SYNTHETIC POLYPEPTIDES

PREPARATION, STRUCTURE, AND PROPERTIES

By

C. H. BAMFORD

A. ELLIOTT W. E. HANBY

Courtaulds Ltd., Research Laboratory, Maidenhead, Berkshire, England

1956

ACADEMIC PRESS INC · PUBLISHERS · NEW YORK

Preface

The study of synthetic polypeptides has proceeded with increasing intensity in the post-war years, and now appears to have reached a stage at which a comprehensive survey is justified and desirable. This is the excuse for writing the present book. It will be unnecessary to tell potential readers that the peculiar interest attached to synthetic polypeptides lies in their relation to the proteins. Although this matter is treated specifically only in the last chapter, it has been continually in our minds during writing and, we hope, will have left its imprint upon the whole work. The interest in polypeptides is not confined to their relation to proteins; they constitute a class of high polymers with unique characteristics, exhibiting with unusual clarity the relationship between properties and molecular structure. Their structure has been studied extensively by infrared and X-ray methods, which are now being applied with increasing precision to high polymers. We have therefore included some account of the fundamental principles of these techniques and their application to the peculiar problems posed by fibrous materials. The theme as we have developed it should, we venture to believe, interest those who work in the difficult but fascinating fields of protein structure and properties, whether on biochemical, medical, or physicochemical aspects, and to all who are concerned generally with high polymers.

In high polymer research, experience shows that advances follow most readily from concerted attacks by all available techniques; we have therefore found it necessary to discuss in detail an unusually wide range of topics. In a book of reasonable size it is clearly impossible to carry the development of each from first principles to the current position. However, we have tried to effect a compromise between this and the omission of all elementary material. The book generally can be read by all who have a basic interest and training in the natural sciences; the chapters on X-ray diffraction studies assume, however, an acquaintance with the nomenclature of crystal geometry.

We are indebted to many of our colleagues in the laboratories of Courtaulds Ltd. for much invaluable help in the preparation of this book. In particular we wish to thank Dr. S. G. Waley of the Nuffield Laboratory of Ophthalmology, Oxford, for writing Chapter XI, Messrs. L. Brown and

I. F. Trotter, Miss E. M. Cant and Dr. J. Crank for reading and criticizing the X-ray chapters, Drs. D. G. H. Ballard, B. R. Malcolm, C. Robinson, and F. J. Weymouth for providing the assistance of their expert knowledge, and Messrs. E. C. Collingwood and W. A. Porter for preparing the diagrams and photographs. Many workers have given us permission to reproduce figures from their publications; they are too numerous to mention individually here but acknowledgment is made in each case in the text. Finally, our thanks are due to the Officers of the Royal Society, the Faraday Society, the National Academy of Sciences (Washington), and the Editors of *Nature* and of the *British Journal of Applied Physics* for permission to reproduce diagrams and photographs from original papers.

March, 1956

C. H. BAMFORD
A. ELLIOTT
W. E. HANBY

Contents

CHAPTER III

The Mechanism of Polypeptide Synthesis from *N*-Carboxy-α-Amino Acid Anhydrides

CHAPTER IV

Chain Configurations in Polypeptides

CHAPTER V

Infrared Spectra

CHAPTER VI

So ne Special Topics in Infrared Spectroscopy

CHAPTER VII

The Study of Synthetic Polypeptides by X-ray Diffraction

CHAPTER VIII

Further Studies of the X-Ray Diffraction Patterns of Alpha Polypeptides

CHAPTER IX

X-Ray Studies of Synthetic Polypeptides in the Beta Form

CHAPTER X

The Properties of Synthetic Polypeptides

CHAPTER XI

Biological Properties of Synthetic Polypeptides

CHAPTER XII

Fibrous Proteins and their Relation to Synthetic Polypeptides

APPENDIX

CHAPTER I

The Role of Synthetic Polypeptides as Protein Models

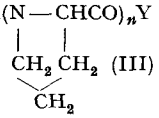
1. Introduction

The polymers which are the subject of this book are the synthetic linear polypeptides derived from α-amino acids. The simplest polypeptides contain only one type of α-amino acid residue, and their general formula is shown in (I).

$$X(NH\cdot CHR\cdot CO)_nY \qquad (I)$$

The side-chain characteristic of the α-amino acid is denoted by R; X and Y are end groups, the nature of which is determined by the method of preparation, and n is the degree of polymerization. In a few cases the hydrogen atoms attached to the nitrogen or α-carbon atoms are replaced by alkyl groups. This is so for example in the polypeptides of sarcosine (II), proline (III), and α-aminoisobutyric acid (IV).

$X(NMeCH_2CO)_nY$ (II)


(III)

$X(NHCMe_2CO)_nY$ (IV)

As far as is known, all the naturally occurring α-amino acids except sarcosine, proline, and hydroxyproline possess an unsubstituted α-NH_2 group, and all have a hydrogen atom attached to the α-carbon atom. The following chapters will deal mainly with the polypeptides symbolized by (I).

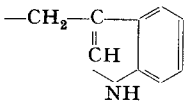
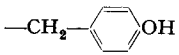
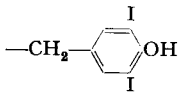
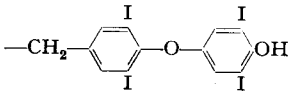
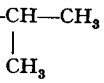
Copolymers of two or more different α-amino acids may be prepared and of course contain R groups of different kinds. A list of the commonest α-amino acids is given in Table 1.1. As is the case with other classes of polymers, copolymers may have either a random arrangement of residues or an arrangement according to some definite pattern. The random copolymers are the easiest to prepare and are made by polymerizing a mixture of the different monomers. The different types of residue are introduced into the

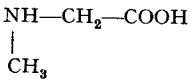
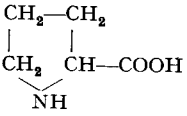
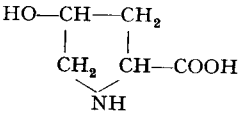
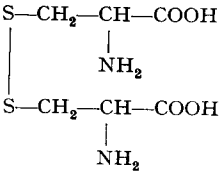
TABLE 1.1

COMMON α-AMINO ACIDS, $NH_2CHRCOOH$

Trivial name	R	Residue weight —NH·CHR·CO—
Alanine	$-CH_3$	71
Arginine	$-CH_2-CH_2-CH_2-NH-C(=NH)-NH_2$	156
Aspartic acid	$-CH_2-COOH$	115
Asparagine	$-CH_2-CONH_2$	114
Glutamic acid	$-CH_2-CH_2-COOH$	129
Glutamine	$-CH_2-CH_2-CONH_2$	128
Cysteine	$-CH_2-SH$	103
Glycine	$-H$	57
Histidine	$-CH_2-C(=CH-NH-)=N-$	137
Leucine	$-CH_2-CH(CH_3)_2$	113
Isoleucine	$-CH(CH_3)-CH_2-CH_3$	113
Norleucine	$-CH_2-CH_2-CH_2-CH_3$	113
Lysine	$-CH_2-CH_2-CH_2-CH_2-NH_2$	128
Methionine	$-CH_2-CH_2-SCH_3$	131
Ornithine	$-CH_2-CH_2-CH_2-NH_2$	114
Citrulline	$-CH_2-CH_2-CH_2-NH-CO-NH_2$	157
Phenylalanine	$-CH_2-C_6H_5$	147
Serine	$-CH_2-OH$	87
Threonine	$-CH(CH_3)-OH$	101

TABLE 1.1 *(Continued)*

Trivial name	R	Residue weight —NH·CHR·CO—
Tryptophane		186
Tyrosine		163
Iodogorgoic acid		415
Thyroxine		759
Valine		99
Norvaline	$-CH_2-CH_2-CH_3$	99

Trivial name	Formula	Molecular weight
Sarcosine*		89
Proline*		115
Hydroxyproline*		131
Cystine		240

* Strictly speaking these are *imino* acids.

polymer at rates determined by the reactivities of the corresponding monomers. The structure of such polymers is not known precisely, since there is a statistical distribution of residues along the polymer chains.

Copolymers of more precise structure may be termed *block* copolymers. In the simplest block copolymer the molecules consist of segments, each of which is composed of one kind of residue only. This type of polymer is represented in (V):

$$X\,A\,A\cdots A\,A\,B\,B\cdots B\,B\,C\,C\cdots C\,C\ldots Y \qquad (V)$$

in which A B C . . . . are different α-amino acid residues. In a second type of block copolymer each segment consists of a few different α-amino acid residues arranged in a definite order, e.g., ABC. The copolymer has the structure (VI).

$$X(ABC)_nY \qquad (VI)$$

Very little is known about the preparation and properties of block copolymers of α-amino acids. The preparation of a high polymer containing a specified order of α-amino acid residues not constructed on the block plan is possible in principle but would be prohibitively laborious and has not been carried out. The remarkable achievements of Work (*see* for example Harris and Work, 1950), du Vigneaud (1953), and their collaborators in the synthesis of polypeptides of known structure are noteworthy in this connection.

2. The Interest in Synthetic Polypeptides: Folding of Polypeptide Chains

The Fischer-Hofmeister theory (Fischer 1902, Hofmeister 1902) that polypeptide chains (often long chains) are a major constituent of proteins is supported by a great weight of evidence, and is now generally accepted as fact. This endows the synthetic polypeptides with a particular interest, since they may be regarded as simple models of proteins. One of the main topics of this book is the folding of polypeptide chains, consequently we shall be mainly concerned with relatively high polymers containing several hundred residues per molecule.

Astbury and his colleagues (1931, 1933, 1935) first demonstrated that in some fibrous proteins the polypeptide chains are not in the simple extended configuration but are folded in a regular manner. The idea of chain folding has subsequently become a constantly recurring theme in protein chemistry and physics. It is now recognized that the biological activity of a protein is

intimately connected with the configuration of its constituent polypeptide chains; moreover, in their native state most proteins (not only fibrous proteins) contain folded chains. If the folding is destroyed the protein usually becomes insoluble and inactive (i.e., denatured). It is hardly necessary to emphasize here the great interest attached to the determination of the chain configurations in proteins.

Some globular proteins (e.g., the hemoglobins, ribonuclease) can be obtained in a highly crystalline state and have been studied intensively by X-ray diffraction. Up to the present this work has not given definite information about the chain configurations. Indeed, although the chemical evidence for the existence of polypeptide chains in globular proteins is quite compelling, the chains are by no means so prominent in the X-ray analysis as might be expected. This situation apparently arises because the chains are folded in a complex manner; it is likely that there is a primary fold, upon which is superimposed some secondary folding. The fact that the molecules can form crystals does not show that the folding is simple, but merely that all the molecules have the same shape and contain chains folded in the same way. The elucidation of the chain configuration would therefore seem to require a complete structure determination. The magnitude of this task will be appreciated when it is realized that the unit cells of these crystalline proteins may contain several hundred atoms.

Obviously no simple polypeptide can be an adequate model of a protein in all respects. The behaviour of a protein molecule is determined both by the configuration of the backbone and by the character of the side chains. We have chosen, as a first step, to study the behaviour of the polypeptide backbone carrying non-polar side-chains, in an attempt to decide whether chain folding is dependent upon some specific biological factor, or upon the nature of the backbone, or upon the character of the side-chains, or upon a combination of all three. We do not wish to suggest that the behaviour of a protein molecule with its complicated pattern of electric charges can be simulated by a simple polypeptide. It would be expected, however, that the synthetic polypeptides would be able to provide information about the way in which the polypeptide chain can fold. We must be careful not to assume without good reason that any fold found in the synthetic materials is the same as that in a protein. Nevertheless the synthetic polymers must almost certainly assist eventually in the interpretation of data on proteins. It will be seen from the following chapters that most synthetic polypeptides can be obtained in α and β forms, containing respectively folded and extended chains. All the folded polypeptides have essentially the same configuration, and the principles underlying the folding are fairly well understood. The determination of this configuration has been possible because the α polymers have proved amenable to the standard techniques of polymer science and can

be obtained as highly oriented fibers or films which are free from secondary folds and are therefore very suitable for X-ray and infrared work. In this respect the synthetic materials have an enormous advantage over the proteins; generally the secondary folds in the latter cannot be undone without producing other structural changes such as disruption of the primary fold.

It is important to realize the difference between chain folding of the type we are considering, and the coiling of polymer chains which occurs generally in solutions of high polymers. This latter, the result of Brownian motion, is a purely random phenomenon. In general all the molecules will have different shapes, which are continually changing with time. The folds in the chains of synthetic polypeptides are perfectly regular and are held in position by strong forces between groups uniformly disposed along the chains. These forces are hydrogen bonds formed between suitably placed peptide (–NHCO–) groups, and the regular folding results from the tendency of the system to take up the configuration of minimum potential energy. The primary folds in protein molecules may be of a similar kind, but the secondary folds may involve other kinds of forces, e.g. salt links or true chemical cross links such as disulfide bonds. If the polypeptide (or protein) is dissolved in a liquid which can form sufficiently strong hydrogen bonds with either the >NH or >CO groups, the interpeptide hydrogen bonds are broken and the regular (primary) folding of the chains is lost. In solution in weakly interacting liquids, however, the characteristic α-fold is retained (Bamford, Hanby and Happey, 1951 *b*; Robinson and Bott, 1951; Doty *et al.*, 1954). Some random coiling may of course also occur in these solutions.

The synthetic polypeptides would also be expected to possess interesting properties as high polymers in their own right, as it were, apart from the specific interest due to their relationship to the proteins. Thus all naturally occurring fibers — wool, hair, the silks, collagen — are polypeptides, and it would be strange if some of the high-molecular-weight synthetic materials were not good fiber formers. In this connection it will be recalled that the nylons and perlons contain the peptide group; the polypeptides may in fact be regarded as derivatives of "nylon 1." The patent literature reveals that several laboratories have prepared polypeptide fibers. The literature affords comparatively little information on the properties of synthetic polypeptides, and practically nothing has been published on mechanical properties. However, some polypeptides are known to possess remarkable self-orienting properties, an account of which is included in Chapter X. These properties must be closely associated with mechanical behavior of the polymers, and it may well appear that they are of importance for understanding other materials besides polypeptides.