PEPTIDOMICS
Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation’s journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!

William J. Pesce  
President and Chief Executive Officer

Peter Booth Wiley  
Chairman of the Board
PEPTIDOMICS

Methods and Applications

Edited by

MIKHAIL SOLOVIEV
Royal Holloway, University of London

CHRIS SHAW
Queen’s University Belfast

PER ANDRÉN
Uppsala University
CONTENTS

Preface xv

Contributors xvii

PART I STUDIES OF NATURALLY OCCURRING PEPTIDES 1

1 Analysis of the Peptidomes of Amphibian Skin Granular Gland Secretions—An Integrated Functional Genomic Strategy 3
Chris Shaw and Tianbao Chen

1.1 Introduction 3
1.2 Historical Perspective 5
1.3 Contemporary Methods of Sample Acquisition 6
1.4 The Integrated Functional Genomic Strategy 7
1.5 How then can All Three Approaches be Integrated? 8
1.6 Limitations of Each Approach 18
1.7 Closing Thoughts 19

2 A Short History of Insect (Neuro)Peptidomics—A Personal Story of the Birth and Youth of an Excellent Model For Studying Peptidome Biology 25
Peter D.E.M. Verhaert, Martijn W.H. Pinkse, Maria C. Prieto-Conaway, and Markus Kellmann

2.1 Introduction 25
2.2 History 29
2.3 Present (and Future): Novel Technologies, New Data . . . 31
  2.3.1 MALDI and (Nano)ESI TOFs and QTOFs 31
  2.3.2 Peptide Displays 31
  2.3.3 Ion Traps and FT-ICRs: Greater Variety of Instruments,
      Greater Variety of Sequences . . . 35
  2.3.4 Orbitrap 35
  2.3.5 FT-ICRs 40
  2.3.6 Combining Peptide Chemical Analysis with Histological
      Localization: Peptidome Mass Spectrometry Imaging 42
2.4 Concluding Remarks 45

3 Peptidomics of Short Linear Cytolytic Peptides from Spider Venom 55
   Sergey A. Kozlov, Alexander A. Vassilevski, and Eugene V. Grishin
   3.1 Introduction 55
   3.2 Peptidomics: Identifying and Sequencing Novel AMPs 57
   3.3 Genomics: Unraveling the Biosynthetic Pathway 60
   3.4 Venomics: AMPs in the Spider Venom 64
      3.4.1 Structural Features 65
      3.4.2 Functional Features 65
      3.4.3 Venomic Significance 66

4 Molecular Cloning Approaches to Peptidomics: The Identification
   of Novel cDNAs Encoding Neurotoxin-like Peptide Pools 71
   Zhensheng Pan, Richard Barry, and Mikhail Soloviev
   4.1 Introduction 71
   4.2 Spider Toxins—Combinatorial Peptide Libraries 73
   4.3 EST Cloning Approach to Cloning Peptide Families 74
   4.4 PCR-Based Amplification and Cloning Approaches 74
   4.5 The Design and Use of Hybrid Primers for the Amplification of
      cDNAs Coding for Large Toxin-like Peptide Pools 77
      4.5.1 Multiple Alignment of the Existing Toxins and Toxin-like
          Sequences 77
      4.5.2 The Design of Hybrid Partially Degenerate Primers 79
          4.5.2.1 The Design of the 3′ Region 79
          4.5.2.2 The Design of the 5′ Region and the Middle
              Part of the Primers 79
          4.5.2.3 The Overall Length and Degeneracy 79
          4.5.2.4 Annealing Temperature Matching 80
      4.5.3 Optimization of the Amplification Conditions 82
      4.5.4 Other Experimental Details 84
      4.5.5 High Throughamplification and Cloning 85
      4.5.6 Sequences Identification and Analysis 86
4.6 Polypeptide Toxins—Novel Applications in Drug and Pesticide Development 88
4.6.1 Analgesic Drugs 88
4.6.2 Insecticides 89
4.7 Concluding Remarks 91
4.7.1 RACE-PCR Issues 91
4.7.2 Sequence Quality Issues 91
4.7.3 Conclusion 92

5 Wheat Antimicrobial Peptides 99
Tatyana I. Odintsova and Tsezi A. Egorov

5.1 Introduction 99
5.2 Materials and Methods 101
5.2.1 Isolation of AMPs 101
5.2.2 Reduction and Alkylation of Peptides [35] 102
5.2.3 MALDI-TOF MS 102
5.2.4 High-Resolution Two-Dimensional Gel Electrophoresis 103
5.3 Isolation of AMPs from T. kiharae Seeds 103
5.4 Sequence Determination of AMPs from T. kiharae Seeds 106
5.5 Comparative Analysis of D Defensins from Triticum and Aegilops Species 111
5.6 Discussion 112
5.7 Concluding Remarks 115

6 Immunopeptidomics: Applications to Dissect Immune Responses Through Proteomic-based Approaches 119
Ernesto Oviedo-Orta and Alexandra Bermudez-Fajardo

6.1 Introduction 119
6.1.1 Innate Immune Responses 120
6.1.2 Adaptive Immune Responses 120
6.2 Naturally Occurring Peptides and MHC-I Molecules: Role in Immune Surveillance and Tolerance 121
6.2.1 Structure and Basic Function of MHC-I Molecules 121
6.2.2 Relationship Between Naturally Occurring Peptides and MHC-I Function: A Peptidomics Approach 122
6.2.2.1 Peptide Analogs as Useful Tools to Dissect Immunological Mechanisms 124
6.2.2.2 Synthetic Peptide Libraries 126
6.3 Presentation of Naturally Processed Peptide Antigens by MHC-II Molecules 126
6.3.1 Peptidomics Characterization of MHC-II-Binding Naturally Occurring Peptides 127
6.3.2 Functional Relationship Between MHC-II Molecules and Naturally Binding Peptides 128
6.4 Current Methodologies Applied to Peptidomics Analysis and the Identification of Naturally Processed Peptides 129
6.5 Potential Clinical Applications of Immunopeptidomic Analyses 132

7 Strategies for Reliable and Improved Identification of Peptides 139
Karl Sköld, Maria Fälth, Marcus Svensson, Anna Nilsson, Per Svenningsson, and Per E. Andréén

7.1 Introduction 139
7.2 Neuropeptides and Hormones 140
7.3 Identification of Neuropeptides 142
7.4 Improved Identification of Peptides Using SwePep 143
7.5 Differential Display of Peptides 144
7.6 Differential Display of Peptides in Models of Parkinson’s Disease 147

PART II PEPTIDOMICS METHODS AND NEW DEVELOPMENTS 153

8 Peptidomics Approach to Proteomics 155
Petra Zürbig and Harald Mischak

8.1 Introduction 155
8.2 Peptidomics Relating to Proteomics 156
8.2.1 Definition of Peptides Versus Proteins 156
8.2.2 Proteases—Producers of Peptides 157
8.3 Proteomics/Peptidomics 158
8.3.1 Background on Proteomics 158
8.3.2 General Objectives of Proteomics 158
8.3.3 Peptidomics 160
8.4 The ‘omics Technologies 161
8.4.1 General Remarks 161
8.4.2 Sample Preparation 161
8.4.3 Basics of Mass Spectrometry 162
8.4.4 Bottom-Up Versus Top-Down Approaches 166
8.4.5 Separation Technologies Coupled MS 167
8.4.5.1 Two-Dimensional Gel Electrophoresis Followed by Mass Spectrometry (2DE-MS) 167
8.4.5.2 Liquid Chromatography Coupled to Mass Spectrometry (LC-MS) 169
8.4.5.3 Surface-Enhanced Laser Desorption/Ionization Coupled to Mass Spectrometry (SELDI-MS) 169
8.4.5.4 Capillary Electrophoresis Coupled to Mass Spectrometry (CE-MS) 170
9 The Importance of Sample Handling in Neuropeptidomics

Karl Sköld, Marcus Svensson, Anna Nilsson, Maria Fälth, Per Svenningsson, and Per Andreén

9.1 Introduction 177
9.2 Issues Using Postmortem Brain Tissue 178
9.3 Enzyme Inactivation Methods 179
  9.3.1 Denaturation 179
  9.3.2 Thermal Denaturation 179
  9.3.3 Proteomics Denaturation Methods 179
  9.3.4 Other Inactivation Methods 180
  9.3.5 Protease Inhibitors 180
9.4 Brain Sample Preparation 181
9.5 Reliable Identification of Peptides in the Brain 181
9.6 Postmortem Time Effects on Peptides and Proteins 184
9.7 Protein Fragments as Degradation Markers 186

10 Affinity Peptidomics Approach to Protein Detection, Quantification, and Protein Affinity Assays: Application to Forensics and Biometrics

Julian Bailes and Mikhail Soloviev

10.1 Introduction 191
  10.1.1 Protein and Peptides: Separation, Staining, and Detection 192
  10.1.2 Affinity-Based Assays and Hardware Realizations 194
  10.1.3 Micro-/Macroarray-Based Protein Assays 196
10.2 Peptidomics Approach to Affinity Assays 198
10.3 Forensic and Biometrics Marker Proteins and Peptide Selection 206
10.4 Protein Markers and Peptide Selection 208
  10.4.1 Myoglobin (MG) 210
  10.4.2 Tumor Necrosis Factor-Alpha (TNF-α) 211
  10.4.3 Fibrinogen (Fb)/D-Dimer 211
  10.4.4 C-Reactive Protein (CRP) 211
  10.4.5 Interferon(s) (IFN) 212
  10.4.6 Interleukin(s) (IL)
    10.4.6.1 IL-1 213
    10.4.6.2 IL-2 213
    10.4.6.3 IL-4 213
    10.4.6.4 IL-6 213
    10.4.6.5 IL-8 214
    10.4.6.6 IL-10 214
  10.4.7 Beta-Enolase (BE) 214
  10.4.8 E-Selectin (ES) 215
  10.4.9 Fibronectin (FN) 215
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.4.10</td>
<td>Haptoglobin (HG)</td>
<td>215</td>
</tr>
<tr>
<td>10.4.11</td>
<td>Heat Shock Protein(s) (HSP)</td>
<td>216</td>
</tr>
<tr>
<td>10.4.12</td>
<td>Insulin-like Growth Factor-1 (IGF-1)</td>
<td>216</td>
</tr>
<tr>
<td>10.4.13</td>
<td>Leptin (Lp)</td>
<td>217</td>
</tr>
<tr>
<td>10.4.14</td>
<td>Prolactin (Prl)</td>
<td>217</td>
</tr>
<tr>
<td>10.4.15</td>
<td>Transforming Growth Factor-Beta (TGF-β)</td>
<td>217</td>
</tr>
<tr>
<td>10.4.16</td>
<td>Adiponectin (Adn)</td>
<td>218</td>
</tr>
<tr>
<td>10.4.17</td>
<td>Insulin (In)</td>
<td>218</td>
</tr>
<tr>
<td>10.4.18</td>
<td>Matrix Metalloproteinases (MMPs 1, 2, 3, 9, 11)</td>
<td>218</td>
</tr>
<tr>
<td>10.4.19</td>
<td>Prostate-Specific Antigen (PSA)</td>
<td>219</td>
</tr>
<tr>
<td>10.4.20</td>
<td>Serum Amyloid-A (SAA)</td>
<td>219</td>
</tr>
<tr>
<td>10.4.21</td>
<td>Sex-Hormone-Binding Globulin (SHBG)</td>
<td>220</td>
</tr>
<tr>
<td>10.4.22</td>
<td>Surfactant-Associated Protein A (SP-A)</td>
<td>221</td>
</tr>
</tbody>
</table>

10.5 Concluding Remarks 224

11 Selective Depletion and Enrichment Methods for the Analysis of Protein and Peptide Pools 233

*Paul Finch and Mikhail Soloviev*

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>11.2</td>
<td>Peptide Mixture Refinement—General Approach to Combinatorial Peptidomics</td>
</tr>
<tr>
<td>11.3</td>
<td>Chemistry and Materials for Selective Reaction of Unmodified Amino Acid Side Chains</td>
</tr>
<tr>
<td>11.4</td>
<td>Selective Enrichment of Posttranslationally Modified Amino Acids</td>
</tr>
<tr>
<td>11.4.1</td>
<td>Phosphorylated Amino Acids</td>
</tr>
<tr>
<td>11.4.2</td>
<td>Selective Collection of Glycosylated Peptides</td>
</tr>
<tr>
<td>11.4.2.1</td>
<td>Phenyl Boronate Resins</td>
</tr>
<tr>
<td>11.4.2.2</td>
<td>Lectins</td>
</tr>
<tr>
<td>11.4.3</td>
<td>Amino Acids Modified by Covalent Attachment of Lipid Groups</td>
</tr>
<tr>
<td>11.5</td>
<td>Concluding Remarks</td>
</tr>
</tbody>
</table>

12 Detection of Target Peptides in Foods and Feeds by Mass Spectrometry 265

*Mireia Fernández Ocaña*

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.1</td>
<td>General Introduction</td>
</tr>
<tr>
<td>12.2</td>
<td>Introduction to GM Crops</td>
</tr>
<tr>
<td>12.3</td>
<td>Food Safety and Consumers Concerns</td>
</tr>
<tr>
<td>12.4</td>
<td>Current Detection Techniques of GM Material</td>
</tr>
<tr>
<td>12.5</td>
<td>Feed Contamination with Animal Protein in the Context of BSE and Control Measures</td>
</tr>
<tr>
<td>12.6</td>
<td>Plant Proteomics</td>
</tr>
<tr>
<td>12.7</td>
<td>Food Proteomics</td>
</tr>
<tr>
<td>12.8</td>
<td>Proteomic Investigations for the Detection of GM Crops</td>
</tr>
</tbody>
</table>
12.9 Case Study: Herbicide-Resistant GM Crops 275
12.9.1 Identification of CP4 EPSPS Peptides as Markers for Herbicide-Resistant GM Maize 275
12.9.2 Differential Protein Expression in GM and Non-GM Soya 277
12.10 Case Study: Gelatine Peptides as Markers for Contamination with Animal Protein 283
12.11 Concluding Remarks 287

13 Quantification of Polypeptides by Mass Spectrometry 291
Pedro R. Cutillas

13.1 Introduction 291
13.2 Development of Quantitative Mass Spectrometry for Polypeptide Analysis 293
13.3 Absolute Quantification of Polypeptides by Mass Spectrometry 295
13.3.1 The External Standard Method 295
13.3.2 The Internal Standard Method 297
13.4 Relative Quantification of Polypeptides by Mass Spectrometry 301
13.4.1 Methods Based on Derivatization 302
13.4.2 Methods Based on Metabolic Labeling 306
13.4.3 Label-Free Methods 307
13.5 Conclusions and Perspectives 311

PART III PEPTIDOMICS APPROACH TO BIOMARKER DISCOVERY 317

14 Biomarker Discovery 319
Petra Zürbig and Harald Mischak

14.1 Introduction 319
14.2 Definition of Biomarkers 320
14.3 Biomarkers as Health Care Devices 321
14.4 Biomarker Discovery 323
14.5 Proteomic and Peptidomic Biomarker Research 323
14.6 Clinical Proteomics/Peptidomics 324
14.6.1 Single Biomarker Versus “Marker Pattern” 326
14.6.2 Clinical Application 327
14.7 CE–MS in Clinical Diagnosis 327
14.7.1 Selection of Source 327
14.7.2 Biomarkers from Blood, Dialysis Fluid, and Cerebrospinal Fluid 330
14.7.3 Urinary Biomarkers 332
14.7.4 Biomarker Identification by MS/MS Sequencing 334
14.8 Concluding Remarks 336
15 Can Peptidomics Provide a Useful Approach for the Identification of Biomarkers of Toxicological Exposure or Effect? 341

Helen Griffiths

15.1 Introduction 341
15.2 Altered Peptide Expression and Secretion Profiles After Exposure to Stress or Toxins in Target Organs 344
15.3 Can Peptidomics Contribute to the Discovery of Biomarkers of Target Organ-Specific Toxin Exposure or Stress? 344
15.3.1 Evidence for Altered Peptide Profiles in Hepatotoxicity 345
15.3.2 Evidence for Altered Peptide Profiles in Nephrotoxicity 346
15.4 Proteolysis During Cytotoxicity 347
15.5 Alterations in Inflammatory Peptides Following Drug/Chemical Induced Toxicity 349
15.6 Concluding Remarks 351

16 Peptidomics in Neuroendocrine Research: A Caenorhabditis elegans and Mus musculus Study 355

Kurt Boonen, Steven J. Husson, Geert Baggerman, Anja Cerstiaens, Walter Luyten, and Liliane Schoofs

16.1 Introduction 355
16.2 Methods 358
16.2.1 Extraction of Peptides 358
16.2.2 Dissection 358
16.2.3 Sample Preparation 359
16.2.4 HPLC 360
16.2.4.1 1D Nano-LC 360
16.2.4.2 2D Nano-LC 360
16.2.4.3 Online 2D Nano-LC 360
16.2.4.4 Off-line 2D Nano-LC 361
16.2.5 Mass Spectrometry 361
16.2.5.1 ESI QTof Mass Spectrometry 361
16.2.5.2 Gas-Phase Fragmentation 364
16.2.5.3 MALDI-TOF Mass Spectrometry 364
16.2.6 Peptide Identification 364
16.2.6.1 Mascot Search 364
16.2.6.2 De Novo Sequencing 365
16.2.7 Neuropeptide Databases 365
16.3 Case Study 1: Peptidomics of C. elegans 367
16.3.1 Introduction 367
16.3.2 Bioinformatics: Prediction of Neuropeptide Genes 367
16.3.3 Biochemical Characterization of Neuropeptides 368
16.3.4 Experimental Procedures 368
16.3.4.1 Sample Preparation 368
16.3.4.2 Peptide Identification 368
16.3.5 Results 369
  16.3.5.1 Identification of 60 Endogenous Peptides by Online 2D Nano-LC QTOF MS/MS 369
  16.3.5.2 Identification of 75 Neuropeptides by Off-Line HPLC MALDI-TOF MS 369
  16.3.5.3 Differential Peptidomic Analysis of C. elegans Strains 372

16.3.6 Conclusions 372

16.4 Case Study 2: Peptidomics of the Murine Pituitary Gland 374
  16.4.1 Introduction 374
  16.4.2 Experimental Procedures 375
    16.4.2.1 Peptide Extraction 375
    16.4.2.2 Peptide Identification 375
  16.4.3 Results 375
    16.4.3.1 Proopiomelanocortin (POMC) 376
    16.4.3.2 Secretogranin I (Chromogranin B) 376
    16.4.3.3 Secretogranin II 376
    16.4.3.4 Secretogranin III 380
    16.4.3.5 Chromogranin A 380
    16.4.3.6 Neuroendocrine Secretory Protein 55 (NESP55) 380
    16.4.3.7 Neurosecretory Protein 7B2 380
    16.4.3.8 ProSAAS 380
    16.4.3.9 Oxytocin Precursor 380
    16.4.3.10 Vasopressin–Neurophysin 2–Copeptin Precursor 381
    16.4.3.11 Peptidyl-Glycine α-Amidating Monooxygenase Precursor 381
    16.4.3.12 Cystatin C 381
  16.4.4 Discussion 381

16.5 General Conclusion 383

PART IV CONCLUSION 387

17 Peptidomics and Biology: Two Scientific Disciplines Driving Each Other 389
  Chris Shaw and Peter D.E.M. Verhaert

Index 397
PREFACE

Only 6 years ago we became living witnesses to the genome revolution, when the human genome was finally sequenced raising hopes and expectations of finding magical cures to eradicate diseases at their “genetic” roots. Since then, the genomes of many other organisms have been sequenced. The central dogma of molecular biology is slowly being replaced as proteins reinstate their position as important players in epigenomics in addition to their well-established roles as structural, enzymatic, motor, information transmission, and other functions. Amino acids have gained additional respect by proving their important functional role as neurotransmitters and neuroregulators underlying our memory and learning abilities (glutamic acid, glycine) as well as playing important roles in the functioning of the cardiovascular and immune systems (homocysteine) in addition to simply being individual building blocks for making proteins. Have not we forgotten about peptides? Polypeptides, being short stretches of amino acids or small proteins, occupy a strategic position between proteins and amino acids and play, for most of the part, fundamental regulatory roles as hormones, cytokines, toxins, and so on. Although peptides have been studied for many decades, it was only recently that the term “peptidomics” has been coined, perhaps out of the desire to add some descriptive nomenclature to the enormous world-wide research effort into the biology and biochemistry of proteins and peptides. One of the most recent and welcomingly disruptive technological developments that has transformed life sciences is mass spectrometry — itself not new, but unlike other more “traditional” techniques, like chromatography and electrophoresis, it has been developed and refined to a level almost beyond belief since its conception, which was at about the same time as chromatography and electrophoresis. Mass spectrometry quickly ploughed through the land of proteins and has forcibly entered the field of peptide research.
The new technological capabilities of this technique have either spawned or greatly accelerated the development of many new disciplines, including peptidomics — the structural and functional inventory of bioactive peptides (the peptidome) in cells, tissues, and organisms.

This book is the first truly comprehensive volume that is aimed at providing a comprehensive and detailed overview of the emerging field of peptidomics: the peptidomics approach to research problem solving for the protein, biological, and biomedical scientist and to those working in related fields, as well as to provide a cross section of modern analytical techniques that can be applied for peptide analysis. The book is written as an essential laboratory tool for students and scientists as well as a guide for those with larger managerial responsibilities. This volume contains examples of successful peptidomic analyses of biological material ranging from plants to mammals and introduces exciting new techniques suitable for researchers formulating high throughout approaches as well as for those preferring more “traditional” approaches (studying one or a few peptides at a time). We believe that this book will provide advice and guidance to the widest range of protein biochemists. We are very grateful to all authors for their contributions to this work in the certainty that their efforts and applications will pave the way for a further explosive forward leap in both the volume and the diversity of peptidomic research.

Mikhail Soloviev
Chris Shaw
Per Andrén
CONTRIBUTORS

Per E. Andrén, Laboratory for Biological and Medical Mass Spectrometry, Department of Pharmaceutical Biosciences, Uppsala University, Box 583 Biomedical Centre, SE-75123 Uppsala, Sweden

Julian Bailes, Centre for Chemical and Bioanalytical Sciences, School of Biological Sciences, Royal Holloway, University of London Egham, Surrey TW20 0EX, United Kingdom

Geert Baggerman, Laboratory of Developmental Physiology, Genomics and Proteomics, KULeuven, Naamsestraat 59, 3000 Leuven, Belgium

Richard Barry, School of Biological Sciences, Royal Holloway, University of London Egham, Surrey TW20 0EX, United Kingdom

Alexandra Bermudez-Fajardo, Faculty of Health and Medical Sciences, University of Surrey, Guildford, GU2 7XH, Surrey, UK

Kurt Boonen, Laboratory of Developmental Physiology, Genomics and Proteomics, KULeuven, Naamsestraat 59, 3000 Leuven, Belgium

Anja Cerstiaens, Laboratory of Developmental Physiology, Genomics and Proteomics, KULeuven, Naamsestraat 59, 3000 Leuven, Belgium

Tianbao Chen, Molecular Therapeutics Research, School of Pharmacy, Queen’s University of Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, Northern Ireland, United Kingdom

Pedro R. Cutillas, Ludwig Institute for Cancer Research, University College London Branch, 91 Riding House St, London W1W 7BS, United Kingdom
Tsezi A. Egorov, Laboratory of Neuroreceptors and Neuroregulators, Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Moscow 117997, Russian Federation

Maria Fälth, Laboratory for Biological and Medical Mass Spectrometry, Department of Pharmaceutical Biosciences, Uppsala University, Box 583 Biomedical Centre SE-75123 Uppsala, Sweden

Mireia Fernández Ocaña, Centre for Chemical and Bioanalytical Sciences, School of Biological Sciences, Royal Holloway, University of London Egham, Surrey TW20 0EX, United Kingdom

Paul Finch, Centre for Chemical and Bioanalytical Sciences, School of Biological Sciences, Royal Holloway, University of London Egham, Surrey, TW20 0EX, United Kingdom

Helen Griffiths, School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, United Kingdom

Eugene V. Grishin, Laboratory of Neuroreceptors and Neuroregulators, Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Moscow 117997, Russian Federation

Steven J. Husson, Laboratory of Developmental Physiology, Genomics and Proteomics, KULeuven, Naamsestraat 59, 3000 Leuven, Belgium

Markus Kellmann, Thermo Fischer Scientific, Bremen, Germany

Sergey A. Kozlov, Laboratory of Neuroreceptors and Neuroregulators, Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Moscow 117997, Russian Federation

Walter Luyten, Faculty of Medicine, KULeuven, Naamsestraat 59, 3000 Leuven, Belgium

Harald Mischak, Mosaiques Diagnostics GmbH, Hannover, Germany

Anna Nilsson, Laboratory for Biological and Medical Mass Spectrometry, Department of Pharmaceutical Biosciences, Uppsala University, Box 583 Biomedical Centre, SE-75123 Uppsala, Sweden

Tatyana I. Odintsova, Laboratory of Plant Genetics, Vavilov Institute of General Genetics, Moscow, Russian Federation

Zhensheng Pan, School of Biological Sciences, Royal Holloway, University of London Egham, Surrey TW20 0EX, United Kingdom

Ernesto Oviedo-Orta, Faculty of Health and Medical Sciences, University of Surrey, Guildford, GU2 7XH, Surrey, UK

Maria C. Prieto-Conaway, Thermo Fischer Scientific, San Jose, California, USA.
Martijn W.H. Pinkse, Analytical Biotechnology Group, Kluyver Laboratory, Department of Biotechnology, Faculty of Applied Sciences, Delft University of Technology, Kluyver Laboratory, Julianalaan 67, 2628BC Delft, The Netherlands

Liliane Schoofs, Laboratory of Developmental Physiology, Genomics and Proteomics, KULeuven, Naamsestraat 59, 3000 Leuven, Belgium

Chris Shaw, Molecular Therapeutics Research, School of Pharmacy, Queen’s University of Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, Northern Ireland, United Kingdom

Karl Sköld, Laboratory for Biological and Medical Mass Spectrometry, Department of Pharmaceutical Biosciences, Uppsala University, Box 583 Biomedical Centre, SE-75123 Uppsala, Sweden

Mikhail Soloviev, Centre for Chemical and Bioanalytical Sciences, School of Biological Sciences, Royal Holloway, University of London Egham, Surrey TW20 0EX, United Kingdom

Marcus Svensson, Laboratory for Biological and Medical Mass Spectrometry, Department of Pharmaceutical Biosciences, Uppsala University, Box 583 Biomedical Centre, SE-75123 Uppsala, Sweden

Per Svenningsson, Department of Physiology and Pharmacology, Karolinska Institutet, SE-17177 Stockholm, Sweden

Alexander A. Vassilevski, Laboratory of Neuroreceptors and Neuroregulators, Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Moscow 117997, Russian Federation

Peter D.E.M. Verhaert, Analytical Biotechnology Group, Kluyver Laboratory, Department of Biotechnology, Faculty of Applied Sciences, Delft University of Technology, Kluyver Laboratory, Julianalaan 67, 2628BC Delft, The Netherlands

Petra Zürbig, Mosaiques Diagnostics GmbH, Hannover, Germany
PART I

STUDIES OF NATURALLY OCCURRING PEPTIDES
ANALYSIS OF THE PEPTIDOMES OF AMPHIBIAN SKIN GRANULAR GLAND SECRETIONS—AN INTEGRATED FUNCTIONAL GENOMIC STRATEGY

CHRIS SHAW AND TIANBAO CHEN

1.1 INTRODUCTION

Amphibians possess an extraordinary anti-predator defence mechanism based upon the stress-induced secretion of noxious and often highly-toxic chemical cocktails from specialised structures, the granular glands, usually located in discrete regions of the tegument. Although the chemical composition of such secretions can vary widely between different taxa, peptides represent the major class of biochemicals in the majority of species studied thus far with the exception of the dendrobatids, the so-called arrow poison frogs of the Neotropics. With several thousand species secreting peptidomes that may contain several hundred different peptides, the potential natural library of novel actives is sizeable requiring the application of modern bioanalytical technologies to the tasks of data acquisition, warehousing and mining. Superimposed upon these problems are the perhaps more difficult issues that gravitate around eco- and etho-friendly sample acquisition from species that are often extremely difficult to obtain or even encounter and that are under considerable survival threats on a global scale. For several years our research
has been directed to solving these problems, both technically and ethically, in a manner that integrates modern technologies with non-invasive, non-destructive sample acquisition and that does not compromise, in any way, the robustness of the scientific data produced. The integrated functional genomic strategy that has resulted from this research stands as a model for unravelling, in a holistic manner, the pharmacological/physiological roles of bioactive peptides that represent the “smart” weaponry encoded by fit-for-purpose evolved genes within the Amphibia.

Amphibian skin has been known to possess potent pharmacological properties for millennia having been utilized in many ways, from traditional ethnic medicines through tipping of arrows for hunting purposes to the more sinister applications in poisons and potions [1,4,5,26,28,29,33]. Even today, the use of amphibian skin per se or secretions derived from the specialized dermal granular glands are used widely across the world especially in Asia and South America [1,4,5,28,29,33]. Amphibians, as a group of animals, are generally rather small, soft-skinned, and slow in movement that renders them a prime source of food for many predatory species among the fish, reptiles, birds, and mammals. In the absence of surface armor, spines, claws, stings, or fangs, the amphibians have developed a highly sophisticated chemical defense against predation in their skin secretions that ranges from the noxious and bitter tasting at one end of the scale to the self-permeabilizing and highly toxic at the other [3,23,24,26,29]. The latter strategy is exemplified by the arrow-poison frogs of the family Dendrobatidae, from the New World [23–25]. Dendrobatid frog skin secretions are rich in alkaloid toxins, that, current thought believes, are derived and sequestered from forest invertebrates such as ants, which constitute, for the most part, the diet of these tiny and often brightly colored frogs [25,31,32]. The skin secretions of typical toads are likewise rich in small alkaloid molecules such as bufodienolides, bufotoxins, and biogenic amines, but this case, such molecules are de novo synthesized within the dermal venom glands that are often concentrated behind the eyes and on the back legs of the animals themselves [23,24]. The vast majority of other frog species and some groups of toads produce, as a result of endogenous gene expression, a highly complex cocktail of bioactive peptides and their biosynthetic precursor-related fragments when subjected to stress such as predator attack [3,26,29]. Typically, there may be as many as 100 different peptides with some species producing as many as 600. This obviously presents the biochemist with a formidable problem in terms of structural analysis and identification even with state-of-the-art analytical hardware and software, and if this is considered as a major barrier to overcome, it is nothing when compared to the establishment of individual biological functions. These barriers to understanding thus necessitate novel approaches that integrate existing elements of analytical strategy with new modifications and applications, and it is the purpose of this chapter to delineate how we have achieved this in part to characterize the peptidomes of amphibian defensive skin secretions.
Amphibian skin peptide research owes its origins to a superlative Italian scientist, Vittorio Erspamer, who was a classical pharmacologist by training [26]. He was intrigued by the obvious biological effects of both invertebrate and vertebrate venoms and was particularly impressed by the effect that the dermal venom of the common toad (Bufo bufo) had upon his dogs when they made a wrong decision to attack such a toad in his garden. This observation by a curious pharmacologist opened a treasure chest of molecular discovery that spanned the best part of half a century of productive ground breaking research. Many early works were focused on the characterization of the biogenic amines and alkaloids present in toad and salamander venoms, but a second chance observation on frog skin extracts led to the discovery of the defensive secretion peptides. Unlike extracts of toad skin, extracts of frog skin that were subjected to protease digestion lost much of their activity, indicating that the endogenous actives were of a proteinaceous character. In the early days of the evolution of chromatographic and electrophoretic techniques, this assertion was confirmed by the use of thin-layer and gel permeation chromatography [27]. Erspamer’s method was for the most part bioassay driven, in that using a series of defined bioassays that included the use of a variety of smooth muscle preparations, actives were identified and purified to homogeneity using the standards of the time [26]. The starting material for these experiments was usually a methanolic extract of air-dried skins that were obtained in the field after sacrifice of often many hundreds if not thousands of specimens of a particular species (phyllokinin—(bradykinyl-IY sulfate) originally purified from 3555 skins of Phyllomedusa rhodei) [2]. While this approach was necessary to produce sufficient material for primary structural analysis and pharmacological workups, remembering that in the early 1960s manual sequencing was the order of the day and peptide synthesis was in its infancy [2], it was highly destructive of animal life and produced an extract that included many elements not derived from the granular gland secretion but rather from additional skin tissues, blood, nerves, muscle, and so on. So while this approach worked well in the beginning with an identifiable bioactivity, it was most inefficient, time-consuming, and would be considered unethical by contemporary standards. For the purposes of contemporary peptidome research, we have the benefits of much more highly advanced analytical technologies, and we are in possession of a greater body of knowledge and an ethical code in the use of live animals for research that can be integrated in a way that results in more rapid, complete, and robust data generation in a manner that is noninvasive and nondestructive to the donor animals [35]. A “classically” prepared skin extract would contain many thousands if not tens of thousands of peptides/proteins that are not components of the granular glands, and hence most of the investigators’ valuable time and resources would be spent in structurally characterizing irrelevant molecules. This has to be overcome to enable a focus on the granular gland secretions themselves.
1.3 CONTEMPORARY METHODS OF SAMPLE ACQUISITION

Since the peptidome of the granular gland is the focused subject of study and analysis, methods of acquisition should ideally provide as pure a sample of the secretion of these structures as possible and ideally without significant damage or mortality to the donor. As the physiological control system for secretion in most species of amphibian relies on application of a stressor [26], the most user-friendly means of applying this has been investigated. Specimens in the field and those recently brought into captivity have a low threshold for such species but some adapt very rapidly to human handling and become readily tame such that more extreme stressors have to be applied after time to acquire defensive secretion. Some species of frog and toad are ideal candidates for study in that they secrete with the mildest of stresses that can be as mundane as lifting from their vivaria while others appear to be resistant to all but one might consider a severe stress. The granular glands are constructed in such a way that they are lined by myoepithelial cells that contract in the presence of stress-related catecholamines, such as norepinephrine, that are located in intrinsic nerve terminals [26,29,34]. This process can thus be mimicked by injection of the catecholamine directly into dorsal lymph sacs after which copious secretion is released from the granular glands [26,29,34]. This technique is reproducible and reliable but involves hypodermic injection of a substance whose effects will be systemic with possible side effects following prolonged usage. Another method involves placing the amphibian into a closed space with ether fumes acting as the stressor [26]. This requires the use of an inflammable and potentially explosive organic solvent that may cause harm to the specimens following prolonged usage. Perhaps the most efficient, reproducible, and harmless method is transdermal electrical stimulation as first described by Tyler et al. [35]. This technique applies a mild electrical stimulus (equivalent to two personal stereo batteries) via platinum electrodes to the moistened skin of the amphibian. Secretion is almost instant, and the whole procedure lasts not more than 1 min in total. It is sufficient to contract the myoepithelial cells and cause granular gland discharge with minimal stress to the whole animal, and it is very well tolerated with frogs even spawning shortly afterward in some cases. The procedure is highly reproducible and can be carried out at monthly intervals on the same individuals. As the stimulator is battery-powered, the technique is amenable to use in the field. Secretion is washed from the skin surface using a stream of deionized water and collected into a chilled glass beaker. The washings are snap-frozen in liquid nitrogen and lyophilized. Lyophilized material is stable indefinitely if stored under vacuum in a freezer at −20°C [6–22,38–41]. If secretions are collected in the field, then addition of trifluoroacetic acid to an approximate final concentration of 0.1% (v/v) is sufficient to stabilize the peptidome for subsequent analysis although lyophilization at the earliest opportunity is recommended [6–22,38–41]. Using this technique, the investigator obtains as complete a granular gland secretion as possible (minus volatiles) permitting objective
peptidomic analysis with minimal contamination of other skin tissue peptides/proteins.

1.4 THE INTEGRATED FUNCTIONAL GENOMIC STRATEGY

This may seem to be somewhat aberrant in a text that focuses on peptidomics, but biochemical analysis should always aim to procure the maximum amount of molecular information that is possible from a given biological sample. This is not a fanciful objective but rather can aid the researcher in producing a holistic understanding of the entire biological system from gene through protein/peptide through action—a necessary prerequisite in assessing academic, therapeutic, or biotechnological potential of new bioactive peptides. Our research on bioactive peptides from amphibian skin basically employs a three-prong approach with entry into the system being possible from each point. The bioassay approach duplicates the original method but tests individual fractions from reverse-phase HPLC fractionated skin secretion rather than crude material as a first pass. Positive fractions are repeated at least in triplicate to confirm effects following which further fractionation is performed if required to a homogeneous peptide, leading ultimately to structural characterization. The peptidomic approach treats the secretion sample like any other complex biological material and separates the components by tandem reversed-phase HPLC/electrospray MS or by off-line reversed-phase HPLC followed by MALDI-TOF analysis of individual fractions. In essence, to provide a more complete and readily accessible data set, both techniques are usually performed. Primary structures of peptides can be determined using MS/MS fragmentation with either online trawling of public databases to establish identity or de novo sequencing where this is not possible. Herein lies the major challenge that perhaps illustrates the major difference in peptidome analysis from exotic species rather than a peptidomic or indeed proteomic study from human, rat, mouse, chicken, Xenopus, zebrafish, and so on, whose genomes and hence translated potential proteomes have been fully or nearly fully sequenced. For the vast majority of species that we study, there is no relevant archive of either gene or protein sequence information publically available to provide a high throughput format that other proteome/peptidome researchers essentially take for granted for their “model” organisms. For this reason, we have to be particularly creative in problem solving to this end, and the finer points of this approach will be explained in greater detail in a following section. The genomic approach is one that is not new in this discrete field, but rather we employ this in a different manner to predecessors. In previous studies, the messenger RNAs encoding the biosynthetic precursors of novel peptides identified in amphibian skin secretions have been cloned from cDNA libraries manufactured from skin taken immediately post mortem from sacrificed specimens [37]. So no matter what good intentions had been met by obtaining the defensive secretion itself
for peptidome analysis, the amphibians were sacrificed in the quest for additional information on genetic coding, precursor organization, propeptide convertase sites, coencoded peptides, and so on. This invariably followed the acquisition of full primary structure of the peptide usually by automated Edman degradation. These data facilitated the generation of degenerate primer sets for the purpose of specifically cloning the precursor of the peptide following PCR amplification of its transcript. Following a brainstorming session over a Monday morning coffee, we had the idea that based on the syncytial nature of the granular gland and our *in cerebro* visualization of its stress-induced rupture and expulsion of entire contents, cytosolic peptide precursor-encoding polyadenylated mRNA might just be present in the skin secretion itself in a quantity and format that might facilitate PCR. The idea was correct. We first cloned from freshly extruded secretion, then lyophilized secretion, then lyophilized skin, and finally from air-dried skin [6–22,38–41]. The mechanism of polyadenylated mRNA protection within the amphibian skin has been proven to be very robust and essentially afforded by electrostatic interaction with the abundant cationic amphipathic peptides present in the secretions [22]. More recently, we have shown that there is sufficient DNA within the lyophilized secretion that after amplification can provide a template for specific gene structural mapping [12]. These discoveries on nucleic acid presence within the secretions meant that a holistic functional genomic study could readily be performed on a single, noninvasively obtained skin secretion sample: Bioactivity from reversed-phase HPLC fractions and establishment of primary structures of both the general peptidome and the specific active, cloning of precursor-encoding cDNA, and mapping of gene organization—all from as little as 10 mg of lyophilized secretion.

1.5 HOW THEN CAN ALL THREE APPROACHES BE INTEGRATED?

The *bioassay approach* gravitates around a whole organism, tissue, or cell that provides a defined target whose interaction with an active molecule in the skin secretion produces a defined and objectively measurable response. Indeed, the assay may be completely reductionist in that the target may simply be another biomolecule, such as a protease if one seeks inhibitors or naked DNA if one seeks nucleic acid interacting peptides. This approach requires no knowledge of the chemical structure of the active in the first instance. Even large pharma, with their considerable financial and human resources, are limited in the number of targets that molecular libraries can be screened against in a high throughput fashion and without doubt many have large numbers of potential actives/drug leads that remain undisclosed because active and appropriate targets have never been personally introduced. A relatively small research team thus has a discovery problem of several orders of magnitude more problematic than the one that necessitates a critical focus in target selection