

SPECIALIST PERIODICAL REPORTS

**Amino - acids
Peptides
and Proteins
VOLUME 4**

A Specialist Periodical Report

**Amino-acids, Peptides, and
Proteins**
Volume 4

**A Review of the Literature Published
during 1971**

Senior Reporter

**G. T. Young, *The Dyson Perrins Laboratory, University
of Oxford***

ISBN: 0 85186 034 6

Library of Congress Catalog Card No. 72-92548

© Copyright 1972

**The Chemical Society
Burlington House, London, W1V 0BN**

Abbreviations

Abbreviations for amino-acids and their use in the formulation of derivatives follow with some exceptions the revised (1971) Recommendations of the I.U.P.A.C.-I.U.B. Commission on Biochemical Nomenclature, which are reprinted in Chapter 5 of this volume.

Other abbreviations which have been used without definition are:

Adoc	adamantylloxycarbonyl
Aoc	t-amyloxycarbonyl
Asu	α -aminosuberic acid
Asx	aspartic acid or asparagine (not yet determined)
ATP	adenosine 5'-triphosphate
Bpoc	2-(4-biphenyl)-isopropoxycarbonyl
BSA	bovine serum albumin
c.d.	circular dichroism
Cha	cyclohexylamine
Cm	carboxymethyl
Cmc	S-carboxymethylcysteine
Dce	2,2-diethoxycarbonyl
Dcha	dicyclohexylamine
DMF	NN-dimethylformamide
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
Dnp	2,4-dinitrophenyl
Dns	1-dimethylaminonaphthalene-5-sulphonyl (dansyl)
Dopa	3,4-dihydroxyphenylalanine
DP	degree of polymerization
Ec	ethylcarbonyl
edta	ethylenediamine tetra-acetate
e.p.r.	electron paramagnetic resonance
e.s.r.	electron spin resonance
Gal	galactose
g.l.c.	gas-liquid chromatography
Glc	glucose
Glp or Pca	pyrrolid-2-one-5-carboxylic acid
Glx	glutamic acid or glutamine (not yet determined)
GTP	guanosine 5'-triphosphate
i.r.	infrared

Man	mannose
NAD	nicotinamide-adenine dinucleotide (NAD ⁺ , oxidized; NADH, reduced)
n.m.r.	nuclear magnetic resonance
ONSu	succinimido-oxy
OPcp	pentachlorophenoxy
OPic	4-picolyloxy
o.r.d.	optical rotatory dispersion
OTcp	2,4,5-trichlorophenoxy
Pca	<i>see</i> Glp
Pipoc	piperidino-oxycarbonyl
Pth-Gly	the phenylthiohydantoin derived from glycine, <i>etc.</i>
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
Ser(P)	<i>O</i> -phosphorylserine
t.l.c.	thin-layer chromatography
u.v.	ultraviolet
Ztf	1-benzyloxycarbonylamino-2,2,2-trifluoro-ethyl

Preface

This fourth Report reviews papers relevant to the chemistry of amino-acids, peptides, and proteins, appearing in the main journals during 1971; the literature continues to increase, and this volume contains nearly 3,000 references. Metal derivatives are reviewed biennially in this series and are not covered this year; volume 5 will survey papers in this field appearing in 1971 and 1972. As in Volume 3, work on the structure and synthesis of cyclic peptides will be found in Chapter 4 (Peptides with Structural Features Not Typical of Proteins). In Chapter 5 we reprint for the convenience of readers the 1971 revision of the recommendations of the I.U.P.A.C.-I.U.B. Commission on Biochemical Nomenclature, 'Symbols for Amino-acid Derivatives and Peptides', together with their recommendations 'Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains'.

For the new reader, we would note that in place of a subject index (the preparation of which would delay publication unduly) there is an extended list of contents from which the sections relevant to a search can be ascertained. Some overlap between sections will be found, and within limits is no doubt desirable.

Finally, I express my gratitude here to the contributors who in these first four volumes have established this series as a service to their colleagues in research.

G. T. YOUNG

Contents

Chapter 1 Amino-acids

By B. W. Bycroft

1 Naturally Occurring Amino-acids	1
A Occurrence of Known Amino-acids	1
B New Natural Free Amino-acids	3
C New Amino-acids from Hydrolysates	5
2 Chemical Synthesis and Resolution of Amino-acids	5
A Introduction and General Methods	5
B Synthesis under Simulated Prebiotic Conditions	7
C Protein and Other Naturally Occurring Amino-acids	7
D C-Alkyl- and Substituted C-Alkyl- α -amino-acids	8
E α -Amino-acids with Aliphatic Hydroxy-groups in the Side-chain	9
F Aromatic and Heterocyclic α -Amino-acids	10
G N-Substituted α -Amino-acids	11
H α -Amino-acids containing Sulphur	12
I A List of α -Amino-acids which have been Synthesized for the First Time	12
J Labelled Amino-acids	14
K Resolution of α -Amino-acids	16
3 Physical and Stereochemical Studies of α-Amino-acids	17
A Crystal Structures of Amino-acids	17
B Nuclear Magnetic Resonance Spectra	18
C Optical Rotatory Dispersion and Circular Dichroism	19
D Mass Spectrometry	21
E Other Physical and Stereochemical Studies	21
4 Chemical Studies of Amino-acids	22
A Introduction	22
B General Reactions	22
C Specific Reactions	24
D Non-enzymic Models of Biochemical Processes involving Amino-acids	26
E Effects of Electromagnetic Radiation on Amino-acids	27

5 Analytical Methods	29
A Gas-Liquid Chromatography	29
B Ion-exchange Chromatography	30
C Thin-layer Chromatography	30
D Other Methods	31
E Determination of Specific Amino-acids	31

Chapter 2 Structural Investigation of Peptides and Proteins

1: Primary Structure and Chemical Modification

By R. N. Perham and J. O. Thomas

1 Introduction	32
2 Methods	32
A Amino-acid Analysis	33
Ion-exchange Chromatography	34
High-voltage Electrophoresis and Thin-layer Chromatography	35
Gas-Liquid Chromatography	36
B End-group Analysis and Sequential Degradation	36
C Mass Spectrometry	39
D Cleavage of Protein Chains	44
Enzymic Cleavage	44
Restriction of Enzymic Cleavage	45
Insolubilized Enzymes	45
E Fractionation Methods	47
Peptide Separation, Detection, and Identification	47
Chromatography	49
Electrophoresis	50
Isoelectric Focusing	55
Affinity Chromatography	56
3 Structural Proteins	61
A The Proteins of Motility	61
B Collagen	62
Primary Structure	63
Cross-links	64
Hydroxylated Residues	65
C Elastin	66
D Fibrinogen	66
E Chromosomal Proteins	67
F Ribosomal Proteins	69
G Serum and Egg Proteins	70

H Miscellaneous	72
Crystallins	72
Keratins and Wool Proteins	73
Casein	75
Other Proteins	75
4 Peptides and Hormones	77
A Pancreatic Hormones	77
B Pituitary Hormones	78
C Other Hormones and Peptides	82
D Toxins	83
E Proteins of the Nervous System	86
5 Enzymes	88
A Proteolytic Enzymes	88
Carboxypeptidases	88
Serine Proteases and Their Inhibitors	90
Neutral Proteases, Acid Proteases, and Thiol Proteases	102
B Lysozyme and α -Lactalbumin	106
C Dehydrogenases	109
D Aldolases	114
E Nucleases	116
F Pyridoxal Phosphate Enzymes	118
G Other Enzymes	121
H Quaternary Structure	128
6 Electron-transport and Oxygen-transport Proteins	128
A Electron-transport Proteins	128
Cytochromes	128
Other Electron-transport Proteins	136
B Oxygen-transport and -storage Proteins	140
Haemoglobin and Myoglobin	140
Haemerythrin	147
7 Immunoglobulins	148
A Light Chains	149
B Heavy Chains	151
C Disulphide Bridges	153
D Antibody Binding Sites	154
E Some Problems of Biosynthesis	156
8 Membrane Proteins	157
A Solubilization and Fractionation	157
B Red Blood Cell Membranes	158
C Mitochondrial and Other Membranes	160

9 Chemical Modification	164
A Amino-groups	164
B Carboxy-groups	166
C Thiol Groups and Disulphide Bridges	167
D Tyrosine	168
E Tryptophan, Arginine, and Histidine	171
F Photo-oxidation	172
G Affinity Labelling	173
10 Conclusion	176
 <i>II: X-Ray Studies</i>	
<i>By T. L. Blundell</i>	
1 Introduction	176
2 Amino-acids and Peptides	178
3 Methods of Protein Structure Analysis	181
4 Globular Proteins	182
A General Structural Principles	182
B Proteases	184
C Nucleases	191
D Glycoside Hydrolases	193
E Carbonic Anhydrase	194
F Enzymes of the Glycolytic Pathway	195
G Dehydrogenases	198
H Regulatory Enzymes	201
I Redox Proteins	203
J Haemoglobins	207
K Concanavalin A	214
L Toxins	215
M Hormones	215
N Plasma Proteins	218
O Immunoglobulins	218
P Muscle Proteins	220
Q Viruses	221
5 Fibrous Proteins	223
A Keratins	223
B Silks	223
C Collagen	224
D Flagella	224

**III: Conformation and Interaction of Peptides and Proteins
in Solution***Edited by R. H. Pain*

1 Theoretical Aspects of Protein Structure	224
<i>Contributed by B. Robson</i>	
A The General Approach using Energy Functions	224
B Solvent Effects	226
C Studies on the Conformation of Single Residues using Energy Functions	226
D Predictions of Local Conformations in Polypeptides	227
E Predictions of the Stability of Globular Proteins	229
2 Spin Labels	230
<i>Contributed by N. C. Price</i>	
A Haemoglobin	230
B Enzyme Binding Sites	232
Creatine Kinase	232
Glyceraldehyde 3-Phosphate Dehydrogenase	232
Glycogen Phosphorylase <i>b</i>	233
Monomeric Enzymes	233
C New Spin Labels	235
3 Fluorescence	236
<i>Contributed by J. R. Brocklehurst</i>	
A Interpretation of Fluorescence	236
B Fluorescent Probes	237
Non-covalently bound	237
Covalently bound	238
C Protein Fluorescence	239
Binding of Small Molecules	240
Protein-Protein Interaction	241
Protein Structure	241
4 Mössbauer Spectroscopy	243
<i>Contributed by C. E. Johnson</i>	
A Haem Proteins	244
B Iron-Sulphur Proteins	244
C Other Proteins	245
5 Nuclear Magnetic Resonance	246
<i>Contributed by H. W. E. Rattle</i>	
A Peptides and Polypeptides	246
B Proteins	247

6 Infrared Spectroscopy	250
<i>Contributed by R. M. Stephens</i>	
A Model Compounds, Amino-acids, and Oligopeptides	250
B Synthetic Polypeptides	251
C Solvation and Structure	252
 7 Circular Dichroism and Optical Rotatory Dispersion	 253
<i>Contributed by P. M. Bayley</i>	
A General	254
Reviews	254
Theory	254
Analysis	256
Instrumental	257
B Small Molecules, Model Compounds, and Synthetic	
Polymers	258
Amino-acids and Derivatives	258
Dipeptides and Oligopeptides	260
Polypeptides	261
C Denaturation	263
Detergents and Neutral Salts	263
Unfolding and Refolding	265
D Proteins	265
Aromatic and Disulphide Chromophores	265
Non-chromophoric Proteins	269
Chromophoric Proteins	271
NAD(P)-dehydrogenases	271
Flavoproteins	271
Pyridoxal enzymes	272
Haem-proteins	273
Metalloproteins	276
Visual pigments	277
Added Extrinsic Chromophores	277
Naphthalenesulphonate derivatives	277
Purine and pyrimidine nucleotides	278
Metals	279
Other extrinsic effects	279
Chemical Modification and Peptide Cleavage	280
E Nucleic Acid-Protein Complexes	281
Model Systems	281
Natural Systems	282
F Immunological Systems	283
Immunoglobulins	283
Antibody-Hapten Interactions	284
G Antibiotics and Hormones	284

H Membrane Systems	285
Artefacts	285
Membrane Proteins	286
Phospholipid Micelles	286
8 Dissociation and Association of Proteins	287
<i>Contributed by G. L. Kellett</i>	
A Analytical Ultracentrifugation Techniques	287
Difference Sedimentation	288
Terminology for Sedimentation and Gel-filtration	
Experiments	289
Active-enzyme Sedimentation	291
B Gel Chromatography	292
Difference Chromatography	293
C Light Scattering	294
D Transport Studies	294
E Kinetic Studies	295
Ligand-induced Association-Dissociation	296
F Subunit Structure of Proteins	297
SDS Gel Electrophoresis	299
Measurements in Guanidine Hydrochloride	300
Hybridization Techniques	301
Isozymes	301
Haemoglobins	302
Bacterial luciferase	302
Cold Inactivation	303
G Protein-Small Molecule Equilibria	303
H Examples of Association-Dissociation Equilibria	304
Enzymes	304
Haemoglobin	305
Cyclic-AMP-dependent Protein Kinases	307
I Comment	308

Chapter 3 Peptide Synthesis

By J. H. Jones and B. Ridge

1 Introduction	309
2 Methods	310
A Protective Groups	310
Established Methods of Amino-group Protection	310
New Methods of Amino-group Protection	317
Protection of Carboxy-groups	324
Protection of Hydroxy-groups	328

Protection of Thiol Groups and Synthesis of	
Cystine Peptides	328
Protection of Histidine Side-chains	334
Final Deprotection	338
Miscellaneous Matters relating to Protective Groups	339
B Formation of the Peptide Bond	340
Activated Esters	341
Coupling Methods involving Acyloxyphosphonium Salts	343
<i>N</i> -Carboxy-anhydrides and <i>N</i> -Thiocarboxy-anhydrides	349
Other Methods	353
C Racemization	354
D Repetitive Methods of Peptide Synthesis	355
E Synthesis of Polymeric Models for Studies in Protein Chemistry	359
Polyamino-acids	359
Sequential Polypeptides	360
F Synthetic Operations with Peptides of Biological Origin	361
With Naked Natural Peptides	361
With Partially Blocked Peptides	362
3 Syntheses Achieved and Structure-Activity Correlations	363
A Calcitonin	365
B Hypothalamic Releasing Factors	366
Growth-hormone Releasing Factor	366
Luteinizing-hormone Releasing Factor and Follicle-stimulating Hormone Releasing Factor	367
Melanocyte-stimulating Hormone Release	370
Thyroid-stimulating Hormone Releasing Factor	371
C Oxytocin	372
D Ribonuclease T ₁	373
E Solid-phase Synthesis of High Molecular Weight Polypeptides	375
Acyl Carrier Protein	376
Basic Pancreatic Trypsin Inhibitor	377
Human Growth Hormone	377
Lysozyme	378
Parathyroid Hormone	379
Ribonuclease A	379
Staphylococcal Nuclease	381
F Scotophobin	382
G Substance P	383

Contents	xiii
4 Appendix A. A List of Syntheses Reported during 1971	383
A Naturally Occurring Peptides, Proteins, Analogues, and Partial Sequences	383
B Sequential Polypeptides	390
C Miscellaneous Peptides	391
5 Appendix B. A List of Some Useful New Synthetic Inter- mediates Described during 1971	393
 Chapter 4 Peptides with Structural Features Not Typical of Proteins	
<i>By J. S. Davies</i>	
1 Introduction	399
2 Cyclic Peptides	400
A 2,5-Dioxopiperazines	403
B Gramicidins	407
C Tyrocidins	410
D Alamethicin	410
E Peptides from <i>Amanita phalloides</i>	410
F Viomycin, Capreomycin, and Tuberactinomycin	412
G Nisin	413
H Other Cyclic Peptides	414
I Synthesis of Homodetic Cyclic Peptides	415
3 Depsipeptides (Heterodetic Peptides)	418
A Actinomycins	418
B Valinomycin	421
C Other Naturally Occurring Depsipeptides	422
D Miscellaneous	424
4 Peptide-Carbohydrate Linkages	425
A Glycopeptides from Bacterial Cell Walls	425
B Glycopeptides from Miscellaneous Sources	426
C Studies on Model Glycopeptide Linkages	428
D Other Carbohydrate-linked Compounds	428
5 Peptides and Amino-acids Linked to Nucleosides and Nucleotides	429
6 Peptide Alkaloids	431
7 Hydrazino Peptides	432
8 Penicillins and Cephalosporins	432
9 Miscellaneous	438

Chapter 5 Further Extracts from the Rules and Tentative Rules of the I.U.P.A.C.—I.U.B. Commission on Biochemical Nomenclature

<i>I Symbols for Amino-Acid Derivatives and Peptides — Recommendations (1971)</i>	441
1 General Considerations	442
2 Symbols for Amino-Acids	443
3 Amino-Acid Residues	445
4 Substituted Amino-Acids	448
5 Symbols for Substituents	450
6 Polypeptides	452
 <i>II Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains</i>	455
Preamble	455
Rule 1. General Principles of Notation	455
Rule 2. The Sequence Rule, and Choice of Torsion Angle	458
Rule 3. The Main Chain (or Polypeptide Backbone)	461
Rule 4. Side Chains	465
Rule 5. Hydrogen Bonds	469
Rule 6. Helical Segments	469
 <i>Appendix</i>	470
Recommendation A. Conformation and Configuration	470
Recommendation B. Definitions of Primary, Secondary, Tertiary, and Quaternary Structure	470
 Author Index	472

The past twelve months have witnessed interesting developments in amino-acid chemistry, but pride of place must be taken by the isolation, characterization, and synthesis of the amino-acid derivative from phenylalanine t-RNA. This work also reflects the growing sophistication in the application of physical methods in general. The interesting advances reported last year on asymmetric synthesis have been extended, and a staggering number of new amino-acids continues to be synthesized for a variety of reasons. The established pattern of coverage for this chapter is maintained, attention being focused on a broad selection of topics with the significant developments highlighted where necessary.

1 Naturally Occurring Amino-acids

A. Occurrence of Known Amino-acids.—A large number of papers which are concerned with the free amino-acid content of a wide variety of living organisms is published annually. Since the emphasis for the majority of this work is on biological aspects, it is not considered appropriate to cover them in this section and only those amino-acids which are rarely encountered or are interesting from the chemical viewpoint have been included.

The presence of amino-acids in extra-terrestrial material as well as their synthesis under simulated prebiotic conditions (see Section 2B) has attracted a considerable amount of attention. Examination of samples from the Apollo 11 and Apollo 12 missions has revealed extremely low concentrations of amino-acids (*ca.* 20—70 p.p.b.);¹⁻³ ultra-sensitive analytical techniques were employed for these investigations and, because of the presence of a considerable number of non-protein amino-acids, the investigators maintain that they are not due to terrestrial contamination. Similar conclusions have been made concerning the presence of amino-acids in the Murray⁴ and

¹ B. Nagy, J. E. Modzeleski, V. E. Modzeleski, M. A. J. Mohammad, L. A. Nagy, W. M. Scott, C. M. Drew, J. E. Thomas, R. Ward, P. B. Hamilton, and H. C. Urey, *Nature*, 1971, 232, 94.

² K. Harada, P. E. Hara, C. R. Windsor, and S. W. Fox, *Science*, 1971, 173, 433.

³ C. W. Gehrke, R. W. Zumwalt, D. L. Stalling, D. Roach, W. A. Aue, C. Ponnampereuma, and K. A. Kvenvolden, *J. Chromatog.*, 1971, 59, 305.

⁴ J. G. Lawless, K. A. Kvenvolden, E. Peterson, C. Ponnampereuma, and C. Moore, *Science*, 1971, 173, 626.

Murchinson⁵ meteorites. In the case of the Murray meteorite, seventeen amino-acids were identified of which seven were conclusively shown to be racemic and eleven to be non-protein in origin. An earlier report⁶ that the amino-acid content of the Orgueil meteorite was due solely to terrestrial contamination has been questioned, and it is now suggested that there are amino-acids indigenous to the meteorite in addition to those present as contaminants.⁷

The presence of amino-acids in the North Atlantic ocean has been the subject of a detailed examination and the distribution appears to be non-uniform and varies qualitatively with depth.⁸ In marine sediments the degree of racemization of free amino-acids shows a progressive increase with the age of the sediment,^{9,10} and the ratio of *allo*-isoleucine to leucine is a reliable indicator of age for samples less than 400 000 years old. In samples older than about 15×10^6 years the amino-acids are completely racemic.

The stereochemistry of the α -hydroxy-analogue of cysteine present in the urine of certain mentally retarded patients has been established by comparison with synthetic material.¹¹ The work reported initially last year on the isolation of the methylated derivatives of arginine and lysine has been extended and it has been noted that, in patients with malignant tumours, the relative urinary level of *guanidino-NN*-dimethylarginine to that of arginine is markedly increased.¹² Bovine and rat brain tissue has been shown to contain appreciable amounts of N^G -monomethylarginine as well as N^GN^G -dimethylarginine.¹³

Further spectral evidence on the aldol condensation product isolated on alkaline hydrolysis of elastin provides strong support¹⁴ for the previously assigned structure (1).

Interest in plants containing L-Dopa continues, and a widespread investigation has shown that several *Mucuna* species contain up to 5% of this important amino-acid.¹⁵ The neurotoxin α -amino- β -oxalylaminopropionic acid has been isolated from *Lathyrus sativus*¹⁶ and the major alkaloid of *Aotus subglauca* has been identified as *S*-(+)- N^{α} -methyltryptophan methyl ester.¹⁷ The previously unidentified amino-acid from *Peganum harmala* is

⁵ J. R. Cronin and C. B. Moore, *Science*, 1971, 172, 1327.

⁶ J. Oro, S. Nakaparksin, H. Lichenstein, and E. Gil-Av, *Nature*, 1971, 230, 107.

⁷ J. G. Lawless, K. A. Kvenvolden, E. Peterson, and C. Ponnampereuma, *Nature*, 1972, 234, 66.

⁸ R. Pocklington, *Nature*, 1971, 230, 374.

⁹ R. O. Brinkhurst, K. E. Chua, and E. Batoosingh, *Limnol. Oceanogr.*, 1971, 16, 555.

¹⁰ J. Wehmiller and P. E. Hare, *Science*, 1971, 173, 907.

¹¹ M. Wälti and D. B. Hope, *J. Chem. Soc. (C)*, 1971, 2326.

¹² S. Akazawa, *Osaka Daigaku Igaku Zasshi*, 1970, 22, 461.

¹³ T. Nakajima, Y. Matsuoka, and Y. Kakimoto, *Biochim. Biophys. Acta*, 1971, 230, 212.

¹⁴ G. Crombie, B. Faris, P. M. Gallop, and C. Franzblau, *Biochemistry*, 1971, 10, 4145.

¹⁵ E. A. Bell, J. R. Nulu, and C. Cone, *Phytochemistry*, 1971, 10, 2191.

¹⁶ K. Bahadar and S. P. Billa, *Indian J. Appl. Chem.*, 1970, 33, 168.

¹⁷ S. R. Johns, J. A. Lamberton, and A. A. Sioumis, *Austral. J. Chem.*, 1971, 24, 439.