

Protein Sequencing Protocols

John M. Walker, SERIES EDITOR

71. **PRINS and *In Situ* PCR Protocols**, edited by John R. Gosden, 1997
70. **Sequence Data Analysis Guidebook**, edited by Simon R. Swindell, 1997
69. **cDNA Library Protocols**, edited by Ian G. Cowell and Caroline A. Austin, 1997
68. **Gene Isolation and Mapping Protocols**, edited by Jacqueline Boulwood, 1997
67. **PCR Cloning Protocols: From Molecular Cloning to Genetic Engineering**, edited by Bruce A. White, 1997
66. **Epitope Mapping Protocols**, edited by Glenn E. Morris, 1996
65. **PCR Sequencing Protocols**, edited by Ralph Rapley, 1996
64. **Protein Sequencing Protocols**, edited by Bryan John Smith, 1997
63. **Recombinant Proteins: Detection and Isolation Protocols**, edited by Rocky S. Tuan, 1996
62. **Recombinant Gene Expression Protocols**, edited by Rocky S. Tuan, 1996
61. **Protein and Peptide Analysis by Mass Spectrometry**, edited by John R. Chapman, 1996
60. **Protein NMR Protocols**, edited by David G. Reid, 1996
59. **Protein Purification Protocols**, edited by Shawn Doonan, 1996
58. **Basic DNA and RNA Protocols**, edited by Adrian J. Harwood, 1996
57. **In Vitro Mutagenesis Protocols**, edited by Michael K. Trower, 1996
56. **Crystallographic Methods and Protocols**, edited by Christopher Jones, Barbara Mulloy, and Mark Sanderson, 1996
55. **Plant Cell Electroporation and Electrofusion Protocols**, edited by Jac A. Nickoloff, 1995
54. **YAC Protocols**, edited by David Markie, 1995
53. **Yeast Protocols: Methods in Cell and Molecular Biology**, edited by Ivor H. Evans, 1996
52. **Capillary Electrophoresis: Principles, Instrumentation, and Applications**, edited by Kevin D. Altria, 1996
51. **Antibody Engineering Protocols**, edited by Sudhir Paul, 1995
50. **Species Diagnostics Protocols: PCR and Other Nucleic Acid Methods**, edited by Justin P. Clapp, 1996
49. **Plant Gene Transfer and Expression Protocols**, edited by Heddwyn Jones, 1995
48. **Animal Cell Electroporation and Electrofusion Protocols**, edited by Jac A. Nickoloff, 1995
47. **Electroporation Protocols for Microorganisms**, edited by Jac A. Nickoloff, 1995
46. **Diagnostic Bacteriology Protocols**, edited by Jenny Howard and David M. Whitcombe, 1995
45. **Monoclonal Antibody Protocols**, edited by William C. Davis, 1995
44. **Agrobacterium Protocols**, edited by Kevan M. A. Garmland and Michael R. Davey, 1995
43. **In Vitro Toxicity Testing Protocols**, edited by Sheila O'Hare and Chris K. Atterwill, 1995
42. **ELISA: Theory and Practice**, by John R. Crowther, 1995
41. **Signal Transduction Protocols**, edited by David A. Kendall and Stephen J. Hill, 1995
40. **Protein Stability and Folding: Theory and Practice**, edited by Bret A. Shirley, 1995
39. **Baculovirus Expression Protocols**, edited by Christopher D. Richardson, 1995
38. **Cryopreservation and Freeze-Drying Protocols**, edited by John G. Day and Mark R. McLellan, 1995
37. **In Vitro Transcription and Translation Protocols**, edited by Martin J. Tymms, 1995
36. **Peptide Analysis Protocols**, edited by Ben M. Dunn and Michael W. Pennington, 1994
35. **Peptide Synthesis Protocols**, edited by Michael W. Pennington and Ben M. Dunn, 1994
34. **Immunocytochemical Methods and Protocols**, edited by Lorette C. Javois, 1994
33. **In Situ Hybridization Protocols**, edited by K. H. Andy Choo, 1994
32. **Basic Protein and Peptide Protocols**, edited by John M. Walker, 1994
31. **Protocols for Gene Analysis**, edited by Adrian J. Harwood, 1994
30. **DNA-Protein Interactions**, edited by G. Geoff Kneale, 1994
29. **Chromosome Analysis Protocols**, edited by John R. Gosden, 1994
28. **Protocols for Nucleic Acid Analysis by Nonradioactive Probes**, edited by Peter G. Isaac, 1994
27. **Biomembrane Protocols: II. Architecture and Function**, edited by John M. Graham and Joan A. Higgins, 1994
26. **Protocols for Oligonucleotide Conjugates: Synthesis and Analytical Techniques**, edited by Sudhir Agrawal, 1994
25. **Computer Analysis of Sequence Data: Part II**, edited by Annette M. Griffin and Hugh G. Griffin, 1994
24. **Computer Analysis of Sequence Data: Part I**, edited by Annette M. Griffin and Hugh G. Griffin, 1994
23. **DNA Sequencing Protocols**, edited by Hugh G. Griffin and Annette M. Griffin, 1993
22. **Microscopy, Optical Spectroscopy, and Macroscopic Techniques**, edited by Christopher Jones, Barbara Mulloy, and Adrian H. Thomas, 1993
21. **Protocols in Molecular Parasitology**, edited by John E. Hyde, 1993
20. **Protocols for Oligonucleotides and Analogs: Synthesis and Properties**, edited by Sudhir Agrawal, 1993
19. **Biomembrane Protocols: I. Isolation and Analysis**, edited by John M. Graham and Joan A. Higgins, 1993
18. **Transgenesis Techniques: Principles and Protocols**, edited by David Murphy and David A. Carter, 1993

METHODS IN MOLECULAR BIOLOGY™

Protein Sequencing Protocols

Edited by

Bryan John Smith

Celltech Therapeutics, Slough, UK

Humana Press



Totowa, New Jersey


© 1997 Humana Press Inc.
999 Riverview Drive, Suite 208
Totowa, New Jersey 07512

All rights reserved.

No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise without written permission from the Publisher. Methods in Molecular Biology™ is a trademark of The Humana Press Inc.

All authored papers, comments, opinions, conclusions, or recommendations are those of the author(s), and do not necessarily reflect the views of the publisher.

For additional copies, pricing for bulk purchases, and/or information about other Humana titles, contact Humana at the above address or at any of the following numbers: Tel.: 201-256-1699; Fax: 201-256-8341; E-mail: humana@interramp.com

This publication is printed on acid-free paper. 
ANSI Z39.48-1984 (American Standards Institute) Permanence of Paper for Printed Library Materials.

Cover illustration: Fig. 1 from Chapter 3, "Two-Dimensional Polyacrylamide Gel Electrophoresis for the Separation of Proteins for Chemical Characterization," by Michael J. Dunn.

Photocopy Authorization Policy:

Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Humana Press Inc., provided that the base fee of US \$5.00 per copy, plus US \$00.25 per page, is paid directly to the Copyright Clearance Center at 222 Rosewood Drive, Danvers, MA 01923. For those organizations that have been granted a photocopy license from the CCC, a separate system of payment has been arranged and is acceptable to Humana Press Inc. The fee code for users of the Transactional Reporting Service is: [0-89603-353-8/97 \$5.00 + \$00.25].

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

Library of Congress Cataloging-in-Publication Data

Main entry under title:
Methods in molecular biology™.

Protein sequencing protocols / edited by Bryan John Smith.

p. cm. — (Methods in molecular biology ; v. 64)

Includes index.

ISBN 0-89603-353-8 (alk. paper)

1. Nucleotide sequence—Laboratory manuals. I. Smith, Bryan

John. II. Series: Methods in molecular biology (Totowa, NJ) ; v. 64

QP625.N89P76 1997

547.87'328—dc20

96-38610

CIP

Preface

If the development of techniques for the labeling of a polypeptide N-terminus and for repetitive N-terminal sequencing mark the beginning of the science of protein structure determination, then the field has just about reached its half century. In more recent times, recombinant DNA techniques have provided powerful means by which to obtain long protein sequences (by theoretical translation from nucleic acid sequences), but rather than replacing the direct, chemical, protein sequencing approach, they have instead added further impetus to the drive towards a better understanding of posttranslational processing and modification events, as well as identification of novel proteins. Recent years have also seen the advent of “biopharmaceuticals” (i.e., pharmaceutical products that are proteins, such as monoclonal antibodies), and this has meant that protein sequencing has found an important new application as a quality control tool.

Over the decades of protein sequencing many new techniques have been introduced, with the basic aim of generating more information from less material. Some techniques have come and gone, but others have been with us for many years. Edman chemistry is perhaps the best example of the latter class, with its basic principles still being applied today in the newest protein sequencer design. Methods for cleaving and modifying peptides are also long-standing, having been adapted over the years to suit progressively smaller amounts of sample. To this end, too, methods for amino acid and peptide separation have been developed to give better sensitivity as well as better resolution—capillary electrophoresis and microbore HPLC are comparatively recent examples from this area. Though repetitive N-terminal sequencing using Edman chemistry has become well-established, sequencing from the C-terminus has proven a more difficult problem to solve. In years past, enzymatic digestion by carboxypeptidase and analysis of products released provided the most practical means to do this, but it had its limitations—not least its (relative) insensitivity and the (usually) short sequence obtained. Various laboratories around the world have attempted to develop reliable chemical methods for C-terminal sequencing and this effort is now, in the mid-1990s, coming to fruition in the form of a more robust chemistry that has been automated and made available on commercial machines. Perhaps the most significant development in recent times, however, has been the adaptation of mass spectrometry to the analysis

of biological macromolecules. This clutch of techniques, particularly electrospray and laser methods, have improved rapidly in accessibility and sensitivity, and now provide a valuable new tool that is complementary to “traditional” sequencing methods. The ideal lab would have all of these techniques in-house, of course, but reality seldom approaches the ideal. Nevertheless, there is great value in keeping abreast with developments in the field and knowing what is technically possible in other labs, this knowledge often being the first step in a fruitful collaboration. Books such as this one are intended to aid in this process.

Just as the whole field has evolved during its first half century, then surely it will continue to change through the next. Inevitably, this continuing process makes any compilation of methods something of a “snapshot” of a field at a single moment in time, but that does not detract from the usefulness of the exercise of pulling together current methods, for handy reference in the lab. That is the aim of *Protein Sequencing Protocols*. It describes techniques in current use for study of protein sequence. Some of them are well-established, others are new. The capital investment in equipment required for some methods in this field can be large—prohibitively so for some labs. In recognition of this, some parts of this book cover manual methods that require less in capital outlay, and which may prove useful where a particular piece of equipment is not available. It is probably true to say, however, that most protein sequencing these days is carried out on automated sequencers and/or with mass spectrometers. Accordingly, much of *Protein Sequencing Protocols* concerns these approaches. Amino acid analysis and peptide mapping are also included here, since, of course, they each reflect protein sequence and both are important techniques in the quality control of biopharmaceutical production. In this area, older techniques such as thin layer chromatography and paper electrophoresis are passed over in favor of HPLC and capillary electrophoresis, since one or both of these is available in the majority of labs today. Though such subjects as these analytical techniques are covered, particularly in later chapters, earlier chapters refer to the (not insignificant) problems of obtaining and handling the protein in the first place. This first stage can prove most challenging, as any experienced protein biochemist will testify. It is unfortunately true that each new protein behaves differently from the last one! The field of protein purification, as a whole, is beyond the scope of this book, but some methods of particular use in the handling of very small amounts of protein are included here, being relevant to the usual position (with novel proteins, at least) of having only micrograms or less to work with. Likewise, protein cleavage and modification techniques acknowledge this problem of working with small samples.

Protein Sequencing Protocols is laid out in the same style as other books in this series. It is intended as a laboratory manual, guiding the experimenter to a successful result. Two appendices are included for ease of reference during analysis of sequence data. These are summaries of masses of amino acids and modifications, and of consensus sequences. The book is, of course, aimed at those with access to all the necessary equipment. It is also intended, however, that those without all the necessary equipment will find this collection of methods valuable, not only for understanding of what the techniques involve, but also as aid in collaboration—how much sample, in what form, is required for a willing colleague to obtain a successful analysis. To all of these experimenters, I would like to dedicate this volume, together with all the contributors and, last but not least, Stef and Paul, without whose support this book would have been impossible.

Bryan John Smith

Contents

Preface.....	v
Contributors	xi
1 Strategies for Handling Polypeptides on a Microscale, <i>Bryan John Smith and Paul Tempst</i>	1
2 SDS Polyacrylamide Gel Electrophoresis for N-Terminal Protein Sequencing, <i>Bryan John Smith</i>	17
3 Two-Dimensional Polyacrylamide Gel Electrophoresis for the Separation of Proteins for Chemical Characterization, <i>Michael J. Dunn</i>	25
4 Electrophotting of Proteins from Polyacrylamide Gels for Chemical Characterization, <i>Michael J. Dunn</i>	37
5 Enzymatic Cleavage of Polypeptides, <i>Bryan John Smith and Colin Wheeler</i>	43
6 Chemical Cleavage of Polypeptides, <i>Bryan John Smith</i>	57
7 Preparation of Glycopeptides, <i>Chris W. Sutton and Jacqui A. O'Neill</i>	73
8 Selective Isolation of the Carboxy Terminal Peptide from a Protein, <i>Tomoko Hayashi Akiyama and Tatsuru Sasagawa</i>	81
9 Selective Isolation of the Amino Terminal Peptide from α -Amino Blocked Protein, <i>Tomoko Hayashi Akiyama and Tatsuru Sasagawa</i>	85
10 Analytical and Micropreparative Capillary Electrophoresis of Peptides, <i>Alan J. Smith</i>	91
11 Reverse-Phase HPLC Purification of Peptides from Natural Sources for Structural Analysis, <i>Chris Shaw</i>	101

12	High-Performance Liquid Chromatography On-Line Derivative Spectroscopy for the Characterization of Peptides with Aromatic Amino Acid Residues, Christoph W. Turck	109
13	Hydrolysis of Samples for Amino Acid Analysis, Ian Davidson	119
14	Amino Acid Analysis: <i>Precolumn Derivatization Methods</i> , G. Brent Irvine	131
15	Postcolumn Amino Acid Analysis, Alan J. Smith	139
16	Fast-Atom Bombardment and Electrospray Mass Spectrometry of Peptides, Proteins, and Glycoproteins, Fiona M. Greer and Howard R. Morris	147
17	Peptide Mass Fingerprinting Using MALDI-TOF Mass Spectrometry, Darryl J. C. Pappin	165
18	N-Terminal Protein Ladder Sequencing, Rong Wang and Brian T. Chait	175
19	The Dansyl-Edman Method for Manual Peptide Sequencing, John M. Walker	183
20	The Dansyl Method for Identifying N-Terminal Amino Acids, John M. Walker	189
21	Identification of PTH-Amino Acids by High-Performance Liquid Chromatography, Gregory A. Grant and Mark W. Crankshaw	197
22	Protein Sequencer Maintenance and Troubleshooting, Bryan Dunbar	217
23	Identifying the Carboxy Terminal Amino Acid of a Protein, Tomoko Hayashi Akiyama and Tatsuru Sasagawa	235
24	Enzymatic and Chemical Methods for Manual C-Terminal Peptide Sequencing, Franca Casagrande and John F. K. Wilshire	243
25	Automated Methods for C-Terminal Protein Sequencing Jerome M. Bailey and Chad G. Miller	259
26	Chemical Modification of Proteins for Sequence Analysis Alex F. Carne and Sally U	271

27	On-Membrane Deblocking of Proteins Hisashi Hirano, Setsuko Komatsu, and Susumu Tsunasawa	285
28	Analysis of Sites of Protein Phosphorylation Alastair Aitken and Michèle Learmonth	293
29	Identification of Sites of Glycosylation Chris W. Sutton and Jacqui A. O'Neill	307
30	Quantification and Location of Disulfide Bonds in Proteins Alastair Aitken and Michèle Learmonth	317
Appendix 1: Letter Codes, Structures, Masses, and Derivatives of Amino Acids James P. Turner and Bryan John Smith		329
Appendix 2: A Library of Consensus Sequences Alastair Aitken		357
Index		369

Contributors

- ALASTAIR AITKEN • *Laboratory of Protein Structure, National Institute for Medical Research, London, UK*
- TOMOKO HAYASHI AKIYAMA • *Toray Research Center, Kamakura, Kanagawa, Japan*
- JEROME M. BAILEY • *Hewlett-Packard, Protein Chemistry Systems, Palo Alto, CA*
- FRANCA CASAGRANDE • *European Molecular Biology Lab, Heidelberg, Germany*
- ALEX F. CARNE • *Institute of Cancer Research, London, UK*
- BRIAN T. CHAIT • *Rockefeller University, New York, NY*
- MARK W. CRANKSHAW • *Washington University School of Medicine, St. Louis, MO*
- IAN DAVIDSON • *Department of Molecular and Cellular Biology, University of Aberdeen, UK*
- BRYAN DUNBAR • *Department of Molecular and Cellular Biology, University of Aberdeen, UK*
- MICHAEL J. DUNN • *Department of Cardiothoracic Surgery, National Heart and Lung Institute, Heart Science Centre, Harefield, UK*
- GREGORY A. GRANT • *Washington University School of Medicine, St. Louis, MO*
- FIONA M. GREER • *M-Scan, Silwood Park, Ascot, UK*
- HISASHI HIRANO • *Yokohama City University, Kihara Institute for Biological Research, Yokohama, Japan*
- G. BRENT IRVINE • *Division of Biochemistry, School of Biology and Biochemistry, Medical Biology Centre, Queen's University of Belfast, Northern Ireland*
- SETSUKO KOMATSU • *Yokohama City University, Kihara Institute for Biological Research, Yokohama, Japan*
- MICHÈLE LEARMONTH • *Laboratory of Protein Structure, National Institute for Medical Research, London, UK*
- CHAD G. MILLER • *Hewlett Packard, Protein Chemistry Systems, Palo Alto, CA*
- HOWARD R. MORRIS • *Department of Biochemistry, Imperial College, London, UK*
- JACQUI A. O'NEILL • *Parke-Davis Neuroscience Research Centre, Cambridge, UK*

- DARRYL J. C. PAPPIN • *Protein Isolation and Cloning Lab, Imperial Cancer Research Fund, London, UK*
- TATSURU SASAGAWA • *Toray Research Centre, Kamakura, Kanagawa, Japan*
- CHRIS SHAW • *Wellcome Research Lab, Royal Victoria Hospital, Belfast, Northern Ireland*
- ALAN J. SMITH • *Beckman Center, Stanford University Medical Center, Stanford, CA*
- BRYAN JOHN SMITH • *Celltech Therapeutics, Slough, UK*
- CHRIS W. SUTTON • *Thermo Bioanalysis, Hemel Hempstead, UK*
- PAUL TEMPST • *Cancer Center, Memorial Sloan Kettering, New York, NY*
- SUSUMU TSUNASAWA • *Kihara Institute for Biological Research, Yokohama City University, Yokohama, Japan*
- CHRISTOPH W. TURCK • *Department of Medicine, Howard Hughes Medical Institute, University of California San Francisco, CA*
- JAMES P. TURNER • *Celltech Therapeutics, Slough, UK*
- SALLY U • *Thermo Bioanalysis, Hemel Hempstead, UK*
- JOHN M. WALKER • *University of Hertfordshire, Hatfield, UK*
- RONG WANG • *Rockefeller University, New York, NY*
- COLIN WHEELER • *Department of Cardiothoracic Surgery, Heart Science Center, National Heart and Lung Institute, Harefield, UK*
- JOHN F. K. WILSHIRE • *Division of Biomolecular Engineering, Melbourne Laboratory and Division Headquarters, CSIRO, Parkville, Australia*