PROTEIN SYNTHESIS

Volume 1

edited by Edwin H. McConkey

Protein Synthesis

a series of advances. **Volume 1**

Edited by

EDWIN H. McCONKEY

DEPARTMENT OF MOLECULAR, CELLULAR AND
DEVELOPMENTAL BIOLOGY
UNIVERSITY OF COLORADO
COLLEGE OF ARTS & SCIENCES
BOULDER, COLORADO

MARCEL DEKKER, INC., New York 1971

MARCEL DEKKER, INC., 95 Madison Avenue, New York, New York 10016

Library of Congress Catalog Card Number: 74-155743

ISBN No.: 0-8247-1459-8

PRINTED IN THE UNITED STATES OF AMERICA

Introduction to the Series

These volumes will present reviews on the structure, functions, biosynthesis, and regulation of the protein-synthesizing apparatus in prokaryotic and eukaryotic cells. Articles on various RNA's will be included regularly because all known functions of RNA are related directly or indirectly to protein synthesis or the regulation of protein synthesis. Although it could be argued that *all* of molecular biology and biochemistry is indirectly related to protein synthesis, this series will maintain a relatively restricted focus. The molecular biology of DNA or the mechanisms of enzyme action, for example, will not be considered as general subjects, although specific genes, such as those that code for ribosomal components, and specific enzymes, such as peptidyl transferase and amino acid: tRNA ligases will obviously be treated in detail. Reviews will be included from time to time on some subjects peripherally related to the biosynthesis of proteins, such as post-translational modification, non-ribosomal peptide biosynthesis, and chemical synthesis of polypeptides.

A unified series of reviews covering all aspects of protein synthesis is much needed. Although the principal reactions of protein synthesis have been identified, the subject is far from exhausted. The number of researchers active in the field is very large, and the rate of publication appears to increase continually. Consider, for example, the *Proceedings of the National Academy of Sciences* and the *Journal of Molecular Biology*, two major sources of literature on protein synthesis. Together these journals published approximately 10,000 pages in 1970; by my interpretation the *Proceedings* contained nearly 1000 pages and the *Journal of Molecular Biology* contained nearly 2000 pages of articles directly pertinent to protein synthesis. This is only a minor part of the total current literature on the subject, of course.

There are few persons fortunate enough to be able to keep abreast of this deluge of new information through personal contacts. Most of those who are either active or interested in the field must depend almost entirely on published articles and many find that their attempts to maintain familiarity with current research are becoming less and less adequate. There is no shortage of reviews; indeed, one can scarcely do a thorough job of reading reviews, even if one almost abandons

original reports. This desperate state of affairs is illustrated by the 1971 Prospectus for *Annual Reviews of Biochemistry*, where *two* separate reviews entitled "Protein Biosynthesis" are announced. How long will it be before it becomes necessary to establish the *Annual Review of All Other Reviews*?

Unfortunately, the intentions and the achievements of reviewers vary as widely as the quality of the research that they review. Excellence appears as rarely in scientific work as in any other profession, and a reviewer who uncritically catalogs the contents of abstracts does his colleagues no service. *Protein Synthesis: A Series of Advances* will provide critical evaluations, rather than mere summaries of recent data. Attempts to reconcile conflicting observations and to clarify confusing terminology will be featured. In order to maximize the usefulness of these reviews as reference sources, full literature citations will be given, alphabetically arranged.

Professional scientists and graduate students should find these volumes helpful in following the development of this complex subject. When possible, several closely related articles will be published in the same volume. The basic mechanisms of protein synthesis, which were exhaustively documented in the 1969 *Cold Spring Harbor Symposium*, are also the subject of all the articles of Volume 1 of *Protein Synthesis*. This subject will undoubtedly generate many future reviews. The intricate molecular anatomy of ribosomes is being investigated vigorously, and the details of intermolecular associations between ribosomal proteins and supernatant factors or between tRNA's and the ligases have barely begun to be elucidated. Soon the sequencers will become dominant, and in due time, we shall almost certainly know every amino acid and every nucleotide in the roughly 150 macromolecules that constitute the protein-synthesizing machinery of *Escherichia coli*.

All manifestations of protein synthesis are not found in bacteria, although the basic mechanisms are remarkably similar throughout the living world. Eukaryotes, however, may possess unique control mechanisms, such as initiation factors specific for certain types of mRNA or repressors that function at the polysome level. The process by which mRNA's reach the cytoplasm is still mysterious, and the eukaryote nucleus has become a jungle of orphan RNA's of all imaginable sizes and properties, on which speculation grows luxuriantly. Interest in these problems has increased tremendously in the past year or two, partly because some of the excitement of exploring bacterial systems

has passed, and partly because the United States government has begun to define health-related research more narrowly.

Volume 2 will include several articles on protein synthesis in eukaryotes. The contents and frequency of appearance of subsequent volumes will depend upon the progress of the field, upon the availability of suitable authors, and of course, upon the reception given Volumes 1 and 2. Suggestions for topics, contributors, and improvement in style would be welcome from all.

Contributors to Volume 1

DAVID APIRION, Department of Microbiology, Washington University School of Medicine, St. Louis, Missouri

MARK S. BRETSCHER, Medical Research Council, Laboratory of Molecular Biology, Cambridge, England

ROGER GARRETT, Laboratory for Physiological Chemistry, State University of Leiden, Leiden, The Netherlands

C. G. KURLAND, Department of Zoology and Laboratory of Molecular Biology, University of Wisconsin, Madison, Wisconsin

ROBERT B. LOFTFIELD, University of New Mexico, School of Medicine, Albuquerque, New Mexico

KIVIE MOLDAVE, Department of Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania

WIM MÖLLER, Laboratory for Physiological Chemistry, State University of Leiden, Leiden, The Netherlands

DAVID SCHLESSINGER, Department of Microbiology, Washington University School of Medicine, St. Louis, Missouri

JACK G. SILER, Department of Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania

Contents of Volume 1

1

The Aminoacylation of Transfer Ribonucleic Acid

ROBERT B. LOFTFIELD

UNIVERSITY OF NEW MEXICO
SCHOOL OF MEDICINE
ALBUQUERQUE, NEW MEXICO

I. ABBREVIATIONS AND APOLOGY

The enzymes discussed in this chapter were originally called amino acid activating enzymes. The discovery of their dual role has led to general use of the term, aminoacyl-tRNA synthetase. The Commission on Enzymes recommends the somewhat more precisely descriptive name, amino acid:tRNA ligase (AMP), which we shall use in abbreviated form. tRNA is transfer RNA, $tRNA^{Ala}_{E.coli}$ is *Escherichia coli* tRNA specific for alanine, and alanyl-$tRNA^{Ala}_{E.coli}$ is the same RNA esterified with alanine. A, C, G, U, Pu (or R), and Py (or Y) represent adenosine, cytidine, guanosine, uridine, purine riboside, and pyrimidine riboside. AA, AA~AMP, and E·(AA~AMP) are amino acid, aminoacyl adenylate, and enzyme-bound aminoacyl adenylate [IUPAC-IUB symbols; *J. Biol. Chem* **24**: 527 (1966)]. ATP:PP_i exchange is used to denote incorporation of ^{32}P-PP_i into ATP by reversal of reaction 3A (see p. 12).

In view of the thousands of articles on this subject published during two decades, I would have to be 'umble as Uriah Heep to apologize for an incomplete presentation. This chapter refers to those areas and articles which I personally feel are experimentally sound, conceptually original, significant in the past, or likely to open new avenues in the future. Even when I have tried to be comprehensive, as in Table 1,

I am sure that I shall have failed. In order that the reader may find more comprehensive treatment of various aspects of the first steps in protein biosynthesis, I have included a number of reviews with brief comments just before the bibliography.

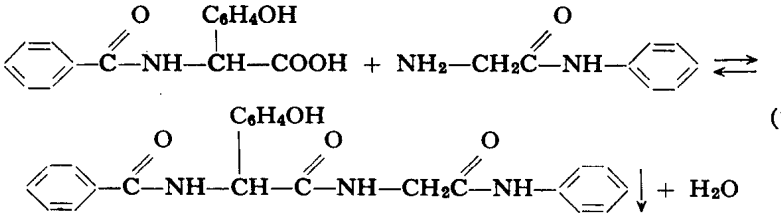
II. HISTORICAL INTRODUCTION

A. Preisotope Background

It is appropriate that the first volume of a series devoted to protein biosynthesis should provide a historical perspective of the field. More important from the point of view of the student in 1970 is a realization of the incredible speed with which concepts have changed in two decades: how apparently irrelevant material has been correlated; how research in vastly different disciplines has been brought to bear on the problem of protein synthesis; and what part has been played by brilliant thought, skillful experimentation, new techniques, and the capitalization on chance observation.

By 1945, proteins had been recognized as a class of organic compounds for a century, and the basic peptide linkage had been known for 50 years, as a result of Fischer's work. Ultracentrifugation studies, electrophoresis, colligative properties, and viscosity measurements showed clearly that proteins vary substantially in molecular weight, electric charge, and shape. X-ray diffraction revealed orderliness that was not interpretable. It was known that enzymes were largely or exclusively protein in composition. Even though several enzymes had been crystallized, the prevailing view was that proteins consist of more or less definite proportions of some 20 amino acids bound indiscriminately together in peptide linkage with no particular sequence dominant. Proteins were, in the words of the period, colloids, our understanding of which was not likely to be an extension of the understanding of simpler molecules and less complex chemistry. Aside from chemical methods, the only means by which a peptide bond had been synthesized used a variety of proteolytic enzymes. For example, chymotrypsin was known to catalyze the hydrolysis of peptides involving aromatic amino acids; thus benzoyltyrosylglycine amide reacts quantitatively with water to yield benzoyl tyrosine and glycine amide. However, enzymes, like

all catalysts, serve only to accelerate the attainment of an equilibrium and if the equilibrium can be displaced by precipitation of one component, synthesis rather than hydrolysis will occur. Taking advantage of this possibility, Bergmann and Fruton (1938) observed that benzoyl tyrosine reacts with glycine anilide in the presence of water to generate a new peptide bond according to the scheme



(1)

The reaction proceeds in the direction of synthesis because of the insolubility of the product. Bergmann and his associates realized that the insolubility of a peptide product could be used only occasionally to drive a reaction toward synthesis. However, they found that various proteolytic enzymes are extremely effective catalysts for transpeptidations. For example Fruton, Hearn, Ingram, Wiggans, and Winitz (1953)

$$\begin{aligned} &3 \text{ alanylphenylalanylamide} \xrightarrow{\text{cathepsin C}} \\ &(\text{alanylphenylalanyl})_2 \text{ alanylphenylalanylamide}\downarrow + 2NH_3 \end{aligned} \qquad (2)$$

were able to form a hexapeptide from a dipeptide amide. Since many proteolytic enzymes were known and since these had widely varying specificities, it seemed possible that a limited number of amidating or esterifying enzymes might suffice to provide "activated" forms of the 20 amino acids. The ubiquitous tripeptide, glutathione (γ-glutamylcysteinyl glycine), was observed to react with a variety of amino acids in the presence of intracellular enzymes to produce new γ-glutamyl peptides plus cysteinyl glycine (Hanes, Hird, and Isherwood, 1950). The concept that enzymic synthesis of a few peptides could be followed by transamidation of these peptides to form protein molecules seemed consistent with the presence of many peptides and enzymes that appeared to have no other biological function. This view was by no means dead in 1960 (Holden, 1962).

The only techniques for determining the amino acid composition of a protein were at best tedious and inaccurate. Only a few of the amino

acids could be isolated directly by crystallization and a few others by the crystallization of insoluble chemical derivatives. For the most part estimates of the quantity of an amino acid present in a hydrolyzate were made on the basis of the dilution of a hydrolyzate that would support growth of a microbe deficient in a particular amino acid. Not surprisingly, none of the analytical data could persuade investigators that a particular protein had an absolutely unique structure.

If the biochemical and organic chemical understanding of protein was primitive in 1945, the knowledge of the structure and function of nucleic acids was more meager. Polymers of high molecular weight containing a high percentage of nitrogen and phosphorus had been known since 1869. Yeast and thymus tissue were excellent sources of nucleic acid. The nucleic acids from these two sources had many properties in common, but the thymus nucleic acid was apparently of higher molecular weight and possessed greater stability to alkali. Generally it was believed that "yeast nucleic acid" (now known as ribonucleic acid, RNA) was characteristic of plant tissues and that "thymus nucleic acid" (now deoxyribonucleic acid, DNA) was unique to animal cells, although evidence was continually accumulating that some RNA was found in animal tissues and some DNA was found in plant tissues. The major base constituents of RNA were identified as uracil, cytosine, adenine, and guanine, each linked to a ribose molecule. The individual ribose residues were linked by phosphoric ester bonds to form a polymer estimated to have a molecular weight of anywhere from 15,000 to several million daltons. Unlike the amino acids which are largely only bifunctional and hence capable of forming only linear polymers, the nucleotides are multifunctional and potentially capable of forming by polymerization a highly branched structure, a possibility that was being actively considered in the 1950's. Analytical data were so poor that it was entirely reasonable to propose that all four bases were present in equal proportions and that the molecule was a repeating tetramer. Except for the replacement of uracil by thymine and the replacement of ribose by deoxyribose, DNA was similar to RNA. Despite evidence to the contrary, "nucleic acid" was generally believed to exist in the nucleus. Except for the division of nucleic acids into DNA and RNA there was no suggestion that there was a native heterogeneity in size, composition, and base sequence of nucleic acids.

This then, was the state of the art about 1945. There had been voices that, in retrospect, were prophetic. Based on nutritional studies Rose (1928) had argued, "If a tissue (protein) is to be formed at all, every

component must be available or be capable of being manufactured by the cell; otherwise the synthesis will not occur." Such a conclusion was surely not to be reconciled easily with a casual assembly of polypeptides from small peptides and enzymes of modest specificity. Borsook and Dubnoff (1940) calculated and determined experimentally the free energy of formation of a peptide bond and argued the necessity of a highly specific coupling with an exergonic reaction. Linderstrøm-Lang (1949) presented evidence for entirely distinct pathways for synthesis and degradation of protein. Lipmann (1941) suggested that some form of high-energy phosphate might be required to "activate" amino acids and Chantrenne (1948) had nominated acyl phosphates as candidates for the role of "activated intermediate." Brachet (1941) had unambiguous histochemical evidence of RNA in animal cells and Caspersson (1941) had correlated rapid protein synthesis with a high concentration of RNA. The X-ray analysis of the dimensions of simple peptides and nucleotides (Corey, 1948) had been undertaken in order to facilitate the interpretation of the X-ray diffraction patterns of proteins and nucleic acids. Mirsky and Pauling (1936) had long since suggested that hydrogen bonds stabilized the structures of macromolecules.

B. Postwar Explosion

For the most part the study of protein biosynthesis was drifting, waiting for new techniques, new leads, new concepts. With the end of the war, radioisotopes became available for biological research. The use of tracers made possible for the first time the measurement of protein synthesis outside the living animal. Quickly demonstrated was the *in vitro* incorporation of labeled amino acids into rat liver protein (Frantz, Loftfield, and Miller, 1947; Melchior and Tarver, 1947; Winnick, Friedberg, and Greenberg, 1947), into tumor protein (Zamecnik, Frantz, Loftfield, and Stephenson, 1948), into erythropoietic tissue (Borsook, Deasy, Haagen-Smit, Keighley, and Lowry, 1950), and into silk gland protein (Zamecnik, Loftfield, Stephenson, and Williams, 1949). Dinitrophenol and other respiratory inhibitors were quickly observed to inhibit protein synthesis, an immediate confirmation of the Borsook and Lipmann arguments that protein synthesis was endergonic (Frantz, Zamecnik, Reese, and Stephenson, 1948).

The development of the theory of partition chromatography (Martin

and Synge, 1941) and its application to paper chromatography (Consden, Gordon, and Martin, 1944) to ion-exchange chromatography (Tompkins, Khym, and Cohn, 1947) and to starch column chromatography (Stein and Moore, 1948) led to enormously more accurate and sensitive determinations of nucleotides (Vischer and Chargaff, 1947; Carter and Cohn, 1949; Cohn, 1950) and of amino acids (Moore and Stein, 1951). These new techniques were immediately applied to the determination of the structures of the relatively small protein molecules, insulin (Sanger and Tuppy, 1951; Ryle, Sanger, Smith, and Kitai, 1955) and β-corticotropin (Howard, Shepherd, Eigner, Davies, and Bell, 1955). For the first time biochemists had convincing evidence that many and perhaps all proteins possessed unique sequences of amino acids and that protein molecules had structures just as specific and definable as those of simple organic molecules. An immediate consequence was to demonstrate that the proteolytic enzymes are unable to discriminate against nonnatural amino acids such as α-aminobutyric acid well enough to account for the apparent precision of protein synthesis (Loftfield, Grover, and Stephenson, 1953). The synthetic and degradative paths were clearly different.

The new analytical techniques were also applied to nucleic acids. Careful work and imaginative correlation quickly led to the conclusion that not all DNA preparations were identical but that there were marked variations in the proportions of adenine, guanine, thymine, and cytosine. Strikingly, however, the content of adenine was essentially equal to that of thymine and the content of guanine was equal to that of cytosine (Chargaff, 1950). It became apparent that both RNA and DNA were linear polymers in which the nucleosides were linked exclusively by phosphate esters between the 3′ and 5′ positions (Cohn and Volkin, 1951, 1952) and that such a structure was consistent with the known chemistry of simpler phosphate esters (Brown and Todd, 1952). These new data and correlations, together with the advances in theory, technique, and interpretation of X-ray diffraction by Wilkins and by Pauling, provided the springboard for the Watson and Crick (1953) proposal for the structure of DNA. Suddenly, the accumulating evidence that DNA might be *the* genetic material, the concept of DNA replication by base pairing, the biological attractiveness of this novel chemical structure—all combined to convert DNA from a functionless wastebin of phosphorus and nitrogen into the most central of all metabolic materials.