THE PROTEINS

Composition, Structure, and Function

SECOND EDITION

Edited by

HANS NEURATH
Department of Biochemistry
University of Washington
Seattle, Washington

VOLUME I

1963

ACADEMIC PRESS
New York and London

ACADEMIC PRESS INC.
111 Fifth Avenue, New York 3, New York

United Kingdom Edition published by
ACADEMIC PRESS INC. (LONDON) LTD.
Berkeley Square House, London W.1

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 63-23023

PRINTED IN THE UNITED STATES OF AMERICA

List of Contributors

CHRISTIAN B. ANFINSEN, *Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts*

ROBERT E. CANFIELD, *Department of Medicine, College of Physicians and Surgeons, Columbia University, New York*

R. CECIL, *Department of Biochemistry, University of Oxford, England*

JOSEPH S. FRUTON, *Department of Biochemistry, Yale University, New Haven, Connecticut*

KLAUS HOFMANN, *Biochemistry Department, University of Pittsburgh, School of Medicine, Pittsburgh, Pennsylvania*

PANAYOTIS G. KATSOYANNIS, *Biochemistry Department, University of Pittsburgh, School of Medicine, Pittsburgh, Pennsylvania*

ALBERT LIGHT, *Department of Biological Chemistry, University of California Medical School, Los Angeles, California*

HAROLD A. SCHERAGA, *Department of Chemistry, Cornell University, Ithaca, New York*

EMIL L. SMITH, *Department of Biological Chemistry, University of California Medical School, Los Angeles, California*

R. H. SMITH, *British Glues and Chemical Ltd., London, England*

G. R. TRISTRAM, *Department of Physiology and Biochemistry, University of St. Andrews, St. Andrews, Fife, Scotland*

Preface to the Second Edition

A decade has passed since the first edition of "The Proteins" was published. During these years the field of protein chemistry has advanced at a revolutionary pace and has reached new heights of achievement and understanding. New approaches to the study of the detailed structure of the protein molecule have been found, and deep new insights have been gained. The protein molecule now presents itself to the observer in a detail approaching atomic dimensions in resolution, and challenges the keenest mind to translate this image into terms which are meaningful for a comprehension of all the biological properties which we attribute or relate to proteins as a group. As a result, the impact of protein chemistry on biological phenomena becomes magnified in scope and depth and the day draws nearer when the mechanism of action of enzymes, antibodies, protein hormones, and the operation of the biosynthetic and genetic mechanisms, all will become understood in terms of the structure and properties of the constituent protein molecules. The question is no longer simply what are proteins made of, but rather how can the composition and structure of proteins be related to their specific biological functions? The subtitle chosen for the second edition is intended to focus attention on these relationships.

A decade ago, the complete amino acid sequence of a protein, insulin, had just been elucidated for the first time. Protein chromatography was still in its infancy and the image of the protein molecule was blurred, figuratively and literally. Today, while still a painstaking undertaking, sequential amino acid analysis has been extended to many proteins, small and large, and sufficiently pinpointed to delineate genetic relationships among proteins, or to determine components of the active center of enzymes. Experimental procedures that seemed feasible then, today are indispensable and often automated tools. Perhaps the most spectacular and far-reaching advances of the past decade relate to the three-dimensional synthesis of the fine structure of crystalline proteins. The models which are emerging are sufficiently detailed, down to angstrom units, to test hypotheses derived from more indirect experimental approaches, and to raise questions and make predictions which were beyond the most hopeful expectations of the recent past. Of no lesser importance are the

successes achieved in the chemical synthesis of peptides which promise to herald the complete synthesis of more complex protein molecules; for the ultimate test of our notions of the structure of molecules, small or large, rests with their chemical synthesis.

In the light of these rapid and fundamental advances of the last decade it appeared neither practical nor desirable merely to expand and update the first edition, chapter by chapter. Instead, an effort has been made to develop new approaches to the subject of protein chemistry and to emphasize wherever possible the relation between protein composition, structure, and function. Prime consideration has been placed on quantitative relationships rather than on encyclopedic coverage of the field.

The preparation of the current edition was initially based on a logical sequence of topics. Practical considerations prevailed in the end, and manuscripts were included in a volume more nearly in order of their receipt by the editor.

It was with deep and sincere regrets that the present editor had to undertake his task without the invaluable counsel and cooperation of his friend and colleague, Kenneth Bailey, the co-editor of the first edition, whose untimely death is being mourned by protein chemists all over the world. The second edition of this treatise is dedicated to his memory.

Once again, it is a pleasure to acknowledge the cooperation of the publisher, Academic Press, in the preparation of this treatise, and to thank Dr. Virginia Dewey for the preparation of the indexes which will accompany each volume.

HANS NEURATH

Seattle, Washington
September, 1963

ERRATA

THE PROTEINS, VOLUME I, *second edition*

(HANS NEURATH, *editor*)

Page 249, line 13, rspect should read respect

Page 475, line 44, Walsh, K. A., Sampath, V. S., Kumar, K. S. V., Bargetzi, J. P., and Neurath, H. should read Walsh, K. A., Sampath-Kumar, K. S. V., Bargetzi, J. P., and Neurath, H.

Page 495, line 9 of footnote 5, netropy should read entropy

Contents

Chapter 3

Chemical Aspects of Protein Synthesis

JOSEPH S. FRUTON

Chapter 4

Concepts and Experimental Approaches in the Determination of the Primary Structure of Proteins

ROBERT E. CANFIELD AND CHRISTIAN B. ANFINSEN

Chapter 5

Intramolecular Bonds in Proteins
I. The Role of Sulfur in Proteins

R. CECIL

Chapter 6

Intramolecular Bonds in Proteins
II. Noncovalent Bonds

HAROLD A. SCHERAGA

Chapter 1

Amino Acid Analysis of Peptides and Proteins

ALBERT LIGHT AND EMIL L. SMITH

I. General Introduction

Recent developments in methods of amino acid analysis now permit rapid analysis with high precision at a micromole level. Application of the methods to the analysis of proteins has led to the elucidation of the composition of many proteins (see Addendum). Furthermore, the newer analytical techniques are being applied to the study of various aspects of protein structure. The place of amino acid analysis is illustrated by an examination of the steps required for the elucidation of the amino acid sequence of a protein[1] (Edsall and Wyman, 1958; Anfinsen and Redfield, 1956; Desnuelle, 1953; Hill *et al.* (1959). A highly purified protein, well-characterized by various methods, is analyzed to calculate the number of residues of each amino acid. If the molar ratios approach integral values, this can be taken as supporting evidence for the homogeneity of the sample. In addition, the absence of one or more amino acids also indicates the purity of the sample, e.g., the complete lack of isoleucine in highly purified samples of human hemoglobin A (Stein *et al.*, 1957).

In his investigations on insulin, Sanger (1959) noted the advantages to be gained from the study of small peptides obtained after enzymic and partial acid hydrolysis. From an analysis of the isolated peptides, the composition is calculated as molar ratios. The purity of the peptide can be judged by relating the molar ratios to theoretical values for the peptide. Finally, the composition of the purified peptide components must be equal to the composition of the protein. Indeed, an accurate estimate of the composition of a protein can be derived from the summation of its component parts. For example, the composition of ribonuclease deduced from the study of peptides formed after digestion with trypsin showed that the analysis of the protein itself was in error by two amino acids (Hirs *et al.*, 1956).

The determination of peptide sequences usually requires several techniques, but common to all is the use of quantitative methods of amino acid analysis. A correlation of the composition and amount of peptide before and after a degradative step clearly indicates the distribution of the resulting fragments. Since degradative reactions are frequently incomplete, the availability of quantitative data adds strength to the conclusions. Inasmuch as methods of separation of peptides are seldom quantitative, the recovery of the isolated peptides can be determined by amino acid analysis, and each purification step can be evaluated with respect to yield.

[1] See Chapter 4.

The mechanism of action of several biologically active proteins and peptides is currently under intensive study (Anfinsen and Redfield, 1956; Anfinsen, 1961; Koshland, 1960; Smith *et al.*, 1962; Stark *et al.*, 1961; Dixon *et al.*, 1958b; Vithayathil and Richards, 1961). The active site and uniquely functional groups can be chemically modified to form derivatives with covalent bonds stable to acid hydrolysis. Analysis of the protein derivative supplies information on the groups involved in the reaction and on the extent of conversion. Two approaches are used to obtain this information. In the first, the amino acid derivative is isolated and estimated by appropriate procedures, whereas the second involves a subtractive approach, by identifying the unique amino acid by its disappearance from the protein.

The scope of the operational problems involved in the analysis of proteins and peptides can be appreciated by considering the individual steps of this process (Martin and Synge, 1945; Block, 1945; Tristram, 1949; Block and Bolling, 1951; Springall, 1954; Block, 1960; Moore and Stein, 1963):

(1) preparation of a sample suitable for hydrolysis and evaluation of ash, moisture, nitrogen, and sulfur content;
(2) hydrolysis in 6 *N* hydrochloric acid;
(3) separation and estimation of amino acids;
(4) determination of tryptophan, methionine, cystine, and cysteine by independent methods;
(5) calculation of the amino acid content as percentage by weight, as residue numbers, and as molar ratios;
(6) calculation of weight and nitrogen recoveries.

The following text will discuss each step, cite various methods, and emphasize the currently preferred procedures and the precision to be expected. Major emphasis will be devoted to quantitative procedures suitable for the determination of those amino acids found in proteins. Many of these same procedures, however, are suitable for the separation and estimation of amino acids found as constituents of naturally occurring peptides and as free amino acids in tissues and body fluids.

II. Preparation of Material for Analysis

A protein suitable for analysis must satisfy the standard tests for homogeneity. An accounting of the protein in terms of its components (nitrogen, sulfur, phosphorus, etc.) and weight is necessary in order to evaluate the data from subsequent analyses for completeness of recovery of all components.

A. Moisture and Ash Determinations

The use of inorganic salts in the purification of many proteins may lead to rather large salt contents in the protein preparation. Before proceeding further the sample must be brought to an isoionic condition by one of the following procedures: (*a*) exhaustive dialysis against 0.1 *M* potassium chloride to displace protein-bound polyions by potassium ions, followed by dialysis against deionized water to remove the potassium chloride; (*b*) passage through mixed beds of anion and cation exchange resins; (*c*) separation of the protein from the salt on a column of dextran gel (Sephadex), which excludes large protein molecules but retains salts.

After removal of salts the protein is recovered by lyophilization. The resulting fluffy powder is hygroscopic and difficult to handle. After extraction of the sample with absolute alcohol and ether the preparation is converted to a finely divided powder. At the same time the organic solvent serves to eliminate traces of contaminating lipids. The sample is allowed to reach constant moisture content by exposure in a room with constant humidity. For convenience, analytical determinations are performed with air-equilibrated samples and subsequently corrected to a dry-weight basis. In this form the preparation is readily handled, in contrast to anhydrous proteins, which may be exceedingly hygroscopic. The protein preparation is pulverized by grinding in an agate mortar, and the moisture content is obtained by drying in a vacuum at 105° to constant weight (Chibnall *et al.*, 1943; Brand and Kassell, 1942). At the same time the ash content is determined by combustion in a platinum crucible over an open flame until constant weight is attained.

B. Elementary Analysis

Standard microanalytical procedures are used for the determination of carbon and hydrogen. Nitrogen is most commonly determined by a micro-Kjeldahl or Dumas procedure; the latter is preferred since it is capable of oxidizing all forms of organic nitrogen to inorganic ammonium salts. Sulfur is usually estimated by the Pregl micromethod of oxidizing the sample; the sulfate produced is measured gravimetrically. Phosphate is measured colorimetrically after digesting the protein with hydrogen peroxide and sulfuric acid.

C. Other Considerations

Conjugated proteins containing metals, organic prosthetic groups, or constituents such as nucleic acid, carbohydrate, and lipid require special analytical methods for the determination of the nonprotein moiety. Analysis of the protein part of the molecule can be affected in one or

more ways by the nonprotein moiety. The presence of ultraviolet-absorbing materials may limit the usefulness of spectrophotometric analysis of aromatic amino acids. Enzymes with prosthetic groups require the separation of the cofactor from the protein by an appropriate procedure.

The presence of nonprotein components, especially carbohydrate and lipids, may increase the amount of destruction of certain amino acids during acid hydrolysis, and removal of these constituents is necessary. Lipids, as noted before, are reduced to a low level by extraction with organic solvents—acetone, alcohol, or ether.

III. Evolution of Amino Acid Analysis—From Macrochemistry to Microchemistry

An appreciation of the present status of amino acid analysis can best be gained from a brief survey of past progress. Furthermore, since future developments always rely on the information of the past, we would like to illustrate the high points of the historical development, emphasizing the salient features at each stage.

A. Gravimetric Procedures

The discovery of the amino acids and evidence for their role as constituents of proteins was pioneered in the early part of the nineteenth century and began as early as 1820 with the evidence for the natural occurrence of glycine and leucine (Vickery and Schmidt, 1931). Within a 75-year span, most of the remaining amino acids were discovered. During this period, the amino acids were isolated directly from protein hydrolyzates by fractional crystallization and after group separation. The amino acid content of a protein was simply based on the yield of the isolated material. The analytical procedures required up to 100 gm. of protein and a great amount of work and time, and suffered from the disadvantages of large losses and inadequate purification. For further details of this fascinating period of history the reader is referred to the superb review by Vickery and Schmidt (1931).

B. Colorimetric Procedures

In contrast to the macroscale required for gravimetric analysis the sensitivity of colorimetry has brought the scale of analysis to the microlevel. The techniques of colorimetric analysis allow rapid and multiple analyses to be made. The basis for the colorimetric determination of amino acids depends on the reaction of a chromophoric group of a suitable reagent with a functional side chain of the amino acid. Conse-

quently, the procedure can offer a high degree of specificity. The quantitative aspects depend on a calibration of the method against appropriate controls so that high precision of measurement is gained from replicate analyses. Use of colorimetric procedures is based on the premise that Beer's law applies.

The first quantitative application of colorimetry to protein analysis was made by Folin and Denis (1912) for the measurement of tyrosine and tryptophan with the use of their phenol reagent. Subsequently, improved procedures were developed (Folin and Ciocalteu, 1927; Folin and Marenzi, 1929). Specific reagents were developed later for other colorimetric analyses, e.g., arginine with the Sakaguchi color reaction (Sakaguchi, 1925a,b,c); histidine by coupling with diazotized sulfanilic acid (Koessler and Hanke, 1919; Hanke and Koessler, 1920), or Knoop's color reaction with bromine for histidine (Kapeller-Adler, 1933); and the color produced by the reaction of hydroxylamine after nitration of phenylalanine, the Kapeller-Adler technique (Kapeller-Adler, 1932).

C. Manometry

The degradation of amino acids by oxidizing agents forming products such as ammonia, carbon dioxide, and aldehydes led to the development of these reactions for quantitative purposes. Oxidation of amino acids by chloramine T leads to the stoichiometric production of carbon dioxide from most amino acids (Dakin, 1917). The oxidative decarboxylation of amino acids by ninhydrin at pH 2.5 releases both ammonia (MacFadyen, 1944; Sobel *et al.*, 1945) and carbon dioxide (Van Slyke and Neil, 1924; Van Slyke *et al.*, 1941; Schott *et al.*, 1944); each can be measured quantitatively. Nitrous acid treatment of amino acids converts the compounds to the corresponding alcohol with release of stoichiometric quantities of gaseous nitrogen, which can be determined manometrically (Van Slyke, 1911, 1912a,b, 1929). The procedure can be applied after separation of the amino acids, but its major use has been to characterize a protein hydrolyzate in terms of amino nitrogen values.

D. Enzymic and Microbiological Methods

Enzymes specific for certain amino acids have been useful for the direct determination of these amino acids in protein hydrolyzates without the necessity of prior isolation. Arginine was determined by conversion of the amino acid to ornithine and urea by the action of arginase (A. Hunter and Dauphinee, 1930; A. Hunter and Pettigrew, 1937). The enzyme is highly specific for arginine and can be employed directly on a protein hydrolyzate. The urea produced can be measured colori-

metrically (Archibald, 1945) or, after urease action, by estimating the carbon dioxide manometrically (A. Hunter and Pettigrew, 1937) or ammonia by various methods.

Gale (1945, 1946) isolated specific decarboxylases from bacterial sources for histidine, lysine, arginine, glutamic acid, tyrosine, and ornithine and applied them to the quantitative determination of this group of amino acids. Krebs (1950) measured aspartic acid after conversion to glutamic acid and oxaloacetic acid with transaminase. The glutamic acid formed was decarboxylated with its specific decarboxylase. The carbon dioxide formed in the decarboxylase reactions was measured manometrically. Although the number of amino acids capable of study by this method is limited, it is nevertheless a simple, specific, and precise procedure.

The determination of D-amino acids is of value, since several peptides of bacterial origin contain one or more of these amino acids (Bricas and Fromageot, 1953). A specific enzyme for the oxidation of D-amino acids is found in aqueous extracts of acetone-dried kidney powder (Krebs, 1935). The action of the enzyme is to split off the α-amino group in an oxidative deamination reaction. The enzyme is highly specific and attacks only the D-amino acids. The reaction can be followed by the uptake of oxygen in a manometric apparatus or by measuring the release of ammonia by a colorimetric procedure.

Microbiological methods for amino acid analysis (Snell, 1945, 1946) depend on knowledge of the nutritional requirements of bacteria and are successful with organisms having an absolute requirement for a particular amino acid to support normal growth. An assay consists of allowing bacteria to grow on a medium deficient in the amino acid under study and observing the rate of growth when graded amounts of the test substance are added. Lactic acid bacteria have been successfully used for the estimation of most amino acids in protein hydrolyzates. In common with enzymic procedures, the assays are stereospecific. As a general procedure, the method does have several pitfalls. Growth response must be due solely to the amino acid under study and not due to other components present that may affect growth. The specificity of response of the organism must not vary during the assay. These factors are usually controlled by performing the assay with several organisms requiring the amino acids under study. Microbiological assays have the advantage of permitting multiple analyses in a relatively short time by very simple techniques. As a method of analysis, the procedure is fairly sensitive, and all amino acids in a protein hydrolyzate can be determined on 10 mg. of protein.