

Chemistry and Biochemistry of the Amino Acids

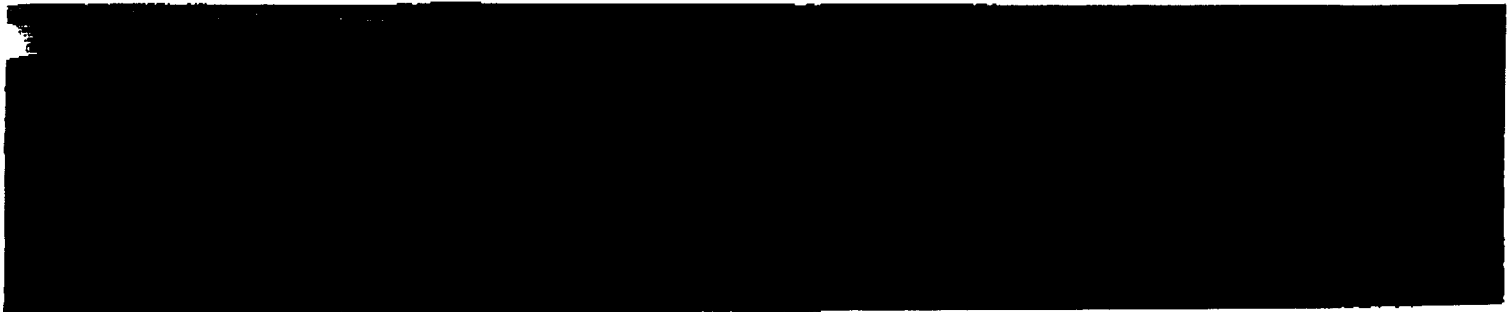
EDITED BY

G.C. Barrett

Oxford Polytechnic, UK

LONDON NEW YORK

Chapman and Hall



First published 1985 by
Chapman and Hall Ltd
11 New Fetter Lane, London EC4P 4EE

Reprinted 1985

Published in the USA by
Chapman and Hall
733 Third Avenue, New York NY 10017

© 1985 Chapman and Hall

Printed in Great Britain
by J.W. Arrowsmith Ltd, Bristol

ISBN 0 412 23410 6

All rights reserved. No part of this book may be reprinted, or reproduced or utilized in any form or by any electronic, mechanical or other means, now known or hereafter invented, including photocopying and recording, or in any information storage and retrieval system, without permission in writing from the publisher.

British Library Cataloguing in Publication Data

Barrett, G.C.
Chemistry and biochemistry of the amino acids.
1. Amino acids
I. Title
547.7'5 QD431

ISBN 0-412-23410-6

Library of Congress Cataloging in Publication Data

Main entry under title:
Chemistry and biochemistry of the amino acids.
Bibliography: p.
Includes index.
1. Amino acids. I. Barrett, G.C., 1935-
QD431.C45155 1985 574.19'245 83-15109
ISBN 0-412-23410-6 (U.S.)

Contributors
Preface

- 1 Nomenclature of Amino Acids
G.C. Barrett
- 2 The Protein Amino Acids
P.M. Hardy
- 3 Beta and Higher Homologues of Amino Acids
C.N.C. Drey
- 4 The Non-Protein Amino Acids
S. Hunt
- 5 Metabolic and Pharmacological Aspects of Amino Acids
D.A. Bender
- 6 The Biosynthesis of Amino Acids
P.J. Lea, R.M. Wallsgrove
- 7 Enzyme Inhibition by Amino Acids
M.J. Jung
- 8 Synthesis of Amino Acids
G.C. Barrett
- 9 Protected Amino Acids
J. Meienhofer
- 10 Resolution of Amino Acids
G.C. Barrett
- 11 Reactions of Amino Acids
G.C. Barrett
- 12 Degradation of Amino Acids
Hydrolysis
S. Hunt
- 13 Racemization of Amino Acids
J.L. Bada

Amino Acids

75) *Ger. Offen.* 2416 048.
206.

u 23, 47; (*Chem. Abs.* 94, 66 040k).
35e).

Am. Chem. Soc. 103, 1857.

Vol. Chem. 180, 307.

Angew. Makromol. Chem. 84, 97.

Vol. Chem. Phys. 182, 9.

Zh. Prikl. Khim. 49, 181; (*Chem.*

ym. Chem. Ed. 17, 1911.

em. 179, 2707.

628, 162.

n. 180, 1419.

180, 1175.

Polym. Sci. 256, 645.

Sci. Polym. Chem. Ed. 16, 2237.

molecules 11, 619.

7.

rg. Chem. 42, 3731.

180) *Chem. Pharm. Bull.* 28, 1899.

t. Chem. 16, 157.

org. Khim. 7, 240.

t. 31, 1073.

Bioorg. Chem. 7, 111.

Experientia 11, 397.

. 687.

081.

c. Perkin Trans. 1, 378.

68.

1361.

Phys. 74, 993.

J. Adv. Mol. Relaxation Interact.

. Soc., Faraday Trans. 1; 73, 476.

; 1301.

Mol. Struct. 39, 175.

Chem. Commun. 338.

334, 2359.

Polym. Sci. 256, 645.

hem. 16, 115.

n. 15, 1.

on Lett. 1813.

rans. 1, 906.

tt. 2359.

431.

1979) *Biomed. Mass Spectrom.* 6,

CHAPTER FOUR

The Non-Protein Amino Acids

S. HUNT

4.1 INTRODUCTION

In 1961 Greenstein and Winitz [1] listed some ninety amino or imino acids of non-protein origin. Scratch a reasonably well-informed biologist today and he might tell you that there are probably about two hundred natural amino acids now known. In fact this would be a wild underestimate since the tables appended to this chapter list close upon seven hundred amino and imino acids or related derivatives. The author makes no claim to have been exhaustive (though exhausted) in his literature search and the pace of discovery of new compounds makes it highly likely that many will have been overlooked. Indeed there are now so many natural amino acids known that 'rediscovery' must be becoming a common event. A comprehensive catalogue of structures, sources, chemical and biological properties is badly needed but this review does not aspire to that function. This is however perhaps the point to apologize for missed molecules, erroneous structures and spurious compounds not yet purged from the literature.

4.2 SCOPE

Organisms are wonderfully diverse in their metabolisms and the range and variety of structures listed herein reflect this diversity.

Non-protein amino acids are those amino acids which are not found in protein main chains either for lack of a specific transfer RNA and codon triplet or because they do not arise from protein amino acids by post-translational modification. In this chapter we are looking at alpha amino acids, that is to say the carboxyl group is separated from the amino or imino nitrogen by one carbon only, the α -carbon atom. Higher homologues are dealt with in another chapter. Built upon this basic structure a vast diversity of organic functional groupings can give rise to a numerous body of molecules which to a greater or lesser extent qualify as amino acids.

Many of these compounds are the end products of secondary metabolism. Their origins are diverse as are their functions although in both respects we are still more ignorant than informed. Many others arise as intermediates upon metabolic pathways or originate from the metabolism or detoxification of foreign compounds. Because of the nature of bacterial metabolism it is possible to produce many new compounds by addition of appropriate organic nutrients to culture media. Such 'artificially' generated amino acids can be found in the literature but have not been collected here. Nor has the question of whether

amino-acids found in meteorites are of natural origin or not been pursued. Aside from these matters I have interpreted what an amino acid is quite liberally.

4.3 SOURCES OF INFORMATION

A lot of molecules are listed in the tables which follow. Given the space available the best one can hope to do for each compound or class of compounds is to provide an entry to the relevant literature.

Many of these amino acids have been grouped before in the tables of authors' reviews. It would not seem to be a terribly effective use of space to have relisted all the primary sources that others have already cited. Accordingly I have in many cases cited secondary sources when these are likely to be readily accessible to the majority of readers. I drew heavily for information upon early reviews and several more recent ones [1-12], upon the Chemical Society Specialist Periodical Reports of Peptides and Proteins [13-25] and upon some general compilations like the *CRC Handbook of Microbiology* [26]. Not unnaturally, *Chemical Abstracts* was one rather important access route. Even so, searching for amino acids has many of the characteristics of wet fly-fishing - chuck it and chance it. Two names deserve special mention in the non-protein amino acid world; E.A. Bell and L. Fowden have probably identified more new plant amino acids than any other pair of workers and have contributed more to the review literature of this subject than anyone else [4-6, 8-10].

4.4 CLASSIFICATION AND TABULATION

Taxonomies are always to some extent artificial. Listing and classifying amino acids is no exception.

In tabulating the amino acids the broad principle has been adopted that a rough division may be made with aliphatic and other amino acids. That the benzenoid aromatics form a distinct group and that heterocyclic amino acids may be crudely separated according to their heteroatoms seems obvious while the presence of sulphur in an amino acid again would seem to distinguish it as does the presence of the analogous selenium. Probably there are sound metabolic arguments to back up these divisions in many cases. It has been convenient to segregate phosphorylated amino acids even though the parent compounds may have been drawn from several different groups and this comment applies also to a broad grouping of α -*N*-substituted amino acids containing those which are *N*-methylated, *N*-acylated or more bizarrely substituted.

4.5 NOVEL AMINO ACIDS; THEIR IDENTIFICATION, ISOLATION, CHARACTERIZATION - TWO CASE HISTORIES

Novel amino acids in bulk tissues or physiological fluids are most frequently first recognized as components of unusual behaviour upon chromatograms. Atypical

not been pursued. Aside acid is quite liberally.

ON

iven the space available ass of compounds is to

in the tables of authors' space to have relisted all ording I have in many readily accessible to the arly reviews and several y Specialist Periodical ne general compilations unnaturally, *Chemical* so, searching for amino chuk it and chance it. amino acid world; E.A. plant amino acids than the review literature of

ATION

g and classifying amino

as been adopted that a amino acids. That the sterocyclic amino acids ns seems obvious while eem to distinguish it as ere are sound metabolic has been convenient to parent compounds may mment applies also to a ing those which are N-

NIFICATION,
ASE HISTORIES

are most frequently first romatograms. Atypical

Rf values in relation to known species or an unusual colour response to ninhydrin may prompt further investigation. New compounds, as components of biologically active molecules such as depsipeptides, will emerge as a logical consequence of structural elucidation as perhaps an extra peak on an amino acid analyser trace. Encounters may thus be planned or fortuitous (hopefully serendipitously). Investigation of a metabolic pathway may have as a basis of experimental rationale the predicted isolation of a hitherto uncharacterized intermediate.

By whatever route a new amino acid is first encountered or suspected, a process of isolation and purification in sufficient quantity will be a necessary preliminary to structural characterization, proof of structure via synthesis and establishment of the conformation at chiral centres in the natural product. Modern physical techniques such as mass spectrometry and nuclear magnetic resonance spectroscopy have rendered characterization relatively simple in the majority of cases while establishment of enantiomeric type has also been made easier by the development of gas-liquid chromatographic techniques capable of separating enantiomers on a systematic basis. The following two examples of characterization of novel amino acids will give an indication of the types of technique now routinely applied.

(a) Dicysteinyldopa [27]

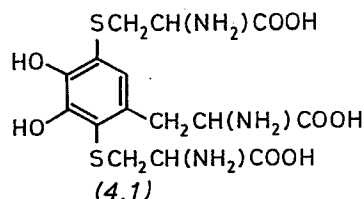
Behind the light-sensitive cells of the retina in the vertebrate eye is a layer called the *tapetum lucidum* whose chemistry is related to its function as a reflector of light. *Lepisosteus* the Alligator Gar (a confusingly named fish – not a reptile) has a pigmented tapetum. A desire to identify the major constituent of this yellow material led to the discovery of a novel sulphur-containing amino acid.

Molar formic acid extraction of dark-adapted tapetum gave an extract which after removal of acid *in vacuo* could be examined by two-dimensional thin layer chromatography. Under ultraviolet light the chromatograms showed at least seven major discrete components of which several were both ninhydrin and ferric chloride positive. The most intense of these spots was obtained in quantity from over a hundred grams of tissue by extraction with 1 N hydrochloric acid and absorption onto Dowex 50 ion-exchange resin. Washing the resin with a molar solution of hydrochloric acid and subsequent elution with 6 N acid gave a product which after removal of acid could be rechromatographed on Sephadex LH20 in methanol/hydrochloric acid. A major peak identified by monitoring at 305 nm was collected and further purified by preparative paper chromatography. A major ultraviolet quenching band was eluted from the paper, converted to its formate salt and rechromatographed on a column of Sephadex G25. A major peak absorbing at 305 nm was collected and freeze-dried ready for structural study.

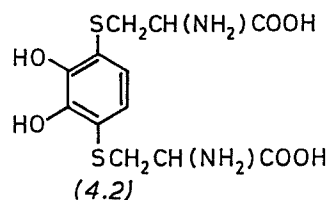
Several factors pointed to this compound being a sulphur-containing ortho-diphenolic amino acid. Reactivity to ferric chloride and ninhydrin, an elemental analysis of $C_{15}H_{21}N_3O_8S_2-HCOOH$, infrared absorption bands at 3600–2200, 1630, 1500 and 1390 cm^{-1} , and ultraviolet absorption spectra with maxima at

303-nm in acid shifting to 316 nm at pH 7 contributed to this view. Proton nuclear magnetic resonance spectra revealed a multiplet for six methylene protons at δ 3.2–3.8, a multiplet for three methine protons at δ 4.3–4.5 and an isolated aromatic proton singlet at δ 7.20.

Reductive hydrolysis in hydriodic acid/red phosphorus yielded cysteine and dihydroxyphenylalanine (dopa) as major products leading to the formulation of Structure (4.1) or *S, S*-dicysteinyl-dopa composed of one molecule of dopa linked to two molecules of cysteine via thio-ether linkages.



This structure was partly confirmed by biological synthesis. Tyrosinase oxidation of L-dopa in the presence of excess L-cysteine yielded the same amino acid together with 5- and 3-*S*-cysteinyl-dopa thus establishing positions 2 and 5 in the aromatic nucleus as being the most likely for substitution. Weight was added to this view by the observation that under the same conditions catechol and cysteine yield 3,5-cysteinylcatechol and 3,6-*S, S*-dicysteinylcatechol (4.2) the latter being a symmetrical structure like the putative Structure (4.1).



The absolute configuration (*2S, 2'S, 2''S*) was inferred from the good agreement of the specific rotation of the natural sample with that of the synthetic one prepared from L-dopa and L-cysteine.

(b) 2,4-Diamino-3-methylbutanoic acid [28]

The amino acid composition of root nodules in *Lotus* species is determined by the *Rhizobium* strain rather than the host plant. Common amino acids are present but ninhydrin-positive compounds with unusual *R_f* values are also evident.

Ethanol (80%) extracts of *Lotus* root nodules were hydrolysed with acid (6 N HCl), chromatographed on Amberlite IR120 ion exchange resin and the amino acid fraction converted to *N*-trifluoroacetyl-*n*-butyl esters (TAB). Gas chromatographic/electron impact mass spectrometry analysis revealed a major component with an aspartate-related retention time of 0.86. Comparison with standard TAB amino acids suggested an unusual amino acid. High resolution mass spectrom-

this view. Proton nuclear methylene protons at δ 4.3–4.5 and an isolated

rus yielded cysteine and ng to the formulation of molecule of dopa linked

l synthesis. Tyrosinase yielded the same amino shing positions 2 and 5 in ution. Weight was added conditions catechol and steinylcatechol (4.2) the Structure (4.1).

m the good agreement of ie synthetic one prepared

ecies is determined by the amino acids are present alues are also evident. ydrolysed with acid (6 nge resin and the amino rs (TAB). Gas chromatoealed a major component rison with standard TAB solution mass spectrom-

etry gave an empirical formula for the derivative of $C_{13}H_{18}O_4F_6$ isomeric with TAB-ornithine and the molecular weight of 380 was confirmed by gas chromatographic/chemical ionization mass spectrometry. Diagnostic ions in the EIMS at m/z 306, 279 and 166 were characteristic of TAB derivatives of basic aliphatic amino acids although there were insufficient mass spectral data to assign an unequivocal structure. The evidence however suggested an ornithine isomer and in consequence the four possible isomers were synthesized and the mass spectra compared. Such comparisons identified the unknown compound as 2,4-diamino-3-methylbutanoic acid whose configuration was established as being 2*R*,3*S* on the basis of the behaviour of the TAB amino acid, in comparison with known compounds, on the gas chromatographic stationary phase Chirasil-Val.

These two examples show slightly differing approaches to the structure determination problem. In one, nuclear magnetic resonance spectroscopy and wet chemical methods are important, in the other, the mass spectrometer provides much of the more significant information.

4.6 BIOLOGICAL SOURCES AND THE RANGE AND DIVERSITY OF NON-PROTEIN AMINO ACIDS

The compounds listed in this chapter are drawn from most groups of living organisms. What constitutes a 'non-protein' amino acid has been, as said above, interpreted quite liberally within the mentioned chemical frames of reference.

The major body of the known non-protein amino acids have as their source the plant world and micro-organisms, the latter secreting a variety of compounds into culture media [26]. Many microbiological products have antibiotic properties and many of these as well as some similar fungal products comprise unusual amino acids incorporated into more complex structures such as depsipeptides [29]. Here both D and L amino acids of common and novel character are linked by peptide bonds to each other and by other linkages to units such as hydroxy acids. Such molecules have proved a rich source of new amino acids.

In the higher plants the unusual amino acids are most frequently present in the free state or as low-molecular weight complexes in isopeptide linkage to glutamic acid. Concentrations of such compounds in plant tissues can be very high indeed but trace amounts of a great variety of compounds may also be present in a single species. A number of unusual plant and animal amino acids constitute parts of pigment structure [30–31].

On the face of things, when one looks at distributions of amino acids, micro-organisms and plants seem to be chemically rather different from animals. It is not that their basic metabolisms differ, for these are broadly similar, but we seem not to find in animals the range of secondary substances like non-protein amino acids, alkaloids, polyketides and phenolics so characteristic of plants and lower organisms. Thus the list of non-protein amino acids distributed in animal tissues and physiological fluids is more limited and less exciting. Even so this may be due

to the fact that most work has concentrated upon mammals while the more diverse invertebrates have been less well investigated.

A high proportion of plant non-protein amino acids have aliphatic structures. Chains do not often exceed six carbon atoms in length although there are some quite large molecules. Diversity is attained by limited branching, hydroxyl, carboxyl and amino substitution and by the inclusion of unsaturation as allenic or allynic groupings. Other types of nitrogenous group include guanidino and cyano functions while there are numerous sulphur or selenium analogues of cysteine, cystine and methionine.

Benzenoid compounds are quite common in both plants and animals and it is therefore surprising to find that the range of aromatic amino acids based upon the benzene ring is relatively limited in comparison with the aliphatics. In view also of the diversity of halogeno aromatic compounds found in marine organisms [32] and the relative ease of substitution of the phenol ring by oxidized halide ions it is surprising that so few free halogenated phenolic amino acids are found. In fact, halogeno amino acids of any type are relatively uncommon outside the group of chlorinated bacterial products, even among the marine algae and marine invertebrates which are known to produce halogenated proteins and other secondary products [33].

In contrast to benzenoid cyclic amino acids, heterocyclics both aromatic and otherwise form a large body of non-protein amino acids. It is perhaps only to be expected that a great proportion of these will be nitrogen heterocycles although there are significant numbers which have oxygen or sulphur in the ring. A number of these nitrogen heterocycles are derived from and based upon the tryptophan indole nucleus. Other nitrogen heterocycles are closely related to proline or are homologues of this amino acid. Once again plant and bacterial sources dominate the heterocyclic scene although some interesting oxygen-nitrogen and nitrogen-sulphur heterocycles are components of animal pigments.

While there are numbers of heterocyclic imino acids, an interesting recent development has been an expansion in the limited range of aliphatic imino acids with pyruvate and amino acid-derived products such as strombine and alanopine being added to octopine as components of anoxic tissue metabolism in invertebrates [34, 35]. The related lysopine and saccharopine are plant metabolites [36-37].

4.7 BIOSYNTHETIC ORIGINS OF NON-PROTEIN AMINO ACIDS

With such diversity of structure it is not surprising that no one clear pattern of biosynthetic origin can be put forward for the non-protein amino acids. Setting aside for the moment environmental and other pressures which have selected for the maintenance of pathways leading to the accumulation of particular components it seems likely that the non-protein amino acids may have arisen in perhaps three or four general ways.

A probable route for many familiar looking products is that of modification of

minals while the more

ve aliphatic structures. Although there are some branching, hydroxyl, unsaturation as allenic include guanidino and selenium analogues of

ts and animals and it is no acids based upon the liphatics. In view also of marine organisms [32] oxidized halide ions it is acids are found. In fact, on outside the group of ine algae and marine ed proteins and other

olics both aromatic and It is perhaps only to be 1 heterocycles although ur in the ring. A number d upon the tryptophan elated to proline or are terial sources dominate -nitrogen and nitrogen- ts.

s, an interesting recent of aliphatic imino acids trombine and alanopine tissue metabolism in ropine are plant meta-

3IN AMINO ACIDS

no one clear pattern of ein amino acids. Setting which have selected for nulation of particular acids may have arisen in s that of modification of

existing amino acids i.e. by mechanisms similar to those involved in post-translational modification of protein amino acids. That is by derivatization or simple change in structure of an existing site on a common amino acid. Dihydroxyphenylalanine [38] production from tyrosine, δ -acetylornithine [39] or *O*-acetylserine [5] synthesis in plants are well-documented examples of such modifications. The appearance of hydroxyproline [40], desmosine and isodesmosine [41] or pyridinoline [42] in mammalian urine are examples of novel amino acids which arise as post-translational modifications of amino acids actually on protein chains but which are then released by metabolic turnover to achieve free status.

Thus we can have potentially many rather simple analogues of well-known amino acids through simple post-synthetic modification of the twenty or so protein amino acids in a variety of ways. In practice certain amino acids participate in this type of modificational activity more than others and while we may encounter many products derived from lysine there may be few originating in valine. Clearly reactivities of side-chains are great determinants. The reactive thiol group of cysteine results in a relatively large range of secondary amino acids (Tables 4.13-4.15).

The formation of the α -amino- α -carboxyl grouping in the alpha amino acids takes place at initial stages in the synthesis of the amino acids by amination of keto acid. If generation of the side-chain function is one involving several enzymically catalysed steps then it is easy to see how modifications to intermediate stages can result in formation of one or even more novel amino acids. Probably many non-protein amino acids originate in this manner. At its very simplest we may see the actual intermediates in the formation of a particular amino acid as amino acids in their own right. In the absence of a metabolic role their existence may be transitory. Examples might be phosphoserine on the route to serine from 3-phosphoglycerate or glutamic- γ -semialdehyde lying between glutamate and proline [43]. Others of this type however may accumulate as does homoserine in the synthetic path to methionine or threonine from aspartate or ornithine on the route to arginine [43]. The latter synthesis has again as a precursor the amino acid glutamyl- γ -semialdehyde while the related aspartate derivative gives rise to homoserine originating in its turn from aspartylphosphate [43]. If we include all derivative intermediates of this type the list of non-protein amino acids becomes quite legally very much greater.

Failure of enzyme production for a particular step in a metabolic path may result in abnormal accumulation of amino acid intermediates. Accumulation of homocysteine in the blood and urine of homocystinurics together with other novel sulphur amino acids points to abnormalities of methionine metabolism between dietary methionine and inorganic sulphate. This can be looked upon basically as a failure to condense homocysteine with serine to form cystathione; a reaction catalysed by cystathione synthase [44,45]. Pyroglutamic acid (5-oxoproline) may appear in abnormal urines as a result of a deficiency in 5-oxoprolinase, a component of the γ -glutamyl cycle [44].

An extension of this particular concept lies with some change in the pathway leading to a common amino acid which gives rise to a final product which is a new amino acid and which is evolutionarily selected. The best-documented example of this class of origin is that of the shikimate to chorismate route. The common amino acids phenylalanine and tyrosine are generated from shikimate (4.3) via chorismate (4.4) and prephenic acid (4.5). The plant amino acids β -(3-carboxyphenyl)alanine (4.6) and β -(3-carboxy-4-hydroxyphenyl)alanine (4.7)

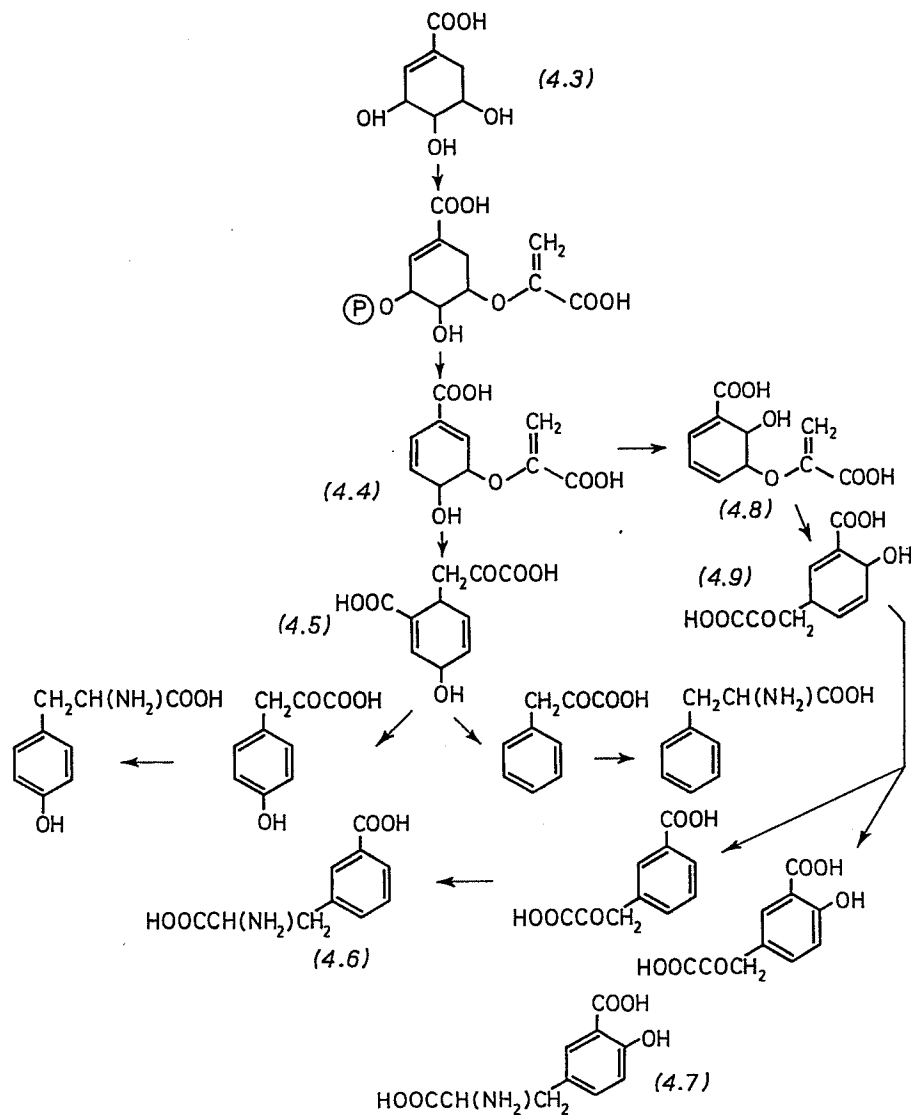
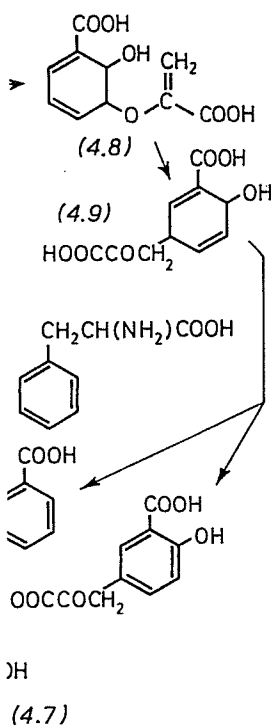


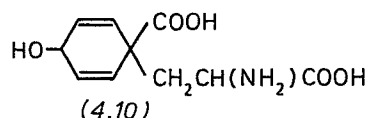
Figure 4.1

change in the pathway product which is a new -documented example te route. The common om shikimate (4.3) via it amino acids β -(3-xyphenyl)alanine (4.7)



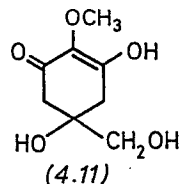
share the same route as far as chorismate but divert from it before prephenate with a conversion of chorismate to isochorismate (4.8) and hence to isoprephenate (4.9) from which the aromatic ring complete with carboxyl group is generated. The modification here may come down merely to removal of a decarboxylation stage [45,47].

Chorismate provides a divergence point for the origin of other novel aromatic amino acids; for example *p*-aminophenylalanine in *Vigna* arises by amination of chorismate, the path thereafter returning to that for the production of phenylalanine [48]. An interesting variant on the theme is the alternative route to tyrosine from chorismate in *Pseudomonas* which takes place via the novel compound pretyrosine or aroenate (4.10) [49].



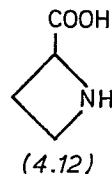
While many novel amino acids undoubtedly do originate from the 'tuning' of metabolic pathways and metabolites as described above it seems equally likely that such routes cannot wholly account for many of the more unusual compounds. Thus there may be a number of biosynthetic routes which like their products are themselves new and atypical. Even some non-protein amino acids whose structures seem to be close analogues of protein amino acids are products of convergence rather than parallel development. *L*- β -Methylaspartic acid of *Clostridium tetanomorphum* arises not from aspartate derivatization but by a novel rearrangement of *L*-glutamate involving 5'-deoxyadenosylcobalamin in an aerobic fermentative reaction [43]. A similar methyl derivative *erythro*-methyl- γ -glutamic acid originates in *Gleditsia triacanthus* from *L*-leucine by methyl group oxidation rather than from glutamic acid itself [50].

At the present time detailed accounts of biosynthesis exist for only a few groups of novel amino acids. For the major body one can only speculate informedly upon the basis of experience of routes leading to protein amino acids and related compounds. The likelihood is that even for many quite exotic structures relationships with known pathways will emerge. For example the mycosporin-like mytilins and related amino acids probably originate from condensation of known amino acids, glycine, serine and threonine with a parent diketone (4.11) originating from the shikimate pathway [51]. Or again strombine and alanopine derive from a NADH-coupled dehydrogenase-mediated reductive condensation



of pyruvate with glycine or alanine [34]; a reaction which has features in common with the formation of glutamate from α -ketoglutarate and ammonium ion by glutamate dehydrogenase and NADH.

Unusual amino acids which have been suggested to arise via quite new routes include the heterocycle azetidine-2-carboxylic acid (4.12) for which a putative synthetic chain from homoserine through 2,4-diaminobutanoate to 4-amino-butanoate and hence by ring closure and dehydration to the azetidine derivative has been proposed. But novel as this may seem, the suggestion has also been made that the synthesis is one made in error by the enzymes of proline synthesis (an amino acid undergoing cyclization before final reduction) [52, 53].



4.8 BIOLOGICAL FUNCTIONS

Novel amino acids may have many roles *in vivo* yet it is a fact that at the present time there is no function obvious for most of these compounds. An important point to bear in mind particularly in relation to plant amino acids is that while these substances often have physiological activity in animals these may be secondary actions rather than prime functions. In some cases there may be no function. As was indicated above some may originate through errors of metabolism, their presence merely illustrating and illuminating an underlying genetic abnormality; they may even be themselves the instruments of further internally directed deleterious action. Such compounds are best known for mammalian species because our knowledge of metabolic error is sparse for invertebrates or even for domestic plant species.

Function is related to the puzzling feature of non-protein amino acid distribution already mentioned in which there is an apparent disparity of occurrence and productivity in animals and plants or micro-organisms. One may speculate that this reflects a tendency in animal evolution towards increasingly efficient excretory systems while in plants the end products of metabolism are not voided but stored and exposed to a broader spectrum of enzymic secondary change. Certainly in lower, less highly evolved animals there seems to be greater versatility of low molecular weight chemical production. In this respect it is a pity that there is such a paucity of information regarding the products of nitrogen metabolism in the protozoa, the amino acid list which follows having few references relating to that group. Perhaps also the evolution of a nervous system and a physiology under delicate hormonal control has militated against the accumulation of numerous active small molecules.

An important aspect of plant metabolism is the need to conserve nitrogen and

which has features in arate and ammonium

e via quite new routes for which a putative utanoate to 4-amino-azetidine derivative on has also been made of proline synthesis (an) [52, 53].

fact that at the present ounds. An important ino acids is that while nimals these may be ases there may be no te through errors of inating an underlying nstruments of further ; are best known for ic error is sparse for

i-protein amino acid apparent disparity of o-organisms. One may i towards increasingly of metabolism are not of enzymic secondary re seems to be greater i this respect it is a pity ; products of nitrogen h follows having few n of a nervous system militated against the conserve nitrogen and

hence to limit nitrogen loss by production of secreted nitrogen compounds. Animals need not practice such economy and hence may produce and excrete nitrogenous compounds. It is therefore no accident that the excretion products of animals are one of the major sources of novel amino acids in that section of the living world.

Nitrogen required for protein synthesis in plants may be stored in large quantities as alanine, asparagine, arginine, acetylornithine, allantoin, citrulline and glutamine and in these cases it is easy to see the route by which entry to the protein synthetic path may be made. But in many cases of high levels of novel amino acid accumulation the route is less easy to see and perhaps these compounds are not intended as nitrogen stores or are unable to participate. Perhaps they may even accumulate for this very reason, having arisen through some change in a route to storage products, they have now replaced these products without a mechanism having yet evolved for their utilization.

4.9 TOXICITY

An important property of many of these products is their toxicity or ability to deleteriously affect metabolism in other organisms.

Many novel plant amino acids are structurally very similar to protein amino acids. In this light one cannot fail to be impressed by the accuracy of a translational system which has such compounds in relative abundance available and yet makes no mistakes. This is particularly so when it is remembered that the genetic code could have had the potential for incorporation of far more than twenty two amino acids into proteins. Yet there is no evidence to suggest that any of these unusual compounds find their way into the protein of their parent organism. Such incorporations are however not wholly impossible. Supply of a non-protein amino acid to a species not normally producing the compound can result in translation into a protein if it is a close enough analogue of a regular protein amino acid. A classic example of this is the incorporation of azetidine-2-carboxylic acid by *Phaseolus aureus* tissues as a replacement for proline residues in proteins [4]. *Convallaria majallis* unlike *Phaseolus* is a source of the amino acid and cannot incorporate it presumably because the enzymes of proline tRNA synthesis have extra specificity.

Incorporation of an abnormal amino acid into the proteins of a species to which the compound is foreign has obvious potential consequences for the properties of those proteins. Here there is a toxic effect which at first sight may seem to be unlikely to be exerted. Yet plants not uncommonly exude low molecular weight metabolites from roots or leaf glands or such products may slowly enter the soil from decomposing litter. Compounds from one species may then enter another via the root system. Plants may thus inhibit the development of other species in their vicinity; a well-documented phenomenon [54, 55]. Competition for nutrients and light may be weighted in favour of a species by such mechanisms operating.

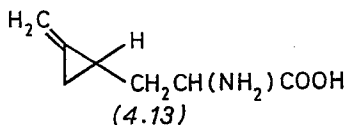
Just as plants may obtain advantage over other plants by exertion of toxic effects their opportunities for survival, range expansion and success may also be affected by toxic influences upon animals. Such influences may be applied to the general metabolism or at specific physiological loci.

The leguminosae are a rich hunting ground for unusual amino acids. The seeds of this group which include peas and beans have frequent importance in the diets of a wide range of world peoples and it is therefore of some consequence that many species accumulate quite seriously toxic amino acids in their seed in appreciable quantity. This fact is made all the more important because the group is particularly useful not only as a staple food producer but also as a source of animal fodder and a soil enrichment crop via the nitrogen fixing bacteria of its root nodules.

Lathyrism is a consequence of certain types of leguminosid toxicity. Seeds of the sweet pea (*Lathyrus odoratus*) are a frequent dietary cause of connective tissue malformation in domestic animals and sometimes humans. Weakened connective tissue may even result in death from aortic rupture. β -Aminopropionitrile present in the seeds as the γ -glutamyl derivative acts as an inhibitor of lysyl oxidase on enzymes required for cross-link formation in connective tissue structural protein. The rubber-like properties of elastin and the inextensibility of collagen both depend upon such lysyl-derived cross-links. β -Aminopropionitrile originates by decarboxylation of β -cyanoalanine in the plant.

β -Cyanoalanine and its γ -glutamyl derivative are reported as neurotoxins present in another legume *Vicia sativa* and in some other *Vicia* species. Here a lesion is manifested as a neurological disorder [57] Cystathionine excretion at unusually high levels in the urine is another manifestation of this condition suggesting an interference with homocysteine production. While the mode of action of β -cyanoalanine and β -aminopropionitrile in these diseases is not wholly clear it should be noted that a link with pyridoxine and pyridoxal phosphate may be a partial basis of the toxicity since pyridoxine deficiency diminishes amine oxidase activity and produces defective elastin *in vivo* while pyridoxal phosphate is not only necessary for the conversion of cystathionine to methionine but also eliminates the neurotoxic effects of β -cyanoalanine [43-44].

Many other toxic effects in humans may be traced to plant amino acids. Hypoglycins A and B (4.13 and 4.14) are responsible for the hypoglycaemia which follow eating unripe Akee fruit (*Blihia sapida*) [58, 59]. Mimosine (4.15) causes loss of hair and wool in cattle and sheep grazing on *Leucaena leucocephala* [60] (a pity an amino acid with antidepilatory properties cannot be found) while the numerous selenium analogues of sulphur amino acids found in a variety of plants have wide ranging toxic effects upon grazing animals.



ino Acids

nts by exertion of toxic
and success may also be
es may be applied to the

l amino acids. The seeds
t importance in the diets
some consequence that
) acids in their seed in
rtant because the group
r but also as a source of
gen fixing bacteria of its

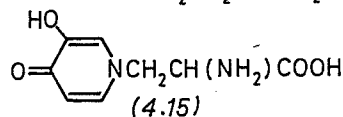
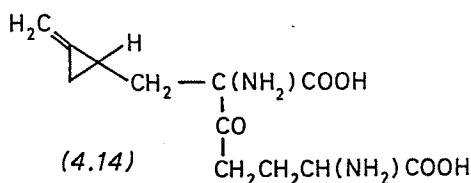
inosid toxicity. Seeds of
ause of connective tissue
mans. Weakened con-
e. β -Aminopropionitrile
as an inhibitor of lysyl
on in connective tissue
and the inextensibility of
s. β -Aminopropionitrile
e plant.

reported as neurotoxins
er *Vicia* species. Here a
stathionine excretion at
tation of this condition
ion. While the mode of
ese diseases is not wholly
pyridoxal phosphate may
iciency diminishes amine
hile pyridoxal phosphate
e to methionine but also
3-44].

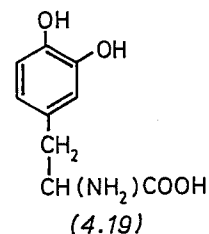
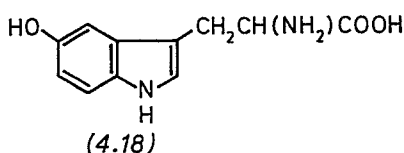
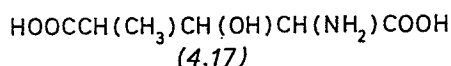
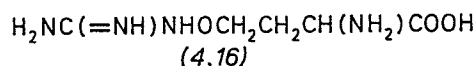
l to plant amino acids.
he hypoglycaemia which
. Mimosine (4.15) causes
aena leucocephala [60] (a
not be found) while the
und in a variety of plants

The Non-Protein Amino Acids

67

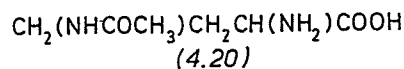


Toxic effects are also exerted in invertebrates. It is possible that insects in particular represent the original targets of some toxins since plants are prone to extensive attack by many groups and species. A static lifestyle may favour evolution of chemical defence against a highly mobile depredator. Canavanine (4.16) and β -hydroxy- γ -methylglutamate (4.17) can act as insect repellents to certain species as well as being toxic while 5-hydroxytryptophan (4.18) and 3,4-dihydroxyphenylalanine (4.19) are toxic to bruchid beetles [61,62] and other insects.



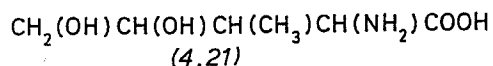
The relationship of (4.18) to a neurotransmitter substance may be of significance since 5-hydroxytryptamine is the major neurotransmitter of insect fore and middle gastrointestinal tract muscle [65]. Dopa is both the precursor of neurotransmitter substances and cuticle cross-linking agents in insects giving rise to dihydroxyphenylalanine and a range of quinonoid and β -substituted ketocatechols [63, 64, 66, 67].

The obvious toxicity of a such wide variety of plant products when present in high concentration leads to the speculation that some human and domestic animal non-specific diseases may originate in low concentrations of toxic amino acids more widely distributed in food plants than presently recognized. For example *N*^ε-acetyl-L- α , γ -diaminobutanoic acid (4.20), which can give rise to α , γ -diaminobutanoic acid, a toxin inhibiting ornithine transcarbamylase in the urea cycle, is present in small quantities in sugar beet [68].



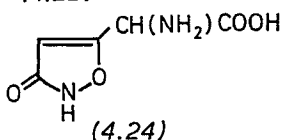
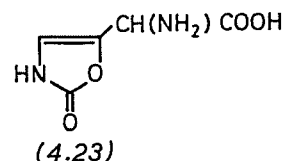
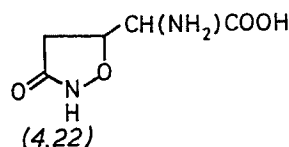
The toxic properties of many unusual amino acid-containing products of micro-organisms are well known, some of their toxicity in mammals being acceptable in view of their even greater toxicity to other micro-organisms. Many unusual peptidic fungal toxins however are physiologically highly active in mammals, usually in a highly deleterious manner [69].

The toxic principles of poisonous toadstools, the genus *Amanita*, are cyclic peptides with unusual amino acid compositions. Although they fall into two groups the phallotoxins and amatoxins [69] they are basically similar with an indole-sulphur bridge derived from tryptophan or hydroxytryptophan and cysteine or cysteine sulphoxide bridging the cyclic structure and with unusual hydroxyamino acids such as L-dihydroxyisoleucine (4.21).



The toxic function in the amatoxins seems to lie in the hydroxyisoleucine component since its replacement by leucine in amanullin results in nontoxicity [69]. The toxic action here as with many of the other fungal products probably lies in the peptide's ability to fit into a symmetric site on RNA polymerase thereby blocking protein synthesis and hence explaining the slow action of the toxin.

Simpler amino acid products than the amatoxins are also responsible for the sinister reputation of fungi. Tricholomic acid (4.22) and muscazone (4.23) are the causes of a variety of lesions to vision, memory and spatial or temporal location in man, while these two from *Tricholoma muscarium* and *Amanita muscaria* together with ibotenic acid (4.24) from *Amanita pantherina* are potent insecticides [69]. All are based upon isoxazole.



The fungi also produce non-protein amino and imino acid phytotoxins. Lycomarasmin (4.25) and the aspergillomarasmins (4.26) occur in *Fusarium* causing leaf wither in tomato by complexation of ferric ion while fusaric acid (4.27) from the same source causes leaf yellowing [69]. Soya leaf necrosis is caused by rhizobitoxin (4.28) of *Rhizobium japonicum* which blocks cystathionine conversion to homocysteine. Interference with sulphur metabolism seems to be a

to Acids

The Non-Protein Amino Acids

69

containing products of
y in mammals being
micro-organisms. Many
ally highly active in

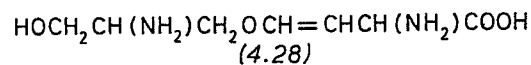
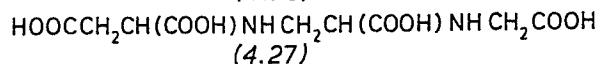
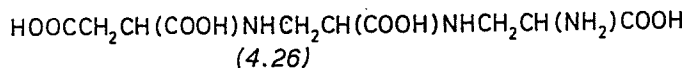
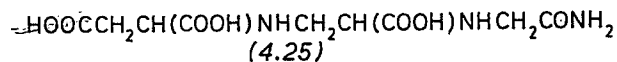
is *Amanita*, are cyclic
igh they fall into two
sically similar with an
droxytryptophan and
ure and with unusual
).

DOH

the hydroxyisoleucine
results in nontoxicity
gal products probably
A polymerase thereby
v action of the toxin.
also responsible for the
muscazone (4.23) are
ial or temporal location
and *Amanita muscaria*
are potent insecticides

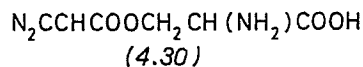
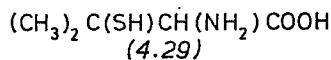
(H₂) COOH

ino acid phytotoxins.
26) occur in *Fusarium*
ion while fusaric acid
a leaf necrosis is caused
blocks cystathionine
stabolism seems to be a

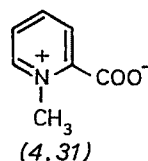


common consequence of toxicity of the unusual amino acids. The locus of action of many antibiotics is still imperfectly understood although undoubtedly many, whose cyclic structures contain novel amino acids or protein amino acids in the D-conformation such as gramicidin or enniatin, act at the level of the bacterial cell membrane affecting ion permeability. Antibiosis is a consequence of competition and evolutionary pressure at the microbiological level [69].

The penicillins derived from penicillamine (4.29) act at the level of peptidoglycan biosynthesis inhibition acting as analogues of D-alanyl-D-alanine and binding to the active site of the bacterial transacylase. The cephalosporins are closely related to the penicillins [43] containing a D- α -amino adipoyl side-chain. While the majority of antibiotics are rather elaborate molecules a few are quite simple. Azaserine (4.30) is one example while L-2-amino-4-(4'-amino-2',5'-cyclohexadienyl)butanoic acid [70] is another.



Animals, in contrast to plants and micro-organisms, produce few toxins which contain novel amino acids. The requirements of toxins of animal origin are usually for a swift and painful response and as we have seen, biotoxic peptides of abnormal content or novel amino acids usually act upon metabolism at a slower rate. Such rapid responses are more usually achieved through the action of neurotransmitter analogues, steroid glycosides and relatively simple peptides which elicit a swift allergic response. Some molluscan venoms however are mixtures of amines, peptides and proteins and these may yet be shown to have novel constituents. Homarin (4.31) which is present in some cone shell venoms, has a curare-like effect [69].



Some tunicates (ascidians or sea squirts) have cytotoxic, non-protein amino acid-containing depsipeptides [71] similar in many respects to depsipeptides of

microbial and fungal origin but this group are atypical in many other respects of animal metabolism. For functions of non-protein amino acids in the animal kingdom we should look rather to physiological roles.

4.10 PHYSIOLOGICAL FUNCTION OF NON-PROTEIN AMINO ACIDS

As we have seen, much of the biological function of novel amino acids in plants and in micro-organisms may be associated not directly with the physiology of the organism itself but with its relationship with other organisms in its environment. We are thus dealing here with physiological (toxic), deterrent (pheromonal) or other modifying roles external to the responsible species.

The question of the physiological functions, within higher plants, of novel amino acids still remains for the majority an uncharted sea, a doubtful and disputed land. The role of such compounds in nitrogen storage has already been touched upon and clearly is of some considerable consequence to the plant; other functions are obscure.

That such compounds may be involved in physiological functions seems to be suggested by observations upon plants made at different stages of development, at different seasons and in different years or under conditions of stress. Such observations show deletions or accumulations of particular novel amino acids under different conditions although the reasons behind this are by no means clear. For example γ -hydroxy- γ -methylglutamic acid is detectable in some years in *Asplenium* but not during others [72]. Canavanine accumulations in seeds may vanish during germination [73]. There is no doubt also that osmotic control in water stress may be achieved by changes in free amino acid concentration over considerable ranges [74]. Free amino acids may also be involved in ion-binding, a factor which could be of consequence in marine and freshwater algae. Cases of the sequestration of metabolically important ions, as for example to avenic acid, have recently been established [75].

Some unusual amino acids lie upon routes to well-established plant metabolites involved in important processes. For example the plant hormone ethylene has been shown to be produced from methionine via *S*-adenosylmethionine and 1-aminocyclopropane-1-carboxylic acid [76].

In invertebrates some unusual amino acid function is associated with energy supply to tissue under anoxic conditions; strombine, alanopine and octopine [34,35] whose biosynthesis has already been referred to function in this manner to maintain a basal rate of energy production.

Certain amino acids in animals also seem to be associated with osmotic regulation. Glycine betaine for example has this function in many molluscs in addition to acting as a donor of the methyl group.

An increasing diversity of phosphorylated amino acids, less well-known than phosphocreatine, function in invertebrates as the latter compound does in the muscle of vertebrates, as phosphagens associated with the maintenance of energy transfer during contraction. Such compounds include *N*-phosphorylarginine, phosphoglycocyamine and lombricine.

ino Acids

n many other respects of
ino acids in the animal

ION-PROTEIN

vel amino acids in plants
with the physiology of the
nisms in its environment.
sterrent (pheromonal) or
ies.

i higher plants, of novel
ted sea, a doubtful and
storage has already been
quence to the plant; other

ical functions seems to be
at stages of development,
onditions of stress. Such
icular novel amino acids
ad this are by no means
detectable in some years
cumulations in seeds may
o that osmotic control in
acid concentration over
e involved in ion-binding,
reshwater algae. Cases of
r example to avenic acid,

l-established plant meta-
ie plant hormone ethylene
-adenosylmethionine and

is associated with energy
, alanopine and octopine
o function in this manner

associated with osmotic
tion in many molluscs in

ids, less well-known than
er compound does in the
the maintenance of energy
le *N*-phosphorylarginine,

The Non-Protein Amino Acids

71

Table 4.1. Some common non-protein α -amino acids and related intermediates of metabolism [43]

Amino acid	Formula	Involvement
<i>N</i> -Acetylglutamic acid	$\begin{array}{c} \text{NHCOCH}_3 \\ \\ \text{HOOCCH}_2\text{CH}_2\text{CHCOOH} \end{array}$	Route to ornithine
<i>N</i> -Acetylglutamic acid semialdehyde	$\begin{array}{c} \text{NHCOCH}_3 \\ \\ \text{OCHCH}_2\text{CH}_2\text{CHCOOH} \end{array}$	Route to ornithine
<i>N</i> -Acetyllysine	$\text{CH}_3\text{CONH}[\text{CH}_2]_4\text{CH}(\text{NH}_2)\text{COOH}$	Lysine degradation
<i>N</i> -Acetylornithine	$\begin{array}{c} \text{NHCOCH}_3 \\ \\ \text{H}_2\text{N}[\text{CH}_2]_3\text{CHCOOH} \end{array}$	Route to ornithine
<i>O</i> -Acetylserine	$\text{CH}_3\text{COOCH}_2\text{CH}(\text{NH}_2)\text{COOH}$	Serine metabolism
<i>S</i> -Adenosylhomocysteine	$\begin{array}{c} \text{Adenosyl} \\ \\ \text{HSCH}_2\text{CH}_2\text{CH}(\text{NH}_2)\text{COOH} \\ + \\ \text{Adenosyl} \end{array}$	Methionine metabolism
<i>S</i> -Adenosylmethionine	$\begin{array}{c} \text{Adenosyl} \\ \\ \text{CH}_3\text{SCH}_2\text{CH}_2\text{CH}(\text{NH}_2)\text{COOH} \\ + \\ \text{Adenosyl} \end{array}$	Methionine metabolism
2-Amino adipic acid	$\text{HOOC}[\text{CH}_2]_3\text{CH}(\text{NH}_2)\text{COOH}$	Lysine degradation
2-Amino adipic acid semialdehyde	$\text{OCH}[\text{CH}_2]_3\text{CH}(\text{NH}_2)\text{COOH}$	Lysine degradation
2-Amino-3-carboxymuconic acid 6-semialdehyde	$\begin{array}{c} \text{COOH} \\ \\ \text{C}=\text{C} \\ / \quad \backslash \\ \text{OHC} \quad \text{NH}_2 \\ \\ \text{COOH} \end{array}$	Tryptophan metabolism
2-Amino-3-ketobutanoic acid	$\begin{array}{c} \text{O} \\ \\ \text{H}_3\text{C}-\text{C}-\text{CH}(\text{NH}_2)\text{COOH} \end{array}$	Threonine metabolism
2-Aminomuconic acid	$\begin{array}{c} \text{HOOC} \\ \\ \text{C}=\text{C} \\ / \quad \backslash \\ \text{COOH} \quad \text{NH}_2 \end{array}$	Tryptophan metabolism
2-Aminomuconic acid 6-semialdehyde	$\begin{array}{c} \text{OCH} \\ \\ \text{C}=\text{C} \\ / \quad \backslash \\ \text{COOH} \quad \text{NH}_2 \end{array}$	Tryptophan metabolism
Argininosuccinic acid	$\begin{array}{c} \text{HOOCCHCH}_2\text{COOH} \\ \\ \text{N} \\ \\ \text{H}_2\text{N}-\text{C}-\text{NH}[\text{CH}_2]_3\text{CH}(\text{NH}_2)\text{COOH} \end{array}$	Ornithine cycle
Aspartylphosphate	$\text{P}^-\text{OOCCH}_2\text{CH}(\text{NH}_2)\text{COOH}$	Aspartate to lysine route
Aspartylsemialdehyde	$\text{OCHCH}_2\text{CH}(\text{NH}_2)\text{COOH}$	Aspartate to lysine route
Betaine	$\text{OOCCH}_2\text{N}^+(\text{CH}_3)_3$	Biosynthesis of glycine