Amino Acids and Peptides

G. C. BARRETT
& D. T. ELMORE

Amino Acids and Peptides

G. C. BARRETT

AND

D. T. ELMORE

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge CB2 1RP, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, United Kingdom
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1998

Printed in the United Kingdom at the University Press, Cambridge

Typeset in Times NR 10/13pt [SE]

A catalogue record for this book is available from the British Library

ISBN 0 521 46292 4 hardback
ISBN 0 521 46827 2 paperback

The authors' objective has been to concentrate on amino acids and peptides without detailed discussions of proteins, although the book gives all the essential background chemistry, including sequence determination, synthesis and spectroscopic methods, to allow the reader to appreciate protein behaviour at the molecular level. The approach is intended to encourage the reader to cross classical boundaries, such as in the later chapter on the biological roles of amino acids and the design of peptide-based drugs. For example, there is a section on enzyme-catalysed synthesis of peptides, an area often neglected in texts describing peptide synthesis.

This modern text will be of value to advanced undergraduates, graduate students and research workers in the amino acid, peptide and protein field.

Amino Acids and Peptides

Foreword

This is an undergraduate and introductory postgraduate textbook that gives information on amino acids and peptides, and is intended to be self-sufficient in all the organic and analytical chemistry fundamentals. It is aimed at students of chemistry, and allied areas. Suggestions for supplementary reading are provided, so that topic areas that are not covered in depth in this book may be followed up by readers with particular study interests.

A particular objective has been to concentrate on amino acids and peptides, as the title of the book implies; the exclusion of detailed discussion of proteins is deliberate, but the book gives all the essential background chemistry so that protein behaviour at the molecular level can be appreciated.

There is an emphasis on the uses of amino acids and peptides, and on their biological roles and, while Chapter 8 concentrates on this, a scattering of items of information of this type will be found throughout the book. Important pharmaceutical developments in recent years underline the continuing importance and potency of amino acids and peptides in medicine and the flavour of current research themes in this area can be gained from Chapter 9.

Supplementary reading
(see also lists at the end of each Chapter)

Standard Student Texts

Standard undergraduate Biochemistry textbooks relate the general field to the coverage of this book. Several such topic areas are covered in

Zubay, G. (1993) *Biochemistry*, Third Edition, Wm. C. Brown Communications Inc, Dubuque, IA

and

Voet, D. and Voet, J. G. (1995) *Biochemistry*, Second edition, Wiley, New York

Typically, these topic areas as covered by Zubay are

Chapter 3: 'The building blocks of proteins: amino acids, peptides and proteins'
Chapter 4: 'The three-dimensional structure of proteins'
Chapter 5: 'Functional diversity of proteins'

Removed more towards biochemical themes, are

Chapter 18: 'Biosynthesis of amino acids'
Chapter 19: 'The metabolic fate of amino acids'
Chapter 29: 'Protein synthesis, targeting, and turnover'

Voet and Voet give similar coverage in

Chapter 24: 'Amino acid metabolism'
Chapter 30: 'Translation' (i.e. protein biosynthesis)
Chapter 34: 'Molecular physiology' (of particular relevance to coverage in this book of blood clotting, peptide hormones and neurotransmitters)

Supplementary reading: suggestions for further reading

(a) Protein structure

Branden, C., and Tooze, J. (1991) *Introduction to Protein Structure*, Garland Publishing Inc., New York

(b) Protein chemistry

Hugli, T. E. (1989) *Techniques of Protein Chemistry*, Academic Press, San Diego, California
Cherry, J. P. and Barford, R. A. (1988) *Methods for Protein Analysis,* American Oil Chemists' Society, Champaign, Illinois

(c) Amino acids

Barrett, G. C., Ed. (1985) *Chemistry and Biochemistry of the Amino Acids*, Chapman and Hall, London
Barrett, G. C. (1993) in *Second Supplements to the 2nd Edition of Rodd's Chemistry of Carbon Compounds*, Volume 1, Part D: Dihydric alcohols, their oxidation products and derivatives, Ed. Sainsbury, M., Elsevier, Amsterdam, pp. 117–66
Barrett, G. C. (1995) in *Amino Acids, Peptides, and Proteins*, A Specialist Periodical Report of The Royal Society of Chemistry, Vol. 26, Ed. Davies, J. S., Royal Society of Chemistry, London (preceding volumes cover the literature on amino acids, back to 1969 (Volume 1))
Coppola, G. M. and Schuster, H. F. (1987) *Asymmetric Synthesis: Construction of Chiral Molecules using Amino Acids*, Wiley, New York
Dawson, R. M. C., Elliott, D. C., Elliott, W. H., and Jones, K. M. (1986) *Data for Biochemical Research*, Oxford University Press, Oxford

Greenstein, J. P., and Winitz, M. (1961) *Chemistry of the Amino Acids*, Wiley, New York (a facsimile version (1986) of this three-volume set has been made available by Robert E. Krieger Publishing Inc., Malabar, Florida)

Williams, R. M. (1989) *Synthesis of Optically Active α-Amino Acids*, Pergamon Press, Oxford

(d) Peptides

Bailey, P. D. (1990) *An Introduction to Peptide Chemistry*, Wiley, Chichester

Bodanszky, M. (1988) *Peptide Chemistry: A Practical Handbook*. Springer-Verlag, Berlin

Bodanszky, M. (1993) *Principles of Peptide Synthesis*, Second Edition, Springer-Verlag, Heidelberg

Elmore, D. T. (1993) in *Second Supplements to the 2nd Edition of Rodd's Chemistry of Carbon Compounds*, Volume 1, Part D: Dihydric alcohols, their oxidation products and derivatives, Ed. Sainsbury, M., Elsevier, Amsterdam, pp. 167–211

Elmore, D. T. (1995) in *Amino Acids, Peptides, and Proteins*, A Specialist Periodical Report of The Royal Society of Chemistry, Vol. 26, Ed. Davies, J. S., Royal Society of Chemistry, London (preceding volumes cover the literature of peptide chemistry back to 1969 (Volume 1))

Jones, J. H. (1991) *The Chemical Synthesis of Peptides*, Clarendon Press, Oxford

Contents

1

Introduction

1.1 Sources and roles of amino acids and peptides

More than 700 amino acids have been discovered in Nature and most of them are α-amino acids. Bacteria, fungi and algae and other plants provide nearly all these, which exist either in the free form or bound up into larger molecules (as constituents of peptides and proteins and other types of amide, and of alkylated and esterified structures).

The twenty amino acids (actually, nineteen α-amino acids and one α-imino acid) that are utilised in living cells for protein synthesis under the control of genes are in a special category since they are fundamental to all life forms as building blocks for peptides and proteins. However, the reasons why all the other natural amino acids are located where they are, are rarely known, although this is an area of much speculation. For example, some unusual amino acids are present in many seeds and are not needed by the mature plant. They deter predators through their toxic or otherwise unpleasant characteristics and in this way are thought to provide a defence strategy to improve the chances of survival for the seed and therefore help to ensure the survival of the plant species.

Peptides and proteins play a wide variety of roles in living organisms and display a range of properties (from the potent hormonal activity of some small peptides to the structural support and protection for the organism shown by insoluble proteins). Some of these roles are illustrated in this book.

1.2 Definitions

The term '*amino acids*' is generally understood to refer to the *aminoalkanoic acids*, $H_3N^+—(CR^1R^2)_n—CO_2^-$ with $n = 1$ for the series of α-amino acids, $n = 2$ for β-amino acids, etc. The term '*dehydro-amino acids*' specifically describes *2,3-unsaturated (or 'αβ-unsaturated')-2-aminoalkanoic acids*, $H_3N^+—(C{=}CR^1R^2)—CO_2^-$.

However, the term '*amino acids*' would include all structures carrying amine and acid functional groups, including simple aromatic compounds, e.g. anthranilic acid,

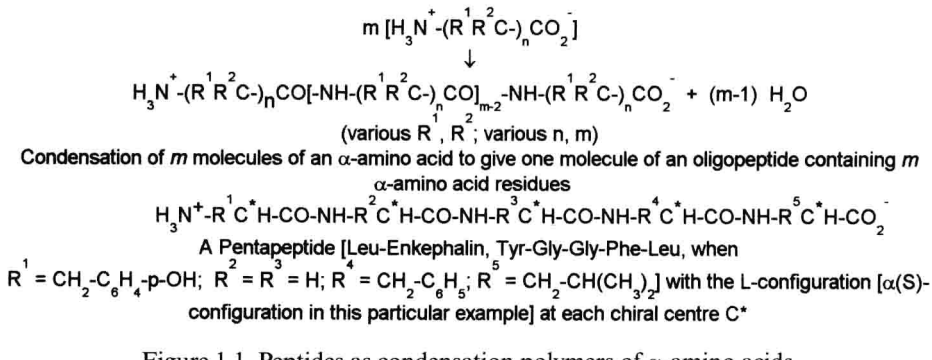

Figure 1.1. Peptides as condensation polymers of α-amino acids.

o-H_3N^+—C_6H_4—CO_2^-, and would also cover other types of acidic functional groups (such as phosphorus and sulphur oxy-acids, H_3N^+—$(R^1R^2C—)_nHPO_3^-$ and R_3N^+—$(R^1R^2C—)_nSO_3^-$, etc). The family of boron analogues $R_3N^{\cdot}BHR^1$—CO_2R^2 ($^{\cdot}$ denotes a dative bond) has recently been opened up through the synthesis of some examples (Sutton *et al.*, 1993); it would take only the substitution of the carboxy group in these 'organoboron amino acids' ($R = R^1 = R^2 = H$) by phosphorus or sulphur equivalents to obtain an amino acid that contains no carbon! However, unlike the amino acids containing sulphonic and phosphonic acid groupings, naturally occurring examples of organoboron-based amino acids are not known.

The term '*peptides*' has a more restricted meaning and is therefore a less ambiguous term, since it covers polymers formed by the condensation of the respective amino and carboxy groups of α, β, γ . . . -amino acids. For the structure with $m = 2$ in Figure 1.1 (i.e., for a dipeptide) up to values of $m \simeq 20$ (an eicosapeptide), the term '*oligopeptide*' is used and a prefix *di-*, *tri-*, *tetra-*, *penta-* (see Leu-enkephalin, a linear pentapeptide, in Figure 1.1), . . . *undeca-* (see cyclosporin A, a cyclic undecapeptide, in Figure 1.4 later), *dodeca-*, . . . etc. is used to indicate the number of *amino-acid residues* contained in the compound. *Homodetic* and *heterodetic* peptides are illustrated in Chapter 7.

Isopeptides are isomers in which amide bonds are present that involve the *side-chain amino group* of an αω-di-amino acid (e.g. lysine) or of a poly-amino acid and/or *the side-chain carboxy-group of an α-amino-di- or -poly-acid* (e.g. aspartic acid or glutamic acid). Glutathione (Chapter 8) is a simple example. Longer polymers are termed '*polypeptides*' or '*proteins*' and the term '*polypeptides*' is becoming the most commonly used general family name (though proteins remains the preferred term for particular examples of large polypeptides located in precise biological contexts). Nonetheless, the relationship between these terms is a little more contentious, since the change-over from polypeptide to protein needs definition. The figure 'roughly fifty amino acid residues' is widely accepted for this. Insulin (a polymer of fifty-one α-amino acids but consisting of two crosslinked oligopeptide

chains; see Figure 1.4 later) is on the borderline and has been referred to both as a *small protein* and as a *large polypeptide*.

Poly(α-amino acid)s is a better term for peptides formed by the self-condensation of one amino acid; natural examples exist, such as poly(D-glutamic acid), the protein coat of the anthrax spore (Hanby and Rydon, 1946). In early research in the textile industry, poly(α-amino acid)s showed promise as synthetic fibres, but the synthesis methodology required for the polymerisation of amino acids was complex and uneconomic.

Polymers of controlled structures made from *N*-alkyl-α-amino acids (Figure 1.1; —NR^n instead of —NH—, $R^1=R^2=H$; $n=1$), i.e. $H_2{}^+NR^n—CH_2CO—[NR^n—CH_2—CO—]_mNR^n—CH_2—CO_2^-$, which are poly(*N*-alkylglycine)s of defined sequence (various R^n at chosen points along the chain), have been synthesised as *peptide mimetics* (see Chapter 9) and have been given the name *peptoids*. These can be viewed as peptides with side-chains shifted from carbon to nitrogen; they will therefore have a very different conformational flexibility (see Chapter 2) from that of peptides and will also be incapable of hydrogen bonding. This is a simple enough way of providing all the correct side-chains on a flexible chain of atoms, in order to mimic a biologically active peptide, but the mimic can avoid enzymic breakdown before it reaches the site in the body where it is needed.

Using the language of polymer chemistry, polypeptides made from two or more different α-amino acids are *copolymers* or irregular poly(amide)s, whereas poly(α-amino acid)s, $H—[NH—CR^1R^2—CO—]_mOH$, are *homopolymers* that could be described as members of the nylon[2] family.

Depsipeptides are near-relatives of peptides, with one or more *amide bonds replaced by ester bonds*; in other words, they are formed by condensing α-amino acids with α-hydroxy-acids in various proportions. There are several important natural examples of these, of defined sequence; for example the antibiotic valinomycin and the family of enniatin antibiotics. Structures of other examples of depsipeptides are given in Section 4.8.

Nomenclature for conformational features of peptide structure is covered in Chapter 2.

1.3 'Protein amino acids', alias 'the coded amino acids'

The twenty L-amino acids (actually, nineteen α-amino acids and one α-imino acid (Table 1.1)) which, in preparation for their role in protein synthesis, are joined *in vivo* through their carboxy group to tRNA to form α-aminoacyl-tRNAs, are organised by ribosomal action into specific sequences in accordance with the genetic code (Chapter 8).

'Coded amino acids' is a better name for these twenty amino acids, rather than 'protein amino acids' or 'primary protein amino acids' (the term 'coded amino acids' is increasingly used), because changes can occur to amino-acid residues after they have been laid in place in a polypeptide by ribosomal synthesis. Greenstein and