 (use the same flow rate of 0.2 mL/min) to prevent bacterial growth.
8. Procedures outlined in the instrument manual must be followed.

9. At this stage one should expect to see different masses in different SEC samples. Larger molecular weight peptides should be eluted faster, whilst later collections should contain lower molecular weight peptides, as shown in **Fig. 16.3**.
10. In order to gain sequence information and identify the progenitor proteins, HMW polypeptides should be proteolytically digested prior to the nanoLC-MS/MS analysis.
11. The amount of trypsin depends on the size of sample. Typical weight ratio for trypsin:peptide is 1:50.
12. The pump flow may be split using a micro-splitter valve to achieve the required flow rate of ~200 nL/min.
13. For searching and identification of the modified peptides, search parameters (i.e. modifications) should be set accordingly.
14. FDR can be determined by performing *SEQUEST* search against a composite database that includes regular as well as reversed protein sequences. FDR = $2 \times n(\text{rev})/(n(\text{rev})+n(\text{forw}))$, where n(forw) and n(rev) are the number of peptides identified in proteins with forward (normal) and reversed sequence, respectively. (20, 21) FDR is an important parameter for the evaluation of mass matching results. Settings should be chosen such that FDR is kept below 5%.

Acknowledgments

This work was supported by National Natural Sciences Foundation of China (No. 20735004), the China High Technology Research Program Grant (2006AA02A309) and the Knowledge Innovation program of DICP to H. Zou.

References

1. Washburn, M.P., Wolters, D. and Yates, J.R. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19**, 242–247.
2. Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**, 198–207.
3. Ludwig, J.A. and Weinstein, J.N. (2005) Biomarkers in cancer staging, prognosis and treatment selection. *Nat. Rev. Cancer* **5**, 845–856.
4. Novak, K. (2006) Biomarkers: Taking out the trash. *Nat. Rev. Cancer* **6**, 92.
5. Soloviev, M. and Finch, P. (2006) Peptidomics: Bridging the gap between proteome and metabolome. *Proteomics* **6**, 744–747.
6. Schrader, M. and Schulz-Knappe, P. (2001) Peptidomics technologies for human body fluids. *Trends Biotechnol.* **19**, S55–S60.
7. Svensson, M., Skold, K., Svenningsson, P. and Andren, P.E. (2003) Peptidomics-based

discovery of novel neuropeptides. *J. Proteome Res.* **2**, 213–219.
8. Petricoin, E.F., Ardekani, A.M., Hitt, B.A., Levine, P.J., Fusaro, V.A., Steinberg, S.M., Mills, G.B., Simone, C., Fishman, D.A., Kohn, E.C. and Liotta, L.A. (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* **359**, 572–577.
9. Villanueva, J., Philip, J., Entenberg, D., Chaparro, C.A., Tanwar, M.K., Holland, E.C. and Tempst, P. (2004) Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. *Anal. Chem.* **76**, 1560–1570.
10. Skold, K., Svensson, M., Kaplan, A., Bjorkesten, L., Astrom, J. and Andren, P.E. (2002) A neuroproteomic approach to targeting neuropeptides in the brain. *Proteomics* **2**, 447–454.
11. Fricker, L.D., Lim, J.Y., Pan, H. and Che, F.Y. (2006) Peptidomics: identification and quantification of endogenous peptides in neuroendocrine tissues. *Mass Spectrom Rev.* **25**, 327–344.
12. Minamino, N., Tanaka, J., Kuwahara, H., Kihara, T., Satomi, Y., Matsubae, M. and Takao, T. (2003) Determination of endogenous peptides in the porcine brain: possible construction of Peptidome, a fact database for endogenous peptides. *J. Chromatogr. B* **792**, 33–48.
13. Liotta, L.A. and Petricoin, E.F. (2006) Serum peptidome for cancer detection: spinning biologic trash into diagnostic gold. *J. Clin. Invest.* **116**, 26–30.
14. Cottingham, K. (2006) Speeding up biomarker discovery. *J. Proteome Res.* **5**, 1047–1048.
15. Traub, F., Jost, M., Hess, R., Schorn, K., Menzel, C., Budde, P., Schulz-Knappek, P., Lamping, N., Pich, A., Kreipe, H. and Tammen, H. (2006) Peptidomic analysis of breast cancer reveals a putative surrogate marker for estrogen receptor-negative carcinomas. *Lab. Invest.* **86**, 246–253.
16. He, F.C. (2005) Human Liver Proteome Project – Plan, progress, and perspectives. *Mol. Cell. Proteomics* **12**, 1841–1848.
17. Hu, L.H., Li, X., Jiang, X.N., Jiang, X.G., Zhou, H.J., Kong, L., Ye, M.L. and Zou, H.F. (2007) Comprehensive peptidome analysis of mouse livers by size exclusion chromatography prefractionation and NanoLC-MS/MS identification. *J. Proteome Res.* **6**, 801–808.
18. Falth, M., Skold, K., Norrman, M., Svensson, M., Fenyo, D. and Andren, P.E. (2006) SwePep, a database designed for endogenous peptides and mass spectrometry. *Mol. Cell. Proteomics* **5**, 998–1005.
19. Baggerman, G., Verleyen, P., Clynen, E., Huybrechts, J., De Loof, A. and Schoofs, L. (2004) Peptidomics. *J. Chromatogr. B* **803**, 3–16.
20. Peng, J.M., Elias, J.E., Thoreen, C.C., Licklider, L.J. and Gygi, S.P. (2003) 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.
21. Xie, H., Bandhakavi, S. and Griffin, T.J. (2005) Evaluating preparative isoelectric focusing of complex peptide mixtures for tandem mass spectrometry-based proteomics: a case study in profiling chromatin-enriched subcellular fractions in Saccharomyces cerevisiae. *Anal. Chem.* **77**, 3198–3207.

Chapter 17

Rat Brain Neuropeptidomics: Tissue Collection, Protease Inhibition, Neuropeptide Extraction, and Mass Spectrometric Analysis

Robert M. Sturm, James A. Dowell, and Lingjun Li

Abstract

Due to the complexity of the mammalian central nervous system, neuropeptidomic studies in mammals often yield very complicated mass spectra that make data analysis difficult. Careful sample preparation and extraction protocols must be employed in order to minimize spectral complexity and enable extraction of useful information on neuropeptides from a given sample. Controlling post-mortem protease activity is essential to simplifying mass spectra and to identifying low-abundance neuropeptides in tissue samples. Post-mortem microwave-irradiation coupled with cryostat dissection has proven to be effective in arresting protease activity to allow detection of endogenous neuropeptides instead of protein degradation products.

Key words: Neuropeptide, extraction, microwave, protease activity, sample preparation, mass spectrometry, HPLC, peptide sequencing/identification.

1. Introduction

Neuropeptides are small (3–100 amino acid residues) endogenous biomolecules that have the ability to act as neurotransmitters, neuromodulators, or neurohormones in the nervous system. These biomolecules play a role in many physiological functions including feeding, sleeping, learning, pain, anxiety, circadian rhythms, and memory (1, 2). The biogenesis of neuropeptides occurs in the cell body of neurons. Here, pre-propeptides are synthesized in the rough endoplasmic reticulum (RER), secreted from the RER after the signal sequence

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_17,

is removed, and packaged into vesicles by the Golgi apparatus. Within these vesicles, propeptides are further processed and undergo post-translational modifications (i.e., glycosylation, C-terminal amidation, acetylation, phosphorylation, and disulfide bond formation) generating bioactive peptides (3).

The important biological role neuropeptides play has made them the target of many investigations. Many neuropeptides have been studied employing traditional techniques such as immunohistochemistry, radioimmunoassay, and Edman degradation. Although these techniques are valuable, mass spectrometry (MS) based techniques do not require a priori knowledge of peptide identity and allow for rapid elucidation of molecular species in a complex mixture. Mass spectrometry offers rapid and sensitive detection of ionizable species, but spectra can be complicated and low-abundance target species (i.e., neuropeptides) can be masked due to salts, lipids, and surfactants (4). Thus, appropriate sample preparation methods that allow preferential identification of neuropeptides with minimal interference are often key to successful MS-based studies.

Important considerations must be taken during sample collection and preparation in order to obtain useful information in any neuropeptidomic studies. The inhibition of active post-mortem proteases is one of the most important considerations a researcher must take into account. Once the animal is sacrificed and the tissue of interest harvested, proteases rapidly degrade larger proteins into smaller fragments that fall into the mass range of neuropeptides. These abundant protein fragments may suppress neuropeptide signals and make mass spectra interpretation very difficult. To minimize the spectra clouding associated with protein degradation, focused microwave-irradiation animal sacrifice (5–9), post-sacrifice microwave-irradiation of tissue (10), and cryostat dissection followed by a boiling extraction buffer (11) methods have all been used. Each of these techniques can effectively minimize the post-mortem protein degradation, but these techniques also possess their own drawbacks. Focused microwave-irradiation animal sacrifice, although an effective means to stop protease degradation, requires an expensive targeted microwave-emitting instrument and can introduce unnecessary stress on the animal which may affect neuropeptide expression. The use of a household microwave for post-mortem tissue fixation allows for the animal to be sacrificed by conventional methods, but may require a number of animals to develop a consistent protocol. Cryostat dissection followed by a boiling extraction buffer inhibits protease activity, but has a longer time gap between sacrifice and protease inhibition which allows some extracellular processing to occur.

Presented in this chapter is a neuropeptidomic procedure that utilizes a conventional microwave to inhibit proteases post-mortem, cryostat dissection to isolate specific tissue, acidified

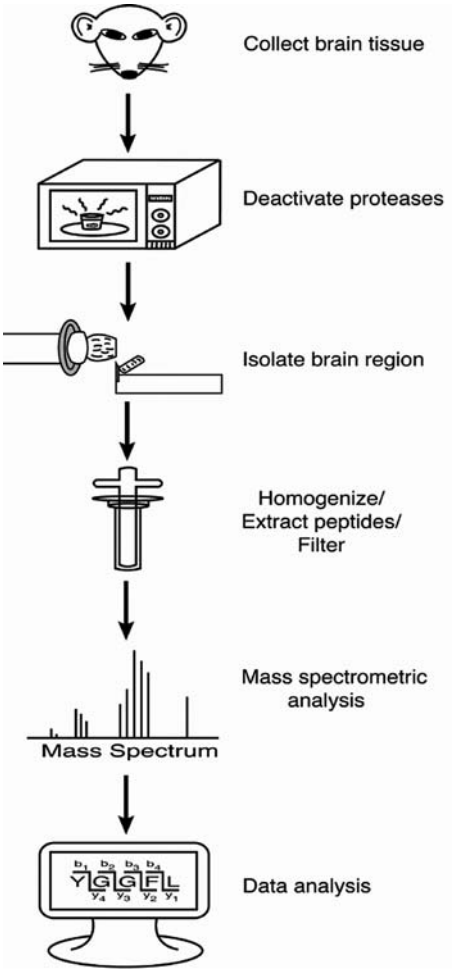


Fig. 17.1. Flowchart of mass spectrometry (MS) based neuropeptide analysis of the mammalian nervous system.

methanol to extract neuropeptides from tissue, and LC-QTOF-MS(/MS) to analyze the tissue extract. **Figure 17.1** depicts a flowchart of the MS-based neuropeptide analysis procedure.

2. Materials

2.1. Rat Dissection

1. Halothane (Sigma, St. Louis, MO)
2. Cryoware cryogenic vials (Nalgene, Rochester, NY)
3. 1.5 kW Microwave (General Electric)
4. Aluminium foil (ordinary kitchen foil is suitable)
5. Ethanol

6. Acidified methanol: 90% methanol, 9% glacial acetic acid, 1% doubly distilled water (v/v/v)
7. Dry ice
8. Cryostat (Leica, Wetzlar, Germany).
9. Optimal cutting temperature compound (OCT) (Sakura, Torrance, CA)
10. Harris micro-punch, 3 mm (Whatman, Clifton, NJ)

2.2. Neuropeptide Extraction

1. Methanol, purge and trap grade (Fisher Scientific, Fair Lawn, NJ)
2. Glacial acetic acid
3. Water, double-distilled by filtration system
4. Formic acid; store in a 4 °C refrigerator
5. 10 kDa molecular weight cut-off tube (Sartorius, Goettingen, Germany)
6. Handheld ground glass homogenizer (Wheaten Science, Millville, NJ)

2.3. Mass Spectrometry Analysis

1. Quadrupole time-of-flight mass spectrometer (QTOF MICRO) (Waters, Milford, MA), equipped with a nanoelectrospray ionization source (nESI) (*see* **Note 1**).
2. Ultra performance liquid chromatography nanoAcquity UPLC (Waters, Milford, MA) (*see* **Note 2**).
3. UPLC trap column Waters Symmetry® C_{18}, 180 μm × 20 mm (Waters, Milford, MA).
4. Analytical column Microtech C_{18}, 75 μm × 15 cm (Micro-Tech Scientific, Vista, CA).
5. Acetonitrile, HPLC grade (Fisher Scientific, Fair Lawn, NJ).
6. Mobile phase A: 0.1% Formic acid in double-distilled water (v/v).
7. Mobile phase B: 0.1% Formic acid in acetonitrile (v/v).
8. *ProteinLynx global server 2.3 (PLGS 2.3)* software for the processing of LC-MS/MS data (Waters).

3. Methods

3.1. Rat Dissection

1. Rats are placed in a plastic cylinder, anesthetized with halothane, and then sacrificed by decapitation.
2. Immediately (<60 s), the brain is removed from the rat and placed in a 15mL cryogenic vial and immersed in warm water (*see* **Note 3**).

3. The cryogenic vial is then placed in the center of a microwave and heated for 7s at full power. This allows the inside of the brain to rapidly reach a temperature of 80 °C (*see* **Notes 4 and 5**).
4. The cryojar is removed from the microwave, water is decanted off, and the brain is allowed to stand at room temperature for 1–3 min (*see* **Note 6**).
5. The brain is then snap-frozen in an ethanol/dry ice bath (*see* **Note 7**).
6. Place the snap-frozen brain on dry ice. At this point the tissue punches may be transferred to –80 °C for storage.
7. Mount the frozen brain tissue on a chilled cryostat chuck (same temperature as the cryostat compartment, –15 to –20 °C) with OCT compound and allow 1 min for the OCT compound to solidify. Add additional OCT compound to the base of the brain to provide additional support for cutting (*see* **Note 8**).
8. Section the brain into 300-μm slices. Advance the objective without slicing until the desired width is achieved (*see* **Note 9**).
9. Use a tissue punch to isolate brain regions of interest. (In our case we used a 3 mm tissue punch to collect hypothalamus and striatum regions of the rat brain.)
10. Place each tissue punch in a separate 1.5 mL microcentrifuge tube and add ice-cold acidified methanol to completely submerge the sample (*see* **Note 10**). At this point the tissue punches may be transferred to –80 °C for storage.

3.2. Neuropeptide Extraction

1. Place the frozen tissue directly into a 1 mL hand homogenizer containing 100–500 μL of ice-cold acidified methanol (*see* **Note 10**).
2. Immediately homogenize the tissue while keeping the temperature of the homogenate at 0 °C (*see* **Note 11, 12**).
3. Place the homogenate in a 1.5 mL microcentrifuge tube and centrifuge the solution at 14,000×*g* for 25 min at 4 °C (*see* **Note 13**).
4. Carefully transfer the supernatant to a clean microcentrifuge tube.
5. Re-extract neuropeptides from the tissue pellet by adding another aliquot of 100–500 μL of ice-cold acidified methanol, vortex and centrifuge for another 25 min.
6. Carefully transfer the supernatant to the microcentrifuge tube containing the first supernatant volume.

7. Filter the supernatants through a 10 kDa molecular weight cut-off (MWCO) microcentrifuge tube by centrifugation at 14,000×*g* for approximately 20 min at 4°C. The flow-through sample contains the neuropeptides (*see* **Note 14**).
8. Vacuum dry the sample.
9. Resuspend the extracted neuropeptides in 20–25 μL of aqueous 0.1% formic acid.
10. Mix by vigorous vortexing.
11. Centrifuge the re-suspended neuropeptides at 14,000×*g* for 3 min.
12. Transfer the supernatant to a clean 0.6 mL microcentrifuge tube.

3.3. Mass Spectrometry and Data Analysis

1. Set up the nanoAcquity UPLC system and load the sample onto a Waters Symmetry® C_{18} trap column. Use isocratic flow of mobile phase B (5%) and the flow rate of 10 μL/min for 1.5 min.
2. Switch the flow rate 200 nL/min, connect the Microtech C_{18} analytical column, and elute peptides into the nESI-QTOF mass spectrometer using a linear mobile phase gradient (A/B): 95/5% to 5/95% over 60 min (*see* **Note 15**).
3. Collect mass spectra for all eluted peptides in MS mode. Switch from the survey MS scan to MS^2 mode if the eluted peptide has an ion count of 15 or greater. Collect tandem mass spectra (MS^2) in data-dependent acquisition mode.
4. To process the LC-MS/MS data, one shall try using a software package supplied with the mass spectrometer. We used *ProteinLynx 2.1* and *Mascot* (*see* **Notes 16 and 17**).
5. To deisotope the data use the "slow" function and convert the results into *.pkl* files.
6. Search the newly created *.pkl* files against the Swiss-Prot database using a database search engine like the online version of *Mascot* (http://www.matrixscience.com) (*see* **Notes 16 and 17**).
7. In *Mascot* select "none" for the protease, set the peptide mass tolerance to 200 ppm and the MS/MS mass tolerance to 0.2 Da.
8. Repeat the search three times using different settings for the amino acid modifications:

   Search 1: Select C-terminal amidation and N-terminal acetylation.

   Search 2: Select methionine oxidation and C-terminal amidation.

   Search 3: Select phosphorylation of tyrosine, threonine, and serine and C-terminal amidation.

9. Peptide identifications that do not yield significant Mascot scores should be verified by de novo sequencing. (*see* **Note 18**).

4. Notes

1. Other mass spectrometers capable of automated acquisition of tandem mass spectra may be used for peptide mass fingerprinting and de novo sequencing.
2. High-performance liquid chromatography (HPLC) can be used in place of UPLC.
3. Protease degradation of proteins occurs immediately after death so the amount of time between sacrifice and protease inhibition should be as short as possible.
4. The microwave power may be different. For example, Che et al. demonstrated effective protease inhibition by placing a whole mouse brain into a 1.38 kW microwave (General Electric) for 5–13 s at full power. The main criterion for the microwave used is that it should have the ability to raise the tissue temperature to >80 °C within 10 s (10). The reproducibility of heating between samples can be further improved if brain tissue samples are placed in the same location inside the microwave.
5. The brain tissue is easier to manipulate if it is allowed to dry for a few minutes directly after microwave-irradiation.
6. Depending on the microwave, some method development may be needed to find the best time and temperature setting. The aim is to raise the temperature of the tissue sample to approximately 80 °C without altering the morphology of the brain.
7. Do not place the brain directly into the dry ice/ethanol bath to preserve tissue integrity. Loosely wrap the brain sample in aluminium foil and place it in the bath. Leave the sample in the bath for approximately 3 min to ensure the inside of the brain completely freezes. Do not unwrap the sample during that incubation.
8. The OCT compound may contain polymer contaminants and could therefore interfere with the mass spectrometric analysis. One should avoid applying OCT to the tissue areas meant for the MS analysis. For example, if one wants to analyze striatum and hypothalamus punches, the brain should be mounted such that the cerebellum is in contact with the cryostat chuck and the OCT. Avoid getting OCT

compound on the cryostat blade. If the blade is contaminated with OCT compound, it should be either cleaned thoroughly with 100% ethanol or replaced with a fresh one.

9. Alternatively, a rat brain matrix (Zivic Instruments; Pittsburgh, PA) may be used to section the brain. If the latter approach is selected, it is recommended that the sectioning is performed on dry ice so the brain does not thaw.
10. Minimal volume of acidified methanol is recommended. The key is to use a volume that allows for complete homogenization without over-diluting the sample. For example, Dowell et al. (11) used an aliquot of 300 μL of ice-cold acetic acid to homogenize a striatum punch weighing 20–30 mg.
11. It is important to immediately homogenize the tissue in order to ensure complete deactivation of all the protease activity.
12. A microsonicator or electric homogenizer may also be used.
13. A Pasteur pipette works well for transferring the homogenate.
14. Prior to use, clean the MWCO tubes with doubly distilled water followed by cold acidified methanol in order to remove all glycerol from MWCO membranes. Spin MWCO tubes for 3 min at 14,000×*g* after each wash and discard the flow-through. Repeat each washing twice.
15. Other MS techniques could be used to detect neuropeptides in brain tissue samples. Offline fractionation of a tissue extract has been used to simplify complex biological samples (11). Matrix-assisted laser desorption/ionization (MALDI) has also proven effective in identifying neuropeptides in mammalian biological samples (9, 11–13). In addition to providing a complementary ionization method, MALDI-based techniques also offer exciting capabilities for MALDI imaging by performing neuropeptide mapping in situ directly from tissue slices (14–17).
16. Bioactive neuropeptides are often found to undergo extensive proteolytic cleavages or post-translational modifications, making it difficult to identify the protein precursor from which a neuropeptide originates. Recently a binary logistic regression model (18) trained on mammalian prohormone cleavages has been developed that helps determine novel bioactive peptides from an organism's genetic sequence information. Once optimized, this bioinformatics tool will minimize the time and effort required to analyze MS data and determine novel bioactive peptides

from genetic sequence information. Until then, searching sequence databases and de novo sequencing are the preferred methods for determining peptide identity.

17. The *SwePep* (19) database (http://www.swepep.org) was developed to increase the throughput of identifying endogenous peptides in complex tissue samples analyzed by MS. This is a good place to start when trying to identify neuropeptides in tissue extracts. Online databases like *Mascot* from Matrix Science (http://www.matrixscience.com) and *SEQUEST* from Thermo Corp. (http://www.thermo.com) are also suitable for the identification of neuropeptides (20, 21).
18. De novo sequence matching of at least three consecutive amino acids of the presumed sequence was previously considered to be sufficient for a positive identification (11). It should be noted that it is often challenging to obtain complete fragmentation for de novo sequencing of neuropeptides. Various chemical derivatization methods could be employed to enhance fragmentation (22, 23).

Acknowledgments

This work was supported in part by National Science Foundation CAREER Award (CHE-0449991), National Institutes of Health through grant 1R01DK071801, and an Alfred P. Sloan Research Fellowship (L.L.). R.M.S. acknowledges the National Institutes of Health Clinical Neuroengineering Training Program Grant 5T90DK070079. J.A.D. acknowledges an American Foundation for Pharmaceutical Education (AFPE) predoctoral fellowship.

References

1. Strand, F.L. (2003) Neuropeptides: general characteristics and neuropharmaceutical potential in treating CNS disorders. *Prog. Drug Res.* **100**, 1–37.
2. Clynen, E., De Loof, A. and Schoofs, L. (2003) The use of peptidomics in endocrine research. *Gen. Comp. Endocrinol.* **132**, 1–9.
3. Strand, F.L. (1994) Models of neuropeptide action. New York Academy of Sciences, New York, N.Y.
4. Li, L. and Sweedler, J.V. (2008) Peptides in the brain: mass spectrometry-based measurement approaches and challenges. *Annu. Rev. Anal. Chem.* **1**, 423–449.
5. Theodorsson, E., Stenfors, C. and Mathe, A.A. (1990) Microwave irradiation increases recovery of neuropeptides from brain tissues. *Peptides* **11**, 1191–1197.
6. Mathe, A.A., Stenfors, C., Brodin, E. and Theodorsson, E. (1990) Neuropeptides in brain: effects of microwave irradiation and decapitation. *Life Sci.* **46**, 287–293.
7. Nylander, I., Stenfors, C., Tan No, K., Mathe, A.A. and Terenius, L. (1997) A comparison between microwave irradiation and decapitation: basal levels of dynorphin and enkephalin and the effect of chronic mor-

phine treatment on dynorphin peptides. *Neuropeptides* **31**, 357–365.
8. Svensson, M., Skold, K., Svenningsson, P. and Andren, P.E. (2003) Peptidomics-based discovery of novel neuropeptides. *J.Proteome Res.* **2**, 213–219.
9. Parkin, M.C., Wei, H., O'Callaghan, J.P. and Kennedy, R.T. (2005) Sample-dependent effects on the neuropeptidome detected in rat brain tissue preparations by capillary liquid chromatography with tandem mass spectrometry. *Anal. Chem.* **77**, 6331–6338.
10. Che, F.Y., Lim, J., Pan, H., Biswas, R. and Fricker, L.D. (2005) Quantitative neuropeptidomics of microwave-irradiated mouse brain and pituitary. *Mol. Cell. Proteomics* **4**, 1391–1405.
11. Dowell, J.A., Heyden, W.V. and Li, L. (2006) Rat neuropeptidomics by LC-MS/MS and MALDI-FTMS: enhanced dissection and extraction techniques coupled with 2D RP-RP HPLC. *J. Proteome Res.* **5**, 3368–3375.
12. Hatcher, N.G., Richmond, T.A., Rubakhin, S.S. and Sweedler, J.V. (2005) Monitoring activity-dependent peptide release from the CNS using single-bead solid-phase extraction and MALDI TOF MS detection. *Anal. Chem.* **77**, 1580–1587.
13. Che, F.Y., Yan, L., Li, H., Mzhavia, N., Devi, L.A. and Fricker, L.D. (2001) Identification of peptides from brain and pituitary of Cpe^{fat}/Cpe^{fat} mice. *Proc. Natl. Acad. Sci. USA* **98**, 9971–9976.
14. Caprioli, R.M., Farmer, T.B. and Gile, J. (1997) Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* **69**, 4751–4760.
15. Stoeckli, M., Staab, D., Staufenbiel, M., Wiederhold, K.H. and Signor, L. (2002) Molecular imaging of amyloid beta peptides in mouse brain sections using mass spectrometry. *Anal. Biochem.* **311**, 33–39.
16. DeKeyser, S.S., Kutz-Naber, K.K., Schmidt, J.J., Barret-Wilt, G.A. and Li, L. (2007) Imaging mass spectrometry of neuropeptides in decapod crustacean neuronal tissues. *J. Proteome Res.* **6**, 1782–1791.
17. Rubakhin, S.S., Hatcher, N.G., Monroe, E.B., Heien, M.L. and Sweedler, J.V. (2007) Mass spectrometric imaging of the nervous system. *Curr. Pharm. Des.* **13**, 3325–3334.
18. Amare, A., Hummon, A.B., Southey, B.R., Zimmerman, T.A., Rodriquez-Zas, S.L. and Sweedler, J.V. (2006) Bridging neuropeptidomics and genetics with bioinformatics: prediction of mammalian neuropeptide prohormone processing. *J. Proteome Res.* **5**, 1162–1167.
19. Falth, M., Skold, K., Norrman, M., Svensson, M., Fenyo, D. and Andren, P.E. (2006) SwePep, a database designed for endogenous peptides and mass spectrometry. *Mol. Cell. Proteomics* **5**, 998–1005.
20. Perkins, D.N., Pappin, D.J.C., Creasy, D.M. and Cottrell, J.S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567.
21. Ducret, A., Van Oostveen, I., Eng, J.K., Yates-III, J.R. and Aebersold, R. (1998) High throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry. *Protein Sci.* **7**, 706–719.
22. Fu, Q. and Li, L. (2005) De novo sequencing of neuropeptides using reductive isotopic methylation and investigation of ESI QTOF MS/MS fragmentation pattern of neuropeptides with N-terminal dimethylation. *Anal. Chem.* **77**, 7783–7795.
23. Ma, M., Kutz-Naber, K.K. and Li, L. (2007) Methyl esterification assisted MALDI FTMS characterization of the orcokinin neuropeptide family. *Anal. Chem.* **79**, 673–681.

Chapter 18

Quantitative Neuroproteomics of the Synapse

Dinah Lee Ramos-Ortolaza, Ittai Bushlin, Noura Abul-Husn, Suresh P. Annangudi, Jonathan Sweedler, and Lakshmi A. Devi

Abstract

An emerging way to study neuropsychiatric or neurodegenerative diseases is by performing proteomic analyses of brain tissues. Here, we describe methods used to isolate and identify the proteins associated with a sample of interest, such as the synapse, as well as to compare the levels of proteins in the sample under different conditions. These techniques, involving subcellular fractionation and modern quantitative proteomics using isotopic labels, can be used to understand the organization of neuronal compartments and the regulation of synaptic function under various conditions.

Key words: Neuroproteomics, subcellular fractionation, presynaptic terminal, mass spectrometry, quantitative proteomics, differential isotopic labelling.

1. Introduction

Neuroproteomics is the study of the proteome, or the collection of proteins encoded by the genes of an organism, in particular that of the central nervous system (CNS). With the development of new techniques and the improvement of those already available, it is now possible not only to identify proteins, but also to determine changes in their abundance under various conditions (1). This is particularly useful in understanding the physiological function of biological systems, as well as determining the functional implication of alterations in proteins in disturbed states, such as those induced by neurodegenerative disorders and/or drugs of abuse (2).

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_18,

In the CNS, synapses are essential for the communication between neurons. Upon stimulation, a presynaptic neuron releases neurotransmitters that bind to receptors in the postsynaptic neuron, which in turn induces a series of events in response to the stimulus. One of the most fascinating properties of the CNS is synaptic plasticity or the ability to reconfigure and/or modulate synapses to accommodate for the wide variety of stimuli that they receive at any given time. The inability to respond adequately may lead to the development of neurodegenerative or addictive disorders.

Neuroproteomic studies have started to identify the proteins present in different compartments of the synapse, including synaptosomes (3) and presynaptic and postsynaptic terminals (4), mainly through subcellular fractionation protocols. This type of approach facilitates the analysis by reducing the complexity of the system. Moreover, it enriches synaptic compartments with less abundant proteins, which are commonly masked by those with the highest abundance. One way to take advantage of this methodology is to select a brain region, isolate the synaptic fraction of interest before and after a treatment (such as exposure to morphine) and identify the proteins in the fraction as well as their relative changes upon treatment (5).

The approach described in this chapter is divided into two major sections. **Section 3.1** describes the details for isolating a brain region and separating it into synaptic fractions using subcellular fractionation. It is expected that (at least) two samples that undergo separate treatments (such as exposure to morphine and a control) are prepared. **Section 3.2** details the steps required to isolate proteins, digest and differentially label them with appropriate isotopic labels and perform mass spectrometry to identify the peptides and hence the proteins from the original sample. The data analysis (also described) yields a list of proteins and specifies the relative change in protein levels upon treatment. Whilst we have used specific examples from our research such as working with synaptic proteins after animal exposure to morphine, these protocols are easily adaptable for use with other animal models, other brain regions or a range of treatments. In addition, while specific fractionation and mass spectrometric equipment have been used, other instrumental platforms can be used with appropriate modifications.

2. Materials

2.1. Sample Preparation

2.1.1. Tissue Acquisition and Storage

1. Dissection tools: forceps, razor blades.
2. Dry ice.
3. Ethanol, 70%.
4. Filter paper (Fisher Scientific, Fairlawn, NJ).

5. Isopropanol (Sigma Aldrich, St. Louis, MO).
6. Phosphate-buffered saline (PBS).

2.1.2. Fractionation

1. Sucrose (0.32 M): 5.4 g sucrose in 50 mL $CaCl_2$. Store at 4°C.
2. Sucrose (1 M): 17 g sucrose in 50 mL $CaCl_2$. Store at 4°C.
3. Sucrose (2 M): 34 g sucrose in 50 mL dH_2O. Store at 4°C.
4. Tris–HCl (1 M), pH 6.
5. Tris–HCl (20 mM)/Triton X-100 (1%), pH 8.
6. Protease inhibitors cocktail (Sigma Aldrich, St. Louis, MO).
7. Phosphatase inhibitors cocktail (Sigma Aldrich, St. Louis, MO).
8. Cold acetone (Sigma Aldrich, St. Louis, MO).
9. SDS, 0.1%.
10. SDS, 1%.
11. Ethanol, 70%.
12. Wheaton glass homogenizer, 7 mL.
13. Amicon Ultra-15 centrifugal filters (Millipore, Bedford, MA).

2.1.3. Protein Estimation

1. BCA protein assay kit (Pierce, Rockford, IL).
2. Bovine serum albumin (BSA) (Sigma Aldrich, St. Louis, MO).
3. Microplate reader, 550 Model (Bio-Rad, Hercules, CA).

2.1.4. SDS-PAGE

1. Acrylamide/Bis solution, 40% (29:1 with 3.3% C) (Bio-Rad, Hercules, CA).
2. TEMED (Sigma Aldrich, St. Louis, MO).
3. Resolving Buffer (4X): 1.5 M Tris–HCl, pH 8.8; 0.4% (w/v) SDS.
4. Stacking buffer (4X): 0.5 M Tris–HCl, pH 6.8; 0.4% (w/v) SDS.
5. Ammonium persulfate, 10%.
6. Running buffer (1X): 25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS.
7. Transfer buffer (1X): 25 mM Tris, 192 mM glycine, 20% (v/v) methanol.
8. Pre-stained molecular weight markers (Dual Color Protein Standards, Bio-Rad, Hercules, CA).
9. Isopropanol (Sigma Aldrich, St. Louis, MO).

10. Loading buffer (6X): 60 mM Tris–HCl/SDS pH 6.8, 10% glycerol, 2% SDS, 20 mM DTT, 0.001% bromophenol blue.
11. Mini-Protean II electrophoresis cell gel system (Bio-Rad, Hercules, CA).

2.1.5. Western Blotting

1. Odyssey Blocking Buffer (Li-Cor Biosciences, Lincoln, NE).
2. Primary antibody solution in Odyssey Blocking Buffer with 0.1% Tween-20 and 0.01% NaN_3.
3. Fluorescently labelled secondary antibody solution in Odyssey Blocking Buffer with 0.1% Tween-20 and 0.01% NaN_3.
4. Tris-buffered saline (1X) with Tween-20 (TBS-T): 20 mM Tris–HCl, pH 7.6, 137 mM NaCl, 0.1% (v/v) Tween-20.
5. Nitrocellulose membrane.
6. Chromatography paper (Fisher Scientific, Fairlawn, NJ).

2.2. Differential Isotopic Labelling and Mass Spectrometry

2.2.1. Protein Digestion and Stable Isotope Labelling

1. Solution of 1 M NaOH in water.
2. Phosphate buffer (pH 8.5).
3. Solution of 200 mM dithiothreitol (DTT) in 100 mM NH_4HCO_3 (pH 8–9).
4. Solution of 1 M iodoacetamide (IAM) in 100 mM NH_4HCO_3.
5. Digestion buffer: 50 mM NH_4HCO_3, 2 M urea (pH 8–10).
6. Sequencing-grade modified trypsin (Promega Co., Madison, WI).
7. Solution of 2 M succinic anhydride in dimethyl sulfoxide (DMSO) (SA-H) (Sigma Aldrich, St. Louis, MO).
8. Solution of 2 M succinic [$^2H_4$] anhydride in DMSO (SA-D) (Sigma Aldrich, St. Louis, MO).
9. Solution of 2.5 M glycine.
10. Solution of 2 M hydroxylamine.

2.2.2. Sample Clean-Up

1. Protein desalting spin columns (Pierce, Rockford, IL).
2. Activating solution: 50% aqueous acetonitrile solution containing 0.1% formic acid (FA) and 0.01% trifluoroacetic acid (TFA) (*see* **Note 1**).
3. Equilibrating solution: 5% aqueous acetonitrile solution containing 0.1% formic acid and 0.01% trifluoroacetic acid.
4. Eluting solution: 70% aqueous acetonitrile solution containing 0.1% formic acid and 0.01% trifluoroacetic acid.

2.2.3. Liquid Chromatography/Mass Spectrometry

1. Micromass™ CapLC system (Micromass, Manchester, UK).
2. HCTUltra-PTM discovery system™ (Bruker Daltonics, Billerica, MA).
3. Manual injector (Valco Instruments Co., Inc., Houston, TX).
4. OPTI-PAK® Trap Column 0.5 μL C18 (Optimize Technologies, Inc., Oregon City, OR).
5. LC Packings™ 300 μm i.d. × 15 cm, C18 PepMap100, 100 Å (LC Packings, San Francisco, CA).
6. Solvent A: 5% acetonitrile in water containing 0.1% formic acid and 0.01% trifluoroacetic acid.
7. Solvent B: 95% acetonitrile in water containing 0.1% formic acid and 0.01% trifluoroacetic acid.
8. Software: Bruker Daltonics Hystar, version 3.2 – integrates the capLC system and the mass spectrometer.
9. Software: Bruker Daltonics EsquireControl, version 6.1 – controls the mass spectrometer.

2.2.4. Data Analysis

1. Software: Bruker Daltonics DataAnalysis, version 3.4 – identifies, deconvolves and quantifies peptides.
2. Software: Bruker Daltonics Biotools, version 3.1 – catalogues and matches MS and MS/MS spectra of peptides with Mascot results.
3. Software: Mascot (Matrix Science, London, UK) – assigns peptide sequence based on the MS/MS spectra and database used.

3. Methods

3.1. Sample Preparation

3.1.1. Tissue Acquisition and Storage

This protocol is based on the use of rodent brain tissue. Once the animal is sacrificed by decapitation, the brain is rapidly removed and the regions of interest, such as striatum and hippocampus, are extracted (6).

1. Prior to tissue acquisition, place a Petri dish on ice.
2. Fill small tubes with isopropanol and place on dry ice. The microtubes used for tissue collection must fit inside these tubes.
3. Remove the brain and place it on a Petri dish lined with a filter paper soaked in ice-cold PBS (*see* **Notes 2 and 3**).
4. Locate the areas of interest and remove them carefully using clean forceps and razor blades (*see* **Note 4**).

5. Place each dissected area, separately, in a clean microtube and freeze immediately by placing the microtube on dry ice in the tube containing isopropanol. If using the dissected tissue immediately, keep the tubes on dry ice until ready to use; otherwise, store at –80°C.

3.1.2. Fractionation

Once the brain region of interest is collected, subcellular fractionation is performed to further simplify the sample for proteomic analysis. The following protocol allows the separation of various synaptic compartments, including synaptosomes, presynaptic and postsynaptic fractions (5, 7).

1. One hour before starting the procedure, place the SW28 rotor and tube holders in the ultracentrifuge (Beckmann L7–65), set to 4°C, 28,000 RPM (141,000×g), 3 h, and turn ON power and vacuum. This allows the centrifuge and tube holders to reach the desired temperature before the samples are ready to be centrifuged.
2. Place centrifuge tubes (Beckman Polyallomer) on ice until ready to use.
3. Weigh total brain samples (at least 200 mg tissue) (*see* **Note 5**).
4. Using a glass homogenizer, homogenize the tissue ~25 times in 3 mL 0.32 M sucrose solution (*see* **Note 6**), 30 μL protease inhibitor cocktail (100X) and 30 μL phosphatase inhibitor cocktail (100X). Store a 200-μl aliquot of this homogenate and transfer the rest to a 50-mL tube (*see* **Note 7**).
5. Add 12 mL 2 M sucrose solution and 5 mL 0.1 mM $CaCl_2$ to the homogenate. This will form a solution with a concentration of 1.25 M sucrose. Mix well (do not vortex) and transfer to an ultracentrifuge tube.
6. Using a 10-mL plastic pipette, overlay slowly and carefully with 1 M sucrose solution until the tube is almost full. Repeat for every sample to be fractionated. This will create a sucrose gradient that allows the isolation of synaptosomes.
7. Place the centrifuge tubes in the tube holders and balance with a 1 M sucrose solution before placing them in the centrifuge
8. Centrifuge in the SW28 rotor at 28,000 RPM for 3 h at 4°C (*see* **Note 8**).
9. After centrifugation, discard myelin (viscous layer floating on top) by suctioning with a glass pipette (**Fig. 18.1**). Harvest the synaptosome band (~3–4 mL) located at the interface between the 1.25 and 1 M sucrose layers. Record the

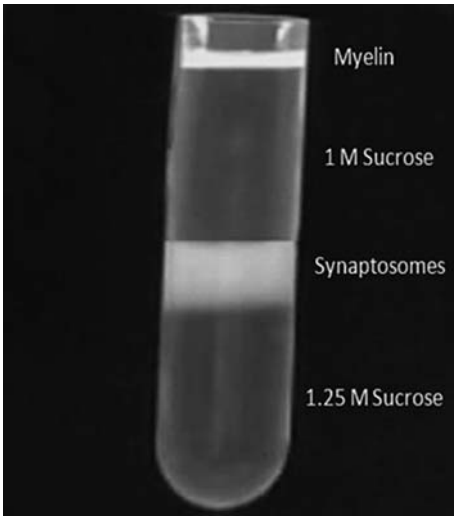


Fig. 18.1. Illustration of layers formed after sucrose gradient centrifugation to isolate synaptosomes.

amount harvested, take a 500-μL aliquot of the synaptosomes fraction and store at –80°C. Transfer the rest to 50-mL polycarbonate tubes (samples can be stored overnight at –20°C after this step) (*see* **Note 9**).

10. Add 0.1 mM $CaCl_2$ at 10X volume of synaptosomes, 1 M Tris–HCl buffer, pH 6 at 1/50th final volume, 10% Triton X-100 at 1/10th final volume and 100X protease and phosphatase inhibitors. Mix by inversion and incubate on shaker (horizontally in ice box or at 4°C at a moderate speed for 20 min; if using the VWR orbital shaker OS-500, the speed should be 2.5.
11. Weigh and balance tubes with $CaCl_2$. Centrifuge at 18,500 RPM (40,000×*g*) on the SS-34 rotor at 4°C for 20 min to pellet synaptic junctions. (This protocol assumes the use of the Sorvall RC 5C Plus centrifuge.)
12. Pour out the supernatant and resuspend the pellet in 2 mL 20 mM Tris–HCl/1% Triton X-100 buffer, pH 8, and 20 μl each of the protease and phosphatase inhibitor cocktails (100X). Homogenize on ice. In this step it is suggested to suck the pellet into the pipette with 1ml of buffer, transfer it to the homogenizer and then add 1 mL of buffer to the tube to wash any residue that might have been left. Then, transfer to the homogenizer and add the protease and phosphatase inhibitors. Homogenize and transfer back to the polycarbonate tube. Take a 200-μL aliquot and store at –80°C. This aliquot contains the synaptic junctions.

13. Add 18 mL of 20 mM Tris–HCl/1% Triton X-100 buffer, pH 8, 180 μL each protease and phosphatase inhibitor cocktails (100X) to the tube containing the synaptic junctions.
14. Mix by inversion and incubate on shaker (horizontally in ice box or at 4°C, vigorously) for 20 min. Weigh and balance with 20 mM Tris–HCl/1% Triton X-100 buffer, pH 8.
15. Centrifuge at 18,500 RPM (40,000×*g*) on the SS-34 rotor at 4°C for 20 min to pellet the postsynaptic density (PSD) fraction. Pour supernatant containing the presynaptic fraction into a separate 50-mL Falcon tube and store PSD pellet at –80°C.
16. While the samples are in the centrifuge, pre-rinse Amicon Ultra-15 centrifugal filters with 10 mL dH_2O. Centrifuge at 3,500 RPM for ~15 min. Discard the flow-through.
17. Add 10 mL of the supernatant containing the presynaptic fraction to the pre-rinsed centrifugal filter and centrifuge at 3,500 RPM, 4°C, for ~20–30 min.
18. Pipette up and down to homogenize the concentrated supernatant. Repeat previous step with the remaining 10-mL supernatant. This step allows concentration of the supernatant containing the presynaptic membrane fraction from 20 to 1 mL.
19. Pour concentrated supernatant into glass centrifuge tube. Add ice-cold acetone to 10X volume to precipitate proteins. Cover with paraffin and store overnight at –20°C (*see* **Note 10**).
20. To pellet the presynaptic membrane fraction, centrifuge the concentrated supernatant stored in acetone at 11,500 RPM (15,000×*g*) at 4°C for 30 min in the SS-34 rotor. If the tubes do not fit properly in the rotor, use rubber tube holders to avoid breaking the glass. Discard supernatant. Let presynaptic pellet air dry and store at –80°C.
21. If using the samples for immunoblotting, dissolve the PSD pellet in 1% SDS and the presynaptic pellet in 0.1% SDS (200 μL each).
22. Add 2 μL protease and phosphatase inhibitor cocktails (100X). Vortex and spin down in the microcentrifuge.
23. Sonicate the samples at 3 Watts by moving the tubes ten times up and down in the sonicator probe.
24. Store the samples at –80°C until ready to use (*see* **Note 11**).

3.1.3. Protein Estimation

This protocol uses the Pierce® BCA Protein Assay for the protein estimation, which uses bicinchoninic acid for colorimetric detection and quantitation of proteins. It allows the detection of protein amounts ranging from 0.02 to 20 μg.

To determine protein amounts of the samples, it is necessary to use a common protein as a reference. This protocol uses bovine serum albumin (BSA) as the reference standard (*see* **Table 18.1** for a guide in preparing the standards). The BSA stock used for the preparation of these curves has a concentration of 1 mg/mL.

1. Use the following formula to determine the amount of working reagent (WR) needed for the estimation:

$$\text{Volume of WR} = (\#\text{ standards} + \#\text{ unknowns}) \times (\#\text{ replicates}) \times (200\mu L)$$

2. Mix 50 parts of BCA Reagent A with 1 part of BCA Reagent B (*see* **Note 12**).

Table 18.1
Preparation of BSA Standard Curve

Tube	Volume of dH_2O	Volume of BSA (1 mg/ml)	Total protein amt in 20 μl
Standard curve (working range: 0.5–10 μg)			
1	80 μL	0	0
2	78 μL	2 μL	0.5 μg
3	76 μL	4 μL	1 μg
4	72 μL	8 μL	2 μg
5	64 μL	16 μL	4 μg
6	56 μL	24 μL	6 μg
7	48 μL	32 μL	8 μg
8	40 μL	40 μL	10 μg
Standard curve (working range: 0.005–0.5 μg)			
1	195 μL	5 μL Stock BSA	0.5 μg
2	100 μL	100 μL Tube 1 dilution	0.250 μg
3	100 μL	100 μL Tube 2 dilution	0.125 μg
4	120 μL	80 μL Tube 3 dilution	0.05 μg
5	100 μL	100 μL Tube 4 dilution	0.025 μg
6	100 μL	100 μL Tube 5 dilution	0.0125 μg
7	120 μL	80 μL Tube 6 dilution	0.005 μg
8	200 μL	0	0

3. This protocol uses a microplate, which allows using a smaller volume of the sample for the protein estimation (*see* **Note 13**).
4. Add 20 μL of each standard to the microplate wells.
5. Add 200 μL of the WR solution to each well and mix on a plate shaker for 30 s.
6. Cover the plate with paraffin and incubate at 37°C for 30 min (*see* **Notes 14 and 15**).
7. Cool plate to room temperature.
8. Measure the absorbance at approximately 562 nm on a plate reader.
9. Prepare a standard curve by plotting the average BSA standards absorbance versus its protein amount in micrograms and use this curve to estimate the concentration of your samples.

3.1.4. SDS-PAGE

These instructions assume the use of a Mini-Protean II electrophoresis cell gel system. It is critical that the glass plates for the gels are scrubbed clean with a detergent after use and rinsed extensively with distilled water.

1. For electrophoresis, prepare two 1.5-mm thick, 7.5% gels by mixing 3.75 mL of 4X resolving buffer (pH 8.8) with 2.8 mL of 40% acrylamide/bis solution, 8.45 mL water, 50 μL 10% ammonium persulfate solution and 7.5 μL TEMED. Pour the gel, leaving space for a stacking gel, and overlay with water-saturated isopropanol. The gel should polymerize in about 45 min.
2. After the gel is polymerized, pour off the isopropanol and rinse the top of the gel twice with water. Blot with paper to remove excess water before adding stacking gel.
3. Prepare the stacking gel by mixing 1.56 mL of 4X stacking buffer (pH 6.8) with 0.56 mL acrylamide/bis solution, 4.13 mL water, 33 μL ammonium persulfate solution and 7.5 μL TEMED. Use about 3 mL of this to pour the stack for each gel and insert the appropriate comb. The stacking gel should polymerize within 30 min.
4. Prepare 1X running buffer by adding 6 g Tris–HCl and 28.8 g glycine to 2L H_2O. Add 20 mL 10% SDS solution and adjust pH to 8–8.9.
5. Once the stacking gel has set, carefully remove the comb and use a syringe to wash the wells with water and then running buffer.
6. Prepare samples in 1% SDS + 6X loading buffer (DTT added). Spin down samples and boil at 100°C for 5 min. Keep samples at 4°C until ready to use.

7. Add the running buffer to the upper and lower chambers of the gel unit and load 36 μL of each sample in a well. Include one well for pre-stained molecular weight markers.
8. Complete the assembly of the gel unit and connect to a power supply. The gel can be run at 100 V for about 1.45 h or until dye front reaches the bottom of the gel.
9. A few minutes before the SDS-PAGE ends, soak filter papers, fiber pads and the nitrocellulose membrane needed for the transfer in cold 1X transfer buffer.
10. When the SDS-PAGE ends, remove the gel from the apparatus and discard the portion corresponding to the stacking gel.
11. Prepare the transfer assembly as follows: On the dark side of the cassette place a fiber pad, two sheets of 3 MM filter paper, the resolving gel, the nitrocellulose membrane, two additional sheets of 3 MM filter paper and a fiber pad. After placing the nitrocellulose membrane on top of the gel, make sure that there are no bubbles between the gel and the membrane. If there are bubbles, remove them by rolling a glass tube on top of the assembly (*see* **Note 16**).
13. Close and lock the cassette carefully to avoid moving the gel and/or the membrane out of place.
14. Place the cassette on its module and put in inside the transfer tank together with the cooling block.
15. Fill the tank with transfer buffer. Since the transfer generates a lot of heat, we suggest using cold transfer buffer (4°C) to help dissipate the heat.
16. Close the lid firmly and connect to a power supply.
17. Transfer at 30 V overnight at room temperature or at 100 V for 1.75–2 h on ice.
18. At the end of the run, remove the cassette from the transfer apparatus and take out the assembly so that the gel is on top and the membrane is behind the gel; cut the membrane to the size and shape of the gel. At this point, the molecular weight marker should be visible on the membrane. Since the colours of the molecular weight markers can fade over time, make a cut in one corner of the membrane to remember the orientation of the samples.

3.1.5. Western Blotting

To validate the fractionation protocol, Western blotting analysis can be performed using the various aliquots obtained during the procedure. Antibodies for presynaptically and postsynaptically enriched proteins, such as Syntaxin 1 and PSD95 respectively, can be used to confirm the separation of the subcellular fractions (**Fig. 18.2**). Since the quality of the sample is critical for accuracy

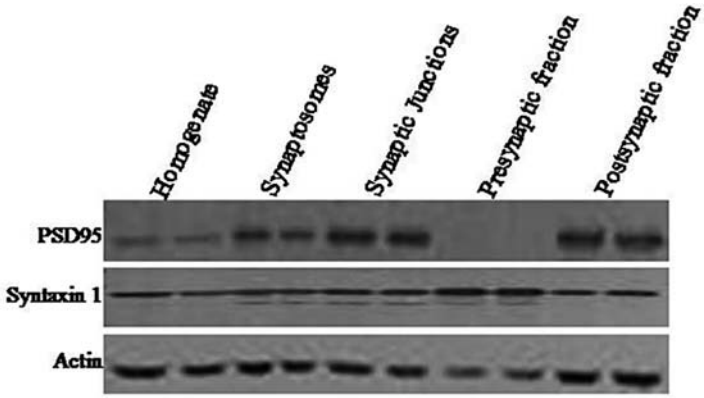


Fig. 18.2. Biochemical validation of the fractionation protocol. To demonstrate the enrichment of proteins in each fraction, equal amounts of protein from each fraction were separated by SDS-PAGE and probed with antibodies for presynaptically (Syntaxin 1) and postsynaptically (PSD95) enriched proteins.

of the data, it is recommended to perform this validation prior to proteomic analysis. The following protocol is based on the use of the Odyssey Infrared Imaging System; therefore, we suggest blocking the membrane in Odyssey Blocking Buffer.

1. Block the membrane in Odyssey Blocking Buffer for 1 h at room temperature on a rocking platform.
2. Add the primary antibody to the membrane and incubate either for 1 h at room temperature or overnight at 4°C with gentle shaking. Add enough antibody solution to cover the whole membrane. Dilution of the antibody solution and incubation time may vary for different antibodies; therefore, it is necessary to optimize these conditions for each particular case (*see* **Note 17**).
3. After incubation with the primary antibody, remove the antibody and wash the membrane three times, 5–10 min each, with TBS + 0.1% Tween-20.
4. Incubate the membrane with the secondary antibody solution for 1 h at room temperature on a rocking platform. Increasing the incubation time can lead to an increase in background. From this point on, the membrane should be protected from light (*see* **Note 18**).
5. Remove the secondary antibody solution and wash the membrane three times, 10–15 min each, with TBS + 0.1% Tween-20.
6. Rinse the membrane with TBS to remove residual Tween-20.
7. Scan the membrane using the Li-Cor imaging system.

3.2. Differential Isotopic Labelling and Mass Spectrometry

3.2.1. Protein Digestion and Stable Isotope Labelling

1. Dissolve 30 μg of extracted proteins in 50 mM $NaHCO_3$ solution to a final volume of 100 μL.
2. Add 5 μL of 200 mM DTT solution to the protein extracts to reduce the disulphide bonds and incubate the solution in a water bath at 40°C for 1 h.
3. To alkylate the reduced proteins, add 5 μL of the 1 M IAM solution and incubate in the dark (cover the vial using aluminium foil) for 40 min.
4. Quench the unreacted IAM using 3 μL of DTT solution and incubate for 1 h at room temperature.
5. Add 500 μL (5 X by volume) of ice-cold acetone to precipitate the alkylated extracts and store at 4°C overnight.
6. Centrifuge the precipitate at 15,000 RPM for 40 min at 4°C and remove the supernatant.
7. Dissolve the pellet in 20 μL of digestion buffer.
8. Add 0.3 μg of trypsin to the dissolved proteins and incubate for 3 h at 40°C.
9. Following digestion, adjust the pH of the solution to 8–9 using 5–10 μL of 1 M NaOH solution.
10. React the contents of the tubes using the isotopic labels (SA-H or SA-D, *see* **Note 19**). For duplicate measurements, ensure that the duplicates are labelled in forward and reverse fashion (Sample A with SA-H/Sample B with SA-D, and then Sample A with SA-D/Sample B with SA-H).
11. Add 5 μL of SA-H/SA-D solution based on the label and incubate at room temperature for 10 min with intermittent vortex and centrifuge.
12. Using 0.1 μL of the solution and pH paper, check the pH. Adjust the pH to ~9 using 1 μL of the 1 M NaOH solution (*see* **Note 20**).
13. Repeat Steps 11 and 12 three times for each sample.
14. Add 5 μL of the 2.5 M glycine solution to the mixture and incubate for 20 min at room temperature or at 4°C overnight.
15. Add 1 μL of the 1 M NaOH solution followed by 2 μL of 2 M hydroxylamine solution and incubate at room temperature for 15 min.
16. Repeat Step 15 again for each sample.
17. Combine the light- and heavy-labelled samples that are to be compared.
18. Vortex the samples and centrifuge using a microcentrifuge.

19. Store the samples at –20°C or proceed to sample clean-up. **Figure 18.3** illustrates the general steps used to prepare the sample, including the protein digestion and isotope labelling, prior to sample clean-up and analysis.

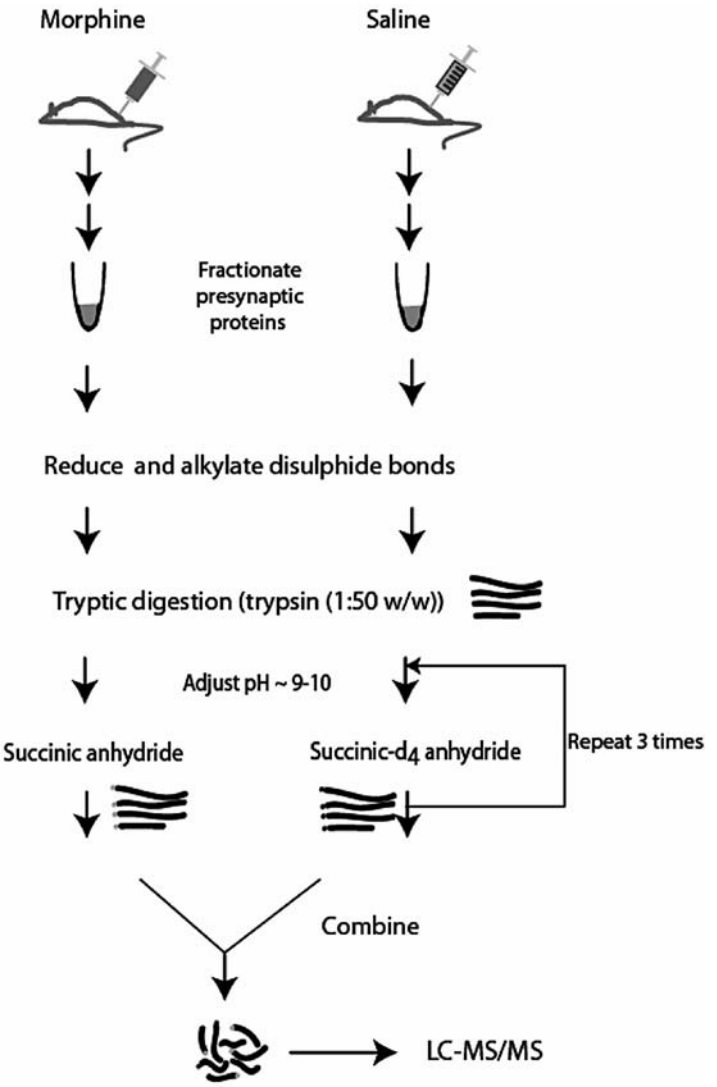


Fig. 18.3. Flow chart illustrating sample processing steps for protein digestion and stable isotopic labelling prior to LC–MS analysis.

3.2.2. Sample Clean-Up

1. Place the spin columns in the centrifuge with 2-mL centrifuge tubes and add 200 μL of activating solution. Centrifuge for 3 min at 1,500 RPM.
2. Repeat Step 1.
3. Add 200 μL of the equilibrating solution to the spin column and centrifuge for 3 min at 1,500 RPM.

4. Repeat Step 3.
5. Discard the solution in the 2-mL centrifuge tube and replace a fresh tube.
6. Add labelled mixture to the spin column and centrifuge for 3 min at 1500 RPM.
7. Reload the eluate to the spin column and centrifuge for 3 min at 1500 RPM.
8. Save the eluate and replace a fresh tube. Label the eppendorf tubes each time a fresh tube is placed to avoid confusion.
9. Add 200 μL of equilibrating solution to wash the salts and centrifuge at 1500 RPM for 3 min.
10. Repeat wash Step 9.
11. Replace with a fresh eluate-collecting tube.
12. Add 30 μL of eluting solution and centrifuge at 1500 RPM for 3 min.
13. Repeat Step 12 twice.
14. Remove the organic solvents in the eluate in a speed vacuum system for 20 min.
15. Reconstitute the residue in 15 μL of equilibrating solution; this solution is directly used for further mass spectrometric analysis.

3.2.3. Liquid Chromatography/Mass Spectrometry (see **Note 21***)*

1. Setup for chromatography using a solvent gradient of solvent A and solvent B; the 70 min gradient run for LC separation includes three steps: 5–80% solvent B in 15–55 min (linear); 80% solvent B for 55–60 min (isocratic); 80–5% solvent B in 60–65 min (linear).
2. Inject samples using a manual injector (Valco Instruments), load onto a trap column (PepMap™, C18, 5 μm, 100 Å, LC Packings) using solvent A and wash for 5 min.
3. Elute the trapped peptides in reverse direction onto a reverse-phased capillary column (LC Packings™ 300 μm i.d. × 15 cm, C18 PepMap100, 100 Å) using a solvent gradient at 2 μL/min flow rate.
4. Use an electrospray system for the chromatographic eluate, with nitrogen as the nebulizing gas, at 15 psi and dry the solvents using heated nitrogen gas (dry temperature 190°C) at 8 L/min.
5. Set the mass spectrometer parameters to a target mass of *m/z* 600, the ion charge control value 200,000 and scan mode set as "standard enhanced" with a speed of 8100 *m/z*/s.

6. Perform MS data acquisition and the subsequent tandem MS (CID) of selected peaks in a data-dependent manner using Esquire Control (Bruker). For each MS scan, select three peptides to be fragmented for 300–500 ms, based on their charge (preferably +2) and intensity.
7. Set the dynamic exclusion of previously fragmented precursor ions to two spectra for a period of 60 s. Perform MS and MS/MS scans in the range of *m/z* 300–1500 and 50–2000, respectively. An example set of labelled mass spectra (and associated MS/MS spectra) are shown in **Fig. 18.4**.

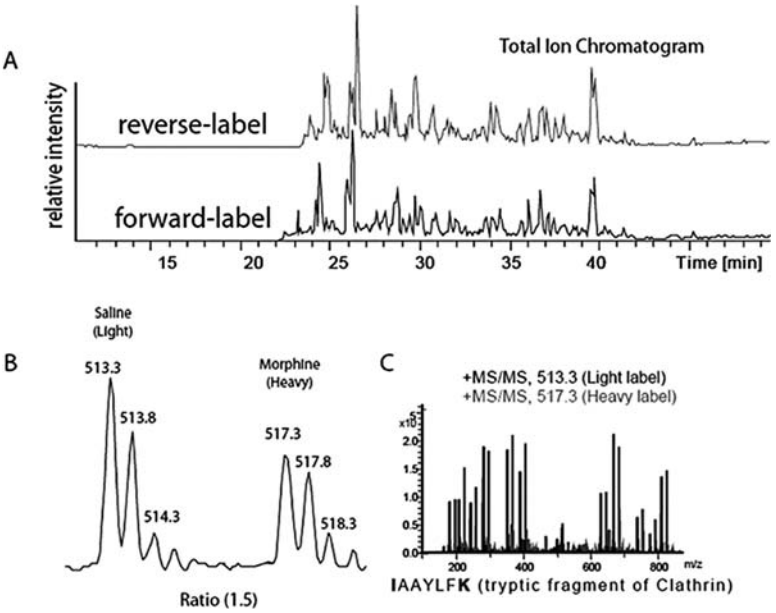


Fig. 18.4. Quantitative proteomic analysis. **A.** Representative total ion**** chromatogram from two samples labelled in forward and reverse fashion. **B**. Combined mass spectra in the range *m/z* 513–519 showing the difference in intensities of the tryptic fragment "IAAYLFK." **C.** Tandem mass spectra of light (*m/z* 513.3) and heavy (*m/z* 517.3) labelled tryptic fragment "IAAYLFK." The fragment shown here contains two sites for the labels: N-terminal and ϵ-amine of lysine residue.

3.2.4. Data Analysis

1. Process the data using the Data analysis software (Bruker). Search the mass spectral data obtained between 20 and 45 min of the LC run for compounds using an automated search option (parameters – intensity threshold of 10^5; retention time window – 0.7 min; fragments qualified by "amino acids").
2. Deconvolve the short-listed compounds with their respective MS/MS scans using the automated feature.
3. Export the deconvoluted spectra to the Biotools software (Bruker) for database searching using an in-house Mascot database search engine. Typical values include a mass tolerance of 0.1% for the MS and 0.5 Da for the MS/MS.

Include in the search parameters a fixed modification for cysteine (carbamidomethyl) and variable modifications for methionine (Met-oxidized), lysine and N-terminal amines (succinic anhydride, succinic [$^2H_4$] anhydride).

4. Consider peptides identified with a Mascot score of ≥50 for each protein for further analysis. Manually inspect MS/MS scans for identified peptides to match the fragment ions.
5. Create the list of positively identified proteins, their corresponding tryptic peptide masses and their retention time (*see* **Note 22**).
6. As each of the identified peptides elute across a span of several MS scans, manually combine appropriate scans and calculate the associated ratio of peak intensities between the heavy and light versions. During data acquisition, in several cases, only one of the labelled peptides may have been selected for fragmentation. In such cases, peptide peak pairs can be manually identified and quantified; heavy and light isotope-labelled peak pairs are separated by 4 Da for singly charged ions, 2 Da for doubly charged ions and 1.3 Da for triply charged ions (*see* **Note 23**).
7. Obtain the average ratios of all the tryptic peptides corresponding to each specific protein.
8. Create the final list of proteins with the average peak intensity ratios representing the two treatments.

4. Notes

1. TFA is known to improve resolution of chromatographic separations through ion-pairing; however, when used at >0.1%, TFA suppresses analyte ionization in a MS run. Formic acid is known to increase analyte ionization. Commercial solvents (Fischer Scientific, Pittsburg, PA) used in this study for sample clean-up and MS analysis have 0.1% FA and 0.01% TFA that is optimum for both ion-pairing and ionization.
2. Change the filter paper each time a different brain sample is dissected.
3. Once an animal is sacrificed, brain dissection should be performed on ice as quickly as possible to avoid protein degradation.
4. Clean dissection tools with 70% ethanol during dissection to avoid contamination of the samples.
5. Everything must be kept on ice at all times, unless otherwise specified.

6. We recommend using fresh solutions and buffers during the fractionation experiment. They may be prepared the day before the experiment starts and stored at 4°C.
7. If larger quantities of tissue are used for the fractionation experiment, amounts of buffers and solutions should be increased in proportion to the amount of tissue.
8. We recommend increasing the centrifugation time during the generation of the sucrose gradient to 4 h or more when using higher amounts of tissue. This allows for a better separation of the layers.
9. When treating the synaptosomes with 1 mM Tris–HCl buffer, pH 6, in order to pellet the synaptic junctions it may be necessary to split the volume of the synaptosomes into more than one centrifuge tube due to the limitation of the total volume that can be added to the tubes. This experiment assumes the use of the Sorvall RC 5C Plus centrifuge and the SS-34 rotor, which holds tubes with a maximum capacity of ~30 mL.
10. Make sure to store the sample with acetone in an appropriate freezer. After covering the tube, make a small hole in the paraffin covering the tube to avoid accumulation of vapours from the acetone.
11. Avoid freezing and thawing the samples too many times, as this can compromise the integrity of the proteins in the samples.
12. It is important to prepare enough WR in order to add 200 μL to each reaction. The equation in this section allows you to determine the amount of solution needed for all the standards and unknowns. To account for pipetting errors, add one extra reaction to the calculations.
13. The sample-to-WR ratio is approximately 1:8 (v/v), with a range of sample volume between 10 and 25 μL. Depending on the abundance of the proteins in the unknown samples, it might be necessary to make dilutions of the samples prior to adding the specific volume to the microplate.
14. Increasing the incubation time and/or temperature can lower the detection level.
15. Colour development continues even after cooling to room temperature, although at a slower rate; therefore, reading of the plate should be performed as quickly as possible.
16. The transfer protocol assumes the use of the Mini Trans-Blot cell transfer system.
17. When using the Odyssey Infrared Imaging System, primary antibodies can be diluted in Odyssey Blocking Buffer with 0.1% Tween-20 and 0.01% NaN_3. Optimum dilution

depends on the antibody; therefore, we suggest determining it individually for each antibody of interest.

18. Two different fluorescently labelled secondary antibodies can be used with the Odyssey System, the Cy5.5 and the IRDye800^{TM}. These secondary antibodies should be diluted in Odyssey Blocking Buffer with 0.1% Tween-20 and 0.01% NaN_3, at a range of 1:2,000–1:10,000, although lower concentrations can be used to detect very small amounts of protein. The diluted secondary antibody can be stored at 4°C and reused.
19. Assuming there are two treatment groups (e.g. morphine or saline treatment), the samples are paired so that one treatment group receives the heavy or light isotopic label (SA-D or SA-H) (8–10). Divide the samples so that each aliquot contains 10–30 μg of protein. Greater or lesser amounts of protein can be used with the appropriate scaling of reagents (11).
20. When adjusting the pH of the sample solutions using 1 M NaOH, use the same amount of solution from samples to be compared, i.e. maintain a constant volume between samples.
21. The peptide separation and mass spectrometric analysis described here were performed using a capLC^{TM} system coupled to an HCTUltra–PTM Discovery system ion-trap mass spectrometer equipped with an electrospray ionization source and the Hystar program was used to define the HPLC and mass spectrometry methods that were used to develop the solvent gradient and MS tune method and integrate the instruments used (12). Other MS and LC systems can be used with appropriate modifications to these protocols (13–15).
22. Protein identification should be based on two or more tryptic peptides.
23. Discard the ratio of peak pairs with overlapping peaks peptides originating from other "unknown peaks."

Acknowledgments

This work is partially supported by the National Institute on Drug Abuse through Awards No. DA018310 and DA017940 to JVS.

References

1. Kim, S.I., Voshol, H., Oostrum, J.V., Hastings, T.R., Cascio, M., and Glucksman, M.J. (2004) Neuroproteomics: An expression profiling of the brain's proteomes in health and disease. *Neurochem. Res.* **29**, 1317–1331.

2. Abul-Husn, N.S. and Devi, L.A. (2006) Neuroproteomics of the synapse and drug addiction. *J. Pharmacol. Exp. Ther.* **318**, 461–468.
3. Schrimpf, S.P., Meskenaite, V., Brunner, E., Rutishauser, D., Walther, P., Eng, J., Aebersold, R., and Sonderegger, P. (2005) Proteomic analysis of synaptosomes using isotope-coded affinity tags and mass spectrometry. *Proteomics* **5**, 2531–2541.
4. Phillips, G.R., Florens, L., Tanaka, H., Khaing, Z.Z., Yates, J.R., 3rd, and Colman, D.R. (2005) Proteomic comparison of two fractions derived from the transsynaptic scaffold. *J. Neurosci.* **81**, 762–775.
5. Morón, J.A., Abul-Husn, N.S., Rozenfeld, R., Dolios, G., Wang, R., and Devi, L.A. (2007) Morphine administration alters the profile of hippocampal postsynaptic density-associated proteins: a proteomics study focusing on endocytic proteins. *Mol. Cell Proteomics* **6**, 29–42.
6. Paxinos, G. and Watson, C. (1986) The rat brain in stereotaxic coordinates. *Academic Press*, San Diego.
7. Phillips, G.R., Huang, J.K., Wang, Y., Tanaka, H., Shapiro, L., Zhang, W., Shan, W.S., Arndt, K., Frank, M., Gordon, R.E., Gawinowicz, M.A., Zhao, Y., and Colman, D.R. (2001) The presynaptic particle web: ultrastructure, composition, dissolution, and reconstitution. *Neuron* **32**, 63–77.
8. Zhang, R., Sioma, C.S., Wang, S., and Regnier, F.E. (2001) Fractionation of isotopically labeled peptides in quantitative proteomics. *Anal. Chem.* **73**, 5142–5149.
9. Ong, S.E. and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* **1**, 252–262.
10. Che, F.Y., Lim, J., Pan, H., Biswas, R., and Fricker, L.D. (2005) Quantitative neuropeptidomics of microwave-irradiated mouse brain and pituitary. *Mol. Cell. Proteomics* **4**, 1391–1405.
11. Gutstein, H.B., Morris, J.S., Annangudi, S.P., and Sweedler, J.V. (2008) Microproteomics: analysis of protein diversity in small samples. *Mass Spectrom. Rev.* **27**, 316–330.
12. Panchaud, A., Hansson, J., Affolter, M., Bel Rhlid, R., Piu, S., Moreillon, P., and Kussmann, M. (2008) ANIBAL, stable isotope-based quantitative proteomics by aniline and benzoic acid labeling of amino and carboxylic groups. *Mol. Cell. Proteomics* **7**, 800–812.
13. Ranish, J.A., Yi, E.C., Leslie, D.M., Purvine, S.O., Goodlett, D.R., Eng, J., and Aebersold, R. (2003) The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* **33**, 349.
14. Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**, 198–207.
15. Romanova, E.V., Annangudi, S.P., and Sweedler, J.V. (2008) Mass Spectrometry of Proteins, In *The New Encyclopedia of Neuroscience*. (Squire, L., ed.), Elsevier Science, Amsterdam, ((in press)).

Chapter 19

Peptidomics Analysis of Lymphoblastoid Cell Lines

Anne Fogli and Philippe Bulet

Abstract

A key challenge in clinics is the identification of sensitive and specific biomarkers for early detection, prognostic evaluation, and surveillance of disease. A biomarker is defined as a biological substance that can be used to specifically detect a disease, measure its progression, or the effect of a treatment. A biomarker should be easily accessible, and ideally sensitivity and specificity must be sufficient to distinguish between false positives, false negatives, and true positives. To be useful for routine clinical evaluation, a biomarker should be detectable in body fluids (e.g., plasma, serum, urine). A biomarker can be a metabolite, a specific post-translational modification, a lipid, a phospholipid, or a protein. Due to technical advances in the analysis of biomolecules by mass spectrometry (MS), investigations of peptide biomarkers have increased. In contrast to genome, the peptidome is dynamic and constantly changing. Elucidating how the peptides complement changes in a cell type in diseases is crucial to understand how these processes occur at a molecular level. Lymphoblastoid cell lines, derived from blood lymphocytes, represent suitable models for biochemical investigations and biomedical applications because of their stability, the ease of amplification, and long-term preservation. Technological improvements of MS and liquid chromatography (LC) during the last 10 years resulted in the development of highly sensitive approaches for proteomic and peptidomic analyses. Here we provide guidelines for the preparation of the lymphoblastoid cell lines, the extraction of the peptides and their purification. We describe a number of technologies which we developed for the peptidomic profiling of lymphoblastoid cell extracts from patients with leukodystrophies, linked to mutations in the genes encoding the eukaryotic initiation factor 2B (eIF2B; eIF2B-related disorders).

Key words: Peptidomics, lymphoblasts, lymphoblastoid cell lines, biomarker, mass spectrometry, molecular mass fingerprints, differential expression.

1. Introduction

The study of peptide expression levels is essential for the analysis of biological processes in normal and pathological conditions. Unlike the genome, protein and peptide expression levels vary; these dynamic patterns contain a wealth of information about the disease processes and may be used to improve our understanding

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_19,

of diseases, improve their diagnostics, and may provide key to new treatments. The key stages of classical proteomic or peptidomic research include the multidimensional separation of complex mixtures of polypeptides and proteins by two-dimensional gel electrophoresis (2-DE) (1) or by multidimensional liquid chromatography (LC) (2) and their identification by mass spectrometry (3). A number of reports covering *proteome* analysis of lymphoblastoid cell extracts have been published to date (4–6), but to our knowledge only one *peptidomics* study of these cells has been published (7).

Lymphoblastoid cell lines, derived from blood lymphocytes, represent suitable models for biochemical investigations and biomedical applications because of their stability, the ease of amplification, and long-term preservation (8). Moreover, they are easy to amplify and to extract and are representative of the genetic diversity. Nevertheless, this cell type must be regarded as a raw biological material of extreme complexity, particularly rich in proteins, polypeptides, and peptides that are all present at different concentrations.

In this chapter, we describe how to perform differential peptidomic analysis of lymphoblastoid cell line extracts. This chapter contains detailed recipes of how to (i) obtain and culture Epstein–Barr Virus (EBV)-immortalized human lymphoblasts, (ii) prepare an extract of cultured lymphoblasts deprived of large proteins, (iii) enrich a fraction in peptides by solid-phase extraction, (iii) fractionate the peptidome using reversed-phase HPLC (RP-HPLC), and finally to (iv) analyze each RP-HPLC fraction using off-line MALDI-TOF mass spectrometry. We developed these procedures for the differential analysis of lymphoblastoid extracts from patients affected with CACH/VWM leukodystrophy (Childhood ataxia with central hypomyelination/vanishing white matter disorder) versus healthy patients (7). The methods detailed here allow performing differential peptidomic analyses from lymphoblasts extracts based on mass fingerprints/peptide profiling. Detailed descriptions of protocols for peptide sequencing with tandem MS/MS and/or Edman degradation can be found in other chapters of this volume and elsewhere (9).

2. Materials

2.1. Blood Collection, Lymphocyte Separation, Cell Immortalization and Culture

1. Microtainer(R) tubes containing adenine-citrate-dextrose or lithium heparin (BD, Franklin Lakes, NJ).
2. Culture medium: RPMI medium (Invitrogen Gibco) supplemented with 10% heat-treated fetal calf serum (Invitrogen Gibco) (*see* **Note 1**).

3. Antibiotics: 100 UI/mL streptomycin B, 2.5 μg/mL amphotericin B, 0.1 UI/mL penicillin, and 4 mM glutamine (Invitrogen Gibco).
4. Leucosep tube (Greiner).
5. Lymphoprep solution (Abcys).
6. Phosphate buffered saline (PBS), pH 7.4 (Gibco).
7. 50-mL Falcon tubes.
8. Cyclosporin A, solution at 10 μg/mL: $^{1}/_{100}$ dilution in RPMI medium of the stock solution (at 1 mg/mL in ethanol from the commercial solution at 1 mg/mL (CSA NOVARTIS 50 mg/mL, NORMAPUR), kept at –20°C).
9. Epstein–Barr virus: Aliquots of 1 mL from supernatants of marmoset monkey B95–8 cell lines infected with EBV.
10. Culture flasks of 25 and 75 cm^2 (Falcon).
11. Incubator 5% CO_2, at 37°C, in humidified atmosphere (Thermo Forma).
12. Centrifuge with rotors suitable for Falcon tubes (15–50 mL).

2.2. Lymphoblastoid Cell Lines Lysis

1. Malassez cell.
2. Phosphate buffered saline (PBS), pH 7.4.
3. Ultrapure water (MilliQ™ or of HPLC quality).
4. 2 M Glacial acetic acid high-purity grade prepared in ultrapure water (*see* **Note 2**).
5. Ultrasonicator for cell lysis (Branson Sonifier cell disruptor B15) (*see* **Note 3**).
6. Ice bucket.
7. Vortex.
8. Automatic peptide and protein analyzer such as Hitachi clinical analyzer 7180 to ensure reproducibility in protein quantification. The technique used is the pyrogallol red technique, adapted for peptide and protein concentrations between 50 and 150 μg/mL. Alternative manual techniques such as the Bradford technique can be used (*see* **Note 4**). Human serum albumin can be used as standard.

2.3. Solid-Phase Extraction

1. Trifluoroacetic acid (TFA), acetonitrile, and methanol are HPLC grade. Water is ultra pure (MilliQ™ or of HPLC quality).
2. Solutions of 0.1% TFA in ultra pure water and of 60% acetonitrile in ultra pure water.
3. Solid-phase extraction cartridges (reversed-phase Sep-Pak C_{18} cartridges, Waters™ or equivalents). The phase quan-

tity should be adapted to the amount of extract (0.3–12 g). The extraction may be performed manually with a syringe or using a multi-position vacuum manifold.

4. Low protein absorption polyethylene tubes (NUNC Immuno tubes, 75 × 12 mm, Roskilde, Denmark).
5. Centrifuge vacuum drier or vacuum lyophilizator (*see* **Note 5**).

2.4. RP-HPLC Profiling of Lymphoblasts Peptides Extracts

The peptide fractionation can be performed either by reversed-phase HPLC (RP-HPLC) coupled directly online to MS via electrospray ionization (ESI-MS) or by RP-HPLC coupled off-line with MS (ESI-MS and/or matrix-assisted laser desorption/ionization time of flight, MALDI-TOF).

1. HPLC system: gradient controller, pump optimized for low flow rates (0.8 mL/min), and a photodiode array detector or a wavelength detector (preferred wavelength 214 or 225 nm). An oven might be used for temperature control in the column and of the solvents delivered.
2. A recorder to plot or record the optical density of the collected fractions.
3. A fraction collector (optional item).
4. Analytical reversed-phase RP-HPLC column (C_{18}). Phase porosity of 300 Å and granulometry of 5 μm are preferred. A guard column may be connected upstream and in series with the main analytical for additional protection.
5. Solvent A: 0.1% TFA acidified ultrapure water (MilliQ^{TM} or HPLC quality) and Solvent B: 90% acetonitrile in 0.1% TFA acidified ultrapure water.
6. Low protein absorption polyethylene tubes (NUNC Immuno tubes, 75 × 12 mm).
7. Centrifuge vacuum drier or vacuum lyophilizator.

2.5. Mass Spectrometry: MALDI-TOF-MS

The molecular mass profiling can be performed either by liquid chromatography coupled to electrospray ionization (ESI) or by matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry (MS). MALDI-TOF MS is preferred for acquiring direct molecular mass fingerprints of complex samples without considering any chromatographic property. The advantage of MALDI ionization compared to the ESI is that mainly single-charged ions are created. Therefore, mass spectra from complex biological mixtures are easier to interpret. With MALDI any single compound would normally yield one signal with the appropriate isotopic pattern on the spectrum, whilst ESI often creates series of molecular ions having multiple charges. Compared to ESI, MALDI is more tolerable to buffers and salts widely used in the preparation of biological samples.

1. UltraFlex MALDI TOF-TOF mass spectrometer (Bruker Daltonics). MALDI-TOF-MS equipment that may operate in either linear or reflector mode and the detection be performed either in positive or negative mode.
2. Appropriate standard peptides/polypeptides for external calibration within the mass range 500–5000 Da in reflector mode and 2000 to 20–30 kDa for the linear mode. Horse heart myoglobin (Sigma-Aldrich) with signals at *m/z* 16,952.5 ($(MH)^+$) and *m/z* 8476.7 ($(M+2H)^{2+}$), and bovine pancreas insulin (Sigma-Aldrich) with signals at *m/z* 5734.57 (M^+) and *m/z* 2,867.78 (M^{2+}) are suitable.
3. A matrix adapted for <30 kDa peptides such as alpha-cyano-4-hydroxycinnamic acid (4-HCCA) (*see* **Note 6**), and the adapted solvents for preparing the different solutions for sample preparation: acetone, TFA, acetonitrile, and water (all HPLC grade).

3. Methods

This chapter has been organized in a way that the reader can start from blood procurement and end with the molecular mass tables for differential analyses. We have chosen to describe the procedures in the order of normal execution, from the lymphocytes culture from blood, immortalization into lymphoblasts, to the molecular mass profiling by MALDI-TOF MS.

3.1. Lymphocytes Culture and Immortalization

3.1.1. Lymphocytes Isolation from Blood

1. Collect 10 mL blood using Microtainer(R) tubes containing adenine-citrate-dextrose or lithium heparin (the anticoagulants that preserve leucocytes integrity).
2. Add 15 mL of Lymphoprep solution in a Leucosep tube, and centrifuge at 500×*g* for 1 min (*see* **Note 7**).
3. Deposit the blood sample in the Leucosep tube with an equal volume of PBS and agitate gently at room temperature.
4. Centrifuge for 20 min at 500×*g*, at room temperature.
5. Put the supernatant, containing the red blood cells, in a new 50-mL Falcon tube (*see* **Note 8**).
6. Centrifuge this tube for 5 min at 500×*g* and at room temperature.
7. Rinse the pellet once with 10 mL of cold-RPMI (4°C-RPMI).
8. Centrifuge for 5 min at 500×*g* at room temperature.

9. Resuspend the pellet in 4 mL of fresh RPMI culture medium and transfer to a 25-cm^2 flask.

3.1.2. EBV-Transformation of Lymphocytes in Lymphoblastoid Cell Lines

1. Add 100 μL of CSA solution (10 μg/mL) to 1 mL of the heated (37°C) aliquot of EBV and then add this solution to the flask containing cells.
2. Place the flask containing cells and the EBV in a vertical position in the incubator (37°C, 5% CO_2).
3. Gently agitate the cells every two days, without changing the culture medium.
4. After a week, change the old culture medium by removing 2 mL of medium above the cell pellet and adding 2 mL of fresh RPMI culture medium.
5. Once the first cells aggregates of lymphoblasts become visible, add ~25 mL of fresh RPMI culture medium.

3.1.3. Amplification of the Lymphoblastoid Cell Line

1. Once the cellular growth is sufficient (cell density is ~10^6 cells per mL for 20 mL), transfer the cell culture from 25-cm^2 flask to a 75-cm^2 flask; add medium to the final volume of 50 mL.
2. Using a Malassez cell estimate the cell density (every 3–4 days) and determine the quantity of fresh medium required to achieve the optimal cell density (*see* **Note 9**).

3.2. Lymphoblastoid Cell Lines Lysis and Peptide Extraction

1. Gently shake the flask to homogenize the cellular suspension.
2. Estimate the cellular density by counting on Malassez cells.
3. Transfer cell culture into a centrifuge flask (approximately 5×10^7 lymphoblasts) and centrifuge for 5 min at 500×*g* (at room temperature).
4. Wash the lymphoblast pellet five times with 10 mL PBS; separate cells by centrifugation as described above. The aim of the washes is to avoid any carry-over contamination with the components of the culture medium.
5. Add 1 mL of a 2 M acetic acid lysis solution to the pellet and vortex vigorously.
6. Sonicate the acidified solutions on ice: three cycles of 20 s each, with a 1 min resting period.
7. Incubate for 45 min at 4°C under gentle stirring and then centrifuge 10 min at 11,000×*g* at 4°C.
8. Divide the supernatant in two aliquots of 100 and 900 μL and store at –80°C until use.

9. Use the smaller aliquot (one 100-μL fraction aliquot) to quantify the total proteins using Hitachi clinical analyzer 7180 or other suitable method.

3.3. Solid-Phase Extraction

1. Acidify the peptide extract with 50 volumes of 0.1% TFA, 45 mL of acidified water per 0.9 mL of extract.
2. Centrifuge 10 min at 11,000×*g*, at room temperature.
3. Solvate the cartridge with methanol and equilibrate with acidified water (0.1% TFA).
4. Load the acidified extract onto two serially connected columns. The pH of the extract should be below 4.0 (*see* **Note 10**).
5. Elute the peptide fraction with 60% acetonitrile in acidified water (*see* **Note 11**). For better recovery of the peptides use 5–10 times the hold-up volume of the cartridge.
6. Remove the organic solvent from the 60% Sep-Pak fraction (SPE extract) by freeze-drying in the concentrator.
7. Reconstitute the SPE extract with 50 μL ultrapure water and store at –25°C until use.

3.4. RP-HPLC Profiling of Lymphoblasts Peptides Extracts

1. Equilibrate analytical C_{18} reversed-phase column with 2% acetonitrile in acidified water (0.1% TFA). Acidify an aliquot of the reconstituted SPE extract (an equivalent of 530 μg of proteins as measured in **Section 3.2**) with 0.1% TFA (60 μL) and load it onto the column.
2. Elute peptides with a linear gradient of acetonitrile in acidified water at a flow rate of 0.8 mL/min. Flow rate of 0.8–1 mL/min would be suitable for a column having internal diameter of 4.6 mm.
3. Fractions are collected manually based on the absorbance measured either at 214 nm (increased sensitivity, but higher background) or at 225 nm (best signal/solvent ratio) (*see* **Note 12**).
4. Vacuum dry the fractions and redissolve them in 50 μL of ultrapure water.
5. Store fractions at –25°C until use.

3.5. Mass Spectrometry (MALDI-TOF MS)

There is a plethora of different matrixes and sample preparation techniques suitable for use with MALDI-TOF MS analyses (10). The choice of matrix and sample preparation depends on the molecular mass of the compounds and the complexity of the sample to be analyzed. Alpha-cyano-4-hydroxycinnamic acid (4HCCA) is the preferred matrix for peptides/polypeptides below 15 kDa, and the sandwich sample preparation can be universally used for molecular mass determination of complex

mixtures. The procedures reported below are the ones we used for studying the peptidome of lymphoblastoid extracts from patients affected with CACH/VWM leukodystrophy (childhood ataxia with central hypomyelination/vanishing white matter disorder) versus healthy patients (7). Sandwich preparation is used with the 4-HCCA matrix types. This method is derived from the fast-evaporation and overlayer method (11).

1. Deposit 0.5 μL of a saturated solution of 4-HCCA in acetone on the stainless steel sample plate. Wait until dry, deposit 0.5 μL 0.1%TFA on the crystallized matrix bed, then add 0.5 μL of sample, and finally add 0.5 μL of a saturated solution of 4-HCCA in 50% acetonitrile prepared in acidified water (0.1% TFA) (*see* **Note 13**).
2. Dry the sample plate under moderate vacuum, do not wash (*see* **Note 14**).
3. Insert the sample plate in the MALDI-TOF mass spectrometer.
4. Mass spectra are recorded using the positive linear mode and the delayed extraction method.
5. Calibration is performed externally with a mixture of bovine pancreas insulin with signals at *m/z* 5,734.57 ($(MH)^+$) and *m/z* 2867.78 ($(M+2H)^{2+}$) and horse myoglobin with signals at *m/z* 16,952.5 ($(MH)^+$) and *m/z* 8476.7 ($(M+2H)^{2+}$).

4. Notes

1. Foetal calf serum is heat-treated 10 min at 60°C, and kept as 1-mL aliquots, and then stored at –80°C until use.
2. The lysis/extraction solution used in our reference study is a 2 M acetic acid solution. We also evaluated three other buffers and extraction procedures: (i) homogenization in 1 mL of HB buffer (45 mM HEPES, pH 7.4, 375 μM magnesium acetate, 75 μM EDTA, 95 mM potassium acetate, 2.5 mg/mL digitonin, microcystin, and 10% (v/v) glycerol); (ii) homogenization in 1 mL HB and sonication (three cycles of 20 s each); and finally (iii) sonication (three cycles of 20 s each, with a 1 min resting period) alone in 1 mL PBS. The optimal conditions for the extraction of the peptides/polypeptides were determined through a three-step procedure. First, we performed direct mass fingerprints of the crude extracts using MALDI-TOF MS, then we analyzed the extracts by RP-HPLC profiling, and finally we did mass spectra analyses of some of the individual

fractions by MALDI-TOF MS. In our hands, extraction of the lymphoblasts with 2 M acetic acid under sonication appeared the most suitable for fractionating complex peptidomes.

3. Optionally, cells can be sheared mechanically using Dounce homogenizer prior to sonication. Such pre-treatment would help to disrupt cell aggregates and improve the efficiency of cell lysis and protein and peptide extraction. Use 2 M acetic acid or other buffers recommended in **Note 2**.
4. Alternative manual technique such as the classical Bradford method can be used to measure the total protein concentration from lymphoblast extracts. We prefer using an automated procedure to ensure best reproducibility.
5. Centrifuge-based vacuum driers (such as Speed-Vac concentrator) are preferred, because the sample can be concentrated and dried at the bottom of the vial. That reduces sample losses and allows subsequent resuspension of such dried sample in a smaller volume.
6. Two types of matrix have been tested and compared for such analyses: (7) saturated 4-HCCA and sinapinic acid at concentration of 10 mg/mL in 27% acetonitrile in acidified water. Although these two matrixes are adapted for proteins analyses, the 4-HCCA matrix is more appropriate for peptides investigations (increased sensitivity).
7. One Leucosep tube could be used to treat up to 30 mL of blood sample.
8. Following centrifugation, the red blood cells should be below the filter of the Leucosep tube. If the red blood cells are not separated, repeat the experiment. Red blood cells could be separated manually by centrifugation: spin blood for 20 min at $500 \times g$ using 15 mL Lymphoprep in a 50-mL Falcon tube; the interface between the two phases contains the red blood cells (between plasma and the platelets), and it should be aspirated gently with a 10-mL pipette and transferred in a fresh new 50-mL Falcon tube.
9. The cellular density is estimated using a Malassez cell. The optimal cellular density is around 0.5×10^6 per mL and this should be achieved by adding fresh medium. The dilution should not exceed twofold. The volume of culture medium added to a 75-cm^2 flask is usually between 50 and 150 mL.
10. When more than one cartridge is used to purify one sample, loading is performed on serially connected columns. But to achieve better recovery of the peptides, each cartridge is treated individually during the elution step.

11. The samples should be acidified before loading onto the SPE support, typically with TFA or acetic acid. A stepwise elution with low, medium, and high percentage of acetonitrile in acidified water could be performed. We prefer to recover all the peptides within a single SPE fraction eluted with 60% acidified acetonitrile to reduce dilution of the material and the associated losses. Having fewer fractions would also simplify interpretation of mass spectra. Salts, sugars, and most hydrophilic proteins are eliminated during the washing cycle, whereas lipids and most hydrophobic proteins are retained irreversibly on the solid phase.
12. Peptide elution is monitored by measuring the UV absorbance at 225 nm (detection of the peptide bond, high sensitivity). It might be easier to collect the eluted peaks manually, rather than by using automatic sample collector to collect entire fraction and then reducing the number of subsequent purification steps. This would also minimize the number of steps required to achieve pure material for structural analysis. Automated systems where the fraction collection in RP-HPLC is triggered by MS detection could also be used, for example Autopurification system (Waters). The latter may be set to collect mass (with m/z specified by the user) in a single fraction.
13. Two sample preparations have been tested, namely the sandwich preparation, described here, and the dried-droplet technique. The latter requires that the sample is mixed with the matrix solution (4-HCCA or sinapinic acid) and typically required more sample material. We chose to use the sandwich technique in order to save sample for further studies, and it worked well in our hands.
14. It is possible to add an additional washing step, by washing the dried sample on the target plate with 1 μL of 0.1% TFA. This shall further decrease MS background and improve the sensitivity of detection.

Acknowledgments

The authors would like to thank the ELA (European Leukodystrophy Association) Foundation for their financial support, Atheris Laboratories (Geneva, Switzerland) for the help with mass spectrometry analyses, and Patricia Combes and Céline Gonthier for the help with cell culture work performed at UMR INSERM U931/CNRS 6247 (Clermont-Ferrand, France).

References

1. Barnouin, K. (2004) Two-dimensional gel electrophoresis for analysis of protein complexes. *Methods Mol. Biol.* **261**, 479–498.
2. Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J., and Gygi, S.P. (2003) 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.
3. Smith, R.D. (2000) Trends in mass spectrometry instrumentation for proteomics. *Trends Biotechnol.* **20**, S3–S7.
4. Joubert-Caron, R., Le Caer, J.P., Montandin, F., Poirier, F., Pontet, M., Imam, N., Feuillard, J., Bladier, D., Rossier, J., and Caron, M. (2000) Protein analysis by mass spectrometry and sequence database searching: a proteomic approach to identify human lymphoblastoid cell line proteins. *Electrophoresis* **21**, 2566–2575.
5. Caron, M., Imam-Sghiouar, N., Poirier, F., La Caër, J.P., Labas, Y., and Joubert-Caron, R. (2002) Proteomic map and database of lymphoblastoid proteins. *J. Chromatogr. B.* **771**, 197–209.
6. Toda, T. and Sugimoto, M. (2003) Proteome analysis of Epstein-Barr virus-transformed B-lymphoblasts and the proteome database. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **787**, 197–206.
7. Fogli, A., Malinverni, C., Thadikkaran, L., Combes, P., Perret, F., Crettaz, D., Tissot, J.D., Boespflug-Tanguy, O., Stöcklin, R., and Bulet, P. (2006) Peptidomics and proteomics studies of transformed lymphocytes from patients mutated for the eukaryotic initiation factor 2B. *J. Chromatogr. B.* **840**, 20–28.
8. Fogli, A. and Boespflug-Tanguy, O. (2006) The large spectrum of eIF2B-related diseases. *Biochem. Soc Trans.* **34**, 22–29.
9. Fields, G.B. (2007) Peptide Characterization and Application Protocols, in Methods in Molecular Biology (Fields, G.B., ed.) Hardcover, vol. **386**. Humana Press Inc., Totowa, New Jersey.
10. Bulet, P. and Uttenweiler-Joseph, S. (2000) A MALDI-TOF mass spectrometry approach to investigate the defense reactions. In *Drosophila melanogaster*, an insect model for the study of innate immunity, in Proteome and Protein Analysis (Kamp, R.M., Kyriakidis, D., and Choli-Papadopoulos, T., ed.). Springer-Verlag, Berlin, pp. 157–174.
11. Li, L., Golding, R.E., and Whittal, R.M. (1996) Analysis of single mammalian cell lysates by mass spectrometry. *J. Am. Chem. Soc.* **118**, 11662–11663.

Chapter 20

Peptidomics: Identification of Pathogenic and Marker Peptides

Yang Xiang, Manae S. Kurokawa, Mie Kanke, Yukiko Takakuwa, and Tomohiro Kato

Abstract

Recent years have seen great advances in mass spectrometry and proteomics, the science dealing with the analysis of proteins, their structure and function. A branch of proteomics dealing with naturally occurring peptides is often referred to as peptidomics. Direct analysis of peptides produced by processing or degradation of proteins might be useful for example for detecting and identifying pathogenic and/or biomarker peptides in body fluids like blood. In this paper, we introduce one of the standard protocols for comprehensive analysis of serum-derived peptides, which consists of methods for purification of serum peptides, detection of peptides, pattern recognition and clustering (bioinformatics), and identification of peptide sequences. Peptide identification should be followed by the investigation of their pathogenic roles using for example synthetic peptides and the establishment of their usefulness as bioclinical markers.

Key words: Peptidomics, body fluid, serum, peptides, mass spectrometry.

1. Introduction

The human body contains a broad spectrum of proteins in intracellular and extracellular compartments. Differential expression of proteins and peptides in disease has been the focus of numerous studies (1–4), which aimed to understand disease pathogenesis, to establish therapeutic targets and relevant diagnostics markers. Blood is a good example of a body fluid which is convenient to obtain and which contains large numbers of various proteins. However, more than 99% of the plasma protein fractions are made of approximately 20 major proteins such as haemoglobin,

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_20,

albumin and globulins (5). Therefore, the remaining ~1% of the plasma proteins should be investigated if disease-related proteins are sought. In fact, biologically important proteins such as hormones, cytokines and chemokines account for only minor parts of the plasma proteins and considerable efforts were spent to detect and identify such low-abundance proteins. Peptides produced by processing and/or degradation of proteins with exopeptidases or specific endopeptidases represent an alternative to protein targets. Some such protein-derived peptides can have biological functions and be disease-related markers. Until recently it has been difficult to detect directly such short peptides in body fluids; however, the dramatic improvement of mass spectrometry made it possible. Many such peptides or their parent proteins have been reported to be associated with pathogenesis of a particular disease and/or to be useful disease markers in cases of cancers (6, 7), diabetes (8, 9), neural diseases (10) and collagen diseases (2). For example, in the field of autoimmune diseases, we reported that complement C3f-des-arginine peptide, detected predominantly in systemic sclerosis (SSc) sera, enhanced proliferation of vascular endothelial cells (2). Mass spectrometry is now universally applied to study various types of body fluids including blood (2, 11) urine (12, 13), and cerebrospinal fluid (14, 15).

In this chapter, we describe protocols for peptidomic analysis of serum peptides. These include 1) collection of peripheral blood from patients 2) separation of sera from the blood 3) purification of peptides from the serum samples 4) detection of individual peptides 5) pattern recognition and clustering (bioinformatics) and 6) sequence identification of the peptides of interest. Biological and pathological functions of the identified peptides and their usefulness as disease markers can be elucidated using synthetic peptides and their parent proteins.

2. Materials

2.1. Collection of Peripheral Blood Samples and Separation of Sera

1. 5-ml syringes, 21 G needles and blood-collecting tubes (Terumo, Tokyo, Japan)
2. A centrifuge (Himac CR21, Hitachi, Tokyo, Japan)
3. Microcentrifuge tubes (Quality Scientific Plastics, Waterford, MI)

2.2. Purification of Serum Peptides

1. Magnetic bead-based hydrophobic interaction chromatography 18 (MB-HIC18, Bruker Daltonics, Ettlingen, Germany). The kit contains MB-HIC binding solution, MB-HIC wash solution and deionized water.

2. 8-strip 0.2-ml thin-wall PCR tubes and caps (Quality Scientific Plastics)
3. Magnetic bead separator (MBS, Bruker Daltonics)
4. Acetonitrile, gradient-grade for liquid chromatography (Merck, Darmstadt, Germany). 50% acetonitrile should be prepared by diluting with deionized water.

2.3. Acquisition of Peptide Peak Spectra with flexControl

1. Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometer (Ultraflex, Bruker Daltonics) and the software (*flexControl*, Bruker Daltonics)
2. Matrix solution (*See* **Note 1**)
3. Target metal plates for mass spectrometry analysis (AnchorChip 600 μm, Bruker Daltonics)
4. ClinProt Standard (CPS, *See* **Note 2**)

2.4. Pattern Recognition and Clustering (Bioinformatics)

1. *ClinProTools* (CPT) software (Bruker Daltonics).

2.5. Identification of Peptides of Interest (MS/MS Analysis)

1. MALDI-TOF/TOF mass spectrometer (Ultraflex TOF/TOF, Bruker Daltonics) and the analysing software (*flexAnalysis*, Bruker Daltonics)
2. Protein search tools (*See* **Note 3**)

3. Methods

3.1. Collection of Peripheral Blood Samples and Separation of Sera (See* Note 4*)

1. Collect 5 ml peripheral blood from patients with a disease of interest and control patients using disposable 5-ml syringes with 21 G needles.
2. Transfer the obtained blood into the blood-collecting tubes.
3. Leave the blood for approximately 30 min or until it makes a clot.
4. Centrifuge the aggregated blood samples at 1500×*g* for 10 min.
5. Recover sera using a pipette and transfer it into the microcentrifuge tubes. The serum samples should be divided into small aliquots to avoid repeated freezing and thawing and stored at –80°C.

3.2. Purification of Serum Peptides (see* Notes 5 and 6*)

1. Shake carefully the magnetic bead MB-HIC18 solution to obtain a homogenous suspension.
2. Transfer 10 μl MB-HIC binding solution and 5 μl serum to a thin-wall microcentrifuge tube.

3. Add 5 μl magnetic bead solution and mix the resultant 20 μl solution carefully by pipetting. Leave the tube for 1 min to allow peptides to bind the beads.
4. Place the tube on MBS and keep it there for 20 s to separate the magnetic beads from solution.
5. Aspirate the solution carefully using a pipette and leave only the beads in the tube (*See* **Note 7**).
6. Transfer the tube from MBS to an ordinary tube stand.
7. Add 100 μl MB-HIC wash solution to wash the peptide-binding beads.
8. Place the tube on MBS and move the tube back and forth 20 times to wash the beads well.
9. Keep the tube on MBS for 20 s to collect the magnetic beads and then discard the solution using a pipette carefully.
10. Repeat washings (Steps 6–9) two more times.
11. Add 5 μl of 50% acetonitrile and mix thoroughly. Incubate the beads with the elution buffer in the tube for 1 min to elute the peptides entirely.
12. Place the tube on MBS and wait for 30 s to collect the magnetic beads.
13. Transfer the peptide solution into a fresh tube.

3.3. Acquisition of Peptide Peak Spectra with flexControl

1. Dilute 1 μl of the peptide solution with 10 μl of the matrix solution (*See* **Note 8**).
2. Apply 1 μl of the peptide/matrix solution to a 600-μm spot on AnchorChip plate and air dry the spot (*see* **Notes 9 and 10**). Leave a few empty 600-μm spots for standards.
3. Mix 1 μl ClinProt Standard (CPS) with 1 μl matrix solution.
4. Apply 1 μl of the above mixture to a 600-μm spot onto the loaded AnchorChip plate and air dry the spot.

3.3.1. Target Preparation

3.3.2. Activation of the flexControl

1. Prepare your own measurement method for peptide peak spectra by modifying one of model methods pre-installed on MALDI-TOF mass spectrometer (Ultraflex). The parameters (i.e. laser powers and shot numbers) should be changed according to the mass range of the targeted peptides.
2. Activate the *flexControl* software and load the selected measurement method.
3. Prepare a run sheet by using an Excel file including the following information: sample names, their position on the anchor chip, turns of shooting, the measurement methods, and the folder in which the data will be saved.

3.3.3. Calibration of the Mass Measurement System Using CPS

1. Insert the prepared AnchorChip into the mass spectrometer (Ultraflex).
2. Move the target AnchorChip to the CPS position for laser irradiation.
3. Tune the measurement system (i.e. the laser power, the number of laser shots, summation of spectra) to get sharp spectra of the standard peptides included in CPS.
4. Calibrate the mass spectrometer using the obtained m/z values of the standard peptides.

3.3.4. Acquisition of Peptide Spectra

1. Click "AutoXecute" tab.
2. Select the prepared measurement method.
3. Click "Run method on current spot" and move the target AnchorChip to the sample position for laser irradiation. Measure the peak spectra of a few samples as trials.
4. If the acquired data are not satisfying, tune the measurement method by changing parameters (i.e. the laser power, the number of laser shots, summation of spectra and the number of targeted positions).
5. Save the modified method, use different filename.
6. Click "Load" button to select and load the prepared run sheet.
7. Click "Start automatic Run" button to measure all the peptide spectra automatically.
8. An example of the results produced is shown in **Fig. 20.1**.

3.4. Pattern Recognition and Clustering (Bioinformatics)

To activate *ClinProTools* click "Start", then "Programs", "Bruker Daltonics" and then "ClinProtTools".

3.4.1. Activation of ClinProTools

3.4.2. Model Generation

1. Prepare a training data set. This should include peptide spectrum data of at least two groups: for example, the data of ten cases with a disease of interest (disease A) and ten control cases or another disease (disease B). The data should be prepared in different folders according to the disease groups (e.g., "folder A" and "folder B").
2. Load the training data sets. From "File" menu, select "Open model Generation Class" and click the browse button of "class 1" to select "folder A" and then click the browse button of "class 2" to select "folder B". At least two classes are needed for this procedure of bioinformatics.
3. Start peak statistics calculation. From "Reports" menu, use "Peak Statistics" command. This includes spectra

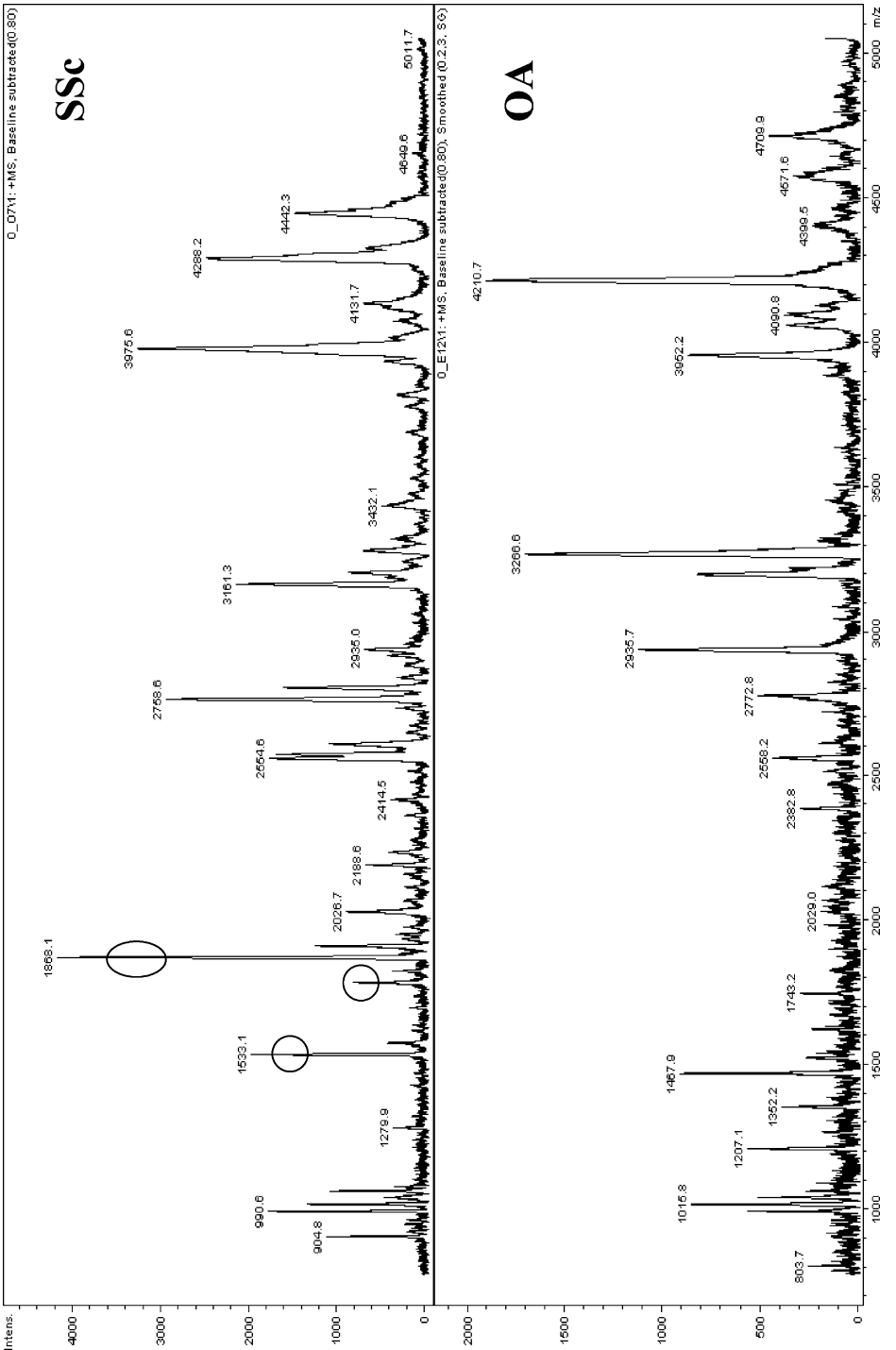


Fig. 20.1. MALDI spectra of peptides detected in sera from SSc and osteoarthritis (OA). Serum peptides were purified by MB-HIC18 and detected by MALDI-TOF mass spectrometer. The peptides specific to SSc sera are indicated by circles.

recalibration, spectra averaging and peak calculation of the loaded data.

4. View and confirm the results of peak statistics calculation. "Peak Statistic" window automatically appears after the calculation, and all the selected peaks listed by the selected sort menu are shown with their statistical data (masses of peaks, averages of peak intensities and P-values). "Spectra View" window shows all the picked up peaks highlighted with their integration regions as light-blue.
5. Select the classification algorithm from the following three algorithms: Genetic Algorithm (GA), Support Vector Machine algorithm (SVM) and QuickClassifier algorithm (QC) (*See* **Note 11**). From "Model Generation" menu, use "New model" command to select one of the algorithm and click "OK" to enter a model name. Repeat this procedure for each of the algorithms (*See* **Note 12**).
6. Start model calculation. From "Model Calculation" menu, use "Calculate" command.
7. View and confirm the results of the model calculation (the generated models are in the model list). From the "Reports" menu, use the "Model List" command to view parameters of the models. From the "Model List View's context-sensitive menu", use the "Show Model" command to open each "Model" report. In the "Spectra View" window, the peptide peaks incorporated in the current model have red highlighted integration regions instead of light-blue ones.
8. If a good model is obtained (i.e. patient data are correctly divided into two groups) save it using the "Save Model As" command from the "Model list View's context-sensitive menu".

3.4.3. Validation

Validate the generated model by using a test data set.

1. Prepare a test data set. The test data set should be different from the training data set used to generate the model. For example, the data of 20 new cases with disease A and that of 20 new cases with disease B should be prepared. These data should be prepared in new and different folders according to the disease groups (e.g. "Folder A2" and "Folder B2").
2. Click "Load model" button and select the model, generated and saved in Step 8 (*see* **Section 3.4.2**).
3. From the "Classification" menu, select the "External Validation" command to start validation.
4. Load the test data set. The data in "Folder A2" should be loaded as "Class 1" and that in "Folder B2" as "Class 2".

5. After the validation, the matrix of "Validation" report is shown. Evaluate the model using the number of true positives (i.e. the cases with disease A that were classified as disease A), false positives (i.e. the cases with disease B that were classified as disease A), true negatives (i.e. the cases with disease B that were classified as disease B), false negatives (i.e. the cases with disease A that were classified as disease B) and the calculated sensitivity and specificity (*See* **Note 13**).

3.4.4. Classification

Classification of peptide MS data for the diagnosis of unknown cases is performed using the validated model obtained (*see* **Section 3.4.3**).

1. Load the model. From the "Model Generation" menu, use the "Load Model" command and select the model generated in **Section 3.4.2** and validated in **Section 3.4.3.**
2. Load the unknown data (these should be saved in a separate folder). From the "Classification" menu, select "Classify" command and load the unknown data.
3. View and confirm the "Classification" report. The "Classification" window automatically opens and lists the classification results. In each case, evaluate the diagnosis obtained by the classification by comparing it with the clinical diagnosis.
4. If the peptide amino acid sequences and their parent proteins need to be identified and incorporated into the model, proceed to the next step of **Section 3.5**.

3.5. Identification of Peptides of Interest (MS/MS Analysis)

See **Section 3.2** for the purification of serum peptides and preparation of the peptide solution (*See* **Note 14**).

3.5.1. Preparation of the Purified Peptide Solution

1. Active *flexControl* software.
2. Select and load "MS/MS method".

3.5.2. MS/MS Analysis

3. Select the parameters for the analysis (i.e. "Mode" as "Reflector").
4. Select the mass spectrum of the peptide of interest for the following secondary mass spectrometry analysis.
5. Measure the precursor ion of the entire peptide mass spectrum.
6. Fragment the selected peptide using TOF/TOF mode of Ultraflex (*See* **Note 15**).
7. Measure all of the peptide-derived fragment ions as "b" ion peaks and "y" ion peaks.
8. Determine amino acid sequence of the peptide from the MS/MS spectra using *flexAnalysis* software. An example of the results produced is shown in **Fig. 20.2**.

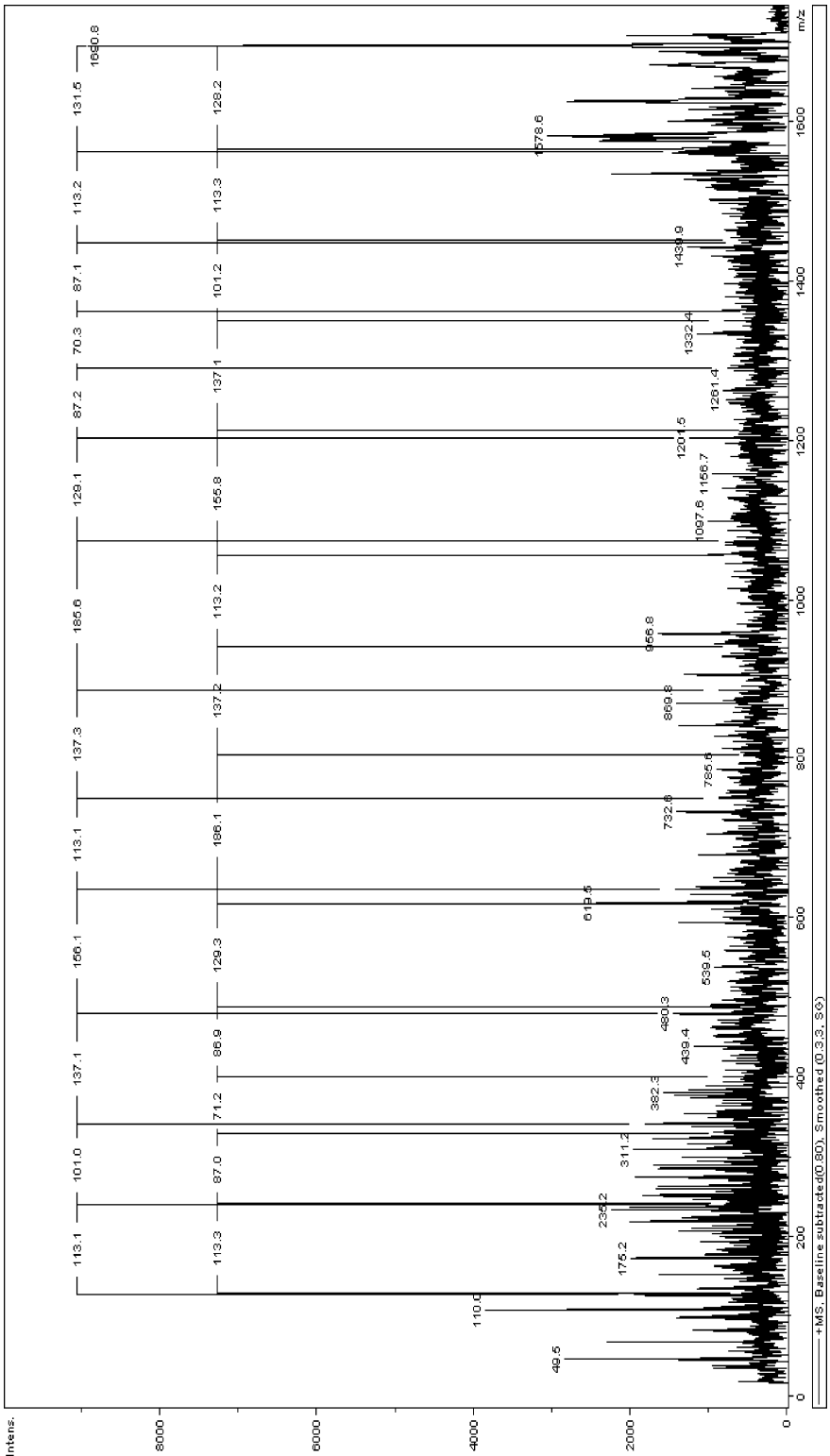


Fig. 20.2. Amino acid sequence analysis of a peptide upregulated in SSc sera by MS/MS analysis. The peptide KITHRIHWESASLL was upregulated in sera from SSc compared to OA and rheumatoid arthritis. This 1690 *m/z* peptide was fragmented using a number of different laser power settings and the mixture of all of the peptide-derived fragments was obtained. Numbers show the mass of each amino acid obtained from "b" ion peaks (the *upper line*) and "y" ion peaks (the *lower line*).

3.5.3. Identification of the Parent Proteins of the Peptides

After the determination of amino acid sequences of the peptides, the parent proteins can be identified using protein search tools.

1. Select and visit one of the protein data banks on their websites such as Mascot or FASTA (*See* **Note 3**).
2. Enter key sequence(s) of the obtained identified peptides.
3. Select the target species.
4. Search the protein database of the target species for the parent protein of the peptide of interest.
5. Evaluate the sequence analysis results (*See* **Note 16**). If only one protein is obtained with a significant score, it is likely to be the parent protein of the peptide of interest.
6. Evaluate the identified parent protein and the peptide of interest whether they play roles in the pathogenesis/pathophysiology of the disease and whether they are useful as disease markers by conducting other kinds of experiments (*See* **Note 17**).

4. Notes

1. To prepare the matrix solution,α-cyano-4-hydroxycinnamic acid (HCCA, Bruker Daltonics), 100% ethanol and 100% acetone are needed. Mix 100% ethanol and 100% acetone as 2:1 and add HCCA to the mixture to prepare its final concentration as 0.3 g/l.
2. ClinProt Standard (CPS) is mixture of several standard peptides and proteins with known mass information. CPS is used for tuning of instrument settings and calibration. Materials for preparing CPS are as follows: 0.1% TFA, 100% acetone, 100% ethanol, 10 mM ammonium acetate (77 mg in 100 ml Milli-Q water), Peptide Calibration Standard (Bruker Daltonics) and Protein Calibration Standard I (Bruker Daltonics). To prepare CPS, first solubilize Peptide Calibration Standard and Protein Calibration Standard I separately in 125 μl of 0.1% TFA for 5 min at RT and vortex for 1 min. Mix 5 μl Peptide Calibration Standard, 20 μl 10 mM ammonium acetate and 25 μl Protein Calibration Standard I for 1 min by vortexing. Use fresh for some days stored at 4°C or store in aliquots at –20°C. This standard is stable for some weeks at –20°C.
3. There are several protein analysis tools freely available online such as Mascot (www.matrixscience.com) and Pro-

teinProspector (prospector.ucsf.edu). Extensive selection of tools with adequate commentaries is available from the ExPASy Proteomics tools website (www.expasy.ch/tools).

4. Biological fluid such as urine, cerebrospinal fluid, bile and synovial fluid requires pre-treatment prior to peptide purification. For example, synovial fluid from patients with arthritis/arthropathy should be filtrated using 50–70 μm meshes because of its high viscosity.
5. To collect short peptides from serum samples effectively, filtration through 3-kDa filters (Microcon; Millipore, Bedford, MA) is sometimes used.
6. The other magnetic beads such as MB-HIC8, MB-IMAC Cu and MB-WCX are also available instead of MB-HIC18 for obtaining peptides. The obtained peptide profiles are different in part between the columns.
7. Keep the pipette tips well away from the magnetic beads to not aspirate the beads.
8. Prepare the fresh matrix solution daily.
9. It is recommended to work continuously from preparing the diluted peptide solution to the application of it because matrix- and several sample-solutions contain very volatile solvents.
10. If necessary, re-crystallization is recommended. In this case, apply 1 μl of the following re-crystallization solution (mixture of 100% ethanol, 100% acetone and 0.1% TFA at ratios of 6:3:1). Dry up the plate by air for 5–10 min.
11. The three kinds of algorithms for generating classification models are different in their methodology and have advantages and drawbacks. Genetic Algorithm (GA) mimics evolution in nature and is used to select the peak combinations which are most relevant for separation. This algorithm is frequently used for the comparison of two groups. Support Vector Machine algorithm (SVM) is used to determine separation planes between the different data classes. Upon the obtained planes, a peak ranking can be calculated. This algorithm is frequently used for comparison of more than two groups. QuickClassifier algorithm (QC) stores the class averages of the peak areas in the model together with some statistical data like p-values at certain peak positions. In this algorithm, the peak areas are sorted per peak and a weighted average over all peaks is calculated for the classification.
12. Try all the three algorithms and select the best in Step 8 in the **Section 3.4.2**.

13. Sensitivity and specificity are calculated by the following formulas: Sensitivity = true positives/(true positives + false negatives). Specificity = true negatives/(true negatives + false positives).
14. When multiple peptide peaks appear around the target peptide peak, identification of the amino acid sequence by MS/MS analysis might be difficult. In such a case two-dimensional liquid chromatography is recommended prior to mass spectrometry analysis. As a result, the multiple peptides are separated into different fractions and the amount of the target peptide can be increased by increasing the initial amount of the purified serum sample for the chromatography.
15. Fragment the peptide of interest using various levels of the laser power. Increase the laser power gradually to achieve the best fragmentation.
16. Typically multiple proteins will be listed. Different sites use different tools and the scores obtained might be different even if the same data sets were submitted. To evaluate the ranking obtained, refer to the description of the scores on the particular website help pages.
17. Often the amount of the parent protein and that of the protein-derived peptides correlate inversely, indicating that the peptides are generated by processing/degradation of the protein. The serum concentration of the parent protein can be measured by using sandwich ELISA etc. There are various ways to examine the function of the peptides. One is to synthesize the peptides and add each of them into a cell culture system, in which cell lines related to the disease of interest (disease A) are used (i.e. vascular endothelial cell lines for vasculitis). In this way, the change of cytokine secretion from the cells in culture supernatant can be detected by sandwich ELISA and the change of any protein expression of the cells can be analysed by proteomics analysis. To detect disease markers, multivariate analysis of the peptide spectrum data is useful (Simca-P+, Umetrics, Kinnelon, NJ). For example, mixture of two patient groups (disease A and disease B) is re-divided into the two original groups by using the multivariate analysis of the peptide spectra. In this case, the peptides which have high magnitude and/or high reliability for the division are selected as excellent candidates for the disease markers. Evaluate individual peptides by their serum concentration in the patients.

Acknowledgments

We greatly appreciate the help and consent from Bruker Daltonics.

References

1. Hudelist, G., Singer, C.F., Pischinger, K.I., Kaserer, K., Manavi, M., Kubista, E., and Czerwenka, K.F. (2006) Proteomic analysis in human breast cancer: identification of a characteristic protein expression profile of malignant breast epithelium. *Proteomics* **6**, 1989–2002.
2. Xiang, Y., Matsui, T., Matsuo, K., Shimada, K., Tohma, S., Nakamura, H., Masuko, K., Yudoh, K., Nishioka, K., and Kato, T. (2007) Comprehensive investigation of disease-specific short peptides in sera from patients with systemic sclerosis: complement C3f-des-arginine, detected predominantly in systemic sclerosis sera, enhances proliferation of vascular endothelial cells. *Arthritis Rheum.* **56**, 2018–2030.
3. Siest, G., Marteau, J.B., Maumus, S., Berrahmoune, H., Jeannesson, E., Samara, A., Batt, A.M., and Visvikis-Siest, S. (2005) Pharmacogenomics and cardiovascular drugs: need for integrated biological system with phenotypes and proteomic markers. *Eur. J. Pharmacol.* **527**, 1–22.
4. Devarajan, P. (2007) Emerging biomarkers of acute kidney injury. *Contrib. Nephrol.* **156**, 203–212.
5. Anderson, N.L. and Anderson, N.G. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics.* **1**, 845–867.
6. Diamandis, E.P. (2006) Peptidomics for cancer diagnosis: present and future. *J. Proteome. Res.* **5**, 2079–2082.
7. Villanueva, J., Martorella, A.J., Lawlor, K., Philip, J., Fleisher, M., Robbins, R.J., and Tempst, P. (2006) Serum peptidome patterns that distinguish metastatic thyroid carcinoma from cancer-free controls are unbiased by gender and age. *Mol. Cell. Proteomics* **5**, 1840–1852.
8. Budde, P., Schulte, I., Appel, A., Neitz, S., Kellmann, M., Tammen, H., Hess, R., and Rose, H. (2005) Peptidomics biomarker discovery in mouse models of obesity and type 2 diabetes. *Comb. Chem. High Throughput Screen* **8**, 775–781.
9. Boonen, K., Baggerman, G., D'Hertog, W., Husson, S.J., Overbergh, L., Mathieu, C., and Schoofs, L. (2007) Neuropeptides of the islets of Langerhans: a peptidomics study. *Gen. Comp. Endocrinol.* **152**, 231–241.
10. Selle, H., Lamerz, J., Buerger, K., Dessauer, A., Hager, K., Hampel, H., Karl, J., Kellmann, M., Lannfelt, L., Louhija, J., Riepe, M., Rollinger, W., Tumani, H., Schrader, M., and Zucht, H.D. (2005) Identification of novel biomarker candidates by differential peptidomics analysis of cerebrospinal fluid in Alzheimer's disease. *Comb. Chem. High Throughput Screen* **8**, 801–806.
11. Tammen, H., Schulte, I., Hess, R., Menzel, C., Kellmann, M., and Schulz-Knappe, P. (2005) Prerequisites for peptidomic analysis of blood samples: I. Evaluation of blood specimen qualities and determination of technical performance characteristics. *Comb. Chem. High Throughput Screen* **8**, 725–733.
12. Jurgens, M., Appel, A., Heine, G., Neitz, S., Menzel, C., Tammen, H., and Zucht, H.D. (2005) Towards characterization of the human urinary peptidome. *Comb. Chem. High Throughput Screen* **8**, 757–765.
13. Norden, A.G., Rodriguez-Cutillas, P., and Unwin, R.J. (2007) Clinical urinary peptidomics: learning to walk before we can run. *Clin. Chem.* **53**, 375–376.
14. Yuan, X. and Desiderio, D.M. (2005) Human cerebrospinal fluid peptidomics. *J. Mass Spectrom.* **40**, 176–181.
15. Lamerz, J., Selle, H., Scapozza, L., Crameri, R., Schulz-Knappe, P., Mohring, T., Kellmann, M., Khameni, V., and Zucht, H.D. (2005) Correlation-associated peptide networks of human cerebrospinal fluid. *Proteomics* **5**, 2789–2798.

Section III

Tools and Approaches

Chapter 21

Peptidomic Approaches to the Identification and Characterization of Functional Peptides in *Hydra*

Toshio Takahashi and Toshitaka Fujisawa

Abstract

Little is known about peptides that control developmental processes such as cell differentiation and pattern formation in metazoans. The cnidarian *Hydra* is one of the most basal metazoans and is a key model system for studying the peptides involved in these processes. We developed a novel peptidomic approach to the isolation and identification of functional signalling peptides from *Hydra* (the *Hydra* peptide project). First, peptides extracted from the tissue of *Hydra magnipapillata* are purified to homogeneity using high-performance liquid chromatography (HPLC). The isolated peptides are then tested for their ability to alter gene expression in *Hydra* using differential display-PCR (DD-PCR). If gene expression is altered, the peptide is considered as a putative signalling peptide and is subjected to amino acid sequencing. Following the sequencing, synthetic peptides are produced and compared to their native counterparts by HPLC and/or mass spectrometry (MS). The synthetic peptides, which are available in larger quantities than their native analogues, are then tested in a variety of biological assays in *Hydra* to determine their functions. Here we present our strategies and a systematic approach to the identification and characterization of novel signalling peptides in *Hydra*. We also describe our high-throughput reverse-phase nano-flow liquid chromatography matrix-assisted laser desorption ionization time-of-flight mass spectrometry (LC-MALDI-TOF-MS/MS) approach, which was proved to be a powerful tool in the discovery of novel signalling peptides.

Key words: *Hydra*, peptide signalling molecules, peptidomics, *Hydra* peptide project, DD-PCR, LC-MALDI-TOF-MS/MS, epithelial *Hydra*.

1. Introduction

Peptides often function as signals for intercellular communication, for example neurotransmitters and hormones. However, little is known about peptides that control developmental processes such as cell differentiation and pattern formation in metazoans. Such peptides are highly diverse structurally and functionally. Once such peptides are obtained, their encoding genes

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_21,

could be easily identified. We are especially interested in small diffusible molecules that regulate development in *Hydra*. To this end we developed a novel systematic approach (the *Hydra* Peptide Project) (1, 2) for studying such functional peptides.

In conventional biochemical approaches, signalling molecules are often isolated using laborious assays and time-consuming procedures and, typically, only one at a time. For example, the first morphogenetic molecule reported in *Hydra*, an undecapeptide called head activator, required 200 kg of sea anemone, *Anthopleura elegantissima* for its purification. Each fractionation step in that work was followed by an assay to test the ability of fractions to enhance tentacle formation during head regeneration of *Hydra* (3). Later the same molecule was also purified from 3 kg of *Hydra attenuate* (3). Because the *Hydra* genome does not contain a gene encoding the head activator, the peptide presence in *Hydra* tissue is doubtful. Another interesting neuropeptide called metamorphosin A, which induces metamorphosis of planula larvae of a marine hydroid, *Hydractinia echinata*, was also purified from a large quantity of starting material (*Anthopleura elegantissima* (4)). We later identified seven more peptides, similar to metamorphosin A (1, 5). These peptides have been shown to originate from a single gene in *Hydra* (6). Using a foot-specific peroxidase assay Hoffmeister and colleagues identified two peptides, Pedibin and Pedin, which enhance foot formation in *Hydra* (7). These peptides have also been identified in the *Hydra* Peptide Project (1, 8). Neuropeptides containing the C-terminal sequence Arg-Phe-NH_2 (RFamide) are ubiquitous in the animal kingdom. In *Hydra*, RFamide-like immunoreactivity was detected in neurons using an antibody specific to RFamide (9, 10). Grimmelikhuijzen and colleagues isolated four RFamides from *Hydra magnipapillata*: Hydra-RFamide I–IV and a variety of other RFamides using antibody-based radioimmunoassays (11). In summary, approaches relying on a combination of standard biological assays and fractionation procedures so far require large amounts of starting material, are labour-intensive, and have yielded a relatively limited number of signalling molecules. This is in stark contrast with our new approach described below, which we used successfully to generate a large number of peptides with a variety of functions from a single species of *Hydra*.

2. Materials

1. Strain 105 – A standard wild-type strain of *H. magnipapillata*.
2. *Hydra* were cultured in plastic dish (35:25:5 = w:d:h, cm) containing ~1.5 L of Modified M (MM) solution: 1 mM

NaCl, 1 mM $CaCl_2$, 0.1 mM KCl, 0.1 mM $MgSO_4$, 1 mM Tris-HCl, pH 7.6 at 18°C.

3. Animals should be fed daily with freshly hatched brine shrimp, *Artemia*. The culture solution should be changed several hours after feeding. Approximately 1 g (wet weight) of *Hydra* tissue can be obtained from about 1000 polyps. For harvesting *Hydra*, collect animals by filtering with a fine mesh (500 μm, will retain polyps but not shrimps). Gently wash *Hydra* several times with fresh culture solution. Transfer the polyps to a plastic bag and freeze immediately in liquid nitrogen, store at –80°C until use (*see* **Note 1**).
4. Blender for tissues homogenization, e.g. Waring kitchen blender; Polytron homogenizer (Biotron) or similar (should be compatible with solvents, such as acetone)
5. Centrifuge (Avanti^{TM}, HP-25, BECKMAN)
6. Rotary evaporator (RE1-N, IWAKI)
7. C18 reverse-phase chromatography cartridges (Analytichem Mega Bond Elut, Varian)
8. CAPCELL PAK analytical C18 column (10 mm × 250 mm; Shiseido, Tokyo, Japan)
9. ODP-50 analytical C18 column (6 mm × 250 mm, Asahipak)
10. cDNA synthesis kit (Ammersham-Pharmacia)
11. DD-PCR primers: Anchor primer d(T_{12}AG), an arbitrarily selected deca-deoxynucleotide 5′ end primer (Operon kit, Operon Inc., CA)
12. ^{33}P-dATP (Ammersham)
13. Automated gas-phase sequencer (PPSQ-10, Shimadzu)
14. Q-TOF-MS/MS analyser (Q-TOF; Micromass)
15. ABI 433A Peptide Synthesizer (Applied Biosystems)
16. Bromodeoxyuridine (BrdU) (Sigma)
17. Maceration fluid – glycerin:acetic acid:water = 1:1:13
18. Phosphate-buffered saline: 0.15 M NaCl, 0.01 M Na_2HPO_4, pH 7.0
19. Blocking solution: 1% horse serum in PBS; ABC-AP; Vector Red substrate (VectorStain ABC kit, Vector Laboratories, CA)
20. Anti-BrdU antibody (Becton and Dickinson)
21. Size exclusion chromatography HPLC column (10 mm × 300 mm, Superdex^{TM} peptide 10/300 GL, GE Healthcare)

22. Capillary Ex-Nano inertsil C-18 reverse-phase column (0.2 × 150 mm; GL Sciences Inc, Torrance, CA)
23. Prespotted AnchorChip (Bruker Daltonics)
24. Ultraflex III mass spectrometer (Bruker Daltonics)

3. Methods

3.1. The Hydra Peptide Project

Conventional biochemical approaches rely on labourious and time-consuming functional assays often carried out after each purification step. These require large amounts of precious samples to be consumed during each assay. We developed a much simplified approach for the isolation of signalling molecules. We focused on the identification of signalling peptides (molecular weight below ~5000 daltons) for the following reasons:

(I) Peptides involved in development are scarcely known, although many important proteins have been uncovered. Furthermore, no precedent study was known to uncover a repertoire of peptides present in a single species of organisms.

(II) Peptides are usually good candidates for being morphogens (*see* **Note 2**).

(III) Peptides may have other interesting functions during the development.

(IV) Genes that encode peptides may be easily obtained and analysed with conventional Molecular Biology methods.

Our initial approach towards the identification of signalling peptides contained four major steps (schematically shown on **Fig. 21.1**; the detailed experimental procedures are described in the following sections):

(I) Peptides extracted from *Hydra* tissue (500 g ~2 kg) are purified to homogeneity in a systematic manner using HPLC; no other structural analysis is performed at this stage (*see* **Note 3**).

(II) Peptide fractions are tested functionally using differential display-PCR (DD-PCR). The selection criterion is the ability of the purified peptide to alter gene expression in *Hydra* (*see* **Notes 4 and 5**).

(III) An aliquot of the signalling peptide is used for determining the amino acid sequence with an automated peptide sequencer and/or a tandem MS.

(IV) Based on the tentative sequences, peptides are chemically synthesized and compared with the native peptides using HPLC or a MS (*see* **Note 6**).

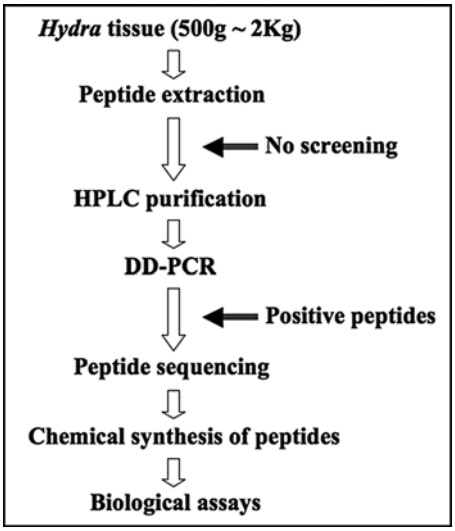


Fig. 21.1. Strategy to identify peptide signalling molecules.

During the course of our research we initiated the *Hydra* EST project; *Hydra* genome database has recently become available as well. These developments prompted us to sequence each purified peptide so it can be compared to EST or genome data, thus making a DD-PCR assay optional. If a matching nucleic acid sequence is identified in sequence databases, the putative precursor is checked for the presence of a signal peptide at the N-terminus, monobasic or dibasic processing sites before or after the peptide sequence, and the amidation motif (i.e. Glycine plus monobasic or dibasic residues) at the C-terminal flanking region (*see* **Note 7**). Our approach to peptidomic research yielded 817 peptides of which we sequenced 527 (2). Based on the results of the DD-PCR assay and the structural characteristics deduced from the precursor proteins, 55 peptides were chosen and chemically synthesized (2). Their expression patterns were studied using in situ hybridization and immunostaining with anti-peptide antibodies (*see* **Note 8**).

3.2. Peptide Extraction and Purification

Peptides were extracted and purified using two methods, differing in their approach to the inactivation of the endogenous peptidases in *Hydra* tissues. Both methods produce comparable results, although the boiling extraction (**Section 3.2.1**) results in a slightly higher rate of peptide degradation.

3.2.1. Acetone Extraction Method

1. Break up frozen tissue of *H. magnipapillata* (~500 g) using a Waring blender and homogenize the minced sample in 2 L of cold acetone using a Polytron homogenizer.

2. Centrifuge the homogenated tissue at 16,000×*g* for 30 min at 4°C.
3. Homogenize the precipitate again in 5 volumes of a 3% acetic acid solution and centrifuge using the same condition (16,000×*g* for 30 min at 4°C).
4. Combine both first and second supernatants and concentrate by rotary evaporation.
5. Add 1/10 volume of 1 M HCl and centrifuge the sample at 16,000×*g* for 30 min at 4°C (*see* **Note 9**).
6. Activate C18 cartridges with 50 ml of 100% methanol in 0.1% trifluoroacetic acid (TFA) (pH 2.2) and then wash them with 30 ml of 0.1% TFA to remove the methanol.
7. Load the clear supernatant obtained in Step 4 (above) onto two large C18 cartridges.
8. Wash the cartridges with 30 ml of 10% methanol in 0.1% TFA.
9. Elute bound peptides with 50 ml of 60% methanol in 0.1% TFA.
10. Dry the eluted samples by rotary-evaporation.
11. Prior to HPLC fractionation, dissolve the peptides in ~5 ml of 0.1% TFA.
12. Load the peptide extracts onto CAPCELL PAK C18 analytical column.
13. Elute peptides with a linear gradient of 0–60% ACN in 0.1% TFA at a flow rate of 1 mL/min for 120 min.
14. Collect fractions (2 mL/fraction), combine fractions if necessary (*see* **Note 10**).
15. Peptides from each fraction can now be purified to homogeneity using further rounds of HPLC (*see* **Note 11**).

3.2.2. Boiling Extraction Method

1. Break up *Hydra* tissue (~150 g) using a Waring blender and boil the preparation in 1.5 L of 5% acetic acid for 5 min and homogenize using a Polytron homogenizer.
2. Centrifuge the homogenated tissue at 16,000×*g* for 30 min at 4°C.
3. Homogenize the precipitate again in 5 volumes of a 5% acetic acid solution and centrifuge using the same condition (16,000×*g* for 30 min at 4°C).
4. Combine both first and second supernatants and concentrate, to reduce volume, by rotary evaporation.
5. Add 1/10 volume of 1 M HCl and centrifuge the sample at 16,000×*g* for 30 min at 4°C (*see* **Note 9**).

6. Activate C18 cartridges with 50 ml of 100% methanol in 0.1% trifluoroacetic acid (TFA) (pH 2.2) and then wash them with 30 ml of 0.1% TFA to remove the methanol.
7. Load the clear supernatant obtained in Step 4 (above) onto two large C18 cartridges.
8. Wash the cartridges with 30 ml of 10% methanol in 0.1% TFA.
9. Elute bound peptides with 50 ml of 60% methanol in 0.1% TFA.
10. Dry the eluted samples by rotary-evaporation.
11. Prior to HPLC fractionation, dissolve the peptides in ~5 ml of 0.1% TFA.
12. Load the peptide extracts onto ODP-50 C18 analytical column.
13. Elute peptides with a linear gradient of 0–50% ACN in 0.1% TFA at a flow rate of 1 mL/min for 100 min.
14. Collect fractions (2 mL/fraction); combine fractions if necessary (*see* **Note 12**).
15. Peptides from each fraction can now be purified to homogeneity using further rounds of HPLC (*see* **Note 13**).

3.3. DD-PCR (see Note 14)

1. Use ~1/4 to 1/2 of the total amount of the purified peptide to make 10^{-9}–10^{-7} M solution. Treat ~20 polyps with the peptide solution. Harvest half of the polyps after 4 h and the other half after 24 h (*see* **Note 15**).
2. Extract total RNA from treated and untreated polyps (each group containing ~10 polyps) (*see* **Notes 16–18**).
3. Make cDNA using a suitable cDNA synthesis kit, use 1 μg of total RNA as a template.
4. Use cDNA as a template for DD-PCR. Use an anchor primer d(T_{12}AG), an arbitrarily selected decadeoxynucleotide 5′ end primer and ^{33}P-dATP. DD-PCR conditions: 3 min of initial denaturation at 94°C, followed by 40 cycles consisting of 94°C for 30 s, 40°C for 30 s, and 72°C for 2 min and one final extension step of 72°C for 5 min (*see* **Note 19**).
5. Separate PCR products on polyacrylamide (6%) gel electrophoresis.
6. Dry the gel and expose it with X-ray film. Compare the band patterns obtained with control and peptide-treated polyp cDNAs. The peptides that caused alterations in mRNA expression (different band patterns) are putative signalling peptides.

7. To confirm the functional effects obtained with native peptide, repeat functional testing (Steps 1–6 above) using relevant synthetic peptides (*see* **Note 20**).
8. Use suitable methods, e.g. Edman degradation or MS analysis to perform peptide sequencing (*see* **Notes 21–23**). A combination of gas-phase sequencing and MS analysis is usually sufficient to determine the entire structure of the purified peptides.
9. Chemically synthesize peptides whose structures have been successfully determined. Use a suitable peptide synthesizer or commercial provider (we use ABI 433A synthesizer). Peptides should be of HPLC purity.
10. At this stage the identity of synthetic peptides can be further confirmed by tandem mass spectrometry or co-chromatography of the synthetic peptides with the native peptides (as in Step 15, **Sections 3.2.1** or **3.2.2** above).

3.4. Biological Assays

Biological assays suitable for testing the purified peptides are described below. All the assays should be carried out at 18°C. Use peptide concentrations between 10^{-8} and 10^{-5} M. In all cases the *Hydra* culture solution (with or without the peptides) should be replaced daily with fresh medium.

3.4.1. Morphogenesis

3.4.1.1. Regeneration Assays

Since *Hydra* has a strong regenerative capacity, morphogenesis can be observed in adult animals. Budless polyps can be cut transversely in the middle of the body column. The lower halves should be used for head regeneration assays and the upper halves for foot regeneration assays. New heads and feet are usually regenerated within 2–3 days.

1. Cut 10 polyps and immediately place upper or lower halves into a plastic dish (diameter 3 cm), containing 10 ml culture solution with or without peptide. Run triplicate experiments for each peptide.
2. For head regeneration, count the number of tentacles formed (daily).
3. For foot regeneration, look for air bubbles at the aboral tip or count polyps attached to the bottom of the culture dish.

3.4.1.2. Budding Assay

Hydra propagate asexually by budding. Bud formation begins with an evagination of tissue in the lower half of the body column. The evaginated tissue elongates into a cylindrical form, which subsequently forms a head (tentacles and mouth) at the apical end and a foot at the basal end. When the foot formation is complete,

the bud detaches from its parent and becomes a solitary polyp. Peptides can be tested for their effect on the budding frequency and bud development rates.

1. Select 10 polyps with their first bud protrusion (stage II polyps) and place them in a 24-well plate containing 1 ml of culture solution with or without peptide.
2. Count the number of buds produced by original polyps daily; continue the experiment for up to 10 days. The buds detached from their parental bodies should be removed, e.g. following each day's observations. The budding frequency can be obtained by calculating the number of buds produced per polyp per day.
3. To estimate and score bud development time, measure the time between the second or third bud protrusion and the foot formation. Foot formation shall be defined as the moment when buds are detached from the parental body, if gently pulled by a pair of tweezers.

3.4.2. Cell Proliferation and Cell Differentiation

The effects of peptides on cell proliferation and cell differentiation should be examined by BrdU labelling.

3.4.2.1. Cell Proliferation Assay

1. Transfer 10 young polyps, detached from parents during the previous 24 h (stage I polyps), in a plastic dish (diameter 3 cm) containing 10 ml culture solution, incubate with or without peptide for 48 h. Replace the solution every 24 h.
2. Pulse-label the animals with 2 mM of bromodeoxyuridine (BrdU) for 1 h (after 4 and 24 or 48 h incubation) (*see* **Note 24**).
3. After 48 h all the polyps should be macerated and detect BrdU-labelled cells (*see* **Sections 3.4.2.3** and **3.4.2.4**) (*see* **Note 25**).

3.4.2.2. Cell Differentiation Assay

1. Following the pulse-labelling, thoroughly flush the gastric cavities with fresh culture solution to remove BrdU and transfer the animals into fresh culture solution containing the same concentration of peptide. Replace the solution every 24 h.
2. The polyps should be macerated at various times after labelling and BrdU should be measured (*see* **Sections 3.4.2.3** and **3.4.2.4**). Determine the fractions of BrdU-labelled epithelial cells, 1s+2s, 4s and nerve cells (*see* **Note 26**).

3.4.2.3. Maceration of Polyps (*see* **Note 27**)

1. Transfer 10 polyps into a sample cup and remove excess of culture medium.
2. Add 250 μL of maceration fluid.

3. Incubate for ~10 min and stir the cup gently (by hand) until the tissue is completely dissociated.
4. Fix the cells by adding 250 μL of 9% formalin.
5. Spread an aliquot of macerated sample (25 μL) on gelatine-coated slide with a micropipette and air-dry for 30 min at 40°C or overnight at room temperature.

3.4.2.4. Detection of BrdU-Labelled Cells

1. Wash the slides containing macerated samples twice with phosphate-buffered saline (PBS) for 5 min.
2. Soak the slides in 3 N HCl for 30 min.
3. Wash the slides three times with PBS-Tween (0.25%) to remove HCl.
4. Block the samples with 100 μL of blocking solution for 30 min.
5. Add 100 μL of monoclonal anti-BrdU antibody diluted with blocking solution (1:30) and incubate for another 30 min.
6. Wash the samples three times with PBS-Tween (for the total of 15 min).
7. Add 100 μL of the secondary antibody (1:200 dilution with PBS) and incubate for another 30 min.
8. Wash the samples three times with PBS-Tween (for the total of 15 min).
9. Add 100 μL of ABC-AP in PBS for 30 min at room temperature.
10. Prepare fresh working solution of Vector Red substrate in 0.1 M Tris/HCl (pH 8.2).
11. Wash samples once with PBS (10 min, room temperature).
12. Add Vector Red substrate working solution, cover slides with aluminium foil, wait until suitable staining develops.
13. Wash the slides with water to stop colour reaction (*see* **Note 28**).

3.4.3. Myoactivity

Both ectodermal and endodermal epithelial cells in *Hydra* are muscle cells with muscle fibres running longitudinally in the ectoderm and circumferentially in the endoderm. Epithelial *Hydra*, which lacks cells of the interstitial cell lineage (with the exception of gland cells), represents an ideal in vivo muscle preparation to study the effect of peptides on muscle cells.

1. Epithelial *Hydra* shall be starved for 24 h prior to the myoactivity analysis.
2. Transfer the polyps into a 24-well plate and incubate for 1 h before adding peptides.

3. Add the peptides to a final concentration of 10^{-6}~10^{-5} M and stir gently.
4. Video record any movements of the polyps using a camera attached to a suitable objective lens (*see* **Note 29**).
5. Quantify the myoactivity of peptides (*see* **Note 30**).

3.5. High-Throughput Analysis of Peptides by LC-MALDI-TOF-MS/MS

The procedure detailed below is summarized in **Fig. 21.2**.

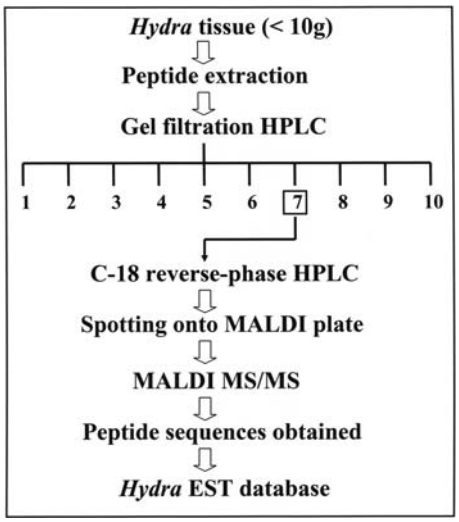


Fig. 21.2. Strategy to identify peptide signalling molecules by automated LC-MALDI-TOF-MS/MS.

1. Use acetone extraction to prepare peptides from ~10 g of *Hydra* (wet weight).
2. Fractions eluted from C18 cartridges (2 mL, as in Steps 6–11, **Sections 3.2.1** or **3.2.2**) should be further fractionated using Superdex^{TM} size exclusion HPLC and isocratic elution with 5% ACN, flow rate of 0.5 mL/min. Discard high molecular weight fractions, start collecting fractions after ~20 min, collect ~10 fractions (1 mL/fraction).
3. Apply each fraction to a C-18 reverse-phase column (Capillary Ex-Nano inertsil), elute with a linear gradient of 4–52% ACN in 0.1% TFA at a flow rate of 2 μL/min.
4. Transfer the eluate onto a Prespotted AnchorChip plate, preferably using a robotic spotter.
5. Take Mass spectra of the spotted peptides. Use 400 laser shots per spot.
6. Analyse the MS/MS data using Mascot (Matrix Science) and the *Hydra* EST database (*see* **Note 31**).

4. Notes

1. Epithelial *Hydra*, which lacks interstitial stem cells and their derivatives (except for gland cells), can be used for the myoactive experiment. These can be produced from strain 105 by colchicine treatment (12) and cultured in MM-solution containing 25 mg/L each of kanamycin and rifampicin (13).
2. It is suggested that small diffusible substances may be involved in patterning process and morphogenesis in *Hydra* (14, 15). The neuropeptide head activator was reported to be a morphogen involved in head formation (3).
3. Typically we isolate 1∼10 nmol of each peptide; this amount would be insufficient for conventional biological assays.
4. DD-PCR (16) is the only assay in our method which relies on the use of native peptide samples.
5. The DD-PCR assay was later omitted because nearly half (45%) of the purified peptides turned out to be positive in the DD-PCR assay (2).
6. The advantage here is that synthetic peptides are available in larger quantities than their native analogues and can be used for a variety of functional bioassays.
7. Since peptides could be processed at unusual processing sites in *Hydra* (and other cnidarians too), these criteria are not always applicable but have proved useful.
8. Our research yielded two groups of peptides – 10 peptides that were derived from the epithelial cells and the other is a group of 18 neuropeptides. The biological activities of these peptides exhibit a wide repertoire of functions, including in morphogenesis, cell differentiation and myoactivity (1, 2, 5, 17–22).
9. For example, 5 mL has to be added to 50 ml of the concentrated sample from previous step.
10. We divided fractions into 15 groups (2–3 fractions/group), as shown on **Fig. 21.3A**. Most of the peptides were eluted between 5 and 37% of ACN.
11. For example, group 8, which eluted at approximately 22% ACN, was subjected to additional rounds of cation-exchange HPLC (**Fig. 21.3B**). Three additional steps of HPLC were required to isolate Hym-54 (1) from the group 8 (**Fig. 21.3C–E**).

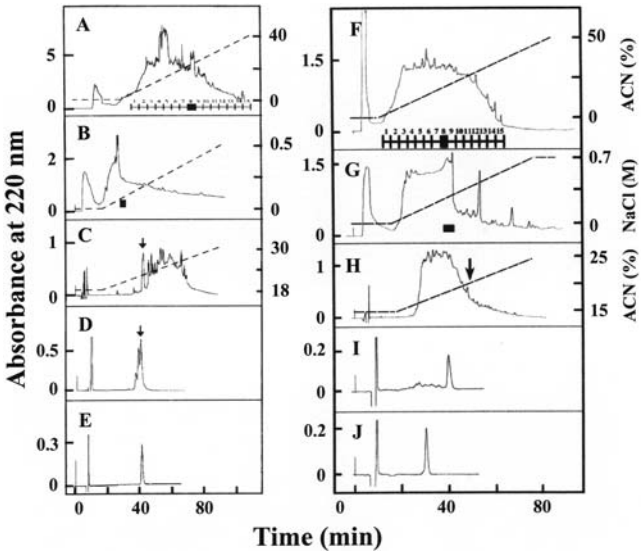


Fig. 21.3. HPLC purification of the Hym-54 (A–E) and the Hym-355 (**F–J**) peptides. Hym-54 purification (Panels A–E). Panel A: C-18 reverse-phase (RP)-HPLC 15 fractions are indicated at the bottom of the elution profile. Panel B: Cation-exchange HPLC of group 8. The fraction indicated by the bar was subjected to the next step. Panel C: C-18 RP-HPLC. Panel D: HPLC with an isocratic elution of 23% ACN. Panel E: A second round of HPLC with an isocratic elution of 23% ACN. Hym-355 purification from boiled extracts (Panels F–J). Panel F: C-18 RP-HPLC. The 15 fractions are shown at the bottom of the elution profile. Panel G: Cation-exchange HPLC of group 8. The fraction indicated by the bar was subjected to the next step. Panel H: C-18 RP-HPLC (ODS-80TM). Panel I: HPLC with an isocratic elution of 20% ACN. Panel J: Final round of HPLC with an isocratic elution of 21% ACN. The arrow (Panels C, D, H) indicates the peak fraction that was selected for subsequent purification step.

12. We divided fractions into 15 groups (2–3 fractions/group), as shown on **Fig. 21.3F**. Most of the peptides were eluted between 5 and 37% of ACN.
13. For example, group 8, which eluted at approximately 20% ACN, was subjected to additional rounds of cation-exchange HPLC (**Fig. 21.3G–J**). Four additional steps of HPLC were required to isolate Hym-355 (19) from the group 8 (**Fig. 21.3I,J**).
14. If peptides are capable of altering gene expression patterns of *Hydra*, they are considered as signalling molecules. DD-PCR developed by Liang and Pardee (16) was employed as a routine functional assay to test isolated and purified peptides.
15. To aid the penetration of some of the peptides into the tissue, dimethyl sulfoxide (DMSO) may be added to a final concentration of 2% during the first hour of peptide

treatment. If DMSO is used, the controls should also be treated with DMSO.

16. Any available method of total RNA isolation can be used, for example Acid Guanidium-Phenol-Chloroform (AGPC) method (23).
17. It is important that culture solution is removed completely from the *Hydra* sample (use a pipette with fine tip) to avoid degradation of RNA. It is also advisable to remove contaminating DNAs by adding RNase-free DNase (5 U/μL) mixed with RNase inhibitor (20 U/μL) and incubating for 30 min at 37°C. Follow this with a standard phenol-chloroform extraction. Typically about 20 μg or more of total RNA should be obtained from ten polyps.
18. Dissolve RNA in 10~20 μL of diethylpyrocarbonate (DEPC)-treated water. Quality and quantity of RNA should be checked by agarose gel electrophoresis or with a spectrometer.
19. We based our DD-PCR procedure on the protocols suggested previously by Liang and Pardee (16).
20. In our hands functional effects of native peptides were almost always reproduced with synthetic analogues so this test was later omitted from the protocol.
21. We used approximately one-fifth of the total amount of the peptide available.
22. Not all peptides can be sequenced by Edman degradation, such as peptides modified at the N-terminus by, for example, acetylation or glutamate cyclization. Such peptides should be analysed by MS.
23. We used approximately one-tenth of the total purified peptide for tandem MS analysis.
24. The pulse-labelling was performed by adding BrdU to the culture solution and then injecting the surrounding solution into the *Hydra* gastric cavity by a fine-tipped polyethylene tubing through the mouth.
25. The *Hydra* body consists of three independent non-interchangeable cell lineages; ectodermal epithelial cell, endodermal epithelial cell and interstitial stem cell lineages. All the three types of cells show constant mitotic capacity and give rise to progenitor cells. The best-studied interstitial stem cells are multipotent stem cells that continually differentiate to nerve cells and nematocytes, and also gland cells and germline cells under certain conditions. Single and pairs of interstitial cells (1s+2s) contain multipotent stem cells and early committed cells. Clusters of four dividing nematoblasts (4s) are the first morphologically recogniz-

able cell type in the nematocyte differentiation pathway. Use fractions, 1s+2s, 4s of BrdU-labelled epithelial cells and gland cells.

26. The times required for stem cells to become 4s and nerves are less than 16 h and 24 h, respectively. Therefore, for a practical reason, labelling indices of 4s and nerve cells can be determined at 24 and 48 h, respectively.
27. Our maceration procedure is based on protocols reported previously (24).
28. To remove any background staining, the samples can be treated with methanol for 5–10 min and then washed in PBS twice for 10 min.
29. We used a Sony CCD camera attached to a binocular (Nikon).
30. To quantify the myoactivity of peptides we measure the length of the body from the head to the foot or tentacles on a monitor screen following the recording. Our results are illustrated on **Fig. 21.4**, which shows two examples of myoactivity for neuropeptides Hym-248 (EPLPIGLWamide) (5) and Hym-176 (APFIFPGP-KVamide) (17). The peptides (10^{-6} M) were applied to epithelial *Hydra*. Hym-248 showed two different effects:

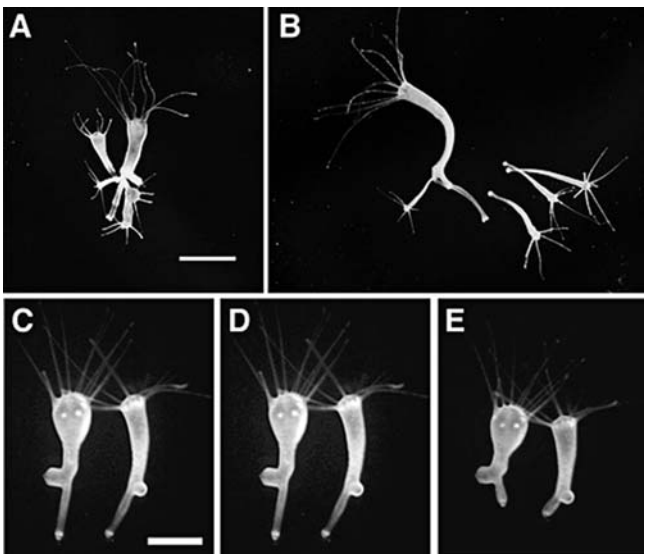


Fig. 21.4. The myoactive effects of neuropeptides Hym-248 and Hym-176 on epithelial *Hydra.* Panels A, C: Morphology of epithelial *Hydra* prior to application of the peptides. Panel B: Bud detachment and elongation of the body column observed within 1 h of application of Hym-248 (10^{-6} M). Panels D, E: Contraction of the foot observed 0.5 and 1 h, respectively, after application of Hym-176 (10^{-6} M). Scale bars, 3 mm (Panel A) and 2 mm (Panel C).

elongation of the body column and bud detachment from the parent (**Fig. 21.4A, B**). The body elongation is caused by contraction of endodermal circular muscle fibres, while ectodermal muscles are relaxed. Bud detachment is caused

Table 21.1
Identification of peptides by LC-MALDI-MS/MS

Fraction No	Sequence	Precursor	Reference
3	ENDQLWFFLQPFKQS GTGFSTN		
4	QWFSGRFGLPNQ	Neuropeptide precursor	25
	QWLSGRFGLTN	Neuropeptide precursor	25
	EEVEMMHQIMP		
	GGAISGSVGGFLANS		
5	MQESPFCT		
	FDLFDRF		
	FPQSFLPRG	Neuropeptide precursor	14
	QWLGGRF	Neuropeptide precursor	25
	QWFNGRF	Neuropeptide precursor	25
6	QWLGGRF	Neuropeptide precursor	25
	QWFNGRF	Neuropeptide precursor	25
	KNSINRF		
	FDLFDRF		
	FPQSFLPRG	Neuropeptide precursor	14
	AAGCQLVATC		
	MVVEMSGLS		
	NVKSCMCKN		
	AVGLSRMLP		
	ASPPARSAPP		
	MALMSKGAFV		
7	QWFNGRF	Neuropeptide precursor	25
	PNWVIAIF		
	LQELGPRF		
8	QWLGGRF	Neuropeptide precursor	25
	QWFNGRF	Neuropeptide precursor	25
	ESIIGMLD		
	NKANPIQSS		
	MIIQVMMS		
	FPQSFLPRG	Neuropeptide precursor	14
	MKSGKNLIY		
9	NPIKMRKL		

by the contraction of ectodermal circular muscle in the basal disc, thereby constricting and detaching the buds. Hym-176 induces contraction only in the foot region (**Fig. 21.4C–E**) (Noro et al., unpublished result). The foot contraction is caused by contraction of ectodermal muscle fibres in the foot while endodermal muscle fibres remain relaxed.

31. A combined total of about 2000 MS/MS spectra were obtained from ten plates. **Table 21.1 s**hows some of our results obtained with LC-MALDI-TOF-MS/MS. Precursor peptides were identified with Mascot analysis. Some peptides were found across different fractions (e.g. Hym-355 in fractions 5, 6 and 8). About 40% of peptides obtained were neuropeptides (both known and novel) and the remaining ones have not been matched, mostly due to insufficient EST coverage of the whole transcriptome.

Acknowledgments

We appreciate invaluable contribution from the members of the *Hydra* Peptide Project: Y. Muneoka, T. Sugiyama, O. Koizumi, H. Shimizu, M. Hatta, S. Yum, N. Harafuji, E. Hayakawa, F., Morishita, O. Matsushima, Y. Kobayakawa, J. Lohmann, T. C. G. Bosch, C. H. David and H. R. Bode. This work was supported by grants from the Ministry of Education, Culture, Sport, Science and Technology (MEXT) of Japan (T.F.) and a Grant-in-Aid for Young Scientists (B) from MEXT (T. T.). T.F. was awarded the Mercator Professorship at University of Heidelberg from 2007 to 2008 by Deutsche Forschungsgemeinschaft (DFG). We appreciate help and suggestions from Prof. Thomas Holstein.

References

1. Takahashi, T., Muneoka, Y., Lohmann, Y., deHaro, M., Solleder, G., Bosch, T. C. G., et al. (1997) Systematic isolation of peptide signal molecules regulating development in Hydra: LWamide and PW families. *Proc. Natl. Acad. Sci. USA* **94**, 1241–1246.
2. Fujisawa, T. (2008) Hydra peptide project 1993–2007. *Dev. Growth Differ.* **50** (Suppl 1), S257–S268.
3. Schaller, H. C. and Bodenmueller, H. (1981) Isolation and amino acid sequence of a morphogenetic peptide from *Hydra*. *Proc. Natl. Acad. Aci. USA* **78**, 7000–7004.
4. Leitz, T., Morand, K., and Mann, M. (1994) Metamorphosin A: a novel peptide controlling development of the lower metazoan *Hydractinia echinata* (Coelenterata, Hydrozoa). *Dev. Biol.* **163**, 440–46.
5. Takahashi, T., Kobayakawa Y., Muneoka Y., Fujisawa Y., Mohri S., Hatta M., et al. (2003) Identification of a new member of the GLWamide peptide family: physiological activity and cellular localization in cnidarian polyps. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **135**(2), 309–24.
6. Leviev, I., Williamson, M., and Grimmelikhuijzen, C. J. P. (1997) Molecular cloning

of a preprohormone from *Hydra magnipapillata* containing multiple copies of Hydra-LWamide (Leu-Trp-NH_2) neuropeptides: Evidence for processing at Ser and Asn residues. *J. Neurochem.* **68**, 1319–25.
7. Hoffmeister, S. A. H. (1996) Isolation and characterization of two new morphologically active peptides from *Hydra vulgaris*. *Development* **122**, 1941–8.
8. Fujisawa, T. (2003) Hydra regeneration and epitheliopeptides. *Dev. Dyn.* **226**, 182–9.
9. Koizumi, O., Wilson, J. D., Grimmelikhuijzen, C. J. P., and Westfall, J. A. (1989) Ultrastructural localization of RFamide-like peptides in neuronal dense-cored vesicles in the peduncle of *Hydra*. *J. Exp. Zool.* **249**, 17–22.
10. Koizumi, O., Mizumoto, H., Sugiyama, T., and Bode, H. R. (1990) Nerve net formation in the primitive nervous system of *Hydra*—an overview. *Neurosci. Res. Suppl.* **13**, S165–70.
11. Moosler, A., Rinehart, K. L., and Grimmelikhuijzen, C. J. P. (1996) Isolation of four novel neuropeptides, the hydra-RFamides I–IV, from *Hydra magnipapillata*. *Biochem. Biophys. Res. Commun.* **229**, 596–602.
12. Campbell, R. D. (1976) Elimination of *Hydra* interstitial and nerve cells by means of colchicine. *J. Cell. Sci.* **21**, 1–13.
13. Nishimiya-Fujisawa, C. and Sugiyama, T. (1993) Genetic analysis of developmental mechanisms in *Hydra*. XX. Cloning of interstitial stem cells restricted to the sperm differentiation pathway in *Hydra magnipapillata*. *Dev. Biol.* **157**, 1–9.
14. Crick, F. (1970) Diffusion in embryogenesis. *Nature* **225**, 420–2.
15. Gierer, A. and Meinhardt, H. (1972) A theory of biological pattern formation. *Kybernetik* **12**, 30–9.
16. Liang, P. and Pardee, A. B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**, 967–71.
17. Yum, S., Takahashi, T., Koizumi, O., Ariura, Y., Kobayakawa, Y., Mohri, S., et al. (1998a) A novel neuropeptide, Hym-176, induces contraction of the ectodermal muscle in *Hydra*. *Biochem. Biophys. Res. Commun.* **248**, 584–90.
18. Grens, A., Shimizu, H., Hoffmeister, S., Bode, H. R., and Fujisawa, T. (1999) Pedibin/Hym-346 lowers positional value thereby enhancing foot formation in *Hydra*. *Development* **126**, 517–24.
19. Takahashi, T., Koizumi, O., Ariura, Y., Romanovitch, A., Bosch, T. C. G., Kobayakawa, Y., et al. (2000) A novel neuropeptide, Hym-355, positively regulates neuron differentiation in *Hydra*. *Development* **127**, 997–1005.
20. Harafuji, N., Takahashi, T., Hatta, M., Tezuka, H., Morishita, F., Matsushima, O., et al. (2001) Enhancement of foot formation in *Hydra* by a novel epitheliopeptide, Hym-323. *Development* **128**, 437–46.
21. Takahashi, T., Hatta, M., Yum, S., Gee, L., Ohtani, M., Fujisawa, T., et al. (2005) Hym-301, a novel peptide, regulates the number of tentacles formed in Hydra. *Development* **132**, 2225–34.
22. Hayakawa, E., Takahashi, T., Nishimiya-Fujisawa, C., and Fujisawa, T. (2007) A novel neuropeptide (FRamide) family identified by a peptidomic approach in *Hydra magnipapillata*. *FEBS J.* **274**, 5438–48.
23. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156–9.
24. David, C. N. (1973) A quantitative method for maceration of *Hydra* tissue. *Wilhelm Roux's Arch. Entw. Mech. Org.* **171**, 259–68.
25. Takahashi, T., Koizumi, O., Hayakawa, E., Minobe, S., Suetsugu, R., Kobayakawa, Y., Bosch, T.C., David, C.N., Fujisawa, T. (2009) Further characterization of the PW peptide family that inhibits neuron differentiation in *Hydra*. *Dev Genes Evol.* **219**, 119–29.

Chapter 22

Immunochemical Methods for the Peptidomic Analysis of Tachykinin Peptides and Their Precursors

Nigel M. Page and Nicola J. Weston-Bell

Abstract

The tachykinins represent the largest known peptide family and are responsible for a range of pleiotropic functions in both vertebrates and invertebrates. Recent research has shown a diversity of mechanisms such as mRNA splicing, precursor processing and post-translation modification that can lead to a complex and continually expanding repertoire of tachykinin peptides. The peptidomic analysis of the tachykinins has been hindered by the lack of specific methodologies to capture, purify and characterise each tachykinin. This chapter summarises some of the methods that have been developed in order to further purify and characterise individual groups of tachykinin peptides from the peptidome.

Key words: Endokinin, hemokinin, tachykinin, antibody, immunoassay.

1. Introduction

The tachykinins represent one of the largest peptide families described in the animal kingdom with over 40 peptides so far identified (1). Traditionally classified as neurotransmitters, they include substance P (SP), neurokinin A (NKA) and neurokinin B (NKB) in mammals which are all short peptides of 10–11 amino acids in length (2). Moreover, each peptide is derived from a larger precursor by the cleavage of flanking dibasic residues and has an adjacent glycine at the C-terminus for amidation. By definition, they all share the same conserved hydrophobic C-terminal region, FXGLM-NH_2, where X is always a hydrophobic residue. This conserved region is crucial to the activation of each of the known mammalian tachykinin receptors, NK_1, NK_2 and NK_3 (3).

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_22,

More recently, we have reported the identification of previously uncharacterised members of the tachykinin family encoded on the human *TAC4* gene that we have named the endokinins (EK) A-D (3). An analogue of the *TAC4* gene also exists in the rat and is predicted to encode the single 10 amino acid residue tachykinin, hemokinin-1 (HK-1) (4). To date, at least 14 different tachykinins including tachykinin gene-related peptides have been extracted from the mammalian peptidome (5). However, the challenge remains on how best to purify and identify further sub-classes of the tachykinin peptides from the complexity of the peptidome. In order to do this we have devised a series of capture antibodies and describe protocols for the raising and purifying of anti-tachykinin antibodies and the development of a two-site immunoassay to aid in the peptidomic analysis of existing and potential novel members of the tachykinin peptide family.

2. Materials

2.1. Raising Antibodies

1. Synthetic peptides: Endokinin A/B decapeptide (GKASQFFGLM-NH_2), hemokinin-1 (SRTRQFC) and the N-terminal TAC4 precursor sequence (AETWEGAGPSIQLQLQEVK, 32-50 of γTAC4) (3).
2. Avian purified protein derivative (PPD) of tuberculin (CVL, Addlestone, UK).
3. γ-maleimidobutyric acid N-hydroxysuccinimide ester (GMBS).
4. Gluteraldehyde (grade II 25% aqueous solution).
5. 0.1 M $NaHCO_3$.
6. 0.9% (w/v) NaCl.
7. Dimethyl sulfoxide (DMSO).
8. Sephadex G-25-50: Soak 1.2 g of Sephadex G-25-50 in 0.9% (w/v) NaCl, pH 8 (use Tris) for 30 min. Pack into an Econo-Column (1.5×10 cm, 18 ml) (Bio-RadLaboratories) and equilibrate with 50 ml of 0.9% (w/v) NaCl pH 8.
9. pH Litmus strips.
10. 0.2 M sodium metabisulphite.
11. Dry ice and isopentane.
12. Freund's complete adjuvant.
13. Freund's incomplete adjuvant.
14. Four sheep per immunogen.

2.2. Antisera Collection

1. 5 μm cellulose nitrate filter (Sartorius AG, Goettingen, Germany).
2. 50-ml Nunc tubes (Nalgene Nunc International, Rochester, USA).

2.3. Preparing Peptide-Coated Exiqon Peptide Immobilizer™ Plates

1. Synthetic peptides: Endokinin A/B decapeptide (GKASQFFGLM-NH_2), hemokinin-1 (SRTRQFC) and the N-terminal TAC4 precursor sequence (AETWEGAGPSIQLQLQEVK, 32-50 of γTAC4).
2. Bicarbonate buffer: 0.1 M Sodium bicarbonate ($Na_2CO_3.10H_20$), 0.1 M sodium hydrogen carbonate ($NaHCO_3$) (*see* **Note 1**).
3. Exiqon Peptide Immobilizer™ plates (Exiqon A/S, Denmark).
4. Wash buffer for plates: 0.9% (w/v) NaCl, 0.01% (v/v) Triton X-100.

2.4. Testing Serum for Peptide Antibody Production

1. Sheep antiserum.
2. Albumin buffer (AB): Dissolve either 3.56 g disodium hydrogen orthophosphate dehydrate ($Na_2HPO_4{\cdot}2H_2O$) or 2.84 g sodium dihydrogenphosphate (Na_2HPO_4) with 0.74 g sodium dihydrogen phosphate dihydrate ($NaH_2PO_4{\cdot}2H_2O$) in 50 ml distilled water. Adjust pH to 7.4 with Tris. Make this solution up to 500 ml with distilled water. Carefully add 2.5 g of protease-free bovine serum albumin onto the surface of the liquid then allow to dissolve without agitation. Add 0.5 g of sodium azide. AB buffer can be stored for up to 1 week at 4°C. For convenience a 10X (0.5 M) phosphate buffer solution (0.4 M $Na_2HPO_4.2H_2O$ or Na_2HPO_4, 0.1 M $NaH_2PO_4.2H_2O$, 0.1% NaN_3) can be kept as stock and then diluted prior to addition of BSA. 0.5 M phosphate buffer (PB): Dissolve either 35.6 g $Na_2HPO_4.2H_2O$ or 28.4 g Na_2HPO_4 with 7.4 g $NaH_2PO_4.2H_2O$ in 50 ml distilled water. Adjust pH to 7.4 using Tris. Make up to 500 ml with distilled water. Add 0.5 g sodium azide. Phosphate buffer will keep for 3–4 months at 4°C. When AB is required simply dilute 0.5 M phosphate buffer 1:10 in distilled water (50 ml phosphate buffer and 450 ml water), add 2.5 g protease-free bovine serum albumin to the surface as above. Once dissolved add a further 0.5 g sodium azide. Prepared AB buffer can be stored for up to 1 week at 4°C.
3. Exiqon Peptide Immobilizer™ plates (Exiqon A/S, Denmark)

4. Wash buffer for plates: 0.9% (w/v) NaCl, 0.01% (v/v) Triton X-100.
5. Anti-sheep IgG (whole molecule) alkaline phosphatase conjugate was prepared as a 1:5000 dilution in 20 ml AB buffer. It is made fresh for each assay and kept at 2–8°C.
6. SIGMA*FAST*™ *p*-nitrophenyl phosphate tablets (Sigma-Aldrich); take one *p*-nitrophenyl (pNPP) tablet and one Tris buffer tablet (Sigma-Aldrich) and dissolve by vortexing in 20 ml distilled water to make the pNPP substrate solution. This gives a solution containing 1.0 mg/ml pNPP, 0.2 M Tris buffer and 5 mM magnesium chloride. The mixture should be used immediately.
7. 2 M NaOH.

2.5. Affinity Purification of Antibodies

1. 300 mg of cyanogen bromide-activated Sepharose 4B resin (Sigma-Aldrich) washed and swelled in 50 ml 1 mM HCl (pH4-4.5) for 10 min
2. 1 M ethanolamine HCl, pH 8.5: 62.5 μl ethanolamine made up to 1 ml with 1 M HCl
3. Saline, azide, EDTA buffer: For 1 l add 9 g NaCl, 1 g NaN_3, 0.372 g ethylenediamine disodium salt. Adjust pH to 7.5 by the addition of Tris crystals. Add 100 μl of triton X-100.
4. Salt washes: High salt: 0.5 M sodium acetate, 20% acetonitrile, use at natural pH. Low salt: 0.05 M sodium acetate, 20% acetonitrile, pH 6. Use formic acid to pH. Elution buffer: 0.05 M sodium acetate, 20% acetonitrile, pH 4. Use formic acid to pH.
5. 0.5 mg of the appropriate immunogenic peptide.

2.6. Biotinylation of Antipeptide Antibodies

1. Biotinamidocaproate *N*-hydroxy succinimide.
2. Dimethyl sulfoxide.
3. 0.1 M $NaHCO_3$.
4. Sephadex G-25-50 (bed volume, 5 ml, 5 g). To prepare, 1.2 g of Sephadex G-25-50 is weighed out and allowed to soak in 0.9% (w/v) NaCl, pH 8 (with tris) for 30 min. The Sephadex is then packed into an Econo-Column and the column pre-equilibrated with 50 ml of 0.9% (w/v) NaCl, pH 8.
5. Coomassie blue G-250-based protein assay reagent (Pierce, Rockford, IL).

2.7. Testing of Antipeptide Biotinylated Antibodies

1. Bicarbonate buffer: 0.1 M Na_2CO_3, 0.1 M $NaHCO_3$.
2. Plate wash buffer: 0.9% (w/v) NaCl, 0.01% (v/v) Triton X-100.

3. 80% Streptavidin horseradish peroxidase (Immunodiagnostic Systems Limited, UK).
4. 3,3′,5,5′-tetramethylbenzidine (TMB) substrate (Europa Bioproducts, UK).
5. 0.5 M HCl

2.8. Two-Site ELISA Development

1. Glaze buffer: 1.05 M 3-[*N*-Morpholino]propane-sulfonic acid, MOPS, 0.95 M 3-[*N*-Morpholino]propane-sulfonic acid mono sodium salt, 0.15 M sucrose, 0.4% (w/v) bovine albumin serum, BSA, pH 6.8–7.4.
2. Plate wash buffer: 0.9% (w/v) NaCl, 0.01% (v/v) Triton X-100.

2.9. Peptide Extraction

1. Peptide extraction buffer: 1 M HCl containing 5% (v/v) formic acid, 1% (w/v) NaCl, 1% (v/v) trifluoroacetic acid. Store at room temperature.
2. Phosphate-buffered saline. Dissolve one tablet in 200 ml of distilled water to yield 0.01 M phosphate buffer, 0.0027 M potassium chloride and 0.137 M sodium chloride, pH 7.4.
3. Methanol.
4. 0.1% (v/v) trifluoroacetic acid in water.
5. 60% acetonitrile in 0.1 (v/v) trifluoroacetic acid in water.
6. 5 mg/ml of mannitol. Store at –20°C.

2.10. Size Exclusion Chromatography

1. 1 mM HCl.
2. 0.1% (v/v) trifluoroacetic acid, 20% (v/v) acetonitrile.
3. 5 mg/ml mannitol.

3. Methods

3.1. Raising Antibodies

1.Couple selected antigenic peptide to 4 mg avian purified protein derivative (PPD) of tuberculin (CVL, Addlestone, UK) (*see* **Note 2**). Cysteine-containing peptides such as hemokinin-1 (SRTRQFC) can be coupled to PPD via γ-maleimidobutyric acid *N*-hydroxysuccinimide ester (GMBS) (Sigma-Aldrich). Alternatively, for non-cysteine-containing peptides such as endokinin A/B decapeptide (GKASQFFGLM-NH_2) and the N-terminal TAC4 precursor sequence (AETWEGAGPSIQLQLQEVK, 32-50 of γTAC4) coupling to PPD can be achieved using gluteraldehyde (grade II 25% aqueous solution).

3.1.1. Immunogen Coupling via GMBS

1. Dissolve 4 mg PPD in 0.5 ml 0.1 M $NaHCO_3$ in a 1.5-ml Eppendorf tube on an end-over-end mixer for 30 min. Add to this 3 mg GMBS dissolved in 10 μl dimethyl sulfoxide. Incubate at room temperature in the dark for 30 min. This will generate PPD-GMB (the PPD replaces the *N*-hydroxysuccinimide part of the ester), but some unbound GMBS may remain in the solution.
2. Meanwhile equilibrate previously swollen Sephadex G-25 in an Econo-column (Bio-Rad Laboratories) or similar with ~5 ml (one column volume) 0.1 M $NaHCO_3$ and dissolve 2 mg of hemokinin-1 (SRTRQFC) peptide in sufficient (300–500 μl) 0.1 M $NaHCO_3$ that the pH is greater than 9.0 (*see* **Note 3**).
3. Remove any un-reacted GMBS from the reaction in Step 1 in **Section 3.1.1** by passing it through the equilibrated Sephadex G-25. As the solution passes through the column un-reacted GMBS will be retained. Watch the solution pass through the column and collect the brown-coloured PPD-GMB fraction, as it emerges, into the pre-dissolved hemokinin-1 peptide. Incubate this PPD-GMB and peptide solution at room temperature in the dark for 30 min. This will generate PPD-peptide (the peptide replaces the GMB) via the sulfhydryl group on the peptide's cysteine residue. Un-reacted peptide may remain in the solution.
4. Meanwhile equilibrate the Sephadex G-25 column with ~5 ml (one column volume) of an 80:20 solution of 0.9% (v/v) NaCl, 0.1 M $NaHCO_3$.
5. Remove any un-reacted peptide from the reaction incubated in Step 3 in **Section 3.1.1** by passing through the NaCl, $NaHCO_3$ equilibrated Sephadex and collecting the brown-coloured PPD-hemokinin-1 fraction into a clean Eppendorf.
6. Dilute gluteraldehyde-coupled immunogens to final volume of 6.25 ml in 0.9% (v/v) NaCl. Aliquot in five 1.25 ml volumes in 5-ml bijoux tubes and flash freeze using dry ice and isopentane. Store at –20°C. This provides sufficient immunogen for four sheep for five months at 0.1 mg peptide per sheep per month.
7. For the first immunisation thaw one aliquot of immunogen and add 2.75 ml Freund's complete adjuvant. Vortex thoroughly (*see* **Note 4**). Use immediately to inject four sheep with 1 ml each of the immunogen, complete adjuvant emulsion.
8. Four weeks later thaw a second 1.25-ml aliquot of immunogen and add 2.75 ml Freund's incomplete adju-

vant. Vortex thoroughly, as above, and inject 1 ml immunogen, incomplete adjuvant emulsion into each of the same four sheep immediately.

9. The sheep can be bled (400–500 ml) 10 day post-injection into glass bottles.
10. Continue with 4-weekly injections of 1.25 ml immunogen emulsified with 2.75 ml Freund's incomplete adjuvant, followed by bleeding 10 days later for as long as is necessary, within ethical constraints, to obtain sufficient purified antibody.

3.1.2. Immunogen Coupling via Gluteraldehyde

1. Dissolve 4 mg PPD in 0.3 ml 0.1 M $NaHCO_3$ in a 1.5-ml Eppendorf tube on an end-over-end mixer for 30 min. Add to this 2 mg of endokinin A/B decapeptide (GKASQFFGLM-NH2) or N-Terminal TAC4 precursor sequence (AETWEGAGPSIQLQLQEVK, 32–50 of γTAC4) dissolved in 0.25 ml 0.1 M $NaHCO_3$.
2. Add 10 μl gluteraldehyde and incubate on an end-over-end mixer at 4°C in the dark for 1 h.
3. Inactivate the gluteraldehyde by adding 100 μl 0.2 M sodium metabisulphite and incubate for 2 h at room temperature.
4. Dilute GMBS-coupled immunogens to final volume of 6.25 ml in 0.9% (v/v) NaCl. Aliquot in five 1.25 ml volumes in 5-ml bijoux tubes and flash freeze using dry ice and isopentane. Store at –20°C. This provides sufficient immunogen for four sheep for 5 months at 0.1 mg peptide per sheep per month.
5. The next steps are performed exactly the same as for immunogen coupling via GMBS by following Steps 7–10 in **Section 3.1.1**.

3.2. Antisera Collection

1. Collect blood from sheep (typically 400–500 ml per bleed) and allow to clot in glass bottles overnight at room temperature.
2. Collect antiserum by pouring into a glass beaker whilst holding the clot in place using a spatula (*see* **Note 5**).
3. Pour serum into 250-ml centrifuge tubes and centrifuge at 3500×*g* for 20–30 min.
4. Filter the antiserum through a 5-μm cellulose nitrate filter (Sartorius AG, Goettingen, Germany) and collect into 50-ml Nunc tubes (Nalgene Nunc International, Rochester, USA).
5. A 1-ml aliquot should also be taken to test the antiserum for peptide antibody production.
6. Store collected antisera at –20°C.

3.3. Preparing Peptide-Coated Exiqon Peptide Immobilizer™ Plates

1. Plates are generally coated with 100–200 ng of synthetic peptide in each well. Therefore, to coat one 96-well Exiqon Peptide Immobilizer™ plate 10–20 μg of peptide is required. It is a good idea at this stage to determine that each peptide can be dissolved in the bicarbonate buffer (*see* **Note 6**).
2. To coat one 96-well Exiqon Peptide Immobilizer™ plate with 200 ng peptide per well, weigh out 20 μg of peptide on an analytical balance (e.g. a Cahn C-31 microbalance).
3. Dissolve the peptide in 100 μl of bicarbonate buffer and then bring the total volume of bicarbonate buffer to 9.9 ml.
4. Dispense 100 μl of this solution into each well of the Exiqon plate; a multichannel pipettor is recommended for this purpose.
5. After gentle agitation, the Exiqon plate should be left in contact with the peptide solution for at least 2 h at room temperature; however, it is recommended that the plate be left to incubate overnight at 4°C.
6. Wash each well of the Exiqon plate three times with 250 μl of plate wash buffer to thoroughly remove the diluted peptide solution.
7. Plates can be stored at 4°C for up to 4 weeks (*see* **Note 7**).

3.4. Testing Serum for Peptide Antibody Production

1. All materials required for the testing of the serum are made ready.
2. The 1 ml frozen serum aliquot to be tested is slowly thawed by leaving the sample at room temperature. The sample should thaw in about 10 min (*see* **Note 8**).
3. On thawing the sample needs gentle agitation to ensure it forms a homogeneous mixture. If there are any visible particles in the thawed serum, the sample should be centrifuged at 10,000×*g* for 30 s and the supernatant used in the next steps.
4. Normally a dilution range from 1:1000 to 1:8000 is sufficient for testing the production of antibodies. Take 100 μl of the thawed serum and add 900 μl of AB buffer. Gently mix by pipetting. This produces a 1:10 dilution of serum. Take 20 μl of the 1:10 diluted serum and add 1980 μl AB buffer to produce a 1:1000 dilution of serum. Subsequent 1:2000, 1:4000 and 1:8000 dilutions are then prepared.
5. Prepare an identical dilution series, this time using normal sheep serum as a control.
6. Once the dilutions of serum have been prepared 100 μl of the diluted 1:1000 to 1:8000 sera can be added to each

well, in duplicate, of an Exiqon plate in which the wells have been pre-coated with 100–200 ng of the original peptide antigen (*see* **Section 3.3**) (*see* **Note 9**).

7. The diluted antisera should be left in each well for at least 3 h at room temperature or if necessary can be left overnight at 4°C.
8. Wash each well of the Exiqon plate three times with 250 μl of plate wash buffer. Either a multi-channel pipette or automatic plate washer can be used for this step.
9. Add 100 μl of prepared anti-sheep IgG (whole molecule) alkaline phosphatase conjugate to each treated Exiqon well. Leave for at least 2 h at room temperature or if necessary can be left overnight at 4°C.
10. Wash each well of the Exiqon plate three times with 250 μl of plate wash buffer to thoroughly remove excess conjugate.
11. Add 200 μl of pNPP substrate solution to each well. Incubate the plate in the dark for 10–15 min at room temperature or until a yellow colour develops.
12. Stop the reaction by adding 50 μl 2 M NaOH straight to the 200 μl of reaction mixture in each well. Do not wash in between.
13. Read the absorbance for the stopped reactions at 405 and 600 nm. Readings at 600 nm should be subtracted as background (*see* **Fig. 22.1a**).

3.5. Affinity Purification of Antibodies

1. Wash and swell 300 mg of cyanogen bromide-activated Sepharose 4B resin (Sigma-Aldrich) in 50 ml 1 mM HCl (pH 4–4.5) for 10 min (*see* **Note 10**).
2. Wash with 100 ml 1 mM HCl through a suction funnel by applying gentle suction to remove supernatant (*see* **Note 11**).
3. The resin should dry on a suction funnel until appearing cracked.
4. Use a spatula to transfer the cracked cyanogen bromide-activated Sepharose 4B resin to a clean LP4 tube (Luckhams, England).
5. Dissolve 1 mg of peptide to be coupled in 300 μl of 0.1 M $NaHCO_3$ pH 8.5. The capacity of the resin is approximately 5–10 mg of peptide per ml of resin (*see* **Note 12**).
6. Add the 300 μl of dissolved peptide to the cracked cyanogen bromide-activated Sepharose 4B resin in the LP4 tube. Add an additional 3 ml of 0.1 M NaHCO3, pH 8.5 (*see* **Note 13**).

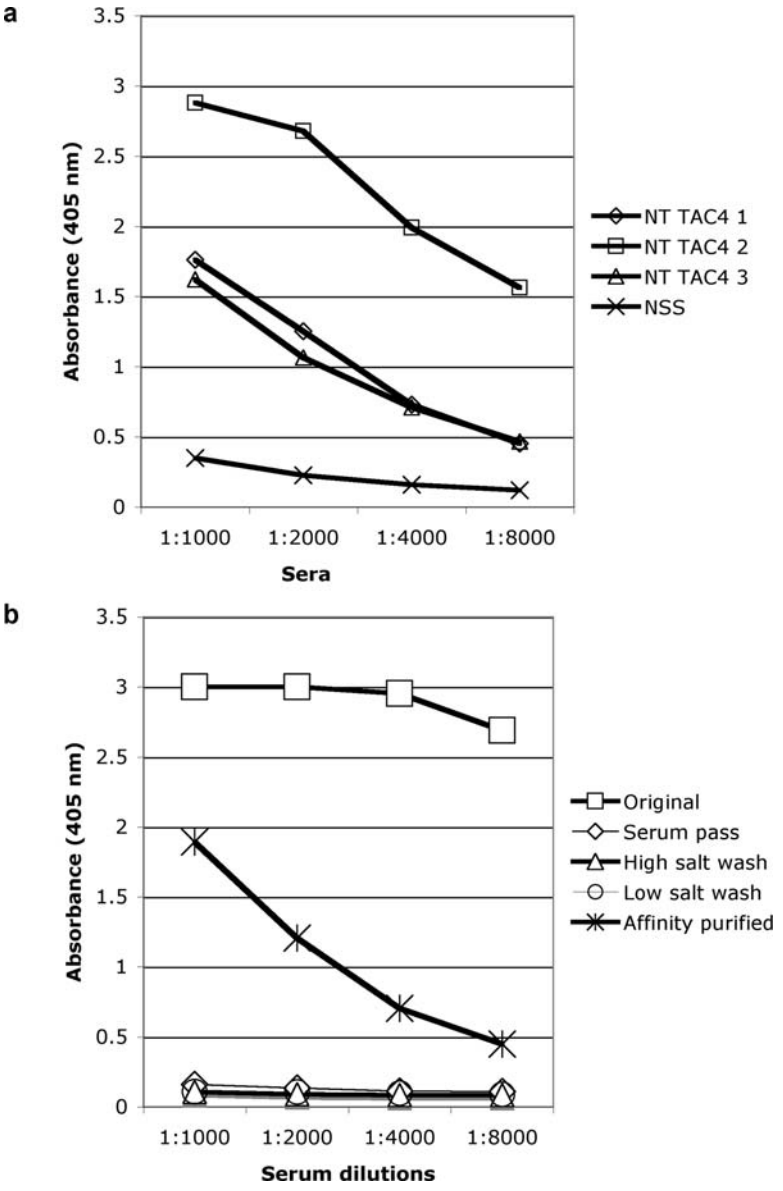


Fig. 22.1. Testing sheep antisera for peptide antibody production. (**a**) Three antisera from three individual sheep were tested in a dilution range from 1:1000 to 1:8000 for antibodies raised to the N-Terminal TAC4 precursor sequence (AETWEGAGPSIQLQLQEVK, 32-50 of γTAC4). The graph demonstrates clear differences in the antibody titre from each sheep. A control for normal sheep serum is also included. (**b**) Affinity purification of antipeptide antibodies involved several steps. The purification of one of the N-Terminal TAC4 precursor sequence antisera is monitored here. It is important to monitor progress at each stage to ensure that optimal capture of antibody is attained. Here, the affinity-purified antibody titre is compared to that of the original serum and to the potential presence of antibody in each of the high- and low-salt washes. A serum pass is included to show the remaining antibody titre that was left in the buffer that passed immediately through the Econo-column.

7. Cap the LP4 tube and mix protein with resin for at least 2–3 h using an end-over-end mixer (*see* **Note 14**).
8. Add 300 μl of 1 M ethanolamine HCl, pH 8.5 and leave overnight mixing end-over-end at room temperature.
9. Unreacted ligand is washed away by first pouring the peptide-coupled cyanogen bromide-activated Sepharose 4B resin into an Econo-Column. The column is then cleaned and equilibrated by passing through 300 ml of saline, azide, EDTA buffer. This is set up by establishing a gradient-controlled reservoir of buffer where the column is fed from the buffer tank by a fine capillary tube.
10. Prepare 100 ml of antiserum for affinity purification. First filter the antiserum through a double filter. On the top is placed a glass microfibre disc, grade GMF3, 25 mm (Sartorius) and on the bottom is placed a cellulose nitrate filter, 5 μm (Sartorius). The filter is attached to a 50-ml syringe so that the sample passes through the glass microfibre disc first (*see* **Note 15**).
11. To 100 ml of filtered antiserum add 900 ml of the saline, azide, EDTA buffer. This solution is then used to establish a gradient-controlled reservoir to feed the Econo-Column with the peptide-coupled cyanogen bromide-activated Sepharose 4B resin. Antibody with affinity for the presented peptide will be bound, while unreactive antibody will pass through the column. The diluted antiserum can be left to pass through the Econo-Column overnight at room temperature. To prevent the Econo-Column from running dry when left overnight part of the fine capillary tubing is looped down to the mid-height of the Econo-Column.
12. Once the diluted antiserum has passed through the Econo-Column add a few ml of the saline, azide, EDTA buffer to the column and gently agitate the Sepharose 4B resin using a glass pipette. The Sepharose 4B resin is now washed and equilibrated by passing at least 100 ml of the saline, azide, EDTA buffer through the Econo-Column. Drain the column of saline, azide, EDTA buffer, but make sure that the Sepharose 4B resin remains moist.
13. Affinity-purified antibodies are removed from the Econo-Column by applying a series of salt washes. First, the application of 10 ml of a high salt wash (0.5 M sodium acetate, 20% acetonitrile) which is left to drip through the Econo-Column, then the application of 10 ml of a low salt wash (0.05 M sodium acetate, 20% acetonitrile, pH 6) which again is left to drip through the Econo-Column. The eluate in each case should be kept for the purpose of

monitoring the pH after each wash. The final elution of the affinity-purified antibody is performed with an elution buffer (0.05 M sodium acetate, 20% acetonitrile, pH 4). This time instead of applying 10 ml of elution buffer all at once, elution is performed using 1 ml at a time. Each eluted 1 ml is individually collected and a small aliquot tested for pH on Litmus paper. To the remainder of the eluant gently mix in 200 µl of saturated $NaHCO_3$ solution and leave until the solution stops fizzing.

14. 10 µl of each eluted fraction is then placed in a microtitre plate well and 200 µl of Coomassie blue G-250-based protein assay reagent (Pierce, Rockford, IL) added. Read the absorbance from a plate reader set to 600 nm. From the Coomassie blue G-250-based protein assay reagent results pool the positively staining elution fractions (*see* **Note 16**) (*see* **Fig. 22.2a**).
15. Dilute the pooled samples 1 in 5 by gently mixing in 0.1 M $NaHCO_3$ and now nitrogen blows down the diluted pooled antibody sample to remove the acetonitrile for 15–20 min (*see* **Note** 17).
16. Take 500 µl of the diluted pooled antibody sample, following removal of acetonitrile, and add to a quartz cuvette and measure the absorbance at 280 nm using a UV spectrometer. Read against a blank quartz cuvette containing 500 µl of water. For calculating the amount of IgG the following formula can be used: 1 absorbance unit = 1.4 mg/ml IgG (*see* **Note** 18).
17. Store the purified antibody in 150–200-µl aliquots in plastic Eppendorf tubes by freezing rapidly using dry ice with isopentane. Store at –70°C (*see* **Note** 19).
18. To re-use or store the Econo-Column add 10 ml 1% (v/v) acetic acid to the peptide-coupled cyanogen bromide-activated Sepharose 4B resin and gently mix by pipetting. Then wash column with 200–300 ml of the saline, azide, EDTA buffer. The column is now ready to be re-used or stored at 4°C.

3.6. Biotinylation of Antipeptide Antibodies

1. Weigh out 1.5 mg biotinamidocaproate *N*-hydroxy succinimide and dissolve in 500 µl of dimethyl sulfoxide.
2. Take 50 µl of this solution and add to 0.5 mg affinity-purified antibody in 950 µl of 0.1 M $NaHCO_3$ (*see* **Section 3.5**). Leave to incubate at room temperature for 2 h.
3. Biotinylated antibodies are purified through Sephadex G-25-50 (bed volume, 5 ml, 5 g). To prepare, 1.2 g of Sephadex G-25-50 is weighed out and allowed to soak in 0.9% (w/v) NaCl, pH 8 (with tris) for 30 min. The Sephadex

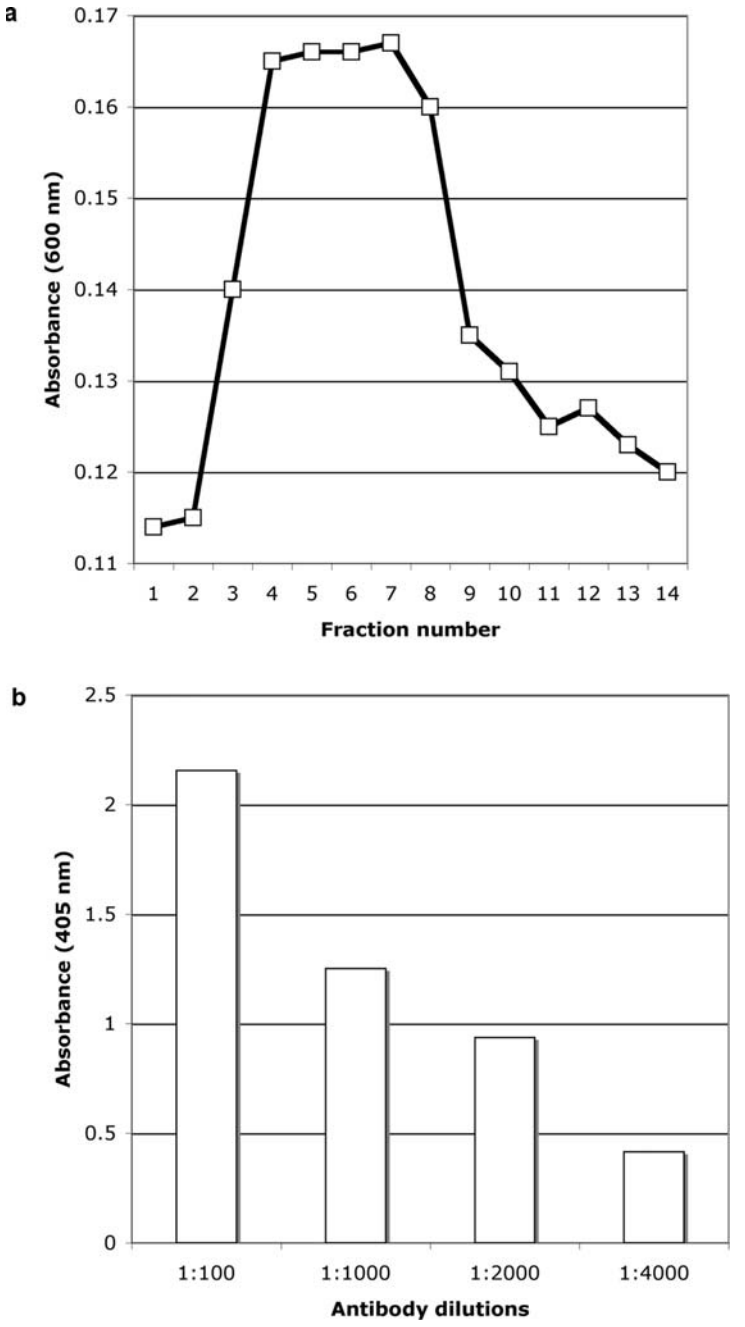


Fig. 22.2. Purification and testing of biotinylated antipeptide antibodies. (**a**) The results of purification through a Sephadex G-25 Econo-column from which consecutive 250 μl fractions are eluted separately. A Coomassie blue G-250-based assay reagent was used to monitor the presence of IgG protein following absorbance measurements at 600 nm. Most of the biotinylated antibody is found to elute in fractions 4–8. (**b**) The successful biotinylation of antibodies was confirmed by coating the biotinylated antibodies onto Exiqon Peptide Immobilizer™ plates in a serial dilution from 1:100 to 1:4000. The amount of antibody was determined following the addition of 80% streptavidin horseradish peroxidase and TMB substrate.

is then packed into an Econo-Column and the column pre-equilibrated with 50 ml of 0.9% (w/v) NaCl pH 8.

4. Add the biotinylated peptide reaction mixture (1 ml) to the column. Collect the first 2 ml in separate 1.5-ml Eppendorf tubes; after this collect 250-μl aliquots in individual 0.5-ml Eppendorf tubes. Check 20 μl of each fraction for protein content by combining with 50 μl Coomassie blue G-250-based protein assay reagent and reading the absorbance at 600 nm on a plate reader (*see* **Fig. 22.2a**).
5. Calculate the total amount of protein in each fraction and pool those that have reasonable protein concentrations.
6. Store the purified biotinylated antibody in 15-μg aliquots in plastic 0.5-ml Eppendorf tubes. Freeze rapidly using dry ice with isopentane. Store at –70°C.

3.7. Testing of Antipeptide Biotinylated Antibodies

1. The successful biotinylation of antibodies is confirmed by coating the biotinylated antibodies on Exiqon plates. This is performed in a similar manner as when peptides are coated on Exiqon plates (*see* **Section 3.3**).
2. To coat Exiqon plate wells with the biotinylated antibodies to be tested prepare a dilution curve of 1:100, 1:1000, 1:2000, 1:4000 dilutions in duplicate in 0.1 M Na_2CO_3, 0.1 M $NaHCO_3$.
3. Dispense 100 μl of each of these dilutions into each well of an Exiqon plate; a multichannel pipettor is recommended for this purpose. After gentle agitation of the plate, the antibody solution should be left in contact with the plate overnight at 4°C. Wrap the plate in cling film to prevent dust entering the wells.
4. Wash each well of the Exiqon plate three times with 250 μl of plate wash buffer (0.9% (w/v) NaCl, 0.01% (v/v) Triton X-100) to remove the diluted antibody solution. If the plates are not to be used immediately they should be stored at 4°C and for up to 4 weeks.
5. Add 100 μl of 80% streptavidin horseradish peroxidase (Immunodiagnostic Systems Limited, UK) to each well to be tested and leave for 30 min at room temperature.
6. The wells are then washed in plate wash buffer three times before the addition of 200 μl of 3,3′,5,5′-tetramethylbenzidine (TMB) substrate (Europa Bioproducts, UK). The TMB substrate is left in each well for 5–10 min until a blue colour appears.
7. At the end of this period the reaction is terminated by the addition of 100 μl of 0.5 M HCl and the absorbance read at 450 nm and 600 nm. Readings at 600 nm should be subtracted as background (*see* **Fig. 22.2b**).

8. Biotinylated antibodies should also be tested to confirm that they can still recognise the original peptide antigen. For this a dilution curve of biotinylated antibodies (1:100, 1:1000, 1:2000 and 1:4000 in AB buffer) is set up by incubating the diluted antibody on peptide antigen-coated Exiqon plates (*see* **Section 3.3**) overnight at 4°C. After this period the Steps from 5 to 7 in **Section 3.7** can now be followed.

3.8. Two-Site ELISA Development

1. Coat Exiqon Peptide Immobilizer™ plates with 200 ng affinity-purified "capture" antibody (*see* **Section 3.5**) using the same protocol from Steps 2 to 4 in **Section 3.7** (*see* **Fig. 22.3**).

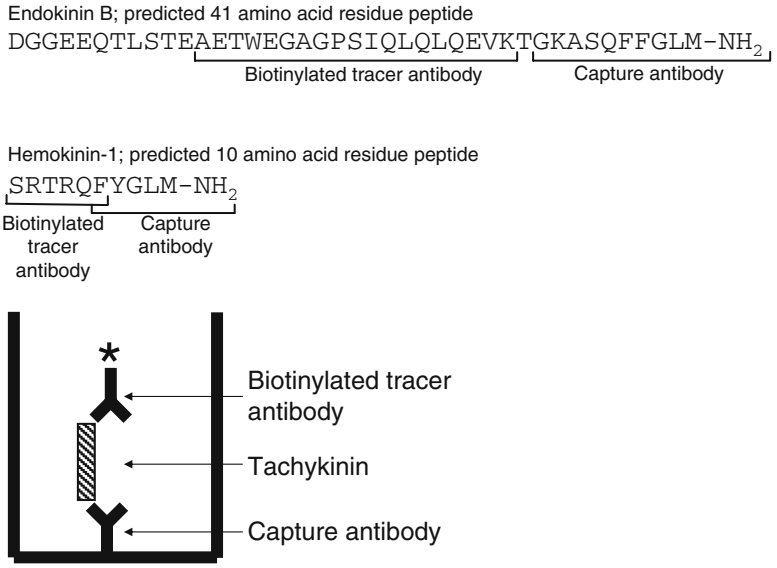


Fig. 22.3. The design strategy for the two-site immunoassays for endokinin B and hemokinin-1. The locations of the biotinylated tracer antibody and capture antibody are shown in each case.

2. Glaze each well with 100 μl glaze buffer (1.05 M 3-[*N*-Morpholino]propane-sulfonic acid, MOPS; 0.95 M 3-[*N*-Morpholino]propane-sulfonic acid mono sodium salt, 0.15 M sucrose, 0.4% (w/v) bovine albumin serum, BSA, pH 6.8–7.4) for at least 1 h at room temperature.
3. Remove the glaze buffer by inverting and flicking plate (do not pipette off) and allow the residue to air dry. If necessary, plates can be stored at 4°C for at least 4 months.
4. Dilute peptides across a concentration range normally 0–25 ng/ml in AB buffer to produce a standard curve by applying 100 μl to each well to be tested. Apply 100 μl of samples, in AB, to be tested alongside. Leave incubating overnight at 4°C.

5. Wash each well of the Exiqon plate three times with 250 μl of plate wash buffer (0.9% (w/v) NaCl, 0.01% (v/v) Triton X-100) to remove unbound peptide.
6. Add 100 μl antipeptide biotinylated "tracer" antibody (*see* **Section 3.7**) using different dilutions in AB buffer and incubate for 3–4 h at 4°C (*see* **Note 20**).
7. Wash each well of the Exiqon Peptide Immobilizer™ plate three times with 250 μl of plate wash buffer (0.9% (w/v) NaCl, 0.01% (v/v) Triton X-100) to thoroughly remove unbound "tracer" antibody solution.
8. Add 100 μl of 80% streptavidin horseradish peroxidase (Immunodiagnostic Systems Limited, UK) to each well to be tested and leave for 30 min at room temperature.
9. Wash the wells in plate wash buffer three times before adding 200 μl of 3,3′, 5,5′-tetramethylbenzidine (TMB) substrate (Europa Bioproducts, UK). Leave the TMB substrate in each well for 5–10 min until a blue colour appears.
10. At the end of this period terminate the reaction by the addition of 100 μl of 0.5 M HCl and read the absorbance at 450 and 600 nm. Readings at 600 nm should be subtracted as background.

3.9. Peptide Extraction

1. Wash extracted tissue in phosphate-buffered saline before cutting up into small cubes roughly 1 cm^2 using a scalpel blade.
2. Add 10 ml of peptide extraction buffer per 10 g of tissue (*see* **Note 21**) and homogenise the sample using a Waring laboratory blender or similar.
3. Centrifuge the homogenate at 3000×*g* for 20 min to remove debris and keep the supernatant. Discard the pellets.
4. Filter the supernatant through a 100-μm nylon mesh (Lockertex, Warrington, UK).
5. Transfer the filtered supernatant to centrifuge tubes suitable for high speeds and centrifuge at 48,000×*g* for 20 min. Retain supernatant and discard pellets. This extra centrifugation step prevents clogging of the Sep Pak 18C cartridges.
6. A 1 ml Sep Pak 18C cartridge should be primed for each tissue sample to be used. Using a 10-ml syringe, push 10 ml methanol through the cartridge, then use a 5-ml syringe to push through 5 ml 0.1% (v/v) trifluoroacetic acid.
7. Next apply 50 ml of the tissue extract supernatant to the cartridge using a 50-ml syringe. Optimum flow rate is

1 ml/min. Increased loading efficiency can be achieved by passing the same 50 ml of supernatant through the cartridge twice.

8. Equilibrate the cartridge with 2 ml of 0.1% (v/v) trifluoroacetic acid, discard the effluent. This is followed by 2 ml of 60% acetonitrile in 0.1% (v/v) trifluoroacetic acid. Collect the elutant into 10-ml polypropylene tubes containing 50 μl of 5 mg/ml of mannitol.
9. A SpeedVac is used to remove the organic solvents and leave a dried precipitate. A high-heat setting can be used for drying for short periods, but for drying overnight a low-heat setting should be chosen.
10. The precipitate can now be frozen at –20°C after capping the tube.
11. The Sep Pak 18C cartridges can be re-used after they have been washed. To wash the cartridge push through using a syringe 2 ml of methanol, then push through 2 ml of 0.1% (v/v) trifluoroacetic acid. The procedure can now be repeated from Step 6 in **Section 3.9**.

3.10. Size Exclusion Chromatography

1. Reconstitute the dried-down tissue precipitate (*see* Step 10 in **Section 3.9**) by adding 200 μl of 1 mM HCl and pipetting gently until the pellet dissolves.
2. Transfer to a fresh Eppendorf tube and mix on an end-over-end mixer for 30 min.
3. Heat shock the mixture by heating to 80°C for 5 min (*see* **Note 22**).
4. Centrifuge at 13,000×*g* for 5 min, remove and keep the supernatant and discard any pellet.
5. A Superdex Peptide® HR 10/30 column (Amersham Pharmacia Biotech, Uppsala, Sweden) with fractionation range 100–7000 Da, or equivalent, should be pre-equilibrated with 0.1% (v/v) trifluoroacetic acid, 20% (v/v) acetonitrile.
6. Peptides are eluted with 0.1% (v/v) trifluoroacetic acid, 20% acetonitrile that is run for approximately 70 min.
7. The column should be run at a pressure of 200 psi and flow rate of 0.25 ml/min. Absorbance is monitored at 215 nm.
8. Upon loading of the sample, immediately begin collecting 0.5 ml (2 min) fractions into 1.5-ml Eppendorf tubes containing 10 μl of 5 mg/ml mannitol. Ideally collect 30–35 such fractions.

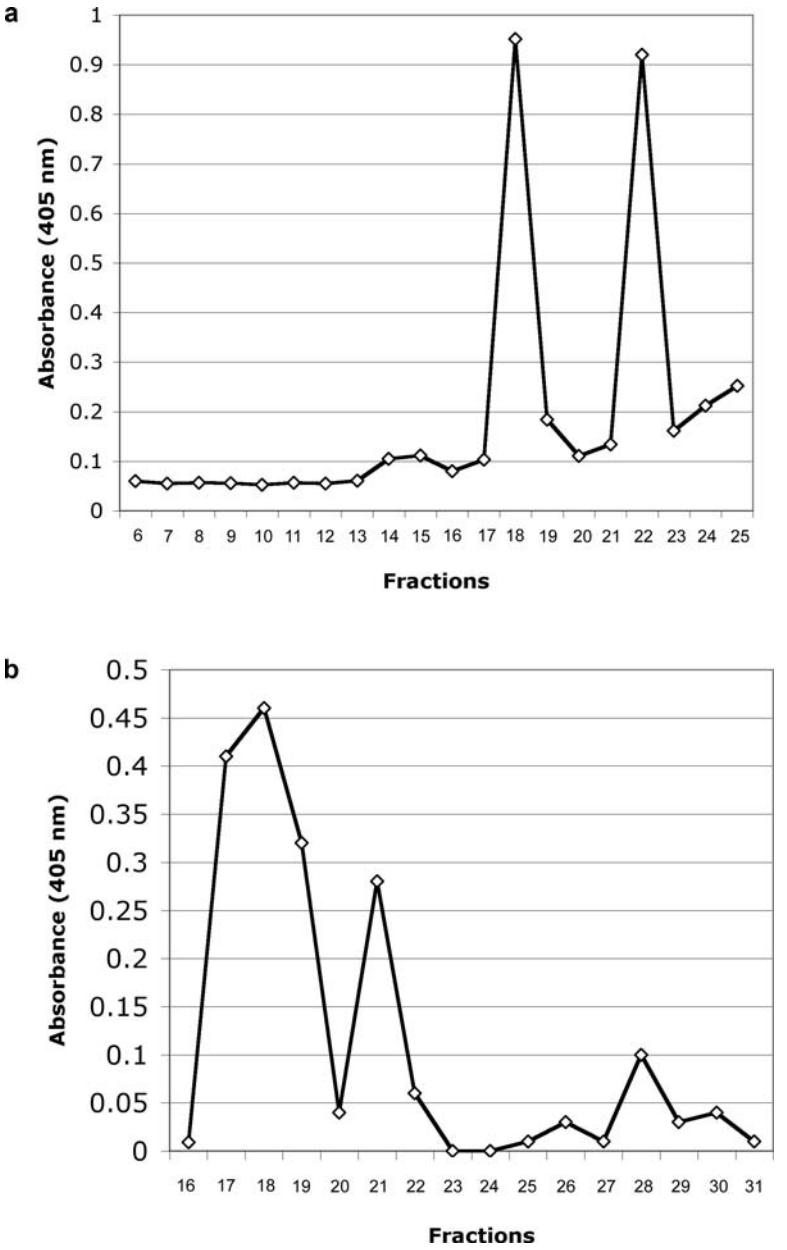


Fig. 22.4. Size exclusion chromatography monitored by the use of two-site immunoassays. (**a**) Extracts from human placenta were separated by size exclusion chromatography through a Superdex® Peptide HR 10/30 column (Amersham Biosciences AB, Uppsala, Sweden). The column was equilibrated with 0.1% (v/v) trifluoroacetic acid and peptides were eluted with 0.1% (v/v) trifluoroacetic acid 20% (v/v) acetonitrile that was run for approximately 70 min. The effluents were monitored at 215 nm in all chromatographic steps and fractions of 0.5 ml collected every 2 min. Fractions were evaporated to dryness in a SpeedVac and reconstituted in 100 μl of AB buffer before being monitored in a two-site ELISA consisting of a biotinylated "tracer" antibody to the N-Terminal TAC4 precursor sequence (AETWEGAGPSIQLQLQEVK, 32-50 of γTAC4) antibody and "capture" antibody designed to the sequence GKASQFFGLM-NH_2. The synthetic peptide TEAETWEGAGPSIQLQLQEVKTGKASQFFGLM-NH_2 was found to elute between fractions 22 and 23. (**b**) shows a similar example but this time using extracts from rat thymus and a two-site immunoassay consisting of biotinylated "tracer" antibody to the N-terminal sequence of hemokinin-1 (SRTRQF) and "capture" antibody designed to the sequence GKASQFFGLM-NH_2. The capture antibody designed to the sequence GKASQFFGLM-NH_2 was used as it was found to display strong cross-reactivity with not only itself but with rat hemokinin-1 (SRTRQFYGLM-NH_2). Synthetic hemokinin-1 was found to elute in fraction 28.

9. Dry the fractions for storage as Steps 9 and 10 in **Section 3.9**.
10. When required reconstitute each dried fraction in 100 μl AB for each detection assay required (*see* **Fig. 22.4a, b**).

4. Notes

1. There is no need to pH this solution.
2. Alternatives to PPD can be used, but the method may need to be adapted in line with the manufacturer's instructions.
3. This step is to prevent subsequent precipitation.
4. A complete emulsion must be achieved for injection.
5. To remove the clot from the bottle, dissolve overnight by adding NaOH pellets at room temperature.
6. If the peptides do not dissolve directly try to dissolve them first in a few microlitres of water or 10 mM HCl.
7. It is possible that plates can be stored for longer without any adverse effects for periods of 2–3 months, but they should be tested before use e.g. with peptide antiserum.
8. Thaw slowly as too fast may denature proteins.
9. It is worth testing with other similar antigens to check for cross-reactivity.
10. When coupling through a cysteine UltraLink® Iodoacetyl Gel can be used instead.
11. Washing is important to remove lactose, which will interfere with binding if present throughout coupling. The use of HCl in this process preserves the activity of the reactive groups.
12. Other buffers can be used, but avoid amine-containing buffers such as Trizma as these will react with the resin-binding sites. If a peptide has not dissolved in the basic 0.1 M $NaHCO_3$ then we have tried to use less peptide around 0.5 mg. We have also dissolved peptides in 1 mM HCl then added 0.1 M $NaHCO_3$ to achieve a pH between pH 7.5 and 8.0 for the initial mixing.
13. This step should be performed immediately as the reactive groups start to hydrolyse in the basic solution.
14. We recommend leaving for 5 h during the day at room temperature. Do not use a magnetic stirrer, as this will grind the resin.

15. It is likely that the filters will need to be changed after the first 50 ml as they could become clogged.
16. Coomassie staining allows the detection of protein (IgG). We normally find that the first five fractions contain the affinity-purified antibody.
17. The volume should decrease by approximately 1–2 ml.
18. This process typically yields affinity-purified antibodies of 0.2–1 mg/ml IgG.
19. It is worth checking the final affinity-purified antibody titres, along with aliquots collected from the passed through buffer, high-salt and low-salt washes (*see* **Fig. 22.1b**).
20. To optimise an assay both "capture" and "tracer" antibodies should be tested thoroughly at a variety of concentrations and incubation times.
21. The acid peptide extraction buffer was used to minimise peptidase activity and maximise the solubilisation of the peptides.
22. This step is included to denature proteases; on heating the solution may appear to go cloudy.

Acknowledgments

The Medical Research Council (UK) and the Biomedical and Pharmaceutical Sciences Research Group (Kingston University London) supported this work.

References

1. Severini, C., Improta, G., Falconieri-Erspamer, G., Salvadori, S., and Erspamer, V. (2002) The tachykinin peptide family. *Pharmacol. Rev.* **54**, 285–322.
2. Page, N.M. (2005) New challenges in the study of the mammalian tachykinins. *Peptides.* **26**, 1356–1368.
3. Page, N.M., Bell, N.J., Gardiner, S.M., Manyonda, I.T., Brayley, K.J., Strange, P.G., and Lowry, P.J. (2003) Characterization of the endokinins: human tachykinins with cardiovascular activity.. *Proc. Natl. Acad. Sci. USA.* **100**, 6245–6250.
4. Kurtz, M.M., Wang, R., Clements, M.K., Cascieri, M.A., Austin, C.P., Cunningham, B.R., Chicchi, G.G., and Liu, Q. (2002) Identification, localization and receptor characterization of novel mammalian substance P-like peptides. *Gene.* **296**, 205–212.
5. Page, N.M. (2004) Hemokinins and endokinins. *Cell Mol. Life Sci.* **61**, 1652–1663.

Chapter 23

Affinity Peptidomics: Peptide Selection and Affinity Capture on Hydrogels and Microarrays

Fan Zhang, Anna Dulneva, Julian Bailes, and Mikhail Soloviev

Abstract

Affinity peptidomics relies on the successfully proven approach used widely in mass-spectrometry-based protein analysis, where protein samples are proteolytically digested prior to the analysis. Unlike traditional proteomic analyses, affinity peptidomics employs affinity detection instead of, or in addition to, the mass-spectrometry detection. Affinity peptidomics, therefore, bridges the gap between protein microarrays and mass spectrometry and can be used for the detection, identification and quantification of endogenous or proteolytic peptides on microarrays and by MALDI-MS. Phage display technology is a widely applicable generic molecular display method suitable for studying protein–protein or protein–peptide interactions and the development of recombinant affinity reagents. Phage display complements the affinity peptidomics approach when the latter is used, e.g. to characterise a repertoire of antigenic determinants of polyclonal, monoclonal antibodies or other recombinantly obtained affinity reagents or in studying protein–protein interactions. 3D materials such as membrane-based porous substrates and acrylamide hydrogels provide convenient alternatives and are superior to many 2D surfaces in maintaining protein conformation and minimising non-specific interactions. Hydrogels have been found to be advantageous in performing antibody affinity assays and peptide-binding assays. Here we report a range of peptide selection and peptide-binding assays used for the detection, quantification or validation of peptide targets using array-based techniques and fluorescent or MS detection.

Key words: Affinity peptidomics, phage display, hydrogel, affinity assays, protein microarrays, MALDI-MS.

1. Introduction

Peptidomics focuses on the fractionation, purification and subsequent characterisation of naturally occurring peptides. Traditional peptidomic studies preferentially deal with peptides that display biological activity such as hormones, cytokines, toxins,

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_23,

neuropeptides and alike, and has by now yielded hundreds of novel peptides with exciting functional properties. Another branch of peptidomic research aims to use peptide-based assays to improve protein detection and quantitative analysis, and to bridge the gap between mass spectrometry (employed mostly for *peptide* analysis) and affinity-based techniques (traditionally limited to the antibody-based *protein* assays). Protein microarrays, first reported just over a decade ago (1–6), were a direct import of DNA technology and were expected to speed up technological development in many areas. But unlike oligonucleotides or DNAs, the antibodies are not so easy to manufacture, preserve, immobilise and assay in a quantitative manner and in highly parallel assays (7). Despite the availability of micro-manufacturing technologies, commercial antibody arrays do not contain more than just a few hundred antibodies, e.g. 224 and 725 antibody arrays (Panorama™ Ab Microarrays, Sigma-Aldrich) or 656 antibody arrays (Spring Bioscience, or Full Moon BioSystems). Unlike nucleic acid samples, protein samples do not make an easy target either. Proteins are susceptible to degradation, denaturation and, unlike complementary DNA strands, antibody–antigen pairs have a widely different range of affinities and specificities. The affinity peptidomics approach to protein arrays allows to at least partially resolve the aforementioned difficulties. In the affinity peptidomics assay the composition of a protein mixture is determined by directly assaying the peptides from crude tryptic or otherwise digested protein preparations, instead of assaying native protein preparations. This simplifies the assays, reduces protein (sample) heterogeneity and allows to simplify affinity reagent selection and preparation. Since the antibody could be easier and cheaper to develop against peptides rather than proteins, such anti-peptide antibodies make perfect capture reagents against relevant proteolytically derived peptides, allowing not only array-based quantitative fluorescence or ELISA-style assays, but also providing a direct link to the mass-spectrometry technologies (7–11). In this chapter we describe our approach to the selection of peptides suitable for use in Affinity peptidomics assays.

Phage display technology overcomes the need of hybridomas and animal immunisation to produce monoclonal antibodies, instead, antibody fragments can be expressed on phage surface (12–14) and their nucleotide and amino acid sequences can be easily determined (15–17). Phage libraries are usually constructed by modifying one of the genes, encoding phage surface proteins. Phage display is one of the simplest "display" technologies capable of encoding a large number of phage transformants (10^9) and suitable for the development of recombinant antibodies (8, 18, 19). However, phages are suited even better for displaying small peptides on their surface, allowing such uses as aptamer selection, antibody characterisation and the studies of protein–peptide

interactions (20–22). Phage display is therefore a technique suitable for use with both peptidomic systems, whether for studying naturally occurring or proteolytically obtained peptides. In this chapter we will exemplify the use of phage display approach for the development of peptide affinity reagents.

Traditionally, DNA and protein assays relied on planar surfaces for the capture reagent immobilisation, but increasingly the emphasis is on 3D surfaces, capable of increased protein loading, controlled orientation, protein stability during storage and the assay, and lacking the problem of non-specific staining. Hydrogels are colloidal gels dispersed in water, and many different kinds have been reported; their uses vary widely from breast implants to nanobiotechnology and include such applications as wound dressings, disposable ECG electrodes, implants, contact lenses, 3D substrates for cell growth, tissue regeneration and tissue engineering, to name just a few. Polyacrylamide-based hydrogels for protein immobilisation have been pioneered by Mirzabekov (1, 2) and have been used by many for both protein and nucleic acid immobilisation (8, 23–29). However the "boom and bust" of the protein microarray "bubble" has, sadly, resulted in the decline of commercial interest in manufacturing polyacrylamide-based hydrogel microarray slides and the inevitable surge in price to end users. A variety of commercial hydrogels of interest to a biomedical researcher can be found for example here (www.xantec.com). We prefer to make our own hydrogel slides, which perform well in routine fluorescence-based peptide affinity capture assays due to their lack of any non-specific protein sorption and the lack of autofluorescence. Here we describe protocols for making polyacrylamide hydrogels and their applications for peptide affinity capture. In the following **Sections 2 and 3** we aim to describe a set of protocols covering all key stages of affinity peptidiomics analyses.

2. Materials

2.1. Selection of Peptides for Anti-peptide Antibody Development

1. *PeptideMass* on-line service for in silico digestion of Uniprot sequences with a chosen proteolytic enzyme (www.expasy.org/tools/peptide-mass)
2. *ProtParam* tool for calculating hydropathicity grand averages (www.expasy.org/tools/protparam)
3. *Antigenic* on-line tool for predicting antigenicity (liv.bmc.uu.se/cgi-bin/emboss/antigenic)

2.2. Peptide Selection Using Phage Display

1. A phage display library, for example, Ph.D.-C7C™ Phage Display Peptide Library Kit (New England BioLabs), 2×10^{10} pfu/μL, 1.2×10^{9} transformants
2. Sequencing Primer: "-96 gIII primer" 5′ – CCC TCA TAG TTA GCG TAA CG – 3′
3. *E. coli* strain: F′ *lacI*^{q} *Δ(lacZ)M15 proA+B+ zzf::Tn10(Tet*^{R}*)/fhuA2 supE thi Δ(lac-proAB) Δ(hsdMS-mcrB)5* (r_k^- m_k^- McrBC$^-$)
4. Antigen protein solutions: 100 μg/ml (*see* **Note 1**)
5. Corning® Universal-BIND™ multi-well plates (Corning Life Sciences)
6. Stratalinker(R) for UV crossling (Stratagen), or a similar UV source
7. LB medium: 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl in dH_2O. Autoclave, store at +4°C
8. IPTG/X-gal: 50 mg/ml IPTG, 40 mg/ml X-gal in dimethylformamide. Store at –20°C in the dark
9. LB/IPTG/X-gal plates: 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, 1.5% (w/v) agar in dH_2O. Autoclave. Cool to below 70°C. Add 0.1% (v/v) IPTG/X-gal and pour onto plates. Store plates at +4°C in the dark.
10. Agarose: 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, 0.1% (w/v) $MgCl_2 \times 6H_2O$, 0.7% (w/v) agarose in dH_2O. Divide into small aliquots (~50 ml). Autoclave. Store at room temperature. Melt in microwave prior to use.
11. Antibiotics: 0.005% (w/v) tetracycline in ethanol. Store at –20°C in the dark.
12. LB-tet plates: 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, 1.5% (w/v) agar in dH_2O. Autoclave. Cool to below 70°C. Add 0.0001% (v/v) tetracycline stock and pour onto plates. Store plates at +4°C in dark.
13. Phosphate buffered saline (PBS): 10 mM phosphate buffer, 2.7 mM potassium chloride, 137 mM sodium chloride, pH 7.4
14. TBS: 10% (v/v) 0.5 M Tris-HCl, 150 mM NaCl, pH 7.5
15. 0.1% TBST: 10% (v/v) 0.5 M Tris-HCl, 150 mM NaCl, 0.1% (v/v) Tween-20, pH 7.5.
16. PEG/NaCl: 20% (w/v) polyethylene glycol-8000, 2.5 M NaCl
17. TE buffer: 1 mM EDTA, 10 mM Tris-HCl, pH 8
18. 2% (w/v) BSA in PBS and 0.1 mg/ml BSA in 0.1% TBST

19. Elution buffer: Glycine-HCl, 1 mg/ml BSA, pH 2.2
20. Phenol: 10 mg of phenol in 1 ml dH_2O, add 10 ml ethanol, mix, allow phases to separate. Collect upper phase. Repeat procedure by adding 10 ml ethanol, mix, allow to separate, collect upper phase. Repeat 2–3 times.
21. Solution of 3 M Sodium Acetate (pH 5.5) and 70% cold ethanol.

2.3. Affinity Peptidomics: Antibody Microarrays

1. Flexys microarray gridding robot (Genomic Solutions Inc.) or another contact microarray spotting robot
2. BioChip microarray Scanner (Packard Bioscience)
3. Standard glass microscope slides
4. Biodyne® Positively charged nylon membrane (0.45 μm) or similar membrane (*see* **Note 2**)
5. Size-exclusion chromatography (SEC) setup: Waters 600E pump and system controller (Waters) and Spectroflow 757 Absorbance detector (Applied Biosystems); Sephadex® G-25 column (5 ml bed volume) (*see* **Note 3**)
6. Rabbit polyclonal anti-peptide antibodies (*see* **Note 4**)
7. Other antibodies: Goat anti-rabbit IgG-Cy3 (Sigma); Total rabbit IgG (Sigma)
8. Proteins: Bovine serum albumin (BSA): 9% (w/v) in water
9. Trypsin inhibitor: 10 mM PMSF in isopropanol, store at –20°C
10. Complete Mini EDTA-free Protease Inhibitor Cocktail Tablets (Roche Applied Science). Prepare 25× stock solution by dissolving one tablet in 400 μl of dH_2O, store at –20°C.
11. SEC running buffer (use PBS): 10 mM phosphate buffer, 2.7 mM potassium chloride, 137 mM sodium chloride, pH 7.4
12. Microarray blocking and assay buffer: 9% BSA, 0.1% Tween 20 in PBS
13. Microarray washing buffer: 0.1% BSA, 0.02% Tween 20 in PBS
14. COOMASSIE® Brilliant Blue G-250
15. Rhodamine B isothiocyanate (RITC)
16. Sequencing-grade Trypsin

2.4. Peptide Assays on Hydrogels

1. Standard glass microscope slides (Menzel-Glaser, Braunschweig, Germany) (*see* **Note 5**)
2. Cleaning solution: 10% sodium hydroxide

3. Solvents: 100% Ethanol (or Methanol)
4. Binding silan: add 3 μl f 3-(Trimethoxysilyl)propyl metharylate to 1 ml of 100% Ethanol. Add 250 μl of 10% (v/v) Acetic acid
5. Hydrogel frames and incubation chambers: adhesive gaskets (1.5 × 1.6 cm, Abgene)
6. Hydrogel components: 1 M Acrylamide; 20 mM *N,N′*-Methylenebisacrylamide; 0.1% (v/v) TEMED; 1 mg/ml ammonium persulfate
7. Hydrogel activation solution: 25% (w/w) glutaraldehyde
8. Micro Bio-Spin 30 Columns, SSC buffer (BioRad)
9. Sample dilution and Hydrogel assay buffer: 0.005% (v/v) Tween 20 in PBS
10. Hydrogel washing buffer (use PBS): 10 mM phosphate buffer, 2.7 mM potassium chloride, 137 mM sodium chloride, pH 7.4
11. Negative controls for hydrogel assays: 1 mg/ml BSA in 0.005% (v/v) Tween 20 in PBS or any other irrelevant protein in the same buffer
12. Labelling buffer: 1 M K_2HPO_4 pH 9.0. Store at +4°C
13. Fluorescent dye for labelling through sylfhydryl groups: 0.1% (w/v) NIR-664-iodoacetamide in 100% Methanol. Store at –20°C in dark
14. Fluorescent dye for labelling through amino groups: 0.1% (w/v) Rhodamine B isothiocyanate (RITC) in 100% Methanol. Store at –20°C in dark
15. Synthetic peptides: Prepare 10% (w/v) stock solution in DMSO, dilute them 1:10 in 10 mM K_2HPO_4, pH 7.1 to yield 1% (w/v) peptide stock solution. Store at –20°C
16. Reducing reagent: 200 mM Butyltriphenylphosphonium bromide (TBP)
17. SEC resin: Sephadex® G-25 (Pharmacia Fine Chemicals) 75% suspension swelled in 10 mM K_2HPO_4. Store at +4°C
18. Disposable 1-ml syringes (BD Plastipak™) and filter paper
19. Falcon centrifuge tubes (15 ml); A suitable centrifuge capable of spinning these at 1000×*g*
20. MALDI matrix #1: 10 mg/ml alpha-cyano-4-hydroxy cinnamic acid (CHCA) in 60% acetonitrile
21. MALDI matrix #2: 7 mg/ml CHCA in 25% acetonitrile, 0.2% trifluoroacetic acid (TFA)
22. BioChip microarray Scanner (Packard Bioscience)
23. Reflex III MALDI-TOF mass spectrometer (Bruker)

3. Methods

3.1. Selection of Peptides for Anti-peptide Antibody Development

Although there are many useful tools for predicting the antigenicity of *proteins* and protein epitopes, no suitable prediction tools are available for predicting which of the *tryptic peptides* would make better antigens for the development of *anti-peptide* antibodies. Many of the tools reported so far rely heavily on protein *structural* information or aim to identify solvent-exposed linear epitopes based on protein structure and may not select sequences which are "antigenic" but are not fully solvent exposed. Such tools have only limited usability in the analysis of tryptic peptides for their antigenicity. Based on our prior experience, however, some limited yet clear correlation exists between the tryptic peptides' hydrophilicity and the ability of the relevant anti-peptide antibodies to capture proteolytic peptides in a MALDI-TOF-MS assay. Parent protein structure, folding and fragment solvent exposure play no role in determining *tryptic* peptides' antigenicities. The approach detailed below should be useful for selecting the best tryptic peptide sequences for anti-peptide antibody development.

1. Enter protein sequence or database accession number of the protein of interest into the *PeptideMass* program (*see* **Notes 6 and 7**).
2. Select "reduced" option for Cysteines, select no acrylamide adducts, no Methionine oxidation, $(M+H)^+$ and monoisotopic masses. Select "Trypsin", choose "no missed cleavages" and choose to display all peptides (i.e. bigger than 0 Da). Choose to sort peptides by peptide masses (*see* **Note 8**).
3. Choose to display all post-translational modification, database conflicts, all polymorphisms and splice variants (*see* **Note 9**).
4. Perform the analysis, the *PeptideMass* program will display a list of predicted tryptic peptides, their masses and any information on splice variants, isoforms and database conflicts. For ease of use, copy the table and paste e.g. into EXCEL datasheet.
5. Select a subset of peptides suitable for chemical synthesis (*see* **Note 10**).
6. Analyse the selected subset for peptides suitable for antibody generation. From all the prediction techniques we tested so far hydrophilicity profiling worked best (*see* **Note 11**).

3.2. Peptide Selection Using Phage Display

Phage display is an example of a molecular display system capable of displaying not just small peptides but also large proteins, including antibodies. Phage display has been widely used and reported since its development in 1985 (12, 18, 20, 30–33). We and others have previously reported the use of recombinant scFvs generated with Phage display for the capture and detection of proteolytically digested proteins in microarray format with fluorescent detection and directly by MALDI-MS. Other uses of the system include the development of peptide aptamers or characterisation of anti-protein or anti-peptide antibody epitopes. A number of phage display kits are now commercially available. These include kits for making own phage display libraries, ready-made libraries of random peptides and cDNA libraries. For example the "T7Select®" kit from Novagen is based on the bacteriophage T7 and is capable of expressing a high copy number of short polypeptide (e.g. 415 copies of up to 50 amino acids long polypeptide per single phage) or lower copy number of larger polypeptides or proteins (e.g. 1 or 5–15 copies of up to 1200 amino acid long protein per single phage). Pre-made cDNA phage display libraries from a variety of human tissues are available from Novagen such as normal and disease T7 Select libraries from brain, breast, colon, heart, liver, lung, stomach, and from Alzheimer's brain, breast tumour, colon tumour, liver tumour and lung tumour. Random peptide phage display libraries are available from New England Biolabs, e.g. Ph.D.-7™ and Ph.D.-12™ (these display 7- and 12-amino acid long random peptides fused with a coat protein (pIII) and expressed on the phage surface) and Ph.D.-C7C™ (this library displays 7 amino acids long random peptide flanked by a pair of cysteine residues). A disulfide cross-link between the two cysteines in the Ph.D.-C7C™ library results in the display of circularised peptides; Ph.D.-7 and Ph.D.-12 phage libraries display linear peptides.

Any phage display experiment includes a number of affinity selection steps which are repeated until the desired specificity of binding is achieved. In each panning step a solid support (e.g. a membrane or a multiwell plate) coated with the target protein or peptide is required; phage display library is incubated with the target, unbound phages are washed away and the bound phages are eluted and amplified and the procedure is repeated, as shown in **Fig. 23.1.** Phage display allows to quickly select affinity pairs and to determine the sequence of the identified epitopes (15–17). We will exemplify the use of phage display with a commercially available Ph.D.-7™ kit from New England Biolabs.

3.2.1. Panning

1. Dilute the target protein stock to 100 μg/ml in PBS (*see* **Note 1**).

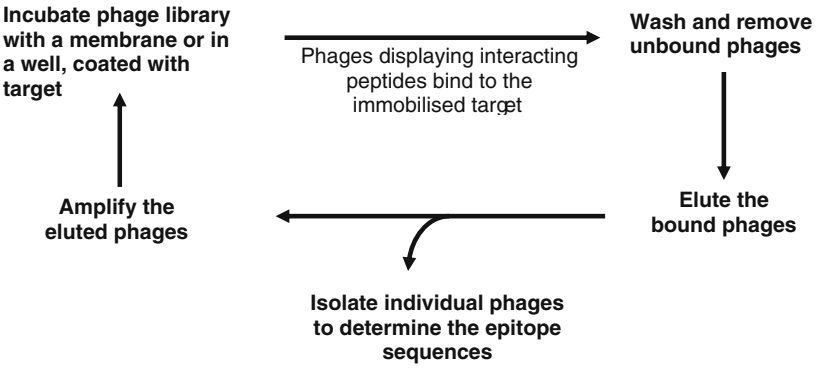


Fig. 23.1. Phage panning scheme. Several panning rounds may be necessary in order to enrich and isolate the phages capable of specific high-affinity binding to the target.

2. To each well of a Costar's multi-well plate, add 300 μl of the diluted protein and incubate for 1 h at room temperature in the dark.
3. Decant the solution, immobilise the protein by exposing the plate to UV using Stratalinker (energy setting 500) or another suitable UV source (*see* **Note 12**).
4. Rinse the plate once with dH_2O, and twice with PBS.
5. Block the plate with 2% BSA (in PBS) for 30 min.
6. Dilute the whole phage library in the required volume of 0.1 mg/ml BSA (in 0.1% TBST) and add to wells containing the target protein. Seal the plate to prevent evaporation and incubate at room temperature overnight (*see* **Note 13**).
7. Inoculate 20 ml LB medium with 2–3 colonies of *E. coli* (streaked at least 1 day beforehand on an LB-tetracycline plate and kept at +4°C in the dark) and incubate at 37°C with 120 rpm shaking for 2 h (*see* **Note 14**).
8. Aspirate the phage solution and wash the plate ten times with 0.1% TBST, remove all washing buffer after each washing step, slap the plate face down on a clean paper towel to remove as much washing buffer as possible after each wash (*see* **Note 15**).
9. Elution of the specifically bound phages can be done by using either of the Step (i) or (ii) outlined below (*see* **Note 16**)
   i. To elute the bound phage specifically, add 100 μl of the free target protein (100 μg/ml) to each well and incubate for 30 min. Collect the eluted phage and store at –20°C.
   ii. To elute the phage nonspecifically, incubate the well with 100 μl of 0.2 M Glycine-HCl (pH 2.2)

containing 1 mg/ml BSA for 8 min. Collect the eluted phage solution and neutralise with 15 μl of 1 M Tris-HCl (pH 9.1). Store the eluates at –20°C.

10. To titer the eluted phage, follow the protocol described in the **Section 3.2.2** below (*see* **Note 17**).
11. Divide the eluted phage sample in two equal aliquots. Half of the material should be stored at –20°C (as a backup or for future use) and the other half could be amplified.
12. To amplify the phage, use the 2 h culture of *E. coli* (from Step 7 above). Transfer an aliquot to sterile 50-ml Falcon tube and dilute it with fresh LB so that the final $OD_{600} = 0.01$ and the final volume is 2 ml. Add the phage eluate (50 μl = 1/2 of the total eluted sample) to the 2 ml diluted *E. coli* culture, incubate in a shaking incubator at +37°C (250 rpm) for 4.5 h (*see* **Note 18**).
13. Transfer the culture to 2.0 ml Eppendorf tubes and centrifuge for 10 min at 10,000×*g* at +4°C. Transfer the supernatant to a fresh tube and centrifuge again (same speed, 10 min).
14. Carefully collect the upper 80% (1.6 ml) of the supernatant and transfer to a fresh tube, add 1/5 volume (400 μl) of PEG/NaCl. Allow phage to precipitate at +4°C overnight (*see* **Note 19**).
15. Precipitate the phage by spinning for 15 min at 10,000×*g* at +4°C. Carefully remove the supernatant (do not discard it until the phage pellet is recovered), re-spin the tubes for 5 s at room temperature, carefully remove the residual supernatant with a fine pipette (*see* **Note 20**).
16. Resuspend the phage pellet in 200 μl TBS + 0.02% NaN_3, incubate the solution on ice for 5 min.
17. Centrifuge for 1 min at 10,000×*g* at +4°C to pellet any remaining insoluble matter. Transfer the supernatant to a fresh tube. This is the amplified phage.
18. Titer the amplified phages as described in **Section 3.2.2** below. Store the rest of the phage solution at –20°C for use in the next round of panning (*see* **Note 21**).
19. A detailed track record of all the panning and titering experiments should be kept for each individual phage library/target protein pair tested. Make a standard flow chart (e.g. as shown in **Fig. 23.2)** and fill the blanks as the work progresses.

3.2.2. Titering

1. Inoculate 20 ml of LB with 2–3 colonies of *E. coli* and incubate at +37°C with 120 rpm shaking until the culture has reached its mid-log phase ($OD_{600} \sim 0.5$).

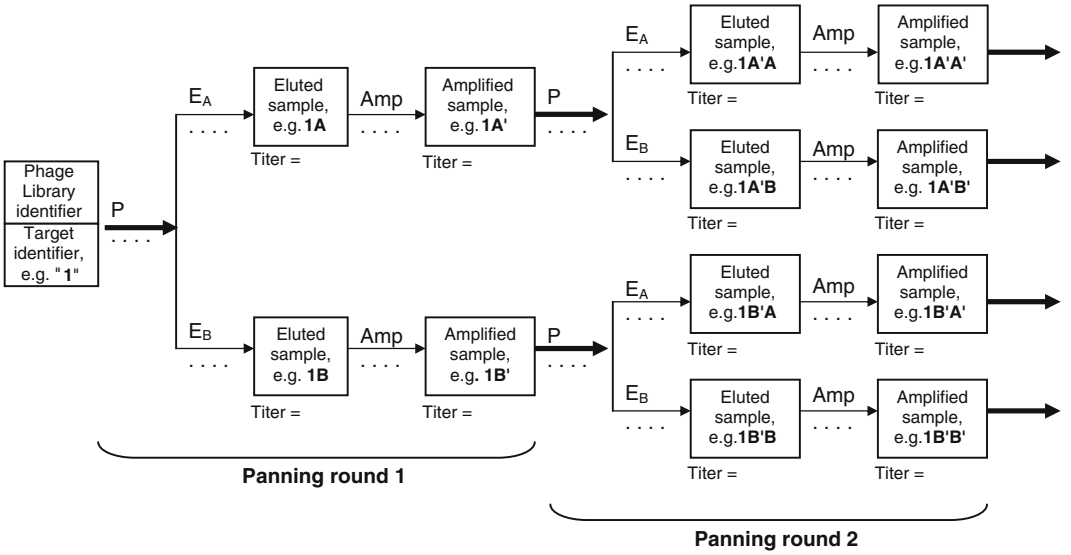


Fig. 23.2. Phage panning flowchart. Label each eluted phage such that the identity of the target and of the panning stage is clear. "P" denotes panning, "E_A" and "E_B" denote Elutions, whilst the superscript denotes elution conditions, e.g. specific and non-specific elution, or soft and stringent elution. "Amp" denotes amplification. Panning round 2 is shown to contain two elutions for each of the two samples, this is optional. For the simplification of the next panning rounds, Phage sample 1B′A′ could be joined with 1A′A′ and 1A′B′ with 1B′B′ Draft flowchart covering four panning rounds should be sufficient for most of the applications. Sample identifies, dates and the titer values should be entered into the table as the work progresses.

2. Pre-warm LB/IPTG/Xgal plates by incubating them at +37°C for at least 1 h (*see* **Note 22**).
3. Whilst the cells are growing, melt the agarose in a microwave and dispense 3-ml aliquots into sterile 15-ml Falcon tubes. Make one Falcon tube per phage dilution. Keep tubes in a +45°C water bath until use (*see* **Note 23**).
4. Prepare dilutions of phage in LB (total volume 10 μl) in sterile 1.5-ml microcentrifuge tubes.
5. Once the culture has reached the mid-log phase, pour 400 μl of the culture into 1.5-ml tubes containing phage dilutions.
6. Vortex and incubate at room temperature for 5 min to allow infection.
7. Transfer the infected cells to tubes containing warm agarose, mix with the pipette tip (do not vortex) and immediately pour onto a pre-warmed LB/IPTG/Xgal plate. Spread the agarose evenly by tilting the plate.
8. Allow plates to cool for ~5 min, or until the top agarose solidifies, invert and incubate overnight at +37°C.
9. The following day inspect the plates and count the plaques. Multiply the number of plaques by the dilution factor for that plate to get the phage titer (in pfu per 10 μl).

3.2.3. DNA Extraction

1. Inoculate 20 ml LB medium with *E. coli*, incubate in a shaking incubator at +37°C (250 rpm) for 2 h.
2. Following that initial incubation, dilute the bacterial culture with fresh LB so that the final $OD_{600} = 0.01$; transfer 2-ml aliquots into sterile 50-ml Falcon tubes. Pick a single blue plaque from the plates used for titering the phage and add it to the 2 ml *E. coli* culture; incubate in a shaking incubator at +37°C (250 rpm) for 4.5 h.
3. Transfer the culture to 2.0-ml microcentrifuge tubes and centrifuge for 10 min at 10,000×*g* at +4°C. Transfer the supernatant to a fresh tube and centrifuge again for 10 min at 10,000×*g* at +4°C.
4. Carefully collect the upper 80% (1.6 ml) of the supernatant and transfer to a fresh tube, add 1/5 volume (400 µl) of PEG/NaCl. Allow phage to precipitate at +4°C overnight.
5. Precipitate the phage by spinning for 15 min at 10,000×*g* at +4°C. Carefully remove the supernatant (do not discard it until the phage pellet is recovered), re-spin the tubes for 5 s at room temperature and carefully remove the residual supernatant with a fine pipette.
6. Resuspend the pellet in 100 µl TE Buffer Mix. Add 100 µl phenol. Vortex.
7. Centrifuge at 10,000×*g* for 5 min at room temperature. Collect 80% of the upper (water) phase to fresh tubes.
8. Add 100 µl of phenol:chloroform (1:1). Vortex 3 min at room temperature. Centrifuge at 10,000 rpm for 5 min at room temperature.
9. Transfer the top 80% of the upper water phase to fresh tubes. Add 70 µl of chloroform. Vortex for 3 min at room temperature. Centrifuge at 10,000×*g* for 5 min at room temperature.
10. Transfer the top 80% of the upper water phase to fresh tubes. Add 1/10 volume of 3 M Sodium Acetate pH 5.5 and 2.5 volumes of ethanol. Vortex briefly. Incubate at –20°C for at least 30 min or overnight.
11. Centrifuge the tubes at 14,000×*g* at +4°C for 20 min.
12. Remove supernatant, wash once with 1 ml of cold 70% ethanol. Dry at room temperature (for approximately 1 h).
13. Add 30 µl of deionised H_2O to suspend the DNA pellets. Store at –20°C. These single-stranded DNAs are ready for sequencing.

3.3. Affinity Peptidomics: Antibody Microarrays

Microarrays allow miniaturisation and multiplexing of affinity-based assays, and a large number of array formats and assays have been reported to date, including for the analysis of serum

samples. The availability and affordability of anti-protein antibodies is often an issue, whilst another typical issue in any protein-based assay is sample stability and preservation (34). Affinity peptidomics relies on the successfully proven approach used widely in mass-spectrometry-based protein analysis, where protein samples are proteolytically digested prior to the analysis. Such treatment removes the need to preserve *protein* samples.

To further streamline the affinity assay, we have chosen to use single-label competitive assays rather than traditional direct binding two-colour assays. The justification of the choice can be found here (11, 35); briefly, this approach allows to avoid repetitive labelling of the experimental samples and compensates for the heterogeneity of the antibody affinities. Our protocols were originally devised for use with recombinant scFv anti-peptide antibodies developed using Phage display (8), but were later adapted for use with traditional anti-peptide polyclonal antibodies. Such peptide affinity assays are widely applicable to the detection and quantification of the proteolytic or naturally occurring peptides.

3.3.1. Proteolysis and Labelling of Serum Protein Samples

1. Aliquot the required amount of sera (e.g. we used 100 μl of each of the serum samples to be tested), add a few microlitres of 1 M K_2HPO_4 or 1 M Tris pH 9 to bring the pH of the sample to pH 8, check pH by spotting a fraction of a microlitre of the buffered serum onto pH paper (*see* **Note 24**).
2. Make one additional pooled serum sample by mixing equal volumes from all serum samples being tested (*see* **Note 25**).
3. Add Trypsin to each sample, including the pooled serum sample, use 1 μg per ~20–50 μg of the total serum protein and incubate at 37°C overnight (*see* **Note 26**).
4. Stop the digestions by adding 20 μl of 10 mM PMSF (*see* **Note 27**).
5. To fluorescently label the pooled serum sample, take a 20-μl aliquot, add 80 μl PBS and add 100 μl of 1% RITC. Incubate at room temperature for 30–60 min.
6. Stop the labelling reaction by adding 20 μl of 1 M Tris pH 8. Proceed with purification (**Section 3.3.2).**

3.3.2. Purification of the Peptides (see **Note 28***)*

1. Calibrate the SEC column by injecting Trypsin diluted in PBS (*see* **Note 29**).
2. Monitor absorbance at 280 nm. Identify the elution peak for trypsin, the end of which will indicate when to commence peptide collection during SEC purification (*see* **Notes 30 and 31**).
3. Load the labelled pooled serum (from Step 6, **Section 3.3.1)** onto the SEC column. Commence the collection

at a time determined in the previous step. Stop collection when the unincorporated Rhodamine peak (slowly moving band) reaches the end of the column. Store the collected peptide on ice (short-term) or freeze –20°C for the longer term storage (*see* **Note 32**).

3.3.3. Microarrays for Fluorescent Detection and Quantification of Peptides (see **Note 33***)*

1. Set up the microarray spotting instrument. The Flexys microarray gridding robot allows for three washing buffers to be used for cleaning the pins and the washing program should be set as follows: 1% Tween 20 wash for 30 s, followed by PBS wash for 10 s, followed by another wash in 1% Tween 20 for 30 s and PBS wash for 10 s. The final wash is in 0.1% BSA in PBS with 0.1% Tween 20 for 30 s (*see* **Note 34**).
2. Fix membranes on glass slides, e.g. using small paper stickers or small pieces of tape and place the slides in the robot holder (*see* **Note 35**).
3. To check pins quality and to match the pins, perform a trial run by spotting the same fluorescently labelled protein and scan the slides to determine the efficiency of protein transfer for each individual pin (*see* **Notes 36, 37, and 38**).
4. To measure sample volumes required for spotting, add an even number of identical ~20-μl aliquots of any sample to the microwell plate, and insert it in the robotic spotter. Samples should have the same protein concentration and buffer as that in the antibody samples to be spotted. Choose the wells (or pins) such that half of the samples are transferred to the membrane, and half are not used. Run a number of transfers (e.g. ~100). Remove the plate from the robot and measure the remaining sample volumes, compare volume in the used and unused wells, average the difference and divide by the number of transfers (*see* **Note 39**).
5. Add the required amount of antibodies to microwell plates, insert them into the robot holder and run the spotting program using the parameters specified and tested in previous steps (*see* **Note 40**).
6. Remove membranes from the robot and transfer them into a sealed chamber containing a few millilitres of 37% formaldehyde. Incubate overnight in a fume hood at room temperature (*see* **Note 41**).
7. Block the membranes using large volume of Microarray blocking and assay buffer (~10 ml per membrane for at least 2 h) (*see* **Note 42**).
8. Assemble the assay mixtures as follows (exemplified for 200 μl final volume sample): Use ~10 μl of the unlabelled

serum digest (or the equivalent amount of the purified proteolytic peptides), add 1 μl of the 25× Protease Inhibitor Cocktail, incubate for 15 min at room temperature. Add 50 μl of the labelled and purified pooled sera digest and 140 μl of the fresh Microarray blocking and assay buffer. Assemble an individual assay mixture for each of the tested sera samples (*see* **Notes 43 and 44**).

9. Trim the array membrane to minimal size. Add 100 μl of the assay mix to a small Petri dish, place the array membrane face down in incubation mix, and add the remaining ~100 μl on top of the membrane. Close the Petri dish; incubate at room temperature in the dark for 2 h.
10. To wash the membranes transfer them to a flask containing ~50 ml of the Microarray washing buffer for 10 s, change buffer and incubate for 5 min, change buffer again and incubate for 10 min. (*see* **Note 45**).
11. Dry membranes on blotting paper (arrayed side up) in darkness (*see* **Note 46**).
12. Mount the dried membranes on glass slides using double-sided adhesive tape and scan using a suitable instrument. We use a BioChip microarray Scanner. The scanner settings (focus, laser intensity and photomultiplier attenuation) should not be changed between the different slides. **Figure 23.3** illustrates a fragment of the scanned microarray, and shows all the normalisation and control spots.
13. Data analysis depends on whether competitive or non-competitive assay was used and also on the set of normalisation spots used. In most cases, however, readouts should be normalised pin-to-pin and array-to-array (*see* **Note 47**).

3.4. Peptide Assays on Hydrogels

Porous membranes provide a convenient support material, which is strong and for which a variety of materials and protocols are available. Hydrogels cannot compete with membranes in terms of strength and durability, but they provide the best 3D support for the immobilisation of test molecules (whether proteins or peptides) in their native functional state in highly porous hydrogel substrate suitable for both functional assays (36) and immunoassays (26, 37, 38). When hydrated, the hydrogels swell, allowing easy access for the molecules and short diffusion times, but when dried, the gel thickness is reduced significantly, resulting in focussing of the trapped fluorescence in a thinner layer. This increases fluorescent readouts (especially on confocal scanners), whilst the background fluorescence remains extremely low (no autofluorescence, no non-specific protein sorption). Furthermore, hydrogels are also suitable for use as MALDI-MS substrates (8, 39). Having an anti-peptide antibody immobilised on

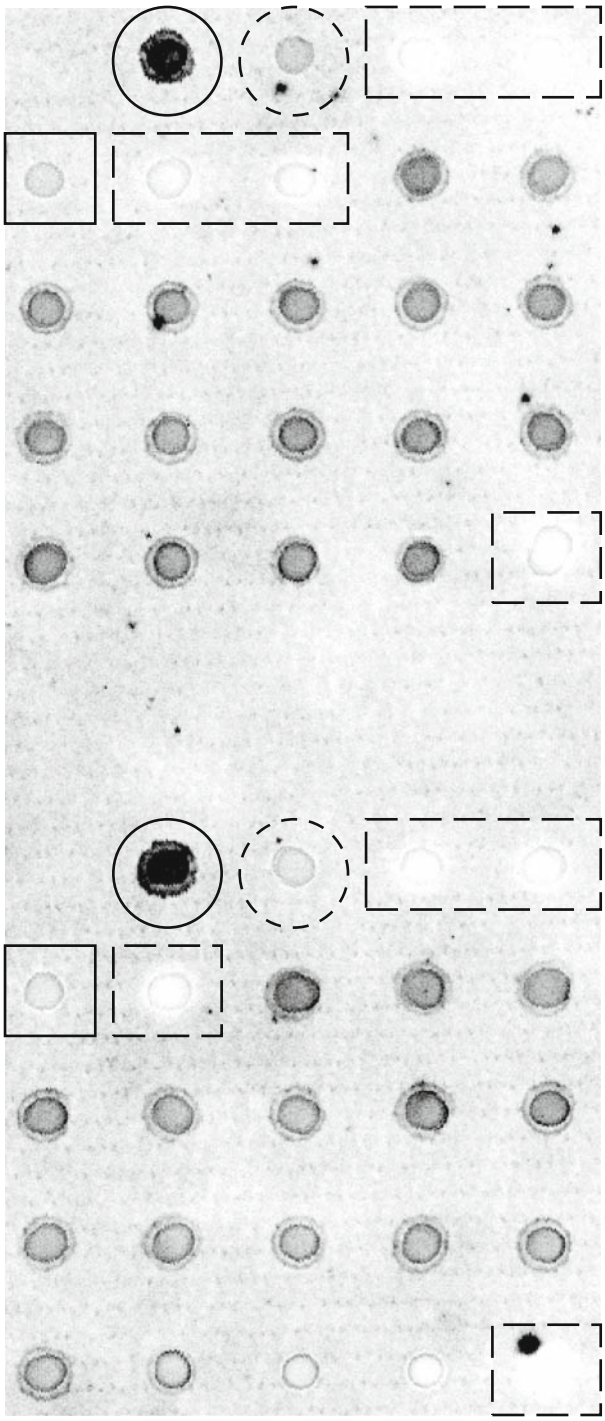


Fig. 23.3. Affinity peptidomics microarray assay. Fluorescent readout at 550 nm of a fragment of a microarray containing two grids with 33 anti-peptide antibodies, spotted onto positively charged nylon membrane and incubated with the proteolytic serum peptides in a competitive binding assay. Shapes denote Coomassie spots (*rectangles, dashed line*), total IgG negative control (*rectangles, solid line*), fluorescent references for pin calibration and grid normalisation – *circle, solid line* (active channel 550 nm), *circle, dashed line* (different channel 650 nm).

a hydrogel slide therefore allows to bypass multidimensional separation stages and to capture peptides for MALDI-MS analysis directly from crude Tryptic digests. We report here a set of distinct hydrogel-based assays, suitable for a variety of applications and also for further method development.

3.4.1. Making Hydrogels

1. Rinse microscope glass slides in 100% ethanol and soak in 10% sodium hydroxide overnight.
2. Rinse the slides four times in deionised water and twice in 100% ethanol.
3. Treat the slides with binding silan solution for 5 min, wash with 100% ethanol and dry at room temperature (*see* **Note 48**).
4. Assemble the adhesive gaskets onto the defined area of the slides. Attach two gaskets to each glass slide (*see* **Note 49**).
5. Make fresh polymerisation mix: 1 M acrylamide, 20 mM *N*,*N*′-Methylenebisacrylamide, 0.1% TEMED. Add 1 mg/ml ammonium persulfate (*see* **Note 50**).
6. Add 80 μl of the assembled polymerisation mix into the frame and carefully seal the frame with the plastic cover slips (provided with frames).
7. When gel is formed and polymerisation is finished, remove the plastic coverslip (leave the gasket on the slide) and wash the hydrogel in water overnight (*see* **Note 51**).
8. Dry the hydrogel slides at room temperature. Store in a dry clean slide box until use.

3.4.2. Target Immobilisation on Hydrogels

1. Activate hydrogels by immersing the slides in 25% glutaraldehyde overnight.
2. Wash the slides with deionised water twice for 5 min and dry the slides at room temperature (*see* **Note 52**).
3. Immobilise the antibodies by spotting 0.5 μl of 1 mg/ml antibody solution onto the hydrogel pads (*see* **Notes 53–55**).
4. Allow the spots to dry fully at room temperature, transfer the slides to a humidified chamber and incubate at +4°C overnight.
5. Rehydrate the hydrogels fully in the Hydrogel assay buffer prior to running affinity assays (*see* **Note 56**).

3.4.3. Affinity Binding Assays on Hydrogels (see **Note 57***)*

Hydrogels can be used to assay a variety of biological targets, including endogenous proteins or peptides (as in traditional Peptidomics applications) (40, 41), proteolytic peptides (as in Affinity Peptidomics) (8, 10), synthetic peptides (e.g. for validation

experiments), glycans or small-molecule ligands (42, 43). We will exemplify hydrogel affinity assays using a simple example of fluorescently labelled synthetic peptide, but the protocol would remain essentially the same for crude peptide digests (*see* **Sections 3.3.1, 3.3.2, and 3.3.3).** The same hydrogels may be probed with MALDI-MS.

1. To fluorescently label peptides obtained after proteolytic digestion of serum, or synthetic peptides having free amino groups (unprotected N-termini, Lysines), follow **Section 3.3.1** (Steps 4 and 5).
2. To fluorescently label synthetic peptides having free sulfhydryl groups (Cysteines) mix 10 μl of 1% peptide solution with 70 μl labelling buffer, add 2 μl of 200 mM TBP (final concentration 5 mM) and incubate the mixture at room temperature for 30 min (*see* **Notes 58 and 59**). Add 30 μl of 0.1% NIR-664-iodoacetamide fluorescent dye to the mixture and incubate at room temperature for 1 h in dark.
3. Whilst incubating the labelling reactions, prepare spin columns for SEC purification of the labelled peptides (one column per labelling reaction). Remove a plunger from 1 ml disposable syringe; cut filter paper to just over twice cross-sectional area of syringe, fold and push to the bottom of the syringe using the plunger; remove the plunger. Load 1 ml of 75% Sephadex® G-25 gel into the syringe column; insert syringe column into a 15 ml Falcon tube and spin at 1000×*g* for 5 min (*see* **Note 60**).
4. Replace the Falcon tube, load the labelled sample (~112 μl) to the centre of the spin column and centrifuge at 1000×*g* for 5 min.
5. Dispose the spin column, transfer the purified peptide sample to a fresh microcentrifuge tube. Store at –20°C.
6. To fluorescently assay proteolytic peptides in competitive assays, e.g. as in Affinity peptidomics assays (as described in **Section 3.3.3,** Steps 8–13), mix the equimolar amounts of the unlabelled peptide test samples and the labelled reference peptide samples, use Hydrogel assay buffer to make up the volume to at least 65 μl per single hydrogel pad (*see* **Note 61**).
7. To fluorescently assay individual peptides, e.g. in validation experiments, prepare two assays for each peptide tested. Use Hydrogel assay buffer to make up the volume to at least 65 μl per single hydrogel pad (*see* **Note 62**).
8. Add the assay mixture to fully hydrated Hydogel pads (*see* **Note 56**), incubate at room temperature overnight in the dark.

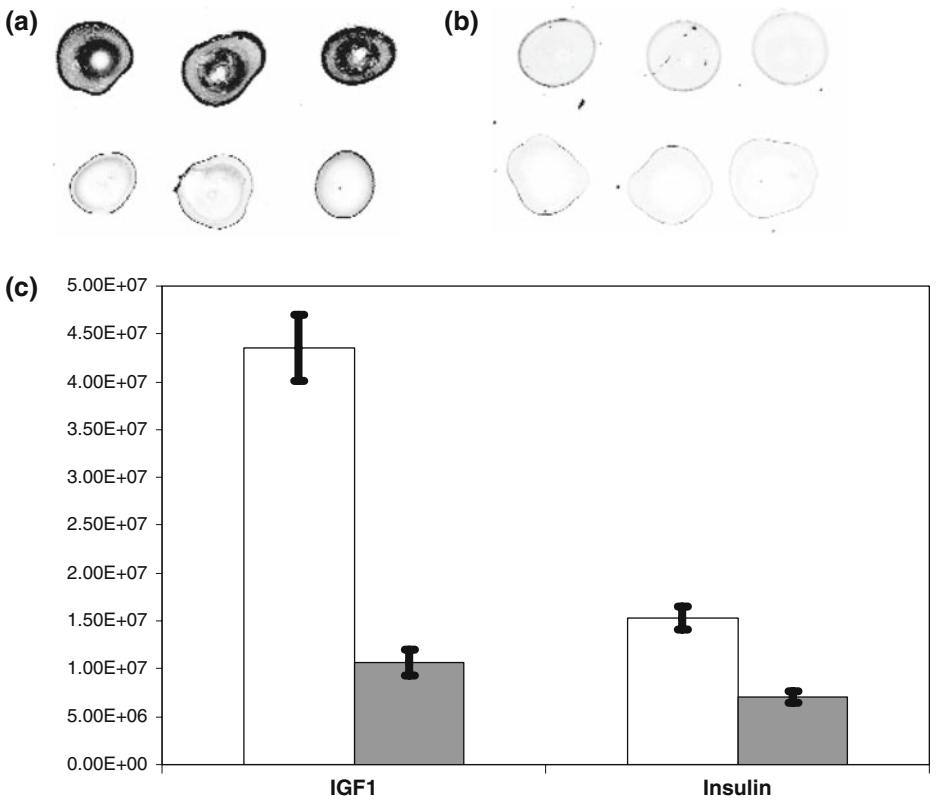


Fig. 23.4. Peptide binding on hydrogels. (**a**) Peptide SALNTPN binds to IGF1 but not to Insulin. (**b**) Same as above, but in the presence of 100-fold excess of the unlabelled peptide SALNTPN. (**c**) Mean values of the fluorescence intensities for the above are shown (±stdev).

9. Wash slides in Hydrogel washing buffer three times ×5 min, and in deionised water twice ×5 min.
10. Air-dry the slides and scan with fluorescent scanners at the appropriate wavelength. **A**ffinity capture of a peptide on hydrogel with the immobilised IGF1 and insulin protein is shown on **Fig. 23.4**. Insulin is used as the negative control (an irrelevant protein).
11. To prepare hydrogels for MALDI-MS, add matrix on top of the hydrogel as follows: apply MALDI matrix #1, air dry; apply MALDI matrix #2, air dry. The hydrogels can be examined on MALDI-TOF MS (*see* **Note 63**).

4. Notes

1. We used a recombinant analogue of human insulin like growth factor 1 (Long R^3 IGF-1, Sigma-Aldrich) to exemplify this protocol, which otherwise is easily adaptable for

use with other proteins. Following the recommendation of the provider, the protein was dissolved in 1 ml of 10 mM HCl. Other protein targets may require different buffers and preparation procedures.

2. Membranes provide 3D porous substrate with very high protein binding capacity and are therefore preferred over flat 2D substrates. Ready-made and commercially available membrane substrates such as immobilised Nitrocellulose (available from multiple suppliers) or FAST™ and CAST™ slides from Schleicher and Schuel could be used, but these would provide a more expensive alternative to ordinary membranes.
3. Liquid chromatography setups vary and any suitable equipment and properly sized columns could be used. Gravity flow may also be used for peptide purification, but care should be taken to properly calibrate the elution times of the protein (Trypsin) fraction, the peptides and the unincorporated RITC. The flow rate will vary if gravity flow is used, so calibration should be done by the volume eluted (weigh each tube containing each sample and subtract the weight of the tube) rather than the elution time.
4. We are currently using 70 sera samples raised against 35 peptide markers (two rabbits per peptide). Antibody sample purity and the protein binding capacity of the microarray substrate material will affect the amount of retained antibodies and therefore the maximum signal obtainable. Surfaces with lower binding capacities may be used with purified antibodies. Total IgGs will require supports with higher protein binding capacities to ensure that sufficient amount of the specific antibody is attached to the membrane.
5. Use a fresh pack of glass slides. Glass slides may not work with MALDI-MS detection, in which case Silicon wafers should be used (8).
6. Entering UniProtKB, Swiss-Prot or TrEMBL accession numbers is the preferred option, since this would allow to also include in the analysis the post-translational modification, database sequence conflicts, alternative splicing variants and polymorphisms.
7. This tool is convenient for the analysis of individual or small sets of proteins. We created a simple proteolytic digestion tool using EXCEL, which we use for in silico digestion and comparison of individual or groups of proteins. Other methods for predicting proteolytic peptides can be used; the choice of the method should not affect the outcome of the predictions.

8. Although mass calculations are not critical at this point, it is worth selecting this and other options, as these would become useful later.
9. Having this information handy will help to avoid errors in the subsequent anti-peptide antibody generation programme, which may be costly and which may cause very long delays, e.g. if an antibody has to be re-made.
10. The first step in anti-protein antibody generation is to make synthetic peptides. There is a list of criteria to bear in mind when selecting suitable peptide sequences, briefly:
    i. Peptide lengths should be between 5 and 30 amino acids. Short peptides are difficult to purify following the synthesis, whilst ~30-mers and above will have lower yields because of the increased error rate. Price will play a significant role too. Often 12–15 amino-acid-long peptides work best for anti-peptide antibody generation.
    ii. Avoid multiple Prolines, Serines, Aspatic Acid and Glycines.
    iii. Avoid the following duplets of amino acids: Ser-Ser, Asp-Gly, Asp-Pro.
    iv. Avoid the following triplets of amino acids: Gly-Asn-Gly, Gly-Pro-Gly.
    v. Avoid charge clustering and fewer than 1 in 5 charged amino acid side chains. The selection of the subset of suitable peptide can be achieved simply by selecting the range of lengths 10–15 amino acids in the EXCEL file, containing the output of the *PeptideMass* program (from the previous step), followed by a quick check for any of the unwanted amino acids (as outlined above). We have entered the above rules into a Visual Basic Macro which is run in Excel, making the selection easy even if multiple proteins are analysed. Having sorted the *PeptideMass* results by mass (Step 2) allows to very easily select a range of peptides of suitable size. We also used truncated tryptic sequences (i.e. just partial peptide sequence, if the predicted peptides were too long).
11. Much has been published on the prediction of antigenic epitopes from protein sequences (44–50). Most of the simpler tools however are based on the amino acid propensity scales, which take into account the hydrophilicity, surface accessibility and segmental mobility of amino acids (51) and are not therefore suitable for selecting tryptic peptides for anti-peptide antibody generation. Surprisingly, the "old ideas" of relying on hydrophobicity scales (52–54) appear to work better than any of the more modern tools

Table 23.1
Five peptides from human vascular cell adhesion molecule (VCAM) chosen for anti-peptide scFv(s) antibody development (8)

Peptide sequence detection	Amino acid length	Hydrophilicity[a]	GRAVY[b]	EMBOSS antigenic[c]	Peptide detection on MALDI-MS
SQEFLEDADR	10	1.44	–1.44	0.998	Very strong
TQIDSPLNGK	10	0.96	–0.96	1.038	Strong
LHIDDMEFEPK	11	0.88	–0.882	No epitope detected, no score	Weak
VTNEGTTSTLT MNPVS FGNEHSY	23	0.58	–0.583	1.031	Very weak
SSEGLPAPE IFWSK	14	0.35	–0.35	1.077	Not detected

[a]Hydrophilicity was calculated using Kyte Doolittle Hydrophilicity scale from (53).
[b]GRAVY indexes were calculated using *ProtParam* tool (www.expasy.org/tools/protparam)
[c]EMBOSS *Antigenic* scores calculated using (liv.bmc.uu.se/cgi-bin/emboss/antigenic) tool.

which we tested. Although the original idea published by Hopp and Woods was to look for hydrophilic regions, because they were the most likely ones to represent surface-exposed fragments, the same principle seems to work for selecting tryptic peptides, although it is not clear why. We use our own ranking tool (a Macro run within EXCEL), which uses Kyte Doolittle hydrophilicities (53). Any similar tools, including on-line tools, might be used, for example the ProtParam tool. Use the grand average of hydropathicity (GRAVY) index. **Table 23.1** shows VCAM peptides used for the generation of single-chain Fv(s) anti-peptide antibodies from a Phage display library (CAT, Melbourne, UK) (8), their calculated GRAVY indexes and their ability to enrich peptides from crude tryptic digests for direct MALDI-MS detection from hydrogel arrays (8). For comparison, the same table shows antigenic scores generated using EMBOSS *Antigenic* prediction tool (liv.bmc.uu.se/cgi-bin/emboss/antigenic). There is a clear correlation between the MALDI-MS detection (directly from the immobilised antibodies) and the peptides' hydrophilicity; but no correlation is observed between MALDI detection and the peptide antigenic scores.

12. When a Universal Covalent surface is used to covalently immobilise protein via an abstractable hydrogen using UV illumination, a calibration step is required to calibrate the UV exposure. Typically, a set of UV-sensitive calibration

labels will be supplied with the pack of Universal Covalent Plates or Strips for that purpose.

13. The required Library dilution will depend on the Library complexity and the phage titer. One library may be sufficient for panning against more than one target. We routinely diluted one Ph.D.-C7C™ Phage Display library in 300 μl of 0.1 mg/ml BSA (in 0.1% TBST) and used 50 μl of the obtained dilution per target, to screen six targets in one experiment.
14. Incubating the LB media with *E. coli* cultures gives best results for phage amplification when it is incubated for 2 h. If $OD_{600}<0.1$ after 2 h, the culture has to be incubated further until the absorbance exceeds 0.1. *E. coli* grown on an LB-tet plate and kept at +4°C in the dark will be alive for 1 week, so new *E. coli* needs to be grown every week. Fresh *E. coli* will reach the mid-log phase much faster (~2 h) compared to the older (end of the week) cultures, which may take up to ~4 h to reach mid-log phase.
15. Use the amount of washing buffer sufficient to completely fill the wells. It is important to make sure here that all unbound phages are washed off. We found that ten changes of the washing buffer are required to completely remove the unbound phages. The total washing time must be kept short.
16. Eluting the bound phages by adding large excess of the target protein appears to provide better results. To be most certain that all the bound phages have eluted, both steps can be done one after another. For example, start by eluting the phage with the free target protein and then elute with low pH. In our experience, specific elution was successful with the most of the protein targets tested so far. The non-specific eluates may be kept as backups, in case the panning has to be repeated.
17. It may be difficult to predict even approximately the concentration of the eluted phages. Therefore a set of different dilutions ranging 10^2–10^8 should be made. For dilution purposes, one may assume that the concentration of the eluted phage will be 1/100,000th of that in the starting material (i.e. of the phage display library in the first panning round, or of the amplified phage solutions in further rounds).
18. Using small volumes of *E. coli* (2 ml) yields the best amplification (~10^6 fold). Using larger volumes of *E. coli* cultures will yield poorer results.
19. In order for the phage particles to precipitate, they need to be incubated with PEG/NaCl for at least overnight or

longer. One PEG/NaCl precipitation is usually sufficient to precipitate all or most of the phages and does not have to be repeated. However, if the phage titer is low the precipitation may need to be repeated (repeat the overnight incubation at +4°C and centrifugation steps).

20. The phage pellet is not necessarily visible after the centrifugation and removal of the PEG/NaCl supernatant. However, if precipitate is visible, this would indicate the total phage content of above $\sim 10^5$ pfu.
21. Although the kit manual recommends to use 10^9–10^{11} pfu of the amplified phage for the subsequent screening rounds, the phage titers as low as 10^5 pfu have proven to work well.
22. The LB/IPTG/X-gal plates need to be pre-warmed at least 1 h before the agarose is being added to prevent the agarose from cooling too quickly and forming lumps on the agar surface. The latter would make the titering results inaccurate.
23. Keeping 3 ml of agarose top in 15-ml Falcon tubes makes sure that the agarose top does not solidify quickly after being taken out of the water bath. Other tubes, such as 50-ml Falcon tubes cool down quicker and are therefore not suitable.
24. This amount (~100 μl) should be sufficient for more than one assay, but much would depend on the volume of the assay chosen by the user.
25. The pooled serum will be used for fluorescent labelling and as a reference sample in a competitive binding assay. We first make a pooled sample and then proteolytically digest it. Alternatively, individually digested samples can be pooled after the proteolysis.
26. It may be assumed that total serum protein concentration is below 10%, hence 100 μl of serum should not contain more than 10 mg protein. Hence 0.2–0.5 mg Trypsin should be added.
27. PMSF will inactivate Trypsin irreversibly. PMSF will hydrolyse in water, especially at high pH, and may not work at high salt concentrations, so if in doubt, samples should be diluted and the pH shall be adjusted to pH7 prior to adding PMSF. Alternatively, trypsin may be inactivated by boiling. However, the high total protein concentration in the sample could result in the formation of protein precipitate which will complicate the extraction of peptides.
28. Crude Tryptic digests may be used for affinity assays with or without additional purification (as long as Trypsin is

inactivated). Fluorescently labelled peptides must be purified from the unincorporated fluorescent molecules. We use SEC on Sephadex® G-25 to separate the labelled peptides from both Trypsin and the unincorporated RITC. The same procedure can be applied to unlabelled tryptic digests.

29. Dilute Trypsin similarly to the dilution used in **Section 3.3.1** (Step 3). Inject the same volume as the volume of the peptide sample to be purified, i.e. ~220 μl, obtained in **Section 3.3.1** (Step 5).
30. The Trypsin calibration sample may be spiked with RITC. Elution can then be monitored by measuring fluorescence on-line or off-line. Inevitably some Trypsin will be labelled but some Rhodamine will remain unincorporated, resulting in that both Trypsin peak and the small-molecule fraction (Rhodamine) will be identified. The gap between the two peaks will determine the elution window for the peptides.
31. Ensure that the column is thoroughly washed and equilibrated with the running buffer after each Trypsin run.
32. The collected eluates may be hand-spotted and scanned for fluorescence to more accurately determine the start and the end of the peptide fraction.
33. Irrespective of the type of spotting instrument used (even if using a hand-held "MicroCaster" spotter, Whatman/Schleicher and Schuell), similar key principles have to be followed:
    i. Spotting should be done at least in triplicate for each individual antibody. The number of replicates is usually not a limiting factor (hundreds or thousands of spots can be made on each array), we found that having six replicates is sufficient in most cases.
    ii. Careful consideration must be given to the array layout: replicates should be spread over the whole array area to minimise staining and scanning artefacts. Our instrument (Flexys robotic spotter) produces blocks of densely arranged spots (grids, having from 5 × 5 to 12 × 12 spots each) whilst each grid is well separated from each other. In such a case each grid may contain only a single copy of any antibody, but the patterns should be replicated at least three (better six) times and be spread over the whole array area.
    iii. Relevant negative controls must be included. For example, if polyclonal rabbit anti-peptide antibodies are used, pre-immunisation sera or just total rabbit IgGs would make a suitable negative control. IgG concentration should be ideally the same as in other (specific)

antibody samples and at least the same number of replicates should be made. These will provide an important reference point for the data analysis; any errors in determining the non-specific background may affect quantification.

iv. Reference spots (fluorescently labelled protein) should be added to each array, we have at least one reference spot per grid of spots. These are necessary for signal normalisation during scanning and for pin calibration (*see* Step 2 of the **Section 3.3.3).**

v. Coloured spots should be added to ease array handling. These can be e.g. Coomassie Brilliant Blue or Coomassie-stained protein. These will help to identify the correct membrane surface, distinguish front from the back of the membrane and identify array borders.

vi. If using contact spotting, pins should be either matched or calibrated. These issues are addressed in Step 2 of the **Section 3.3.3.**

34. Pin washing and reconditioning is very important for the avoidance of carry-over contaminations and for achieving high reproducibility of spotting. Pin washing procedures and buffers differ significantly from DNA gridding protocols.

35. We use positively charged nylon. Other membranes such as supported nitrocellulose membranes or immobilised membranes may also be used. Using the unsupported nitrocellulose membranes should be avoided (very fragile nature of nitrocellulose makes it nearly impossible to handle). Any tape can be used. Having some overhanging tape facilitates handling of the membrane strips.

36. If a large number of pins is available to the user, the simplest way would be to select those which result in the identical efficiency of protein transfer from the microwell plates to the membrane (array). If this is not possible, pins should be calibrated (by measuring the fluorescence in each spot) from multiple replicates and the values should be taken into the account when interpreting the main assay results. Alternatively, calibration controls (fluorescence reference spots) should be included for each individual pin when spotting the antibodies.

37. Multiple transfers should be made for each spot (i.e. the material spotted repeatedly onto the same spot on the membrane). This will dramatically increase the reproducibility of antibody transfer and increase the amount of the spotted antibodies (leading to the stronger and more reproducible signals and lesser variability between spots).

We routinely use between 6 and 10 transfers per spot. Further increases are counterproductive as the procedure becomes very long and sample evaporation becomes an issue.

38. High humidity should be maintained inside the robot whilst spotting, especially for longer runs.
39. When using contact spotting, the volume transferred by the pins will depend on many parameters, such as sample viscosity, surface tension, cleanliness of the pins, contact time and the material and porosity of the membrane. These are difficult to predict but easy to measure. We typically have values of ~20 nl per single transfer per pin.
40. Making small batches of arrays (up to 10 arrays per batch) works best in our hands. Increasing the number of arrays further increases variations in the efficiency of transfer. This is likely due to the built up of dry residue on the pins, which causes the changes. As a rule keep the total number of transfers between pin washes below ~50.
41. Because protein cross-linking with formaldehyde occurs slowly, long incubation time is necessary. This will also ensure better reproducibility of the cross-linking. Blocking the unreacted groups with glycine or Tris buffer is optional; we found no clear evidence for including this step, perhaps because blocking might be accomplished during the subsequent steps during incubation of the membranes in the blocking and assay buffers containing high concentration of BSA.
42. Ensure that membranes do not adhere to each other, otherwise blocking may be incomplete. Ideally, block individual membranes in separate vessels: 15-ml Falcon tubes or flat-bottom scintillation tubes or similar work well.
43. Because of the competitive nature of the assay, higher concentration of the unlabelled peptide (test sample) will yield weaker fluorescent staining (higher degree of displacement of the labelled reference).
44. The protocol described here is most suitable for running a number of different affinity assays and for relative quantification of the peptide levels. The pooled serum sample will serve as a good reference sample. Alternatively any one of the samples can be used, e.g. any normal serum sample. The concentration (or the dilution) of the unlabelled proteolytic peptides should be approximately equivalent to the concentration of pooled labelled peptides. This will provide the most accurate measurements. Before running large series, it is worth running a pilot experiment to check that addition of the unlabelled test sample does not reduce the

fluorescent signal more than twice. Use two identical slides, make the assay mixture for two arrays, but only add unlabelled serum to one of the arrays (use equivalent volume of 9% BSA in PBS for the other array).

45. We use 50-ml Falcon tubes for washes. For convenience and to avoid handling mistakes, we use sets of three tubes for each array, filled with 50 ml of the washing buffer. The membranes are transferred from one flask to another at preset intervals. Optionally membranes can be rinsed in water prior to the next step.
46. It may take up to an hour to dry the membranes completely. The filters may be left to dry overnight.
47. In competitive assays a higher readout would indicate lower competition for the immobilised binding site from the unlabelled sample and therefore lower concentration of the competing unlabelled peptide. Lower fluorescence would indicate increased competition for binding sites (higher concentration of the matching peptide in the test sample).
48. The slides have to be dried completely to achieve better attachment of the hydrogel.
49. If the specified gaskets (1.5 × 1.6 cm) are used, two can be fitted on a single microscope slide. This adds the advantage of running binding and displacement assays for the same target on the same slide.
50. The volume of ammonium persulfate has to be adjusted experimentally, to allow sufficient handling time yet to ensure fast polymerisation (within ~30 min). Making acrylamide hydrogels is very similar to making SDS-PAGE gels, except that no SDS should be present. Pre-made Acrylamide:*N*,*N*′-Methylenebisacrylamide mixtures can be used. Glycerol may be added to the gel up to 40% final concentration. It improves mechanical properties of the gel, aids handling but requires longer washing times and does not significantly improve the binding assays to justify its use. However, if photochemical polymerisation is used instead of the chemical (TEMED/persulphate), the addition of glycerol is beneficial (26).
51. It is important to wash the hydrogel pads thoroughly in order to remove any unpolymerised acrylamide.
52. The slides can be dried in an incubator (~20–25°C), but the drying time should not exceed 10 min.
53. Any protein or peptide containing free amino groups could be immobilised. Importantly, the target should not be in Tris buffer (otherwise, the buffer should be changed e.g. to

phosphate buffer using SEC0. Micro Bio-Spin 30 Columns from BioRad are suitable for desalting ~30–70 μl samples).

54. If robotic spotting is sought, follow **Section 3.3.3.**
55. Normally three spots for each of the target protein/antibody, negative and/or positive controls are sufficient. The gaskets used (1.5 × 1.6 cm) allow up to 4 × 4 hand spots but significantly higher number of spots if robotic arrayer is used (up to 30 × 30 of 250 μm spots).
56. The blocking step is not necessary for hydrogels (unlike membrane-based blots and arrays), but adding ~0.01% BSA to the Hydrogel assay buffer (PBST) may help to further reduce any background, especially if home-made hydrogels are used. No BSA shall be used if MALDI-MS detection is sought.
57. In our experience peptide assays with fluorescent detection on hydrogels often outperform ELISA-based assays despite the lack of signal amplification. We attribute this to the advantages of the hydrogel 3D matrix and target protein immobilisation.
58. The final concentration of sulfhydryl groups in the peptide labelling reaction should be below 5 mM.
59. The fluorescence dye, NIR-664-iodoacetamide, labels peptides through cysteine side chains. Final concentration of TBP in the sample should be below 5 mM, but TBP should be in molar excess to the sulfhydryl groups.
60. If air becomes trapped in the syringe during loading, dilute the gel medium slightly. The volume of the settled gel after spinning the columns should be no less than 0.7 ml.
61. At least two samples should be assayed, so relative concentrations of the assayed peptides can be compared between the two samples, or between one unknown sample and one known or pooled reference sample. Labelled peptides' concentrations may be high, ideally above their binding K_D values.
62. One assay mixture should contain only labelled peptides, but no unlabelled peptide should be added. Another assay mixture (displacement assay) should also contain a 100-fold excess of the unlabelled peptide. The unlabelled and labelled peptides should be mixed prior to the incubation with the target protein spotted on hydrogels.
63. The sample holder may need to be modified to accommodate the hydrogel slides. The latter should be cast on silicon wafers and be mounted on standard MALDI plates.

References

1. Guschin, D., Yershov, G., Zaslavsky, A., Gemmell, A., Shick, V., Proudnikov, D., Arenkov, P. and Mirzabekov, A. (1997) Manual manufacturing of oligonucleotide, DNA, and protein microchips. *Anal. Biochem.* **250**, 203–211.
2. Vasiliskov, A.V., Timofeev, E.N., Surzhikov, S.A., Drobyshev, A.L., Shick, V.V. and Mirzabekov, A.D. (1999) Fabrication of microarray of gel-immobilized compounds on a chip by copolymerization. *Biotechniques* **27**, 592–594.
3. Lueking, A., Horn, M., Eickhoff, H., Bussow, K., Lehrach, H. and Walter, G. (1999) Protein microarrays for gene expression and antibody screening. *Anal. Biochem.* **270**, 103–111.
4. MacBeath, G. and Schreiber, S.L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science* **289**, 1760–1763.
5. de Wildt, R.M., Mundy, C.R., Gorick, B.D. and Tomlinson, I.M. (2000) Antibody arrays for high-throughput screening of antibody–antigen interactions. *Nat. Biotechnol.* **18**, 989–994.
6. Arenkov, P., Kukhtin, A., Gemmell, A., Voloshchuk, S., Chupeeva, V. and Mirzabekov, A. (2000) Protein microchips: use for immunoassay and enzymatic reactions. *Anal. Biochem.* **278**, 123–131.
7. Soloviev, M., Barry, R. and Terrett, J. (2004) Chip based proteomics technology, in Molecular Analysis and Genome Discovery (Rapley R., Harbron S., ed.), John Wiley and Sons Ltd, New York, pp. 217–249.
8. Scrivener, E., Barry, R., Platt, A., Calvert, R., Masih, G., Hextall, P., Soloviev, M. and Terrett, J. (2003) Peptidomics: a new approach to affinity protein microarrays. *Proteomics* **3**, 122–128.
9. Barry, R. and Soloviev, M. (2004) Peptidomics approaches to proteomics research. *BIOforum Eur.* **1**, 34–36.
10. Soloviev, M. and Finch, P. (2005) Peptidomics, current status. *J. Chromatogr. B.* **815**, 11–24.
11. Soloviev, M. and Terrett, J. (2005) Practical guide to protein microarrays: assay systems, methods and algorithms, in Protein Microarrays (Schena M., ed.), Jones and Bartlett Publishers, Boston, MA, pp. 43–56.
12. Clarckson, T., Hoogenboom, H.R., Griffiths, A.D. and Winter, G. (1991) Making antibody fragments using phage display libraries. *Nature* **352**, 624–628.
13. Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Pluckthun, A. and Virnekas, B. (2000) Fully synthetic human combinatorial antibody libraries (HuCAL). *J. Mol. Biol.* **296**, 57–86.
14. Binz, K.H. and Plückthun, A. (2005) Engineered proteins as specific binding reagents. *Curr. Opin. Biotechnol.* **16**, 459–469.
15. Messmer, B.T., Benham, C.J. and Thaler, D.S. (2000) Sequential determination of ligands binding to discrete components in heterogeneous mixtures by iterative panning and blocking (IPAB). *J. Mol. Biol.* **296**, 821–832.
16. Kay, B.K., Kasanov, J. and Yamabhai, M. (2001) Screening phage-displayed combinatorial peptide libraries. *Methods* **24**, 240–246.
17. Noren, K.A. and Noren, C.J. (2001) Construction of high-complexity combinatorial phage display peptide libraries. *Methods* **23**, 169–178.
18. Smith, G.P. (1985) Filamentous fusion phage: novel expression vectors that display clones antigens on the virion surface. *Science* **228**, 1315–1317.
19. Ditzel, H.J., Barbas, S.M., Barbas, C.F., 3rd and Burton, D.R. (1994) The nature of the autoimmune antibody repertoire in human immunodeficiency virus type 1 infection. *Proc. Natl Acad. Sci. USA.* **91**, 3710–3714.
20. Scott, J.K. and Smith, G.P. (1990) Searching for peptide ligands with an epitope library. *Science* **249**, 386–390.
21. Cwirla, S.E., Peters, E.A., Barrett, R.W. and Dower, W.J. (1990) Peptides on phage: a vast library of peptides for identifying ligands. *Proc. Natl Acad. Sci. USA.* **87**, 6378–6382.
22. Devlin, J.J., Panganiban, L.C. and Devlin, P.E. (1990) Random peptide libraries: a source of specific protein binding molecules. *Science*. **249**, 404–406.
23. Barry, R., Scrivener, E., Soloviev, M. and Terrett, J. (2002) Chip-based proteomics technologies. *Int. Genomic Proteomic Technol.* Feb, 14–22.
24. Quanjun, L., Wei, Z., Hong, W., Jujun, Z., Yujie, Z. and Zuhong, L. (2002) Detection YMDD mutation of HBV with a polyacrylamide film immobilized molecular beacon array. 8th International conference on Electronic Materials, IUMRS-ICEM 2002, Xi'an, China, pp. 387–39.2
25. Brueggemeier, S.B., Kron, S.J. and Palecek, S.P. (2004) Use of protein-acrylamide copolymer hydrogels for measuring protein

concentration and activity. *Anal. Biochem.* **329**, 180–189.
26. Burnham, M.R., Turner, J.N., Szarowski, D. and Martin, D.L. (2006) Biological functionalization and surface micropatterning of polyacrylamide hydrogels. *Biomaterials* **27**, 5883–5891.
27. Salata, O. (2007) Nanotechnology in therapeutics: hydrogels and beyond. *J. Nanobiotechnol.* **5**, 5.
28. Jacchetti, E., Emilitri, E., Rodighiero, S., Indrieri, M., Gianfelice, A., Lenardi, C., Podestà, A., Ranucci, E., Ferruti, P. and Milani, P. (2008) Biomimetic poly(amidoamine) hydrogels as synthetic materials for cell culture. *J. Nanobiotechnol.* **6**, 14.
29. Tang, J. and Xiao, P. (2009) Polymerizing immobilization of acrylamide-modified nucleic acids and its application. *Biosens. Bioelectron.* **24**, 1817–1824.
30. Barbas, C.F., 3rd, Kang, A.S., Lerner, R.A. and Benkovic, S.J. (1991) Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proc. Natl Acad. Sci. USA.* **88**, 7978–7982.
31. Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P. and Winter, G. (1991) Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic Acids Res.* **19**, 4133–4137.
32. Griffiths, A.D., Williams, S.C., Hartley, O., Tomlinson, I.M., Waterhouse, P., Crosby, W.L., Kontermann, R.E., Jones, P.T., Low, N.M., Allison, T.J., Prospero, T.D., Hoogenboom, H.R., Nissim, A., Cox, J.P.L., Harrison, J.L., Zaccolo, M., Gherardy, E. and Winter, G. (1994) Isolation of high affinity human antibodies directly from large synthetic repertoires. *EMBO J.* **13**, 3245–3260.
33. Vaughan, T.J., Williams, A.J., Pritchard, K., Osbourn, J.K., Pope, A.R., Earnshaw, J.C., McCafferty, J., Hodits, R.A., Wilton, J. and Johnson, K.S. (1996) Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nat. Biotechnol.* **14**, 309–314.
34. Rai, A.J., Gelfand, C.A., Haywood, B.C., Warunek, D.J., Yi, J., Schuchard, M.D., Mehigh, R.J., Cockrill, S.L., Scott, G.B., Tammen, H., Schulz-Knappe, P., Speicher, D.W., Vitzthum, F., Haab, B.B., Siest, G. and Chan, D.W. (2005) HUPO Plasma Proteome Project specimen collection and handling: towards the standardization of parameters for plasma proteome samples. *Proteomics* **5**, 3262–3277.
35. Barry, R., Diggle, T., Terrett, J. and Soloviev, M. (2003) Competitive assay formats for high-throughput affinity arrays. *J. Biomol. Screen.* **8**, 257–263.
36. Heo, J. and Crooks, R.M. (2005) Microfluidic biosensor based on an array of hydrogel-entrapped enzymes. *Anal. Chem.* **77**, 6843–6851.
37. Zubtsov, D.A., Savvateeva, E.N., Rubina, A.Y., Pan'kov, S.V., Konovalova, E.V., Moiseeva, O.V., Chechetkin, V.R. and Zasedatelev, A.S. (2007) Comparison of surface and hydrogel-based protein microchips. *Anal. Biochem.* **368**, 205–213.
38. Thomas, G., El-Giar, E.M., Locascio, L.E. and Tarlov, M.J. (2006) Hydrogel-immobilized antibodies for microfluidic immunoassays: hydrogel immunoassay. *Methods Mol. Biol.* **321**, 83–95.
39. Parker, L.L., Brueggemeier, S.B., Rhee, W.J., Wu, D., Kent, S.B., Kron, S.J. and Palecek, S.P. (2006) Photocleavable peptide hydrogel arrays for MALDI-TOF analysis of kinase activity. *Analyst* **131**, 1097–1104.
40. Rubina, A.Y., Dyukova, V.I., Dementieva, E.I., Stomakhin, A.A., Nesmeyanov, V.A., Grishin, E.V. and Zasedatelev, A.S. (2005) Quantitative immunoassay of biotoxins on hydrogel-based protein microchips. *Anal. Biochem.* **340**, 317–329.
41. Tanaka, H., Hanasaki, M., Isojima, T., Takeuchi, H., Shiroya, T. and Kawaguchi, H. (2009) Enhancement of sensitivity of SPR protein microarray using a novel 3D protein immobilization. *Colloids Surf. B Bioint.* **70**, 259–265.
42. Dyukova, V.I., Dementieva, E.I., Zubtsov, D.A., Galanina, O.E., Bovin, N.V. and Rubina, A.Y. (2005) Hydrogel glycan microarrays. *Anal. Biochem.* **347**, 94–105.
43. Charles, P.T., Taitt, C.R., Goldman, E.R., Rangasammy, J.G. and Stenger, D.A. (2004) Immobilization strategy and characterization of hydrogel-based thin films for interrogation of ligand binding with staphylococcal enterotoxin B (SEB) in a protein microarray format. *Langmuir* **20**, 270–272.
44. Hofmann, H.J. and Hädge, D. (1987) On the theoretical prediction of protein antigenic determinants from amino acid sequences. *Biomed. Biochim. Acta* **46**, 855–866.
45. Pellequer, J.L. and Westhof, E. (1993) PREDITOP: a program for antigenicity prediction. *J. Mol. Graph.* **11**, 204–210.
46. Ponomarenko, J.V. and Bourne, P.E. (2007) Antibody–protein interactions: benchmark datasets and prediction tools evaluation. *BMC Struct. Biol.* **7**, 64.

47. Flower, D.R. (2007) Immunoinformatics and the in silico prediction of immunogenicity. An introduction. *Methods Mol. Biol.* **409**, 1–15.
48. El-Manzalawy, Y., Dobbs, D. and Honavar, V. (2008) Predicting linear B-cell epitopes using string kernels. *J. Mol. Recognit.* **21**, 243–255.
49. Chen, J., Liu, H., Yang, J. and Chou, K.C. (2007) Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. *Amino Acids* **33**, 423–428.
50. Ponomarenko, J., Bui, H.H., Li, W., Fusseder, N., Bourne, P.E., Sette, A. and Peters, B. (2008) ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics* **9**, 514.
51. Van Regenmortel, M.H. and Pellequer, J.L. (1994) Predicting antigenic determinants in proteins: looking for unidimensional solutions to a three-dimensional problem?. *Pept. Res.* **7**, 224–228.
52. Hopp, T.P. and Woods, K.R. (1981) Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl Acad. Sci. USA.* **78**, 3824–3828.
53. Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–312.
54. Hopp, T.P. (1993) Retrospective: 12 years of antigenic determinant predictions, and more. *Pept. Res.* **6**, 183–190.

Chapter 24

In Situ Biosynthesis of Peptide Arrays

Mingyue He and Oda Stoevesandt

Abstract

Polypeptide and protein arrays enable high-throughput screening capabilities for studying molecular interactions and profiling of biomarkers, and provide a powerful functional screening tool for peptidomics. To overcome the limitations of conventional arraying methods, we have exploited cell-free systems for generating arrays of polypeptides by direct on-chip biosynthesis from DNA templates. Here we describe two protocols: (i) Protein In Situ Array (PISA), which allows the generation of polypeptide arrays in a single reaction by spotting cell-free lysate together with PCR DNA on a glass surface pre-coated with a capturing reagent, and (ii) DNA Array to Protein Array (DAPA), which is capable of producing multiple copies of a polypeptide array from a single DNA array template. The main advantage of these methods is in using an in vitro coupled transcription and translation system which circumvents the need to synthesise and purify individual polypeptides. Our methods allow making polypeptide arrays using amplified linear DNA fragments.

Key words: Peptide array, cell-free protein synthesis.

1. Introduction

Peptidomics requires technologies for high-throughput, multiplexed interaction assays. Peptide arrays can be used for simultaneous analysis of a large number of protein–peptide interactions and protein signalling pathways in a time- and cost-effective manner (1). One of the major bottlenecks in making peptide arrays is ensuring the supply of a large number of peptides for immobilisation. Chemical synthesis of peptides remains an expensive option, while expression and purification of large numbers of polypeptides or proteins in heterologous hosts is a time-consuming process. Cell-free synthesis can be used to overcome these problems

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_24,

(2–8). It directs the synthesis of polypeptides and proteins from added PCR DNA templates without the need for bacterial cloning, providing a rapid and economic means for conversion of genetic information into polypeptides. DNA fragments encoding peptides can be synthesised routinely, rapidly and inexpensively.

By coupling cell-free synthesis and in situ protein immobilisation on the array surface, we have developed two cell-free methods, termed *PISA* and *DAPA*, for making protein arrays on demand directly from DNA molecules (2, 4, 8). These approaches eliminate the need for separate expression, purification and printing of individual proteins, and help to avoid the risk of deterioration in protein function during storage, as protein arrays can be produced in a matter of hours immediately prior to their application, as shown in **Fig. 24.1**. Our methods can also be used for arraying functional full-length proteins (5).

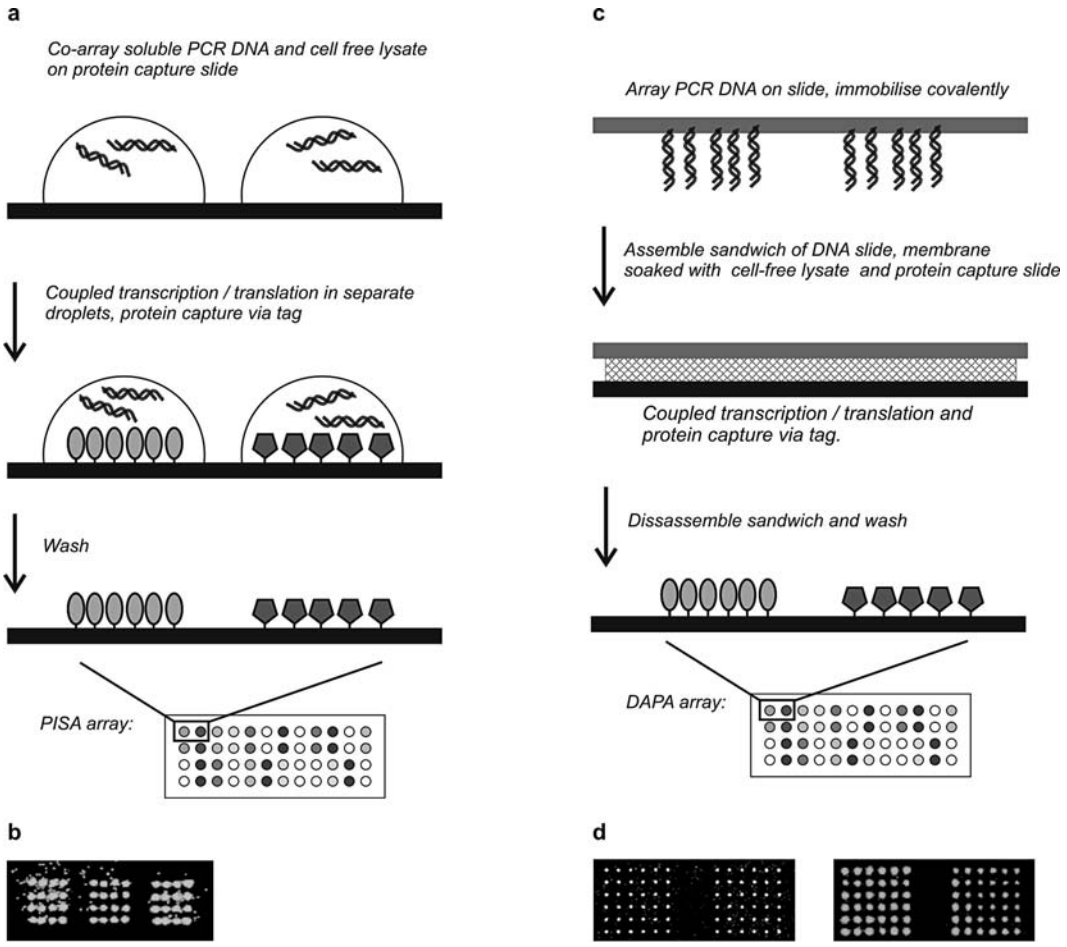


Fig. 24.1. Principle of PISA and DAPA. (**a**) Scheme of Protein In Situ Array procedure (*PISA*). (**b**) Example of a PISA protein array. (**c**) Scheme of DNA Array to Protein Array procedure (*DAPA*). (**d**) Example of a template DNA array (*left*) and the synthesised DAPA protein array (*right*).

2. Materials

2.1. Cell-Free Expression

1. Primers for the amplification of DNA template for use with Rabbit Reticulocyte Lysate System:
   - "T7/back(R)": 5′-GCAGC*TAATACGACTCACTATA G*GAA CAGACCACCATG-3′, an upstream primer containing T7 promoter (italics) and Kozak sequence (underlined) and the start codon ATG (doubly underlined).
   - "G/back(R)": 5′-TAGGAACAGACCACCATG $(N)_{15–25}$-3′, an upstream primer for PCR amplification of target genes. It contains a sequence overlapping with "T7/back (R)" primer and 15–25 nucleotides (N) matching the 5′ sequence of the gene of interest (*see* **Note 1**).
   - "G/for":5′-CACCGCCTCTAGAGCG$(N)_{15–25}$-3′, a downstream primer for PCR amplification of target genes. It contains a sequence overlapping with a PCR fragment encoding a C-terminal region of the expression construct and 15–25 nucleotides complementary to the 3′ region of a target gene (*see* **Note 1**).
2. Primers for the amplification of DNA template for use with *E. coli* S30 Extracts:
   - "RTST7/back": 5′-GATCTCGATCCCGCG-3′, an upstream primer for the amplification of T7 fragment (in combination with the "RTST7/for" primer or a full-length construct in combination with "T-term/for" primer).
   - "RTST7/for": 5′-CATGGTATATCTCCTTCTTAAAG-3′, a downstream primer for the amplification of T7 fragment in combination with the "RTST7/back" primer.
   - "G/back(E)": 5′-CTTTAAGAAGGAGATATACCATG $(N)_{15–25}$-3′, an upstream primer for the amplification of target genes. It contains a sequence overlapping with the T7 fragment and 15–25 nucleotides from the 5′ sequence of the gene of interest (*see* **Note 1**).
   - "G/for": 5′-CACCGCCTCTAGAGCG$(N)_{15–25}$-3′, a downstream primer for the amplification of a target gene. It contains a sequence overlapping with a PCR fragment encoding a C-terminal region of the expression construct and 15–25 nucleotides complementary to the 3′ region of the target gene (*see* **Note 1**).

3. Primers for the amplification of the C-terminal region of the expression construct:
   – "Linker-tag/back": 5′-GCTCTAGAGGCGGTGGC-3′, an upstream primer for the amplification of a termination region in combination with the "T-term/for" primer.
   – "T-term/for": 5′-TCCGGATATAGTTCCTCC-3′, a downstream primer for the amplification of the termination region in combination with the "Linker-tag/back" primer or the amplification of the full-length construct in combination with one of the "RTST7/back" or "T7/back(R)" primers.
4. Cy5 and NH_2-modified primers
   – "Cy5-RTST7/back": 5′-Cy5-GATCTCGATCCCGCG-3′, an upstream primer for the amplification of the full-length construct in combination with the "NH_2-Tterm/F" primer (*see* **Note 2**).
   – "NH_2-T-term/for": 5′-NH_2-TCCGGATATAGTTCCTCC-3′, a downstream primer for PCR generation of the full-length construct in combination with "Cy5-RTST7/B" primer (*see* **Note 3**).
5. T7 regulatory fragment for *E. coli* cell-free expression: 5′-GATCTCGATCCCGCGAAATTAATACGACTCACTATA GGGAGACCACAACGGTTT*CCCTCTAGAAATAATTT* TGTTTAACTTTAAGAAGGAGATATACCATG-3′. T7 promoter is underlined; the ribosome binding site (underlined italics) and the start codon ATG (doubly underlined) are indicated (*see* **Note 4)**.
6. C-terminal region regulatory fragment for *E. coli* cell-free expression: 5′-GCTCTAGAggcggtggctctggtggcggttctgg-cggtggcaccggtggcggttctggcggtggcAAACGGGCTGATGC-TGC ACATCACCATCACCATCACTCTAGAGCTTGGC GTCACCCGC CAGTTCGGTGGTCACCACCACCACC ACCACTAATAA$(A)_{28}$CCGCTGAGCAAT*AACTAGCA T-AACCCCTTGGGGGCCTCTAAACGGGTCTTGAGGG GTTTTTTG*CTGAAAGGAGGAACTATATCCGGA-3′. This fragment encodes a C-terminal region, composed of a flexible 19 amino acid linker (lower case), a double $(His)_6$ tag (underlined), two consecutive stop codons (doubly underlined), a poly(A) tail and a transcription termination region (shown in italics) (*see* **Note 5)**.
7. Cell-free systems, molecular biology reagents and kits: Rabbit Reticulocyte T&T T7 Quick for PCR DNA (Promega, UK); RTS100 *E. coli* HY (Roche Molecular Biochemicals, UK); GenElute™ Gel Extraction kit (Sigma, UK); GenElute™ PCR Clean-Up kit (Sigma, UK);

8. Arrays and slides: Nexterion^{TM} slide E (epoxysilane coated, Schott Nexterion, UK); Durapore 0.22-μm membrane filters (Millipore, UK); Ni-NTA-coated microscope slides (Xenopore, USA)
9. PBS: Phosphate-buffered saline, pH 7.4
10. Wash buffer 1: PBS, 300 mM NaCl, 20 mM imidazole, pH 8.0
11. Wash buffer 2: PBS, 0.05% Tween20
12. 6× spotting buffer: 300 mM sodium phosphate, pH 8.5
13. Saturated NaCl solution: 30% NaCl, boil and cool down to make saturated solution
14. Quenching buffer: 0.1 M Tris-HCl, pH 9.0. Add ethanolamine to a final concentration of 50 mM immediately before use.
15. Other buffers: 100 mM magnesium acetate; 0.1% Tween-20 in H_2O; 1 mM HCl; 100 mM KCl.

3. Methods

3.1. Amplification of cDNA Constructs for Cell-Free Expression

PCR-amplified DNA fragments make suitable templates especially for short polypeptide synthesis using cell-free systems. The PCR construct should contain the essential regulatory elements for transcription and translation. These include a promoter (usually T7), translation initiation site and sequences for transcription and translation termination. The translation initiation site for eukaryotic systems is different to that for prokaryotic *E. coli* S30 extracts. A poly(A) tail should also be included after the stop codon. For in situ immobilisation of polypeptides on a surface, an affinity tag sequence should be placed at either N- or C-terminus of the polypeptide (*see* **Note 6**).

To simplify the generation of templates for cell-free expression, these common sequence elements should be made and cloned. Plasmids make convenient templates for PCR amplification. **Figure 24.2 s**ummarises the process of generating the DNA fragments. The T7 promoter and translation initiation site should be present upstream of the target cDNA. These can be introduced either by using a long primer containing the required sequences (an approach most suitable for the rabbit reticulocyte system) or by using a PCR-amplified DNA derived from the cloned T7 fragment (the approach more suitable for the *E. coli* expression system). A DNA fragment encoding C-terminal immobilisation tag and containing transcription and translation

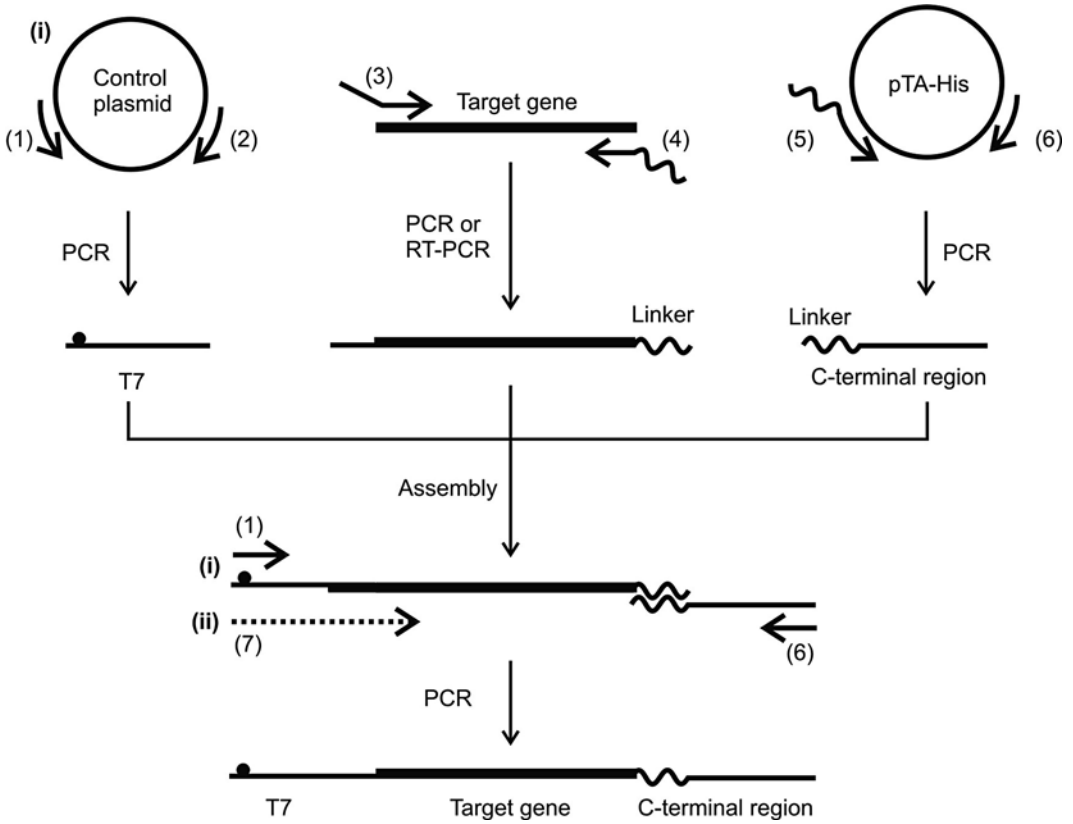


Fig. 24.2. PCR strategy for the generation of constructs for cell-free expression. The primers used are (1) "RTST7/back", (2) "RTST7/for", (3) "G/back", (4) "G/for", (5) "Linker-tag/back", (6) "T-term/for", (7) "T7/back (R)". (i) PCR amplification strategy for *E. coli* cell-free system. (ii) PCR amplification strategy for rabbit reticulocyte lysate.

termination sequences should be placed downstream of the target DNA.

3.1.1. Amplification of Target Genes, the C-Terminal Region and the T7 Domain (see Note 7)

1. Set up a standard 50 μl PCR reaction using e.g. Qiagen Taq system. Use "G/for" primer together with either "G/back(R)" primer (for rabbit reticulocyte lysate system) or with "G/back(E)" primer (for *E. coli* cell-free system). Carry out thermal cycling for 30 cycles (94°C for 30 s, 54°C for 30 s and 72°C for 1 min).
2. Set up standard 50 μl PCR reaction using e.g. Qiagen Taq system. Use primers "Linker-tag/back" and "T-term/for" to amplify the C-terminal region (*see* **Section 2**, Step 7). Carry out thermal cycling for 30 cycles (94°C for 30 s, 54°C for 30 s and 72°C for 1 min).
3. For the *E. coli* expression system only, set up a standard 50 μl PCR reaction using e.g. Qiagen Taq system. Use

primers "RTST7/back" and "RTST7/for" to amplify the T7 domain from the control plasmid. Carry out thermal cycling for 30 cycles (94°C for 30 s, 54°C for 30 s and 72°C for 1 min).

4. Analyse the amplified fragments on 1% agarose gel. Isolate the expected fragments using GenElute™ or similar gel extraction kit.

3.1.2. Assembly of the DNA Constructs by PCR

1. For the rabbit reticulocyte lysate system, set up a PCR reaction using e.g. Qiagen Taq system; mix the target gene and the C-terminal region in equimolar ratios (total DNA 50–100 ng), no oligonucleotide primers needed at this stage, total volume 25 μl. Carry out thermal cycling for eight cycles (94°C for 30 s, 54°C for 1 min and 72°C for 1 min) to assemble the fragments (*see* **Note 8**).
2. To further amplify the assembled product, transfer 2 μl of the assembled construct (from the Step 1 above) to another standard PCR reaction mix, add primers "T7/back(R)" and "T-term/for" and amplify for 30 cycles (94°C for 30 s, 54°C for 1 min and 72°C for 1.2 min).
3. For the *E. coli* system, set up a PCR reaction using e.g. Qiagen Taq system; mix T7 domain, target gene and the C-terminal region in equimolar ratios (total DNA 50–100 ng), no oligonucleotide primers needed at this stage, total volume 25 μl. Carry out thermal cycling for 8 cycles (94°C for 30 s, 54°C for 1 min and 72°C for 1 min) to assemble the fragments (*see* **Note 8**).
4. To further amplify the assembled product, transfer 2 μl of the assembled construct (from the Step 3 above) to another standard PCR reaction mix, add primers "RTST7/back" and "T-term/for" and amplify for 30 cycles (94°C for 30 s, 54°C for 1 min and 72°C for 1.2 min).
5. Analyse the amplified fragments on 1% agarose gel. Isolate the expected fragments using GenElute™ or similar gel extraction kit (*see* **Note 9**).
6. Confirm the construct identity by PCR mapping (*see* **Note 10**). The resulting PCR construct, either purified or unpurified, is ready for use for peptide arrays. The construct may be stored at –20°C for at least 6 months.

3.1.3. Assembly of the Fluorescently Labelled DNA Construct for Use with E. coli Cell-Free Expression System

1. Assemble a PCR reaction using e.g. Qiagen Taq system; mix T7 domain, target gene and the C-terminal region in equimolar ratios (total DNA 50–100 ng), no oligonucleotide primers needed at this stage, total volume 25 μl. Carry out thermal cycling for eight cycles (94°C for 30 s,

54°C for 1 min and 72°C for 1 min) to assemble the fragments (*see* **Note 8**).

2. To further amplify the assembled product, transfer 2 μl of the assembled construct (from the Step 3 above) to another standard PCR reaction mix, add primers "Cy5-RTST7/back" and "T-term/for" and amplify for 30 cycles (94°C for 30 s, 54°C for 1 min and 72°C for 1.2 min).
3. Analyse the amplified fragments on 1% agarose gel. Isolate the expected fragments using GenElute™ or similar gel extraction kit (*see* **Note 9**).
4. Measure the concentration and purity of the PCR product by absorption at 260 nm and 280 nm or by gel electrophoresis. DNA concentration of 100 ng/μL is recommended for spotting (*see* **Note 11**).

3.2. In Situ Peptide Arrays on Nickel-Coated Glass Slides

1. To set up T&T reaction using Rabbit Reticulocyte Lysate system, mix the following kit components: Rabbit Reticulocyte Lysate T&T system for PCR DNA (40 μl), 1 mM Methionine (1 μl), 100 mM magnesium acetate (1 μl), assembled cDNA expression construct (50–100 ng), H_2O (to 50 μl final volume) (*see* **Note 12**).
2. To set up T&T reaction using RTS100 *E. coli* HY, mix the following kit components: *E. coli* lysate (12 μl), Reaction mix from the kit (10 μl), Amino acids (12 μl), Methionine (1 μl), Reconstitution buffer (5 μl), assembled cDNA expression construct (50–100 ng), H_2O (to 50 μl final volume) (*see* **Note 13**).
2. Spot the T&T mixture onto a Ni-NTA-coated glass slide (40 nl per spot) (*see* **Note 14**).
3. Incubate the slide in a humidified chamber (*see* **Note 15**) at 30°C for 2 h (*see* **Note 16**).
4. Wash three times with the wash buffer 1 (*see* **Note 17**) or with the wash buffer 2, followed by a final wash with 100 μl PBS, pH 7.4.

3.3. DNA Array to Protein Array

DNA Array to Protein Array (*DAPA*) is achieved using cell-free synthesis of polypeptides within a membrane held between the surfaces of two glass slides. One of the slides carries an array of immobilised PCR molecules, the other slide is coated with a reagent to capture the newly synthesised polypeptides. After synthesis within the membrane, individual polypeptides bind to the capturing surface, creating a polypeptide array with the layout mirroring that of the DNA array. We use epoxysilane-activated slides for DNA immobilisation, *E. coli* cell-free system for polypeptide synthesis, and Ni-NTA-coated slides for capturing His-tagged polypeptides.

3.3.1. Making DNA Array Template on Epoxysilane Slides

1. Add 1 volume of 6× spotting buffer to 5 volumes of the assembled DNA PCR product (**3.1.3.** Step 4).
2. Spot DNA samples on the epoxysilane slide (*see* **Note 18**) with spot-to-spot distances of 1 mm and volumes per spot of 2–3 nl. Incubate spotted slides in a humidified chamber at room temperature for 1 h. (*see* **Note 15**).
3. Incubate slides at 60°C for 30 min.
4. Wash the slides once with 0.1% Tween-20 for 5 min, twice with 1 mM HCl for 2 min, once with 100 mM KCl for 10 min and once with ddH_2O for 1 min (all washes should be performed at room temperature).
5. Quench the remaining epoxy groups by incubating slides in 0.1 M Tris-HCl pH 9.0, 50 mM ethanolamine at 50°C for 15 min. Rinse slides with ddH_2O for 1 min and dry either by pressurised air or by centrifugation at 2000 rpm for 1 min.
6. Scan the slides in a suitable microarray scanner to confirm immobilisation of Cy5-labelled DNA. The slides should stored in the dark at 4°C until use.

3.3.2. Printing Polypeptide Array Using the DNA Array Template

1. Use a slide holder similar to the prototype shown in **Fig. 24.3**.

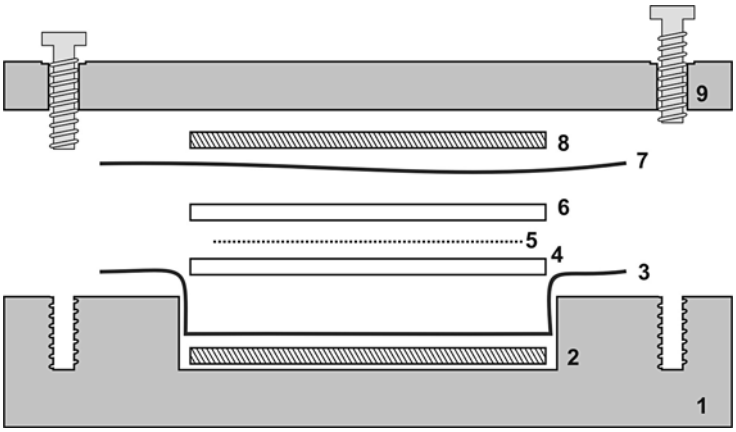


Fig. 24.3. Schematic cross-section of DAPA assembly. The numbering of components is the same as in **Section 3.3.2** (Step 3).

2. Cut a Durapore membrane filter large enough to cover the area of the DNA template array. Prepare *E. coli* cell-free lysate, make 10 μl of the lysate per 1 cm^2 of the membrane area.
3. Assemble the slide holder in the following order (as shown in **Fig. 24.3**):

(i) Put a rubber spacer (**Fig. 24.3**, #2) in the bottom plate (**Fig. 24.3**, #1), followed by a layer of parafilm (**Fig. 24.3**, #3);

(ii) Place a Ni-NTA-coated slide (**Fig. 24.3**, #4) with the capturing surface facing up onto the parafilm (*see* **Note 18**);

(iii) Spread the required volume of *E. coli* cell-free lysate on the surface of the Ni-NTA slide (**Fig. 24.3**, #4), cover with the membrane filter (**Fig. 24.3**, #5) allowing it to soak up the lysate (*see* **Note 19**);

(iv) Position the DNA template slide (**Fig. 24.3**, #6) with DNA surface facing down on the membrane filter. Cover with another layer of parafilm (**Fig. 24.3**, #7) and another rubber spacer (**Fig. 24.3**, #8) (*see* **Note 20**);

(v) Close the slide holder with the top plate (**Fig. 24.3**, #9); ensure even pressure by carefully tightening screws.

4. Incubate the assembled slide holder at 30°C for 2–4 h.
5. Disassemble the slide holder and wash the Ni-NTA slide (peptide arrays) three times with washing buffer 2. At this stage the peptide array is ready for use in downstream applications.
6. Rinse the DNA template slide with ddH_2O, dry and store at 4°C; the DNA array can be used for making more than one peptide array.
7. A standard direct binding immunoassay can be used to detect immobilised polypeptides on the array or for quality control purposes (*see* **Note 21**).

4. Notes

1. Sequence of the fragment marked as "$(N)_{15–25}$" will depend on the particular target gene used and on the position along that sequence and has to be devised by the user.
2. Cy5 fluorescent label allows detection and quantification of the immobilised PCR product.
3. The coupled NH_2 group allows immobilisation of the PCR product on epoxy-activated slides.
4. This fragment can be obtained from the control plasmid included with the RTS100 *E. coli* HY kit (Roche).
5. The encoded double-$(His)_6$ tag has shown an order of magnitude or greater affinity for Ni-NTA modified surfaces compared to a conventional single-$(His)_6$ tag (2, 6).

6. The location of a tag can be at both the N- and C-termini of the polypeptide, although C-terminal immobilisation tags are preferable, as their presence ensures that the entire polypeptide is expressed.
7. The C-terminal region and the T7 domain can be produced in large quantities by PCR and stored at –20°C until use.
8. Alternatively, long oligonucleotides (about 100 bases) encoding peptides can be synthesised and then assembled with the 5′ T7 domain and the C-terminal domain by PCR.
9. If multiple PCR bands are generated, the expected PCR fragment with the correct size should be isolated by gel extraction and used as template for PCR re-amplification. In general, unpurified PCR fragments can be directly used for protein synthesis in cell-free systems. However, if purification is needed, a Sigma GenElute^{TM} PCR Clean-Up kit can be used.
10. A construct can be confirmed by PCR mapping, which is performed by using a combination of primers annealing at different positions along the construct. If all PCR products give the expected size, it suggests the correct construction.
11. If the eluted PCR product is below this range, it can be concentrated in a vacuum centrifuge.
12. Magnesium acetate added to rabbit reticulocyte lysate TNT mixture during translation was found to improve protein expression. We produced a better yield for single-chain antibodies and other protein when additional Mg^{2+} concentrations ranging from 0.5 to 2 mM were included in this system.
13. RTS100 *E. coli* HY can yield 3–25 μg of protein or polypeptide in a 50 μl reaction.
14. The Ni-NTA-coated glass slides are capable of capturing His-tagged polypeptides.
15. A humidified chamber can be prepared using a box containing saturated NaCl solution.
16. Depending on the polypeptide and the planned downstream application, the time can vary. For rabbit reticulocyte lysate T&T system 1–2 h is most suitable, or 1–4 h for RTS *E. coli* HY System.
17. Rabbit reticulocyte lysate contains large amounts of haemoglobin which sometimes binds to Ni-coated slides. More washes may be required to remove haemoglobin from the slides.
18. Mark glass slides and their orientation with a diamond-tipped pen. Any possible glass splinters or dust from the slide surfaces can be removed by using pressurised air.

19. The soaking should take just a few seconds. It is important to avoid drying the cell-free lysate within the membrane filter.
20. The parafilm must form an airtight seal around the slide sandwich (as shown in **Fig. 24.3**) in order to prevent evaporation of cell-free lysate.
21. Fluorescently labelled antibodies or signal amplification, e.g. with horseradish peroxidase/tyramide-Cy3 system can be used with commonly available microarray scanners, most of which are capable of fluorescence detection in the Cy3 and Cy5 range (550 and 650 nm, respectively).

Acknowledgements

We thank Hong Liu for technical assistance. Research at the Babraham Institute is supported by Biotechnology and Biological Sciences Research Council (BBSRC), UK.

References

1. Uttamchandani, M. and Yao, S.Q. (2008) Review: Peptide microarrays: next generation biochips for detection, diagnostics and high-throughput screening. *Curr. Pharm. Des.* **14**, 2428–2438.
2. He, M. and Taussig, M.J.(2001). Single step generation of protein arrays from DNA by cell-free expression and in situ immobilization (PISA method). *Nucleic Acids Res.* **29**, e73.
3. Ramachandran, N., Hainsworth, E., Bhullar, B., Eisenstein, S., Rosen, B., Lau, A.Y., Walter, J.C., and LaBaer, J. (2004) Self-assembling protein microarrays. *Science* **305**, 86–90.
4. Angenendt, P., Kreutzberger, J., Glokler, J., and Hoheisel, J.D. (2006) Generation of high density protein microarrays by cell-free in situ expression of unpurified PCR products. *Mol. Cell. Proteomics* **5**, 1658–1666.
5. He, M. and Taussig, M.J.(2003) DiscernArray™ technology: a cell-free method for the generation of protein arrays from PCR DNA. *J. Immunol. Methods* **274**, 265–270.
6. Khan, F., He, M., and Taussig, M.J. (2006) A double-His tag with high affinity binding for protein immobilisation, purification, and detection on Ni-NTA surfaces. *Anal. Chem.* **78**, 3072–3079
7. He, M., Stoevesandt, O., Palmer, E.A., Khan, F., Ericsson, O., and Taussig, M.J. (2008) Printing protein arrays from DNA arrays. *Nat. Methods* **5**, 175–177.
8. He, M., Stoevesandt, O., and Taussig, M.J. (2008) In situ synthesis of protein arrays. *Curr. Opin. Biotechnol.* **19**, 4–9.

Chapter 25

Bioinformatic Approaches to the Identification of Novel Neuropeptide Precursors

Elke Clynen, Feng Liu, Steven J. Husson, Bart Landuyt, Eisuke Hayakawa, Geert Baggerman, Geert Wets, and Liliane Schoofs

Abstract

With the entire genome sequence of several animals now available, it is becoming possible to identify *in silico* all putative peptides and their precursors in an organism. In this chapter we describe a searching algorithm that can be used to scan the genome for predicted proteins with the structural hallmarks of (neuro)peptide precursors. We also describe how to use search strings such as the presence of a glycine residue as a putative amidation site, dibasic cleavage sites, the presence of a signal peptide, and specific peptide motifs to improve a standard BLAST search and make it suitable for searching (neuro)peptides in EST data. We briefly explain how bioinformatic tools and *in silico* predicted peptide precursor sequences can aid experimental peptide identification with mass spectrometry.

Key words: Bioinformatics, BLAST, expressed sequence tags, Mascot, mass spectrometry, neuropeptide prediction, Sequest.

1. Introduction

1.1. Prediction of Peptide Precursor Genes from the Genome

Since the advent of genome projects, computational methods have become especially important in predicting novel putative peptides and their precursor genes. The genome of an organism may be screened for peptide-coding genes based on sequence similarity to known peptide genes from other organisms using Basic Local Alignment Searching Tool (*BLAST*). For example, *BLAST* helped to identify 36 peptide genes found in *Drosophila melanogaster* (1, 2). However, for in silico prediction of peptide precursor genes, the performance of the BLAST tool is limited because putative peptide precursor sequences, for which no homologous biologically

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_25,

active peptides or their precursors have been identified as yet, will not be revealed. Peptide precursors are a special class of proteins because they undergo extensive posttranslational processing before producing final mature peptides. In most cases, only a short conserved motif might be responsible for the function of a particular peptide. The remainder of the peptide precursor sequence may be essentially irrelevant and show no significant sequence similarity (3). Due to the limited sequence conservation between peptides or their precursors, the *BLAST* tool is also not very effective at identifying new members of known peptide families. *BLAST* is suitable for scanning databases for protein sequences in which the sequence similarity is expected along the entire or most part of the sequences (global alignment) or when the similarity is limited to a specific domain (local alignment). But, it is far less efficient at finding similarity to short conserved regions spanning only few amino acids. For large peptide precursors which are between 50 and 500 amino acids in length and for which the biologically conserved regions are limited, the relevant motifs are often masked by random matches with long but unrelated sequence regions. This is because for any two random large protein sequences, *BLAST* usually can find a relatively long local alignment. That alignment is likely to be longer than any typical conserved peptide motif, and therefore *BLAST* would assign higher scores to long "random" alignments rather than to the short peptide conservative motifs. If a pair of homologous proteins share only a short mature peptide sequence, *BLAST* may not be able to detect the homology because the short alignment makes the pairwise sequence alignment less likely to obtain a significant *BLAST* score (e.g., *e*-value < 0.001) (4, 5). Many bioactive peptides have been sequenced by now, several of these are short and no precursors are yet known for these. There is a growing need to take advantage of these mature peptides in identifying homologous peptides and peptide precursors.

Here we describe a searching algorithm for systematic search and identification (*in silico*) of all peptide precursor proteins in a specific species. Our method uses *BLAST* but also relies on the detection of additional structural hallmarks of peptides and their precursor sequences. The original study was performed in *D. melanogaster*, where 76 additional putative secretory peptide genes were discovered in addition to 43 known sequences (6). This bioinformatic study opens perspectives for the genome-wide analysis of (neuro)peptide genes in other eukaryotic model organisms.

1.2. Prediction of Peptide Precursor Proteins from EST Data

In many organisms (neuro)peptide research is hampered by the absence of genomic information. In its absence the Expressed Sequence Tag (EST) databases can be interrogated for (neuro)peptide precursors. ESTs are short, single-read cDNA

sequences, usually between 200 and 500 nucleotides, derived from a particular tissue and/or a particular stage in development. One disadvantage of ESTs is the high sequencing error rate. Also, information at the transcriptome level varies in time and place. However, for many organisms EST libraries comprise the largest pool of sequence data available and often contain portions of transcripts from many uncharacterized genes.

Here, we describe how EST databases can be searched for (neuro)peptide precursors using a simple *BLAST* search. For such a search, one has to take into account the peptide's small size and its very limited sequence similarity, and to take advantage of the structural hallmarks of peptide precursor sequences. To validate our approach we searched an EST database of the locust *Locusta migratoria*, which contained 12,161 clustered Unigenes, and compared our predictions against known locust neuropeptide precursors (7). Using neuropeptide precursors from *D. melanogaster* as a query, we annotated six novel neuropeptide precursors.

1.3. Identification of Peptides by Mass Spectrometry (MS)

Expression of the predicted peptides in different tissues can be confirmed with mass spectrometric techniques. This not only shows which peptides are cleaved from the precursor proteins but could also reveal their posttranslational modifications. A number of databases and web tools exist that can be used to speed up the process of identifying endogenous peptides analyzed by mass spectrometry. One way to identify peptides in a biological extract is by matching experimental peptide masses against theoretically calculated masses in a database, both with and without annotated posttranslational modifications, using a selected mass tolerance based on the mass accuracy of the mass spectrometer used. The *SwePep* database consists of approximately 4200 annotated endogenous peptides originating from 394 different species, which are divided into three classes (i) biologically active peptides, (ii) potential biologically active peptides, and (iii) uncharacterized peptides (8). Another database *PeptideDB* represents the most complete collection of metazoan peptides, peptide motifs, and peptide precursor proteins identified to date (9). It contains 20,027 peptides that are processed from 19,438 precursor proteins. The peptides include neuropeptides, growth factors, peptide toxins, and antibacterial peptides and have currently been retrieved from 2820 different metazoan species. However, a peptide identity based solely on the observed molecular mass is only a suggestion and needs to be confirmed by sequence analysis of the corresponding tandem mass spectra (MS/MS). We here describe how sequence information can be retrieved by MS/MS ion searches and de novo sequencing.

2. Materials

1. Personal computer installed with *SAS* (Statistical Analysis System – a statistical software package), a web browser, and Internet access
2. *Swepep* database (www.swepep.org)
3. *PeptideDB* (www.peptides.be)
4. *Uniprot* protein database
5. *SignalP* (www.cbs.dtu.dk/services/SignalP)
6. *NCBI BLAST* (www.ncbi.nlm.nih.gov/blast/Blast.cgi)
7. *tBLASTn* (www.ncbi.nlm.nih.gov/blast/Blast.cgi)
8. *TMpred* (www.ch.embnet.org/software/TMPRED_form.html)
9. *SOSUI* (bp.nuap.nagoya-u.ac.jp/sosui/sosuiG/sosuigsubmit.html)
10. *Translate tool* (www.expasy.ch/tools/dna.html)
11. *ClustalW* (www.ebi.ac.uk/clustalw/)
12. MS/MS fragmentation data – peak list files
13. *Mascot* (www.matrixscience.com)
14. *Sequest* (fields.scripps.edu/sequest)
15. *Peaks* (www.bioinformaticssolutions.com)
16. *PepNovo* (proteomics.bioprojects.org/MassSpec/)
17. *ProP* software tool (www.cbs.dtu.dk/services/ProP/)
18. *NeuroPred* (neuroproteomics.scs.uiuc.edu/neuropred.html)

3. Methods

3.1. Prediction of Peptide Precursor Genes from the Genome

The existence of the common structural characteristics of known peptide precursors (*see* **Note 1**) allows to devise a sensitive searching procedure capable of identifying peptide genes. We have originally developed such program for *D. melanogaster*, but because the structural hallmarks of peptide precursor sequences are highly conserved across phyla, the established searching algorithm can be easily adapted for the genome-wide analysis of peptide precursor genes in other animal model systems that have their genome sequenced (*see* **Note 1**). The general principles of our algorithm are exemplified below for *D. melanogaster*. The same steps can be used in relation with other species.

1. From *D. melanogaster* genome database select protein sequences that are shorter than 500 amino acids and that contain a signal peptide sequence.
2. Cleave these proteins *in silico* at typical cleavage sites (*see* **Note 1)** and use *BLAST* to compare these polypeptide fragments (subsequences) against full-length protein sequences to identify proteins which match at least two similar polypeptide fragments.
3. Compare the fragments obtained at the Step 2 (above) with all known bioactive peptide sequences from all metazoan organisms (*see* **Note 2**).
4. Based on the sequence comparison results, two types of screening procedures can be constructed (*see* **Note 3**):
   i. Finding the precursor proteins which encode multiple highly related putative peptides
   ii. Finding the precursors containing a single putative peptide or multiple unrelated putative peptides that share conserved motifs with known bioactive peptides.

The program is implemented in *SAS* – a powerful integrated software for accessing, management, and analysis of large datasets (*see* **Note 4**). External tools such as *SignalP, BLAST, TMpred*, and *SOSUI* need to be run independently. Text files are used to exchange the data between the different programs. The *SAS* program includes a few subprograms listed below.

3.1.1. Protein.SAS

This subprogram is the first part of the *SAS* program, and it serves to select a subset of candidate protein sequences from any given species. For example, in *D. melanogaster*, the input of the subprogram consists of the Uniprot protein database file and additional *D. melanogaster* genes at GenBank identified by Hild et al. (10). The algorithm and the operational procedure are outlined below.

1. The relevant information for each of the proteins, such as accession number, protein name, gene name, protein sequence, signal peptide information, length, and mass, is entered into *SAS*. The first 70 amino acids of every protein sequence serve as output to a text file in FASTA format, which is used as the input for *SignalP*.
2. *SignalP* for eukaryotes is then run to predict the presence and location of a signal peptide in each protein sequence (11).
3. The subprogram reads the output file from *SignalP*, and another *SAS* dataset is created that includes the predicted signal peptide information for each protein.
4. The dataset containing the predicted signal peptide information is then checked against all the proteins, and the pro-

teins are retained if they are either annotated to have signal peptides in *Uniprot* or predicted to have signal peptides by *SignalP*. The result is a dataset of proteins having amino-terminal signal peptides.

5. From this dataset, only proteins that are shorter than 500 amino acids are retained.

3.1.2. Cleavage.SAS

This subprogram is used to cleave the protein sequences in the *D. melanogaster* protein dataset into polypeptide fragments following the removal of the signal peptide sequences. A number of conserved precursor proteins cleavage motifs have been reported (12). These are GKR, GRK, GRR, GKK, KR, RK, RR, KK, GR, GK (*see* **Note 5**). **Table 25.1 c**ompares the frequency of occurrence of these motifs in the proteome of *D. melanogaster* and compares that with the frequencies that these basic sites are actually used as cleavage sites in all of the annotated peptides from *D. melanogaster*. A similar analysis is shown for the vertebrate *Mus musculus* in **Table 25.2**. All the protein fragments, obtained by cleavage through these cleavage motifs form the *D. melanogaster* subsequence dataset (*see* **Notes 6 and 7**). Flow chart shown in **Fig. 25.1** summarizes *Protein.SAS* and *Cleavage.SAS* procedures.

Table 25.1
Frequencies of known consensus cleavage sites in known peptides in *D. melanogaster*[a]

	GKR	GRK	GRR	GKK	KR	RK	RR	KK	GR	GK	R	K
Cleaved sites[b]	18(C)	1(C)	6(C)	0	35(N) 13(C)	2(N) 3(C)	16(N) 10(C)	3(N)	11(N) 17(C)	1(N) 5(C)	6(N) 2(C)	1(N)
Uncleaved sites[c]	1	3	1	4	12	30	25	29	21	13	260	305
Percentage (%)[d]	94.7	25.0	85.7	0	80.0	14.3	51.0	9.4	57.1	31.6	3.0	0.3

[a]The numbers are based on the analysis of 146 annotated peptides in *D. melanogaster*. The total number of amino acids in all these peptides is 7346 (the flanking basic cleavage sites not included).
[b]Cleaved sites: Number of consensus sites at which cleavage process occurs. The (N) or (C) following the number indicates whether the cleavage site is located at the amino- or carboxy-terminus of the peptide sequence.
[c]Uncleaved sites: Number of consensus sites at which no cleavage occurs.
[d]Percentage (%): The number of sites at which cleavage occurs relative to the total number of consensus sites found (expressed in %): $\frac{\text{Cleaved sites}}{\text{Cleaved sites}+\text{Uncleaved sites}} \times 100$.

3.1.3. Peptide.SAS and the BLAST Analysis

This subprogram searches the *UniProt* database for all the annotated bioactive peptides from all metazoan organisms. The summary of *Peptide.SAS* is shown in **Fig. 25.2**. The algorithm and the operational procedure are outlined below.

1. All proteins from Metazoa, which function as mature peptides or peptide precursor proteins, are assembled into a dataset of peptides and precursors. A protein sequence has characteristics of a peptide or peptide precursor if its name

Table 25.2
Frequencies of known consensus cleavage sites in known peptides in *Mus musculus*[a]

	GKR	GRK	GRR	GKK	KR	RK	RR	KK	GR	GK	R	K
Cleaved sites[b]	23(C)	0(C)	19(C)	8(C)	88(N) 29(C)	6(N) 1(C)	44(N) 17(C)	6(N) 4(C)	1(N) 6(C)	1(N) 4(C)	48(N) 17(C)	6(N) 1(C)
Uncleaved sites[c]	11	7	53	14	165	225	230	160	170	169	2416	2251
Percentage (%)[d]	67.6	0	26.4	36.4	41.5	3.0	21.0	5.9	4.0	2.9	2.6	0.3

[a]The numbers are based on the analysis of 595 annotated peptides in *Mus musculus*. The total number of amino acids in all these peptides is 54,621 (the flanking basic cleavage sites not included).
[b]Cleaved sites: Number of consensus sites at which cleavage process occurs. The (N) or (C) following the number indicates whether the cleavage site is located at the amino- or carboxy-terminus of the peptide sequence.
[c]Uncleaved sites: Number of consensus sites at which no cleavage occurs.
[d]Percentage (%): The number of sites at which cleavage occurs relative to the total number of consensus sites found (expressed in %): $\frac{\text{Cleaved sites}}{\text{Cleaved sites}+\text{Uncleaved sites}} \times 100$.

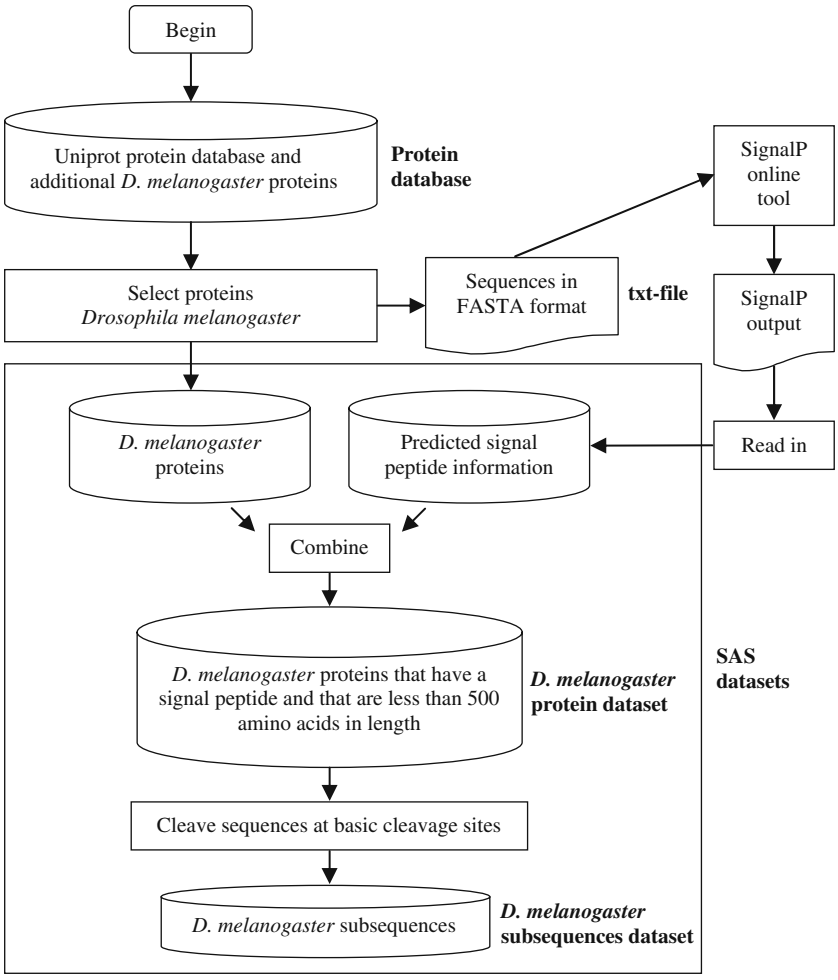


Fig. 25.1. Protein.SAS and Cleavage.SAS, Modified from 6.

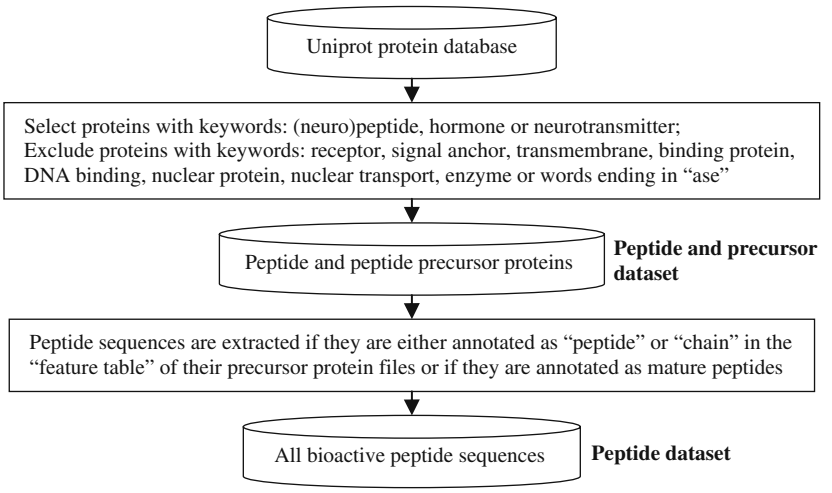


Fig. 25.2. Peptide.SAS, Modified from 6.

contains typical peptide keywords or if it is annotated with peptide keywords in the "Keywords" line in *UniProt*. The bioactive peptide keywords include (neuro)peptide, hormone, and neurotransmitter.

2. Proteins defined as membrane proteins (as indicated in *UniProt*) or proteins having the keywords such as receptor, signal anchor, transmembrane, binding protein, DNA binding, nuclear protein, nuclear transport, enzyme, or words ending in "ase" are excluded.
3. Bioactive peptide sequences are then extracted *in silico* from each precursor protein present in the peptide and precursor dataset. Peptides are extracted if they are annotated with the keyword peptide or chain in the "Feature table" of their corresponding precursor protein files in the *UniProt*. The endpoint specifications "from" and "to" indicate beginning and the end of the peptide fragments. The conserved basic cleavage sites flanking the peptides are also extracted (*see* **Note 8**).
4. Database entries from the peptide and precursor dataset that represent mature peptide sequences are also retained in the peptide dataset (*see* **Note 9**).
5. All the selected metazoan peptides (the above dataset) are exported as a single FASTA formatted file "peptide.txt".
6. All the selected amino acid sequences in the *D. melanogaster* dataset (from the **Sections 3.1.1** and **3.1.2**) are exported as a single FASTA formatted file "subsequence.txt".
7. Standalone *BLAST* is applied to compare the two sequence files. The score matrix "PAM30" is used, and the expectation value (*e*-value) and the parameter "word size" are set

to 6 and 2, respectively, in order to find short but strong similarities (*see* **Note 10**).

3.1.4. Extract.SAS, Motif.SAS, and Shift.SAS

These subprograms are used to screen the alignment results output by *BLAST* and determine the biologically significant matches. The summary of the procedure is shown in **Fig. 25.3**.

1. *Extract.SAS* reads the alignments between *D. melanogaster* fragments (*see* **Section 3.1.2**) with themselves and extracts the proteins that have at least two similar subsequences within the protein.
2. *Motif.SAS* reads the alignments between *D. melanogaster* fragments and known peptide sequences as well as the alignments among peptide sequences themselves and identifies the fragments that contain conserved peptide motifs.
3. *Shift.SAS* reads the alignment results and computes the shift value. The shift value is the minimal distance between the amino- or carboxy-termini of the aligned sequence and the matching amino acids (sequence tags) in the sequence. The shift value is set to be no larger than 3 in the program (*see* **Note 11**).

3.1.5. The Implementation of Screening Procedures

The subprograms described in the previous section (*Extract.SAS, Motif.SAS, and Shift.SAS*) facilitate peptide screening procedures, the principles of which were described in **Section 3.1**.

1. The first procedure searches for proteins which contain the following sequence pattern:

$$\cdots[\text{cleavage1}]-\text{x1}(3,60)-[\text{cleavage2}]-\cdots$$
$$-[\text{cleavage3}]-\text{x2}(3,60)-[\text{cleavage4}]\cdots$$

   In this formula, "x1 (3, 60)" and "x2 (3, 60)" are two similar fragments which are between 3 and 60 amino acids long (*see* **Note 12**). "[cleavage1–4]" can be any conventional cleavage site. The fragments do not need to be adjacent within the precursor, and the matching amino acid sequence should be present close to the amino- or carboxy-termini of at least one of the fragments (shift value ≤ 3). To implement such screening procedure, the file "*subsequence.txt*" should be compared with itself using *BLAST*. Then the subprograms *Extract.SAS* and *Shift.SAS* should be used to select those proteins which match the first structural pattern of a putative peptide precursor (containing multiple highly related putative peptides).
2. The second procedure looks for proteins that meet other structural characteristics of peptide precursors (containing

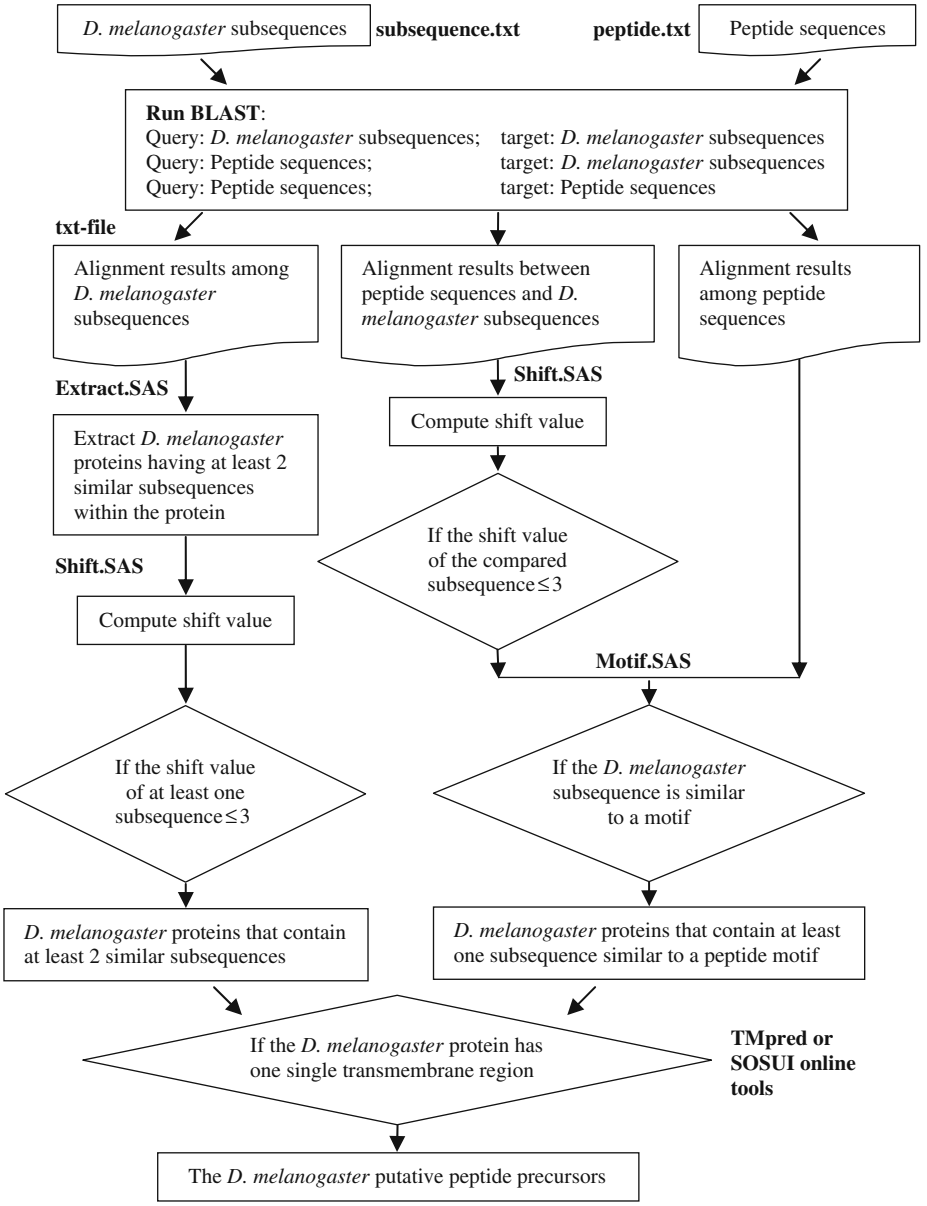


Fig. 25.3. BLAST analysis and screening of the alignment results, Modified from 6.

a single putative peptide or multiple unrelated putative peptides):

i. The protein should contain at least one fragment that shares at least 60% amino acid sequence identity with a known peptide sequence, and the identical amino acids are situated close to the amino- or carboxy-termini of that fragment (shift value ≤ 3).

ii. The identical amino acids should be similar to a conserved motif present in other known peptide sequences.

This screening procedure involves the following sequence comparison by *BLAST*:

i. The sequences in the "*peptide.txt*" are compared with each other and the obtained similar amino acid sequence tags are considered as possible conserved peptide motifs.

ii. The file "*peptide.txt*" is compared with the "*subsequence.txt*" and those *D. melanogaster* fragments that display sequence similarities to any peptide motifs from the previous step are retained.

3.1.6. TMpred and SOSUI

These tools are available online and could be used to identify the presence of a single transmembrane region at the amino-terminus of a protein. The length of the hydrophobic part of the transmembrane region should be set to between 17 and 33 amino acids. For the *TMpred* program a score above 500 for both inside to outside as well as outside to inside helixes is considered to be significant for the presence of the amino-terminal transmembrane region. A score of 250 is considered to be significant for the presence of an inside to outside helix of any second or third transmembrane region. A putative peptide precursor is retained if any of the programs predicts a single transmembrane region at the protein amino-terminus. When both programs predict the absence of an amino-terminal transmembrane region, the protein sequence is removed from the list (*see* **Note 13).**

3.2. Prediction of Peptide Precursor Proteins from EST Data

1. Go to *NCBI BLAST* website and select *tBLASTn* (*see* **Note 14**). Enter the query sequence (*see* **Note 15**). Select the expressed sequence tags (EST) database from the pull-down menu and limit the search by specifying the correct species or entering the EST accession numbers under "entrez query" (*see* **Note 16**). Other *BLAST* parameters are left at their default values.
2. When a significant match is found (e.g., *e*-value < 0.001) the corresponding EST sequences should be further analyzed for the presence of start and stop codons and for the typical peptide precursor features (*see* **Note 1).** For this the EST sequence must be translated. In the *BLAST* result page, click on the accession number to display the full EST sequence. Go to www.expasy.ch/tools/dna.html, paste the EST sequence in FASTA format, and select to translate the sequence. Specify the correct frame. The frame number is displayed on the *BLAST* result page. The "+1" corresponds to 5′3′ frame 1, "–1" corresponds to 3′5′ frame 1, and so on. Select the amino acid sequence between the first start (Met) and stop codon. Paste this sequence into *ClustalW*. Replace

all Met by M and delete the word STOP. Align this sequence with the homologous peptide precursor(s) (the ones that were used to perform the *BLAST* search) and perform a multiple sequence alignment. Analyze the results. Look for the presence of cleavage sites and the conservation of motifs (*see* **Note 17**).

3.3. Identification of Peptides in an MS/MS Ion Search

The types of peptide fragment ions observed in an MS/MS spectrum depend on many factors including the primary sequence, the mode of energy introduction, and the charge state. Fragments can only be detected if they carry at least one charge. If this charge is retained on the amino-terminal fragment, the ions are classified as either a, b, or c. If the charge is retained on the carboxy-terminal fragment, the ions are classified as either x, y, or z. The subscripts indicate the number of residues in a certain fragment (summarized in **Fig. 25.4**). In a typical MS/MS ion search, all MS/MS data of every peptide selected for fragmentation during a liquid chromatography (LC)-MS/MS run are combined in a single peak list file. This simple type of file contains the monoisotopic masses and associated intensity values of all the parent ions and their corresponding fragmentation ions, and can be used for further bioinformatics analyses such as MS/MS ion searches and de novo sequencing. The peptides can be identified by comparing the experimentally obtained fragmentation spectra with the theoretical fragmentation spectra in databases (13).

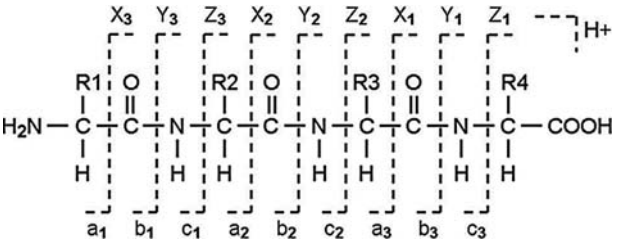


Fig. 25.4. Possible peptide fragmentation patterns.

1. The peak list files can be used to query MS/MS data using *Mascot* and *Sequest* tools. Settings for use of endogenous peptides should be as follows: variable modifications; carboxy-terminal amidation; oxidation of methionine; and pyro-glu (N-term Q). Set enzyme to "none".
2. A *FASTA* protein database containing all the (in silico) identified putative peptide precursors should be constructed and loaded onto the *Mascot* or *Sequest* server and be used for the identification of peptides using an MS/MS ion search (*see* **Note 18**).

3. Posttranslational peptide modifications may hinder confident identification with MS/MS data. To overcome this problem one should use tools like *Peaks* and *PepNovo* (14) which assist in determining the amino acid sequence of peptides from the raw MS/MS data. De novo sequencing will further increase the number of reliable protein identifications.

4. Notes

1. Structural hallmarks of peptide precursor sequences can be arranged in three major categories:
   i. Almost all known peptide precursors are less than 500 amino acids in length and contain one single transmembrane region at the amino-terminus corresponding to the signal peptide that directs them into the secretory pathway of the cell.
   ii. The precursor is processed into bioactive peptides by a series of enzymatic steps. After cleavage of the signal peptide, prohormone convertases break the precursor protein into smaller peptides by cleaving mainly at paired dibasic residues. Carboxy-terminal basic residues are subsequently removed by carboxypeptidases, and peptides with a carboxy-terminal glycine are converted into the amide by peptidylglycine α-amidating monooxygenase thereby stabilizing the C-terminus. Also other posttranslational modifications occur, e.g., N-terminal glutamine residues often cyclize resulting in pyroglutamate.
   iii. Many peptide precursors encode multiple bioactive peptides that are often highly related. Peptide genes encoding multiple, unrelated bioactive peptides or genes encoding just a single bioactive peptide also occur.
2. The direct alignment of short *D. melanogaster* polypeptide fragments and the metazoan peptide sequences increases the sensitivity of finding and matching short (conserved) peptide motifs and thus overcomes the shortcomings of *BLAST* when searching *long* sequences for *short* matching fragments.
3. Because each *D. melanogaster* protein sequence was cleaved into a number of subsequences and because all of these subsequences were subsequently compared with each other or with all known metazoan peptide sequences, a very large

number of alignments were obtained, with a high score. Because similarity does not imply homology, only the alignments which were filtered by the screening procedure were considered as candidate putative peptide precursors.

4. The software tool is not limited to *SAS*. Any software, which is capable of dealing with large datasets, can be applied to implement the program.
5. The cleavage of peptide precursors does not occur at every basic site (as evident from **Tables 25.1** and **25.2**). In the described program, we cleave a protein sequence into short fragments at every position where motifs GKR, GRK, GRR, GKK, KR, RK, RR, KK, GR, and GK occur. This results in the maximum possible number of candidate fragments. A statistical analysis on all known peptides in Metazoa shows that the minimal distance between the amino- or carboxy-termini of a peptide sequence and the conserved region (motif) in the peptide sequence is usually small (the distance is defined as "shift" in **Section 3.1.4**). This means that the conserved peptide motif should be close to a cleavage site in the peptide precursor. Based on this observation, our program identifies a sequence as a potential peptide if the sequence possesses a conserved peptide motif near its amino- or carboxy-terminus. We do not consider monobasic sites R and K as cleavage sites because of the low probability of their occurrence –3.5% (260/7346) and 4.1% (305/7346), respectively, as seen from **Table 25.1**. Furthermore, many conserved peptide motifs contain the amino acids R and K, such as, for example, the motif "[LVMI]-[MLIV]-R-F" from the peptide families "FMR Famide and related neuropeptides" and "K-[KN]-[YF]-G-G-F-M" motif from adrenocorticotropic hormone domain and opioids neuropeptides (3).
6. It has been suggested that the cleavage process also depends on the amino acids that are at the proximity of the cleavage site. For example, aliphatic amino acids (leucine, isoleucine, valine, methionine) are rarely present immediately after the consensus cleavage site of the subtilisin/kexin-like proprotein convertases [R/K]-(X)*n*-[R/K] ↓ (where X is any amino acid except cysteine and *n* is equal to 0, 2, 4 or 6) (15). Several prediction tools have been developed to identify putative cleavage sites in peptide precursors. For example, the *ProP* software and online tool *NeuroPred* predict basic cleavage sites of peptide precursors based on biochemical sequence data (16, 17). *NeuroPred* also has the capability to calculate the mass of the neuropeptides resulting from the predicted cleavages. The resulting mass list aids the discovery and confirmation of

new neuropeptides using MS techniques. The *ProP* and *NeuroPred* prediction tools can be used as the alternative to the *Cleavage.SAS* program (6).

7. In addition to the basic cleavage sites, peptide precursors may cleave at other non-basic sites (18). It will be a challenge to consider the existence of these unconventional cleavage sites in the further refinement of this method.
8. If the residues flanking the peptides are a combination of a few consecutive K or R, the combination is extracted as the cleavage site.
9. Many bioactive mature peptides are identified by direct protein sequencing techniques and their precursor proteins are unknown.
10. The expected value (*e*-value) is set to 6 because of the short length of the sequence fragments being compared.
11. Based on the statistical analysis of the peptide precursors in the peptide and precursor dataset, the shift value should be low. This means that the motifs should be in the close vicinity of a cleavage site.
12. For the majority (~98%) of the known peptide precursors that encode such multiple related peptides, the length of the fragments does not exceed 60 amino acids.
13. We predicted 76 additional putative secretory peptide genes in *D. melanogaster* (6). Some of these predicted novel precursors contain two or more fragments that share significant sequence similarities and others share conserved peptide motifs with known vertebrate or invertebrate peptides. These similarities could not be discovered by *BLAST* scanning of the whole *D. melanogaster* genome. Only one of the characterized peptide precursors in *D. melanogaster* was not identified by our method, i.e., the diuretic hormone precursor CG8348, because it has four transmembrane regions. Our procedure yielded four false positives (6).
14. Since EST sequences are not annotated, no protein translations are available for the *BLAST* search of EST databases. Hence the *tBLASTn* search is the only way to search for these potential coding regions at the protein level. *TBLASTn* compares protein query sequences against a nucleotide sequence database dynamically translated in all reading frames.
15. The *BLAST* search program is not suitable for the detection of small peptides. To circumvent this problem, one could combine several peptide isoforms and (posttranslational) processing sites in a single sequence query. For example,

(neuro)peptide sequences that are expected to originate from a single precursor should be flanked by typical processing sites [(G)KR, (G)RK, (G)(R)R or (G)(K)K] and combined into a single sequence; all possible combinations should be entered. The *BLAST* query sequence box accepts a number of different types of inputs and it will automatically determine the format used.

16. EST sequences reside in a specific division within GenBank, the dbEST database. For example, the ESTs of *L. migratoria* are deposited in the GenBank database under the accession numbers C0819675–C0832059 and C0832067–C0865130. These ESTs can be searched by selecting the EST database and entering "CO819675:CO832059[accn] OR CO832067:CO865130[accn]" in the "entrez query" box.
17. In the *L. migratoria* study, some of the known neuropeptide precursors were not found when searching EST databases (7). It is possible that these sequences were not present in the EST database. Alternatively, because ~3% of ESTs are estimated to contain sequencing errors, these could easily mask or disrupt short peptide alignments.
18. *Mascot* or *Sequest* should be set and run locally. Online versions of *Mascot* and *Sequest* are available, but are limited to using large databases like Swiss-Prot or NCBInr. These are less suitable for peptide searches. Also, most proteomic identification tools, including *Mascot*, are designed to identify a protein from several individual peptides or fragments originating from the same protein. The protein score in a peptide summary is derived from the ion scores of the individual peptides. Many peptide precursors give rise to only one or a very limited number of bioactive peptides, and because peptidomic experiments focus on the peptides themselves rather than on the peptide precursor proteins, only peptide scores can be taken into the account for peptide identification. In addition, because the exact processing mechanisms involved in the production of any particular peptide are unknown, no cleavage enzyme can be selected for identification. All these features of naturally occurring peptides should be considered to allow peptide identification. Our research indicates that improved success rate of identification of secretory peptides could be achieved using restricted databases of predicted peptide precursor proteins.

Acknowledgments

This work was supported by grants of the Fund for Scientific Research (FWO)-Flanders (1.5.137.06) and the Institute for the Promotion of Innovation through Science and Technology (IWT)-Flanders (SBO 335605). The authors also acknowledge Prometa, the Interfacultary Centre for Proteomics and Metabolomics at K.U. Leuven. E. Clynen and S.J. Husson are postdoctoral fellows of the FWO-Flanders and B. Landuyt is a postdoctoral fellow of the IWT-Flanders.

References

1. Hewes, R.S. and Taghert, P.H. (2001) Neuropeptides and neuropeptide receptors in the *Drosophila melanogaster* genome. *Genome Res.* **11**, 1126–1142.
2. Vanden Broeck, J. (2001) Neuropeptides and their precursors in the fruitfly, *Drosophila melanogaster*. *Peptides* **22**, 241–254.
3. Liu, F., Baggerman, G., Schoofs, L., and Wets, G. (2006) Uncovering conserved patterns in bioactive peptides in Metazoa. *Peptides* **27**, 3137–3153.
4. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.
5. Altschul, S.F., Bundschuh, R., Olsen, R., and Hwa, T. (2001) The estimation of statistical parameters for local alignment score distributions. *Nucleic Acids Res.* **29**, 351–361.
6. Liu, F., Baggerman, G., D'Hertog, W., Verleyen, P., Schoofs, L., and Wets, G. (2006) In silico identification of new secretory peptide genes in *Drosophila melanogaster*. *Mol. Cell. Proteomics* **5**, 510–522.
7. Clynen, E., Huybrechts, J., Verleyen, P., De Loof, A., and Schoofs, L. (2006) Annotation of novel neuropeptide precursors in the migratory locust based on transcript screening of a public EST database and mass spectrometry. *BMC Genom.* **7**, 201.
8. Fälth, M., Skold, K., Norrman, M., Svensson, M., Fenyo, D., and Andren, P.E. (2006) SwePep, a database designed for endogenous peptides and mass spectrometry. *Mol. Cell. Proteomics* **5**, 998–1005.
9. Liu, F., Baggerman, G., Schoofs, L., and Wets, G. (2008) The construction of a bioactive peptide database in Metazoa. *J. Proteome Res.* **7**, 4119–4131.
10. Hild, M., Beckmann, B., Haas, S.A., Koch, B., Solovyev, V., Busold, C., Fellenberg, K., Boutros, M., Vingron, M., Sauer, F., Hoheisel, J.D., and Paro, R. (2003) An integrated gene annotation and transcriptional profiling approach towards the full gene content of the *Drosophila* genome. *Genome Biol.* **5**, R3.
11. Bendtsen, J.D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**, 783–795.
12. Veenstra, J.A. (2000) Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. *Arch. Insect Biochem. Physiol.* **43**, 49–63.
13. Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567.
14. Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., and Lajoie, G. (2003) PEAKS: powerful software for peptide *de novo* sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **17**, 2337–2342.
15. Rholam, M., Brakch, N., Germain, D., Thomas, D.Y., Fahy, C., Boussetta, H., Boileau, G., and Cohen, P. (1995) Role of amino acid sequences flanking dibasic cleavage sites in precursor proteolytic processing. The importance of the first residue C-terminal of the cleavage site. *Eur. J. Biochem.* **227**, 707–714.
16. Duckert, P., Brunak, S., and Blom, N. (2004) Prediction of proprotein convertase cleavage sites. *Protein Eng. Des. Sel.* **17**, 107–112.

17. Southey, B.R., Amare, A., Zimmerman, T.A., Rodriguez-Zas, S.L., and Sweedler, J.V. (2006) NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. *Nucleic Acids Res.* **34**, W267–W272.
18. Seidah, N.G., Benjannet, S., Wickham, L., Marcinkiewicz, J., Jasmin, S.B., Stifani, S., Basak, A., Prat, A., and Chretien, M. (2003) The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): liver regeneration and neuronal differentiation. *Proc. Natl. Acad. Sci. USA* **100**, 928–933.

Chapter 26

Bioinformatic Identification of Plant Peptides

Kevin A. Lease and John C. Walker

Abstract

Plant peptides play a number of important roles in defence, development and many other aspects of plant physiology. Identifying additional peptide sequences provides the starting point to investigate their function using molecular, genetic or biochemical techniques. Due to their small size, identifying peptide sequences may not succeed using the default bioinformatic approaches that work well for average-sized proteins. There are two general scenarios related to bioinformatic identification of peptides to be discussed in this paper. In the first scenario, one already has the sequence of a plant peptide and is trying to find more plant peptides with some sequence similarity to the starting peptide. To do this, the Basic Local Alignment Search Tool (BLAST) is employed, with the parameters adjusted to be more favourable for identifying potential peptide matches. A second scenario involves trying to identify plant peptides without using sequence similarity searches to known plant peptides. In this approach, features such as protein size and the presence of a cleavable amino-terminal signal peptide are used to screen annotated proteins. A variation of this method can be used to screen for unannotated peptides from genomic sequences. Bioinformatic resources related to *Arabidopsis thaliana* will be used to illustrate these approaches.

Key words: Peptide, peptidomics, bioinformatics, *Arabidopsis thaliana*.

1. Introduction

Plant peptides can be defined as small proteins, below an arbitrary molecular weight or length cut-off (1). Peptides can be generated either from a gene encoding a small open reading frame, or they can be produced from a larger protein that undergoes post-translational proteolytic cleavages that give rise to one or more smaller peptides. Proteolytically produced peptides may be bioactive functional peptides or they may represent non-functional turnover of formerly active proteins. The cleavage sites of plant

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4_26,

proteolytic enzymes are not well established, so it is not possible to look at primary amino acid sequences and identify which proteins will be processed or which mature peptides will result.

Nine plant signalling peptides have been intensively characterized by biochemical and molecular genetic experiments (2–13). These founding peptides coupled with the availability of genome sequences have led to the identification of additional peptides through bioinformatics analyses (9, 10, 14–18). In addition, based on the collective properties of identified plant peptides, some general properties have emerged. This information has been exploited to find additional peptides (19). The purpose of this review is to suggest how to use available tools and resources with the goal of identifying plant peptides of interest for further investigation.

BLAST (20, 21) is a useful and well-known bioinformatic tool that can be used to find additional members of a gene family, if a founding member is available to use as a query. For example, many plant peptides that were originally identified through genetic or biochemical studies were found to belong to families of genes encoding similar peptides (9, 15). Using *BLAST* with default settings is not ideal for plant peptide studies. Various parameters involved in *BLAST* searches will be discussed as well as the rationale for changing them. The overall goal of the specific parameters suggested is to increase sensitivity. Following these changes, one will greatly increase the odds of finding meaningful similar sequences in the database when searching with short queries. At the same time spurious matches will increase, so healthy scepticism, sound judgement and further investigation will be required.

2. Materials

1. A personal computer with a web browser installed and Internet access are required.
2. *TAIR BLAST* at The Arabidopsis Information Resource (http://www.arabidopsis.org/Blast).
3. *SignalP3.0* (http://www.cbs.dtu.dk/services/SignalP).
4. *TMHMM2.0* (http://www.cbs.dtu.dk/services/TMHMM).

5a. *TAIR* bulk data retrieval and analysis tools (http://www.arabidopsis.org/tools/bulk/index.jsp).

5b. *TAIR* bulk protein search page (http://www.arabidopsis.org/tools/bulk/protein/index.jsp).

5c. *TAIR* gene description search and download page (http://www.arabidopsis.org/tools/bulk/genes/index.jsp).

5d. *TAIR* sequence bulk download and analysis page (http://www.arabidopsis.org/tools/bulk/sequences/index.jsp).

6a. The *Arabidopsis* Unannotated Secreted Peptide Database *AUSP* (http://peptidome.missouri.edu).

6b. *AUSP* search page (http://peptidome.missouri.edu/cgi-bin/getprotein.cgi).

7. The *Arabidopsis* Transcriptome Genome Expression Database *ATGED* (http://signal.salk.edu/cgi-bin/atta).

8. NCBI *BLAST* (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

9. *BLASTCLUST*(http://toolkit.tuebingen.mpg.de/blastclust#;).

10. *REPRO* (http://zeus.cs.vu.nl/programs/reprowww/).

11a. Multiple Em for Motif Search *MEME* (http://meme.sdsc.edu/meme/).

11b. Motif Alignment & Search Tool *MAST* (http://meme.sdsc.edu/meme/cgi-bin/mast.cgi).

3. Methods

3.1. Identifying Peptides with Sequence Similarity to Another Peptide Sequence

When conducting *BLAST* searches over the internet, The *Arabidopsis* Information Resource TAIR *BLAST* tool is a good choice. It contains plant-specific datasets to search and usually returns the results faster than NCBI *BLAST*. To access the additional parameters discussed below that may not be displayed initially, click on the "+" sign on the Advanced *BLAST* Parameter Options line on the webpage.

1. Select the appropriate *BLAST* Program. To start with, select *BLASTP* – it is often the best choice (*see* **Notes 1–4**).
2. Select suitable BLAST Dataset. To search *Arabidopsis* alone, select the *A. thaliana GB all* database; to search other higher plants, the *Green plant GB all* database should be chosen (*see* **Notes 5–7**).
3. Enter the sequence query. Full or partial sequence may be used. To start with, use the full precursor sequence (*see* **Notes 8** and **9**).
4. Turn off all filters on the *BLAST* input page by un-checking the "Filter query" box (*see* **Note 10**).

5. Select appropriate weight matrix. PAM70, PAM30 or BLOSUM45 are suitable matrices for peptide searches because they are less stringent for scoring amino acid similarities (*see* **Note 11**).
6. Choose appropriate word size. For peptide searches, setting the word size to 2 makes the search more sensitive, albeit slower (*see* **Note 12**).
7. Increase the "expect score" value to 1000 (*see* **Note 13**).
8. Select the scoring penalties (gap opening and gap extension penalties). The default gap opening penalty is 11 and the gap extension penalty is 1. To achieve better sensitivity with short sequence queries, lower gap opening and extension penalties should be selected. Reasonable choice would be 7–10 for the gap opening penalty and 1–2 for the gap extension penalty (*see* **Note 14**).
9. Increase the number of reported scores ("Max scores") to 500 (*see* **Note 15**).
10. Analysis of the matched sequence

.

3.2. Analysis of Structural Features in the Matched Putative Peptide Sequences

Analysis of structural features in the found sequences may be very useful for discriminating potentially meaningful matches, identified, e.g. using **Section 3.1** (see above). The choice as to which structural characteristics to analyse would depend on the target peptide and is likely to be different in each individual case. Two common examples are described here.

1. If the query sequence represents a secreted peptide, one might expect the matched putative peptide also to be secreted. Signal peptide and transmembrane domain prediction tools, such as *SignalP3.0* and *TMHMM2.0*, respectively, should be used to evaluate this possibility (*see* **Note 16**).
2. The length of the database sequence might be expected to conform to size limits seen in known plant peptide precursors. For example, the largest of known plant peptide precursors is 200 amino acid long Systemin from tomato (2). This is a sensible upper limit to have in mind when evaluating the peptides and protein precursors identified (*see* **Note 17**).

3.3. Peptide Identification Based on Peptide-Associated Characteristics

In addition to the use of sequence similarity searches, peptides may be identified through the analysis of features common to known plant peptides (19). Most such peptides are produced from larger precursor proteins. The precursor proteins typically have ~200 or fewer amino acids. Additionally, most peptide precursors have cleavable amino-terminal signal peptides that direct the protein to the secretory pathway. *TAIR* Bulk Data Retrieval and Analysis Tools offer a simple solution to the identification

of all of the annotated genes in *Arabidopsis* that are predicted to encode proteins with the aforementioned qualities (i.e. small size and having a signal peptide).

3.3.1. Bulk Protein Search

1. Open *TAIR* bulk protein search page (http://www.arabidopsis.org/tools/bulk/protein/index.jsp).
2. The format should be set to "html", output boxes should be unchecked, the predicted molecular weight range should be set to "0 to 22,000" (based on 110 Da per amino acid and an estimated 200 amino acids maximum precursor length). "Secreted Proteins" should be checked under the choice of "Predicted sub-cellular location."
3. After clicking "Get Protein Data", approximately 2000 gene IDs will be returned. Select all and copy the gene IDs onto clipboard, just as one would copy text in a word processor program.
4. Open the *TAIR* gene description search and download page, and paste the gene IDs into the search box. Run the search. One can peruse the gene descriptions or follow the hypertext links to obtained detailed information about the gene.
5. Open *TAIR* sequence bulk download and analysis page; paste the gene IDs to download all of the amino acid sequences in *FASTA* format for further evaluation.

3.3.2. Arabidopsis Unannotated Secreted Peptide Database Searches

Given that small genes are often poorly annotated, many potential peptides are not included in the *TAIR* annotated list (19). A complementary resource that can be used to address this issue is the *Arabidopsis* Unannotated Secreted Peptide Database *AUSP*. This is a searchable database of more than 30,000 unannotated open reading frames that may encode small secreted proteins.

1. Open the *AUSP* search page.(http://peptidome.missouri.edu/cgi-bin/getprotein.cgi).
2. Set the Chromosome to "All", Select "Both Strands" and click "Search and View."
3. Evaluation of the peptide expression levels. Some expression data are available in *AUSP*, but more extensive expression data are available from the *Arabidopsis* Transcriptome Genome Expression Database *ATGED*.

3.3.3. Other Structural Considerations – Internal Repeats

There exists one example of a plant peptide precursor that contains two bioactive peptides which share sequence similarity to one another (8). In invertebrates, this is a common finding (22). For example, multiple FMRFamide peptides may be encoded by a single precursor protein. If this pattern can be extrapolated to additional plant peptides, they may be putatively identified by

looking for internally repeated sequences. A bioinformatics tool called *REPRO* is available on the web to identify internal repeats.

3.3.4. Other Structural Considerations – Sequence Patterns

Multiple Em for Motif Elicitation *MEME* is an algorithm that can be used to find patterns among a group of peptides (18, 23). The patterns found by *MEME* can be used to search with Motif Alignment & Search Tool *MAST* for other proteins sharing this pattern.

4. Notes

1. *BLASTP* is generally the best choice for *BLAST* program. *BLASTP* is used with an amino acid query to search a database of proteins. *BLASTP* is a better choice for this application than *BLASTN* because conservation of sequence similarity at the amino acid level is higher than nucleic acid sequence similarity. It is worth noting that genome annotations are not static and results of searches may vary over time as the database is updated. Therefore, it is important to pay attention to which database and which version of that database is available
2. In some circumstances it may be advantageous to choose *TBLASTN* over *BLASTP*. Many small genes are not well annotated; in such cases, the protein encoded by that gene may not be included in the protein database. *TBLASTN* will deal with this by searching genomic DNA databases translated in all six reading frames, using a protein query.
3. Position-specific iterated *BLAST* (*PSI-BLAST*) is another variant of *BLAST* that can be used to find additional members of a plant peptide family (14, 18). *PSI-BLAST* searches are not available at the *TAIR* website, but are available through NCBI *BLAST*.
4. Many plant peptides belong to gene families, rather than being "singletons." *BLASTCLUST* is a useful way to group a large list of peptide precursors in *FASTA* format into groups that have sequence similarity. *BLASTCLUST* uses single-linkage clustering. This may be a useful filter for screening *TAIR* proteins. If the "cluster" box is checked at *AUSP*, one can see whether the peptides in the search results have sequence similarity to other peptides in *AUSP*.
5. Database selection (referred to as "Datasets" on *TAIR* website) could greatly influence the results. There are eight protein databases available at *TAIR*. Larger databases are normally most suitable if searching for peptides. Increasing

the size of the dataset by selecting more than one library makes it more likely to find a match based on sequence similarity alone. However, if the sequence is not in the database, the match can never be found.

6. If the same sequence fragment is found in two databases, the expect value for the same peptide sequence match in this larger dataset will be higher. This means that the probability of the match being significant will be lower.
7. The protein databases at *TAIR* are organized to be mutually exclusive for searching *Arabidopsis* or non-*Arabidopsis* higher plants. Currently there is no protein database at *TAIR* that would combine *Arabidopsis* and other higher plant sequences in a single search.
8. Search using the full-length precursor sequence typically produces fewer false positives and the results obtained are usually easier to interpret.
9. One might suspect that the mature peptide sequence would make a better query because this sequence might be better conserved. This is not always the case. For example, the signal peptides within the same gene family could be more similar to each other than the rest of the protein sequences (14).
10. These filters mask out part of your query which might eliminate potential matches.
11. The matrix chosen for the *BLAST* search can affect results. Many matrices have been optimized for searching proteins not peptides and are only suitable for detecting high sequence similarity between the query and database subjects.
12. The *BLAST* algorithm requires an initial exact match with a "word" which is your query sequence broken into small chunks. Increasing the word size parameter speeds up *BLAST* search, but makes it less sensitive and will yield fewer hits.
13. The *BLAST* "expected score" indicates the statistical significance of the sequence matches found. These depend on the degree of sequence identity (better match results in lower "expect" values) and on the size of the database (in larger databases, the chance of having a random match is higher). When searching with very short queries, it makes sense to sacrifice selectivity for sensitivity and to consider "expect score" values. Some such matches may be biologically significant and meaningful, but one must regard these with a sceptical eye.

14. When a query sequence and a database sequence are aligned, inserting gaps in one of these sequences may increase the apparent quality and the degree of alignment. However, inserting a gap will reduce the overall alignment score. Similarly, extending the gap length may improve the alignment but will also decrease the alignment score. If the overall improvement in the alignment outweighs the gap penalties incurred, the alignment is accepted by the system as the preferred fit.
15. The default setting is 50. Changing this value will not affect the search, but will simply increase the number of alignments reported. One line identifiers can be checked quickly.
16. As with many prediction tools, care should be taken in using *SignalP3.0* and *TMHMM2.0* prediction tools and interpreting the results obtained (24, 25).
17. When analysing the proteins, one should not rely exclusively on the short descriptors of the protein sequences. These may appear to be useful, but they can also be misleading, in that annotation of gene function can be speculative in some cases.
18. In *ATGED* gene expression data are displayed graphically. These can be used to evaluate expression of the peptides from *AUSP*, as well as any other genes from *TAIR*. The annotation track *AUSP* links to the *Arabidopsis* Unannotated Secreted Peptide sequences.

References

1. Farrokhi, N., Whitelegge, J.P., and Brusslan, J.A. (2007) Plant peptides and peptidomics. *Plant Biotechnol. J.* **6**, 105–134.
2. McGurl, B., Pearce, G., Orozco-Cardenas, M., and Ryan, C.A. (1992) Structure, expression, and antisense inhibition of the systemin precursor gene. *Science* **255**, 1570–1573.
3. Matsubayashi, Y., and Sakagami, Y. (1996) Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of *Asparagus officinalis* L. *Proc. Natl. Acad. Sci. USA* **93**, 7623–7627.
4. Fletcher, J.C., Brand, U., Running, M.P., Simon, R., and Meyerowitz, E.M. (1999) Signaling of cell fate decisions by CLAVATA3 in *Arabidopsis* shoot meristems. *Science* **283**, 1911–1914.
5. Kondo, T., Sawa, S., Kinoshita, A., Mizuno, S., Kakimoto, T., Fukuda, H., and Sakagami, Y. (2006) A plant peptide encoded by CLV3 identified by in situ MALDI-TOF MS analysis. *Science* **313**, 845–848.
6. Schopfer, C.R., Nasrallah, M.E., and Nasrallah, J.B. (1999) The male determinant of self-incompatibility in *Brassica*. *Science* **5445**, 1697–1700.
7. Takayama, S., Shiba, H., Iwano, M., Shimosato, H., Che, F.S., Kai, N., Watanabe, M., Suzuki, G., Hinata, K., and Isogai, A. (2000) The pollen determinant of self-incompatibility in *Brassica campestris*. *Proc. Natl. Acad. Sci. USA* **97**, 1920–1925.
8. Pearce, G., Moura, D.S., Stratmann, J., and Ryan, C.A. (2001) Production of multiple plant hormones from a single polyprotein precursor. *Nature* **411**, 817–820
9. Pearce, G., Moura, D.S., Stratmann, J., and Ryan, C.A. (2001) RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development. *Proc. Natl. Acad. Sci. USA* **98**, 12843–12847.
10. Butenko, M.A., Patterson, S.E., Grini, P.E., Stenvik, G.E., Amundsen, S.S., Mandal, A., and Aalen, R.B. (2003) Inflorescence deficient in abscission controls floral organ

abscission in *Arabidopsis* and identifies a novel family of putative ligands in plants. *Plant Cell* **15**, 2296–2307.
11. Huffaker, A., Pearce, G., and Ryan, CA (2006) An endogenous peptide signal in *Arabidopsis* activates components of the innate immune response. *Proc. Natl. Acad. Sci. USA* **103**, 10098–10103.
12. Stenvik, G., Tandstad, N.M, Guo, Y., Shi, C., Kristiansen, W., Holmgren, A., Clark, S.E., Aalen, R., and Butenko, M.A. (2008) The EPIP Peptide of INFLORESCENCE DEFICIENT IN ABSCISSION is sufficient to induce abscission in *Arabidopsis* through the receptor-like kinases HAESA and HAESA-LIKE2. *Plant Cell* **20**, 1805–1817.
13. Cho, S.K., Larue, C.T., Chevalier, D., Wang, H., Jinn, T., Zhang, S., and Walker, J.C. (2008). Regulation of floral organ abscission in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **105**, 15629–15634.
14. Vanoosthuyse, V., Miege, C., Dumas, C., and Cock, J.M. (2001) Two large *Arabidopsis* thaliana gene families are homologous to the *Brassica* gene superfamily that encodes pollen coat proteins and the male component of the self-incompatibility response. *Plant Mol. Biol.* **46**, 17–34.
15. Cock, J.M., and McCormick, S. (2001) A large family of genes that share homology with CLAVATA3. *Plant Physiol.* **126**, 939–942.
16. Yang, H., Matsubayashi, Y., Nakamura, K., Sakagami, Y. (2001) Diversity of *Arabidopsis* genes encoding precursors for phytosulfakine, a peptide growth factor. *Plant Physiol.* **127**, 842–851.
17. Olsen, A.N., Mundy, J., and Skriver, K. (2002) Peptomics, identification of novel cationic *Arabidopsis* peptides with conserved sequence motifs. *In Silico Biol.* **2**, 441–451
18. Oelkers, K., Goffard, N., Weiller, G.F., Gresshoff, P.M., Mathesius, U., and Frickey, T. (2008) Bioinformatic analysis of the CLE signaling peptide family. *BMC Plant Biol.* **8**, 1.
19. Lease, K.A., and Walker, J.C. (2006) The *Arabidopsis* unannotated secreted peptide database, a resource for plant peptidomics. *Plant Physiol.* **142**, 831–838.
20. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
21. Poole, R.L. (2007). The TAIR Database. *Methods Mol. Biol.* **406**, 179–212.
22. Baggerman, G., Cerstiaens, A., De Loof, A., and Schoofs, L. (2002). Peptidomics of the larval *Drosophila melanogaster* central nervous system. *J. Biol. Chem.* **277**, 40368–40374.
23. Baggerman, G., Liu, F., Wets, G., and Schoofs, L. (2006) Bioinformatic analysis of peptide precursor proteins. *Ann. NY Acad. Sci.* **1040**, 59–65.
24. Bendtsen, J.D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**, 783–787.
25. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580.

INDEX

A

M. Soloviev (ed.), *Peptidomics*, Methods in Molecular Biology 615, DOI 10.1007/978-1-60761-535-4,

B

C

D

E

F

G

H

I

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z