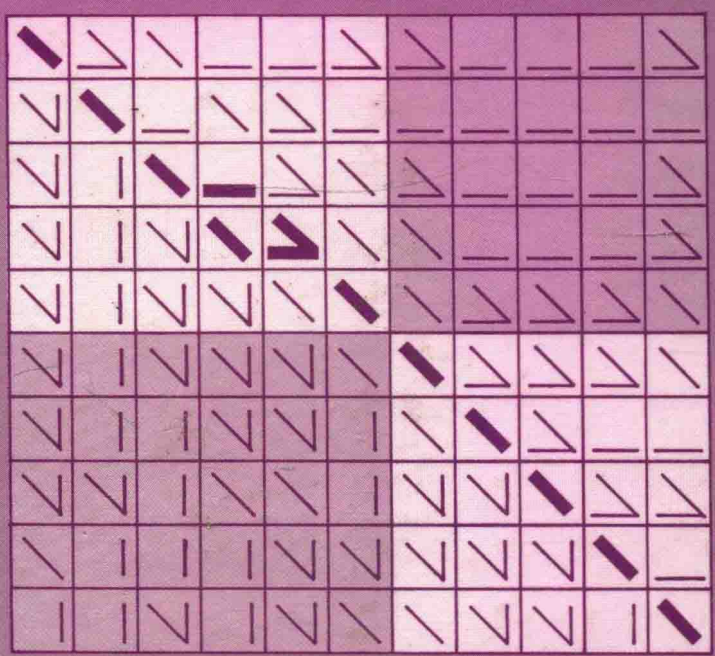
Macromolecular Sequencing and Synthesis

Selected Methods and Applications

Edited by

David H. Schlesinger



Alan R. Liss, Inc., New York

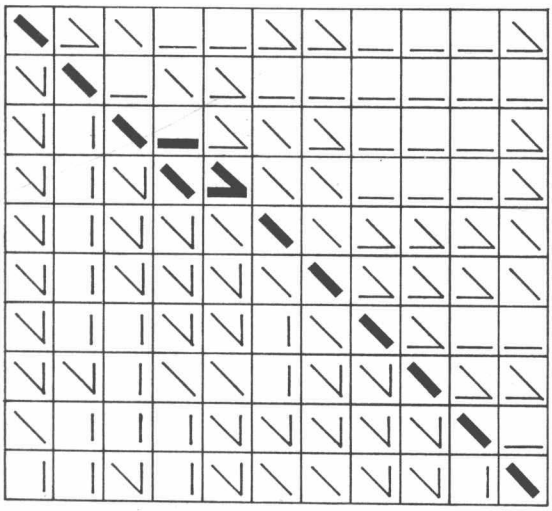
Macromolecular Sequencing and Synthesis

Selected Methods and Applications

Edited by

David H. Schlesinger

Department of Experimental Medicine and Cell Biology
and Member of The Kaplan Cancer Center
New York University Medical Center
New York, New York



Alan R. Liss, Inc., New York

Address all Inquiries to the Publisher
Alan R. Liss, Inc., 41 East 11th Street, New York, NY 10003

Printed in the United States of America

Library of Congress Cataloging in Publication Data

Macromolecular sequencing and synthesis : selected methods and applications / editor, David H. Schlesinger.
p. cm.
Includes index.
ISBN 0-8451-4246-1
1. Amino acid sequence. 2. Nucleotide sequence. 3 Peptides—
—Synthesis. 4. Nucleotides—Synthesis. I. Schlesinger, David H.
[DNLM: 1. Amino Acid Sequence. 2. Macromolecular Systems.
3. Peptides—chemical synthesis. 4. Proteins—chemical synthesis.
QU 68 M174]
QP551.M23 1988
574.19'245—dc19
DNLM/DLC
for Library of Congress 87-33920
CIP

Macromolecular Sequencing and Synthesis

Selected Methods and Applications

Dedication

This book is dedicated to my two daughters, Sarah and Karen, whose curiosity and interest in current methods in molecular biology have made the compilation of this volume a stimulating and rewarding challenge and pleasure.

D.H.S.

Title page illustration:

Path matrix for optimal alignment of amino acid sequences

(from George et al., chapter 12, page 139).

Contributors

Tapan Audhya, Department of Immunobiology, Biotech Division, Ortho Pharmaceutical Corporation, Raritan, NJ 08869-0602 **[75,101]**

Winona C. Barker, Protein Identification Resource, National Biomedical Research Foundation, Georgetown University Medical Center, Washington, DC 20007 **[127]**

Hal Beilan, Chemical Research and Development, Applied Biosystems, Inc., Foster City, CA 94404 **[199]**

B. John Bergot, Chemical Research and Development, Applied Biosystems, Inc., Foster City, CA 94404 **[199]**

Michael S. Bernatowicz, Cellular and Molecular Research Laboratory and Cardiac Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114; present address: Milligen Division, Millipore Corporation, Bedford, MA 01730 **[195]**

Steven A. Carr, Department of Physical and Structural Chemistry, Smith Kline and French Laboratories, King of Prussia, PA 19406-0939 **[83]**

Ben M. Dunn, Departments of Biochemistry and Molecular Biology, University of Florida College of Medicine, J.H. Miller Health Center, Gainesville, FL 32610 **[243]**

J. William Efcavitch, Department of Research and Development, Applied Biosystems, Inc., Foster City, CA 94404 **[221]**

Tim Geiser, Chemical Research and Development, Applied Biosystems, Inc., Foster City, CA 94404 **[199]**

David G. George, Protein Identification Resource, National Biomedical Research Foundation, Georgetown University Medical Center, Washington, DC 20007 **[127]**

Mark E. Hemling, Department of Physical and Structural Chemistry, Smith Kline and French Laboratories, King of Prussia, PA 19406-0939 **[83]**

Charles S. Hollander, Department of Medicine, New York University Medical Center, New York, NY 10016 **[75]**

Richard A. Houghten, Department of Molecular Biology MB 14, Scripps Clinic and Research Foundation, La Jolla, CA 92037 **[185]**

Michael W. Hunkapiller, Department of Research and Development, Applied Biosystems, Inc., Foster City, CA 94404 **[45,117]**

Lois T. Hunt, Protein Identification Resource, National Biomedical Research Foundation, Georgetown University Medical Center, Washington, DC 20007 **[127]**

E.T. Kaiser, Laboratory of Bioorganic Chemistry and Biochemistry, Rockefeller University, New York, NY 10021-6399 **[235]**

William R. Kem, Departments of Pharmacology and Therapeutics, University of Florida College of Medicine, J.H. Miller Health Center, Gainesville, FL 32610 **[243]**

Debra A. Kendall, Laboratory of Bioorganic Chemistry and Biochemistry, Rockefeller University, New York, NY 10021-6399 **[235]**

The number in brackets is the opening page number of the contributor's article.

Gary R. Matsueda, Cellular and Molecular Research Laboratory, and Cardiac Unit, Massachusetts General Hospital; and Department of Pathology, Harvard Medical School, Boston, MA 02114 **[195]**

Kenneth M. Otteson, Chemical Research and Development, Applied Biosystems, Inc., Foster City, CA 94404 **[199]**

Michael W. Pennington, Departments of Biochemistry and Molecular Biology, University of Florida College of Medicine, J.H. Miller Health Center, Gainesville, FL 32610 **[243]**

Gerald D. Roberts, Department of Physical and Structural Chemistry, Smith Kline and French Laboratories, King of Prussia, PA 19406-0939 **[83]**

Julie Ivory Rushbrook, Department of Biochemistry, College of Medicine and Dentistry, Rutgers Medical School, New Jersey, Piscataway, NJ 08854; present address: Department of Biochemistry, SUNY-Health Science Center at Brooklyn, Brooklyn, NY 11203 **[3]**

Andrew V. Schally, Department of Medicine, School of Medicine, Tulane University Medical Center, New Orleans, LA 70112 **[25]**

David H. Schlesinger, Department of Experimental Medicine and Cell Biology, and Member of The Kaplan Cancer Center, New York University Medical Center, New York, NY 10016 **[xi,25,35,59,65,75,251]**

Kathryn L. Stone, Department of Protein and Nucleic Acid Chemistry, Yale University, New Haven, CT 06510 **[7]**

James P. Tam, Department of Biochemistry, Rockefeller University, New York, NY 10021 **[153]**

Kenneth R. Williams, Department of Protein and Nucleic Acid Chemistry, Yale University, New Haven, CT 06510; present address: Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06510 **[7]**

Preface

Protein chemists and molecular biologists of the future will look back upon the 1980s as the renaissance of protein and DNA chemistry, a period marked by an explosive improvement of their respective technologies in sensitivity, rapidity, and automation. In contrast, the late 1970s and early 1980s may be regarded as the rise and fall of protein sequencing due to the rapidity and ease in the sequencing of DNA, a technique that convinced some researchers that protein sequencing was obsolete and on the verge of its demise.

The change in strategy in protein structure determination from the 1970s to the mid-1980s reflects this change from imminent demise to renaissance. The approach for the complete structure determination of a large protein in the 1970s was to isolate many peptide fragments and to sequence and align them in the protein chain by identity with overlapping segments. By contrast, a commonly used strategy of the mid-1980s is to determine the amino terminal or internal sequence of a large protein; chemically synthesize this peptide segment; prepare antibodies to the synthetic peptide, or prepare a cDNA probe coding for this segment; screen a gt10 or gt11 library, for example, with the cDNA probe or antibody to isolate the gene coding for the entire protein; sequence this DNA and rapidly deduce the entire structure of the protein of interest. Thus, in many ways the two techniques are in synergy.

Most investigators today look with optimism to this growing synergism between the determination of DNA and protein structure. For example, an amino acid sequence of a protein can be deduced from the sequence of its genes, but this deduction indicates nothing about the posttranslational processing required for determining the functionality of many mature proteins. For the role of a mature protein to be fully understood, the structure of the fully processed, active protein must be determined by protein chemical methods.

This volume is divided into two major parts. The first is devoted to protein and DNA sequencing. It is introduced by a short chapter on protein isolation methods. Initially, a section on electroblotting was planned along with a short chapter on electroelution in purifying peptides and proteins but, since there still remain numerous problems with electroblotting, such as low capacity of the glass filters and low recoveries, this section has been omitted. Also in this first half of the book, one chapter is included on the structure determination of relatively complex glycoproteins using mass spectrometry.

With respect to DNA structure determination, both current enzymatic and chemical methods are described, with the very new automated DNA sequencer being highlighted. This machine enables researchers to consider for the first time the possibility of sequencing the entire human genome. The first half of the book ends with a discussion of computer analysis of both DNA and protein structures.

The second part of this volume is itself divided in half, with one section devoted to peptide synthesis (examples and techniques) and the other to chemical synthesis of DNA probes and genes. Recent improvements of both peptide and DNA synthesis have paralleled the rapid advances in protein and DNA sequence determination. Examples of rapid synthesis of large polypeptides with up to 140 amino acids are now beginning to appear in the literature. In addition, novel methods have also been developed recently for the extremely rapid and simultaneous synthesis of multiple peptides and improvements in the cleavage of peptides from their support and removal of protecting groups have been advanced. Similarly, the improvement and refinement of the chemical synthesis of DNA probes have resulted in the ability to prepare large DNA probes in excess of 150 nucleotides in length in only two or three days (with cycles of nucleotide addition of only 20 minutes) at very low cost (about $4.00/nucleotide).

Methods for purifying these probes have also been improved and simplified while chemical methods for condensing the DNA segments have recently been developed so that synthesis of entire genes is now feasible. Finally, the technique of site-directed mutagenesis enables scientists to design new genes coding for proteins of altered or improved function such as enzymes with altered specificity and kinetics. Advances such as these may make possible the use of gene therapy in treating human disease.

This volume concludes with selected examples in which these techniques have been combined to attack complex biological problems (some with potential clinical implications) which it has not, until now, been possible to approach. For these reasons we have chosen this time to prepare this volume, aware that it is truly an exciting time for science.

David H. Schlesinger
New York

Contents

PURIFICATION OF PEPTIDES AND PROTEINS

Macromolecular Sequencing and Synthesis
Selected Methods and Applications, pages 3–6

1

Electroelution of the Subunits of Beef Heart NAD-Dependent Isocitrate Dehydrogenase

Julie Ivory Rushbrook

INTRODUCTION

NAD-isocitrate dehydrogenase is a complex enzyme responding to numerous allosteric effectors and to modulation by its own substrates. The regulatory molecules, which include NADH, NADPH, ATP, and ADP (Chen and Plaut, 1963), reflect the energy and redox states of the cell, in keeping with the role of the enzyme as a control point for the tricarboxylic acid cycle. While kinetic and binding studies have been carried out for many years, it is relatively recently that the complexities of the protein's subunit composition have become apparent.

Early work established that the enzymes from beef heart and pig heart contained subunits of 40,000 daltons which aggregated to form multiples of four units (Giorgio et al., 1970; Shen et al., 1974; Fan et al., 1975). A consideration of the numerous allosteric effectors suggests that these subunits might well be nonidentical. This was shown to be the case for the enzymes from both sources—four nonidentical species of molecular weights close to 40,000 being identified in the enzyme from beef heart (Rushbrook and Harvey, 1978), three in that from pig heart (Ramachandran and Coleman, 1980).

In the case of the beef-heart enzyme, heterogeneity was first apparent on SDS-gel electrophoresis of a highly purified preparation (Rushbrook and Harvey, 1978). Four closely spaced bands were found, separated each from the other by approximately 1 mm, present in the ratio 1:1:0.9:0.4. Nondenaturing first-dimension polyacrylamide gel elec-

Department of Biochemistry, SUNY-Health Science Center at Brooklyn, Brooklyn, NY

This work was carried out while the author was a graduate student in the Department of Biochemistry, College of Medicine and Dentistry, Rutgers Medical School, New Jersey, Piscataway, NJ

trophoresis at several pH values followed by second-dimension SDS-gel electrophoresis showed that the enzymatically active band in the first dimension gave rise to the four bands, confirming these as subunits of the enzyme. The subunits were subsequently shown to differ by electrophoresis in urea and by two-dimensional peptide-mapping (Rushbrook and Harvey, 1978). The electroelution procedure described here provided purified subunits for peptide-mapping.

MATERIALS AND METHODS

Chemicals were from the sources described previously (Rushbrook and Harvey, 1978). Methods other than electroelution were carried out as previously described (Rushbrook and Harvey, 1978).

Electroelution was carried out in a Buchler Polyanalyst tube gel apparatus (Haake-Buchler Instruments, Inc., 244 Saddle River Rd., Saddle Brook, NJ, 07662). Units supplied by Biorad (Biorad Laboratories, 1414 Harbour Way South, Richmond, CA 94804) or Hoeffer (Hoeffer Scientific Instruments, 654 Minnesota St., San Francisco, Ca 94107) may also be used. See also the Discussion section.

ELECTROELUTION PROCEDURE

Separation of the subunits in a 10% polyacrylamide SDS–slab gel is shown in Figure 1–1. For preparative runs, 3–mm–thick slab gels were used and approximately 40 μg of protein was applied/0.8-mm-length well, for a total of 0.51 mg (3.2 nmol). The gel was stained briefly for 10 minutes in 0.25% Coomassie Brilliant Blue in 20% trichloracetic acid and destained for 10 minutes in 7.5% acetic acid, 5% *tert*-butyl alcohol. The bands were cut out with a razor blade; like bands were combined and inserted into electrophoresis tubes whose lower ends were securely covered with nylon net (Fig. 1–2). A piece of dialysis tubing was knotted securely three times at one end and the other end was placed over the net covering the gel tube. The upper end of each assemblage was inserted into a grommet in the upper buffer compartment of a Buchler Polyanalyst tube gel apparatus, and the free end of the dialysis tubing was securely fastened to the upper part of the gel tube.

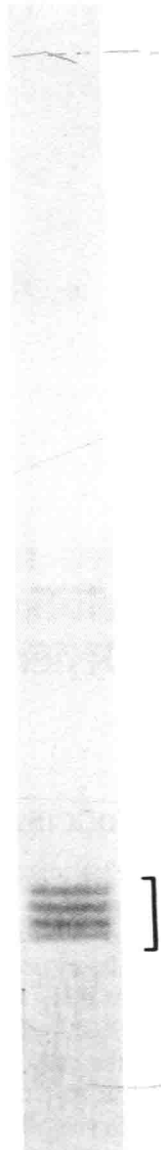


Fig. 1–1. *Separation of the four subunits of NAD-dependent isocitrate dehydrogenase by SDS-gel electrophoresis. The subunits (bracket) were separated on a 10% polyacrylamide gel; 1.8 μg protein/well.*

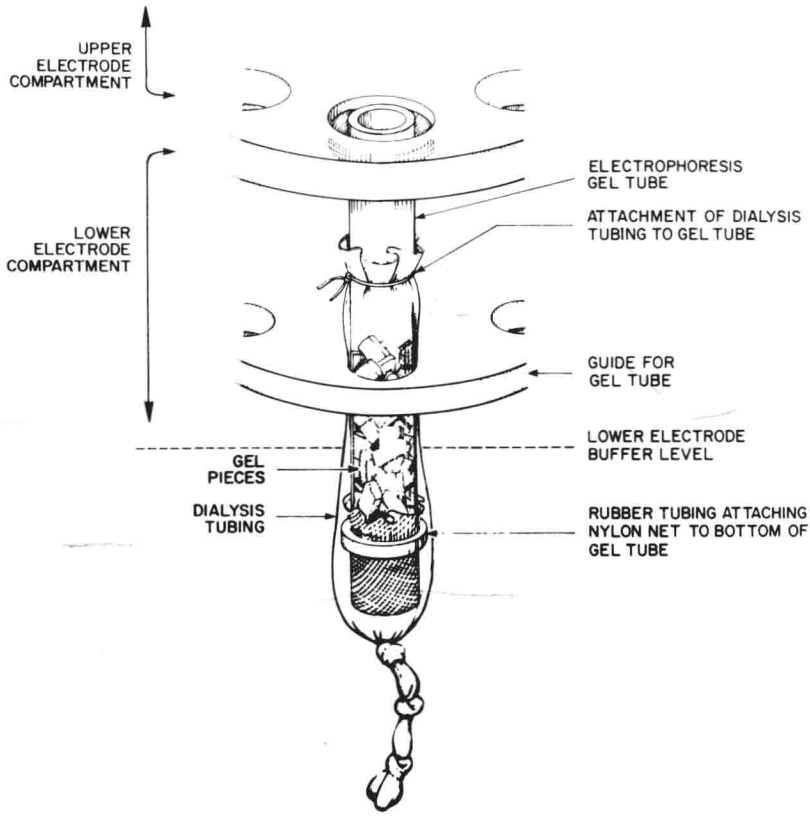


Fig. 1–2. *Modification to gel tube for electrophoretic elution. Gel pieces containing protein are retained in the tube by nylon net securely fastened to the bottom of the tube. Dialysis tubing, knotted three times, encloses the lower end of the tube and serves as the reservoir into which protein elutes. The lower buffer level must be well below the top of the dialysis tubing.*

Electrophoresis buffer (0.05 M Tris chloride, 0.5% SDS, 0.5% mercaptoethanol, pH 8.0) was placed in the upper and lower buffer compartments, care being taken that the lower buffer level was well below the top of the dialysis tubing around each gel tube. Electrophoresis was carried out at room temperature for 21 hours at 80 V constant voltage. Starting and ending currents were approximately 24 and 9 mA respectively. The integrity of each dialysis tubing receiving chamber was indicated by the presence of the Coomasie Brilliant Blue dye used to stain the protein.

The Gel tube assemblies were removed from the apparatus. Elution buffer above the gel slices in each tube was discarded and that inside the dialysis tubing was carefully collected. Eluted protein was precipitated by the addition of 10 volumes of a cold acidified acetone solution (1:40, 1 N HCl:acetone), separating protein from both SDS and Coomassie Brilliant Blue. The subunits were reduced and alkylated in guanidine hydrochloride and subjected to peptide-mapping as described (Rushbrook and Harvey, 1978). From a comparison of the intensity of individ-