# ADVANCES IN PROTEIN CHEMISTRY

EDITED BY

M. L. ANSON
*Continental Foods, Hoboken*

JOHN T. EDSALL
*Harvard Medical School, Boston*

VOLUME I

1944
ACADEMIC PRESS, INC.
NEW YORK

---

Second Printing 1946
Third Printing 1948

Printed in the United States of America
The Murray Printing Company
Wakefield, Massachusetts

CONTRIBUTORS TO VOLUME I

KENNETH BAILEY, *University of Cambridge, England*

ERWIN CHARGAFF, *Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York, N. Y.*

BACON F. CHOW, *The Squibb Institute for Medical Research, New Brunswick, N. J.*

DAVID M. GREENBERG, *Division of Biochemistry, University of California Medical School, Berkeley, Cal.*

JESSE P. GREENSTEIN, *National Cancer Institute, National Institute of Health, Bethesda, Maryland*

DONALD S. PAYNE, *Office of Distribution, War Food Administration, Washington, D. C.*

FRANCIS O. SCHMITT, *Department of Biology and Biological Engineering, Massachusetts Institute of Technology, Cambridge, Mass.*

L. S. STUART, *Office of Distribution, War Food Administration, Washington, D. C.*

HENRY P. TREFFERS, *Department of Comparative Pathology and Tropical Medicine, Harvard Medical School, Boston, Mass.*

# Editors' Preface

In the last generation, protein chemistry, which was once a relatively narrow branch of organic and biological chemistry, has spread out into the most varied fields of physics, chemistry, and biology. Enzymes, viruses, and many substances of immunological importance, are now known to be proteins. The techniques now used for the study of proteins range from the most elaborate form of X-ray analysis to quantitive measurements of antibodies. Workers in the most diverse fields of science have not only contributed to the development of techniques, but have become interested themselves in applying the techniques they helped develop in the study of the problems of protein chemistry. With the great progress in the knowledge of proteins, the industrial and medical applications of this knowledge have increased greatly and promise to increase far more.

The rapid pace of the advances in protein chemistry, the varied character of the work being done, and of its practical applications to industry and medicine, have given rise to an increasing need for thoughtful and critical evaluation of the results achieved, and of their implications. We hope that this series of volumes will give the opportunity to workers in special subjects to present their views in more organized form than is possible in the regular journals, and also to express their peronal judgment on problems which are still unsettled. We hope, too, that as the reviews accumulate, they will provide a useful and comprehensive picture of the changing and growing field of protein chemistry, and a stimulus to its further development.

In this first volume, special emphasis is laid on proteins as they occur in nature, as components of complex biological systems. In the second volume, which we expect to appear in 1945, there will be a group of contributions which reflect the increased interest in protein nutrition stimulated by the war. These will include discussions of the estimation of amino acids by chemical and bacterial growth methods, of the amino acid contents of protein foods, of protein nutrition in man, and of the relation of protein nutrition to antibody formation.

Since the physical chemistry of protein systems was extensively treated in a symposium which appeared in "Chemical Reviews" in 1942, and has also been discussed in several recent monographs, it has received relatively little emphasis in the present volume. Extensive and critical discussions of recent advances in the physical chemistry of amino acids, peptides, and proteins, however, will appear in later volumes.

The circumstances of war have inevitably imposed severe restrictions on possible contributors at the present time. In many countries, those who might be interested in contributing to such an enterprise are totally inaccessible to us; in the United States and England, others who expressed keen interest in the project have been unable to contribute because of the pressure of more urgent work. The authors whose work is presented here have all prepared their contributions under difficult conditions. We, and we believe our readers also, are greatly indebted to them for their presentation of significant developments in fields with which they are intimately familiar.

M. L. Anson
John T. Edsall

## CONTENTS

**Some Contributions of Immunology to the Study of Proteins**

By Henry P. Treffers, *Department of Comparative Pathology and Tropical Medicine, Harvard Medical School, Boston, Mass.*

**The Interaction between the Alkali Earth Cations, Particularly Calcium, and Proteins**

By David M. Greenberg, *Division of Biochemistry, University of California Medical School, Berkeley, Cal.*



**The Purification and Properties of Certain Protein Hormones**

By Bacon F. Chow, *Division of Pharmacology, The Squibb Institute for Medical Research, New Brunswick, N. J.*

**Soybean Protein in Human Nutrition**

By Donald S. Payne and L. S. Stuart, *Office of Distribution, War Food Administration, Washington, D. C.*



**Nucleoproteins**

By Jesse P. Greenstein, *National Cancer Institute, National Institute of Health, U. S. Public Health Service, Bethesda, Maryland*

## The Proteins of Skeletal Muscle

By Kenneth Bailey, *University of Cambridge, England*

# Lipoproteins

By ERWIN CHARGAFF

*Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York, N.Y.*

## Contents



## I. Introductory Remarks

This review will deal largely with compounds which are traditionally regarded as impure. It therefore may be appropriate briefly to examine the conception of purity as it applies to many of the complicated substances that are isolated from biological material (compare the review by Pirie, 1940). Although not often clearly formulated, the disparity in the con-

ceptions of chemical purity and biological purity has beset the biochemist from the very beginnings of his science.

Workers are generally agreed on the criteria for the chemical purity of proteins, carbohydrates, lipids, of biological origin, *e.g.*, the constancy of chemical composition, physical characteristics, etc. To these requisites one more usually is added implicitly, *viz.*, that the types of linkages supporting the architecture of the particular substance be clearly definable in terms of our present knowledge. This requirement is obviously not fulfilled in a number of compounds of great biological importance, *e.g.*, certain enzymes, viruses, etc. Yet, many of these substances show a functional homogeneity of the highest degree. If the cytoplasm or certain cytoplasmic inclusions of a cell were to be isolated in their unaltered states, they would certainly be found extremely impure chemically, even though they were to exhibit a high order of biological purity. While it may be hoped that with a more profound understanding of the types of linkages prevailing in these complex structures the borderline between chemical and biological purity will disappear, it is clear that the fear of offending the chemical proprieties has frequently led to the most far-fetched model experiments, in which the true objects of biochemistry (which, after all, is not thanatochemistry) have become obscured. The study of biologically important substances in their natural environment will certainly form a most fascinating subject of biochemical research.

This article cannot attempt to offer a complete treatment of the chemistry of lipid-protein complexes which probably are ubiquitous components of living matter; it will rather be its purpose to point to a number of instances where the occurrence of at least fairly well defined lipoproteins has been shown to be probable. Representative compounds present inside the cell and in the extracellular fluids will be discussed.

Some of the historical aspects of our knowledge of the manner in which the lipids occur in living matter have been discussed by Sandor (1934) and by Macheboeuf (1937). A very adequate review of the present state of chemical information was recently published by Lovern (1942). There exists a vast amount of histological observations on the occurrence of lipid-protein complexes in tissues, and in the innumerable speculations on the nature of protoplasm the function of lipoproteins has often been debated. Some of the monographs and articles mentioned in the bibliography may be found useful in this connection (Biedermann, 1924; Bourne, 1942; Degkwitz, 1933; Guilliermond, 1934; Kiesel, 1930; Lepeschkin, 1938; Lison, 1936; Sponsler and Bath, 1942.)

## II. Definition of Term and Classification

The term lipoprotein is used for compounds between proteins and lipids; the latter are in turn defined as a group of naturally occurring fatty acid

derivatives that are soluble in organic solvents. This group comprises the fatty acids, glycerides, glycerol ethers, waxes, phosphatides, phosphatidic acids, cerebrosides, and esters of xanthophylls and sterols. Sometimes the uncombined steroids, carotenoids and related substances are likewise included. A desirable classification of the lipoproteins could be based on the nature of the lipid residue contained in the conjugated protein: one would speak of lecithoproteins, cephaloproteins, etc. Actually, except in the case of a few synthetic lipoproteins which will be discussed later, this cannot yet be done, since the conjugated proteins isolated from natural sources until now were all found to contain more than one lipid species. If the nature of the protein moiety is known, the prefix lipo- could be used in these cases; for instance, lipovitellin for the lipoprotein from egg yolk (Chargaff, 1942, a).

Not infrequently, terms, as *e.g.*, lecithovitellin, are employed for protein preparations which had been freed of phosphatides. This is confusing; it would certainly appear advantageous to reserve such prefixes for the conjugated proteins.

## III. Modes of Linkage between Lipid and Protein

A schematic survey of the types of primary valence bonds that could exist between lipids and proteins may be helpful for an understanding of the lipoprotein problem.

1. Fatty acids: Carboxyl group, electrostatic (salt); covalent (ester, amide, etc.).
2. Triglycerides: None. (Mono- and diglycerides could, of course, form covalent ester links through their free hydroxyl groups.)
3. Lecithin: Trimethyl ammonium group, electrostatic (salt). Phosphoric acid group, electrostatic (salt); covalent (ester, etc.).
4. Cephalin: Amino group, electrostatic (salt); covalent (amide). Phosphoric acid group, electrostatic (salt); covalent (ester, etc.).[1]
5. Phosphatidyl serine: Amino group, electrostatic (salt); covalent (amide). Phosphoric acid and carboxyl groups, electrostatic (salt); covalent (ester, amide, etc.).
6. Sphingomyelin: Trimethyl ammonium group, electrostatic (salt). Phosphoric acid group, electrostatic (salt); covalent (ester). Hydroxyl group, covalent (ester).
7. Phrenosin and kerasin: Hydroxyl groups, covalent (ester).
8. Xanthophyll and sterol esters: None.

It is clearly necessary to distinguish between genuine lipid-protein complexes and mixtures or loose adsorption systems which can be separated into their component parts by mild methods of fractionation or extraction.

[1] The phosphatidic acids, glycerol ethers, and acetal phosphatides probably behave similarly to groups 1, 2, and 4 respectively.

The term lipoprotein, therefore, connotes, or should connote, a group of compounds with properties (biological reactivity, solubility, color, optical and other physical constants) different from those of the sum of their components. It is, at the present state of information, not possible to offer a strict formulation of the type of linkages prevailing between a lipid as a prosthetic group and a protein. The introduction of the term "symplex" (Willstätter and Rohdewald, 1934) has, as a purely terminological innovation, by no means advanced our understanding. The inspection of the polar characteristics of the various lipids enumerated above could, however, lead to the following provisional classification (compare also Przylecki, 1939; Lovern, 1942).

I. *Covalent Compounds* (Groups 1 and 3 to 7).
II. *Electrostatic Compounds* (Groups 1 and 3 to 6).
III. *Secondary Valence Compounds* (Groups 1 to 8).

There is little evidence of the occurrence of covalent lipoproteins (*e.g.*, esters, amides, etc.) in nature. It is, of course, true that only very rarely was a search made for the presence of fatty acids in hydrolyzates of thoroughly defatted proteins. There are, however, some indications of the occurrence of compounds of this type in certain species of pathogenic bacteria, *e.g.*, in human tubercle bacilli (Anderson, Reeves, and Stodola, 1937) and in diphtheria bacteria (Chargaff, 1931). It often will be difficult to distinguish experimentally between substances belonging to this class and those of salt-like nature.

The electrostatic compounds would be represented by salts in which the attraction is due to ionic forces between the lipid and the protein. The synthetic lipoproteins and lipoprotamines to be discussed later certainly fall into this group. Substances of this type will be stable within a certain pH range only. It is therefore to be assumed that there will be a significant difference in the reactivities of lecithin and sphingomyelin and of the more acidic phosphatides (cephalin, phosphatidyl serine, phosphatidic acids, etc.).

Lecithin and sphingomyelin, which contain the strong base choline, may be considered as internally neutralized compounds. The isoelectric point of lecithin has been found at pH 6.7, not much lower than the value required by the theory, *viz.*, pH 7.5 (Chain and Kemp, 1934; Bull and Frampton, 1936). The admixture of cephalin appears to bring about a considerable lowering of the isoelectric point of lecithin (Bull and Frampton, 1936; Cohen and Chargaff, 1940). Unfortunately, the information with respect to cephalin is less satisfactory. Most physical measurements on cephalin were carried out with preparations from brain which, as we know now (Folch and Schneider, 1941), contained a large proportion of phosphatidyl serine. This compound is doubtless markedly acidic and it is not unlikely

that most properties assigned to cephalin in the literature, especially those distinguishing it from lecithin, actually relate to the serine-containing compound. This will be particularly true of preparations obtained from brain and lungs, whereas egg yolk phosphatides are practically free of this serine derivative (Chargaff, Ziff, and Rittenberg, 1942). In its polar characteristics, cephalin itself (*i.e.*, ethanolamine phosphoryl diglyceride) is probably much nearer to lecithin than was previously supposed. It may be concluded that at the physiological pH lecithin can hardly be expected to form salts with tissue proteins, whereas cephalin and especially phosphatidyl serine will be able to combine with protamines and certain basic proteins (*e.g.*, histone) by means of ionic bonds to form insoluble products; in other cases (*e.g.*, globin and cephalin at pH 7), the resulting salt may show a higher solubility than the uncombined protein (Chargaff, 1938; Chargaff and Ziff, 1939).

Most lipoproteins occurring in nature probably have to be classified as secondary valence complexes which are held together by van der Waals forces. This assumption certainly is necessary for those lipids (Groups 2 and 8 in the scheme given above) that lack centers of attachment which could give rise to the establishment of covalent or electrostatic bonds. The situation may be further complicated by the possible existence of solid solutions of lipids in the prosthetic lipid portion of a lipoprotein, *e.g.*, glycerides or steroids dissolved in the cephalin part of a cephaloprotein.

The distinction between electrostatic and coordination compounds is probably not sharp in the field of conjugated proteins, as the coordination centers are presumably often represented by electrically charged groupings in the protein and the prosthetic substance. These complexes are stable, and move intact in an electric field, at a pH well above the isoelectric points of their component parts (compare Cohen and Chargaff, 1940; 1941 b; Chargaff, Ziff, and Moore, 1941); on treatment with certain organic solvents, as for instance ethyl alcohol, which would not be expected to disrupt electrostatic compounds, the linkage between lipid and protein tends to break. The action of heparin (Chargaff, Ziff, and Cohen, 1940, b) in displacing the lipids from certain lipid-protein compounds (see Section V) could equally well be explained as the formation of protein salts of a solubility lower than that of the original lipid-protein compounds. Generally, it may be assumed that any action which results in the displacement or the distortion of the centers of attachment in the protein, thereby changing the critical spacing of these centers, will bring about the cleavage of the conjugated protein.[2]

Very little can be said about the nature of the centers of attraction and

[2] The theory of the antigen-antibody reaction has dealt exhaustively with problems of this nature, and reference should be made to recent treatments of this subject (Heidelberger, 1939; Landsteiner, 1936; Marrack, 1938; Pauling, 1940a).

of the forces supporting such secondary valence structures. In some cases the formation of hydrogen bonds may play a rôle (Pauling, 1940, b). The action of alcohol in destroying the links between protein and lipid appears to be irreversible.[3] Because of the lack of reliable amino acid analyses of the protein moieties of naturally occurring lipoproteins, it is not even possible to establish a correlation between the number of basic amino acid residues in these proteins and their tendency to combine with lipids. Such a correlation seems to exist, however, in the salts between basic proteins and lipids which will be discussed in the next section.

Weiss (1941, 1942) has proposed an interesting theory of the structure of organic molecular compounds between substances possessing electronegative groups and unsaturated hydrocarbons (or their derivatives) containing conjugated double bonds. The formation of these essentially ionic complexes is ascribed to an electron transfer from the unsaturated hydrocarbon to the electronegatively charged molecule. Compounds of this type will possibly be encountered among the naturally occurring complexes.

## IV. Chemical Properties of Lipoproteins

### 1. *Synthetic Compounds*

There have been numerous attempts to prepare artificial lipoproteins. Since in most cases very little attention was paid to the polar characteristics of the lipids or the proteins employed, it is not surprising that many of these experiments were not successful. The effort of combining anything with everything will not lead to the creation of a homunculus, even if the resulting mixtures are termed symplexes or coacervates.

In a number of studies the preparation of complexes between serum albumin or egg albumin and lecithin was attempted (Galeotti and Giampalmo, 1908; Liebermann, 1893; Mayer and Terroine, 1907; von Przylecki and Hofer, 1936; Went and von Kúthy, 1934). The competition between lecithin and cholesterol for the serum proteins was investigated by Theorell (1930). The interactions between lecithin and serum globulin (Chick, 1914), zein (Galeotti and Giampalmo, 1908), and caseinogen (Parsons, 1928) were likewise studied. Other workers followed the influence of proteins on the flocculation of lecithin sols (Feinschmidt, 1912; Handovsky and Wagner, 1911; Went and Faragó, 1931). The almost exclusive use of

[3] Macheboeuf and Sandor (1932) have attempted to explain the structure of lipoproteins by the mutual attraction of the hydrophobic groupings of the protein and the lipids. These complexes are assumed to be surrounded by a nimbus of water, held in position by their hydrophilic groups, which prevents the access of ether and other solvents immiscible with water, but not of alcohol. It is doubtful whether this explanation could be applied to the dehydrated complexes.

lecithin in these experiments probably explains the rather inconclusive results since lecithin, in contrast to the more acidic phospholipids, is an inappropriate model substance.

The difference in the behavior of lecithin and the acidic phosphatides of the cephalin group was emphasized in studies on the formation of compounds between these lipids and protamines or basic proteins. With the highly basic protamine salmine (isoelectric point at pH 12) cephalin preparations from brain, which presumably were rich in phosphatidyl serine, formed water-insoluble salts over a wide pH range, *viz.*, from pH 2 to 11 (Chargaff, 1938). These products, which had a P:N ratio of 1:4 or 1:5 and were composed of about 80% of cephalin and 20% of salmine, were soluble in organic solvents, could be recovered unaltered when their solutions in hot ethyl acetate were cooled, and did not change their composition following treatment of their solutions in ether with dilute acids or repre-

TABLE I

*Composition of Cephalin-Histone Compounds*

(From Chargaff and Ziff, 1939)

| pH of reaction | Composition of compounds | | | Cephalin-binding capacity |
|---|---|---|---|---|
| | P | N | Cephalin in compound on basis P value | |
| | *per cent* | *per cent* | *per cent* | *m.eq. per g. protein* |
| 3.7 | 2.5 | 6.7 | 71 | 2.78 |
| 4.7 | 2.1 | 7.9 | 60 | 1.71 |
| 7.2 | 1.6 | 9.4 | 50 | 1.03 |

cipitation with acetone. In one such compound 4.7 milliequivalents of cephalin per g. of salmine were found instead of 5.0, as calculated from the acid-binding capacity of this protamine. With lecithin, on the other hand, no compound formation was observed, except at a very high pH, *viz.*, 10 and 11. The same was true of sphingomyelin.

A study of the compound formation between lecithin and cephalin (rich in phosphatidyl serine) and basic proteins, *viz.*, histone from calf thymus and globin from cattle hemoglobin, showed that histone formed insoluble compounds with cephalin between pH 2 and 7, whereas with globin appreciable formation of insoluble cephalin compounds was observed only below pH 4; lecithin formed no compounds with globin, but did so with histone between pH 7 and 8 (Chargaff and Ziff, 1939). The cephalin-histone and cephalin-globin compounds, whose composition is summarized in Tables I and II were, in contrast to the corresponding protamine salts, insoluble in organic solvents.