COMPREHENSIVE BIOCHEMISTRY

EDITED BY

MARCEL FLORKIN

AND

ELMER H. STOTZ

VOLUME 7

PROTEINS (Part 1)

COMPREHENSIVE BIOCHEMISTRY

EDITED BY

MARCEL FLORKIN

Professor of Biochemistry, University of Liège (Belgium)

AND

ELMER H. STOTZ

Professor of Biochemistry, University of Rochester, School of Medicine and Dentistry, Rochester, N.Y. (U.S.A.)

VOLUME 7

PROTEINS (PART 1)

ELSEVIER PUBLISHING COMPANY

AMSTERDAM · LONDON · NEW YORK

1963

SOLE DISTRIBUTORS FOR THE UNITED STATES AND CANADA:
AMERICAN ELSEVIER PUBLISHING COMPANY, INC.
52, VANDERBILT AVENUE, NEW YORK 17 N.Y.

SOLE DISTRIBUTORS FOR GREAT BRITAIN
ELSEVIER PUBLISHING COMPANY LIMITED
12B, RIPPLESIDE COMMERCIAL ESTATE
RIPPLE ROAD, BARKING, ESSEX

Library of Congress Catalog Card Number 62–10359

With 26 illustrations and 36 tables

PRINTED IN THE NETHERLANDS BY
DRUKKERIJ MEIJER – WORMERVEER AND AMSTERDAM

COMPREHENSIVE BIOCHEMISTRY

COMPREHENSIVE BIOCHEMISTRY

COMPREHENSIVE BIOCHEMISTRY

SECTION I (VOLUMES 1–4)

PHYSICO-CHEMICAL AND ORGANIC ASPECTS OF BIOCHEMISTRY

SECTION II (VOLUMES 5–11)

CHEMISTRY OF BIOLOGICAL COMPOUNDS

SECTION III

BIOCHEMICAL REACTION MECHANISMS

SECTION IV

METABOLISM

SECTION V

CHEMICAL BIOLOGY

GENERAL INDEX

CONTRIBUTORS TO THIS VOLUME

ROBERT L. BALDWIN, B.A., PH.D.
Professor of Biochemistry, Stanford University, School of Medicine,
Palo Alto, Calif. (U.S.A.)

H. FRAENKEL-CONRAT, PH.D., M.D.
Professor of Virology, Department of Virology and Virus Laboratory,
University of California, Berkeley 4, Calif. (U.S.A.)

E. GORDON YOUNG, B.A., M.SC., PH.D., D.SC., F.R.S.C., F.C.I.C.
Director, Atlantic Regional Laboratory, National Research Council of Canada,
Halifax (Canada)

D. W. GREEN, B.SC. (CAMBRIDGE), PH.D. (LONDON)
Davy Faraday Research Laboratory, The Royal Institution, 21, Albemarle Street,
London, W.1 (Great Britain)

KAI O. PEDERSEN, PH.D., M.D. (H.C.)
Associate Professor of Physical Chemistry, Institute of Physical Chemistry,
University of Uppsala (Sweden)

HAROLD A. SCHERAGA, B.S., A.M., PH.D. (DUKE UNIV.),
SC.D. (HON.) (DUKE UNIV.)
Professor of Chemistry and Chairman of the Department, Cornell University,
Ithaca, N.Y. (U.S.A.)

RODES TRAUTMAN, B.E.E., PH.D.
Principal Research Physicist, Plum Island Animal Disease Laboratory, Animal Disease and Parasite Research Division, Agricultural Research Service, U.S. Department of Agriculture, P.O.Box 848, Greenport, Long Island, N.Y. (U.S.A.)

GENERAL PREFACE

The Editors are keenly aware that the literature of Biochemistry is already very large, in fact so widespread that it is increasingly difficult to assemble the most pertinent material in a given area. Beyond the ordinary textbook the subject matter of the rapidly expanding knowledge of biochemistry is spread among innumerable journals, monographs, and series of reviews. The Editors believe that there is a real place for an advanced treatise in biochemistry which assembles the principal areas of the subject in a single set of books.

It would be ideal if an individual or small group of biochemists could produce such an advanced treatise, and within the time to keep reasonably abreast of rapid advances, but this is at least difficult if not impossible. Instead, the Editors with the advice of the Advisory Board, have assembled what they consider the best possible sequence of chapters written by competent authors; they must take the responsibility for inevitable gaps of subject matter and duplication which may result from this procedure.

Most evident to the modern biochemist, apart from the body of knowledge of the chemistry and metabolism of biological substances, is the extent to which he must draw from recent concepts of physical and organic chemistry, and in turn project into the vast field of biology. Thus in the organization of Comprehensive Biochemistry, the middle three sections, Chemistry of Biological Compounds, Biochemical Reaction Mechanisms, and Metabolism may be considered classical biochemistry, while the first and last sections provide selected material on the origins and projections of the subject.

It is hoped that sub-division of sections into volumes will not only be convenient, but will find favour among students concerned with specialized areas, and will permit easier future revisions of the individual volumes. Toward the latter end particularly, the Editors will welcome all comments in their effort to produce a useful and efficient source of biochemical knowledge.

Liège/Rochester
March 1962

M. FLORKIN
E. H. STOTZ

PREFACE TO SECTION II

(Volumes 5–11)

Section II on the Chemistry of Biological Compounds deals with the organic and physical chemistry of the major organic constituents of living material. A general understanding of organic and physical chemistry is presumed, but the reader will find the special topics in Section I of value in the fuller understanding of several parts of Section II. The Editors have made special effort to include a sound treatment of the important biological high polymers, including sections on their shape and physical properties. A number of substances peculiar to plants, certain isoprenoids, flavonoids, tannins, lignins, and plant hormones, often omitted from text-books of biochemistry, are included. Nevertheless, it is inevitable that some omissions, hopefully minor ones, have occurred. The only intentional omission is the chemistry of the coenzymes and certain components of biological oxidation, which will be covered in connection with their function in Section III.

The previous policy of dividing the section into smaller volumes has been continued, resulting in seven volumes for Section II. Two of the volumes each contain a complete area, namely Carbohydrates (Volume 5) and Sterols, Bile Acids and Steroids (Volume 10). Comments from readers will be appreciated by the Editors and be most helpful for possible future revisions.

Liège/Rochester
December 1962

M. Florkin
E. H. Stotz

CONTENTS

VOLUME 7

PROTEINS – PART I

Chapter I. Occurrence, Classification, Preparation and Analysis of Proteins

by E. Gordon Young

Chapter II. Chemistry of Proteins

by H. FRAENKEL-CONRAT

Chapter III. Acid–Base Properties and Electrophoresis of Proteins

by R. TRAUTMAN

Chapter IV. Reactions of Proteins; Denaturation

by H. A. SCHERAGA

Chapter V. Thermodynamic Properties of Proteins Found From Osmotic Experiments

by R. L. BALDWIN

Chapter VI. Sedimentation, Diffusion and Partial Specific Volume

by K. O. PEDERSEN

Chapter VII. Spacial Configuration in Proteins

by D. W. GREEN

Chapter I

Occurrence, Classification, Preparation and Analysis of Proteins

E. GORDON YOUNG

Atlantic Regional Laboratory, National Research Council, Halifax (Canada)

1. Introduction

Of all known chemical compounds proteins are the most complex and, at the same time, the most characteristic of living matter. All viable cells contain them. They are the compounds which, as nucleoproteins, are essential to the process of cell division, and, as enzymes and hormones, control many chemical reactions in the metabolism of both plant and animal cells. As the major, if not the only, constituent of viruses, nucleoproteins are synonymous with the most elementary form of living matter.

The word protein first appeared in scientific literature in 1838 in an article by the Dutch chemist, G. J. Mulder. It is derived from the Greek adjective πρωτειοζ (*prote*–first and *eidos*–like) which may be translated as of the first rank or position. While Mulder was investigating an organic substance which contained nitrogen, sulfur and phosphorus and appeared to be widely distributed in biological material, he received a letter from the great Swedish chemist Berzelius who suggested the term to him[1,2]. Thus as early as 1838 the fundamental importance of protein in living matter was recognized. Liebig, a few years later, investigated the composition of several plant and animal proteins. Later, the German term *Eiweisskörper* and the English modification *proteid* came into general use. The latter has now disappeared.

During the second half of the nineteenth century and the early years of the twentieth investigators were concerned with the separation of proteins from their complex environment and the determination of their constituent amino acids. The names of Osborne, Abderhalden, Kossel, Ritthausen, Hammarsten, Schmiedeberg, Emil Fischer, Levene, Sörensen, and Van Slyke, may be recalled as the founders of the modern chemistry of the proteins. In those days the determination of the amino acids in a protein required esterification, fractional distillation, and frequently isolation of each amino acid —a task which required many weeks, in contrast to the present mechanized procedure which requires less than a day.

The first protein to be crystallized was haemoglobin, in 1840, by evaporation of the blood of the earthworm[3]; probably the second was the globulin[4] from the Brazil nut in 1877, and the third ovalbumin[5] in 1889. The first enzyme to be crystallized was urease, isolated by Sumner[6] in 1926. Abel[7] crystallized the first protein hormone, insulin, in 1926, and Stanley[8] the first virus, tobacco mosaic virus, in 1935. Crystalline proteins now are numbered by the score.

In the earlier part of this century investigators turned their attention to the physico-chemical characteristics of a few proteins. These studies have formed the basis for our present conceptions of colloidal behavior, hydrodynamic properties, size and shape of protein particles. The invention of the ultracentrifuge in 1923 by Svedberg and the extensive modification of the electrophoretic cell by Tiselius in 1930 were significant milestones in the progress of protein chemistry. The advent of modern chromatography has made possible a new approach to the structure of proteins in the classical study of the sequence of amino acids in insulin by Sanger[9] in 1955. The initial attempts of Emil Fischer to synthesize polypeptides have now led to a synthesis of the pituitary hormones, oxytocin and vasopressin, by du Vigneaud[10] and his colleagues. So rapid has been the advance and so voluminous the amount of work that there is a yearly volume *Advances in Protein Chemistry* dating back to 1944, a comprehensive treatise on proteins in four volumes[11], the classical monograph of Cohn and Edsall[12], and several smaller text books[13].

2. Occurrence

Proteins are believed to occur in all living matter – viruses, bacteria, plants and animals. In addition they are found, sometimes in high concentration, in such animal products as urine, blood, lymph, milk, eggs, in epidermal structures such as hair, feathers, and hoofs, in internal secretions such as the digestive juices, cerebrospinal, ocular, and synovial fluids, and in external secretions, such as uterine, seminal, prostatic, and testicular. Rather special cases are the secretion of the silkworm and the cobweb of the spider. In plants they occur in solution in the circulating fluids such as the cell sap. In the undissolved state, sometimes as crystals, they are found in seeds, spores, tubers, bulbs, and roots. The concentration in seeds is very high, especially in the endosperm and this is of particular importance to man in nuts, in grains and in some legumes, such as peas and beans. A very poisonous protein, ricin, occurs in the castor bean. The greater part of the protein in leaves or fronds is in a rather insoluble form.

Within the structure of the cell, proteins constitute part of the cell wall or membrane and all of cilia or flagella. Both in solution and in granular or crystalline form they are present in the cytoplasm and in vacuoles. They probably constitute the major portion of mitochondria and plastids. It is especially in the nucleus and nucleolus that they attain their greatest biological importance as the substance of chromosomes. In the submicroscopic structures known as phages or viruses they constitute all or the greater part of the particle[14,15].

3. Classification

Classifications, as efforts to systematize knowledge in science, are useful. In biology this approach is axiomatic and from such efforts came the theory of evolution. In organic chemistry classification is indispensable. It is therefore not surprising that in the early days of biochemistry proteins were classified, and, indeed, national committees were appointed to standardize the nomenclature in 1907. At that time little was known about proteins and the most distinctive properties were differences in solubility, coagulability, and the presence or absence of a "prosthetic" group as distinct from that portion made up of amino acids. The differences between the two classifications proposed in 1907 were minor and have no significance today. These classifications are now used to a limited degree because of the many exceptions which are known. They are reproduced below. Proteins are so classified because, basically, they are polypeptides made up of α-amino acids. But when does a polypeptide become a protein with increasing size of particle? There is no authoritative answer to this question but general practice places it at a molecular weight of about 5000, which includes the simplest native group, the protamines, but excludes, for example, oxytocin and vasopressin, the hormones of the posterior pituitary gland, as polypeptides.

A subdivision of proteins[16] into simple and conjugated was proposed in 1907 on the basis of the following definitions, (1) simple proteins are those which yield only α-amino acids or their derivatives on hydrolysis; (2) conjugated proteins are those which contain a protein molecule united to some other molecule or molecules otherwise than as a salt. The definition of a conjugated protein should now be modified to read that the other or "prosthetic" molecule be organic in character and possibly in combination with the protein moiety as a salt. The French use the terms holoproteins and heteroproteins to distinguish these main subdivisions.

The distinctive physical and chemical characteristics of these groups will now be described.

Simple proteins	*Conjugated proteins*
Protamines	Phosphoproteins
Histones	Mucoproteins (glycoproteins)
Albumins	Chromoproteins (metalloproteins)
Globulins	Nucleoproteins
Prolamins (alcohol-soluble)	Lipoproteins
Glutelins	
Scleroproteins (albuminoids)	

(*a*) *Simple proteins*

(*i*) *Protamines*

These are the simplest native proteins. They occur in mature spermatozoa and have been isolated mainly from fish sperm. The first to be recognized was discovered by Miescher in 1874 in the spermatozoa of the salmon and named by him *protamin*[17]. Kossel[18] and his collaborators carried out many investigations of this group. The protamines are distinctive for their low molecular weights of about 5000, a limited variety of amino acids with a preponderance of the basic ones, especially arginine. These properties account for their relatively high value of nitrogen (18–25 %), their basic isoelectric points (pH 10–12) and the absence of any sulfur. They are moderately soluble in water and ammonia and basic in character but difficult to prepare as the free base. They form insoluble sulfates and chloroplatinates which are frequently crystalline. They are not coagulable by heat. Because of their basicity they form protein–protein compounds and combine with nucleic acids to form nucleoproteins[19]. This property has found practical application in the form of protamine insulinate for which salmine, scombrine, clupeine, and cyclopterine have been used as they form insoluble compounds with insulin. They are hydrolyzed by trypsin, trypsin-kinase, cathepsin, and papain but not by pepsin. They dissociate in concentrated aqueous salt solutions.

Salmine and clupeine have been most thoroughly studied but the most recent observations indicate heterogeneity in the preparations[20]. The distribution of amino acids as percentages of the total protein is as follows.

	Ala	*Arg*	*Gly*	*Ileu*	*Pro*	*Ser*	*Thr*	*Val*
Salmine	1.5	88.4	3.3	1.2	7.9	7.0	0.0	4.1
Clupeine	4.7	87.1	0 0	1.0	8.2	3.4	1.9	3.6