SPECIALIST PERIODICAL REPORTS

Amino-acids Peptides and Proteins

VOLUME 6

THE CHEMICAL SOCIETY

Specialist Periodical Reports

The enormous increase in the rate of accumulation of knowledge witnessed during the past few years now makes any attempt at comprehensive coverage of chemistry within the compass of a single annual volume, such as The Chemical Society *Annual Reports*, unattainable both in terms of reasonable size and cost.

It was with this in mind that in 1967 the Society launched the *Specialist Periodical Reports* series. The aim of the series is to provide systematic and comprehensive review coverage of the progress in the major areas of chemical research. The series has now reached some 34 titles, each being published in annual, or in some cases biennial, volumes.

The Society has been fortunate in obtaining the services of many leading experts in their specialist fields and the series thus provides a unique service for the active specialist chemist—critical in-depth accounts of the progress in their areas by acknowledged authorities, which usually appear less than twelve months after the period of literature coverage.

As the series grows towards its full complement of titles, it is intended that *Annual Reports on the Progress of Chemistry* will continue its shift in emphasis towards provision for the general reader of critical coverage of significant advances in all the major areas of chemistry.

For further details of related *Specialist Periodical Reports* see the back cover of this jacket. For a complete list and full information on all titles in the series write to:

The Marketing Officer
The Chemical Society
Burlington House
London W1V 0BN

ISBN 0 85186 054 0
Library of Congress Catalog No. 72-92548

A Specialist Periodical Report

Amino-acids, Peptides, and Proteins

Volume 6

A Review of the Literature Published during 1973

Senior Reporter

R. C. Sheppard, *MRC Laboratory of Molecular Biology, Cambridge*

ISBN: 0 85186 054 0

Library of Congress Catalog Card No. 72-92548

The Chemical Society
Burlington House, London, W1V 0BN

Reporters

G. C. Barrett, *Oxford Polytechnic*
J. Bridgen, *MRC Laboratory of Molecular Biology, Cambridge*
B. Bycroft, *University of Nottingham*
R. Camble, *ICI Pharmaceuticals Division, Alderley Park, Macclesfield*
E. A. Carrey, *University of Newcastle*
C. Chothia, *Institut Pasteur, Paris*
A. S. Dutta, *ICI Pharmaceuticals Division, Alderley Park, Macclesfield*
A. T. Hagler, *The Weizmann Institute, Rehovot, Israel*
C. E. Johnson, *University of Liverpool*
J. H. Jones, *University of Oxford*
P. Knowles, *University of Leeds*
G. L. E. Koch, *MRC Laboratory of Molecular Biology, Cambridge*
R. H. Pain, *University of Newcastle*
G. R. Penzer, *University of York*
H. W. Rattle, *Portsmouth Polytechnic*
B. Ridge, *University of Exeter*
B. Robson, *University of Newcastle*
R. M. Stephens, *Portsmouth Polytechnic*

Organic formulae composed by Wright's Symbolset method

PRINTED IN GREAT BRITAIN BY JOHN WRIGHT AND SONS LTD., AT THE STONEBRIDGE PRESS, BRISTOL

Preface

This sixth Report reviews papers relevant to the chemistry of amino-acids, peptides, and proteins published in the main journals during 1973. Metal derivatives are reviewed biennially in this series and are not included this year. Likewise, next year's Report will include reviews of circular dichroism and optical rotatory dispersion studies and association–dissociation reactions of proteins covering the literature for 1973 and 1974. The other main areas surveyed this year are the same as those in Volume 5.

I thank most warmly the many contributors to this volume for their time and effort devoted to its preparation.

R. C. SHEPPARD

Contents

Abbreviations

Abbreviations for amino-acids and their use in the formulations of derivatives follow, with some exceptions, the various Recommendations of the I.U.P.A.C.–I.U.B. Commission on Biochemical Nomenclature which have been reprinted in Volumes 4 and 5 in this series.

Other abbreviations which have been used are:

Adoc	adamantyloxycarbonyl
Aoc	t-amyloxycarbonyl
Asu	α-aminosuberic acid
Asx	aspartic acid or asparagine (not yet determined)
ATP	adenosine 5′-triphosphate
Bpoc	2-(4-biphenylyl)-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
DTNB	5,5′-dithiobis-(2-nitrobenzoic acid)
Ec	ethylcarbamoyl
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	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
OPfp	pentafluorophenoxy
ONSu	succinimido-oxy
OPcp	pentachlorophenoxy
ONp	*p*-nitrophenoxy
ONp(*o*)	*o*-nitrophenoxy
OPic	4-picolyloxy
o.r.d.	optical rotatory dispersion
OTcp	2,4,5-trichlorophenoxy
Pipoc	piperidino-oxycarbonyl
Pth-Gly	the phenylthiohydantoin derived from glycine, *etc.*
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
Ser(P)	*O*-phosphorylserine
t.l.c.	thin-layer chromatography
u.v.	ultraviolet
Ztf	1-benzyloxycarbonylamino-2,2,2-trifluoro-ethyl

1
Amino-acids

BY G. C. BARRETT

1 Introduction

More new amino-acids, and more new syntheses, reactions, and properties of amino-acids, have been reported in the year under review. However, the present chapter, intended to be a thorough coverage but excluding many biological aspects, is based on a similar number of references to its predecessors, and is about as long. The 'enormous increase in the rate of accumulation of knowledge' which feeds these Specialist Periodical Reports is apparent in some physical and analytical aspects of the study of amino-acids, and in methods of synthesis, but is far more evident in the biological aspects.

Textbooks and Reviews.—Series One of the MTP International Review of Science includes a Volume with reviews of amino-acids [1] partly overlapping the coverage of Volumes 3 and 4 of the present series; a textbook on microbial synthesis of amino-acids has appeared.[2] Various reviews are cited in the appropriate sections below.

2 Naturally Occurring Amino-acids

Occurrence of Known Amino-acids.—The implications of the distribution of non-protein amino-acids in plants, and their role, have been discussed.[1a] A more restricted review of physiologically active amino-acids in plants concentrates on mimosine and indospicine.[3] Paleobiochemical aspects of evolution are considered in a review of the amino-acid content of black cherts of different geological age.[4] Extraterrestrial distribution, in meteorites [5] and in the moon soil brought back by the Apollo 14 mission,[6] has been reviewed; improved analytical methods give more emphatic

[1] MTP International Review of Science, Organic Chemistry, Series One, Vol. 6, ed. D. H. Hey and D. I. John, 1973: (*a*) 'Amino-acids of Natural Origin', E. A. Bell, p. 1; (*b*) 'Synthesis, Structural Properties, and Reactions of Amino-acids', A. Thomson, p. 17.

[2] 'Microbial Production of Amino-acids', ed. K. Yamada, Kodansha, Tokyo, 1972.

[3] M. P. Hegarty, *Australas. J. Dermatol.*, 1973, **14**, 35.

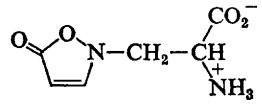
[4] W. Heller, *Naturwiss.*, 1973, **60**, 460.

[5] E. Anders, R. Hayatsu, and M. H. Studier, *Science*, 1973, **182**, 781.

[6] C. W. Gehrke, R. W. Zumwalt, K. Kuo, W. A. Aue, D. L. Stalling, K. A. Kvenvolden, and C. Ponnamperuma, 'Proceedings of the 3rd Lunar Science Conference', ed. D. Heymann, Massachusetts Institute of Technology, Cambridge, Mass., 1972, Vol. 2, p. 2119.

answers concerning the moon samples than reported [7] for samples from earlier missions (amino-acids are present to the extent of *ca.* 20—70 p.p.b.[7]). No amino-acids are present in moon soil [8] within the detection limits of ion-exchange and g.l.c. methods – between 300 pg and 1 ng for the protein amino-acids. The range of amino-acids in certain meteorites is more positively identified.[5]

Studies with pea seedlings (*Pisum sativum*) continue to provide information on early stages in biosynthetic processes. Isolation of *N*-malonyl-D-alanine and γ-L-glutamyl-D-alanine from this source has been reported,[8] the first time derivatives of D-alanine have been found in higher plants. Enzyme preparations from this source, or from *Pisum arvense*, convert *O*-acetyl-serine and isoxazolin-5-one into β-(2-isoxazolin-5-onyl)alanine (1).[9] T.l.c. of extracts of pea seedlings exposed to $H_2{}^{35}S$ reveals the presence



(1)

of a radioactive substance with all the properties of thio-threonine (α-amino-β-mercaptobutyric acid).[10] Pyroglutamic acid, found in large amounts in young pea seedling extracts, is nevertheless a glutamine artefact.[11] Maturing seeds of lima beans biosynthesize *Se*-methyl-selenocysteine, while selenomethionine accumulates as the principal product of selenate assimilation in the leaves of this plant.[12]

Other plant and fungal sources, and the more notable amino-acids present, are: *Sedum acre* (N^{ε}-methyl-lysine);[13] *Cannabis* seeds (L-isoleucine betaine);[14] a New Guinea *Boletus* (L-2-amino-4-methylhex-5-enoic acid).[15] (2*S*,3*R*,4*R*)-4-Hydroxyisoleucine, known as a component of γ-amanatin but not previously reported to be present in higher plants, is found in *Trigonella foenum-graecum* seeds [16] (the 2*R*,3*R*,4*R*-diastereoisomer is a minor component). *N*-Feruloylglycyl-peptides in barley seed globulins are considered to arise through the presence of *N*-feruloylglycine as a 'starter' of protein biosynthesis.[17]

[7] 'Amino-acids, Peptides, and Proteins', ed. G. T. Young (Specialist Periodical Reports), The Chemical Society, London, 1972, Vol. 4, p. 1.
[8] T. Ogawa, M. Fukuda, and K. Sasaoka, *Biochim. Biophys. Acta*, 1973, **297**, 60; M. Fukuda, A. Tokumura, T. Ogawa, and K. Sasaoka, *Phytochemistry*, 1973, **12**, 2593.
[9] I. Murakoshi, F. Kato, J. Haginiwa, and L. Fowden, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 918.
[10] J. Schnyder and K. H. Erismann, *Experientia*, 1973, **29**, 232.
[11] H. G. Wager and F. A. E. Porter, *J. Sci. Food Agric.*, 1973, **24**, 69.
[12] S. N. Nigam and W. B. McConnell, *Phytochemistry*, 1973, **12**, 359.
[13] E. Leistner and I. D. Spenser, *J. Amer. Chem. Soc.*, 1973, **95**, 4715.
[14] C. A. L. Bercht, R. J. J. C. Lousberg, F. J. E. M. Kuppers, and C. A. Salemink, *Phytochemistry*, 1973, **12**, 2457.
[15] E. Gellert, B. Halpern, and R. Rudzats, *Phytochemistry*, 1973, **12**, 689.
[16] L. Fowden, H. M. Pratt, and A. Smith, *Phytochemistry*, 1973, **12**, 1707.
[17] C. F. Van Sumure, H. de Pooter, H. Ali, and M. Degrauw-van Bussel, *Phytochemistry*, 1973, **12**, 407.

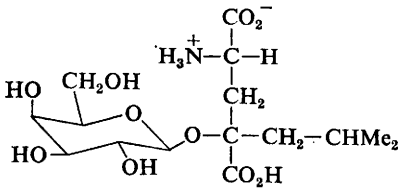
An increasing number of side-chain methylated arginines (N^G-monomethyl-, N^GN^G-dimethyl-, $N^GN^{G'}$-dimethyl-) and lysines (N^ϵ-methyl-, N^ϵ-dimethyl-, and N^ϵ-trimethyl-) has been found in bovine brain proteins;[18] only the first two of these are present in myelin basic proteins from a wide variety of species.[19] Hypusine, N^6-(4-amino-2-hydroxybutyl)-2,6-diaminohexanoic acid, is present in bovine brain protein.[20] 2,4-Diaminobutyric acid is a component of the murein of seventeen strains of coryneform bacteria.[21] Formation of L-2-amino-4-methoxy-*trans*-but-3-enoic acid by *Pseudomonas aeruginosa* grown on straight-chain paraffins is notable amongst microbiological studies;[22] this amino-acid was previously detected in culture fluids from the same source grown on cerelose.[23] Subtilin, a peptide produced by *Bacillus subtilis*, contains two residues of dehydroalanine and one of β-methyldehydroalanine, and is similar in this respect to nisin.[24]

Microbial synthesis of amino-acids offers scope for large-scale production of compounds important in medicine.[2, 25] Of course, L-dopa production is pre-eminent in such studies, and its formation, with L-tyrosine, from DL-serine and catechol by *Erwinea herbicola*,[26] from pyruvate, ammonia, and catechol by the same organism,[27] or from tyrosine by *Vibrio tyrosinaticus*,[28] has been described. Other microbial syntheses recently reported concern tryptophan,[29] L-threonine,[30] and N^δ-acetyl-L-ornithine.[31]

β-Alanine betaine is present in the adductor muscle of the fan mussel *Atrina pectinata japonica*.[32]

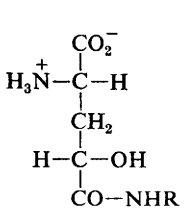
New Natural Free Amino-acids.—Many of the new amino-acids reported this year are hydroxylated versions of long-known amino-acids, *e.g.* (2*S*,4*R*)-4-(β-D-galactopyranosyloxy)-4-isobutylglutamic acid (2) from flowers of *Reseda odorata*.[33] The assignment of 4*R*-configuration is

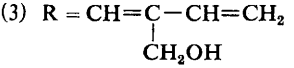
[18] Y. Matsuoka, *Seikagaku*, 1972, **44**, 353.
[19] G. E. Deibler and R. E. Martenson, *J. Biol. Chem.*, 1973, **248**, 2387.
[20] N. Imaoka and T. Nakajima, *Biochim. Biophys. Acta*, 1973, **320**, 97.
[21] F. Fiedler and O. Kandler, *Arch. Mikrobiol.*, 1973, **89**, 51.
[22] U. Sahm, G. Knobloch, and F. Wagner, *J. Antibiotics*, 1973, **26**, 389.
[23] L. H. Sello, T. Williams, and A. Stempel, *J. Antibiotics*, 1972, **25**, 122.
[24] E. Gross and H. H. Kiltz, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 559.
[25] S. Abe, *Kagaku Kogyo*, 1973, **24**, 389.
[26] H. Enei, H. Matsui, H. Nakazawa, S. Okumura, and H. Yamada, *Agric. and Biol. Chem.* (*Japan*), 1973, **37**, 493.
[27] H. Enei, H. Nakazawa, S. Okumura, and H. Yamada, *Agric. and Biol. Chem.* (*Japan*), 1973, **37**, 725.
[28] H. Yoshida, Y. Tanada, and K. Nakayama, *Agric. and Biol. Chem.* (*Japan*), 1973, **37**, 2121.
[29] H. Nakazawa, H. Enei, S. Okumura, and H. Yamada, *Agric. and Biol. Chem.* (*Japan*), 1972, **36**, 2523.
[30] T. Hirakawa, T. Tanaka, and K. Watanabe, *Agric. and Biol. Chem.* (*Japan*), 1973, **37**, 123; T. Hirakawa, *ibid.*, p. 243.
[31] K. Araki and J. Nakajima, *Agric. and Biol. Chem.* (*Japan*), 1973, **37**, 2639.
[32] S. Konosu and K. Watanabe, *Nippon Suisan Gakkaishi*, 1973, **39**, 645.
[33] P. O. Larsen, H. Soerensen, D. W. Cochran, E. W. Hagaman, and E. Wenkert, *Phytochemistry*, 1973, **12**, 1713.

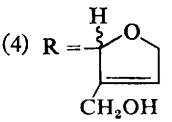


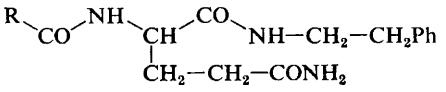
(2)

tentative. This is the first example of a glycoside from a higher plant in which the carbohydrate residue is linked to an aliphatic hydroxy-group of an amino-acid. Further details of structural studies reported for pinnatanine (3) in 1971 [34] are available,[35] together with details of a new analogue from the same source, seeds of *Staphylea pinnata* L., oxypinnatanine (4).[35] *N*-Acylated L-glutaminyl 2-phenylethylamines (5), and





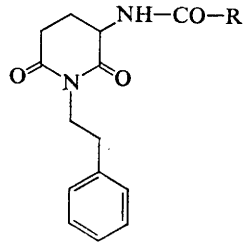




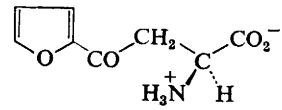
(5) a; R = Pr^i
b; R = (*R*)-EtCHMe

related glutarimides (6), have been isolated from *Croton humilis*,[36] the novel feature as far as the glutarimide is concerned being the *R*-configuration of the 2-methylbutanoyl substituent, the first natural appearance of the *R*-enantiomer of this acid; the *S*-analogue of (6b), julocrotine, has been known for some time.[37] N^5-(2′-Hydroxybenzyl)-allo-4-hydroxy-L-glutamine and N^5-(4′-hydroxybenzyl)-L-glutamine are new amino-acid amides from buckwheat seeds.[38] L-3-(2-Furoyl)alanine (7) from buckwheat seeds [39]

[34] M. D. Grove, M. E. Daxenbichler, D. Weisleder, and C. H. VanEtten, *Tetrahedron Letters*, 1971, 4477.
[35] M. D. Grove, D. Weisleder, and M. E. Daxenbichler, *Tetrahedron*, 1973, **29**, 2715.
[36] K. L. Stuart, D. McNeill, J. P. Kutney, G. Eigendorf, and F. K. Klein, *Tetrahedron*, 1973, **29**, 4071.
[37] T. Nakano, C. Djerassi, R. A. Corral, and O. O. Orazi, *J. Org. Chem.*, 1961, **26**, 1184.
[38] M. Koyama, Y. Tsujizaki, and S. Sakamura, *Agric. and Biol. Chem. (Japan)*, 1973, **37**, 2749.
[39] A. Ichihara, H. Hasegawa, H. Sato, M. Koyama, and S. Sakamura, *Tetrahedron Letters*, 1973, 37.



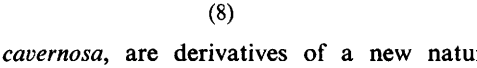
(6) a; R = Pr^i
b; R = (*R*)-EtCHMe



(7)

and from Green Gram seeds[40] may be an artefact since it is formed[41] from ascorbalamic acid during treatment with hot hydrochloric acid. The amino-acid (7) from buckwheat seeds has the L-configuration since it gives L-aspartic acid on ozonolysis,[39] and (7) is readily synthesized from 2-acetylfuran,[39] from L-3-chloroalanine and ascorbic acid,[41] or from furacyl bromide by the acetamidomalonate route.[41]

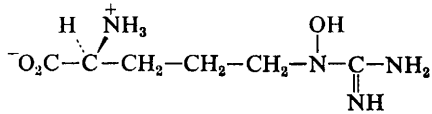
4-Hydroxy-3-hydroxymethylphenylalanine and 3-hydroxymethylphenylalanine are major constituents of seeds of *Caesalpinia tinctoria*.[42] *N*-Acylated 2-methylene-β-alanine methyl esters (8; n = 13—17) from the sponge



(8)

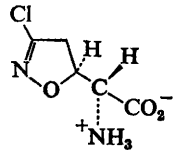
Fasciospongia cavernosa, are derivatives of a new naturally occurring β-amino-acid.[43]

Amino-acids with fungal or bacterial origins have been reported. Another new acetylenic amino-acid from *Tricholomopsis rutilans* is 2-amino-3-hydroxyhex-4-ynoic acid, $MeC{\equiv}C{-}CH(OH){-}CH(NH_3^+){-}CO_2^-$ (stereochemistry not yet elucidated).[44] N^5-Hydroxy-L-arginine (9) from *Nannizzia gypsea*[45] and from an unidentified bacillus[46] may be a

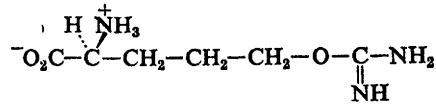


(9)

[40] T. Kasai, Y. Kishi, M. Sano, and S. Sakamura, *Agric. and Biol. Chem.* (*Japan*), 1973, **37**, 2923.
[41] R. Couchman, J. Eagles, M. P. Hegarty, W. M. Laird, R. Self, and R. L. M. Synge, *Phytochemistry*, 1973, **12**, 707.
[42] R. Watson and L. Fowden, *Phytochemistry*, 1973, **12**, 617.
[43] Y. Kashman, L. Fishelson, and I. Ne'eman, *Tetrahedron*, 1973, **29**, 3655.
[44] S. I. Hatanaka, Y. Niimura, and K. Taniguchi, *Z. Naturforsch.*, 1973, **28C**, 475.
[45] B. Fischer, W. Keller-Schierlein, H. Kniefel, W. A. Koenig, W. Loeffler, A. Mueller, R. Muntwyler, and H. Zaehner, *Arch. Mikrobiol.*, 1973, **91**, 203.
[46] H. Maehr, J. F. Blount, D. L. Pruess, L. Yarmchuk, and M. Kellett, *J. Antibiotics*, 1973, **26**, 284.



(10)

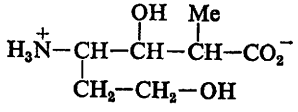


(11)

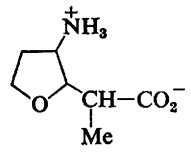
valuable antimicrobial agent, but even more important perhaps will be the antitumour property of the antibiotic (10) from *Streptomyces sviceus*;[47] 5-(*O*-isoureido)-L-norvaline (11) is a new arginine antagonist from a bacterial culture.[48]

New Amino-acids from Hydrolysates.—A review, covering amino-acids reported since 1931 to be components of native proteins, describes the less common protein constituents.[49]

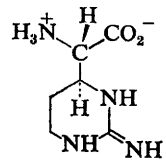
Antibiotics, inhibitory peptides, and antimetabolites are major sources of new amino-acids. This year, Actinomycin Z_1 has come under structural scrutiny;[50] one of its constituent amino-acids is 3-hydroxy-4-oxo-5-methylproline. Actinomycin Z_5 contains *cis*-5-methylproline.[51] Antibiotics YA-56 of the phleomycin–bleomycin group contain an extraordinary variety of amino-acids;[52, 53] in addition to β-amino-β-(4-amino-6-carboxy-5-methyl-2-pyrimidinyl)propionic acid, β-aminoalanine, L-erythro-β-hydroxyhistidine,[54] and 2-[2-(2-aminoethyl)-Δ^2-thiazolin-4-yl]thiazole-4-carboxylic acid, all previously reported, the presence of β-hydroxy-L-valine,[53] 4-amino-3,6-dihydroxy-2-methylhexanoic acid (12),[53] and several unidentified amino-acids, has been reported. The last-named amino-acid appears in hydrolysates as stereoisomers of (13), formed from it by cyclization.[53]



(12)



(13)

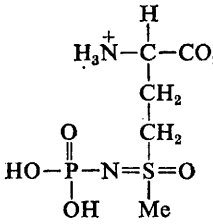


(14)

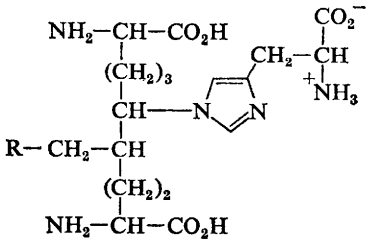
[47] D. G. Martin, D. J. Duchamp, and C. G. Chidester, *Tetrahedron Letters*, 1973, 2549.
[48] W. A. Koenig, H. Kneifel, E. Bayer, G. Mueller, and H. Zaehner, *J. Antibiotics*, 1973, **26**, 44.
[49] H. B. Vickery, *Adv. Protein Chem.*, 1972, **26**, 81.
[50] H. Brockmann and E. A. Stahler, *Tetrahedron Letters*, 1973, 3685.
[51] E. Katz, K. T. Mason, and A. B. Mauger, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 819.
[52] Y. Ohashi, H. Abe, and Y. Ito, *Agric. and Biol. Chem. (Japan)*, 1973, **37**, 2277.
[53] Y. Ohashi, H. Abe, and Y. Ito, *Agric. and Biol. Chem. (Japan)*, 1973, **37**, 2283.
[54] G. Koyama, H. Nakamura, Y. Muraoka, T. Takita, K. Maeda, H. Umezawa, and Y. Iitaka, *J. Antibiotics*, 1973, **26**, 109.

The chymostatins are a group of tetrapeptides of microbial provenance, important as inhibitors of chymotrypsin and papain. They all contain two modified phenylalanine residues and (2*S*,3*S*)-α-(2-iminohexahydro-4-pyrimidyl)glycine (14); chymostatin A also contains L-leucine, while chymostatin B contains L-valine, and chymostatin C contains L-isoleucine.[55] The basic amino-acid (14) is a diastereoisomer of capreomycidine, previously established[56] to be a component of capreomycins.

L-(N^5-Phosphono)methionine-(*S*)-sulphoximine (15) is a component of the tripeptide (15)-L-Ala-L-Ala, an antimetabolite of L-glutamine produced by an unclassified *Streptomycete*.[57]



(15)



(16) a; R = OH

b; R = NH—CH_2—CH(OH)—$(CH_2)_2$—CH(^-O_2C)—H_3N^+

Further information on the aldol-histidine cross-link (16a) of collagen (see Vol. 5, p. 6) is now available.[58] Histidino-hydroxymerodesmosine (16b), formed from aldol-histidine and hydroxylysine, is a newly discovered cross-link of collagen, capable of uniting three or four polypeptide chains. These cross-links, whose structures were largely established by mass

[55] K. Tatsuta, N. Mikami, K. Fujimoto, S. Umezawa, H. Umezawa, and T. Aoyagi, *J. Antibiotics*, 1973, **26**, 625.

[56] B. W. Bycroft, D. Cameron, L. R. Croft, and A. W. Johnson, *Chem. Comm.*, 1968, 1301.

[57] D. L. Pruess, J. P. Scannell, H. A. Ax, M. Kellett, F. Weiss, T. C. Demny, and A. Stempel, *J. Antibiotics*, 1973, **26**, 261.

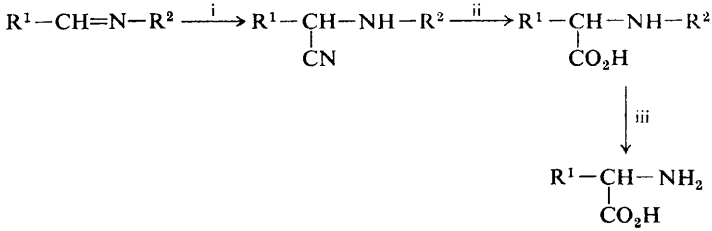
[58] M. L. Tanzer, T. Housley, L. Berube, R. Fairweather, C. Franzblau, and P. M. Gallop, *J. Biol. Chem.*, 1973, **248**, 393.

spectrometry, are surprising discoveries since the pyridinium cross-links present in elastin might have been expected.[58]

3 Chemical Synthesis and Resolution of Amino-acids

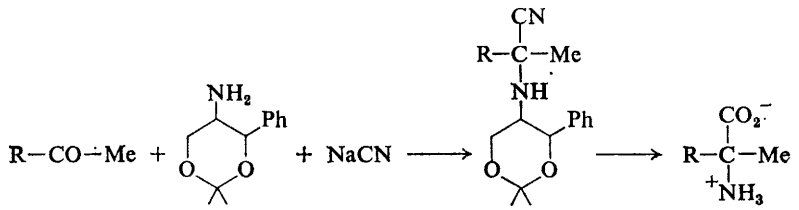
General Methods of Synthesis.—*Asymmetric Synthesis.* The ingenuity in approaches to novel asymmetric syntheses of α-amino-acids seems limitless, and more papers have come into the reckoning for this chapter. Further results on the hydrogenolytic asymmetric transamination approach, involving a resolved α-amino-acid and an α-keto-acid, have been reported.[59] D-α-Phenylglycine undergoes this reaction with pyruvic acid and with α-ketoglutaric acid to give D-alanine and D-glutamic acid, respectively, and a clearer picture of the structural requirements for maximum optical purity is emerging.[59] Schiff bases of aliphatic aldehydes with an optically active benzylamine give optically active N-substituted amino-nitriles with hydrogen cyanide[60,61] or with benzoyl cyanide[62] which, on hydrolysis, give predominantly D-amino-acids when the Schiff base involves the *R*-enantiomer of the α-alkylbenzylamine. Optical purities of the products are in the range 22—58% when hydrogen cyanide is used, and 15—37% when the benzoyl cyanide route is employed (Scheme 1).

A novel asymmetric Strecker synthesis[63] applied to the synthesis of α-methyl-α-amino-acids (Scheme 2) employs a methyl ketone and



Reagents: i, HCN; ii, H_3O^+; iii, H_2–$Pd(OH)_2$–C

Scheme 1

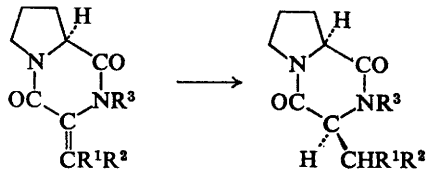


Scheme 2

[59] K. Harada, T. Iwasaki, and T. Okawara, *Bull. Chem. Soc. Japan*, 1973, **46**, 1901.
[60] K. Harada and T. Okawara, *J. Org. Chem.*, 1973, **38**, 707.
[61] K. Harada, T. Okawara, and K. Matsumoto, *Bull. Chem. Soc. Japan*, 1973, **46**, 1865.
[62] K. Harada and T. Okawara, *Bull. Chem. Soc. Japan*, 1973, **46**, 191.
[63] K. Weinges and B. Stemmle, *Chem. Ber.*, 1973, **106**, 2291.

(4*S*,5*S*)-5-amino-2,2-dimethyl-4-phenyl-1,3-dioxan. L-Amino-acids predominate in the product when the starting methyl ketone has an even number of carbon atoms.[63]

Dioxopiperazine formation from *N*-alkyl α-bromopropionamides containing a chiral *N*-alkyl group favours one enantiomer to a small extent, hydrolysis and hydrogenolysis of the dioxopiperazines giving alanine of optical purity up to 27%; chirality and optical purity depend upon the structure and chirality of the *N*-alkyl group.[64] In a related approach,[65] arylidene derivatives (17) of glycyl-L-proline anhydride give corresponding L-phenylalanyl-L-proline anhydrides (18) on hydrogenation, and the route



(17) (18)

has been used to prepare L-phenylalanine and L-dopa. *N*-Methylation of the starting compound to give (17; R^3 = Me) is readily achieved, and the synthesis then becomes a route to L-(*N*-methyl)phenylalanines. There is no asymmetric induction with alkylidene analogues (17; R^1 = R^2 = Me, or R^1 = H, R^2 = $CHMe_2$),[65] and no doubt this will arouse curiosity.

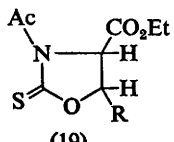
L-Phenylglycine, L-alanine, L-phenylalanine, and L-leucine are obtained with optical purities 27, 33, 62, and 75%, respectively, by Neber rearrangement of corresponding (−)-menthyl *N*-chloroimidates.[66]

β-Hydroxy-α-amino-acids prepared by the reaction of aldehydes with glycine derivatives at high pH are enriched in one enantiomer if the N-substituted glycine is part of a resolved cobalt(III) complex.[67] The (+)-enantiomer of potassium bis-(*N*-salicylideneglycinato)cobaltate(III), obtained by resolution of the racemate with brucine, reacts with acetaldehyde at pH 11.2 to give a mixture of D-threonine and D-allothreonine in asymmetric yields 19—46% and 55—64%, respectively.[67]

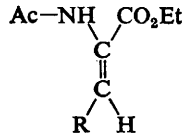
Syntheses using α-Isocyano- and α-Isothiocyanato-esters. Variations of the method (see Vol. 5, p. 12) using oxazolin-2-one-4-carboxylates to synthesize β-hydroxy-α-amino-acids have broadened the scope of the route. Like the oxazolin-2-ones, *N*-acetyl oxazolin-2-thiono-4-carboxylates (19) may be cleaved by base (Bu^tOK in the present case [68]) to give derivatives of

[64] T. Okawara and K. Harada, *Bull. Chem. Soc. Japan*, 1973, **46**, 1869.
[65] H. Poisel and U. Schmidt, *Chem. Ber.*, 1973, **106**, 3408.
[66] Y. Nogami, Y. Kawazoe, and T. Taguchi, *Yakugaku Zasshi*, 1973, **93**, 1058.
[67] Y. N. Belokon, M. M. Dolgaya, N. I. Kuznetsova, S. V. Vitt, and V. M. Belikov, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1973, 156.
[68] D. Hoppe, *Angew. Chem. Internat. Edn.*, 1973, **12**, 656.

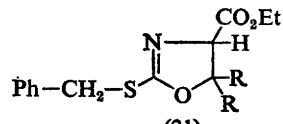
α-aminoacrylic acid (20) from which a variety of β-functionalized α-amino-acids may be prepared.[68] A similar route, starting likewise with α-isothiocyanato-acetates, but proceeding *via* 2-benzylthio-oxazolin-4-carboxylates (21), gives N-protected α-aminoacrylates.[69]



(19)

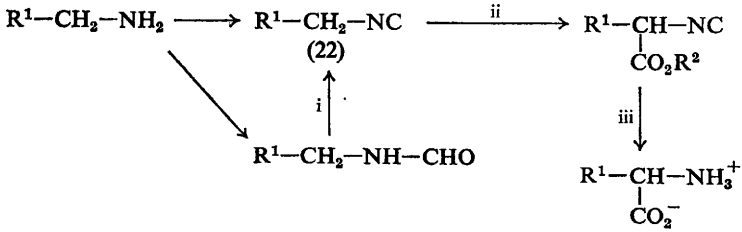


(20)



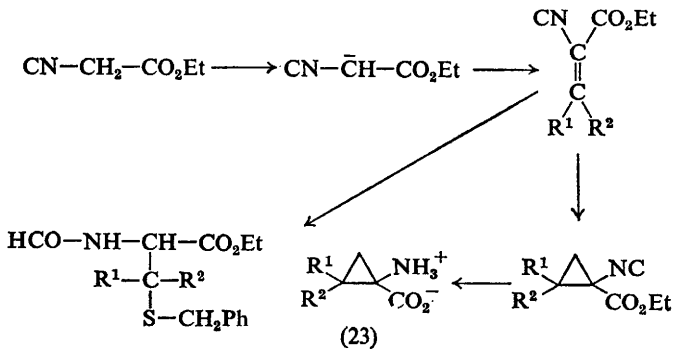
(21)

A new synthesis of α-amino-acids (Scheme 3) involving α-carboxylation of amines with chloroformates or dialkyl carbonates,[70, 71] after their conversion into isocyanides (22), should have useful applications in cases where side-chains can withstand the successive reactions. Ethyl isocyanoacetate $NCCH_2CO_2Et$ gives α-isocyanoacrylates by reaction with ketones, from which β,β-disubstituted *N*-formyl *S*-benzylcysteines (Scheme 4) may



Reagents: i, $COCl_2$–NEt_3; ii, $ClCOOR^2$ or R^2OCOOR^2; iii, aq. HCl

Scheme 3



(23)

Scheme 4

[69] D. Hoppe, *Angew. Chem. Internat. Edn.*, 1973, **12**, 658.
[70] W. Vaalburg, J. Strating, and M. G. Woldring, *Synthetic Comm.*, 1972, **2**, 423.
[71] K. Matsumoto, M. Suzuki, and M. Miyoshi, *J. Org. Chem.*, 1973, **38**, 2094.

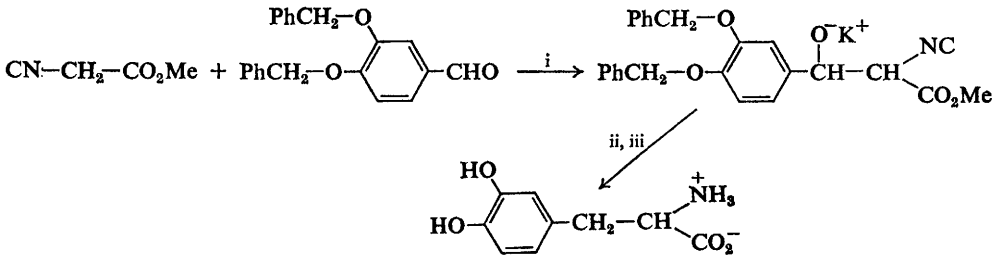
be obtained by addition of benzyl mercaptan;[72] alternatively, cyclopropane amino-acids (23) may be obtained by their reaction with dimethylsulphonium methylide (Scheme 4).[73]

The same general scheme has been employed in a new synthesis of DL-dopa (Scheme 5) in a one-pot process.[74] Methyl α-alkyl isocyanoacetates give oxazolines with aromatic aldehydes and α-acyl-α-amino-acids with acid chlorides, from which β-hydroxy-α-amino-acids are obtained, the two routes favouring *threo*- and *erythro*-isomers, respectively (Scheme 6).[75, 76]

Syntheses by Classical Methods. α-Halogeno-amines or -acids, on treatment with cyanide ion or amines, respectively, provide intermediates from which corresponding α-amino-acids may be obtained.[14, 77, 78] The former route, using $CF_3 \cdot CHCl \cdot NH \cdot CO \cdot CH_2Ph$, provides a new synthesis of trifluoroalanine.[79] Reaction of $p\text{-}ClC_6H_4 \cdot S \cdot NH \cdot CO \cdot OCH_2Ph$ with an α-bromo-acid gives the corresponding *N*-benzyloxycarbonyl-DL-amino-acid after removal of the *N*-(*p*-chlorophenylsulphenyl) group from the condensation product by treatment with *p*-chlorothiophenol.[80]

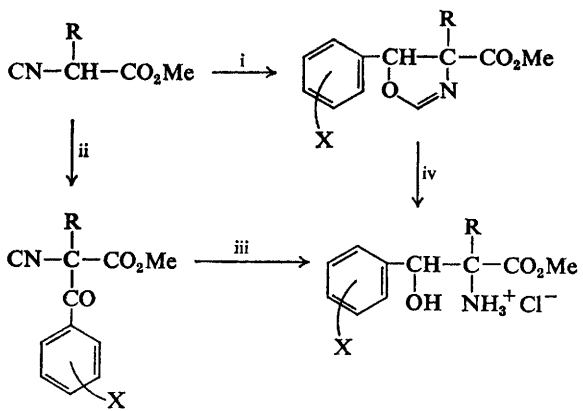
Alkylation of α-acylamino-malonic esters, then hydrolysis of the product, remains a standard route to α-amino-acids.[81–84] One modification includes a half-ester stage, permitting decarboxylation (in refluxing dioxan) to the *N*-acylamino-acid ester to be effected, so that an enzymic hydrolysis step (chymotrypsin or subtilisin) to give a mixture of *N*-acyl-L-α-amino-acid and *N*-acyl-D-α-amino-acid esters may be incorporated.[85] A modified Curtius reaction with malonic acid half-esters, $HO_2C \cdot CHR \cdot CO_2Et$, using diphenylphosphoryl azide with triethylamine, followed by addition of benzyl alcohol provides [86] *N*-benzyloxycarbonylamino-acid esters, including derivatives of α,α-disubstituted series not accessible directly through the acylamino-malonate route. Alkylation of methyl α-nitroacetate, $MeOCO\text{-}CH_2NO_2$, followed by reduction and hydrolysis, is a related route used to prepare amino-alkyl glycines and representative protein amino-acids;[87] *NN*-bis(trimethylsilyl)glycine trimethylsilyl esters, $(Me_3Si)_2NCH_2CO_2Si\text{-}$

[72] U. Schollkopf and D. Hoppe, *Annalen*, 1973, 799.
[73] U. Schollkopf, R. Harms, and D. Hoppe, *Annalen*, 1973, 611.
[74] R. Damico and J. M. Nicholson, *J. Org. Chem.*, 1973, **38**, 3057.
[75] M. Suzuki, T. Iwasaki, K. Matsumoto, and K. Okumura, *Chem. and Ind.*, 1973, 228.
[76] M. Suzuki, T. Iwasaki, M. Miyoshi, K. Okumura, and K. Matsumoto, *J. Org. Chem.*, 1973, **38**, 3571.
[77] J. E. Baldwin, J. Loliger, W. Rastetter, N. Neuss, L. L. Huckstep, and N. de la Higuera, *J. Amer. Chem. Soc.*, 1973, **95**, 3796; corrigendum, *ibid.*, p. 6511.
[78] M. Gacek and K. Undheim, *Tetrahedron*, 1973, **29**, 863.
[79] A. Uskert, A. Neder, and E. Kasztreiner, *Magyar Kém. Folyóirat*, 1973, **79**, 333.
[80] T. Taguchi and T. Mukaiyama, *Chem. Letters*, 1973, 1.
[81] Y. K. Lee and T. Kaneko, *Bull. Chem. Soc. Japan*, 1973, **46**, 2924.
[82] M. L. Sethi, G. S. Rao, and G. J. Kapadia, *J. Pharm. Sci.*, 1973, **62**, 1802.
[83] H. R. Bosshard and A. Berger, *Helv. Chim. Acta*, 1973, **56**, 1838.
[84] K. Balenovic and A. Deljac, *Rev. Trav. chim.*, 1973, **92**, 117.
[85] A. Berger, M. Smolarsky, N. Kurn, and H. R. Bosshard, *J. Org. Chem.*, 1973, **38**, 457.
[86] S. Yamada, K. Ninomiya, and T. Shioiri, *Tetrahedron Letters*, 1973, 2343.
[87] E. Kaji and S. Zen, *Bull. Chem. Soc. Japan*, 1973, **46**, 337.



Reagents: i, Bu^tOK–MeOH; ii, HCl–MeOH 17 h; iii, Pd–H_2–H_3O^+

Scheme 5



Reagents: i, XC_6H_4CHO; ii, XC_6H_4COCl; iii, H_2–Pd; iv, aq. HCl

Scheme 6

Me_3, may be alkylated as the α-sodio-derivatives, illustrated by the preparation of DL-2-amino-6-nitrocaproic acid, a masked lysine for use in peptide synthesis.[88]

Reductive cleavage of phenylhydrazones of α-keto-acids, readily available through the Japp–Klingemann reaction starting from active methylene compounds, would be more widely used with a convenient hydrogenolysis procedure. Whereas their treatment with zinc dust and acid at 0 °C is ineffective, hydrogenation of aqueous suspensions over palladium gives α-amino-acids in high yields.[89] An improved method for converting α-hydroxy-acids into *N*-phthaloylamino-acids with inversion of configuration employs phthalimide, triphenylphosphine, and diethyl azodicarboxylate.[90] Treatment of a bisulphite addition product of an aldoxime with sodium cyanide, followed by hydrolysis, can be vaunted as a new synthesis of amino-acids from amines,[91] since aldoximes can now be prepared from amines, *e.g.*

$$HOCH_2CH_2NH_2 \longrightarrow HOCH_2CH{=}NOH \longrightarrow HOCH_2CH(NH_2)CO_2H$$

Routes to amino-acids other than α-amino-acids are not easily generalized since various strategies must be chosen to provide the required location of the functional groups. 4*S*-Amino-3*S*-hydroxy-6-methylheptanoic acid $Me_2CH{\cdot}CH(NH_2){\cdot}CH(OH){\cdot}CH_2{\cdot}CO_2H$, a constituent of pepstatin A, may be synthesized from L-leucine,[92] all four stereoisomers being available by straightforward methods.[93] The Arndt–Eistert reaction gives optically active β-amino-acids starting from α-amino-acids.[94]

Synthesis under Simulated Prebiotic Conditions.—Reviews on abiogenic synthesis of amino-acids have appeared.[95, 96] A 20 mA silent discharge through a 0.14 : 0.21 : 1 mixture of methane, carbon dioxide, and nitrogen gives a 1% overall yield of glycine, alanine, norvaline, serine, aspartic acid, α-amino-isobutyric acid, and other unidentified amino-acids.[97] High temperatures reduce the number of amino-acids and favour the formation of β-amino-acids when a 2 : 1 mixture of water-saturated methane and ammonia is passed over quartz sand at temperatures between 900 and 1060 °C.[98] Amino-acid precursors (methane, nitrogen, carbon monoxide

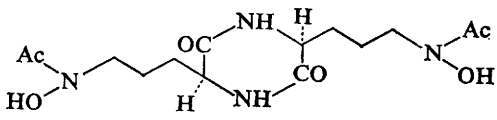
[88] E. Bayer and K. Schmidt, *Tetrahedron Letters*, 1973, 2051.
[89] N. H. Khan and A. R. Kidwai, *J. Org. Chem.*, 1973, **38**, 822.
[90] M. Wada, T. Sano, and O. Mitsunobu, *Bull. Chem. Soc. Japan*, 1973, **46**, 2833.
[91] G. Natta and I. Pasquon, *Chimie et Industrie*, 1973, **55**, 323.
[92] H. Morishima, T. Takita, and H. Umezawa, *J. Antibiotics*, 1973, **26**, 115.
[93] M. Kinoshita, S. Aburaki, A. Hagiwara, and J. Imai, *J. Antibiotics*, 1973, **26**, 249.
[94] Y. Seto, T. Yamada, K. Niwa, S. Miwa, F. Tanaka, S. Kuwata, and H. Watanabe, *Chem. Letters*, 1973, 151.
[95] R. M. Lemmon, *Environ. Biol. Med.*, 1973, **2**, 1.
[96] F. Balestic, *J. Chim. phys.*, 1973, **70**, 169.
[97] E. F. Simonov and V. B. Lukyanov, *Vestnik Moskov Univ., Khim.*, 1973, **14**, 118 (*Chem. Abs.*, 1973, **79**, 5557c).
[98] J. G. Lawless and C. D. Boynton, *Nature*, 1973, **243**, 405.

and dioxide, carbon disulphide, hydrogen cyanide) are present in moon soil brought back by Apollo 14 and earlier missions, at 20—70 ng g^{-1} levels.[99]

With a substantial accumulation of data on the precursors and conditions required for chance synthesis of α-amino-acids, interest is shifting towards the next stage, the chance coupling of such amino-acids to form peptides. Reaction of amino-acids with polyphosphates in the prebiotic soup in the presence of Mg^{2+} can give phosphoramidates, $R^1O \cdot PO_2^- \cdot NH \cdot CHR^2 \cdot CO_2^-$, as derivatives suitably activated towards peptide bond formation.[100] Peptide bond formation induced by hydroxyapatite or orthophosphates is enhanced by cyanate,[101] formed as a consequence of oligomerization of hydrogen cyanide and subsequent hydrolysis of the resulting non-peptide polymer.[102] Acid hydrolysis of the polymer gives amino-acids, citrulline being a major constituent of the hydrolysate;[102] further study of the cyanide oligomerization should reveal new chemistry of more general interest,[103] and, already, the rate of oligomerization has been shown to be independent of added nucleophiles.[104]

Protein and Other Naturally Occurring Amino-acids.—New syntheses described in the preceding sections are illustrated by syntheses of some of the well-known protein amino-acids.

Syntheses of alanosine (β-hydroxynitrosamino-L-alanine) have been reported by two groups.[105, 106] One synthesis [105] starts with DL-β-chloroalanine, which on reaction with sodium *anti*-benzaldoximate gives the *N*-hydroxyamine, $HONHCH_2CH(NH_2)CO_2H$, after cleavage of the benzylidene group; DL-alanosine is obtained on nitrosation. The L-isomer is obtained [106] starting from ethyl 2,3-dibromopropionate and *N*-tosyl-*O*-benzylhydroxylamine, a resolution stage being included late in the synthesis. A further synthesis of δ-hydroxy-L-ornithine has been reported,[107] and N^{δ}-acetyl-N^{δ}-benzyloxy-L-ornithine for use in the synthesis of rhodo-

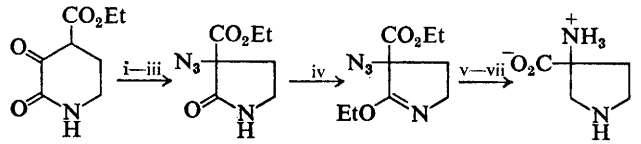


(24)

[99] S. W. Fox, K. Harada, and P. E. Hare, 'Proceedings of the 3rd Lunar Science Conference', ed. D. Heymann, Massachusetts Institute of Technology, Cambridge, Mass., 1972, Vol. 2, p. 2109.
[100] R. Lohrmann and L. E. Orgel, *Nature*, 1973, **244**, 418.
[101] J. J. Flores and J. O. Leckie, *Nature*, 1973, **244**, 436.
[102] J. P. Ferris, D. B. Donner, and A. P. Lobo, *J. Mol. Biol.*, 1973, **74**, 499.
[103] J. P. Ferris and T. J. Ryan, *J. Org. Chem.*, 1973, **38**, 3302.
[104] J. P. Ferris, D. B. Donner, and A. P. Lobo, *J. Mol. Biol.*, 1973, **74**, 511.
[105] C. N. Eaton, G. H. Denney, M. A. Ryder, M. G. Ly, and R. D. Babson, *J. Medicin. Chem.*, 1973, **16**, 289.
[106] Y. Isowa, H. Kurita, M. Ohmori, M. Sato, and K. Mori, *Bull. Chem. Soc. Japan*, 1973, **46**, 1847.
[107] G. Tomlinson and T. Viswanatha, *Canad. J. Biochem.*, 1973, **51**, 754.

torulic acid (24) has been prepared [108] from diethyl α-(3-bromopropyl)-α-acetamidomalonate by reaction with $PhCH_2ONHAc$ followed by conventional elaboration and deprotection steps.

In a new synthesis (Scheme 7) of (±)-cucurbitine (25), a novel ring-contraction provides ethyl 3-azido-2-oxopyrrolidine-3-carboxylate, reduc-



Reagents: i, Br_2; ii, NaN_3–boiling 1,2-dimethoxyethane; iii, peroxyacetic acid; iv, $Et_3O^+BF_4^-$; v, diborane; vi, H_2–Pt; vii, H_2O

Scheme 7

tion of which gives (±)-(25).[109] The (−)-isomer occurs in several species of Cucurbitaceae.

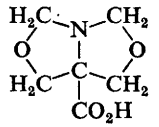
α-Alkyl- and α-Aralkyl-α-amino-acids.—Continued exploration in the synthesis of α-alkyl-α-amino-acids is referred to in preceding paragraphs.[63, 75, 86] *N*-Acetyl-α-benzylphenylalanine ethyl ester is obtained by Schmidt rearrangement of ethyl dibenzylacetoacetate, and converted without difficulty into its *N*-*o*-nitrophenylsulphenyl derivative for use in peptide synthesis.[110]

α-Hydroxy-α-amino-acids and Amino-acids with Aliphatic Hydroxy-groups in the Side-chain.—Amino-acids with α-hydroxy- or α-alkoxy-groups are of renewed interest, partly because of the possible value in medicine of 6-methoxy-penicillins and cephalosporin analogues. *N*-Phenylacetylamino-acid esters give *N*-chloro-derivatives with t-butyl hypochlorite, which on dehydrochlorination give acylimines $PhCH_2CON{=}CRCO_2Me$, to which an alcohol may be added to give α-alkoxy-α-amino-acids, *e.g.* $PhCH_2{\cdot}CO{\cdot}NH{\cdot}CPh(OMe){\cdot}CO_2Me$.[111, 112] *N*-Acyl dehydroalanine esters give analogous α-alkoxy-*N*-acylalanine esters through acid-catalysed addition of alcohols.[113]

β-Hydroxy-α-amino-acids may be obtained by reaction between a glycine derivative and an aldehyde or ketone, usually in alkaline solution, and asymmetric synthesis possibilities have been demonstrated.[67] A similar route uses α-isocyanoacetates.[75] The prebiotic significance of the analogous reactions of glycine itself, in mildly alkaline solution, to give

[108] T. Fujii and Y. Hatanaka, *Tetrahedron*, 1973, **29**, 3825.
[109] H. J. Monteiro, *J.C.S. Chem. Comm.*, 1973, 2.
[110] G. C. Barrett, P. M. Hardy, T. A. Harrow, and H. N. Rydon, *J.C.S. Perkin I*, 1972, 2634.
[111] J. E. Baldwin, F. J. Urban, R. D. G. Cooper, and F. L. Jose, *J. Amer. Chem. Soc.*, 1973, **95**, 2401.
[112] G. A. Koppel and R. E. Koehler, *J. Amer. Chem. Soc.*, 1973, **95**, 2403.
[113] G. Lucente and D. Rossi, *Chem. and Ind.*, 1973, 324.

serine, hydroxymethylserine, sarcosine, iminodiacetic acid, and alanine, has been discussed.[114] Bis-(L-serinato)copper(II) reacts with excess formaldehyde at pH 7—9 to give the bis(amino-acidato)copper(II) complex involving (26), from which α-hydroxymethylserine is obtained on treatment with hydrogen sulphide.[115] An optically active intermediate can be

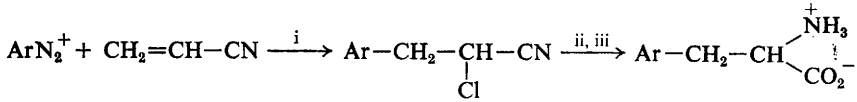


(26)

detected in the hydroxymethylation stage, indicating initial hydroxymethylation on nitrogen.

γ-Hydroxy-α-amino-acids are an important group of naturally occurring compounds for which improved synthetic methods are being developed. Photochlorination of α-amino-acids carrying a γ-hydrogen atom gives corresponding α-amino-γ-lactones after hydrolysis;[116] an alternative route to such lactones uses diethyl allyl acetamidomalonate as starting material,[81] which by addition of Br_2 and treatment with 48% hydrobromic acid gives diastereoisomeric 2-amino-5-bromo-4-valerolactone hydrobromides which after separation and hydrolysis give DL-γ-hydroxyproline and the allo-isomer.[81] Synthesis of all stereoisomers of γ-hydroxyglutamic acid is reported by the same workers.[117]

Aromatic and Heterocyclic Amino-acids.—Among numerous routes to the phenylalanine analogues which have been synthesized recently, that employing acrylonitrile (Scheme 8) is of wide potential use.[118] The oxazolone route,[119, 120] and the acetamidomalonate route [82] or its benzyloxycarbonylamino variant [83] have been used for phenylalanine analogues.



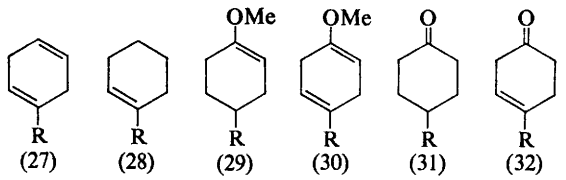
Reagents: i, $CuCl_2$–NaOAc; ii, NH_3; iii, H_3O^+

Scheme 8

[114] A. S. Subbaraman, Z. A. Kazi, and A. S. U. Choughuley, *Indian J. Biochem. Biophys.*, 1972, **9**, 268.
[115] J. R. Brush, R. J. Magee, M. J. O'Connor, S. B. Teo, R. J. Geue, and M. R. Snow, *J. Amer. Chem. Soc.*, 1973, **95**, 2034.
[116] H. Faulstich, J. Dolling, K. Michl, and T. Wieland, *Annalen*, 1973, 560.
[117] Y. K. Lee and T. Kaneko, *Bull. Chem. Soc. Japan*, 1973, **46**, 3494.
[118] M. Y. Mogilevskii, N. T. Morozova, and O. E. Antropova, *Zhur. obshchei Khim.*, 1973, **43**, 1822.
[119] K. Karpavicius, G. Prasmickiene, L. Gurviciene, and O. V. Kildisheva, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1973, 1887.
[120] B. Sila, J. Wojtanis, and T. Lesiak, *Roczniki Chem.*, 1973, **47**, 1281.

Aryl-substitution procedures include a new synthesis of L-dopa from L-tyrosine;[121] Friedel–Crafts reaction of L-tyrosine with acetyl chloride gives 3-acetyl-L-tyrosine, from which L-dopa is obtained by treatment with alkaline H_2O_2. Previously described '2,4-dibromo-L-phenylalanine' is shown to be the 2,3-isomer;[122] separation of the bromination products of L-phenylalanine on Sephadex LH-20 using aqueous methanol gives 2,3-dibromo- (2%), 2,5-dibromo- (13%), 3,4-dibromo- (5%), 2-bromo- (29%), and 4-bromo-L-phenylalanine (36%).[122] 5-Fluoro-L-dopa (*i.e.* 3,4-dihydroxy-5-fluoro-L-phenylalanine) and its [5-^{18}F]-analogue have been synthesized *via* the Schiemann reaction,[123] and a variety of nitro-imidazole analogues of L-histidine is now available.[124] N^{α}-Benzoyl-L-histidine gives 2,4-bis(arylazo) coupled products with arenediazonium salts, whereas the *N*-acetyl analogue gives predominantly the 2-arylazo-imidazole, which gives 2-amino-L-histidine by catalytic hydrogenolysis and hydrolysis.[125] Since the previously synthesized 4-fluoro-L-histidine showed interesting *in vivo* and enzymic properties, *N*-acetyl-2-amino-L-histidine methyl ester was diazotized and irradiated in HBF_4 solution, to give 2-fluoro-L-histidine after enzymic hydrolysis and enzymic deacylation.[126]

Birch reduction of L-phenylalanine gives L-3-[1-(cyclohexa-1,4-dienyl)]-alanine (27),[127] while *O*-methyl-L-tyrosine gives cyclohexenes (28) and (29), and the cyclohexadiene (30), by reduction with sodium in liquid ammonia;[128]



(27) (28) (29) (30) (31) (32)

R = $-CH_2-CH(NH_3^+)CO_2^-$

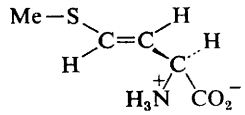
two of these, (29) and (30), are enol ethers, and give the cyclohexanone (31) and the cyclohexenone (32), respectively, on acid hydrolysis.

N-Substituted Amino-acids.—*N*-Carboxymethylation of α-amino-acids can be brought about by treatment with glyoxal in an acetate buffer at 100 °C during 30 min.[129] A useful *N*-methylation procedure is illustrated by the

[121] H. Bretschneider, K. Hohenlohe-Oehringen, A. Kaiser, and U. Wolcke, *Helv. Chim. Acta*, 1973, **56**, 2857.
[122] H. Faulstich, H. O. Smith, and S. Zobeley, *Annalen*, 1973, 765.
[123] G. Firnau, C. Nahmias, and S. Garnett, *J. Medicin. Chem.*, 1973, **16**, 416.
[124] W. Tantz, S. Teitel, and A. Brossi, *J. Medicin. Chem.*, 1973, **16**, 705.
[125] W. Nagai, K. L. Kirk, and L. A. Cohen, *J. Org. Chem.*, 1973, **38**, 1971.
[126] K. L. Kirk, W. Nagai, and L. A. Cohen, *J. Amer. Chem. Soc.*, 1973, **95**, 8389.
[127] G. R. Nagarajan, L. Diamond, and C. Ressler, *J. Org. Chem.*, 1973, **38**, 621.
[128] K. Kaminski and T. Sokolowska, *Roczniki Chem.*, 1973, **47**, 1091.
[129] N. V. Chuyen, T. Kurata, and M. Fujimaki, *Agric. and Biol. Chem. (Japan)*, 1973, **37**, 2209.

one-step conversion of an *N*-benzyloxycarbonylamino-acid into its *N*-methyl derivative with methyl iodide and sodium hydride.[130] Surprisingly, even *N*-benzyloxycarbonyl-methionine could be selectively *N*-methylated.[130] Use of N^{in}-formyl-L-tryptophan, as its N^{α}-Boc-derivative, in solid-phase peptide synthesis has been studied.[131]

Amino-acids containing Sulphur.—Interest in α-mercapto-α-amino-acids is a consequence of the presence of the $-NH-\overset{|}{C}(S-)-CO-$ group in gliotoxins and related natural products. *N*-Acyl α-chloro-α-amino-acids, $R^1CONHCR^2ClCO_2H$, obtained by addition of hydrogen chloride in acetic acid to *N*-acyl αβ-dehydroamino-acids, give α-mercapto-, α-*S*-acylmercapto-, and α-*S*-benzhydrylthio-*N*-acylamino-acids.[132] An alternative approach[133] using methyl t-butyl 2-benzylthio-malonate, $Bu^tO_2C-CH(SCH_2Ph)CO_2Me$, as starting material, gives *N*-Boc DL-α-benzylthioglycine. The dipeptide of this amino-acid could be cyclized to the dioxopiperazine and debenzylated, but conversion into the 3,6-epi-dithiodioxopiperazinedione was not successful.[133]

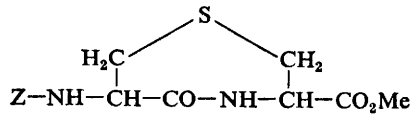


(33)

(+)-(*S*)-*E*-2-Amino-4-methylthiobut-3-enoic acid (33; dehydro-L-methionine) has been synthesized[84] starting with 3-acetamido-3,3-bis(ethoxycarbonyl)propionaldehyde, $AcNHC(CO_2Et)_2CH_2CHO$. The mercaptal, formed with methanethiol, was saponified, decarboxylated, and converted into the methyl ester $AcNHCH(CO_2Me)CH_2CH(SMe)_2$; an interesting use of the newer sulphur chemistry followed, whereby the mercaptal gives the monosulphoxide with sodium metaperiodate, which on pyrolysis gives dehydro-L-methionine as its *N*-acetyl methyl ester, as a mixture of *E*- and *Z*-isomers. Separation and resolution gave the (+)-*E*-isomer; no doubt there will be many potential amino-acid antagonists of this type synthesized in the future.

A lanthionine synthesis starting with *N*-benzyloxycarbonyl-*S*-benzoyl-L-cysteinyl-*O*-benzoyl-L-serine methyl ester, *via* the cyclo-lanthionine derivative (34) followed by hydrolysis, is also suitable for the synthesis of the *meso*-isomer.[134] *S*-Adenosyl-L-homo-cysteine and nucleoside analogues, and their sulphoxides, have been prepared as potential inhibitors of transmethylases.[135]

[130] J. R. McDermott and N. L. Benoiton, *Canad. J. Chem.*, 1973, **51**, 1915.
[131] D. Yamashiro and C. H. Li, *J. Org. Chem.*, 1973, **38**, 2594.
[132] S. M. Patel, J. O. Currie, and R. K. Olsen, *J. Org. Chem.*, 1973, **38**, 126.
[133] T. Petrzilka and C. Fehr, *Helv. Chim. Acta*, 1973, **56**, 1218.
[134] L. Zervas and N. Ferderigos, *Experientia*, 1973, **29**, 262.
[135] V. N. Rekunova, I. P. Rudakova, and A. M. Yurkevich, *Tetrahedron Letters*, 1973, 2811.



(34)

A List of α-Amino-acids which have been Synthesized for the First Time [a]

Compound	*Ref.*
3-(2-Furoyl)alanine [b]	39, 41
2-Amino-6-nitrocaproic acid [b]	88
γ-Hydroxyglutamic acid [b]	117
3,4-Dehydromethionine [b]	84
threo-β-Amino-DL-asparagine	136
erythro-β-Amino-DL-asparagine	136
threo-β-Phenyl-DL-asparagine	136
erythro-β-Phenyl-DL-asparagine	136
threo-β-(*N*-4-Hydroxyphenyl)amino-L-asparagine	136
erythro-β-(*N*-4-Hydroxyphenyl)amino-L-asparagine	136
S-Inosyl-L-homocysteine	135
S-Inosyl-L-homocysteine sulphoxide	135
S-Guanosyl-L-homocysteine	135
S-Guanosyl-L-homocysteine sulphoxide	135
S-Uridyl-L-homocysteine	135
S-Uridyl-L-homocysteine sulphoxide	135
S-Cystidyl-L-homocysteine	135
S-Cystidyl-L-homocysteine sulphoxide	135
DL-γ-Hydroxyproline	81
DL-Allo-γ-hydroxyproline	81
DL-2-Aminoadamantane-2-carboxylic acid	137
Pentamethylphenylalanine [b]	138
DL-3,4,5-Trimethoxyphenylalanine	82
DL-4-Hydroxy-3,5-dimethoxyphenylalanine	82
DL-3-Hydroxy-4,5-dimethoxyphenylalanine	82
DL-3,4-Dihydroxy-5-methoxyphenylalanine	82
L-(*p*-Pentafluorophenyl)phenylalanine	83
DL-(3,4-Dihydroxy-5-fluoro)phenylalanine	123
DL-(3,4-Dihydroxy-[5-^{18}F]fluoro)phenylalanine	123
L-*o*-Methylphenylalanine [b]	85
L-β-(2-Naphthyl)alanine	85
L-β-(6-Quinolyl)alanine	85
DL-*p*-Bis-(2-chloropropylamino)phenylalanine	119
DL-5,7-Disubstituted 2-coumaronylalanines	120
DL-3-Amino-1,2,3,4-tetrahydrocarbazole 3-carboxylic acid	139
DL-3-Amino-6-hydroxy-1,2,3,4-tetrahydrocarbazole 3-carboxylic acid	139

[136] P. K. Chang, L. J. Sciarini, and R. E. Handschumacher, *J. Medicin. Chem.*, 1973, **16**, 1277.
[137] H. T. Nagasawa, J. A. Elberling, and F. N. Shirota, *J. Medicin. Chem.*, 1973, **16**, 823.
[138] G. I. Tesser, H. G. A. Slits, and J. W. van Nispen, *Internat. J. Peptide Protein Res.*, 1973, **5**, 119.
[139] Y. Maki, T. Masugi, T. Hiramitsu, and T. Ogiso, *Chem. and Pharm. Bull.* (*Japan*), 1973, **21**, 2460.

List of α-Amino-acids (*cont.*)

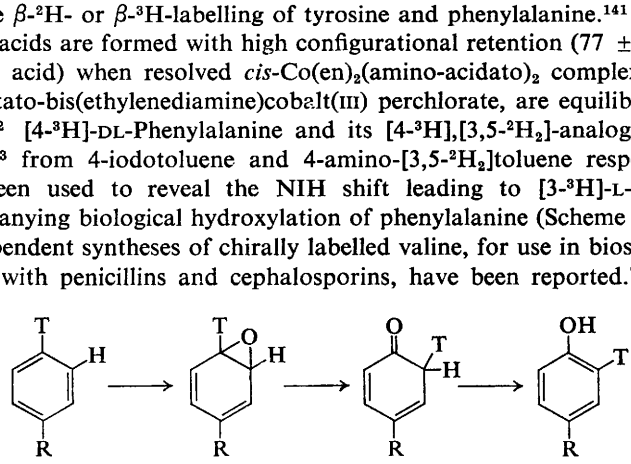
Compound	*Ref.*
DL-3-Amino-6-methoxy-1,2,3,4-tetrahydrocarbazole 3-carboxylic acid	139
DL-3-Amino-6-benzyloxy-1,2,3,4-tetrahydrocarbazole 3-carboxylic acid	139
L-2-Aminohistidine	125
L-2-Fluorohistidine	125
DL-1-Methyl-2-nitrohistidine	124
L-1-Methyl-4-nitrohistidine	124
L-1-Methyl-5-nitrohistidine	124

[a] Other new amino-acids, and labelled analogues of known amino-acids, are mentioned in this chapter.
[b] All stereoisomers synthesized and separated.

Amino-acids with Amino-alkyl Side-chains.—Further details of the synthesis of all four stereoisomers of 2,3-diaminobutyric acid from threonine and allothreonine have been given.[140] Reference is made elsewhere in this chapter to the synthesis of amino-acids with modified amino-alkyl side-chains.[105–108]

Labelled Amino-acids.—In some areas, the ability to conduct increasingly sophisticated biosynthetic studies is dependent upon the availability of specifically labelled α-amino-acids. Full details have been given of stereoselective β-^{2}H- or β-^{3}H-labelling of tyrosine and phenylalanine.[141] [2-^{2}H]-Amino-acids are formed with high configurational retention (77 ± 2% for aspartic acid) when resolved *cis*-Co(en)$_2$(amino-acidato)$_2$ complexes, *e.g.* L-aspartato-bis(ethylenediamine)cobalt(III) perchlorate, are equilibrated in ^2H_2O.[142] [4-^{3}H]-DL-Phenylalanine and its [4-^{3}H],[3,5-$^2H_2$]-analogue, prepared[143] from 4-iodotoluene and 4-amino-[3,5-$^2H_2$]toluene respectively, have been used to reveal the NIH shift leading to [3-^{3}H]-L-tyrosine accompanying biological hydroxylation of phenylalanine (Scheme 9).

Independent syntheses of chirally labelled valine, for use in biosynthetic studies with penicillins and cephalosporins, have been reported.[77, 144, 145]



Scheme 9

[140] E. Atherton and J. Meienhofer, *Z. physiol. Chem.*, 1973, **354**, 689.
[141] G. W. Kirby and J. Michael, *J.C.S. Perkin I*, 1973, 115.
[142] W. E. Keyes and J. I. Legg, *J. Amer. Chem. Soc.*, 1973, **95**, 3431.
[143] W. R. Bowman, W. R. Gretton, and G. W. Kirby, *J.C.S. Perkin I*, 1973, 218.
[144] D. J. Aberhardt and L. J. Lin, *J. Amer. Chem. Soc.*, 1973, **95**, 7859; for related synthetic studies, see R. K. Hill, S. Yan, and S. M. Arfin, *ibid.*, p. 7857.
[145] H. Kluender, C. H. Bradley, C. J. Sih, P. Fawcett, and E. P. Abraham, *J. Amer. Chem. Soc.*, 1973, **95**, 6149.

A straightforward route (Scheme 10) to (2*RS*,3*S*)-[4,4,4-$^2H_3$]valine is notable for employing a 'mild' Strecker synthesis at the last stage.[144] Rather longer routes (Schemes 11[77] and 12[145]) give (2*RS*,3*R*)-[4-^{13}C]-valine[77] and (2*S*,3*S*)-[4-^{13}C]valine.[145]

[5-^{3}H]-L-Ornithine and [6-^{3}H]-L-lysine are available through catalytic tritiation of ω-cyano-L-aminobutyric acid and ω-cyano-L-norvaline, respectively.[146]

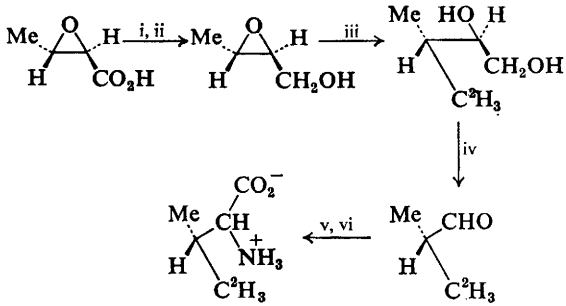
A clever application of the reversibility of the tyrosine phenol-lyase reaction allows [^{14}C]phenol to be incorporated into L-tyrosine in high yield, and using mushroom tyrosinase, [^{14}C]catechol can be prepared from [^{14}C]phenol and [^{14}C]-L-tyrosine can be converted into [^{14}C]-L-dopa. Preparation of [^{14}C]-L-dopa with two different labelling patterns is then possible, since [^{14}C]catechol can be incorporated into [^{14}C]-L-dopa using tyrosine phenol-lyase; ring [4-^{14}C]- and ring [3,4-$^{14}C_2$]-labelling of L-dopa are achieved in this way.[147]

A commercial sample of DL-phenylalanine, non-specifically labelled with ^{3}H, gave $PhCH_2CO_2H$ retaining 96% of the label, on degradation;[148] electrophilic substitution studies indicated that 45% of this residual label was at the *p*-position, and 26% in each of the *o*-positions.[148]

Resolution of Amino-acids.—The claim[149] that preferential adsorption and polymerization of the L-enantiomer of a DL-amino-acid takes place on edge faces of kaolinite has been disputed for the original test case of DL-phenylalanine in aqueous solutions at pH 2 or 6.[150] Since the claim provides a start towards explaining differential availability of D- and L-enantiomers of amino-acids, and chance formation of proteins composed of L-amino-acid residues, no doubt its refutation will come under further scrutiny. Complete resolution of DL-tryptophan using bovine serum albumin attached to succinoylaminoethylagarose as an affinity chromatographic matrix has been reported;[151] as should be expected, the D-isomer emerges first from the column. The first optical resolution by liquid–liquid chromatography of a racemate by differential complexation has been demonstrated with DL-α-amino-acids.[152] The host–guest molecule principle is employed, based upon chiral cyclic polyethers; considerable planning is required to devise the right host molecule for a particular amino-acid, but then complete optical resolution of valine is easily achieved.[152, 153]

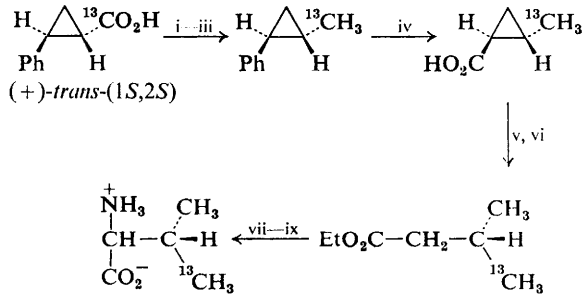
Resolution by analytical g.l.c. is discussed in Section 5 of this chapter. More conventional procedures include a preferential crystallization

[146] M. Havranek and I. Mezo, *Acta Chim. Acad. Sci. Hung.*, 1973, **77**, 341.
[147] B. E. Ellis, G. Major, and M. H. Zenk, *Analyt. Biochem.*, 1973, **53**, 470.
[148] R. B. Herbert and I. T. Nicholson, *J. Labelled Compounds*, 1973, **9**, 567.
[149] T. A. Jackson, *Experientia*, 1971, **27**, 242.
[150] W. A. Bonner and J. Flores, *Biosystems*, 1973, **5**, 103.
[151] K. K. Stewart and R. F. Doherty, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2850.
[152] R. C. Helgeson, K. Koga, J. M. Timko, and D. J. Cram, *J. Amer. Chem. Soc.*, 1973, **95**, 3021.
[153] R. C. Helgeson, J. M. Timko, D. J. and Cram, *J. Amer. Chem. Soc.*, 1973, **95**, 3023.



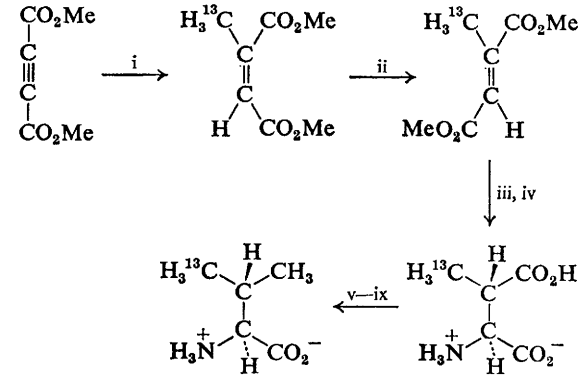
Reagents: i, CH_2N_2; ii, $NaBH_4$; iii, C^2H_3Li; iv, $NaIO_4$; v, NH_4OH–NaCN, 0 °C; vi, reflux conc. HCl

Scheme 10



Reagents: i, $LiAlH_4$; ii, mesyl chloride; iii, $LiAlH_4$; iv, O_3; v, diazoethane; vi, Li–NH_3; vii, Br_2; viii, NH_3; ix, H_3O^+

Scheme 11



Reagents: i, $^{13}CH_3Cu$; ii, *hν*/trace of Br_2; iii, H_3O^+; iv, β-methylaspartase; v, reduce *N*-trifluoroacetyl α-methyl ester with B_2H_6; vi, mesyl chloride; vii, NaI; viii, reduction; ix, H_3O^+

Scheme 12

procedure applied to DL-amino-acid arenesulphonate salts;[154] resolution of β-(2-pyridyl)-DL-alanine as the tartrate salt;[155] and enzymic resolution (using papain) of *N*-alkoxycarbonyl-DL-amino-acids.[155a]

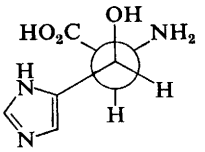
4 Physical and Stereochemical Studies of Amino-acids

Further study of 1-amino-cycloalkane-1-carboxylic acids has resolved disagreement in favour of the preferred adoption of equatorial and axial dispositions, respectively, by carboxy- and amino-groups in such compounds.[156]

Crystal Structures of Amino-acids.—A survey of published crystal structures of phenylalanine and tyrosine derivatives [156a] pays special attention to conformational regularities through the series.

Precision neutron diffraction studies of α-amino-acids (*i.e.* definitive hydrogen locations and conformational features) continue with structures assigned to L-arginine dihydrate,[157] L-tyrosine,[158] L-tyrosine hydrochloride,[158] L-glutamine,[159] 4-hydroxy-L-proline,[160] L-serine monohydrate,[161] and DL-serine.[161]

Absolute configurational assignments have been made to 4*S*-amino-3*S*-hydroxy-6-methylheptanoic acid, through *X*-ray analysis of its *N*-*p*-bromobenzoyl derivative,[162] confirming chemical correlation, and to *erythro*-β-hydroxy-L-histidine.[54] The *gauche* conformation (35), resulting from the presence of the hydroxy-group, is not adopted by other histidine derivatives in the crystal state.[54]



(35)

[154] S. Yamada, M. Yamamoto, and I. Chibata, *J. Org. Chem.*, 1973, **38**, 4408.
[155] L. N. Veselova and E. S. Chaman, *Zhur. obshchei. Khim.*, 1973, **43**, 1637.
[155a] J. L. Abernethy, R. Bobeck, A. Ledesma, and R. Kemp, *J. Org. Chem.*, 1973, **38**, 1286.
[156] Y. Maki, T. Masugi, and K. Ozeki, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 2466.
[156a] V. Cody, W. L. Duax, and H. Hauptmann, *Internat. J. Peptide Protein Res.*, 1973, **5**, 297.
[157] M. S. Lehmann, J. J. Verbist, W. C. Hamilton, and T. F. Koetzle, *J.C.S. Perkin II*, 1973, 133.
[158] M. N. Frey, T. F. Koetzle, M. S. Lehmann, and W. C. Hamilton, *J. Chem. Phys.*, 1973, **58**, 2547.
[159] T. F. Koetzle, M. N. Frey, M. S. Lehmann, and W. C. Hamilton, *Acta Cryst.*, 1973, **B29**, 2571.
[160] T. F. Koetzle, M. S. Lehmann, and W. C. Hamilton, *Acta Cryst.*, 1973, **B29**, 231.
[161] M. N. Frey, M. S. Lehmann, T. F. Koetzle, and W. C. Hamilton, *Acta Cryst.*, 1973, **B29**, 876.
[162] H. Nakamura, H. Morishima, T. Takita, H. Umezawa, and Y. Iitaka, *J. Antibiotics*, 1973, **26**, 255.

Other *X*-ray studies concern the diethylphosphoric acid salt of arginine,[163] DL-aspartic acid,[164] orthorhombic L-cysteine,[165] L-dopa,[166] DL-isoleucine,[167] L-methionine,[168] L-norleucine,[168] γ-aminobutyric acid at low temperature,[169] L-penicillamine hydrochloride monohydrate,[170] N^5-hydroxy-L-arginine hydrobromide,[46] DL-penicillaminato-methylmercury(II),[171] $\alpha S,5S$-α-amino-3-chloro-2-isoxazoline-5-acetic acid,[47] and the bis(amino-acidato)copper(II) complex of (26).[115]

N.M.R. Spectrometry.—Apart from the routine use of proton n.m.r. in support of structural assignments to amino-acids and their derivatives, most of the more notable results reported this year relate to conformational studies. Careful correlation with conformation of the solvent-dependent n.m.r. of the phenylalanine anion in mixed aqueous solvents as a function of temperature indicates the progressive disappearance of less-favoured conformers as the proportion of water in such solvents decreases.[172] Similar studies with histidine, and its *im*-benzyl, N^{α}-acyl, and methyl ester derivatives in ^2H_2O at various pD, suggest a preferred conformation for histidine in basic solutions in which the imidazole and carboxylate groupings are close together.[173]

Stability constants for complexes of neodymium(III) with amino-acids have been determined by n.m.r. and by potentiometric titration, and are relevant to the use of Nd^{3+} complexes as shift reagents for aqueous solution studies.[174]

Proton magnetic relaxation times, now increasingly recognized as capable of providing additional structural information, have been determined for aqueous solutions of glycine and a series of aliphatic amino-acids;[175] T_1 values are pH-independent, but T_2 values have a minimum value near pH 6 for this series. Proton spin–lattice relaxation times at 90 and 270 MHz for several amino-acids show no frequency dependence, while some T_2 values for certain peptides are frequency-dependent;[176] conformational inferences are discussed.

^{19}F N.m.r. reveals preferential binding of ring-substituted *N*-trifluoroacetyl-DL-phenylalanines to chymotrypsin; separate resonances for the

[163] S. Furberg and J. Solbakk, *Acta Chem. Scand.*, 1973, **27**, 1226.
[164] S. T. Rao, *Acta Cryst.*, 1973, **B29**, 1718.
[165] K. A. Kerr and J. P. Ashmore, *Acta Cryst.*, 1973, **B29**, 2124.
[166] J. W. Becker, Y. T. Thathachari, and P. G. Simpson, *Proc. Indian Acad. Sci.* (*A*), 1973, **77**, 99.
[167] E. Benedetti, C. Pedone, and A. Sirigu, *Acta Cryst.*, 1973, **B29**, 730.
[168] K. Torii and Y. Iitaka, *Acta Cryst.*, 1973, **B29**, 2799.
[169] E. G. Steward, R. B. Player, and D. Warner, *Acta Cryst.*, 1973, **B29**, 2038, 2825.
[170] S. N. Rao, R. Parthasarathy, and F. E. Cole, *Acta Cryst.*, 1973, **B29**, 2373.
[171] Y. S. Wong, P. C. Chieh, and A. J. Carty, *J.C.S. Chem. Comm.*, 1973, 741.
[172] J. M. Purcell, J. E. Ramirez, and J. R. Cavanaugh, *J. Phys. Chem.*, 1973, **77**, 1501.
[173] R. J. Weinkam and E. C. Jorgensen, *J. Amer. Chem. Soc.*, 1973, **95**, 6084.
[174] A. D. Sherry, C. Yoshida, E. R. Birnbaum, and D. W. Darnall, *J. Amer. Chem. Soc.*, 1973, **95**, 3011.
[175] D. D. Eley, A. S. Fawcett, and M. J. Hey, *J.C.S. Faraday I*, 1973, **69**, 399.
[176] H. B. Coates, K. A. McLauchlan, I. D. Campbell, and C. E. McColl, *Biochim. Biophys. Acta*, 1973, **310**, 1.

D- and L-enantiomers appear in the presence of the enzyme, with a chemical shift difference related to the inhibitor : enzyme ratio.[177] Assignment of an envelope conformation, with axial fluorine, for *cis*- and *trans*-4-fluoro-L-proline in aqueous solution has been made using ^{1}H and ^{19}F n.m.r. spectra.[178]

All six carbon resonances suffer shifts with varying ionization patterns of the amino, carboxy, and imidazole groupings of L-histidine;[179] related ^{13}C n.m.r. studies with 1-methyl- and 3-methyl-histidines in comparison with histidine itself indicate the predominance of the 1-H-tautomer for L-histidine in basic solutions.[180]

Tentative interpretation is made of ^{15}N chemical shift data for several 95%-enriched ^{15}N-amino-acids;[181] an indication of the pH-dependence of lineshapes, chemical shifts, nuclear Overhauser enhancements, and nuclear relaxation rates has been obtained for aqueous solutions of ^{15}N-glycine and its ethyl ester.[182]

O.R.D. and C.D. Spectra.—The 210 nm $n \rightarrow \pi^*$ Cotton effect, diagnostic of absolute configuration at the α-carbon atom in simple α-amino-acids (positive Cotton effect = L-configuration) has been used to assign absolute configuration to α-methyl-α-amino-acids prepared by asymmetric Strecker synthesis.[63] The same relationship between sign of Cotton effect and absolute configuration holds also for simple *N*-methyl-,[183] *NN*-dimethyl-L-amino-acids,[184] and L-amino-acid betaines.[78] Calculations using semi-empirical CNDO–MO methods for various conformations of L-alanine and related α-amino-acids give non-empirical support[185] to Jorgensen's sector rule,[186] but some parameters used in the calculations are not sufficiently accurate to permit more precise c.d.–conformation correlations.[185]

Prototype c.d. spectrometers are capable of reaching to *ca.* 160 nm,[187] and have shown the presence of a negative Cotton effect in the range 168—172 nm, in addition to features at longer wavelengths, for five aliphatic L-amino-acids.[187] Proline differs from open-chain analogues in showing sigmoid $n \rightarrow \pi^*$ c.d.[187] Thin films of aliphatic α-amino-acids show four absorption bands in the 140—200 nm wavelength region, deriving from transitions in the carboxylate chromophore;[188] band locations are somewhat structure-dependent.[188]

[177] B. C. Nicholson and T. M. Spotswood, *Austral. J. Chem.*, 1973, **26**, 135.
[178] J. T. Gerig and R. S. McLeod, *J. Amer. Chem. Soc.*, 1973, **95**, 5725.
[179] M. H. Freedman, J. R. Lyerla, I. M. Chaiken, and J. S. Cohen, *European J. Biochem.*, 1973, **32**, 215.
[180] W. F. Reynolds, I. R. Peat, M. H. Freedman, and J. R. Lyerla, *J. Amer. Chem. Soc.*, 1973, **95**, 328.
[181] J. A. Sogn, W. A. Gibbons, and E. W. Randall, *Biochemistry*, 1973, **12**, 2100.
[182] R. A. Cooper, R. L. Lichter, and J. D. Roberts, *J. Amer. Chem. Soc.*, 1973, **95**, 3724.
[183] J. Shoji, *J. Antibiotics*, 1973, **26**, 302.
[184] C. J. Hawkins and G. A. Lawrence, *Austral. J. Chem.*, 1973, **26**, 1801.
[185] J. Webb, R. W. Strickland, and F. S. Richardson, *Tetrahedron*, 1973, **29**, 2499.
[186] E. C. Jorgensen, *Tetrahedron Letters*, 1971, 863.
[187] P. A. Snyder, P. M. Vipond, and W. C. Johnson, *Biopolymers*, 1973, **12**, 975.
[188] T. Inagaki, *Biopolymers*, 1973, **12**, 1353.

The already extensive study of the c.d. of the perturbed indole chromophore, as it exists in L-tryptophan and its derivatives, is supplemented by extensive data on the series of negative Cotton effects in the 180—215 nm region.[189] There are four or five Cotton effects in this region, none of which arises by chromophore coupling between indole and carboxylate moieties. The magneto-c.d. spectra of tyrosine and tryptophan have been compared,[190] with special reference to pH-dependence.

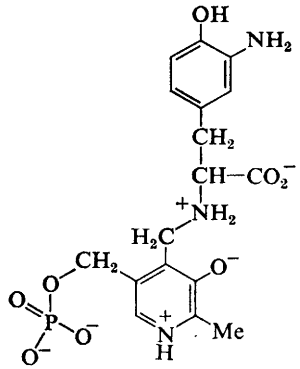
The c.d. spectra of several common sulphur-containing amino-acids have been studied in an attempt to correlate various features with conformation-dependent c.d. characteristics of the sulphur chromophores.[191] The temperature-dependent c.d. of *NN'*-diacetyl-L-cystine bis(dimethylamide) has been analysed in terms of the chirality of the disulphide chromophore.[192]

β-Amino-acids adopt different spatial relationships of carboxy chromophore to substituents at the asymmetric centre, compared with α-amino-acids, and a less secure relationship between sign of $n \rightarrow \pi^*$ Cotton effect and absolute configuration is found. The pH-dependence of the o.r.d. of compounds of this series is a satisfactory basis for absolute configurational assignments,[94] while the sign of the $n \rightarrow \pi^*$ Cotton effect of their *N*-dithioalkoxycarbonyl derivatives fails to correspond unambiguously with absolute configuration.

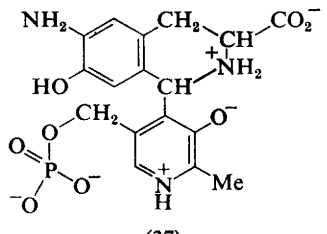
Hydantoins derived from L-α-amino-acids show a negative Cotton effect in the 230 nm wavelength region,[193] while *N*-methylthiohydantoins have some value in the assignment of absolute configuration to the α-amino-acids from which they are derived, provided that racemization can be avoided during their preparation.[194]

Combined Use of Physical Methods; Amino-acid Derivatives in Solution.—Solution conformation and antibody studies of the haptens *N*-(5-phosphopyridoxyl)-3'-amino-L-tyrosine (36) and its cyclic analogue (37) take account of fluorescence spectrometric, n.m.r., and c.d. data.[195] A thorough appraisal of the conformational behaviour in solution of N^{α}-acetylamino-acid *N*-methylamides (38) and proline analogues involves i.r.,[196] n.m.r.,[197] dipole moment,[198] effective molecular weight,[199] and u.v., c.d., and o.r.d.

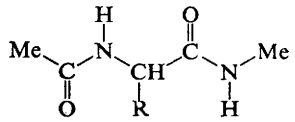
[189] H. E. Auer, *J. Amer. Chem. Soc.*, 1973, **95**, 3003.
[190] M. Gabriel, D. Larchier, H. Rinnert, and C. Thirion, *Compt. rend.*, 1973, **276**, *B*, 39.
[191] G. Jung, M. Ottnad, and M. Rimpler, *European J. Biochem.*, 1973, **35**, 436.
[192] T. Takagi, R. Okano, and T. Miyazawa, *Biochim. Biophys. Acta*, 1973, **310**, 11.
[193] T. Suzuki, K. Igarashi, K. Hase, and K. Tuzimura, *Agric. and Biol. Chem.* (*Japan*), 1973, **37**, 411.
[194] T. Suzuki and K. Tuzimura, *Agric. and Biol. Chem.* (*Japan*), 1973, **37**, 689.
[195] V. Raso and B. D. Stollar, *J. Amer. Chem. Soc.*, 1973, **95**, 1621.
[196] E. S. Efremov, L. B. Senyavina, V. N. Zheltova, A. G. Ivanova, P. V. Kostetskii, V. T. Ivanov, E. M. Popov, and Y. A. Ovchinnikov, *Khim. prirod Soedinenii*, 1973, **9**, 322.
[197] V. T. Ivanov, P. V. Kostetskii, T. A. Balashova, S. L. Portnova, E. S. Efremov, and Y. A. Ovchinnikov, *Khim. prirod Soedinenii*, 1973, **9**, 339.
[198] E. S. Efremov, P. V. Kostetskii, V. T. Ivanov, E. M. Popov, and Y. A. Ovchinnikov, *Khim. prirod Soedinenii*, 1973, **9**, 348.
[199] E. S. Efremov, P. V. Kostetskii, V. T. Ivanov, E. M. Popov, and Y. A. Ovchinnikov, *Khim. prirod Soedinenii*, 1973, **9**, 354.



(36)



(37)



(38)

studies.[200] N^{α}-Acetyl-N^{α}-methyl-alanine *N*-methylamide exists predominantly in the *trans*-acetamide form in all solvents, but 15% in the *cis*-form in CCl_4 and 26% in the *cis*-form in $(C^2H_3)_2SO$.[197] The i.r. study reveals a high degree of coiling through intramolecular hydrogen-bonding in $CHCl_3$ and CCl_4,[196] while osmometry[199] indicates a tendency to form hydrogen-bonded dimers in concentrated solutions.

Mass Spectrometry.—Pyrolytic fragmentation of amino-acid betaines $Me_3N^+(CH_2)_nCO_2^-$ (n = 1—5) in the mass spectrometer has been rationalized in terms of formation of $Me_2N^+{=}CH_2$ (base peak at *m/e* 58) together with cyclized fragments depending upon structure.[201] Identification of ten of the amino-acids present in soil samples by quadrupole mass spectrometry has been demonstrated for nanogram quantities of their *N*-trifluoroacetyl n-butyl esters, with results comparable with those from the amino-acid analyser.[202] Chemical ionization mass spectra of amino-acids and peptides and their ester and amide derivatives have been reported using either isobutane[203] or methane[204] as reactant plasma. Free amino-acids show loss of HCO_2H with isobutane at 440—630 °C, in addition to fragments MH^+, M_2H^+, $(M + 57)^+$, and $(M + 39)^+$;[203] lower temperatures cause insufficient fragmentation.

[200] V. T. Ivanov, P. V. Kostetskii, E. A. Meshcheryakova, E. S. Efremov, E. M. Popov, and Y. A. Ovchinnikov, *Khim. prirod. Soedinenii*, 1973, **9**, 363.
[201] K. Undheim and T. Laerum, *Acta Chem. Scand.*, 1973, **27**, 589.
[202] W. E. Pereira, Y. Hoyano, W. E. Reynolds, R. E. Summons, and A. M. Duffield, *Analyt. Biochem.*, 1973, **55**, 236.
[203] M. Meot-Ner and F. H. Field, *J. Amer. Chem. Soc.*, 1973, **95**, 7207.
[204] P. A. Leclercq and D. M. Desiderio, *Org. Mass Spectrometry*, 1973, **7**, 515.

Where amino-acid sequences in peptides are not obtainable by mass spectrometric study of derivatized peptides, then mass spectrometry can be employed to identify the amino-acid derivative released in the cleavage step of a conventional peptide-sequencing procedure. The last residue in the Edman degradation of a peptide is obtained as its *N*-phenylthiocarbamoyl derivative, or as its methyl analogue, and these isomerize thermally in the mass spectrometer to corresponding *N*-phenyl- or *N*-methyl-thiohydantoins, and can be identified as such.[205] Neopentylidene amino-acid ethyl esters are useful volatile derivatives for mass spectrometry of amino-acids.[15]

Determination of Absolute Configuration.—In addition to results noted in preceding paragraphs, assignment of *S*-configuration to (+)-2-amino-2-phenylbutyric acid by chemical correlation with (−)-(*S*)-2-amino-2-phenylbutane has been confirmed by molecular rotation data,[206] and L-configuration has been assigned to the β-centre in the β-methyllanthionine component of nisin.[207] *N*-Methyl-alloisoleucine found in quinoxaline antibiotics is the L-enantiomer, as found in the Actinomycins.[183]

5 Chemical Studies of Amino-acids

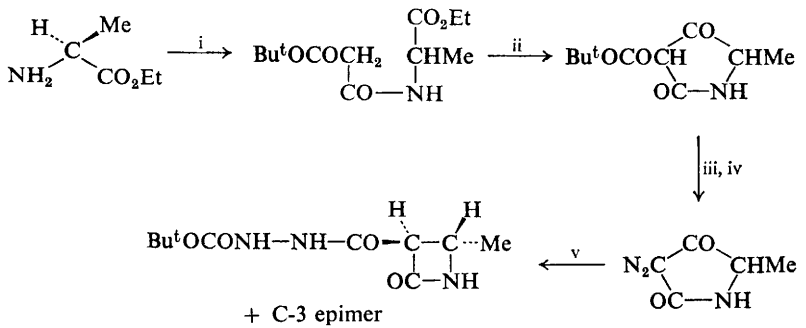
Racemization and Inversion.—The stirring of interest last year in the geochronological information which can be inferred from the degree of racemization suffered by fossil amino-acids has been developed into a subject in its own right with its own jargon. The half-life for the racemization of L-isoleucine giving D-allo-isoleucine (more than 100 000 years at ambient temperature) has been determined as a function of temperature; if the age of a fossil can be assessed by radiocarbon dating, then the average temperature to which the fossil has been subjected can be calculated from the L-isoleucine : D-allo-isoleucine ratio. That for a particular bone of age 40 000 years is between 0.42 and 0.46, suggesting that the average temperature during this period at the site of the fossil was 26.5°C (the average temperature in recent times is 28 °C).[208] Rates of racemization of amino-acids in modern bone samples are in the order aspartic acid > alanine ~ glutamic acid > isoleucine ~ leucine,[209] so that more accurate data should be obtainable for paleotemperature studies with aspartic acid; the rise in temperature during the Earth's last glaciation period was about 4 °C for the Mediterranean coast, and 5—6 °C in East Africa, based upon the amount of D-aspartic acid found in fossil bones from these regions.[210] Temperatures deduced are claimed to be reliable to within 1 °C, and,

[205] T. Fairwell, S. Ellis, and R. E. Lovins, *Analyt. Biochem.*, 1973, **53**, 115.
[206] J. A. Garbarino, J. Sierra, and R. Tapia, *J.C.S. Perkin I*, 1973, 1866.
[207] J. L. Morell and E. Gross, *J. Amer. Chem. Soc.*, 1973, **95**, 6480.
[208] J. L. Bada, R. Protsch, and R. A. Schroeder, *Nature*, 1973, **241**, 394.
[209] J. L. Bada and R. Protsch, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1331.
[210] R. A. Schroeder and J. L. Bada, *Science*, 1973, **182**, 479.

although the 'right' answers are being derived from these studies, uncertainties concerning solid-state catalysis or inhibition of the racemization process may need to be considered.

Conversion of (+)-*R*-α-methylphenylalanine into its enantiomer has been achieved[211] *via* the *R*-isocyanide after conversion of the carboxy-group into $-CH_2OAc$; the *R*-isocyanide gives the *S*-cyanide with 75.6% retention when heated in diphenyl ether at 280 °C, and further elaboration gives the (−)-*S*-amino-acid.

General Reactions.—Pride of place should go to an efficient synthesis of optically active *N*-benzyloxycarbonyl-aziridin-2-ones from *N*-benzyloxycarbonyl-L-amino-acids, using phosgene, thionyl chloride, or phosphorus oxychloride in THF at −20 °C to −30 °C, with addition of triethylamine to maintain neutral conditions.[212] Some of the derivatives are crystalline; they may be used in peptide synthesis, aminolysis proceeding without racemization.[213] The other small-ring system of major importance is azetidin-2-one (β-lactam); the series is accessible through addition of *N*-benzyloxycarbonylglycyl chloride to imines in the presence of triethylamine,[214] or through a novel ring-contraction route (illustrated in Scheme 13 for the 4-methyl series obtainable from L-alanine ethyl ester).[215]



Reagents: i, $Bu^tO_2CCH_2CO_2H$–DCCI; ii, Bu^tOK; iii, heat in xylene, 1.5 h; iv, $MeSO_2N_3$–NEt_3, −10 °C; v, *hν*, $Bu^tO_2CNHNH_2$

Scheme 13

Reactions at the α-amino- or α-imino-group of representative compounds include *N*-nitrosation of L-proline and other imino-acids with nitrosyl tetrafluoroborate,[216] and conversion of phenylalanine into its α-diazo-ester

[211] M. Shibasaki, S. Terashima, and S. Yamada, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 552.
[212] M. Miyoshi, *Bull. Chem. Soc. Japan*, 1973, **46**, 212.
[213] M. Miyoshi, *Bull. Chem. Soc. Japan*, 1973, **46**, 1489.
[214] A. K. Bose, H. P. S. Chawla, B. Dayal, and M. S. Manhas, *Tetrahedron Letters*, 1973, 2503.
[215] G. Lowe and D. D. Ridley, *J.C.S. Perkin I*, 1973, 2024.
[216] H. T. Nagasawa, P. S. Fraser, and D. L. Yuzon, *J. Medicin. Chem.*, 1973, **16**, 583.

en route to *cis*- or *trans*-cinnamate esters, for which a stereoselective route has been demonstrated; treatment of the diazo-ester with a sodium alkoxide gives the *trans*-series whereas the *cis*-series is reached using boron trifluoride diethyletherate.[217] Silylurethanes $Me_3SiOCONH \cdot CHR^1CO_2R^2$ are obtained by reaction between an amino-acid ester and t-butyl trimethylsilyl carbonate;[218] the same reaction under forcing conditions gives corresponding trimethylsilyl esters.

Side-chain studies include reaction of alkylating agents with *N*-acetyl methylamides of amino-acids with side-chain nucleophilic centres,[219] and intramolecular reactions accompanying N-protection of 4-chloro- and 4-hydroxy-lysines.[220]

Significant observations have been made concerning the formation of peptide esters through the condensation of amino-acid esters in the presence of Cu^{II} [221] or Pt^{II}.[222]

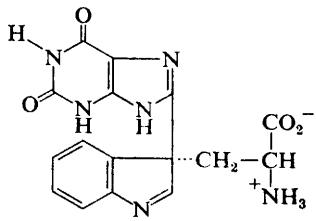
Specific Reactions and Interactions.—Discussion of side-chain protection and deprotection procedures with amino-acids is appropriate in this chapter where there is a broad significance in the chemistry involved. Removal of Boc groups with trifluoroacetic acid from protected peptides is commonly conducted in the presence of mercaptoethanol to prevent modification of tryptophan residues, but 1,2-ethanedithiol is superior for the purpose.[223] *S*-(*p*-Methoxybenzyl)-protected cysteine residues elsewhere in the protected peptide are unaffected by this Boc-removal procedure. *O*-(*o*-Bromobenzyloxycarbonyl) protection for tyrosine is recommended (removal by HF),[224] *O*-benzyl-protection being unsatisfactory since Boc-removal at the *N*-terminus can be accompanied by conversion of the tyrosine residue into 3-benzyl-tyrosine.[225] Blocking of the guanidino function by N^{δ},N^{ω}-bis(isobornyloxycarbonyl)ation is advocated when the arginine side-chain must be protected.[226]

A kinetic study of deuteriation at position 2 of the imidazole ring of histidine and related compounds, as models for the exchange behaviour in proteins, shows that the NH_3^+ group increases the rate of deuteriation three times compared with that for glycylhistidine, and seven times compared with the rate for histidine in alkaline solution (unprotonated NH_2).[227] Mono-*N*-hydroxymethylation of the imidazole ring of *N*-acetyl-

[217] N. Takamura, T. Mizoguchi, and S. Yamada, *Tetrahedron Letters*, 1973, 4267.
[218] Y. Yamamoto, D. S. Tarbell, J. R. Fehlner, and B. M. Pope, *J. Org. Chem.*, 1973, **38**, 2521.
[219] C. C. Price, H. Akimoto, and R. Ho, *J. Org. Chem.*, 1973, **38**, 1538.
[220] S. Clarke, R. C. Hider, and D. I. John, *J.C.S. Perkin I*, 1973, 230.
[221] S. Terashima, M. Wagatsuma, and S. Yamada, *Tetrahedron*, 1973, **29**, 1487, 1497.
[222] W. Beck, B. Purucker, and E. Strissel, *Chem. Ber.*, 1973, **106**, 1781.
[223] J. J. Sharp, A. B. Robinson, and M. D. Kamen, *J. Amer. Chem. Soc.*, 1973, **95**, 6097.
[224] D. Yamashiro and C. H. Li, *J. Org. Chem.*, 1973, **38**, 591.
[225] B. W. Erickson and R. B. Merrifield, *J. Amer. Chem. Soc.*, 1973, **95**, 3750.
[226] G. Jager and R. Geiger, *Annalen*, 1973, 1928.
[227] J. H. Bradbury, B. E. Chapman, and F. A. Pellegrino, *J. Amer. Chem. Soc.*, 1973, **95**, 6139.

histidine is brought about with formaldehyde in alkaline solution; further *N*-hydroxymethylation can occur in acidic solution.[228]

L-Tryptophan forms a complex with inosine, with equilibrium constant 10^3 larger than that for other tryptophan nucleoside complexes;[229] L-tryptophan reacts with the powerful oncogen 3-acetoxy-xanthine to give the condensation product (39) and its C-3 epimer,[230] adding a further

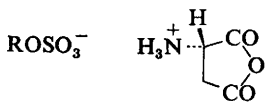


(39)

example to the list of compounds potentially capable of modifying this amino-acid irreversibly when it is a component of peptides and proteins. *trans*-4,5-Dehydro-lysine, prepared from ethyl *trans*-2-acetamido-2-ethoxycarbonyl-6-phthalimidohex-4-enoate, resists lactonization in acidic solution,[231] correcting a report that diastereoisomeric 4-hydroxy-DL-lysine lactones are formed during attempted synthesis.

Difficulties in the preparation of phenylthiohydantoins of β-hydroxy-α-amino-acids were noted in the early development of the Edman sequencing procedure; β-elimination during the preparation of methylthiohydantoins can be avoided by keeping the pH of solutions high.[232]

The latest improved route to L-aspartic acid anhydride uses concentrated sulphuric acid in ethyl acetate at room temperature during 48 h as a means of taking L-aspartic acid into solution, and is followed by treatment with an alcohol, ROH, then cyclization with acetic anhydride; the anhydride is isolated as its alkyl hydrogen sulphate salt (40).[233]

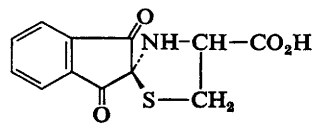


(40)

[228] P. Dunlop, M. A. Marini, H. M. Fales, E. Sokoloski, and C. J. Martin, *Bio-org. Chem.*, 1973, **2**, 235.
[229] I. Ibanez, M. Pieber, and J. Toha, *Z. Naturforsch.*, 1973, **28c**, 385.
[230] G. Stohrer, G. Salemnick, and G. B. Brown, *Biochemistry*, 1973, **12**, 5084.
[231] A. L. Davis, M. B. Cavitt, T. J. McCord, P. E. Vickrey, and W. Shive, *J. Amer. Chem. Soc.*, 1973, **95**, 6800.
[232] M. M. Amirkhanyan and V. M. Stepanov, *Zhur. obshchei Khim.*, 1973, **43**, 1630.
[233] Y. Ariyoshi, T. Yamatani, and Y. Adachi, *Bull. Chem. Soc. Japan*, 1973, **46**, 2611.

Sulphur functional groups in amino-acid side-chains have received attention. Oxidation of methionine to the sulphone with trichloroisocyanuric acid is rapid, but all other amino-acids are more or less rapidly degraded by this reagent.[234] Aqueous chlorine or bromine reacts with cysteine, cystine, alanine-3-sulphinic acid, and cystine *SS*-dioxide to give cysteic acid and cysteinylcysteic acid.[235] L-Methionine gives the L$^{\alpha}$-(*S*)-sulphoxide with an equimolar amount of $HAuCl_4$ within a few minutes;[236] the mechanism of this stereospecific oxidation involving $Au^{III} \rightarrow Au^{I}$ is not clear, but rapid formation of a methionine–$AuCl_3$ complex may be followed by stereospecific attack of a second molecule of methionine before oxidation. Hydrolysis rates of *NS*-diacetylcysteinamide and *N*-acetyl-*S*-benzoylcysteinamide are some twenty times faster than expected on the basis of p*K* values, and the rate for *N*-benzyloxycarbonyl-*S*-acetyl-L-cysteinyl-L-threonine ethyl ester is five to six times faster still, implying anchimeric assistance by functional groups neighbouring the thiolester grouping.[237] Electrochemical reduction of *S*-methyl-methionine gives mainly α-amino-butyric acid and methionine, with α-amino-γ-butyrolactone as minor product.[238]

The ninhydrin colour reaction is a reliable work-horse for amino-acid analysis, but there are still aspects for study, such as the dependence of colour formation upon hydrindantin concentration.[239] The hitherto accepted 1,4-thiazine structure for the cysteine–ninhydrin condensation product must now be replaced by the spiro-isomer (41);[240] analogous products are formed with penicillamine and cysteamine.



(41)

Effects of Electromagnetic Radiation on Amino-acids.—Recent work in this area concerns photochemical transformations of aromatic amino-acids, particularly tryptophan and tyrosine, the processes being relevant to possible mechanisms for the photoinactivation of enzymes. Photo-oxidation of tyrosine and histidine in the presence of tryptophan is shown to involve singlet oxygen, with formylkynurenin as the most probable photosensitizer.[241] Photo-oxidation of tryptophan is dependent upon

[234] M. Z. Atassi, *Tetrahedron Letters*, 1973, 4893.
[235] P. G. Gordon, *Austral. J. Chem.*, 1973, **26**, 1771.
[236] E. Bordignon, L. Cattalini, G. Natile, and A. Scatturin, *J.C.S. Chem. Comm.*, 1973, 878.
[237] D. G. Clark and E. H. Cordes, *J. Org. Chem.*, 1973, **38**, 270.
[238] T. Iwasaki, M. Miyoshi, M. Matsuoka, and K. Matsumoto, *Chem. and Ind.*, 1973, 1163.
[239] P. J. Lamothe and P. G. McCormick, *Analyt. Chem.*, 1973, **45**, 1906.
[240] G. Prota and E. Ponsiglione, *Tetrahedron*, 1973, **29**, 4271.
[241] P. Walraut, R. Santus, and M. Bazin, *Compt. rend.*, 1973, **276**, *C*, 149.

solvent, and the effects of co-solutes (3,4-benzpyrene, caffeine, sodium dodecyl sulphate) have been studied.[242] Photoconversion of tyrosine into bityrosine can be followed by fluorescence intensity changes at 400 nm, permitting the different mechanisms operating in acidic and alkaline solutions to be studied.[243] Data on quantum yields and fluorescence lifetimes of tyrosine and tryptophan in ^2H_2O–H_2O and glycerol–H_2O have been reported;[244] a comparative study of the photochemistry of phenylalanine, tyrosine, and dopa shows phenylalanine to be the most easily photolysed in dilute aqueous solution at 254 nm, giving alanine, glycine, and four other amines in the absence of air, and giving tyrosine, and its *o*- and *m*-isomers, under aeration.[245] Tyrosine and dopa, under the latter conditions, are converted into melanin.[245]

Photodecarboxylation of *N*-phthaloyl amino-acids in acetone solution,[246] and of amino-acids themselves in the presence of transition-metal ions,[247] has been studied. The range of products (stereoisomers of propenyl sulphides) formed from *S*-(*cis*-1-propenyl)-L-cysteine under γ-irradiation in oxygen-free aqueous solutions[248] is different from that (numerous thiophens, aldehydes, thiols) formed under u.v.-irradiation,[249] although 1-propenylthiyl radicals are implicated in both processes.

6 Analytical Methods

As is customary in this Section, many of the reports are cited without discussion, since they derive from methodology established in recent years, which will be generally familiar to readers.

Gas–Liquid Chromatography.—The range of uses to which quantitative g.l.c. is now being applied illustrates the slowly evaporating conservatism which has prevented many analysts from considering anything but their ion-exchange amino-acid analyser for routine amino-acid assay. Comparisons of results from the two techniques are reported and the special potential of g.l.c. in the determination of enantiomeric composition of small samples is illustrated.

Volatile derivatives of amino-acids in current use are *N*-trifluoroacetyl methyl esters[250, 251] or n-butyl esters[6, 202, 252–254] following earliest prece-

[242] G. Reske and H. Bauer, *Z. Naturforsch.*, 1973, **28c**, 390.
[243] O. Shimizu, *Photochem. and Photobiol.*, 1973, **18**, 125.
[244] R. McGuire and I. Feldman, *Photochem. and Photobiol.*, 1973, **18**, 119.
[245] C. Hasselmann and G. Laustriat, *Photochem. and Photobiol.*, 1973, **17**, 275.
[246] Y. Sato, H. Nakai, T. Mizoguchi, M. Kawanishi, and Y. Kanaoka, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 1164.
[247] R. Poupko, I. Rosenthal, and D. Elad, *Photochem. and Photobiol.*, 1973, **17**, 395.
[248] H. Nishimura and J. Mizutani, *Agric. and Biol. Chem. (Japan)*, 1973, **37**, 213.
[249] T. Hanzawa, H. Nishimura, and J. Mizutani, *Agric. and Biol. Chem. (Japan)*, 1973, **37**, 2393; *Tetrahedron Letters*, 1973, 343.
[250] W. S. Gardner and G. F. Lee, *Environ. Sci. Technol.*, 1973, **7**, 719.
[251] A. J. Cliffe, N. J. Berridge, and D. R. Westgarth, *J. Chromatog.*, 1973, **78**, 333.
[252] C. W. Gehrke and H. Takeda, *J. Chromatog.*, 1973, **76**, 63.
[253] H. Hediger, R. L. Stevens, H. Brandenberger, and K. Schmid, *Biochem. J.*, 1973, **133**, 551.
[254] C. W. Gehrke and H. Takeda, *J. Chromatog.*, 1973, **76**, 77.

dent, though *N*-heptafluorobutyroyl n-propyl esters,[255] isoamyl esters,[256] and *N*-acetyl n-butyl esters [257] have their proponents. The twenty protein amino-acids can be separated on a single column with Apiezon M as stationary phase.[252] Use of 2.7M-HCl in n-butanol as an anhydrous reagent for esterification of amino-acids has been studied; it is the routine method as part of the derivatization procedure, but should not be conducted above 100 °C [258] since substantial amounts of water, chlorobutane, and dibutyl ether are formed in the reagent at higher temperatures. Pyrolysis of *N*-neopentylidene amino-acid trimethylanilinium salts in the injector port of the gas chromatograph gives volatile derivatives characteristic of the parent amino-acid;[259] Schiff bases of amino-acid esters with benzaldehyde or with pentane-2,4-dione have been proposed [260] for quantitative g.l.c. Acetylation of phenylthiohydantoins gives derivatives more suited to g.l.c. analysis,[261] while silylated thiohydantoins are also readily identified by g.l.c. and by mass spectometry.[262] Short glass capillary columns permit the separation of silylated methylthiohydantoins and phenylthiohydantoins, nineteen of twenty protein amino-acids being separated during 35 min in one pass when cysteine is *S*-methylated and arginine converted into ornithine in the sample preparation sequence.[263] The twentieth amino-acid, histidine, can be eluted from the column as its silylated *N*-methyl- or *N*-phenyl-thiohydantoin using a rather higher oven temperature.[263] Trimethylsilyl derivatives of glycine, lysine, and serine are suitable for g.l.c.–mass spectrometric analysis,[264] and comparison of results obtained by this technique with those of conventional amino-acid analysis for these amino-acids in Devonian black slate has been reported.[264]

Improvements in the g.l.c. technique have been proposed. Losses during sample preparation and transfer may be avoided using an aluminium or gold micro-reactor for all steps.[265] Problems with tryptophan have been discussed,[254] and qualitative use of the technique to identify asparagine, glutamine, and pyroglutamic acid in total enzymic hydrolysates, distinguishing them from their parent amino-acids by esterification rates, has been explored.[253] The sensitivity of the method (0.1 ng) [256] is well established, though progress still has to be made in terms of simpler, possibly automated, sample preparation procedures.

Separation of enantiomers of amino-acids as their *N*-trifluoroacetylisopropyl esters using *N*-trifluoroacetyl-L-α-aminobutyryl-L-α-aminobutyric

[255] J. Jonsson, J. Eyem, and J. Sjoquist, *Analyt. Biochem.*, 1973, **51**, 204.
[256] J. P. Zanetta and G. Vincendon, *J. Chromatog.*, 1973, **76**, 91.
[257] P. G. Vincent and J. Kirksey, *J. Assoc. Offic. Analyt. Chemists*, 1973, **56**, 158.
[258] J. P. Hardy, S. L. Kerrin, and S. L. Manatt, *J. Org. Chem.*, 1973, **38**, 4196.
[259] K. M. Williams and B. Halpern, *Analyt. Letters*, 1973, **6**, 839.
[260] P. W. D. Mitchell, *J. Chromatog.*, 1973, **76**, 236.
[261] A. S. Inglis and P. W. Nicholls, *J. Chromatog.*, 1973, **86**, 117.
[262] M. Rangarajan, R. E. Ardrey, and A. Darbre, *J. Chromatog.*, 1973, **87**, 499.
[263] J. Eyem and J. Sjoquist, *Analyt. Biochem.*, 1973, **52**, 255.
[264] W. Heller, W. A. Koenig, and K. Schmidt, *Chromatographia*, 1973, **6**, 327.
[265] B. Kolb and W. Hoser, *Chromatographia*, 1973, **6**, 28.

acid cyclohexyl ester as stationary phase has been demonstrated;[266] correspondingly substituted L-alanine, L-norvaline, or L-norleucine dipeptides were less satisfactory as optically active stationary phases. The first successful use of an optically active mesophase, smectic carbonyl bis-(D-leucine isopropyl ester) at 55—100 °C, for g.l.c. separation of enantiomers has been reported.[267] The alternative approach, conversion of the DL-amino-acid into a diastereoisomeric derivative, *e.g.* its *N*-trifluoroacetyl s-butyl ester,[268] has been further exploited in a novel quantitative analysis of the extent of degradation of solid amino-acid samples by photolysis, radiolysis, or electron bombardment; a known weight of one enantiomer is added and the mixture converted into a volatile diastereoisomer mixture for g.l.c. analysis.[269]

Ion-exchange and Partition Chromatography.—Progress continues towards automation [270–272] and computer-controlled data-acquisition [271] of the ion-exchange amino-acid analyser. Single-column sub-micro analysis of all protein amino-acids can be achieved with a pH-gradient technique,[272, 273] or by employing three sodium citrate buffers [274, 275] (*e.g.* [275] pH 5.25 for the third buffer) to give improved resolution of basic amino-acids. Lithium citrate buffers permit the separation of more than fifty amino-acids on a single column.[276] Modifications to standard techniques permit quantitative analysis of N^{ϵ}-methyl-lysine [277] and other methylated basic amino-acids,[19] and permit simultaneous analysis of three samples in a little over 6 h.[278] A modified commercial amino-acid micro-analyser can be operated with picomole amounts of amino-acids.[279]

An amino-acid analyser based upon the fluorescamine detection method, which is two orders more sensitive than the ninhydrin colorimetric method,[280] has been developed. The procedure has been illustrated with one microgram of protein.[274, 280] The failure of imino-acids (prolines) to respond directly to the fluorescence-forming reaction can be circumvented by their oxidative decarboxylation into imines with halogenating

[266] W. Parr and P. Y. Howard, *Analyt. Chem.*, 1973, **45**, 711.
[267] C. H. Lochmuller and R. W. Souter, *J. Chromatog.*, 1973, **87**, 243.
[268] F. Raulin and B. N. Khare, *J. Chromatog.*, 1973, **75**, 13.
[269] W. A. Bonner, *J. Chromatog. Sci.*, 1973, **11**, 101.
[270] A. M. C. Davies, *Lab. Practice*, 1973, **22**, 627.
[271] C. P. Hohberger, B. Soucek, R. L. Chase, and D. Potter, *Brookhaven National Laboratory Report* BNL-17677, 1973; *Nuclear Sci. Abs.*, 1973, **28**, 7626.
[272] H. Tschesche, C. Frank, and H. Ebert, *J. Chromatog.*, 1973, **85**, 35.
[273] J. L. Young and M. Yamamoto, *J. Chromatog.*, 1973, **78**, 349.
[274] A. G. Georgiadis and J. W. Coffey, *Analyt. Biochem.*, 1973, **56**, 121.
[275] L. G. Gurtler, *J. Chromatog.*, 1973, **76**, 255.
[276] J. L. Young and M. Yamamoto, *J. Chromatog.*, 1973, **78**, 221.
[277] H. W. Lange, R. Lower, and K. Hempel, *J. Chromatog.*, 1973, **76**, 252.
[278] J. P. Ellis and J. B. Garcia, *J. Chromatog.*, 1973, **87**, 419.
[279] A. M. Gressner, *Analyt. Biochem.*, 1973, **56**, 532.
[280] S. Stein, P. Bohlen, J. Stone, W. Dairman, and S. Udenfriend, *Arch. Biochem, Biophys.*, 1973, **155**, 203.

agents;[281, 282] these give primary amines on hydrolysis, which are susceptible to the fluorescamine reaction. Thus, all natural amino-acids can now be analysed using the fluorimetric amino-acid analyser, and wide adoption of the technique can be expected. Modified fluorescence reactions are being considered; *e.g.* reaction of the amino-acid mixture with *o*-phthalaldehyde and 2-mercaptoethanol.[283]

A rapid ligand-exchange chromatographic method for the separation of amino-acids from peptides, employing Chelex 100, should be useful for studies of partial hydrolysates or for solving problems due to incomplete acid hydrolysis.[284] The effect which residual hydrochloric acid can have in modifying elution times and resolution of many amino-acids on ion-exchange columns has been demonstrated.[285]

Liquid chromatography of amino-acids in aqueous solution is feasible using poly(glycine) bonded to resin-coated glass beads, Porasil C, or Corasil II as stationary support.[286]

Thin-layer and Paper Chromatography.—The use of preparative t.l.c. on 1 mm cellulose layers for isolating amino-acids from mixtures is illustrated well for the isolation of cyclopent-2-en-1-yl glycine, a potent growth inhibitor from *Hydnocarpus anthelminthica*.[287]

A stringent test of analytical chromatographic methods is the separation of leucine and isoleucine. Ascending paper chromatography using n-butanol–acetic acid–water (4 : 1 : 5), the doyen of solvents for amino-acid paper chromatography, can separate these amino-acids if a temperature gradient of 5.88 °C cm^{-1} is applied;[288] their dansyl derivatives can be resolved on silica gel layers by multiple development with $CHCl_3$–MeOH (95 : 5).[289] Chromatographic data for ion-exchange and paper chromatography of *N*-methylamino-acids have been reported.[290]

Phenylthiohydantoins can be separated by short-run two-dimensional t.l.c. on silica gel and detected on t.l.c. plates at 10^{-10} mol l^{-1} levels.[291] Separations of histidine and arginine phenylthiohydantoins [292] and diphenylindonyl-substituted thiohydantoins [293] have been described. Dansylation of amino-acid mixtures, followed by two-dimensional polyamide t.l.c.

[281] M. Weigele, S. de Bernardo, and W. Leimgruber, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 352.
[282] A. M. Felix and G. Terkelsen, *Analyt. Biochem.*, 1973, **56**, 610; *Biochem. Biophys. Res. Comm.*, 1973, **50**, 352.
[283] M. Roth and A. Hampai, *J. Chromatog.*, 1973, **83**, 353.
[284] J. F. Bellinger and N. R. M. Buist, *J. Chromatog.*, 1973, **87**, 513.
[285] H. D. Spitz, *Analyt. Biochem.*, 1973, **56**, 66.
[286] E. Grushka and R. P. W. Scott, *Analyt. Chem.*, 1973, **45**, 1627.
[287] F. Spener and M. Dieckhoff, *J. Chromatog. Sci.*, 1973, **11**, 661.
[288] C. Liteanu and A. Constantinescu, *Rev. Roumaine Chim.*, 1973, **18**, 155.
[289] R. S. Fager and C. B. Kutina, *J. Chromatog.*, 1973, **76**, 268.
[290] T. K. Audhya and D. W. Russell, *J. Chromatog.*, 1973, **84**, 361.
[291] M. C. Solal and J. L. Bernard, *J. Chromatog.*, 1973, **80**, 140.
[292] T. Inagami, *Analyt. Biochem.*, 1973, **52**, 318.
[293] C. P. Ivanov and I. N. Mancheva, *J. Chromatog.*, 1973, **75**, 129.

and elution and quantitation by fluorimetry and scintillation counting, is sensitive to 3×10^{-12} mol l^{-1} levels.[294]

Other Analytical Methods.—Fluorophotometric methods continue to show great promise, and recent papers, in addition to those cited in preceding paragraphs, cover highly fluorescent derivatives obtained from amino-acids by treatment with 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole,[295] and with pyridoxal.[296] However, the latter technique is not significantly more sensitive than ninhydrin colorimetry, and is not suitable for assay of proline, hydroxyproline, or tryptophan.[296]

A kinetic method based upon the retardation by an amino-acid of the Cu^{II}-catalysed oxidation of catechol violet at pH 8.65 with hydrogen peroxide allows several amino-acids to be assayed at the 2 μmol level.[297] At the other extreme, 3—8 mg samples of amino-acids can be assayed by treatment with 2-nitrobenzenesulphenyl thiocyanate, followed by iodometric titration of the released thiocyanate ion.[298] A microbiological assay [299] and further studies of assays based on the potential developed in a cell with immobilized L-amino-acid oxidase electrodes [300] have been reported.

An isotope dilution principle is employed for the determination of amino-acids in biological samples, where the amino-acid to be determined competes with the added labelled amino-acid for binding to a tRNA, catalysed by an amino-acyl tRNA-synthetase.[301] Independent studies have been reported [302, 303] of a sensitive double-isotope derivative assay using ^{3}H-labelled dansyl chloride and ^{14}C-labelled amino-acids as internal standards, so that the $^{3}H : {}^{14}C$ ratio in the dansyl derivatives depends upon the ratio of added ^{14}C-labelled amino-acid to the natural amino-acid concentration in the sample. The method needs no more than 20 pmol of amino-acid,[302] and has been used [303] to measure the release of amino-acids from microgram samples of brain tissue.

The most sensitive method yet reported appears to be the least likely to be routinely applied; tunnelling measurements of vibrational spectra of amino-acids give characteristic features with less than a monolayer of sample on an area 0.1—1.0 mm^2.[304]

[294] J. Airhart, S. Sibiga, H. Sanders, and E. A. Khairallah, *Analyt. Biochem.*, 1973, **53**, 132.
[295] R. S. Fager, C. B. Kutina, and E. W. Abrahamson, *Analyt. Biochem.*, 1973, **53**, 290.
[296] M. Maeda, A. Tsuji, S. Ganno, and Y. Onishi, *J. Chromatog.*, 1973, **77**, 434.
[297] T. J. Janjic and G. A. Milovanovic, *Analyt. Chem.*, 1973, **45**, 390.
[298] S. I. Obtemperanskaya, N. N. Kalinina, and M. N. Sizoi, *Zhur. analit. Khim.*, 1973, **28**, 399.
[299] H. Itoh, K. Kawashima, and I. Chibata, *Agric. and Biol. Chem. (Japan)*, 1973, **37**, 2227.
[300] G. G. Guilbault and G. Nagy, *Analyt. Letters*, 1973, **6**, 301.
[301] R. Parrilla, M. S. Ayuso-Parrilla, and M. N. Goodman, *Analyt. Biochem.*, 1973, **54**, 362.
[302] J. P. Brown and R. N. Perham, *European J. Biochem.*, 1973, **39**, 69.
[303] S. R. Snodgrass and L. L. Iversen, *Nature New Biol.*, 1973, **241**, 154.
[304] M. G. Simonsen and R. V. Coleman, *Nature*, 1973, **244**, 218.

Determination of Specific Amino-acids.—Titration methods have been reported for glutamic acid [305] and cystine,[306] and fluorimetric assays for dopa [307] and phenylalanine.[308] The dopa assay (1 ng ml^{-1}) is one hundred times more sensitive than existing methods. Determinations of cysteine or cystine in intracellular fluids have been described,[309] and details given for microbiological assay of the antitumour agent L-alanosine;[310] L-canaline [311] and *O*-ureido-L-homoserine [312] have been studied from the aspects of isolation, synthesis, and assay.

[305] K. Bahadur and I. Saxena, *Microchem. J.*, 1973, **18**, 358.
[306] J. Willemot and G. Parry, *Ann. pharm. franç*, 1973, **31**, 249.
[307] J. C. Johnson, G. J. Gold, and D. H. Clouet, *Analyt. Biochem.*, 1973, **54**, 129.
[308] T. M. Andrews, R. Goldthorp, and R. W. E. Watts, *Clinica Chim. Acta*, 1973, **43**, 379.
[309] B. States and S. Segal, *Clinica Chim. Acta*, 1973, **43**, 49.
[310] R. F. Pittillo and C. Woolley, *Antimicrobial Agents and Chemotherapy*, 1973, **3**, 739.
[311] G. A. Rosenthal, *Analyt. Biochem.*, 1973, **51**, 354.
[312] G. A. Rosenthal, *Analyt. Biochem.*, 1973, **56**, 435.

2

Structural Investigations of Peptides and Proteins

BY J. BRIDGEN, G. L. E. KOCH, C. CHOTHIA, AND R. H. PAIN

PART I: Primary Structure and Chemical Modifications *by J. Bridgen and G. L. E. Koch*

1 Introduction

The amount of information accumulating on protein primary structure has been increasing almost exponentially over the past few years, creating substantial difficulties for reviewers hoping to convey the spirit of protein chemistry today and also to cover the field comprehensively. In this Report we have tried to assess as much of the information as possible which will aid those researchers in the process of sequencing proteins as well as those anticipating doing so in the near future. Since the inevitable prerequisite is the availability of the proteins concerned in adequate amounts we have treated the subject of protein isolation in some detail. As the feasibility of sequencing also depends on the size of a protein, information on molecular weights has been collected. Advances in the methodology for the selective modification and sequencing of proteins have been treated comprehensively and, finally, the results of the sequencing studies performed over the past year are reviewed.

2 Protein Isolation and Characterization

It is clear from Table 1 (see p. 49) that a considerable number of proteins have been isolated and characterized during the past year and it is not possible to discuss each of these in detail. It should be emphasized that not all the reports included in Table 1 are concerned with novel isolations but they do involve new information either on the isolation procedures or on the subunit structures of the proteins concerned. Many reports which are only concerned with sizes or subunit structure of proteins have been included in Table 1a (see p. 67). Since one of the objectives of the table is to aid those assessing the feasibility of sequencing studies on such proteins, an indication is given of the availability of the amino-acid composition of the proteins concerned.

Techniques for Isolating Proteins.—Several new approaches to the isolation of proteins have been described. Of special interest are those which are concerned with the extraction and solubilizing of membrane proteins. It

was found[1] that about 45% of the proteins of the erythrocyte membrane can be extracted with dilute hydrochloric acid at pH values lower than 3. The proteins remain soluble and are therefore easily separable although they are probably highly aggregated. The general effect of the hydrochloric acid seems to be to split off vesicular structures which are attached to the membrane by hydrophilic interactions. The organic mercurial mersalyl has been used to solubilize erythrocyte membrane proteins with some success.[2] One advantage of this method over the more conventional sodium dodecylsulphate procedure is that the slower-moving components are very well resolved during polyacrylamide gel electrophoresis. Furthermore, the proteins are not resolved purely on the basis of size, which makes it a useful complementary system. The binding of mersalyl to proteins appears to depend on interaction with sulphydryl groups and it is possible that selectivity in the extraction results from the variation in the sulphydryl content of proteins.

A detailed study of the precipitation of proteins with polyethyleneglycol has been reported in connection with the development of procedures for the continuous-flow precipitation of proteins.[3] Optimum fractionation of proteins requires the use of low total protein concentrations, but total precipitations can be carried out at high protein concentrations. The partitioning of proteins in two-phase systems containing charged polyethyleneglycol could prove useful for purification on a large scale.[4] When a mixture of positively charged glycol is made with dextran, polyethyleneglycol, and water, the charged glycol is found mainly in the upper phase of the biphasic system. The presence of the positively charged component causes negatively charged proteins to pass into the upper layer. The purely electrostatic nature of the phenomenon is emphasized by the damping effect of salts on separation. The main use of this method is in large-scale batch operations, although counter-current distribution can be used for greater refinement. The method was tested on a yeast lysate and resolved several glycolytic enzymes. The phase-partitioning method has also been used in the isolation of the hydrophobic membrane protein phospholipase A1 from *E. coli*.[5]

Several new types of matrices for the fractionation of proteins by column chromatography have been described. Many of these have been developed as a result of unexpected observations made during studies with other techniques. Thus a new technique called 'hydrophobic chromatography' has evolved from studies on affinity chromatography when it was found that the agarose gels with the usual spacers attached prior to ligand coupling had very useful properties for protein purification. In a study on the

[1] H. Schiechl, *Biochim. Biophys. Acta*, 1973, **307**, 65.
[2] A. C. Cantrell, *Biochim. Biophys. Acta*, 1973, **311**, 381.
[3] P. R. Foster, P. Dunnill, and M. D. Lilly, *Biochim. Biophys. Acta*, 1973, **317**, 505.
[4] G. Johansson, A. Hartman, and P. A. Albertsson, *European J. Biochem.*, 1973, **33**, 379.
[5] P. A. Albertsson, *Biochemistry*, 1973, **12**, 2525.

purification of glycogen synthetase a range of ω-aminoalkyl-agaroses was prepared and their affinity for proteins was tested.[6] Glycogen synthetase from rabbit muscle was retained quite strongly on aminobutyl-agarose and was only eluted with a salt gradient. Higher members of the series (*e.g.* $n = 6$) bound enzyme too tightly for elution without denaturation. In contrast, glycogen phosphorylase was only retarded by the longer alkyl derivatives. The fact that there is such a strong dependence on the length of the alkyl chain shows that the properties of the column are not simply due to ion-exchange effects. Thus the retention power probably depends on hydrophobic interactions as well. It is likely that it will be necessary to consider this phenomenon when designing affinity columns with aminoalkyl spacers in future. The conventional approach of attaching the spacer to the matrix before coupling the ligand could result in the generation of uncoupled spacers, which will affect the binding capacity of such columns in an unspecific way. Therefore from the point of view of specificity it is better to attach the spacer to the ligand and then to couple the adduct to the matrix. An analogous study to that described above was carried out on plant aminoacyl-tRNA synthetases.[7] The involvement of hydrophobic forces in the binding of these enzymes was confirmed by methylating the aminohexyl-Sepharose and observing that binding was unimpaired. One interesting observation is that large molecules are retained more strongly than small ones, indicating that there is a higher proportion of hydrophobic 'pockets' in larger proteins. The very high affinity of some aminoalkyl-agaroses for enzymes has actually been used to prepare columns of immobilized enzymes.[8] Thus virtually irreversible binding of several enzymes was achieved with retention of activity and without recourse to the more drastic conditions used for the covalent attachment of enzymes to matrices for immobilization.

A new approach to the separation of strongly interacting macromolecules, called 'intervent dilution chromatography', has been developed.[9] Conventional techniques are not very effective with interacting systems since they actually rely on the independent behaviour of the components. The object in intervent dilution chromatography is to introduce reagents called 'intervents' which actually reduce the binding between the interacting molecules, and thereby enhance separation. This is achieved by using a gel column to drive the interacting macromolecules repeatedly across an intervent concentration boundary and cause a repeated shift in the equilibria involved. Under appropriate conditions this continual change in equilibrium constant leads to separation. The system has been applied to the separation of ribosomal proteins from one another and from ribosomal RNA.

[6] S. Shaltiel and Z. Er-El. *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 778.
[7] H. Jakubowski and J. Pawelkiewicz, *F.E.B.S. Letters*, 1973, **34**, 150.
[8] B. H. J. Hofstee, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 1137.
[9] S. Kirkegaard and C. C. Agee, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2424.

Another chromatographic technique which relies on the hydrophobic characteristics of proteins is called 'phosphate-induced chromatography'.[10] It is based on the observation that in the presence of high concentrations of phosphate (1.0 mol l^{-1}), over half the proteins in a cell-free extract of *E. coli* were absorbed on columns of agarose to which valine was attached. Elution with decreasing concentrations of phosphate caused most of the absorbed proteins to be eluted, and the order of elution correlated with the solubility of the proteins in ammonium sulphate. The columns have a high capacity for protein and are relatively easy to prepare.

Affinity Chromatography.—There is little doubt that the most powerful procedure available for the specific isolation of biological molecules is that known as 'affinity chromatography'. The strategy of this technique is to exploit the fact that most, if not all, molecules which exist in biological systems interact more or less specifically with other molecules in such systems. Thus enzymes react with specific substrates, antibodies with specific antigens, receptors with specific hormones, *etc.* The general approach which has been developed involves the attachment of the appropriate complementary molecule to a suitable matrix, to bind the impure sample to the matrix, usually in a column, and to elute the desired protein in as specific a manner as possible. Because of the innate power of this approach it is likely that it will supersede most other purification procedures in the near future. We shall therefore consider the advances and applications of the technique during 1973 in some detail in this Report.

Hitherto, the question of the feasibility of using affinity chromatography for purifying a protein has been a largely empirical matter which could become quite protracted. Recently, however, an analytical method has been developed which permits the prediction of the result of an affinity chromatography experiment without actually performing the experiment.[11] The technique is called *crossed immuno-affinoelectrophoresis* and is an extension of ordinary immuno-electrophoresis. It is shown that the binding properties of serum proteins to Con A-Sephadex are predictable by this technique. However, it should be noted that quite a lot of effort could be involved if the arrangements for immuno-electrophoresis are not already available.

The use of a suitable matrix for the preparation of affinity columns is important, and comparisons have been made of some of the common choices. Agarose and cellulose have been compared as supports for affinity chromatography of dehydrogenases and kinases.[12] It was found that the cyanogen bromide treatment of cellulose tends to introduce ion-exchange groups which usually have a deleterious effect on the columns. In contrast, no such effect was observed when agarose was similarly treated,

[10] R. A. Rimerman and G. W. Hatfield, *Science*, 1973, **182**, 1264.
[11] T. C. Bog-hansen, *Analyt. Biochem.*, 1973, **56**, 480.
[12] C. R. Lowe, M. J. Harvey, D. B. Craven, and P. D. G. Dean, *Biochem. J.*, 1973, **133**, 499.

showing that it is the preferred matrix when CNBr treatment is involved. This article [12] also describes studies on other aspects of affinity chromatography. Thus it was found that an extension arm 0.8—1 nm long gave maximum binding and that the binding of enzyme was independent of equilibration time and flow rate. Also, the concentration of immobilized ligand is more important than the concentration of enzyme. An accompanying paper [13] considers the general applicability of affinity chromatography to NAD(P)-dependent dehydrogenases, and develops methods for increasing the selectivity of the process.

A rather unexpected property of agarose has been detected during studies with avidin.[14] It was found that there was a temperature-dependent association of the avidin subunits even after covalent attachment to the agarose matrix, which implies that the agarose matrix is itself more mobile than previously anticipated. It was estimated that the avidin subunits could move through distances greater than 20 nm. It will be important to consider this mobility when agarose matrices are used for studies on oligomeric enzymes, in which it is often assumed that covalently bound subunits are incapable of interacting with one another. Alternatively, agarose which has been cross-linked with divinylsulphone may be used to reduce this mobility.

One rather irritating feature of affinity columns is the tendency towards a small but constant leakage of the ligand from the column, which limits the use of such columns for the isolation of very small amounts of proteins. In order to circumvent this a procedure has been described for coupling polylysine, polyornithine, or polyvinylamine to agarose.[15] These derivatives differ from conventional ones by virtue of a multipoint attachment of the polylysine to the matrix, thus enhancing the stability of the 'spacer' relative to the usual alkyl spacers, which are only univalently attached to the matrix. Studies with dansyl adducts of polylysyl-Sepharose showed complete stability over 3 months and this compares favourably with N_{ε}-DNP-lysine coupled directly to agarose, which loses 15% of the dinitrophenyl groups over the same period. The polylysyl-agarose has one disadvantage since all the amino-groups cannot be removed after coupling. However, it is claimed that this has not prevented success with the univalent derivatives and should therefore not pose serious problems.

Although agarose has been used most extensively as the matrix for affinity chromatography, other supports which do not suffer from some of its disadvantages, *e.g.* compressibility, are being used increasingly. Of these, glass beads are likely to be quite promising, and a good example of their value is evident in a report on the isolation of lipoamide dehydrogenase.[16] A column was prepared by allowing the surface of glass beads to

[13] C. R. Lowe, M. J. Harvey, D. B. Craven, M. A. Kerfoot, M. E. Hollows, and P. D. G. Dean, *Biochem. J.*, 1973, **133**, 507.
[14] N. M. Green and E. J. Toms, *Biochem. J.*, 1973, **133**, 687.
[15] M. Wilchek, *F.E.B.S. Letters*, 1973, **33**, 70.
[16] W. H. Scouten, F. Torok, and W. Gitomer, *Biochim. Biophys. Acta*, 1973, **309**, 521.

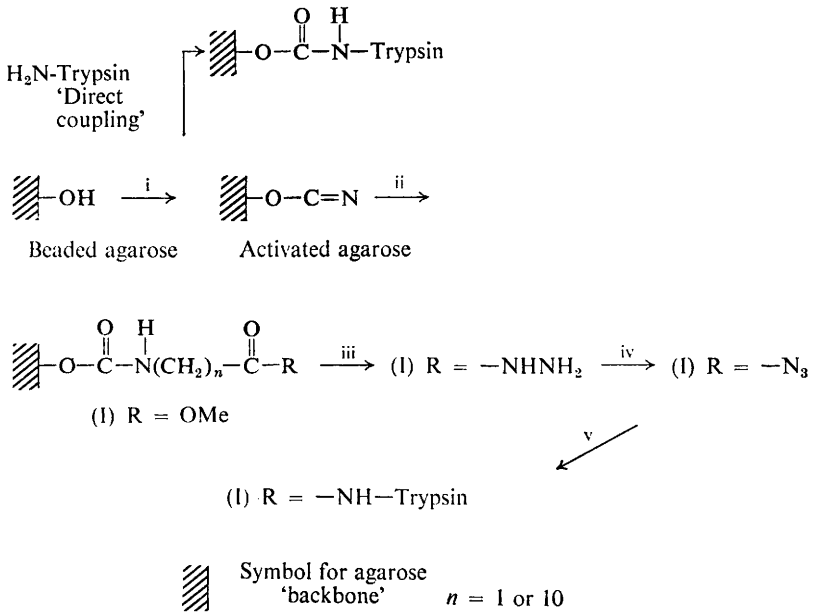
react with γ-aminopropyltriethoxysilane and further coupling the product to lipoyl chloride. The excess amino-groups could be blocked with acetic anhydride. Apart from the physical stability of the matrix, glass bead columns should prove very useful in those cases in which a chemically inert material is required. For example, it should be possible to couple amino-acids through the carboxy-group using N-blocked amino-acids. The removal of the N-blocking groups often requires rather drastic conditions which would damage agarose and cellulose matrixes very severely. On the other hand, glass beads would be quite adequate for this purpose providing the linkage is secure.

A new procedure for the preparation of agarose hydrazides for coupling nucleotides and nucleosides has been described.[17] This is achieved in a single step by coupling adipic acid hydrazide by the CNBr procedure. Periodate-oxidized nucleotides can then be coupled to the agarose hydrazide under mild conditions. The binding of nucleotide to the agarose hydrazide can be controlled by the amount of periodate used. A further advantage of agarose hydrazide is that the unreacted hydrazide groups do not impose ion-exchange properties on the matrix at neutral and high pH.

Another detailed study on the preparation and use of hydrazide derivatives of agarose and porous glass has appeared.[18] The hydrazides were prepared by coupling methyl-11-aminoundecanoate or glycine methyl ester to activated agarose and then converting the esters into the corresponding hydrazides. Just before use the hydrazides are converted into the more reactive azides, which can react with nucleophilic groups on protein molecules and therefore be used to produce affinity columns. Both agarose and beaded glass have been used as supports with comparable results. Scheme 1 shows a summary of the procedures used.

One of the most successful general methods based on the principle of affinity chromatography is that involving the use of immunoabsorption (see Table 2, p. 73). Many applications have been manipulated so as to utilize the readily available antidinitrophenol antibody. Thus, proteins associated with the surface of the erythrocyte membrane have been isolated after specific labelling with sodium 2,4,6-trinitrobenzenesulphonate.[19] The trinitrophenylated proteins are easily purified by reverse immunoabsorption using anti-DNP antibody-agarose columns. Minor modifications permit the use of the same approach for the study of other cell types and this approach promises to be very useful in the future. The only disadvantage, from the point of view of isolating surface proteins, is that the reagent (TNBS) is not totally non-permeable and would therefore tend to affect the placement of specific proteins. Another general procedure for protein isolation based on the use of anti-DNP antibody-agarose has been reported.[20] The approach was applied to the isolation of trypsin in

[17] R. Lamed, Y. Levin, and M. Wilchek, *Biochim. Biophys. Acta*, 1973, **304**, 231
[18] L. J. Loeffer and J. V. Pierce, *Biochim. Biophys. Acta*, 1973, **317**, 20.
[19] G. Tarone, M. Prat, and P. M. Comoglio, *Biochim. Biophys. Acta*, 1973, **311**, 214.
[20] M. Wilchek and M. Gorecki, *F.E.B.S. Letters*, 1973, **31**, 149.



Reagents: i, BrCN, pH 10; ii, $H_2N(CH_2)_nCO_2Me$, pH 9.0; iii, H_2NNH_2–MeOH; iv, $NaNO_2$–0.1N-HCl; v, trypsin, pH 8.0–9.0

Scheme 1

the following way. Soybean trypsin inhibitor was dinitrophenylated and mixed with trypsin to form the trypsin–inhibitor complex. The complex was then passed down the antibody column, whereupon the complex was absorbed specifically and eluted subsequently. Dissociation of the complex gave pure trypsin. It should be noted that in its present state the method is rather drastic, with corresponding loss of efficiency, but suggestions for improvement are provided.

Reverse immunoabsorption has also been used in the isolation of the elusive antiviral protein interferon.[21] A substantial purification was achieved in a single step using antibodies to unpurified interferon and it is likely that this can be improved substantially by using enriched preparations. An interesting variation of the conventional approach to affinity chromatography has been employed in the isolation of the serum i antigen.[22] The antigen was bound to an antibody-sepharose column at low temperature (4 °C) and eluted at 37 °C. This is a fairly gentle method of elution compared with many others and should prove useful.

Analogous with the antibody-agarose systems are the lectin-agarose columns which have been developed for the isolation of various glycoproteins. The lectins are particularly useful for this purpose since they

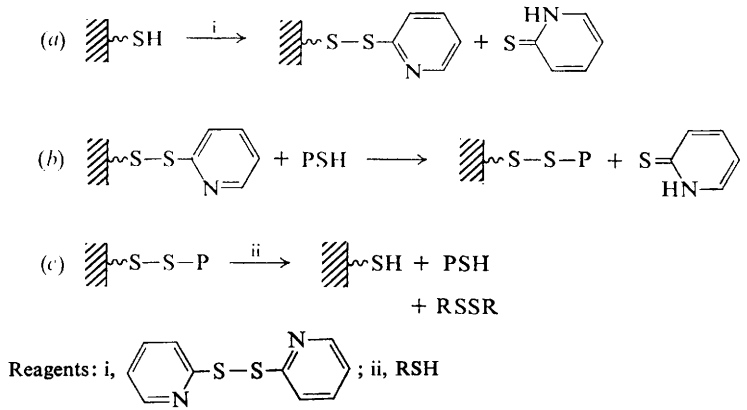
[21] J. D. Sipo, J. De Maeyer-Grignard, B. Fauconnier, and E. De Maeyer, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1037.

[22] A. G. Cooper and M. C. Brown, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 297.

have a very high affinity for sugars. The columns can often be used in the presence of denaturants, which is a substantial asset since these conditions are often obligatory when dealing with membrane and other glycoproteins. A novel use of lectin columns has been reported in studies on the insulin receptors.[23] It was found that lectins like concanavalin A have profound insulin-like effects on metabolic processes such as the inhibition of adenyl cyclase activity. Use of a concanavalin A-agarose column showed that there was direct competition between the lectins and insulin for isolated insulin receptors. This is consistent with earlier reports that the insulin receptor is a glycoprotein.

It has been known for some time that glycosidases are very sensitive to inhibition by the corresponding sugar lactones, suggesting that the inhibitors might be very useful as ligands for affinity chromatography of glycosidases. This has now been borne out by a study using β-glucuronidase from bovine liver.[24] Saccharo-1,4-lactone was coupled to aminohexyl-agarose through the free carboxy-group using a water-soluble carbodi-imide. The success of the method suggests that sugar lactone affinity columns should be one of the simplest ways of isolating glycosidases. A slightly unusual feature of β-glucuronidase is that it appears to form irreversible complexes with the lactone-agarose in the absence of the aminohexyl arm. The reason for this phenomenon is obscure.

Several extensions of the principles of affinity chromatography have been developed recently. One of these, called 'covalent chromatography',[25] has been used with papain and is illustrated in Scheme 2. It involves the



Scheme 2

[23] P. Cuatracasas and G. P. E. Tell, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 485.
[24] R. G. Harris, J. J. M. Rowe, P. S. Stewart, and D. C. Williams, *F.E.B.S. Letters*, 1973, **29**, 189.
[25] K. Brocklehurst, J. Carlsson, M. P. J. Kierstan, and E. M. Crook, *Biochem. J.*, 1973, **133**, 573.

preparation of a glutathione-2-pyridyl disulphide derivative of agarose which reacts specifically with the active-site sulphydryl group of papain, forming a covalent disulphide linkage between the protein and the matrix. The protein can then be eluted with a reducing agent, which cleaves the disulphide linkage.

Estimation of Protein Concentration.—Several reports have been concerned with the monitoring of protein in column effluents during purification. An ultrasensitive protein assay which is reported to be about 500 times more sensitive than the conventional Lowry procedure has been described.[26] The method is based on the complexing of copper to the protein and measuring the amount of bound copper by a very sensitive assay. The procedure utilizes five distinct steps but most of these are carried out in the same tube. The assay is not affected by many of the standard buffers but reducing agents and edta interfere seriously. It is linear between 0.01 and 0.2 μg protein. With the increasing use of detergents in protein isolation, protein assays which can be applied to samples containing such compounds are highly desirable. One method which is based on the biuret technique and uses propane-1,2-diol to solubilize the copper–protein complex has been described.[27] The method also works in the presence of urea and guanidine hydrochloride. The presence of sulphydryl reagents is a common obstacle to the use of many protein assays, but a way of circumventing this is to swamp out the thiol with an alkylating agent, *e.g.* iodoacetic acid. A procedure for combining this with the Lowry protein assay has been described.[28]

A rapid and sensitive assay for protein in dilute solution which should have many applications is based on the staining techniques used for acrylamide gels.[29] The approach is to precipitate the protein with trichloroacetic acid in the presence of sodium dodecylsulphate, filter onto a Millipore membrane, and stain with Amidoschwarz 10B. The protein–dye complex is eluted and its absorbance measured. The assay is very reproducible and linear from 3 to 30 μg protein. One major advantage is the lack of interference by most, if not all, of the reagents commonly used in protein isolation.

A detailed study has been made of the use of the fluorescent reagent fluorescamine for the estimation of proteins.[30] The manual assay can be applied to 500 ng of protein, and a semiautomated procedure which separates protein from amines of low molecular weight can be used with 10—50 ng of protein. The main limitation of this very useful technique is that it cannot be used in the presence of amine buffers in the usual way.

[26] M. L. Goldberg, *Analyt. Biochem.*, 1973, **51**, 240.
[27] S. Futterman and M. H. Rollins, *Analyt. Biochem.*, 1973, **51**, 443.
[28] E. Ross and G. Schatz, *Analyt. Biochem.*, 1973, **54**, 304.
[29] W. Schaffer and C. Weissmann, *Analyt. Biochem.*, 1973, **56**, 502.
[30] P. Bohlen, S. Stein, W. Dairman, and S. Udenfriend, *Arch. Biochem. Biophys.*, 1973, **155**, 213.

However, the introduction of a gel-filtration step in the semiautomatic procedure permits its application even in the presence of otherwise interfering small molecules.

Kirschenbaum has extended the lists of the molar absorptivities of pure proteins [31–33] and has also compiled available information on the amino-acid analyses of many proteins.[34–36]

Molecular Weight Determination.—Improvements continue to be made in the techniques used to determine the sizes of proteins and of their subunits. A method for the analysis of elution data of proteins obtained from gel filtration has been described.[37] It was found that polynomial relationships represent elution data for proteins with a much greater degree of precision than linear equations. Proteins labelled with the fluorescent reagent dansyl chloride behave in a very similar fashion to unlabelled proteins during gel filtration in 6M guanidine hydrochloride.[38] Consequently, molecular weight estimations of dissociated polypeptide chains can be determined with very small amounts of protein and the sensitivity can probably be increased further by using radioactive dansyl chloride. A major advantage of this approach over the analogous method using SDS gel electrophoresis is that the excess reagent does not have to be removed from the protein and actually serves as a useful reference.

A rapid and accurate method for the determination of the molecular weights of homogeneous proteins in the analytical ultracentrifuge has been described.[39] The data are collected with a photoelectric scanner during approach to equilibrium in short columns of solution. The method is accurate and reproducible for pure proteins in the range 13 000—40 000 dalton. Only about 40 μg of protein and a total of about 3 h are required if a computer and teletype are available. The theoretical upper size limit is estimated to be about 400 000 dalton.

SDS-Gel electrophoresis is unquestionably the simplest and probably the most accurate method available for the estimation of the molecular weights of dissociated proteins. It has, however, been known for some time that some types of proteins, *e.g.* glycoproteins, can behave anomalously in this system. It is now reported that α-amylase from *B. subtilis* migrates in SDS-gels with a mobility which is completely unrelated to the molecular weight of its monomers.[40] Thus the mobility corresponds to a molecular weight of 154 000 instead of 24 000. It was shown that this is

[31] D. Kirschenbaum, *Internat. J. Peptides Protein Res.*, 1973, **5**, 49.
[32] D. Kirschenbaum, *Analyt. Biochem.*, 1973, **55**, 166.
[33] D. Kirschenbaum, *Analyt. Biochem.*, 1973, **56**, 237.
[34] D. Kirschenbaum, *Analyt. Biochem.*, 1973, **56**, 208.
[35] D. Kirschenbaum, *Analyt. Biochem.*, 1973, **53**, 223.
[36] D. Kirschenbaum, *Analyt. Biochem.*, 1973, **52**, 234.
[37] A. K. Grover and M. Kapoor, *Analyt. Biochem.*, 1973, **51**, 163.
[38] M. Yoshino and T. Takagi, *Analyt. Biochem.*, 1973, **54**, 290.
[39] R. M. Baurain, J. C. Moreux, and F. Lamy, *Biochem. Biophys. Acta*, 1973, **295**, 18.
[40] E. D. Mitchell, P. Riquetti, R. H. Loring, and K. L. Carraway, *Biochim. Biophys. Acta*, 1973, **295**, 314.

Table 1 *Isolation and subunit structures of proteins*

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
APS-Sulphohydrolase	Bovine liver	68 000	68 000	+	41
Alkaline protease	*S. aureus*	12 500	—	−	42
Arylsulphatase A_β	Human urine	125 000	—	−	43
Amylase	Human pancreas	—	$(57\,000)_n$	+	44
β-L-Arabinosidase	*Cajanus indicus*	25 00	—	−	45
Adenosine deaminase conversion factor	Human lung	140 000	—	−	46
Arginine kinase	*Panulirus longipes*	38 000	—	+	47
Aspartate transcarbamylase	Yeast	138 000	$(21\,000)_n$	−	48
Aspartate aminotransferase	Bovine brain	95 000	—	−	49
Agglutinin	Wheat germ	—	—	+	50
Amelogenins	Bovine dental enamel	4 000 16 000	—	+	51
Apoferritin	Horse spleen	443 000	24 × 20 000	−	52
Aldolase (Class I)	*E. coli*	140 000	4 × 35 000	+	53
Aldolase (Class II)	*E. coli*	70 000	2 × 35 000	+	53
Acetyl-CoA synthetase	Ox heart mitochondria	57 000	—	+	54
Aldolase	Human muscle and liver	—	—	+	55

[41] A. M. Stokes, W. H. B. Denmer, F. A. Rose, and K. S. Dodgson, *Biochim. Biophys. Acta*, 1973, **302**, 64.
[42] S. Arvidson, T. Holme, and B. Lindholm, *Biochim. Biophys. Acta*, 1973, **302**, 135.
[43] R. L. Stevens, M. Hartman, A. L. Fluharty, and H. Kihara, *Biochim. Biophys. Acta*, 1973, **302**, 338.
[44] D. J. Stiefel and P. J. Keller, *Biochim. Biophys. Acta*, 1973, **302**, 345.
[45] P. M. Dey, *Biochim. Biophys. Acta*, 1973, **302**, 393.
[46] H. Nishihara, S. Ishikawa, K. Shinkai, and H. Akeda, *Biochim. Biophys. Acta*, 1973, **302**, 429.
[47] P. Masiar and D. C. Shaw, *Biochim. Biophys. Acta*, 1973, **303**, 308.
[48] D. M. Aitken, A. R. Bhatti, and J. G. Kaplan, *Biochim. Biophys. Acta*, 1973, **309**, 50.
[49] M. L. Krista and M. L. Fonda, *Biochim. Biophys. Acta*, 1973, **309**, 83.
[50] A. K. Allen, A. Neuberger, and N. Sharon, *Biochem. J.*, 1973, **131**, 155.
[51] F. M. Eggert, G. A. Allen, and R. C. Burgess, *Biochem. J.*, 1973, **131**, 471.
[52] R. R. Crichton, R. Eason, A. Barclay, and C. F. A. Bryce, *Biochem. J.*, 1973, **131**, 855.
[53] D. Stribling and R. N. Perham, *Biochem. J.*, 1973, **131**, 833.
[54] J. C. Londesborough, S. L. Yuan, and L. T. Webster, *Biochem. J.*, 1973, **133**, 23.
[55] P. A. M. Eagles and M. Iqbal, *Biochem. J.*, 1973, **133**, 429.

Table 1 (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
α_1-Antitrypsin	Human	54 000	—	+	56
Abrin	*Abrus precatorium*	260 000	4 × 35 000 4 × 30 000	−	57
α_1-Antitrypsin	Human	51 000	1 × 51 000	+	58
AMV subunits with DNA polymerase and RNase H	AMV	—	$(65\,000)_n$ $(105\,000)_n$	−	59
Alveolar proteinosis glycoproteins	Human	36 000 62 000	— —	+	60
Avian oncornavirus protein precursor		—	76 000 12 000	−	61
Arabinose dehydrogenase	Pig liver	245 000	—	−	62
Alcohol dehydrogenase	*P. aeruginosa* membrane	~300 000	—	−	63
Apiitol dehydrogenase	*M. lysodiecticus*	110 000	—	−	64
Angistensin 1-converting-enzyme	Hog lung	210 000	3 × 70 000	−	65
Adenosine deaminase	Bay scallop	130 000	—	−	66
α-Amylase	*B. licheniformis*	22 500	—	−	67
Aldolase	Human cardiac	160 000	4 × 40 000	+	68
Agglutinin receptor	Mouse leukaemia L1210	~50 000	—	−	69
Adenosine triphosphatase	*B. megaterium*	380 000	6 × 65 000	−	70
Aryl sulphatase isozyme	*A. oryzae*	100 000	—	−	71
Aspartyl-tRNA synthetase	Yeast	100 000	1 × 100 000	−	72
α-Actinin	Pig muscle	—	—	+	73
α-Actinin	Pig muscle	190 000	2 × 95 000	+	74
Agglutinin	Castor beans	125 000	33 000 30 000 27 500	−	75
Aminopeptidase	*C. histolyticum*	340 000	6 × 60 000	−	76
Arginine kinase	*Drosophila melanogaster*	40 000	—	−	77
Apovitellenin 1	Emu egg yolk	10 000	1 × 10 000	+	78
Aspartic β-semialdehyde dehydrogenase	Yeast	156 000	4 × 41 000	−	79

4-Aminobutyrate-2-ketoglutarate transaminase	Mouse brain	109 000	—	—	80
Acetylcholinesterase	*Bungarus fasciatus*	126 000	2 × 60 000	—	81
ATP:RNA adenyltransferase	*E. coli*	58 000	1 × 58 000	—	82
Abrin	*Semen jequisti*	65 000	1 × 30 000 1 × 35 000	—	83
APTase (unstimulated)	*M. lysodeikticus*	330 000—350 000	52 500 47 000 41 500 28 500	—	84

[56] S. K. Chan, J. Luby, and Y. C. Wu, *F.E.B.S. Letters*, 1973, **35**, 79.
[57] A. Macpherson and A. Rich, *F.E.B.S. Letters*, 1973, **35**, 257.
[58] I. P. Crawford, *Arch. Biochem. Biophys.*, 1973, **156**, 215.
[59] D. P. Grandgenett, G. F. Gerard, and M. Green, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 230.
[60] M. A. Passero, R. W. Tye, K. H. Kilburn, and W. S. Lynn, *Proc. Nat. Acad. Sci., U.S.A.*, 1973, **70**, 973.
[61] V. M. Vogt and R. Eisenman, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1734.
[62] A. G. Maijub, M. A. Pecht, G. R. Miller, and W. R. Carper, *Biochim. Biophys. Acta*, 1973, **315**, 37.
[63] J. P. Tassin, C. Celier, and J. P. Vandecastecle, *Biochim. Biophys. Acta*, 1973, **315**, 220.
[64] R. Hanna, M. Picken, and J. Mendicino, *Biochim. Biophys. Acta*, 1973, **315**, 259.
[65] T. Nakajima, G. Oshima, H. S. J. Yeh, R. Igic, and E. G. Erdos, *Biochim. Biophys. Acta*, 1973, **315**, 430.
[66] G. R. Harbison and J. R. Fisher, *Arch. Biochem. Biophys.*, 1973, **154**, 84.
[67] N. Saito, *Arch. Biochem. Biophys.*, 1973, **155**, 290.
[68] B. L. Allen, R. W. Gracy, and B. G. Harris, *Arch. Biochem. Biophys.*, 1973, **155**, 325.
[69] V. K. Jansons and M. M. Burger, *Biochim. Biophys. Acta*, 1973, **291**, 127.
[70] R. Mirsky and V. Barlow, *Biochim. Biophys. Acta*, 1973, **291**, 480.
[71] M. Rasburn and C. H. Wynn, *Biochim. Biophys. Acta*, 1973, **293**, 191.
[72] J. Gangloff and G. Dirheimer, *Biochim. Biophys. Acta*, 1973, **294**, 263.
[73] A. Suzuki, D. E. Goll, M. H. Stromer, I. Singh, and J. Temple, *Biochim. Biophys. Acta*, 1973, **295**, 188.
[74] R. M. Robson and M. G. Zeece, *Biochim. Biophys. Acta*, 1973, **295**, 208.
[75] L. G. Grutler and H. J. Horstmann, *Biochim. Biophys. Acta*, 1973, **295**, 582.
[76] E. Kessler and A. Yaron, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 405.
[77] T. Wallimann and H. M. Eppenberger, *European J. Biochem.*, 1973, **38**, 180.
[78] R. W. Burley, *Biochemistry*, 1973, **12**, 1464.
[79] M. J. Holland and E. W. Westhead, *Biochemistry*, 1973, **12**, 2264.
[80] A. Schousloe, J. Y. Wu, and E. Roberts, *Biochemistry*, 1973, **12**, 2868.
[81] V. Kumar and W. B. Elliot, *European J. Biochem.*, 1973, **34**, 586.
[82] A. E. Sippel, *European J. Biochem.*, 1973, **37**, 31.
[83] S. Olsnes and A. Sikl, *European J. Biochem.*, 1973, **35**, 179.
[84] S. M. Andreu, J. A. Albendea, and E. Munoz, *European J. Biochem.*, 1973, **37**, 505.

Table 1 (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
Activator for cAMP phospho-diesterase	Bovine heart	19 000—27 000	—	—	85
Anthranilate synthetase	*P. putida*	63 000 18 000	—	—	86
Arginine binding protein	*E. coli*	27 500	—	—	87
Anthranilate synthetase	*N. crassa*	300 000	1 × 80 000 1 × 200 000	—	88
Aldolase (Class I)	*Micrococcus aerogenes*	33 000	1 × 33 000		89 90
Arginine decarboxylase	*E. coli*	296 000	4 × 70 000	—	91
Acyl carrier protein	*Euglena gracilis*	10 400	—	—	92
Adrenodoxin reductase	Bovine adrenal glands	54 000	1 × 54 000	—	93
Antihaemophilic factor	Human	1 120 000	5 × 240 000	—	94
Actin	Human platelets	45 000	1 × 45 000	—	95
Ascorbate oxidase	Zucchini	140 000		—	96
Apolipoprotein	Pig high-density lipoprotein		26 000	—	97
Crystallin (FM)	Calf	14 500	—	—	98
Cytochromes-*c* type	*Nitrosomonas europeae*	F_1, 12 600 F_2, 34 400 F_3, 52 000 F_4, 77 000	—	— — — —	99
Cathepsin B1	Human	25 000	—	—	100
Cytochrome oxidase	Beef heart mitochondria	69 000	14 000 11 500	—	101
Carboxylesterase	Pig liver microsomes	180 000	3 × 61 000	—	102
Colicin K	*E. coli*	45 000	—	—	103
Calcium-binding protein	Bovine intestine	11 000	—	—	104
Complement ClS	Human plasma	180 000	—	—	105
Cytochrome P-450	Bovine adrenocortex	800 000	—	—	106
Carotenoprotein	*Rhodospirillium rubrum*	—	12 000	—	107
Co-lipase	Porcine pancreas	1, 9 870 2, 8 720	—	— —	108

Collagen	Chick embryonic notochord	100 000	—	–	109
Collagenolytic protease	*Uea pugilator*	—	25 000	–	110
Carboxypeptidases	*Penaeus setiferus*	A, 30 000	—	–	111
		B, 34 200		–	
γ-Carboxymuconolactone decarboxylase	*P. putida*	93 000	12 500	–	112
Carboxylesterase (E_1)	Rat liver microsomes	177 000	61 500	–	113

[85] T. S. Teo, T. H. Wang, and J. H. Wang, *J. Biol. Chem.*, 1973, **248**, 588.
[86] S. W. Queener, S. F. Queener, J. R. Meeks, and I. C. Gunsalus, *J. Biol. Chem.*, 1973, **248**, 151.
[87] B. P. Rosen, *J. Biol. Chem.*, 1973, **248**, 1211.
[88] A. Arroya-Begorich and J. A. DeMoss, *J. Biol. Chem.*, 1973, **248**, 1262.
[89] H. G. Lebherz and W. J. Rutter, *J. Biol. Chem.*, 1973, **24,8** 1650.
[90] H. G. Lebherz, R. A. Bradshaw, and W. J. Rutter, *J. Biol. Chem.*, 1973, **248**, 1660.
[91] W. H. Wu and D. R. Morris, *J. Biol. Chem.*, 1973, **248**, 1687.
[92] R. K. Dinello and M. L. Ernst-Fonberg, *J. Biol. Chem.*, 1973, **248**, 1707.
[93] J. W. Chu and T. Kimura, *J. Biol. Chem.*, 1973, **248**, 2089.
[94] M. E. Legoz, G. Schmer, R. B. Counts, and E. W. Davie, *J. Biol. Chem.*, 1973, **248**, 3946.
[95] F. M. Booyse, T. P. Horeke, and M. E. Rafelson, *J. Biol. Chem.*, 1973, **248**, 4083.
[96] M. H. See and C. R. Dawson, *J. Biol. Chem.*, 1973, **248**, 6596.
[97] R. L. Jackson, H. N. Baker, O. D. Taunton, L. C. Smith, C. W. Garner, and A. M. Gotto, *J. Biol. Chem.*, 1973, **248**, 2639.
[98] W. G. M. Van Den Broek, J. N. Leget, and H. Bloemendal, *Biochim. Biophys. Acta*, 1973, **310**, 278.
[99] D. A. Trouson, G. A. F. Ritchie, and D. J. D. Nicholas, *Biochim. Biophys. Acta*, 1973, **310**, 331.
[100] A. J. Barrett, *Biochem. J.*, 1973, **131**, 809.
[101] H. Komai and R. A. Capaldi, *F.E.B.S. Letters*, 1973, **30**, 273.
[102] K. C. Aune, *Arch. Biochem. Biophys.*, 1973, **156**, 115.
[103] W. F. Goebel, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 854.
[104] C. S. Fullmer and R. H. Wasserman, *Biochim. Biophys. Acta*, 1973, **317**, 172.
[105] K. Okamura, M. Muramatu, and S. Fujii, *Biochim. Biophys. Acta*, 1973, **295**, 252.
[106] M. Shikita, P. F. Hall, and S. Isaka, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 289.
[107] V. Schwenber and G. Gingras, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 94.
[108] M. F. Maylie, M. Charles, M. Astier, and P. Desnuelle, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 291.
[109] T. F. Linsenmayer, R. L. Trelstad, and J. Gross, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 39.
[110] A. Z. Eisen, K. O. Henderson, J. J. Jeffrey, and R. A. Bradshaw, *Biochemistry*, 1973, **12**, 1814.
[111] B. J. Gates and J. Travis, *Biochemistry*, 1973, **12**, 1867.
[112] D. Parke, R. B. Meagher, and L. N. Ornston, *Biochemistry*, 1973, **12**, 3537.
[113] R. Arndt, E. Heymann, W. Junge, K. Krisch, and H. Hollandt, *European J. Biochem.*, 1973, **36**, 120.

Table 1 (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
Cytochrome *c* oxidase	Baker's yeast	215 000	42 000, 14 000 34 500, 12 500 23 000, 9 500	–	114
Ceramide trihexosidase (A)	Human plasma	95 000	22 000	–	115
Catalase	Baker's yeast	248 000	61 000	–	116
Cytochrome b_{560}	*Ascaris lumbricoides*	13 000	—	–	117
Calcium-binding protein	Bovine adrenal medulla	11 900	—	–	118
Carboxypeptidase (D-Ala)	*Bacillus subtilis*	—	50 000	–	119
Cholinergic receptor	*Electrophorus electricus*	260 000	44 000	–	120
Chromatin non-histone proteins	Calf thymus	19 000 17 000	—	+ –	121
Cardiotoxic protein (Volvatoxin)	*Volvariella volvacea*	A_1, 50 000 A_2, 24 000	—	– –	122
Carboxypeptidase C_N	*Citrus natsudaidai*	93 000	—	–	123
Dextranase	*Penicillium funiculosum*	44 000	—	–	124
DNA modification enzyme	Phage P1	—	76 000 45 000	–	125
Dopamine-β-hydroxylase	Bovine adrenal glands	300 000	75 000	–	126
DNA polymerase	Human breast tumour	—	—	–	127
Diphenylenediamine	*Tecoma* leaves	200 000	—	–	128
2,4-Diaminopentanoic acid dehydrogenase	*Clostridium stricklandii*	72 000	35 000	–	129
DNA polymerases	*Lactobacillus acidophilus*	78 000 92 000	—	–	130
DNA polymerase	Calf thymus chromatin	45 520	44 000	–	131
DNA ligase	*E. coli*	74 000	74 000		132
edta-sensitive protease	*Staphylococcus aureus*	28 000	—	–	133
Extracellular protease	*Aspergillus oryzae*	25 000	—	–	134
Enolase	Swine kidney	90 000	52 000	–	135
Enterokinase	Porcine duodenal mucosa	195 000	134 000 62 000	–	136

Enolase	*Thermus aquaticus* YT-1	352 000	44 000	–	137
Enolase	*Thermus aquaticus* X-1	404 000	48 000	–	138
Enterotoxin	*Vibrio cholerae*	—	66 000 36 000	–	139
Factor XI	Human	90 000—100 000	72 000	–	140
Ferrodoxins	*Clostridium pasteurianum*	110 000	Azo, 55 000 Mo, 50 000 60 000	–	141
Fibroin (silk)	*Bombyx mori*	365 000	—	–	142

[114] T. L. Mason, R. O. Poyton, D. C. Wharton, and G. Schatz, *J. Biol. Chem.*, 1973, **248**, 1202.
[115] C. A. Mapes, C. H. Suelter, and C. C. Sweeley, *J. Biol. Chem.*, 1973, **248**, 2471.
[116] T. C. M. Seah and J. G. Kaplan, *J. Biol. Chem.*, 1973, **248**, 2889.
[117] K. S. Cheah, *J. Biol. Chem.*, 1973, **248**, 4101.
[118] J. C. Brooks and F. L. Sigel, *J. Biol. Chem.*, 1973, **248**, 4189.
[119] J. N. Umbreit and J. L. Strominger, *J. Biol. Chem.*, 1973, **248**, 6759.
[120] G. Biesecker, *Biochemistry*, 1973, **12**, 4403.
[121] G. H. Goodwin and E. W. Johns, *European J. Biochem.*, 1973, **40**, 215.
[122] J. Y. Lin, T. W. Seng, C. C. Chen, G. Y. Shi, and T. C. Teeng, *Nature*, 1973, **246**, 524.
[123] Y. Kubota, S. Shoji, T. Funakoshi, and H. Veki, *J. Biochem.*, 1973, **74**, 757.
[124] M. Sugiura, A. Ito, T. Ogiso, K. Kato, and H. Asano, *Biochim. Biophys. Acta*, 1973, **309**, 357.
[125] J. P. Brockes, *Biochem. J.*, 1973, **133**, 629.
[126] E. F. Wallace, M. J. Krantz, and W. Lovenberg, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2253.
[127] B. I. Gerwin and R. H. Bassin, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2453.
[128] C. Kandaswami and C. S. Vaidyanathan, *Arch. Biochem. Biophys.*, 1973, **155**, 247.
[129] R. Somack and R. N. Costilow, *J. Biol. Chem.*, 1973, **248**, 385.
[130] M. F. Goodman and M. J. Bessman, *J. Biol. Chem.*, 1973, **248**, 815.
[131] L. M. S. Chang, *J. Biol. Chem.*, 1973, **248**, 3789.
[132] P. Modrich, Y. Anraku, and I. R. Sehman, *J. Biol. Chem.*, 1973, **248**, 7495.
[133] S. Arvidson, *Biochim. Biophys. Acta*, 1973, **302**, 149.
[134] B. F. Klapper, D. M. Jameson, and R. M. Mayer, *Biochim. Biophys. Acta*, 1973, **304**, 505.
[135] S. K. Oh and J. M. Brewer, *Arch. Biochem. Biophys.*, 1973, **157**, 491.
[136] J. Baratti, S. Maroux, D. Louvard, and P. Desnuelle, *Biochim. Biophys. Acta*, 1973, **315**, 147.
[137] E. Stellwagen, M. M. Crossland, and L. D. Barnes, *Biochemistry*, 1973, **12**, 1552.
[138] L. D. Barnes and E. Stellwagen, *Biochemistry*, 1973, **12**, 1559.
[139] P. Cuatracasas, I. Parikh, and M. D. Hollenberg, *Biochemistry*, 1973, **12**, 4253.
[140] W. Kisiel and D. J. Hanahan, *Biochim. Biophys. Acta*, 1973, **304**, 103.
[141] J. S. Chen, J. S. Multani, and L. E. Mortenson, *Biochem. Biophys. Acta*, 1973, **310**, 51.
[142] T. Sasaki and H. Noda, *Biochim. Biophys. Acta*, 1973, **310**, 76.

Table 1 (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
Fibrinogen fragment d	Human	46 000	14 000	–	143
Factor X	Human plasma	56 000	—	–	144
Factor X	Pig serum	39 000	—	–	145
Four-iron ferrodoxin	*Desulphovibrio desulphuricans*	6 000	—	–	146
Factor X	Bovine	48 000	30 000 17 000	–	147
Factor IIIm	*Clostridium thermoaceticum*	—	27 000	–	148
Ferritin	Tadpole	478 000	19 000	–	149
Factor X	Bovine	56 000	44 000 15 000	–	150
Factor IX	Bovine	55 400	—	–	151
β-Glucosidase	*Dictyostelium discoideum*	225 000	—	–	152
α-Glucosidase	*Dictyostelium discoideum*	150 000	84 000	–	153
γ-Glutamylcysteine synthetase	Rat liver	138 000	—	–	154
Glucoside transferase	Yeast	122 000 85 000 70 000	—	–	155
	Rabbit	120 000	—	–	
Glutathione peroxidase	—	84 000	21 000	–	156
Glucose-6-phosphate isomerase	Ehrlich ascites tumour cells	110 000	—	–	157
Glyceraldehyde-3-phosphate dehydrogenase	*Thermus aquaticus*	—	36 000	–	158
1,4-α-Glucan phosphorylase	Sweet corn	315 000	150 000	–	159
Glucoamylase	Human intestine	210 000	—	–	160
γ-Glutamyl cyclotransferase	*Musca domestica*	30 000	—	–	161
Glycine reductase component A	*Clostridium stricklandii*	12 000	—	–	162
Glucose receptor	Rat kidney cortex	60 000	30 000	–	163
Glycerol-3-phosphate dehydrogenase	Rabbit	75 000	37 500	–	164
Glyceraldehyde-3-phosphate dehydrogenase	*Bacillus cereus*	144 000	36 000	–	165

α_1-Glycoprotein	Bullfrog serum	200 000	—	–	166
Glutamate dehydrogenase (NADP specific)	*Neurospora* sp.	288 400	48 800	–	167
β-Glucosidase	*Aspergillus fumigatus*	40 800	—	–	168
β-Galactosidase	Bovine testes	68 000	—	–	169
Galactose dehydrogenase	*Pseudomonas saccharophilus*	103 000	25 000	–	170
HL-A Antigens	Human platelets	> 200 000	15 000—30 000	–	171
HL-A Antigens	Human lymphocytes	A_2, 30 000 A_7, 11 000	—	–	172

143 M. Furhan and E. A. Beck, *Biochim. Biophys. Acta*, 1973, **310**, 205.
144 M. P. Esnouf, P. H. Lloyd, and J. Jesty, *Biochem. J.*, 1973, **131**, 781.
145 R. J. Dupe and R. M. Howell, *Biochem. J.*, 1973, **133**, 311.
146 J. A. Zabieta, R. Mason, and J. R. Postgate, *Biochem. J.*, 1973, **133**, 851.
147 R. D. Radcliffe and P. G. Barton, *Arch. Biochem. Biophys.*, 1973, **155**, 381.
148 L. G. Ljringdahl, J. Lelgall, and J. P. Lee, *Biochemistry*, 1973, **12**, 1802.
149 E. C. Theil, *J. Biol. Chem.*, 1973, **248**, 622.
150 S. P. Bajaj and K. G. Mann, *J. Biol. Chem.*, 1973, **248**, 7729.
151 A. Fujikawa, A. R. Thompson, M. E. Legaz, R. G. Meyer, and E. W. Davie, *Biochemistry*, 1973, **12**, 4938.
152 D. Every and J. M. Ashworth, *Biochem. J.*, 1973, **133**, 37.
153 See ref. 152.
154 J. S. Davis, J. B. Balinsky, J. S. Harrington, and J. B. Shepherd, *Biochem. J.*, 1973, **133**, 667.
155 E. Y. C. Lee and J. H. Carter, *F.E.B.S. Letters*, 1973, **32**, 78.
156 L. Flohe, W. A. Gunzler, and H. A. Schock, *F.E.B.S. Letters*, 1973, **32**, 132.
157 K. H. Rohm and F. Schneider, *F.E.B.S. Letters*, 1973, **33**, 89.
158 J. D. Hocking and J. I. Harris, *F.E.B.S. Letters*, 1973, **34**, 280.
159 E. Y. C. Lee and J. J. Braun, *Arch. Biochem. Biophys.*, 1973, **156** 276.
160 J. J. Kelly and D. H. Alpers, *Biochim. Biophys. Acta*, 1973, **315**, 115.
161 R. P. Bodnaryk and L. Meljirr, *Biochim. Biophys. Acta*, 1973, **315**, 352.
162 D. C. Turner and T. C. Stadtman, *Arch. Biochem. Biophys.*, 1973, **154**, 366.
163 L. Thomas, *Biochim. Biophys. Acta*, 1973, **291**, 454.
164 D. L. Warkenstein and T. P. Fondy, *European J. Biochem.*, 1973, **36**, 97.
165 K. Suzuki and K. Imahori, *J. Biochem.*, 1973, **73**, 97.
166 H. Nagano, T. Shimada, and R. Shukuya, *J. Biol. Chem.*, 1973, **248**, 5774.
167 K. M. Blumenthal and E. L. Smith, *J. Biol. Chem.*, 1973, **248**, 6002.
168 M. J. Rudick and A. D. Elbein, *J. Biol. Chem.*, 1973, **248**, 6506.
169 J. J. Distler and G. W. Jourdian, *J. Biol. Chem.*, 1973, **248**, 6772.
170 F. Wengenmeyer, K. H. Uberschar, G. Kurz, and H. Sund, *European J. Biochem.*, 1973, **40**, 49.
171 A. Dautigny, I. Bernier, J. Colombani, and P. Jolles, *Biochim. Biophys. Acta*, 1973, **298**, 783.
172 P. Cresswell, M. J. Turner, and J. L. Strominger, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1603.

Table 1 (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
Hexose oxidase	*Chondrus crispus*	130 000	—	–	173
Hyaluronidase		56 000	14 000	–	174
Hydroxykynureninase	Mouse liver	85 000	—	–	175
Histidinal phosphate phosphatase	Yeast	38 000	—	–	176
L(+)-3-Hydroxybutyryl-CoA dehydrogenase	*Clostridium kluyveri*	220 000	30 000	–	177
Hexokinase	Pig heart	297 000	—	–	178
L-Histidinal phosphate amino-transferase	*Salmonella typhimurium*	74 000	37 000	–	179
Histones	*Neurospora crassa*	14 000 8 000	—	–	180
L-3-Hydroxyacyl-CoA dehydrogenase	Pig heart	75 000	31 000	–	181
Haemoglobin	*Candida mycoderma*	50 000	–	–	182
Intrinsic factor	Human	—	30 000 60 000	–	183
Isolectins	Wax Bean	120 000	30 000		184
Intrinsic factor	Human	63 000	—	–	185
Initiation factor	*Cocos nucifera*	76 000	—	–	186
Isocitrate dehydrogenase	*Thermus aquaticus*	~77 500	—	–	187
Isocitrate dehydrogenase	*Bacillus subtilis*-168	~95 000	—	–	188
Isocitrate dehydrogenase	*Chlamydomonas reinhardtii*	~95 000	—	–	189
Immunoglobulin IgG(Ud)	Human	176 000	—	–	190
Isoleucyl-tRNA synthetase	*E. coli*	102 000	102 000	–	191
Initiation factor	Rabbit liver	—	M_1, 96 000 M_{26}, 21 000	–	192
Initiation factor	*E. coli*	98 000 83 000	—	–	193
Isoprothrombin	Rat	85 000	—	–	194
Intrinsic factor	Hog	~55 000	—	–	195
J chain	Human	24 000	—	–	196
κ-Casein	Goat	18 400	—	–	197
Lysophospholipase	Bovine pancreas	65 000	—	–	198

Lysophospholipase	*Penicillium notatum*	116 000	—	–	199
β-Lactamase	*Klebsiella aerogenes*	23 000	—	–	200
	Enterobacter cloacae	49 000		–	
Lactate dehydrogenase	*Leuconostoc mesenteroides*	64 000	—	–	201
Lymphotoxin	Human	~95 000	—	–	202
Lipopolysaccharide protein	*E. coli*	—	14 000	–	203

[173] J. D. Sullivan and M. Ikawa, *Biochim. Biophys. Acta*, 1973, **309**, 11.
[174] A. Y. Khorlin, I. V. Vikha, and A. N. Milishnikov, *F.E.B.S. Letters*, 1973, **31**, 107.
[175] C. E. McDermott, D. A. Casciano, and F. H. Gaertner, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 813.
[176] R. H. Millay and L. L. Houston, *Biochemistry*, 1973, **12**, 2591.
[177] V. K. Madan, P. Hillmer, and G. Gottschalk, *European J. Biochem.*, 1973, **32**, 51.
[178] J. S. Easterby and M. J. O'Brien, *European J. Biochem.*, 1973, **38**, 201.
[179] G. B. Hendesson and E. E. Snell, *J. Biol. Chem.*, 1973, **248**, 1906.
[180] M. W. Hsiang and R. D. Cole, *J. Biol. Chem.*, 1973, **248**, 2007.
[181] B. E. Noyes and R. A. Bradshaw, *J. Biol. Chem.*, 1973, **248**, 3052.
[182] R. Oshino, T. Asakura, K. Takio, N. Oshino, B. Chance, and B. Hagihara, *European J. Biochem.*, 1973, **39**, 581.
[183] J. M. Christensen, E. Hippe, H. Olesen, M. Rye, E. Haber, L. Lee, and J. Thomson, *Biochim. Biophys. Acta*, 1973, **303**, 319.
[184] B. A. Sela, H. Lis, N. Sharon, and L. Sachs, *Biochim. Biophys. Acta*, 1973, **310**, 273.
[185] K. Uisuri and R. Grasbeck, *Biochim. Biophys. Acta*, 1973, **310**, 508.
[186] A. Ganguly, A. Das, H. Mondal, R. K. Mondal, and B. B. Biowas, *F.E.B.S. Letters*, 1973, **34**, 27.
[187] R. F. Ramaley and M. O. Hudock, *Biochim. Biophys. Acta*, 1973, **315**, 22.
[188] See ref. 187.
[189] See ref. 187.
[190] K. Bobrzecka, L. Konieczny, J. Rybarska, J. Bogdel, and I. Urasinski, *Biochim. Biophys. Acta*, 1973, **295**, 564.
[191] A. Durekovic, J. Flossdorf, and M. R. Kula, *European J. Biochem.*, 1973, **36**, 528.
[192] D. J. Picciano, P. M. Bichard, W. C. Merrick, D. A. Shafritz, H. Graf, R. G. Crystal, and W. F. Anderson, *J. Biol. Chem.*, 1973, **248**, 204.
[193] M. J. Miller and A. J. Wahba, *J. Biol. Chem.*, 1973, **248**, 1084.
[194] J. J. Morrissey, J. P. Jones, and R. E. Olsen, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 1075.
[195] R. H. Allen and C. S. Mehlmen, *J. Biol. Chem.*, 1973, **248**, 3660.
[196] K. Kobayashi, J. P. Vaerman, and J. F. Heremans, *Biochim. Biophys. Acta*, 1973, **303**, 105.
[197] B. C. Richardson, L. K. Creamer, and R. E. Munford, *Biochim. Biophys. Acta*, 1973, **310**, 111.
[198] J. G. N. deJong, H. van den Bosch, A. J. Aarsman, and L. L. M. van Deenan, *Bio chim. Biophys. Acta*, 1973, **296**, 94.
[199] N. Kawasaki and K. Saito, *Biochim. Biophys. Acta*, 1973, **296**, 426.
[200] G. W. Ross and M. G. Boulton, *Biochim. Biophys. Acta*, 1973, **309**, 430.
[201] R. C. Garland, *Arch. Biochem. Biophys.*, 1973, **157**, 36.
[202] G. A. Granger, E. C. Laserna, W. P. Kolb, and F. Chapman, *Proc. Nat. Acad. Sci., U.S.A.* 1973, **70**, 27.
[203] M. C. Wu and E. C. Heath, *Proc. Nat. Acad. Sci., U.S.A.*, 1973, **70**, 2572.

Table 1 (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
β-Lysine mutase	*Clostridium stricklandii*	170 000	32 000 50 000	–	204
Lactate dehydrogenase	*E. coli* membranes	72 000	—	–	205
α-Lactalbumin	Bovine	—	—	–	206
Leucoagglutinin	Kidney bean	126 000	31 000	–	207
Lactose phosphotransferase	*Staphylococcus aureus*	~80 000	—	–	208
Lactose specific phospho carrier	*Staphylococcus aureus*	35 000	—	–	209
Leucyl-tRNA synthetase	Baker's yeast	120 000	60 000	–	210
Lecithin β-hydroxybutyrate dehydrogenase	Beef heart mitochondria	—	32 000		211
Lysins	*Megathura crenulata*	57 250 53 688 58 000	—	–	212
Lipoproteins (high density)	*Macacus rhesus*	27 000 8 500	—	–	213
Myosin fragment	Rabbit	26 000	—	–	214
α-Mannosidase	*Dictyostelium discoideum*	280 000	—	–	215
Monellin	*Dioscoreophyllum cumminssii*	11 500	—	–	216
Malic enzyme isozymes	Pome fruits	4.55*S* 5.45*S*	—	–	217
Messenger discriminating factor	*E. coli*	α, 23 000 β, 21 000	—	–	218
Mammalian cell protein (P8)	Mouse Human	—	28 000	–	219
Myosin light chains	Scallop	18 000	—	–	220
M. protein (RNA polymerase)	*E. coli*	85 000	—	–	221
Myrosinases	*Brassica napus*	135 000	65 000	–	222
Myosin A	*Ascaris lumbricoides*	6*S*	—	–	223
Myelin basic protein	Foetal calf brain	18 100	—	–	224
Myosin rods	Rabbit	205 000	—	–	225
Muconate lactonizing enzyme	*Pseudomonas putida*	190 000	40 000	–	226
Muconolactone isomerase	*Pseudomonas putida*	—	12 000	–	227

Messenger-specific initiation factor	Krebs ascites cells	—	53 000	—	228
α-Mannosidase	Hog kidney	—	42 000	—	229
Malic enzyme (NAD)	*E. coli*	200 000	55 000	—	230
Methyltetrahydrofolate dehydrogenase	*Clostridium thermoaceticum*	55 000	25 000	—	231
Glycine methyltransferase	Rabbit liver	123 000	~30 000	—	232
Melilotate hydroxylase	*Psuedomonas* sp.	250 000	65 000	—	233

[204] J. J. Baker, C. van der Drift, and T. C. Stadtman, *Biochemistry*, 1973, **12**, 1054.
[205] R. C. Haberkorn and G. L. Cantoni, *Biochemistry*, 1973, **12**, 2389.
[206] T. E. Barman, *European J. Biochem.*, 1973, **37**, 86.
[207] V. Rasanen, T. H. Weber, and R. Grasbeck, *European J. Biochem.*, 1973, **38**, 193.
[208] R. D. Simoni, T. Nakazawa, J. B. Hays, and S. Roseman, *J. Biol. Chem.*, 1973, **248**, 932.
[209] J. B. Hays, R. D. Simoni, and S. Roseman, *J. Biol. Chem.*, 1973, **248**, 941.
[210] J. G. Chirikjian, K. Kanagalingam, E. Lau, and J. R. Fresco, *J. Biol. Chem.*, 1973, **248**, 1074
[211] H. M. Menzel and G. G. Hammes, *J. Biol. Chem.*, 1973, **248**, 4885.
[212] E. Heller and M. A. Raftery, *Biochemistry*, 1973, **12**, 4106.
[213] A. M. Scanu, C. Edelstein, L. Vitello, R. Jones, and R. Weisler, *J. Biol. Chem.*, 1973, **248**, 7648.
[214] D. Stone and S. V. Perry, *Biochem. J.*, 1973, **131**, 127.
[215] D. Every and J. M. Ashworth, *Biochem. J.*, 1973, **133**, 37.
[216] K. Van Der Wel and K. Loeve, *F.E.B.S. Letters*, 1973, **29**, 181.
[217] H. Komai and R. A. Capaldi, *F.E.B.S. Letters*, 1973, **30**, 273.
[218] S. L. Huang and S. Ochoa, *Arch. Biochem. Biophys.*, 1973, **156**, 84.
[219] R. L. Tsai and H. Green, *J. Mol. Biol.*, 1973, **73**, 307.
[220] A. G. Szent-Gyorgyi, E. M. Szentkiralyi, and J. Kendrick-Jones, *J. Mol. Biol.*, 1973, **74**, 179.
[221] T. Ramakrishnan and H. Echols, *J. Mol. Biol.*, 1973, **78**, 675.
[222] B. Lonnerdal and J. C. Janson, *Biochim. Biophys. Acta*, 1973, **315**, 421.
[223] M. Yamaguchi, T. Nakanura, H. Oya, and T. Sekine, *Biochim. Biophys. Acta*, 1973, **317**, 312.
[224] L. R. Hegstrand and S. E. Kornguth, *Biochim. Biophys. Acta*, 1973, **317**, 380.
[225] C. Driol, M. F. Landon, and N. van Thoai, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 1023.
[226] R. B. Meagher and L. N. Dinston, *Biochemistry*, 1973, **12**, 3523.
[227] See. ref. 226.
[228] D. J. Wigle, *F.E.B.S. Letters*, 1973, **35**, 11.
[229] T. Okumura and I. Yamashina, *J. Biochem.*, 1973, **73**, 131.
[230] See ref. 229.
[231] W. E. O'Brien, T. M. Brewer, and L. G. Ljungdahl, *J. Biol. Chem.*, 1973, **248**, 403.
[232] J. E. Heady and S. J. Kerr, *J. Biol. Chem.*, 1973, **248**, 69.
[233] S. Strickland and V. Massey, *J. Biol. Chem.*, 1973, **248**, 2944.

Table 1 (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
Myosin	*Acanthamoeba castellanii*	180 000	140 000 16 000 14 000	–	234
NAD nucleosidase	*Fusarium nivale*	27 000	—	–	235
Neutral proteinase	*Staphylococcus aureus*	21 000	—	–	236
NADH dehydrogenase	*Halobacterium cutirubrum*	64 000	—	–	237
5′-Nucleotidase	*Micrococcus radiodurans*	44 000	—	–	238
5′-Nucleotidase	Mouse liver plasma membranes	160 000	70 000	–	239
Non-histone protein	Histone F_1	20 000	—	–	240
N-Acetyl hexoamidase	*Dictyostelium discoideum*	168 000	65 000 51 000	–	241
Nitrite reductase (NADH)	*Azotobacter chroococcum*	67 000	—	–	242
α,β-*N*-Acetyl glucosaminidase	*Patella vulgata*	—	82 000 54 000	–	243
Nucleotide phosphotransferase	*E. coli*	29 000	—	–	244
Nuclease, single strand specific	*Aspergillus oryzae*	32 000	—	–	245
Neutral protease (collagenase)	Human granulocytes	76 000	42 000 32 000	–	246
NADPH-sulphite reductase	*E. coli*	670 000	—	–	247
NADH-cytochrome b_5 reductase	Calf liver	43 000	—	–	248
Neurotensin	Bovine hypothalamus	1 600	—	–	249
α-Ornithine aminomutase	*Clostridium stricklandii*	180 000	90 000	–	250
Pentose-rich glycoprotein	Bovine vitreous body	25 000	—	–	251
Prolylcarboxypeptidase angiotensinase C	Hog kidney cortex	210 000	23 500	–	252
Phycoerythrins	Red algae	250 000 44 000	20 000 23 000	–	253
Phytohaemagglutinin	Sun hemp seeds	120 000	—	–	254
Palmityl-CoA synthetase	Rat liver microsomes	168 000	27 000	–	255
Pepsinogen	Chicken	36 000	—	–	256
Pepsin	Chicken	34 000	—	–	
Purine nucleoside phosphorylase	Mammalian	84 000	28 000	–	257

Phenylethanolamine-*N*-methyl transferase	Rat adrenal gland	40 000	—	–	258
Post-γ-globulin	Human	11 500	—	–	259
Prephenoloxidase-activating enzyme	*Bombyx mori*	—	35 000	–	260
Pre-keratin		—	47 000 58 000	–	261
Phytomitogen	*Phytolacca esculenta*	32 000	—	–	262
Progesterone–testostesterone binding protein	Guinea pig serum	—	82 800	–	263

[234] T. D. Pollard and E. D. Korn, *J. Biol. Chem.*, 1973, **248**, 4682.
[235] D. Stathakos, I. Isaakidou, and H. Thomou, *Biochim. Biophys. Acta*, 1973, **302**, 80.
[236] S. Arvidson, T. Hohme, and B. Lundholm, *Biochim. Biophys. Acta*, 1973, **302**, 135.
[237] L. I. Hochstein and B. P. Dalton, *Biochim. Biophys. Acta*, 1973, **302**, 216.
[238] R. E. J. Mitchel, *Biochim. Biophys. Acta*, 1973, **309**, 116.
[239] W. H. Evans and J. W. Gurd, *Biochem. J.*, 1973, **133**, 189.
[240] J. A. Smith and L. A. Stocken, *Biochem. J.*, 1973, **131**, 859.
[241] D. Every and J. M. Ashworth, *Biochem. J.*, 1972, **133**, 37.
[242] J. M. Vega, M. G. Guerrero, E. Leadbetter, and M. Losada, *Biochem. J.*, 1973, **133**, 701.
[243] J. V. Bannister and P. J. R. Phizackerley, *F.E.B.S. Letters*, 1973, **34**, 120.
[244] E. F. Brenngraber and E. Chargaff, *Biochemistry*, 1973, **12**, 3005.
[245] V. M. Vogt, *European J. Biochem.*, 1973, **33**, 192.
[246] K. Ohlsson and I. Ohlsson, *European J. Biochem.*, 1973, **36**, 473.
[247] L. M. Siegel, M. J. Murphy, and H. Kanin, *J. Biol. Chem.*, 1973, **248**, 251.
[248] L. Spatz and P. Strittmatter, *J. Biol. Chem.*, 1973, **248**, 793.
[249] R. Carraway and S. E. Leeman, *J. Biol. Chem.*, 1973, **248**, 6854.
[250] R. Somack and R. N. Costilow, *Biochemistry*, 1973, **12**, 2597.
[251] W. S. Allen and A. H. Wardi, *Biochim. Biophys. Acta*, 1973, **297**, 1.
[252] T. Kakimoto, G. Oshima, H. S. J. Yeh, and E. G. Erdos, *Biochim. Biophys. Acta*, 1973, **302**, 178.
[253] H. H. Van der Velde, *Biochim. Biophys. Acta*, 1973, **303**, 246.
[254] B. Ersson, K. Aspberg, and J. Porath, *Biochim. Biophys. Acta*, 1973, **310**, 446.
[255] J. Bar-Tana and G. Rose, *Biochem. J.*, 1973, **131**, 443.
[256] M. L. Green and J. M. Llewellin, *Biochem. J.*, 1973, **133**, 105.
[257] Y. H. Edwards, P. A. Edwards, and D. A. Hopkinson, *F.E.B.S. Letters*, 1973, **32**, 235.
[258] L. A. Poharecky and B. S. Baliga, *Arch. Biochem. Biophys.*, 1973, **156**, 703.
[259] J. Cejka and L. E. Fleischmann, *Arch. Biochem. Biophys.*, 1973, **157**, 168.
[260] K. Dohke, *Arch. Biochem. Biophys.*, 1973, **157**, 210.
[261] H. P. Baden, L. A. Goldsmith, and B. Fleming, *Biochim. Biophys. Acta*, 1973, **317**, 303.
[262] H. Tokuyama, *Biochim. Biophys. Acta*, 1973, **317**, 338.
[263] O. A. Lea, *Biochim. Biophys. Acta*, 1973, **317**, 351.

Table 1 (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
Phosvitin kinase	Calf brain	—	41 000	–	264
Prohistidine decarboxylase	*Lactobacillus* 30a	—	37 000	–	265
Pyruvate phosphate dikinase	Maize leaf	387 000	94 000	–	266
Phoratoxin	*Phoradendron tomentorum*	50 000	—	–	267
Phosphodiesterase	*Bothrops atrox*	130 000	—	–	268
6-Phosphogluconate dehydrogenase	Sheep liver	94 000	45 000	–	269
6-Phosphogluconate dehydrogenase	*Neurospora*	115 000	57 000	–	270
Phosphofructokinase	Yeast	570 000	130 000 96 000	–	271
Pyocin S2	*Pseudomonas aeruginosa*	78 000	—	–	272
Progesterone-binding protein	Guinea pig	77 500	—	–	273
Protocollagen proline hydroxylase	Chick embryos	230 000	60 000 64 000	–	274
Phenylalanine hydroxylase stimulating protein	Rat liver	51 500	12 500	–	275
Poly(A) polymerase	Calf thymus	150 000	—	–	276
Pyruvate kinase	Bovine skeletal muscle	230 000	57 000	–	277
Phycocyanin	*Chroomonas* sp.	50 000	10 000 16 000	–	278
Phosphohydrolases	*Bacillus subtilis*	150 000	75 000	–	279
Pyruvate kinase	Pig liver	202 000	50 000	–	280
Prothrombin	Bovine	72 000	—	–	281
Progesterone-induced glycoprotein	Pig uterine fluid	32 000	—	–	282
Phosphofructokinase	Chicken liver	400 000	60 000	–	283
RNA polymerase I	*Cocos nucifera*	730 000	175 000 150 000 95 000 80 000	–	284
RNA polymerase	*Physarum polycephalium*	—	A, 185 000, 135 000 B, 175 000, 145 000	–	285

Ribonuclease H	Calf thymus	64 000	—	–	286
RNA polymerase	*Bacillus amyloliquefaciens*	428 000	170 000 48 000 40 000	–	287
RNA polymerase II	Maize	—	200 000 160 000 35 000 25 000 20 000 17 000	–	288
Ribonuclease (hybrid specific)	Calf thymus	74 000	—	–	289

[264] O. Walinder, *Biochim. Biophys. Acta*, 1973, **293**, 140.
[265] P. A. Recsei and E. E. Snell, *Biochemistry*, 1973, **12**, 365.
[266] T. Sugiyama, *Biochemistry*, 1973, **12**, 2862.
[267] S. T. Mellstrand and G. Samuelsson, *European J. Biochem.*, 1973, **32**, 143.
[268] A. M. Frischauf and F. Eckstein, *European J. Biochem.*, 1973, **32**, 479.
[269] M. Silverberg and K. Dalziel, *European J. Biochem.*, 1973, **38**, 229.
[270] W. A. Scott and T. Abramsky, *J. Biol. Chem.*, 1973, **248**, 3535.
[271] W. Diezel, H. J. Bohme, K. Nissler, R. Freger, W. Heilmann, G. Kopperschlager, and E. Hofmann, *European J. Biochem.*, 1973, **38**, 479.
[272] I. Ohkawa, M. Kagoyama, and F. Egami, *J. Biochem.*, 1973, **73**, 281.
[273] E. Milgrom, P. Allouch, M. Atger, and E. E. Baulieu, *J. Biol. Chem.*, 1973, **248**, 1106.
[274] R. A. Berg and D. J. Prockop, *J. Biol. Chem.*, 1973, **248**, 1175.
[275] C. Y. Huang, E. E. Max, and S. Kaufman, *J. Biol. Chem.*, 1973, **248**, 4235.
[276] M. A. Winters and M. Edwards, *J. Biol. Chem.*, 1973, **248**, 4756.
[277] J. M. Cardenas, R. D. Dyson, and J. J. Strandholm, *J. Biol. Chem.*, 1973, **248**, 6931.
[278] R. MacColl, W. Habig, and D. S. Berns, *J. Biol. Chem.*, 1973, **248**, 7080.
[279] J. C. Le Hegarat and C. Anagnotopoulos, *European J. Biochem.*, 1973, **39**, 525.
[280] C. Kutzbach, H. Bischofberger, B. Hess, and H. Zimmermann-Telschow, *Z. physiol. Chem.*, 1973, **354**, 1473.
[281] S. P. Bajaj and K. G. Mann, *J. Biol. Chem.*, 1973, **248**, 7729.
[282] T. T. Chen, F. W. Bazer, J. J. Cetorelli, W. E. Pollard, and R. M. Roberts, *J. Biol. Chem.*, 1973, **248**, 8560.
[283] N. Kono, K. Uyeda, and R. M. Oliver, *J. Biol. Chem.*, 1973, **248**, 8592.
[284] A. Ganguly, A. Das, H. Mondal, R. K. Mondal, and B. B. Biowas, *F.E.B.S. Letters*, 1973, **34**, 27.
[285] A. Hildebrandt and H. W. Sauer, *F.E.B.S. Letters*, 1973, **35**, 41.
[286] W. Wickner, R. Sheckman, K. Geider, and A. Kornberg, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1764.
[287] M. Holland and H. R. Whiteley, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2234.
[288] K. P. Mullinix, G. C. Strain, and L. Bogorad, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2386.
[289] R. C. Haberkorn and G. L. Cantoni, *Biochemistry*, 1973, **12**, 2389.

Table 1 (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
Ribose phosphate isomerase	*Candida utilis*	105 000	25 000	–	290
Rhodopsinoprotein	Squid	—	49 000	–	291
RNA polymerase	HeLa cells	500 000	220 000 140 000 35 000 25 000	–	292
RNA polymerase	*E. coli* (*T4* infected)	105 000	—	–	293
Ribonuclease A, B, and C	Ovine pancreas	—	—	–	294
SFV-membrane protein precursor	Semliki forest virus	—	—	–	295
Sucrose synthetase	Rice	400 000	100 000	–	296
Supernatant initiation factor	*Artemia saliva*	148 000	74 000	–	297
Squalene/sterol carrier protein	—	—	16 000	–	298
Sulphatase B	Ovine brain	60 000	—	–	299
Structural proteins	Spring viremia virus	—	150 000 70 000 40 000 19 000	–	300
Seryl-tRNA synthetase	*E. coli* B	103 000	53 000	–	301
Superoxide dismutase	Spinach leaves	32 000	16 000	–	302
Specific basic protein	Rat testes	—	13 000	–	303
Superoxide dismutase	*E. coli*	38 700	20 000	–	304
Tropomysin	Rabbit muscle	—	32 000 36 000	–	305
Troponin inhibitor protein	Rabbit muscle	46 000	23 000	–	306
Threonine deaminase	*E. coli*	190 000	50 000	–	307
Threonine deaminase	*E. coli*	204 000	51 000	–	308
Tryptophanase	*Aeromonas liquefaciens*	216 000	—	–	309
Tropoelastin	Chick	—	—	–	310
Tetanus toxin	—	—	95 000 55 000	–	311

Toxin	Castor beans	55 000	32 000 28 000	—	312
Thiamine-binding protein	*E. coli*	39 000	—	—	313
α-Tocopherol binding protein	Rat liver	—	—	—	314
Tryptophan oxygenase	*Pseudomonas acidovorans*	—	—	—	315
Tubulin	Mouse brain	—	57 000	—	316
α-Toxin	*Staphylococcus aureus*	28 000	—	—	317
Trypsin (anionic)	Human pancreas	25 800	—	—	318

[290] G. F. Domagk, K. M. Doering, and R. Chilla, *European J. Biochem.*, 1973, **38**, 259.
[291] F. M. Hagins, *J. Biol. Chem.*, 1973, **248**, 3298.
[292] B. Sugden and W. Keller, *J. Biol. Chem.*, 1973, **248**, 3777.
[293] P. R. Chakraburty, P. Sarkar, H. H. Huang, and U. Maitra, *J. Biol. Chem.*, 1973, **248**, 6637.
[294] R. R. Becker, J. L. Holbrook, and C. H. W. Hirs, *J. Biol. Chem.*, 1973, **248**, 7826.
[295] K. Simons, S. Keranen, and L. Kaarianien, *F.E.B.S. Letters*, 1973, **29**, 87.
[296] T. Nomura and T. Akazawa, *Arch. Biochem. Biophys.*, 1973, **156**, 644.
[297] M. Zasloff and S. Ochoa, *J. Mol. Biol.*, 1973, **73**, 65.
[298] M. C. Ritter and M. E. Dempsey, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 265.
[299] W. S. Bleszynski and A. B. Roy, *Biochim. Biophys. Acta*, 1973, **317**, 164.
[300] G. Lenoir, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 895.
[301] D. E. Boeker, A. P. Hays, and G. L. Cantoni, *Biochemistry*, 1973, **12**, 2379.
[302] K. Asada, M. Vrano, and M. Takahashi, *European J. Biochem.*, 1973, **36**, 257.
[303] W. S. Kistler, M. E. Geroch, and H. G. Williams-Ashman, *J. Biol. Chem.*, 1973, **248**, 4532.
[304] F. J. Yost and I. Fridovich, *J. Biol. Chem.*, 1973, **248**, 4905.
[305] P. Cummin and S. V. Perry, *Biochem. J.*, 1973, **133**, 765.
[306] R. S. Mani, W. D. McCubbin, and C. M. Kay, *F.E.B.S. Letters*, 1973, **29**, 243.
[307] H. Grimminger, I. Rahimi-Laridjani, K. Sloerner, and F. Lingens, *F.E.B.S. Letters*, 1973, **35**, 273.
[308] J. L. Cowell, K. Moser, and R. D. De Moss, *Biochim. Biophys. Acta*, 1973, **315**, 449.
[309] R. B. Rucker, W. G. Riemann, and K. Tarn, *Biochim. Biophys. Acta*, 1973, **317**, 193.
[310] C. J. Craven and D. J. Dawson, *Biochim. Biophys. Acta*, 1973, **317**, 277.
[311] L. G. Gurtler and H. J. Horstmann, *Biochim. Biophys. Acta*, 1973, **295**, 582.
[312] A. Matsimura, A. Imashima, and Y. Nose, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 241.
[313] O. V. Rajaram, P. Fatterpaker, and A. Sreenivasen, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 459.
[314] E. M. Diamond, A. Jedeikin, and J. E. Kench, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 679.
[315] N. O. Hinman, J. L. Morgan, N. W. Seeds, and J. R. Cann, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 752.
[316] H. R. Six and S. Harshman, *Biochemistry*, 1973, **12**, 2672.
[317] P. A. Mallory and J. Travis, *Biochemistry*, 1973, **12**, 2847.
[318] R. Tsukui and S. Ebashi, *J. Biochem.*, 1973, **73**, 1119.

Table 1 (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
Troponin	Bovine cardiac muscle	—	41 000 28 000 18 000	–	319
Thioredoxin	Rat Novikoff ascites hepatoma	11 400	—	–	320
Troponin	Rabbit muscle	—	37 000 24 000 20 000	–	321
Tryptophan synthetase B protein	*Bacillus subtilis*	82 000	41 000	–	322
Tropomyosin	Chick embryo brain	—	30 000	–	323
Thionein	Rat and chicken liver	12 000	—	–	324
UDP apiose/UDP xylose synthetase	Parsley	101 000	—	–	325

[319] E. C. Herrmann and E. C. Moore, *J. Biol. Chem.*, 1973, **248**, 1219.
[320] M. L. Greaser and J. Gergerly, *J. Biol. Chem.*, 1973, **248**, 2125.
[321] S. O'N. Hoch, *J. Biol. Chem.*, 1973, **248**, 2992.
[322] D. H. Calhoun, R. A. Rimerman, and G. W. Hatfield, *J. Biol. Chem.*, 1973, **248**, 3511.
[323] R. E. Fine, A. L. Blitz, S. E. Hitchcock, and B. Kaminer, *Nature New Biol.*, 1973, **245**, 182.
[324] V. Meser, H. Rupp, F. Donay, F. Linnermann, W. Voelter, W. Voetsch, and G. Jung, *European J. Biochem.*, 1973, **39**, 127.
[325] D. Baron, U. Streitberger, and H. Grisebach, *Biochim. Biophys. Acta*, 1973, **293**, 526.

Table 1a *Molecular weights and subunit structures*

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
Collagenase	Human skin	63 000	—	–	326
Collagenase	Rheumatoid synovium	32 000	—	–	327
Haemocyanin	*Cancer magister*	—	76 000 84 000	–	328
Erythrocyte glycoprotein	Human	—	64 000 31 000	–	329
Phosphorylase kinase	Rabbit muscle	1.3×10^6	118 000 108 000 41 000	–	330
Transferrins	Bovine	77 300	—	+	331
	Rabbit	76 700	—	+	
	Equine	79 100	—	+	
	Porcine	76 400	—	+	
l-Asparagine	*Proteus vulgaris*	120 000	$4 \times 30\,000$	+	332
Glycogen phosphorylase *a*/*b*	Shark	190 000	$2 \times 95\,000$	+	333
Phenylalanine ammonia-lyase	Maize	—	—	+	334
	Potato	—	—		
Inosinic acid dehydrogenase	*E. coli*	250 000	62 000 44 000	–	335

[326] D. E. Woolley, R. W. Glanville, and J. M. Evanson, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 729.
[327] See ref. 326.
[328] J. S. Hoehr and H. S. Mason, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 741.
[329] L. S. G. Marton and J. E. Garvin, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 1457.
[330] T. Hayakawa, J. P. Perkins, D. A. Walsh, and E. G. Krebs, *Biochemistry*, 1973, **12**, 567.
[331] B. G. Hudson, M. Ohno, W. J. Brockway, and F. J. Castellino, *Biochemistry*, 1973, **12**, 1047.
[332] T. Tora, R. Sano, K. Yamamoto, M. Nakamura, and I. Chibata, *Biochemistry*, 1973, **12**, 1075.
[333] S. A. Assaf and A. A. Yunis, *Biochemistry*, 1973, **12**, 1423.
[334] E. A. Havier and K. R. Hanson, *Biochemistry*, 1973, **12**, 1583.
[335] G. F. Powell, *Biochemistry*, 1973, **12**, 1592.

Table 1a (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
Butyrylcholinesterase	Horse serum	440 000	4 × 110 000	−	336
Acetylcholinesterase	Eel electric tissue	—	60 000 75 000	+	337
	Bovine erythrocyte membrane	200 000	126 000 75 000	+	337
Aspartokinase	*E. coli*	200 000 100 000	4 × 48 000 2 × 48 000	−	338
Haemocyanin	*Cancer magister*	—	76 000 83 000	+	339
Immunoglobulin Ig-J chain	—	15 000	—	−	340
Sucrase–isomaltase complex	Rabbit intestine	220 000	2 × 110 000	+	341
Phosphorylase kinase	Rabbit muscle	1 280 000	145 000 128 000 45 000	−	342
Amidase	*P. aeruginosa*	200 000	6 × 33 000	+	343
Glycogen phosphorylase	Human muscle	180 000	2 × 94 500	+	344
Malate–lactate dehydrogenase	*Veillonella alcalescens*	70 000	2 × 35 000	−	345
Acetylcholinesterase	*Electrophorus electricus*	134 000	2 × 65 000	−	346
Elongation factors	*E. coli*	Tu-GDP 47 000 Ts 35 000	47 000 35 000	+	347
Monellin	*Dioscorcophyllum cumminsii*	10 000	10 000	+	348
Phosphofructokinase	Rabbit muscle	—	80 000	+	349
Inorganic pyrophosphatase	Yeast	—	2 × 32 000	+	350
Aspartokinase	*B. polymyxa*	—	43 000 17 000	−	351
Anthranilate synthetase	*Acinetobacter caleoaceticus*	86 000	70 000 14 000	−	352
Ribonucleoside diphosphate reductase	*E. coli*	245 000	80 000 78 000	+	353

α-Ketoglutarate dehydrogenase	*E. coli*	190 000	2 × 95 000	+	354
Dihydrolipoyl transuccinylase	*E. coli*	1 000 000	24 × 42 000	+	355
Dihydrolipoyl dehydrogenase	*E. coli*	112 000	2 × 56 000	−	356
Thyroglobulin	Calf	—	235 000 215 000 190 000 30 000	+	357
Nucleoside triphosphate–diphosphate transphosphorylase	Yeast	102 000	6 × 17 500	+	358
Tryptophanase	*Aeromonas liquefaciens*	210 000	4 × 54 000	−	359

[336] J. C. Lee and J. A. Hargst, *Biochemistry*, 1973, **12**, 1622.
[337] J. D. Bennan, *Biochemistry*, 1973, **12**, 1710.
[338] E. G. Niles and E. W. Westhead, *Biochemistry*, 1973, **12**, 1715.
[339] D. E. Carpenter and K. E. Van Holde, *Biochemistry*, 1973, **12**, 2231.
[340] C. E. Wilde and M. E. Koshland, *Biochemistry*, 1973, **12**, 3218.
[341] A. Cogoli, A. Eberle, H. Sigrist, C. Joss, E. Robinson, H. Morimann, and G. Semenza, *European J. Biochem.*, 1973, **33**, 40.
[342] P. Cohen, *European J. Biochem.*, 1973, **34**, 1.
[343] P. R. Brown, M. J. Smyth, P. H. Clarke, and M. A. Rosemeyer, *European J. Biochem.*, 1973, **34**, 177.
[344] S. A. Assaf and A. A. Yunis, *European J. Biochem.*, 1973, **35**, 282.
[345] S. H. G. Allen, *European J. Biochem.*, 1973, **35**, 338.
[346] D. B. Millar, M. A. Grafins, D. A. Palmer, and G. Millar, *European J. Biochem.*, 1973, **37**, 425.
[347] K. Asai, M. Kawakita, Y. Kaziro, T. Kondo, and N. Ui, *J. Biochem.*, 1973, **73**, 1095.
[348] J. A. Morris, R. Martenson, G. Deibler, and R. H. Cagan, *J. Biol. Chem.*, 1973, **248**, 534.
[349] C. J. Coffee, R. P. Aaronson, and C. Frieden, *J. Biol. Chem.*, 1973, **248**, 1381.
[350] R. L. Heinrikson, R. Sterner, C. Noyes, B. S. Cooperman, and R. H. Bruckmann, 1973, **248**, 2521.
[351] C. Biswas and H. Paulus, *J. Biol. Chem.*, 1973, **248**, 2894.
[352] R. V. Sawala and I. P. Crawford, *J. Biol. Chem.*, 1973, **248**, 3573.
[353] L. Thelander, *J. Biol. Chem.*, 1973, **248**, 4591.
[354] F. H. Pettit, L. Hamilton, P. Munk, G. Mamihira, M. H. Elay, C. R. Willis, and L. J. Reed, *J. Biol. Chem.*, 1973, **248**, 5282.
[355] See ref. 354.
[356] See ref. 354.
[357] M. J. Spiro, *J. Biol. Chem.*, 1973, **248**, 4446.
[358] R. Palmieri, R. H. Yue, H. K. Jacobs, L. Maland, L. Wu, and S. A. Kuby, *J. Biol. Chem.*, 1973, **248**, 4486.
[359] J. L. Cowell and R. D. De Moss, *J. Biol. Chem.*, 1973, **248**, 6262.

Table 1a (*cont.*)

Protein	*Source*	*Molecular weight*	*Subunit structure*	*Amino-acid analysis*	*Ref.*
Serine transhydroxymethylase	—	200 000	4 × 53 000	+	360
Microtubule protein	Calf brain	—	54 000	+	361
Mitochondrial adenosine triphosphatase	Rat liver	—	62 500 57 000 36 000	−	362
Polynucleotide phosphorylase	*E. coli*	216 000	2 × 92 000	−	363
Aspartokinase III	*E. coli*	—	2 × 50 000	−	364
Dopamine-β-hydroxylase	Bovine adrenals	290 000	4 × 76 000	+	365
Extracellular haemoglobin	*Lumbricus terrestris*	3 230 000	12 000 14 000 16 000 19 000 31 000 36 000	+	366
Ferritin	Horse	—	19 000	−	367
	Human	—	10 500	−	
	Rat	—	7 500	−	
	Rabbit	—	—	−	

[360] L. Schirch, M. Edmiston, M. S. Chen, D. Barra, F. Bossa, L. Hinds, and P. Fasella, *J. Biol. Chem.*, 1973, **248**, 6456.
[361] J. C. Lee, R. P. Frigon, and S. N. Timasheff, *J. Biol. Chem.*, 1973, **248**, 7253.
[362] W. A. Catterall, W. A. Coty, and P. L. Pedersen, *J. Biol. Chem.*, 1973, **248**, 7427.
[363] C. Portier, R. van Rapenbusch, M. N. Thang, and M. Grunberg-Manago, *European J. Biochem.*, 1973, **40**, 77.
[364] C. Richaud, J. P. Mazat, C. Gros, and J. C. Patte, *European J. Biochem.*, 1973, **40**, 619.
[365] J. E. Craine, G. H. Daniels, and S. Kaufman, *J. Biol. Chem.*, 1973, **248**, 7838.
[366] J. M. Schlom and S. V. Vinogradov, *J. Biol. Chem.*, 1973, **248**, 7904.
[367] R. D. Dyson and J. M. Cardenas, *J. Biol. Chem.*, 1973, **248**, 8482.

Table 2 *Affinity chromatography of proteins*

Protein	*Ligand*	*Matrix*	*Eluant*	*Ref.*
Neutral proteinase	Caproyl-Gly-Leu	Agarose	$CaCl_2$, pH 10	368
Acetylcholinesterase	*m*-Carboxyphenyltrimethyl-ammonium ion	Agarose	—	369
Glucagon-like protein	Anti-glucagon antibody	Agarose	4.25M-Ethanol in 4mM-HCl	370
Glucose-6-phosphate dehydrogenase	NADP	Agarose	—	17
Alkaline protease	Chicken ovoinhibitor	Agarose	0.2M-Acetic acid	371
Lipoamide dehydrogenase	Propyl-lipoamide	Glass beads	0.05M-Phosphate, pH 6.8	16
Phytohaemagglutinin	—	Acid-treated agarose	—	372
Erythrocyte proteins (TNBS-labelled)	Anti-DNP antibody	Agarose	0.1M-Acetic acid and 20% formic acid	19
Dihydrofolate reductase	Methotrexate	Agarose	Dihydrofolate	373
Papain	Glutathione-2-pyridyl disulphide	Agarose	Cysteine	25
Virus glycoproteins	Phytohaemagglutinin	Agarose	α-Methyl mannoside	374
β-Glucuronidase	Saccharo-1,4-lactone, Aminophenyl glucuronide	Agarose	Acetic acid	24
Gentamicin inactivating enzyme	Gentamicin	Agarose	Sodium chloride	375
Cathepsin B1	Aminophenyl mercuric acetate	Agarose	Mercuric chloride	376
Threonine deaminase	Isoleucine	Agarose	Phosphate	377

[368] L. G. Sparrow and A. B. McQuade, *Biochem. Biophys. Acta*, 1973, **302**, 90.
[369] H. I. Yamamura, D. W. Reichard, T. L. Gardner, J. D. Morrisett, and C. A. Broomfield, *Biochim. Biophys. Acta*, 1973, **302**, 305.
[370] R. F. Murphy, K. D. Buchanan, and D. T. Elmere, *Biochim. Biophys. Acta*, 1973, **303**, 118.
[371] G. Feinstein and A. Gertler, *Biochim. Biophys. Acta*, 1973, **309**, 196.
[372] B. Ersson, K. Aspberg, and J. Porath, *Biochim. Biophys. Acta*, 1973, **310**, 446.
[373] J. Gauldie, L. Marshall, and B. L. Hillcoat, *Biochem. J.*, 1973, **133**, 349.
[374] M. J. Hayman, J. J. Skehel, and M. J. Crumpton, *F.E.B.S. Letters*, 1973, **29**, 185.
[375] F. Le Goffic and M. Moreau, *F.E.B.S. Letters*, 1973, **29**, 289.
[376] H. Keilova and V. Tomasek, *F.E.B.S. Letters*, 1973, **29**, 335.
[377] I. Rahini-Laridjani, H. Grimminger, and F. Lingens, *F.E.B.S. Letters*, 1973, **30**, 185.

Table 2 (*cont.*)

Protein	*Ligand*	*Matrix*	*Eluant*	*Ref.*
Ribonuclease	Aminophenyl uridine phosphate	Agarose	Sodium chloride	378
α_1-Antitrypsin	Concanavalin A	Agarose	α-Methyl glucoside	379
Brain glycoproteins	Concanavalin A	Agarose	α-Methyl mannoside	380
α-Fetoprotein	Oestradiol	Agarose	—	381
Phytohaemagglutinin	Concanavalin A	Agarose	α-Methyl mannoside	382
Rennin	Pepstatin	Agarose	6M-Urea	383
Lactate dehydrogenase	Oxamate	Agarose	NADH	384
Folate-binding protein	Folic acid	Agarose	Acetic acid	385
Acetylcholinesterase	1,6-Trimethylammoniumhexyl amine	Agarose	Trimethylammonium hexylamine	386
Lactate dehydrogenase	Aminohexyl AMP	Agarose	NAD	387
Threonine deaminase	Valine, Isoleucine	Agarose	Potassium chloride	
Insulin receptor	Concanavalin A, Wheat-germ agglutinin	Agarose	α-Methyl mannoside, *N*-Acetylglucosamine	23
Interferon	Anti-interferon antibody	Agarose	Acetic acid and sodium chloride	21
Glycopeptide binding protein	Gycopeptides	Agarose		388
Glutathione reductase	Glutathione	Agarose	—	389
DNase II	Anti-DNase antibody	Agarose	Acetic acid	390
Thioredoxin	Anti-thioredoxin antibody	Agarose	Acetic acid	391
β-Lactamase	Ampicillin	Agarose	Sodium chloride	392
α-Fetoprotein	Concanavalin A	Agarose	Glucose	393
Histone-F3	Organomercurial	Agarose	Cysteine	394
Trypsin/thrombin	*p*-Aminophenylguanidine, *p*-Aminobenzamidine, *m*-Aminobenzamidine	Agarose	—	395
Wheat-germ agglutinin	2-Deoxyglucoside	Agarose	—	396
Penicillinase	Cephalosporin C	Agarose	—	397

Dihydroneopterin triphosphate synthetase	Periodate-treated GTP	Agarose	GTP	398
α-Antitrypsin	Concanavalin A	Agarose	α-Methyl glucoside	399
Thiamine-binding protein	Thiamine pyrophosphate	Agarose	Urea	312
tRNA Methylase	$tRNA_{Glu}$	Agarose	Potassium chloride	400
Wheat-germ agglutinin	Acetamido-(aminocaproyl)-deoxyglucosamine	Agarose	Acetic acid	401
Tubulin	Deacetyl-/isodeacetyl-colchicine	Agarose	Sodium chloride	315
Haemoglobin	Haptoglobin	Agarose	Guanidine hydrochloride	402
RNA Polymerase	DNA	Cellulose	Potassium chloride	403

[378] R. K. Wierenga, J. D. Huizinga, W. Gaastea, G. W. Welling, and J. J. Beintema, *F.E.B.S. Letters*, 1973, **31**, 181.
[379] R. J. Murthy and A. Hercz, *F.E.B.S. Letters*, 1973, **32**, 243.
[380] J. P. Susz, H. I. Hof, and E. G. Bunngraber, *F.E.B.S. Letters*, 1973, **32**, 289.
[381] R. Arnon, E. Feicher, M. Bustin, and M. Sela, *F.E.B.S. Letters*, 1973, **32**, 335.
[382] W. Bessler and I. J. Goldstein, *F.E.B.S. Letters*, 1973, **34**, 58.
[383] P. Corval, C. Devaux, and J. Menard, *F.E.B.S. Letters*, 1973, **34**, 189.
[384] H. Spielmann, R. P. Erickson, and C. J. Epstein, *F.E.B.S. Letters*, 1973, **35**, 19.
[385] J. Selhub and N. Grosswicz, *F.E.B.S. Letters*, 1973, **35**, 76.
[386] W. H. Hopff, G. Riggio, and P. G. Waser, *F.E.B.S. Letters*, 1973, **35**, 220.
[387] P. Brodelius and K. Mosbach, *F.E.B.S. Letters*, 1973, **35**, 223.
[388] G. Sepulcre and E. Moczar, *J. Chromatog.*, 1973, **75**, 114.
[389] J. J. Harding, *J. Chromatog.*, 1973, **77**, 191.
[390] K. W. Ryder and M. E. Hodes, *J. Chromatog.*, 1973, **80**, 128.
[391] B. M. Sjoberg and A. Holmgren, *Biochim. Biophys. Acta*, 1973, **315**, 176.
[392] F. Le Goffic, R. Labia, and J. Andrillon, *Biochim. Biophys. Acta*, 1973, **315**, 439.
[393] C. J. Smith and P. J. Kelleher, *Biochim. Biophys. Acta*, 1973, **317**, 231.
[394] A. Ruiz-Carrillo and V. G. Allfrey, *Arch. Biochem. Biophys.*, 1973, **154**, 185.
[395] H. F. Hixson and A. H. Nishihawa, *Arch. Biochem. Biophys.*, 1973, **154**, 501.
[396] J. H. Shazer, R. Barker, and R. L. Hill, *Analyt. Biochem.*, 1973, **53**, 564.
[397] L. J. Crane, G. E. Bettinger, and J. O. Lampen, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 220.
[398] R. J. Jackson, R. M. Wolcott, and T. Shiota, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 428.
[399] I. E. Liener, O. R. Garrison, and Z. Pravda, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 436.
[400] Y. Taya and S. Nishimura, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 1062.
[401] R. Lotan, A. E. S. Gussin, H. Lis, and N. Sharon, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 656.
[402] M. Klein and C. Mihaesco, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 774.
[403] R. Mukai and Y. Lida, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 134.

Table 2 (*cont.*)

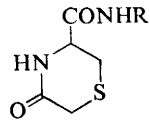
Protein	*Ligand*	*Matrix*	*Eluant*	*Ref.*
Thermolysin	Acetylphenylalanine, Succinylleucine	Agarose	pH 9 or 6M-guanidine hydrochloride	403*a*
Low density lipoproteins	Concanavalin A	Agarose	Glucose, mannose, α-methyl glucoside, α-methyl mannoside	404
Carboxypeptidase A and B	Arginine	Agarose	—	111
Phosphodiesterase	Nitrophenyl phenyl thiophosphate	Agarose	Sodium chloride	405
Aminoacyl-tRNA synthetase	Phosphate	Cellulose	tRNA	406
Acetylcholinesterase	Trimethyl(*p*-aminophenyl)-ammonium chloride	Agarose	Idrophonium chloride	81
Collagenases	Collagen	Agarose	Acetic acid	246
Phosphofructokinase	Cibracion Blue F_3G-A	Sephadex G100	—	270
Protocollagen proline hydroxylase	Collagen	Agarose	$(Pro\text{-}Gly\text{-}Pro)_n$	407
Isoprothrombin	Isoprothrombin antibody	Agarose	Sodium dodecylsulphate	194
Rennin	Pepstatin	Polyacryl-amide (AE)	Salt	408
Serum albumin	Oleyl- and palmitoyl-	Agarose	Alcohol, pH 3.0	409
α-Galactosidase	*p*-Aminophenyl melobioside	Agarose	Triton × 100	410
Vitamin B_{12}-binding proteins	Vitamin B_{12}	Agarose	Guanidine hydrochloride	195
Antihaemophilic factor				94
Glutamate dehydrogenase	*N*-Carboxymethyl glutamate	Agarose	Glutamate	167
Plasminogens	Lysine	Agarose	—	411
β-Galactosidase	Thiogalactoside	Agarose	Tris, pH 7.5	412
Acetylcholine receptor	Cobrotoxin	Agarose	Hexamethinium chloride	413
Leucyl-tRNA synthetase	$tRNA_{leu}$	Agarose	Salt and pH gradient	414
Complement $C1_q$	Anti $C1_q$ antibody	Agarose	1,4-diaminobutane	415
Enterotoxin	Albumin/poly(Lys-Ala) ganglioside	Agarose	Guanidine hydrochloride	416
i Antigen	Anti i antibody	Agarose	Increase in temperature	22
Factor IX	Heparin	Agarose	Salt	417

[403d] M. K. Pangburn, Y. Burstein, P. H. Morgan, K. A. Walsh, and H. Neurath, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 371.
[404] V. G. Shore and B. Shore, *Biochemistry*, 1973, **12**, 502.
[405] A. M. Frischauf and F. Eckstein, *European J. Biochem.*, 1973, **32**, 479.
[406] F. Von der Haar, *European J. Biochem.*, 1973, **34**, 84.
[407] R. A. Berg and D. J. Prockop, *J. Biol. Chem.*, 1973, **248**, 1175.
[408] K. Murakami, T. Inagami, A. M. Michelakis, and S. Cohen, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 482.
[409] T. Peters, H. Taniuchi, and C. B. Anfinsen, *J. Biol. Chem.*, 1973, **248**, 2447.
[410] C. A. Mapes and C. C. Sweeley, *J. Biol. Chem.*, 1973, **248**, 2461.
[411] L. Summaria, L. Arzadon, P. Bernabe, and K. C. Robbins, *J. Biol. Chem.*, 1973, **248**, 6522.
[412] J. J. Distler and G. W. Jourdian, *J. Biol. Chem.*, 1973, **248**, 6772.
[413] R. P. Klett, B. W. Fulpins, D. Cooper, M. Smith, E. Reich, and L. P. Bossani, *J. Biol. Chem.*, 1973, **248**, 6841.
[414] H. Hayashi, *J. Biochem.*, 1973, **74**, 203.
[415] C. R. Sledge and D. H. Bing, *J. Immunol.*, 1973, **111**, 661.
[416] V. Sica, E. Nola, I. Parikh, G. A. Puca, and P. Cuatracasas, *Nature New Biol.*, 1973, **244**, 36.
[417] G. Biesecker, *Biochemistry*, 1973, **12**, 4403.

the result of the failure of the native protein to bind SDS in sufficient amounts to give it a charge to mass ratio similar to that of standard proteins. It should be noted that the anomaly is not evident if the protein is denatured with guanidine hydrochloride. This illustrates the need to attempt confirmation of SDS-gel data with an independent method.

3 Chemical Modification

Methods.—*Thiol Groups.* It is common practice in protein chemistry to block thiol groups with iodoacetate, thus preventing formation of disulphide bridges. Bradbury and Smyth now report[418] that under acid conditions the *S*-carboxymethylcysteine derivatives formed may undergo intramolecular cyclization to the lactam (1). This compound possesses no

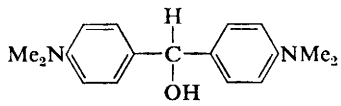


(1)

free amino-group and would therefore block attempts at peptide sequencing or at ninhydrin quantitation of sulphydryl groups on the amino-acid analyser. The reaction of thiols with 3-bromopropionic acid gives the carboxyethyl derivative, and since this does not undergo cyclization the reagent is recommended as an alternative to iodoacetate. Presumably, acrylonitrile, which also reacts with cysteine to give a carboxyethyl derivative, would be equally effective.

For the reduction prior to alkylation, dihydrolipoic acid bound to an inert support has been used.[419] In this scheme lipoic acid is bound to aminoethyl-Sephadex *via* its succinimide ester. The substituted polymer is then reduced with sodium borohydride to give the dihydrolipoate derivative. Shaking with the polymer than gives quantitative reduction of disulphides and avoids an excess of thiol reducing agent in solution.

For the quantitative determination of sulphydryl groups, a new reagent, 4,4′-bis(dimethylaminophenyl)carbinol (2), has been proposed.[420] The reaction can be followed by monitoring the loss in absorbance at 612 nm but urea interferes, possibly as a result of cyanate contamination.



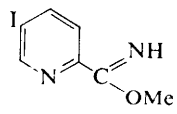
(2)

[418] A. F. Bradbury and D. G. Smyth, *Biochem. J.*, 1973, **131**, 635.
[419] M. Gorecki and A. Patchornik, *Biochim. Biophys. Acta*, 1973, **303**, 36.
[420] M. S. Rohrbach, B. A. Humphries, F. J. Yost, W. G. Rhodes, S. Boatman, R. G. Hiskey, and J. H. Harrison, *Analyt. Biochem.*, 1973, **52**, 127.

Methionine. The reaction of this amino-acid with an equimolar amount of $HAuCl_4$ is claimed to give quantitative oxidation to the sulphoxide.[421] The reaction is performed in water and may be of value as a means of reversible modification of this amino-acid.

Tyrosine. Iodoacetamide, a reagent normally used for thiol-group modification, has been found[422] to react with tyrosine in both glucagon and riboflavin-binding protein. The reaction is pH-dependent so that at pH 8.5 half of the available tyrosines are alkylated whereas at pH 10.0 the protein is completely derivatized. A study has been made of the reaction between tyrosine-containing peptides and α-nitroso-β-naphthol.[423] These authors note that the condensation products are considerably retarded on Sephadex and suggest that this might be useful as a selective method of purification for these peptides.

Amino-groups. Imido-esters have for some time been used for the specific modification of amino-groups in proteins (see, for instance, the structure determination of carbonic anhydrase, 'Complete Sequences' section). Riley and Perham have now introduced[424] an iodinated imido-ester, 5-iodopyridine-2-carboximidate (3), which has the added advantage of



(3)

introducing a heavy atom into the protein. This is of value to crystallographers, where the preparation of isomorphous heavy-atom derivatives is a necessary adjunct to solution of the protein tertiary structure. One problem is that a good isomorphous derivative should satisfy the criterion of being bound with high occupancy at only one or two sites in the protein, and a reagent that reacts with amino-groups is unlikely to satisfy this requirement. These authors suggest that this difficulty may be avoided by nitration of specific tyrosine residues with tetranitromethane and reduction of the introduced nitro-groups with dithionite to form amino-tyrosine derivatives. Since the pK_a of this aromatic amino-group is about 5.0, reaction with the imido-ester at this pH should be relatively specific.

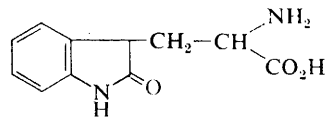
Carboxy-groups. In an attempt to gain specificity for modification of α, β, or γ carboxy-groups in proteins the effect of various substituted diboranes has been investigated.[425] On reaction with lysozyme or myoglobin,

[421] E. Bordignon, L. Cattalini, G. Natile, and A. Scatturin, *J.C.S. Chem. Comm.*, 1973, **22**, 878.
[422] R. C. Cotner and C. O. Clagett, *Analyt. Biochem.*, 1973, **54**, 170.
[423] R. Håkanson, A. L. Rönnberg, and K. Sjölund, *Analyt. Biochem.*, 1973, **51**, 523.
[424] M. Riley and R. N. Perham, *Biochem. J.*, 1973, **131**, 625.
[425] M. Z. Atassi, A. F. Rosenthal, and L. Vargas, *Biochim. Biophys. Acta*, 1973, **303**, 379.

disiamylborane specifically reduced accessible glutamic acid side-chains and unhindered *C*-terminal carboxy-groups. 9-Borabicyclo-[3,3,1]nonane, however, also reacted to some extent with the β-carboxy-groups of aspartic acid. As well as chemical modification these reagents may find some use as conformational probes.

An alternative modification is the condensation of the carboxy-group with ammonia using 1-ethyl-3-dimethylaminopropylcarbodi-imide as condensing agent.[426] This results in the formation of asparagine and glutamine residues. Side-reactions of carbodi-imides include cross-linking and reactions with tyrosine and serine residues, so that the results of this modification would have to be interpreted with caution.

Iodination. Although iodination is generally used for modification of tyrosine residues, it may also react with tryptophan, methionine, histidine, cysteine, and cystine. In a study of the reaction of various iodinating species with small tryptophan-containing peptides,[427] oxidation to oxindole was observed in the pH range 4.0—7.5. In the pH range 4.0—5.0 significant cleavage of tryptophanyl bonds was observed, and these authors stress the need for caution in interpreting the results of iodination. A separate publication [428] discusses the effect on the c.d. spectra of this modification and concludes that the major species produced is β-3-oxindolyl-L-alanine (4), although other related compounds may be formed.



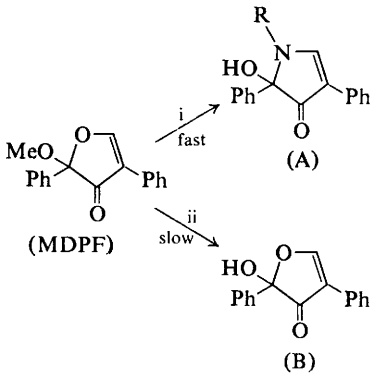
(4)

An alternative and possibly less destructive approach is to iodinate an acylating reagent and then allow this to react with the protein. For instance, various peptide hormones have been labelled to high specific activity using radio-iodinated 3-(4-hydroxyphenyl)propionic acid *N*-hydroxysuccinimide ester.[429] The method is particularly useful for the development of radioimmunoassay methods, where direct iodination may decrease the affinity of the labelled antigen for antibody, giving a corresponding decrease in sensitivity of the assay.

Fluorescent Labelling. Hoffmann La Roche have introduced a new compound, 2-methoxy-2,4-diphenyl-3(2*H*)-furanone (MDPF), for the fluorescent labelling of proteins.[430] As is the case with fluorescamine, neither the

[426] S. D. Lewis and J. A. Schafer, *Biochim. Biophys. Acta*, 1973, **303**, 292.
[427] N. M. Alexander, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 614.
[428] E. H. Stickland, M. Wilchek, and C. Billups, *Biochim. Biophys. Acta*, 1973, **303**, 28.
[429] A. E. Bolton and W. M. Hunter, *Biochem. J.*, 1973, **133**, 529.
[430] M. Weigele, S. De Bernardo, W. Leimgruber, R. Cleeland, and E. Grunberg, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 899.

compound nor its hydrolysis product (B, in Scheme 3) is fluorescent, so that purification of the resulting conjugates is unnecessary. The advantage over the former compound, however, is that the fluorescent products are stable over a wider pH range. The authors suggest that MDPF will be very useful in immunofluorescence studies, but as no preparative method is given for the reagent it will be difficult to examine this prediction.



Reagents: i, RNH_2; ii, H_2O

Reaction of MDPF with primary amines to form fluorescent 3,5-*diphenyl*-5-*hydroxy*-2-*pyrrolin*-4-*ones* (A)

Scheme 3

Cross-linking Reagents. The synthesis of a cleavable cross-linking compound, methyl 4-mercaptobutyrimidate, has been described.[431] This bifunctional ester is used in a three-step process. First, protein amino-groups are allowed to react with the imidate function. Second, mild oxidation gives S—S linked dimers or higher oligomers. Third, after resolution of the cross-linked products, *i.e.* by SDS gels, reductive cleavage restores the original protein monomers, which may then be identified. The two advantages claimed for this reagent over other cleavable cross-links are that the cleaved proteins have their original charge and that the 14.6 Å N—N separation of the compound is sufficient for effective cross-linking. An application in the study of ribosome topography is discussed below.

General Applications.—It is not possible to discuss all the conventional applications of chemical modification methods here, and this information has been summarized in Table 3.

Affinity Labelling.—*Immunoglobulins.* Affinity labelling with Dns-lysine has allowed the distance between the hapten cross-linking site and the single S—S bond linking the heavy chains in an antidansyl antibody to be

[431] R. R. Traut, A. Bollen, T.-T. Sun, J. W. B. Hershey, J. Sundberg, and L. R. Pierce, *Biochemistry*, 1973, **12**, 3266.

Table 3 *Chemical modification of peptides and proteins*

Protein	*Source*	*Label*	*Residue modified*	*Comments*	*Ref.*
Acetylcholinesterase	Eel	Pinaconyl methylphosphono-fluoridate	Ser	Peptide isolated	432
Acetylcholinesterase	Eel	Photochromic ligand	—	Hydrophobic probe	433
Alcohol dehydrogenase	Horse liver	$HCHOBH_4^-$	Lys		434
Alcohol dehydrogenase	Rat liver	Iodoacetate	Cys		435
Alcohol dehydrogenase	Yeast	Butyl isocyanate	Cys	Active-site label	436
Alcohol dehydrogenase	Yeast	Iodoacetamide	Cys		437
Alcohol dehydrogenase	Yeast	Coenzyme analogue	His		602
Aldolase	Rabbit muscle	Chloroacetol phosphate	Cys	Substrate analogue	438
α-Amylase	*B. subtilis*	*p*-nitrophenyl acetate	Lys		439
α-Amylase	*B. subtilis*	—	—		440
A protein	*S. aureus*	Iodoacetate	Cys		441
A protein	*S. aureus*	Tetranitromethane (TNM)	Tyr		442
Apyrase	Potato	I_2, TNM, *N*-acetylimidazole	—		443
Arginase	Rat kidney	—			444
Asparaginase	*E. coli*	Photo-oxidation	His		445
Aspartate aminotransferase	Pig heart	Iodoacetate	Cys	Cys-45, 82, 252	446
Aspartate aminotransferase	Pig heart	—	Cys, Tyr		447
Aspartate aminotransferase	Pig heart	TNM	Cys, Tyr	Cys-390	448
Aspartate aminotransferase	Pig heart	Bromopyruvate	Cys	Affinity label	590
Aspartate aminotransferase	Pig heart	β-Chloroalanine	Lys	Affinity label	589
Aspartic β-semialdehyde dehydrogenase	Yeast	Thiol reagents	Cys		449
ATPase	Beef heart	Various	— Cys, Tyr		450
Benzylamine oxidase	Pig plasma	Hydrazine derivatives	—	Reacts with cofactor	451
Carbonic anhydrase	Human and bovine	Affinity labels	His	Peptides isolated	452
Carbonic anhydrase	Bovine	Phenyl-*N*-methylacetimidate	Lys		453
Carbonic anhydrase	Human and bovine	2-Hydroxy-5-nitrobromo-benzene	Trp		454
Carboxylesterase	Chicken/ox liver	Bromoacetophenone	His	Active-site label	455

Carboxypeptidase A	—	Butanedione	Arg		456
Carboxypeptidase B	—	Cadmium	—	Replaces zinc	457
Carboxypeptidase Y	Yeast	DFP	Ser		458
Catechol-*O*-methyl transferase	Rat liver	Affinity label	—	Substrate protects	459
Chymotrypsin	—	Affinity label	Met	Met-192	460
Chymotrypsin	Bovine	Dinitrophenylbromobutanone	His		461

[432] N. K. Shafter, H. O. Michel, and A. F. Bridges, *Biochemistry*, 1973, **12**, 2946.
[433] K. T. Galley, M. De Sorgo, and W. Prins, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 300.
[434] H. Jörnvall, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 1069.
[435] H. Jörnvall, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 1227.
[436] J.-S. Twu and F. Wold, *Biochemistry*, 1973, **12**, 381.
[437] J.-S. Twu, C. C. Q. Chin, and F. Wold, *Biochemistry*, 1973, **12**, 2856.
[438] D. W. Salter, I. L. Norton, and F. C. Hartman, *Biochemistry*, 1973, **12**, 1.
[439] I. Wabe, H. Narigo, and H. Ohada, *Biochim. Biophys. Acta*, 1973, **302**, 73.
[440] J. Hora, *Biochim. Biophys. Acta*, 1973, **310**, 264.
[441] I. Sjöholm and M. Ljimgdahl, *Immunochemistry*, 1973, **10**, 287.
[442] I. Sjöholm, A. Bjerkén, and J. Sjöquist, *J. Immunol.*, 1973, **110**, 1562.
[443] M. A. Valenzuela, G. D. Campo, E. Marin, and A. Traverso-Cori, *Biochem. J.*, 1973, **133**, 755.
[444] G. A. Kaysen and H. J. Strecker, *Biochem. J.*, 1973, **133**, 779.
[445] H. Makino and Y. Inada, *Biochim. Biophys. Acta*, 1973, **295**, 543.
[446] R. A. Zufarova, M. M. Dedyukina, L. V. Memelova, and Y. M. Torchinsky, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 127.
[447] A. L. Bocharov, T. V. Demidkina, M. Y. Karpeisky, and O. L. Polyanovsky, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 377.
[448] W. Birchmeier, P. E. Zaovalek, and P. Christen, *Biochemistry*, 1973, **12**, 2874.
[449] M. J. Holland and E. W. Westhead, *Biochemistry*, 1973, **12**, 2276.
[450] A. E. Senior, *Biochemistry*, 1973, **12**, 3622.
[451] A. Lindström and G. Pettersson, *European J. Biochem.*, 1973, **34**, 564.
[452] D. L. Cybulsky, S. I. Kandel, M. Kandel, and A. G. Gornall, *J. Biol. Chem.*, 1973, **248**, 3411.
[453] Y. Pocker, M. W. Beug, and F. A. Beug, *Biochemistry*, 1973, **12**, 2483.
[454] S. Lindskog and A. Nilsson, *Biochim. Biophys. Acta*, 1973, **295**, 177.
[455] P. Willadsen, J. de Jersey, and B. Zerner, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 620.
[456] J. F. Riordan, *J. Biol. Chem.*, 1973, **242**, 3915.
[457] N. Zisapel and M. Sokolovsky, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 722.
[458] R. Hayashi, S. Moore, and W. H. Stein, *Biochemistry*, 1973, **248**, 8366.
[459] R. T. Borchardt and D. Thakker, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 1233.
[460] R. Alazaid, J. J. Béchet, A. Dupaix, and J. Yon, *Biochim. Biophys. Acta*, 1973, **309**, 379.
[461] S. M. Grey, G. R. Nurse, and L. Visser, *Biochim. Biophys. Acta*, 1973, **52**, 687.

Table 3 (*cont.*)

Protein	*Source*	*Label*	*Residue modified*	*Comments*	*Ref.*
Chymotrypsin	—	Butyl and octyl isocyanates	Ser	Ser-195	462, 463
Chymotrypsin	—	Affinity label	His	His-57	464
Chymotrypsin	—	Photoaffinity label	His		465
Chymotrypsin	Bovine	Carbodi-imide	—	Modifies carboxy-groups	466
Chymotrypsinogen	Bovine	Carbodi-imide	—	Modifies carboxy-groups	467
Chymotrypsinogen	Bovine	Dimethyl 2-hydroxy-5-nitrophenyl sulphonium bromide	Trp	Trp-215	468
Chymotrypsinogen	Bovine	Carbodi-imide	—	Modifies carboxy-groups	466
Chymotrypsinogen	—	Glutaraldehyde	—	Cross-link	469
Cobratoxin	Taiwan cobra	Various	Trp		470
Concanavalin A	—	Succinic anhydride, acetic anhydride	Lys		471
Creatine kinase	Rabbit muscle	2-Hydroxy-5-nitrobromobenzene 2-Methoxy-5-nitrobromobenzene	Cys		472
Creatine kinase	Rabbit muscle	Dansyl aminoethyl phosphate	—	Interacts at nucleotide site	472
Cytochrome *c*	Horse heart	Bromine adduct of NPS-skatole	Trp	Bond cleavage	473
Deoxyribonuclease	Porcine spleen	Iodoacetate	His	Reaction at pH 4.6	474
Dihydrofolate reductase	Sarcoma 180 cells	Active-site labels	—	Hydrophobicity probes	587
Elastase	—	Isocyanates	Ser	Active-site label	462, 463
Enolase	Rabbit muscle	Glycidol phosphate	Glu	Active-site label	586
Erythrocyte membranes	Human	Iodine	—	*via* lactoperoxidase	475
Erythrocyte membranes	Human	Iodine	—		476
Erythrocyte membranes	Human	Dansyl chloride	—		477
Erythrocyte membranes	Human	Fluorodinitrobenzene	—	Cross-link to Ⓟ-lipid	478

Erythrocyte membranes	Human	Dimethyl suberimidate	—	Cross-link to Ⓟ-lipid	478
Erythrocyte membranes	Human	Dimethyl adipimidate	—	Cross-links glycoproteins	479
Fatty acid synthetase	Pigeon/rat liver	Phenylmethanesulphonyl fluoride	—		480
Fibrinogen	Bovine	Various	Lys, Tyr		481, 482
Fibrinoligase	Human plasma	Iodoacetamide	Cys		483
Ficin	—	Bromoacetone	His		484
Glucagon	—	Various	His		485

[462] W. E. Brown and F. Wold, *Biochemistry*, 1973, **12**, 828.
[463] W. E. Brown and F. Wold, *Biochemistry*, 1973, **12**, 835.
[464] J.-J. Béchet, A. Dupaix, J. Yon, M. Wakselman, J.-C. Robert, and M. Vilkas, *European J. Biochem.*, 1973, **35**, 527.
[465] H. Nakayama and Y. Kanaoka, *F.E.B.S. Letters*, 1973, **37**, 200.
[466] A. R. Fersht and J. Sperling, *J. Mol. Biol.*, 1973, **74**, 137.
[467] J. C. Wriston, *F.E.B.S. Letters*, 1973, **33**, 93.
[468] G. S. Katrukha, N. Zylber, and V. Keil-Dlouhá, *F.E.B.S. Letters*, 1973, **29**, 25.
[469] W. Gaffield, Y. Tomimatsu, A. C. Olsen, and E. F. Jensen, *Arch. Biochem. Biophys.*, 1973, **157**, 405.
[470] C. C. Chang and C. C. Yang, *Biochim. Biophys. Acta*, 1973, **295**, 595.
[471] G. R. Gunther, J. L. Wang, I. Yahara, B. A. Cunningham, and G. E. Edelman, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1012.
[472] C. Roustan, A. Brevet, and L.-A. Pradel, *European J. Biochem.*, 1973, **39**, 371.
[473] A. Fontana, C. Vita, and C. Tonislo, *F.E.B.S. Letters*, 1973, **32**, 139.
[474] R. G. Oshima and P. A. Price, *J. Biol. Chem.*, 1973, **248**, 7522.
[475] E. Reichstein and R. Blostein, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 494.
[476] J. P. Segrest, I. Kahane, R. L. Jackson, and V. T. Marchesi, *Arch. Biochem. Biophys.*, 1973, **155**, 167.
[477] R. Schmidt-Ullrich, H. K. Knufermann, and D. F. H. Wallach, *Biochim. Biophys. Acta*, 1973, **307**, 353.
[478] G. V. Marinetti, R. Baumgarten, D. Sheeley, and S. Gordesky, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 302.
[479] T. H. Ji, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 508.
[480] S. Kumar, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 334.
[481] H. M. Phillips and J. L. York, *Biochemistry*, 1973, **12**, 3637.
[482] H. M. Phillips and J. L. York, *Biochemistry*, 1973, **12**, 3642.
[483] C. G. Curtis, P. Stenberg, C.-H. J. Chou, A. Gray, K. L. Brown, and L. Lorand, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 51.
[484] J. M. Gleisner and T. E. Liener, *Biochim. Biophys. Acta*, 1973, **317**, 482.
[485] R. M. Epand, R. F. Epand, and V. Grey, *Arch. Biochem. Biophys.*, 1973, **154**, 132.

Table 3 (*cont.*)

Protein	*Source*	*Label*	*Residue modified*	*Comments*	*Ref.*
Glucagon	—	Iodine	—	$ICl/Na^{125}I$	486
Glucagon	—	Iodoacetamide	Tyr		422
Glucose-6-phosphatase	Rat liver	Thiol reagents	Cys		487
Glutamate dehydrogenase	Bovine liver	Glycine + carbodi-imide	Asp, Glu		488
Glutamate dehydrogenase	Bovine liver	4-Iodoacetamidosalicylate	Cys, Met		489
Glutamate dehydrogenase	Bovine/porcine/ chicken liver	4-Iodoacetamidosalicylate	Lys	Lys-126	490
Glutamate dehydrogenase	Bovine liver	Diethylpyrocarbonate	His	Activates	491
Glutamate dehydrogenase	—	Photo-oxidation	His		492
Glutamate dehydrogenase	—	Glyoxal/tetrazole	Lys		493
Glutamate dehydrogenase	—	Azidoisophthalic acid	—	Photoaffinity labels	494
Glutamine synthetase	Sheep brain	Dimethyl sulphone	—	Reversible	495
Glutamine synthetase	*E. coli*	$1,N^6$-ethenoadenosine 5′-triphosphate	Tyr	Fluorescent conformational probe	496
Glyceraldehyde-3-phosphate dehydrogenase	Rabbit muscle	Benzylamine	Cys		497
Glyceraldehyde-3-phosphate dehydrogenase	Rabbit muscle	Ethacrynic acid	Cys		498
Glyceraldehyde-3-phosphate dehydrogenase	Rabbit muscle	Glyceryl trinitate	Cys		498
Glyceraldehyde-3-phosphate dehydrogenase	Rabbit muscle	Photo-oxidation	His	His-38	499, 500
Glyceraldehyde-3-phosphate dehydrogenase	Rabbit muscle	Pyridoxal phosphate	Lys	Lys-212	501
Glyceraldehyde-3-phosphate dehydrogenase	Rabbit muscle	NAD analogue	Cys		603
Glyceraldehyde-3-phosphate dehydrogenase	Pig muscle	ATP	Cys		502
Glyceraldehyde-3-phosphate dehydrogenase	Rabbit muscle	Phenylhydrazine	'Cys'	Reacts with oxidized form	503
Glyceraldehyde-3-phosphate dehydrogenase	Rabbit muscle	Phenyldi-imide	Cys		504

Glyceraldehyde-3-phosphate dehydrogenase	Rabbit muscle	—	Cys		
Glycogen phosphorylase	Rabbit muscle	Aliphatic aldehydes	—		505
Glycogen phosphorylase	Rabbit muscle	AMP analogues	Tyr	Peptide isolated	506
Growth hormone	Bovine	Tetranitromethane	Tyr		507
Growth hormone	Ovine	Tetranitromethane	Tyr		507
Haemocyanin	*Helix pomatia*	Diethyl pyrocarbonate	His		508
Haemoglobin	Human	Aspirin	—	HbS	509
Haemoglobin	Human	Pyridoxal derivatives	—	React at *N*-termini	510
Haemoglobin	Human	Azo dyes	Cys, Lys	Cross-links	511
Haemoglobin	Human	Cross-link	α-NH_2	HbA, HbS mixture	601

[486] W. W. Bromer, M. E. Boucher, and J. M. Patterson, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 134.
[487] W. Colilla and R. C. Noidlié, *Biochim. Biophys. Acta*, 1973, **309**, 328.
[488] H. Suaisgood and M. Natake, *J. Biochem.*, 1973, **74**, 77.
[489] N. L. Rosen, L. Bishop, J. B. Burnett, M. Bishop, and R. F. Colman, *J. Biol. Chem.*, 1973, **248**, 7359.
[490] J. J. Holbrook, P. A. Roberts, and R. B. Wallis, *Biochem. J.*, 1973, **133**, 165.
[491] R. B. Wallis and J. J. Holbrook, *Biochem. J.*, 1973, **133**, 183.
[492] F. Hucho, U. Markau, and H. Sund, *European J. Biochem.*, 1973, **32**, 69.
[493] W. Deppert, F. Hucho, and H. Sund, *European J. Biochem.*, 1973, **32**, 76.
[494] W. E. White and K. L. Yielding, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 1129.
[495] W. B. Rowe and A. Meister, *Biochemistry*, 1973, **12**, 1578.
[496] P. B. Chock, C. Y. Huang, R. B. Timmans, and E. R. Stadtman, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3134.
[497] W. S. Allison, L. V. Benitez, and C. L. Johnson, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 1403.
[498] S. B. Jackschick and P. Needleman, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 539.
[499] S. H. Francis, B. P. Meriweather, and J. H. Park, *Biochemistry*, 1973, **12**, 346.
[500] S. H. Francis, B. P. Meriweather, and J. H. Park, *Biochemistry*, 1973, **12**, 355.
[501] M. C. Zapponi, G. Fevri, B. G. Forcina, and S. Ronchi, *F.E.B.S. Letters*, 1973, **31**, 286.
[502] J. Ovadi, M. Nuridsany, and T. Keleti, *Biochim. Biophys. Acta*, 1973, **302**, 191.
[503] W. S. Allison, L. C. Swain, S. M. Tracy, and L. V. Benitez, *Arch. Biochem. Biophys.*, 1973, **155**, 400.
[504] W. S. Allison, L. C. Swain, S. M. Tracy, and L. V. Benitez, *Arch. Biochem. Biophys.*, 1973, **155**, 405.
[505] H. C. Ho, E. Wirch, and J. H.-C. Wang, *Biochim. Biophys. Acta*, 1973, **317**, 462.
[506] M. A. Shatsky, H. C. Ho, and J. H.-C. Wang, *Biochim. Biophys. Acta*, 1973, **303**, 298.
[507] C. B. Glaser, T. A. Bewley, and C. H. Li, *Biochemistry*, 1973, **12**, 3379.
[508] Y. Engelborghs and R. Lontie, *European J. Biochem.*, 1973, **39**, 335.
[509] I. M. Klotz and J. W. O. Tam, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1313.
[510] R. E. Benesch, S. Yung, T. Suzuki, C. Bauer, and R. Benesch, *Proc. Nat. Acad. Sci., U.S.A.*, 1973, **70**, 2595.
[511] H. Fasold, C. Meyer, and G. Steinkopff, *European J. Biochem.*, 1973, **32**, 63.

Table 3 (*cont.*)

Protein	*Source*	*Label*	*Residue modified*	*Comments*	*Ref.*
Hexokinase	Yeast	Diethyl pyrocarbonate	His	Not essential	512
Histone f_1	Calf thymus	Maleic anhydride	Lys		513
Immunoglobulins	—	Maleic anhydride	—		514
Immunoglobulins	Rabbit	Dansyl-lysine	—	Affinity label	579
Immunoglobulins	Turtle and duck	*m*-Nitrobenzenediazonium fluoroborate	—	Affinity label	515
Immunoglobulins	Goat	Dinitrophenyl bromoacetate	—	Affinity label	516
Immunoglobulins	Mouse	DNP-Ala-CHN_2	Lys	Photoaffinity label	591, 592
Immunoglobulin	Mouse	2,4-Dinitrophenylazide	—	Photoaffinity label	591, 592
Immunoglobulin	Mouse	*m*-Nitrobenzenediazonium fluoroborate and difluoro-dinitrobenzene	—	Affinity label	517
Inorganic pyrophosphatase	Yeast	Iodoacetamide	Met	Reaction at pH 5.5	518
Inorganic pyrophosphatase	Yeast	Various	—		519
Insulin	Bovine	Iodoacetate	His	Reaction at pH 5.6	520
Insulin	Bovine	Ethyl thiotrifluoracetate	NH_2	For n.m.r. spectroscopy	521
Insulin	Bovine	Triethyloxonium tetrafluoroborate	Asp, Glu	For n.m.r. spectroscopy	522
Insulin	Bovine	*p*-Nitrophenyl *p*-chloromercuribenzoate	NH_2	For n.m.r. spectroscopy	523
Isocitrate dehydrogenase	Pig heart	Carbodi-imide	Glu	Substrate protects	524
α-Lactalbumin	Bovine milk	Dithioerythritol	Cys		525
α-Lactalbumin	Bovine milk	Maleic anhydride	Lys		526
α-Lactalbumin	Bovine milk	Sodium trinitrobenzene sulphonate	Lys		526
Lactate dehydrogenase	Human	Acetic anhydride	—		527
Lactate dehydrogenase	Pig heart	Diethyl pyrocarbonate	His		528

Lactogenic hormone	Ovine	*o*-Nitrophenylsulphenyl chloride	Trp	Trp-149, Trp-90	529
Lactoferrin	Human	Tetranitromethane	Tyr, Trp		530
Lactoperoxidase	—	3-Amino-1,2,4-triazole	His, Tyr		531
Lactoperoxidase	Skim milk	Phenylhydrazine γ-phenyldi-imide	—		504
Lysozyme	—	Carbodi-imide	—	Modifies carboxy-groups	532
Lysozyme	Hen	Citraconic anhydride	Lys		533
Lysozyme	Hen/turkey	Dimethyl 2-hydroxy-5-nitro-benzene-sulphonium bromide	—	Substrate protects	534*a*
Lysozyme	Hen	Affinity label	Asp	Asp-52	588

[512] M. Grouselle, A. A. Thiam, and J. Pudles, *European J. Biochem.*, 1973, **39**, 431.
[513] J. Burnotte, B. D. Stollar, and G. D. Fasman, *Arch. Biochem. Biophys.*, 1973, **155**, 428.
[514] G. L. Mayers, A. L. Grossberg, and D. Pressman, *Immunochemistry*, 1973, **10**, 37.
[515] G. W. Litman, S. L. Chartrand, C. L. Finstad, and R. A. Good, *Immunochemistry*, 1973, **10**, 323.
[516] Y. Weinstein, S. Shaltiel, D. Givol, and P. H. Strausbauch, *Immunochemistry*, 1973, **10**, 337.
[517] M. Hadler and H. Metzger, *Immunochemistry*, 1973, **10**, 455.
[518] Y. Yano, T. Negi, and M. Irie, *J. Biochem.*, 1973, **74**, 67.
[519] B. S. Cooperman and N. Y. Chiu, *Biochemistry*, 1973, **12**, 1676.
[520] I. Covelli, L. Frati, and J. Wolff, *Biochemistry*, 1973, **12**, 1043.
[521] D. Levy and R. A. Paselk, *Biochim. Biophys. Acta*, 1973, **310**, 398.
[522] D. Levy, *Biochim. Biophys. Acta*, 1973, **310**, 406.
[523] D. Levy, *Biochim. Biophys. Acta*, 1973, **317**, 473.
[524] R. F. Colman, *J. Biol. Chem.*, 1973, **248**, 8137.
[525] Y. Schechter, A. Patchornik, and Y. Burstein, *Biochemistry*, 1973, **12**, 3407.
[526] B. Kitcher and T. E. Banran, *Biochim. Biophys. Acta*, 1973, **298**, 861.
[527] J. F. Burd and M. Vsategui-Gomez, *Biochim. Biophys. Acta*, 1973, **310**, 238.
[528] J. J. Holbrook and V. A. Ingram, *Biochem. J.*, 1973, **131**, 729.
[529] H. Kawauchi, T. A. Beuley, and C. H. Li, *Biochemistry*, 1973, **12**, 2124.
[530] B. Teuwissen, P. L. Masson, P. Osinski, and J. F. Heremans, *European J. Biochem.*, 1973, **35**, 366.
[531] J. Y. Chang and W. A. Schroeder, *Arch. Biochem. Biophys.*, 1973, **156**, 475.
[532] K. J. Kramer and J. A. Rupley, *Arch. Biochem. Biophys.*, 1973, **156**, 414.
[533] M. Z. Atassi, A. F. S. A. Habeeb, and K. Ando, *Biochim. Biophys. Acta*, 1973, **303**, 203.
[534] (*a*) D. K. Jorkasky, S. E. Pearson, and C. L. Borders, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 987; (*b*) B. A. Humphries, M. S. Rohrbach, and J. H. Harrison, *ibid.*, 1973, **50**, 493.
[535] P. H. Fishman, P. G. Pentchev, and J. M. Bailey, *Biochemistry*, 1973, **12**, 2490.

Table 3 (*cont.*)

Protein	*Source*	*Label*	*Residue modified*	*Comments*	*Ref.*
Malate dehydrogenase	Pig	Di(*p*-dimethylaminophenyl)-carbinol	Cys	NAD^+ protects	534*b*
Monoamine oxidase	Rat liver	Glyceryl trinitrate	Cys		498
Monoamine oxidase	Rat liver	Ethacrynic acid	Cys		498
Mutarotase	Bovine kidney				535
Myosin	—	*N*-Ethylmorpholine	Cys	Also S-1 fragment	536
Myosin	Rabbit muscle	*p*-Nitrothiophenol	Cys	Forms S—S bond	537
Myosin	Rabbit muscle	*p*-Chloromercuribenzoate	Cys		538
Nerve growth factor	Mouse	Tetranitromethane	Tyr	Not essential	539
Nerve growth factor	Mouse	*N*-Bromosuccinimide	Trp	Trp-21, Trp-76, Trp-99	539
Octopine dehydrogenase	*Pecten maximus*	Photo-oxidation	His	Substrate protects	540
Oestradiol dehydrogenase	Human placenta	Iodoacetoxyoestrone	His	Affinity label	541, 542
Ornithine δ-aminotransferase	Rat liver	Various	—		543
Papain	*Papaya* latex	2,2′-Dipyridyl disulphide	Cys		544
Papain	*Papaya* latex	Phenyldi-imide	Cys		504
Papain	*Papaya* latex	Fluorescent probes	Cys		545
Parvalbumin	Pike	Cyclohexane-1,2-dione	Arg		546
Penicillinase	*B. cereus*	Affinity label	—		547
Pepsin	Pig	Carbodi-imide	—	Modifies carboxy-groups	548*a*
Pepsin	Pig	Iodine	Lys	Peptides isolated	548*b*
Phosphofructokinase	Rabbit muscle	ATP analogue	Cys	Reversed by thiols	604
6-Phosphogluconate dehydrogenase	Sheep liver	DTNB, *p*-Hydroxymercuri-benzoate	Cys	Substrate/NADPH protect	549
Phosphoglucose isomerase	Yeast	Affinity label	Glu		586
Phosphoglycerate kinase	Yeast	Carbodi-imide	Glu		550
Phosphoglycerate kinase	Yeast	Diethyl pyrocarbonate	His		550

Phospholipase A_2	*Crotalus adamanteus*	Various	Trp, Lys		551
Phosphorylase *b*	Rabbit muscle	Fluorodinitrobenzene	Cys		552
Plasminogen	Human	Active site label	His		553
Platelet membrane glycoprotein	Human	Iodine	—	Major glycoprotein labelled	554
Platelet membrane glycoprotein	Human	DTNB, *p*-CMB	Cys		555
Platelet membrane glycoprotein	Rat	Iodine	—		
Pyrocatechase	*Brevibacterium fuseum*	Mercuric chloride, *N*-bromosuccinimide	Trp	Affects Fe^{3+} binding	556
Pyruvate carboxylase	Sheep kidney	Sodium trinitrobenzene-sulphonate	Lys		557

[536] J. L. Daniel and D. J. Hartshorne, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 125.
[537] R. G. Wolcott and P. D. Boyer, *Biochim. Biophys. Acta*, 1973, **303**, 292.
[538] R. Heazlitt, G. Conway, and J. Montag, *Biochim. Biophys. Acta*, 1973, **317**, 316.
[539] W. A. Frazier, R. A. Hogue-Angeletti, R. Sherman, and R. A. Bradshaw, *Biochemistry*, 1973, **12**, 3281.
[540] F. Thomé-Blau and A. Olomucki, *European J. Biochem.*, 1973, **39**, 557.
[541] M. Pons, J. C. Nicolas, A. M. Boussieux, B. Descomps, and A. C. de Paulet, *F.E.B.S. Letters*, 1973, **31**, 256.
[542] M. Pons, J. C. Nicolas, A. M. Boussieux, B. Descomps, and A. C. de Paulet, *F.E.B.S. Letters*, 1973, **36**, 23.
[543] T. Matsuzaua and M. Nishiyama, *J. Biochem.*, 1973, **73**, 481.
[544] K. Brocklehurst and G. Little, *Biochem. J.*, 1973, **133**, 67.
[545] G. Allen and G. Lowe, *Biochem. J.*, 1973, **133**, 679.
[546] G. Gosselin-Rey, N. Bernard, and C. Gerday, *Biochim. Biophys. Acta*, 1973, **303**, 90.
[547] G. V. Patil and R. A. Day, *Biochim. Biophys. Acta*, 1973, **293**, 490.
[548] (*a*) L. F. Matyash, O. G. Ogloblina, and V. M. Stepanov, *European J. Biochem.*, 1973, **35**, 540; (*b*) G. Mains, R. H. Burchell, and T. Hofmann, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 275.
[549] M. Silverberg and K. Dalziel, *European J. Biochem.*, 1973, **38**, 229.
[550] A. Brevet, C. Roustan, G. Desvages, L.-A. Pradel, and N. van Thoai, *European J. Biochem.*, 1973, **39**, 141.
[551] M. A. Wells, *Biochemistry*, 1973, **12**, 1086.
[552] H. G. Baumert, H. Fasold, F. Keller, M. Halbach, and F. Ostanderl, *F.E.B.S. Letters*, 1973, **31**, 23.
[553] K. C. Robbins, P. Bernabe, L. Arzadon, and L. Summaria, *J. Biol. Chem.*, 1973, **248**, 1631.
[554] R. L. Nachman, A. Hubbard, and B. Ferris, *J. Biol. Chem.*, 1973, **248**, 2928.
[555] Y. Ando and M. Steiner, *Biochim. Biophys. Acta*, 1973, **311**, 26.
[556] K. Nagami, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 364.
[557] L. K. Ashman, J. C. Wallace, and D. B. Keech, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 294.

Table 3 (*cont.*)

Protein	*Source*	*Label*	*Residue modified*	*Comments*	*Ref.*
Pyruvate carboxylase	Rat liver/yeast	Sodium trinitrobenzene-sulphonate	Lys		558
Pyruvate kinase	Rabbit muscle	Affinity label	Lys	Reversible	559
Riboflavin-binding protein	Chicken eggs	Iodoacetamide	Tyr		422
RNA polymerase	*E. coli*	*p*-Mercuribenzoate	Cys		560
RNA polymerase	*E. coli*	*p*-CMB, DTNB	Cys		561
RNA polymerase	*E. coli*	Affinity labels	—	B and B′ subunits	593
Ribosomal proteins	*B. stearothermophilus*	Tos-PheCH_2Cl	—	S1 + S3	562
Ribosomal proteins	*E. coli*	Fluorescein isothiocyanate	—	50*S* subunit	563
Ribosomal proteins	*E. coli*	DTNB	Cys		564
Ribosomal proteins	*E. coli*	Iodine	—	30*S* subunit	565, 566
Ribosomal proteins	*E. coli*	S1-16S RNA cross-link	NH_2	30*S* subunit	597
Ribosomal proteins	*E. coli*	Diethyl malonimidate	—	Cross-links subunits	
Ribosomal proteins	*E. coli*	Methyl 4-mercaptobutyrimidate	Lys, Cys	Cleavable cross-link	431
Ribosomal proteins	*E. coli*	TNM cross-link	—	30*S* subunit	596
Ribosomal proteins	*E. coli*	Bromoamphenicol	—	Labels L_2, L_{27}	583*a*
Ribosomal proteins	*E. coli*	Iodoamphenicol	—	Labels L_{16}	581
Ribosomal proteins	*E. coli*	Bromoacetyl-Phe-$tRNA^{Phe}$	—	Labels L_2, L_{27}	583*b*
Ribosomal proteins	*E. coli*	Cross-links	—	Cross-links EF-G to 50*S* subunit	567
Serum albumin	Bovine	Phenyl *N*-methylacetimidate	Lys		453
Subtilisin	*B. subtilis*	*p*-Nitrophenyl diazoacetate	—	Photoaffinity label	594
Superoxide dismutase	Bovine	Photo-oxidation	His	Apoenzyme	568
Thermolysin	*B. thermoproteolyticus*	Diethyl pyrocarbonate	His	Superactive form	569
Thioredoxin	*E. coli*	*N*-Bromosuccinimide	Trp		570
Thrombin	Bovine	*N*-Acetylimidazole	Tyr		571
Tobacco mosaic virus	—	Methyl picolinimidate	Lys	Coat protein mutants	572

tRNA Synthetase-Ile	*E. coli*	*N*-Bromoacetyl-Ile-tRNA	—	Affinity label	584
tRNA Synthetase-Ser	*E. coli*	Thiol reagents	Cys		573
tRNA Synthetase-Val	Bovine liver	Alkylation	—		574
Triose phosphate isomerase	Rabbit muscle	Glycidol phosphate	Glu	Affinity label	586
Triose phosphate isomerase	Human	Chloroacetol phosphate	Glu	Affinity label	575
Trypsin	Bovine	Various	Lys	ε-Guanidated trypsin	576, 577
Trypsin	—	Octyl isocyanate	—	Not inhibited	462
Trypsin	—	Butyl isocyanate	—	Not inhibited	462
Trypsin inhibitor	(Kunitz)	Phenylglyoxal	Arg		578
Trypsinogen	Bovine	Guanidination	Lys		576

[558] M. C. Scrutton and M. D. White, *J. Biol. Chem.*, 1973, **248**, 5541.
[559] F. Davidoff, S. Carr, M. Lanner, and J. Leffler, *Biochemistry*, 1973, **12**, 3017.
[560] J. D. Harding and S. Beychok, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 711.
[561] B. H. Nicholson and A. M. Q. King, *European J. Biochem.*, 1973, **37**, 575.
[562] J. Jonák, J. Sedláchek, and I. Rychlík, *Biochim. Biophys. Acta*, 1973, **294**, 322.
[563] M. Hsiung and C. R. Cantor, *Arch. Biochem. Biophys.*, 1973, **157**, 125.
[564] A. S. Acharya and P. B. Moore, *J. Mol. Biol.*, 1973, **76**, 207.
[565] R. V. Miller and P. S. Sypherd, *J. Mol. Biol.*, 1973, **78**, 527.
[566] R. V. Miller and P. S. Sypherd, *J. Mol. Biol.*, 1973, **78**, 539.
[567] A. S. Acharya, P. B. Moore, and F. M. Richards, *Biochemistry*, 1973, **12**, 3108.
[568] H. J. Forman, H. J. Evans, R. L. Hill, and I. Fridovich, *Biochemistry*, 1973, **12**, 823.
[569] S. Blumberg, B. Holmquist, and B. L. Vallee, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 987.
[570] A. Holmgren, *J. Biol. Chem.*, 1973, **248**, 4106.
[571] R. L. Lundblad, J. H. Harrison, and K. G. Mann, *Biochemistry*, 1973, **12**, 409.
[572] R. N. Perham, *Biochem. J.*, 1973, **131**, 119.
[573] R. M. Waterson, S. J. Clarke, F. Kalowsek, and W. H. Konigsberg, *J. Biol. Chem.*, 1973, **248**, 4181.
[574] L. Y. Frdova, G. K. Ksvaleva, M. B. Agalavova, and L. L. Kisseler, *F.E.B.S. Letters*, 1973, **34**, 213.
[575] F. C. Hartman and R. W. Gray, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 388.
[576] N. C. Robinson, H. Neurath, and K. H. Walsh, *Biochemistry*, 1973, **12**, 414.
[577] N. C. Robinson, H. Neurath, and K. H. Walsh, *Biochemistry*, 1973, **12**, 420.
[578] J. F. DeLarco and I. E. Liener, *Biochim. Biophys. Acta*, 1973, **303**, 274.

measured.[579] In one case the hapten was non-covalently bound in the combining site of the antibody and in another case it was bound in the combining site of the pepsin Fab′ fragment of anti-Dns (Figure 1). By measuring the efficiency of energy transfer between these donor chromophores and an acceptor molecule, difluoresceinthiocarbamyl cysteine bound

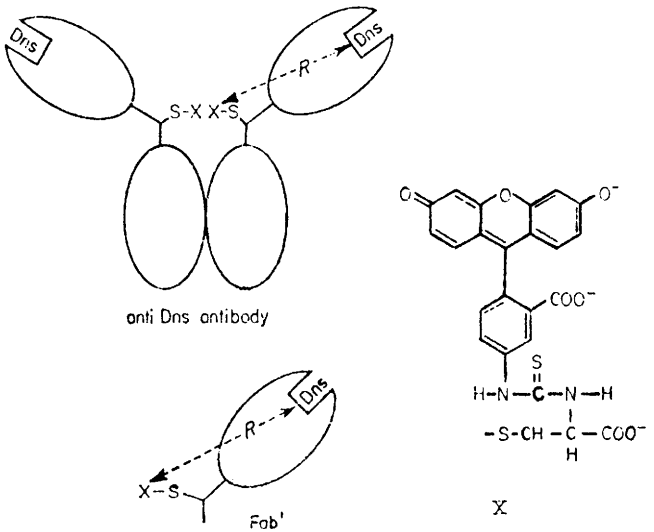


Figure 1 *Schematic representation of antibody–dye preparations. The donor, DnsLys, is bound specifically in the combining sites of rabbit anti-Dns antibody and Fab′ fragment; X is the acceptor, Fl-D,L-Cys, specifically attached at the thiols generated by selective reduction of the interheavy chain disulphide bond. The distance between the donor and acceptor chromophore is shown by* R
(Reproduced by permission from *J. Mol. Biol.*, 1973, **77**, 223)

to a thiol generated by selective reduction of the inter-heavy chain S—S bridge, an average separation of 81 Å was calculated. Taking crystallographic evidence[580] into account, this places the combining site at, or very close to, the tip of the Fab fragment and the inter-heavy chain disulphide bond at, or near the edge of, the C_L–CP_H I domain.

Ribosomal Proteins. Of the chemical methods available for determining the function and spatial distribution of ribosomal proteins, general chemical modification (see Table 3) rarely provides useful information. Cross-linking experiments (see below) do provide some details of the ribosomal topography but one would expect specific affinity labelling of individual ribosomal proteins to be relatively unambiguous. However, recent attempts to label the peptidyltransferase site in *E. coli* ribosomes

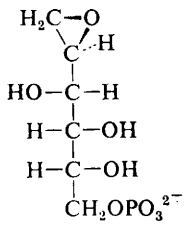
[579] J. R. Bunting and R. E. Cathou, *J. Mol. Biol.*, 1973, **77**, 223.
[580] R. J. Poljak, L. M. Amzel, H. P. Avey, L. N. Becka, and A. Nisonoff, *Nature New Biol.*, 1972, **235**, 137.

have not yielded clear answers. Monoiodoamphenicol, an analogue of chloramphenicol which has been known for many years to be a strong but reversible inhibitor of protein biosynthesis, has recently been shown[581] to be an irreversible inhibitor of protein synthesis in *E. coli* ribosomes. Almost all of the label was found to be bound to protein L16, and this conclusion has been supported by reconstitution experiments.[582] Hence, the inference is that this protein forms part of the peptidyltransferase site. However, similar experiments by a group at the Weizmann Institute,[583a] using bromamphenicol, show that most of the label is attached to proteins L2 and L27 and hardly any to L16. This result is supported by a New York group,[583b] who find an essentially similar labelling pattern using bromoacetylphenylalanyl-tRNA^{Phe}.

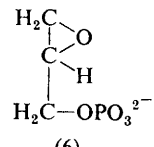
This difference in labelling could be simply a steric effect or it could be that L2 and L27 both possess thiol groups that are reactive to the bromofunction.

Aminoacyl-tRNA Synthetases. A similar reagent to the tRNA derivative above has been used for labelling isoleucyl-tRNA synthetase from *E. coli*.[584] *N*-Bromoacetyl-[^{3}H]Ile-tRNA was found to bind irreversibly to the enzyme, and several lines of evidence suggest that in this case it acts as a genuine affinity label. First, under comparable conditions, bromoacetate does not bind to the enzyme. Secondly, whereas tRNA^{Ile} does protect the enzyme from alkylation, tRNA^{Phe} does not. Thus, it appears that there is a specific interaction between the synthetase and tRNA moiety followed by covalent bond formation between the alkylating group and an, as yet unidentified, amino-acid residue of the protein.

Isomerases. Phosphoglucose isomerase catalyses the abstraction and intramolecular transfer of the C-2 proton of glucose-6-phosphate to form fructose-6-phosphate. A competitive inhibitor, 1,2-anhydro-D-mannitol-6-phosphate (5), has been shown[585] to bind to a glutamate residue in the



(5)



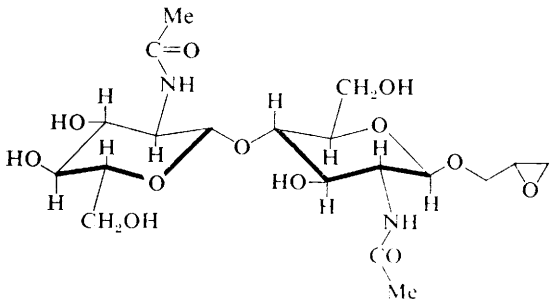
(6)

[581] O. Pongs, R. Bald, and V. A. Erdmann, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2229.
[582] D. Nierhaus and K. H. Nierhaus, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2224.
[583] (*a*) N. Sonenberg, M. Wilchek, and A. Zarnie, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1423; (*b*) H. Oen, M. Pellegrini, D. Eilat, and C. R. Cantor, *ibid.*, p. 2799.
[584] D. V. Santi and W. Marchant, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 370.
[585] E. L. O'Connell and I. A. Rose, *J. Biol. Chem.*, 1973, **248**, 2225.

enzyme in a 1 : 1 molar ratio. The label may be removed by hydroxylamine or alkali, indicating an ester linkage. Affinity labelling of triose phosphate isomerase with glycidol phosphate (6) has also indicated[586] a glutamic acid residue at the active site but from very limited sequence information on the phosphoglucose isomerase there appears to be no homology around this presumptive active-site residue in the two enzymes.

Dihydrofolate Reductase. Various 5-alkyldiaminopyrimidines have been used as affinity labels to probe the hydrophobicity of the active site of this enzyme.[587] When the substituent in the 5-position was a methyl or ethyl group no inhibition was observed. Other C-3 and C-4 substituents caused binding to the enzyme, mainly through van der Waals forces, and hydrophobic bonding only became important when the side-chains were of five or more carbon atoms. Good correlation was found between hydrophobicity and affinity for the enzyme except that cyclohexyl- and adamantyl-substituted pyrimidines had anomalously high affinities. These workers suggest that this may be due to the rigidity of the ring systems compared to the conformational freedom of the straight-chain hydrocarbons.

Lysozyme. Total enzymic hydrolysis has indicated[588] the point of attachment of the affinity label (7) to be aspartic acid-52. Peptic digestion of the



(7)

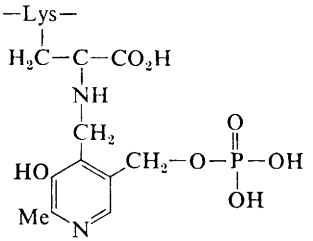
enzyme labelled with radioactive inhibitor produced a single derivatized peptide:

Asn-Thr-Gln-Ala-Thr-Asn-Arg-Asn-Thr-Asp-Gly-Ser-Thr-Asp-Tyr

Further digestion with the broad-specificity aminopeptidase from *Clostridium histolyticum* (see Methods section) produced a labelled peptide, Gly-Ser-Thr-Asp-Tyr. This corresponds to residues 49—53 in the sequence, and the point of attachment must therefore be aspartate-52, a conclusion in accord with recently published crystallographic evidence.

[586] K. J. Schray, E. L. O'Connell, and I. A. Rose, *J. Biol. Chem.*, 1973, **248**, 2214.
[587] Y. K. Ho, S. F. Zakrzewski, and L. H. Mead, *Biochemistry*, 1973, **12**, 1003.
[588] Y. Eshdat, J. F. McKelvy, and N. Sharon, *J. Biol. Chem.*, 1973, **248**, 5892.

Aspartate Aminotransferase. Several papers have appeared on the chemical modification of this enzyme (see Table 3) and a preliminary report on the complete amino-acid sequence has been given (see Table 6). A Japanese group [589] have now shown that the active site of the enzyme may be labelled by syncatalytic inactivation with β-chloro-[^{14}C]alanine. After borohydride reduction, carboxymethylation, and tryptic digestion, a radioactive peptide (8) has been isolated with a covalently bound cofactor (pyridoxal-5-phosphate). Sequence work shows the phosphopyridoxal group to be



(8)

bound to a lysine residue, and comparison with the full primary structure indicates this to be lysine-258.

The same workers have also inactivated this enzyme with bromopyruvate.[590] Loss of activity was only observed in the presence of L-cysteine sulphinate or L-aspartate and there was a simultaneous conversion from the pyridoxamine form of the cofactor into the pyridoxal form. Thus, bromopyruvate also acts as a keto-acid substrate in this reaction. Using a radioactive affinity label, inactivation was found to result from alkylation of a cysteine residue.

Photoaffinity Labelling.—*Immunoglobulins.* Attempts to locate and identify the antibody-combining regions of immunoglobulins with hapten-based affinity labels are generally limited by the reactivity spectrum of the reagent. Photoaffinity reagents which can be activated to give species such as carbenes, ketens, and nitrenes can insert into C—O, C—N, C—S, C—H, and N—H bonds to form covalent linkages between reagent and protein. Thus, those amino-acid residues without particularly reactive side-chains but which nevertheless participate in the antibody-combining site may be identified.

The ability of a photoaffinity-labelling reagent to react *in situ* at the active site will depend on several factors: (1) the spectrum of reactivity of the reactive group; (2) the distance between the reactive group and the binding moiety on the affinity reagent; and (3) equilibrium and rate considerations.

In a study of this type, *N*-(2,4-dinitrophenyl)alanyl diazoketone and 2,4-dinitrophenyl azide have been shown to have high affinities for protein

[589] Y. Morino and M. Okamoto, *Biochem. Biophys. Res. Chem.*, 1973, **50**, 1061.
[590] M. Okamoto and Y. Morino, *J. Biol. Chem.*, 1973, **248**, 82.

460, a mouse IgA myeloma protein with DNP-binding activity.[591, 592] In the dark there is no reaction and the binding of DNP-lysine is competitively inhibited. On photolysis, both reagents become covalently bound, with concomitant loss of DNP-lysine binding sites. DNP-azide generates a nitrene on photolysis (Figure 2) and residues within a sphere of 3.1 Å from

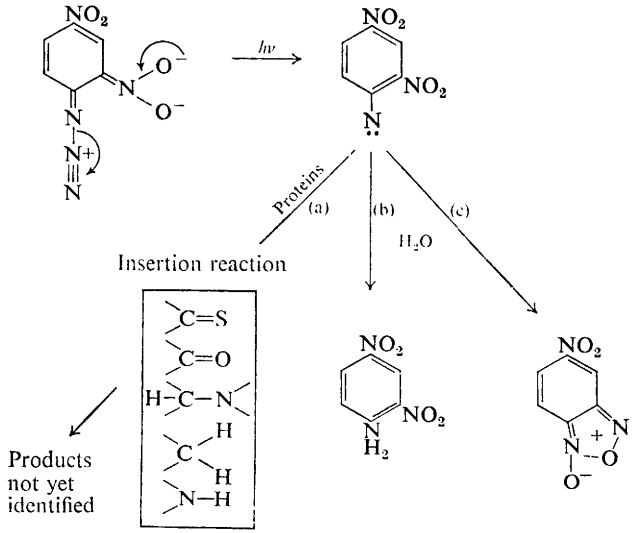


Figure 2 *Generation of* 2,4-*dinitrophenylnitrene and its possible reaction products* (Reproduced by permission from *Biochemistry*, 1973, **12**, 4679)

the centre of the DNP ring are capable of reacting when the DNP group is in the combining site. Irradiation of DNP-alanine diazoketone produces a carbene (Figure 3) which is capable of direct insertion into various covalent bonds. However, the authors propose that the carbene undergoes a rearrangement to produce a keten (Figure 3) which then reacts by acylation of nucleophiles. The position of the labels in the protein has been identified. The diazoketone derivative was bound predominantly to lysine-54 of the light chain, the same residue that reacts with conventional affinity labels. However, only 15% of the azide reagent was found in the light chain, and this was distributed along residues 29—58, 62—77, and 78—108. The remainder were located in two variable-region peptides of the heavy chain.

RNA Polymerase. Irradiation of *E. coli* RNA polymerase in the presence of the substrate analogue 4-thiouridine triphosphate or the template

[591] M. Yoshioka, J. Lifter, C.-L. Hew, C. A. Converse, M. Y. K. Armstrong, W. H. Konigsberg, and F. F. Richards, *Biochemistry*, 1973, **12**, 4679.

[592] C.-L. Hew, J. Lifter, M. Yoshioka, F. F. Richards, and W. H. Konigsberg, *Biochemistry*, 1973, **12**, 4685.

analogue poly(deoxy-4-thio-thymidylic acid) has been found to produce covalent substitution and subsequent irreversible inactivation.[593] SDS-Gel electrophoresis indicates that the β and β' subunits are the sole sites of modification.

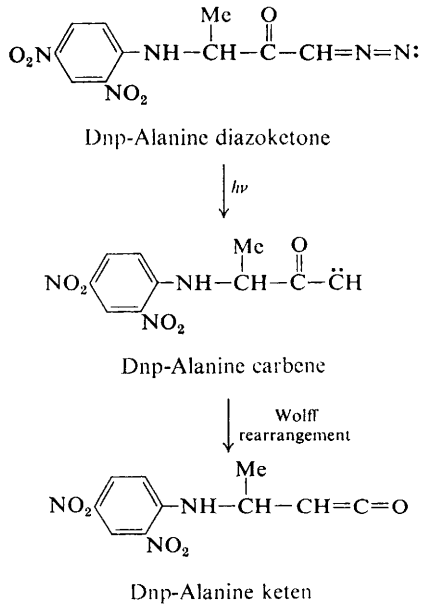
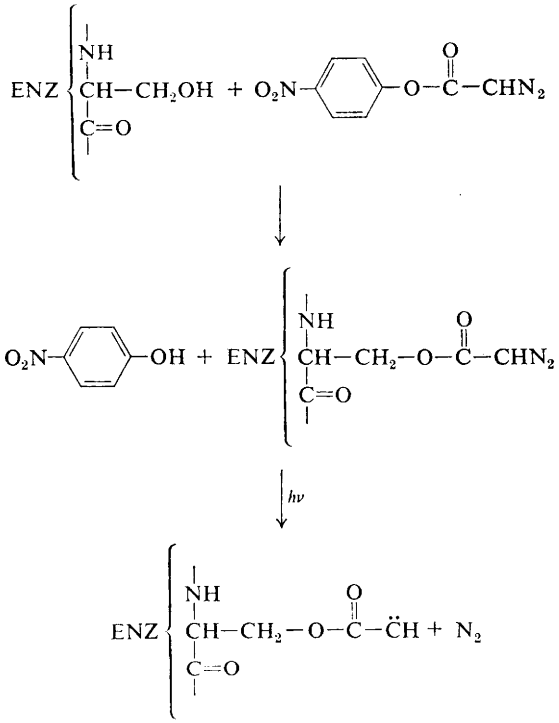


Figure 3 *Photolysis of Dnp-AD to dinitrophenylalanylcarbene and subsequent rearrangement to dinitrophenylalanylketen*
(Reproduced by permission from *Biochemistry*, 1973, **12**, 4685)

Subtilisin. The active sites of both trypsin and chymotrypsin have been mapped previously by the preparation of diazoacyl derivatives followed by photolysis and carbene formation. The carbene will then insert into residues in the active-site region. This approach has now been extended[594] to another serine esterase, subtilisin, using *p*-nitrophenyl diazoacetate. After reaction at pH 7.8 the diazoacetyl enzyme is unstable but may be isolated by rapidly reducing the pH to 5.0. Photolysis of the derivative (Scheme 4) followed by hydrolysis showed that several components had reacted, and these are presently being characterized.

Finally, a group of American workers has investigated the mechanism of photoaffinity labelling and has produced some rather unexpected results.[595] These authors show that inhibition of acetylcholinesterase with

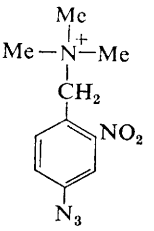
[593] A. M. Frischauf and K. H. Scheit, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 1227.
[594] Y. Stepanovsky and F. H. Westheimer, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1132.
[595] A. E. Rusho, H. Kiefer, P. E. Roeder, and S. J. Singer, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2567.



Preparation and photolysis of diazoacetyl subtilisin

Scheme 4

compound (9) satisfied all the normal criteria of genuine photoaffinity labelling but that there was also a fairly high level of non-specific reaction. On the assumption that this arose from excess photolysed reactive inhibitor



(9)

in solution, a scavenger molecule, *p*-aminobenzoate, was added. This successfully produced a ten-fold reduction in the level of bound inhibitor but, surprisingly, it also prevented the photolytic inactivation. Comparable results were obtained when small proteins were used as scavengers and

also in a similar photoaffinity-labelling experiment with DNP-antibodies. This would suggest that many apparent photoaffinity-labelling reactions are in fact occurring by ordinary affinity labelling, and that the bulk of the inhibitor is in free solution prior to photolysis rather than interacting at the active site.

Cross-linking.—*Ribosomal Proteins.* Tetranitromethane was originally introduced for the modification of tyrosine residues in proteins. However, several reports have appeared that under certain conditions it can introduce intermolecular cross-links, and this reaction has now been observed [596] for the proteins of the 30*S* ribosomal subunit from *E. coli.* After treatment with the reagent, four proteins, S11, S18, S21, and S12, disappeared from two-dimensional gel-electrophoresis fingerprints. The first three of these proteins were found to be cross-linked together but the fate of S12 has not yet been determined. Since prior incubation with iodoacetate had no effect on the reaction, the cross-link is presumably between two adjacent tyrosine residues rather than a reaction with thiol groups. This hypothesis is supported by the loss of tyrosine on amino-acid analysis. Proteins S18 and S21 have previously been cross-linked using the bifunctional reagent *NN'-p*-phenylenedimaleimide, and this lends support to a recently published [596a] assembly map of 30*S* ribosomal proteins (Figure 4) which suggests that these two, together with S11, are neighbouring proteins. The assembly map is based on the earlier work of Mizushima and Nomura, with slight subsequent modifications, but the nomenclature used is that of Wittmann *et al.* The thick arrow from 16*S* RNA to S4 indicates that S4 (P4a) binds directly to 16*S* RNA in the absence of other proteins. The thin arrow from 16*S* RNA to S7 indicates that S7 (P5) binds weakly to 16*S* RNA in the absence of other ribosomal proteins. Thin arrows pointing towards S7 from S4 (P4a), S8 (P4b), S20 (P14), S9 (P8), and S19 (P13) indicate that the latter proteins all help the binding of S7 (P5) to RNA. The thick arrow from S7 to S19 indicates that, in the absence of S7 (P5), S19 (P13) fails to bind the complexes containing 16*S* rRNA even in the presence of all other proteins under the standard reconstitution conditions. The arrow to S11 (P7) from the large box with a dashed outline indicates that S11 (P7) binding depends on some of the proteins enclosed in the box; it is not known exactly which proteins. The binding of S2 (P2) and S12 (P10) takes place at a later stage in the assembly sequence, but the exact position of these proteins in the map is not known. S1 (P1) does not bind under the conditions used. S16 and S17 were previously studied as a mixture; S17 (P9b), but not S16 (P9a), binds directly to 16*S* RNA. The specificity of this binding has not been proved and some investigators have not observed the direct binding of S13 to 16*S* RNA. Proteins above the dotted line in Figure 4 are those either required for the formation of RI* particles or found in the isolated 21*S* RI particles.

[596] C. T. Shih and G. R. Craven, *J. Mol. Biol.*, 1973, **78**, 651.
[596a] M. Nomura, *Science*, 1973, **179**, 864.

A nucleic acid–protein cross-link has been introduced between the 16*S* RNA and a protein tentatively identified as S1.[597] Oxidation of the 3′-ribosyl moiety of the RNA with $NaIO_4$ produced a dialdehyde, which reacted with an adjacent protein amino-function to form a Schiff's base.

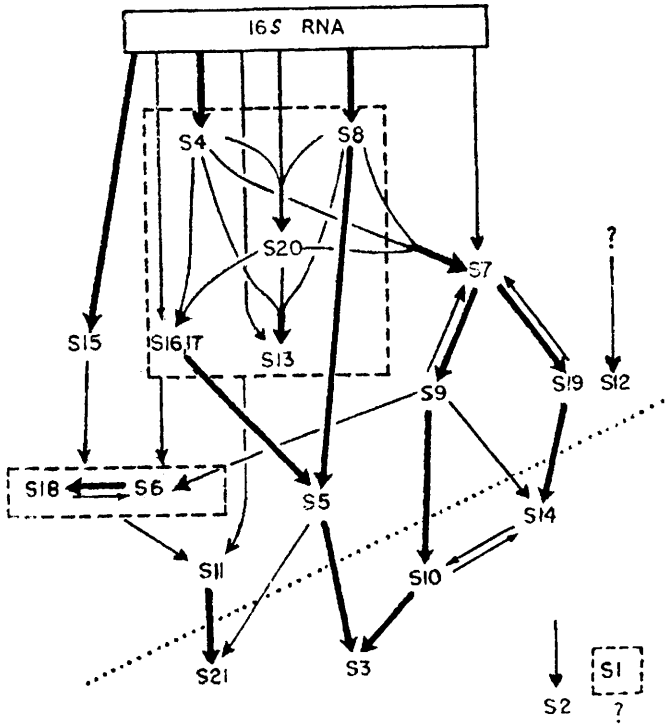


Figure 4 *The assembly map of 30S ribosomal proteins. Arrows between proteins indicate the facilitating effect on binding of one protein on another; a thick arrow indicates a major facilitatory effect*
(Reproduced by permission from *Science*, 1973, **179**, 864)

Reduction with borohydride produced a covalent cross-link, and SDS gels showed that the protein was bound to the 16*S* RNA. Identification of the protein component on this one-dimensional system was complicated by side-reactions of the periodate, and its identification as S1 must be regarded as tentative.

Treatment of complete 70*S* ribosomes with the bifunctional imido-ester diethyl malonimidate has been shown to cross-link the two ribosomal subunits.[598] Dissociation was observed only at low ionic strength (0.1mM-Mg^{2+}) but the cross-linked proteins were not identified.

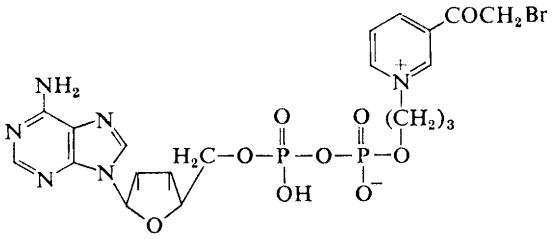
[597] R. A. Kenner, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 932.
[598] Y. Nakamura and A. Wada, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 35.

A preliminary investigation of the effect of a new reagent, methyl 4-mercaptobutyrimidate, on the 30*S* subunit from *E. coli* ribosomes has been published.[431] The advantage of this reagent is that it is easily cleaved by exposure to thiols, and since no charges are introduced, the cross-linked proteins are easily identified on two-dimensional gel systems. Thus, reaction with the 30*S* proteins followed by reduction and carboxymethylation gave the same two-dimensional fingerprint as a carboxymethylated control.

Erythrocyte Membranes. Reports have appeared on the use of difluorodinitrobenzene and dimethyl suberimidate to cross-link phospholipids to erythrocyte membrane proteins[599] and on the use of dimethyl adipimidate to cross-link sialoglycoproteins from similar membranes.[600] In neither case were the cross-linked species characterized.

Haemoglobin. Haemoglobin S is the variant haemoglobin found in erythrocytes of individuals carrying the genetic trait for sickle-cell anaemia. Since, under physiological conditions, haemoglobin exists in a tetramer ⇌ dimer equilibrium it should be possible to isolate an asymmetric hybrid of the type $\alpha_2^A\beta^A\beta^S$. This has not proved possible, presumably owing to a rapid exchange of β^S subunits. However, a recent report[601] shows that cross-linking with 4,4′-difluoro-3,3′-dinitrodiphenylsulphone will fix the various tetrameric haemoglobin species, allowing the isolation of the $\alpha_2^A\beta^A\beta^S$ hybrid. The concentration of this cross-linked form is greater than that of either of the homologous hybrids, which suggests that the $\alpha_2^A\beta^A\beta^S$ tetramer is the preponderant component in sickle-cell trait haemoglobin at equilibrium. It has previously been suggested that the low probability of sickling in erythrocytes of sickle-cell heterozygotes, compared with homozygotes, is due to difficulties in incorporating hybrid tetramers into the aggregates. This demonstration of a relatively high concentration of asymmetric hybrids lends weight to this suggestion.

Coenzyme Analogues.—*NAD.* A structural analogue of NAD^+, [3-(3-bromoacetylpyridinio)propyl]adenosine pyrophosphate (10), has been



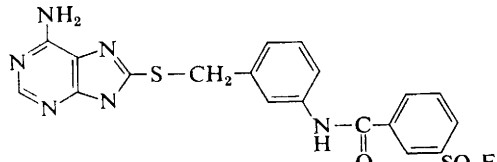
(10)

[599] G. V. Marinetti, R. Baumgarten, D. Sheeley, and S. Gordesky, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 302.
[600] T. H. Ji, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 508.
[601] R. M. Macleod and R. J. Hill, *J. Biol. Chem.*, 1973, **248**, 100.

found to inactivate alcohol dehydrogenase (ADH) from yeast[602] and glyceraldehyde-3-phosphate dehydrogenase (GPD) from rabbit muscle.[603] Initially a reversible dehydrogenase–inactivator complex is formed but after treatment with dithionite the acetylpyridinium ring is converted into a dihydropyridine system, and the inhibitor becomes covalently bound. Using radioactive analogues it could be shown that ADH incorporated 1 mole per mole of subunit whereas GPD was inactivated after incorporation of only 0.5 mole of inhibitor per mole of subunit. To identify the binding sites on the enzymes the dinucleotide part of the label was removed with hydrogen peroxide and the proteins were hydrolysed. In the case of the alcohol dehydrogenase all the radioactivity was found in carboxymethylhistidine whereas in GPD only cysteine was labelled.

ATP. Phosphofructokinase, which normally requires ATP for catalytic activity, has been shown[604] to function with an alternative phosphoryl donor, 6-mercapto-9-β-D-ribofuranosyl-purine 5′-triphosphate (sRTP). At high concentrations of this analogue (>0.4 mmol l^{-1}) the enzyme is inhibited, an observation which has also been made for the natural cofactor. However, incubation of the sRTP-inhibited protein with thiols restored activity and caused release of the inhibitor. This release could not be effected by dialysis or TCA precipitation, suggesting that an S—S bridge had been formed. Investigation of the binding stoicheiometry showed 6—8 mole of bound sRTP per 380 000 g of phosphofructokinase, but whether or not the binding is truly at the active site must await further investigation.

AMP. In order to investigate the nature of the AMP binding site of rabbit muscle glycogen phosphorylase *b*, various adenine and hypoxanthine derivatives have been evaluated[605] as competitive inhibitors towards AMP. A direct correlation with hydrophobicity was found and one derivative, 8-[*m*-(*m*-fluorosulphonylbenzamido)benzylthio]adenine (11), was found not only to bind irreversibly but also to yield an enzyme form with 24% of the control activity. Since the inhibitor contains neither a ribose ring



(11)

[602] C. Woenckhaus, M. Zoltobrocki, J. Berghäuser, and R. Jeck, *Z. physiol. Chem.*, 1973, **354**, 60.
[603] C. Woenckhaus, R. Jeck, E. Schättle, G. Dietz, and G. Jentsch, *F.E.B.S. Letters*, 1973, **34**, 175.
[604] D. P. Bloxham, M. G. Clark, P. C. Holland, and H. A. Lardy, *Biochemistry*, 1973, **12**, 1596.
[605] R. A. Anderson and D. J. Graves, *Biochemistry*, 1973, **12**, 1895.

nor a phosphate group, this activation is rather unexpected. The enzyme covalently binds 1 mole of analogue per mole.

In a subsequent paper [606] these authors have described the isolation of a pentapeptide containing the bound inhibitor and with composition Gly_2, Ala_2, Tyr. The presumed site of attachment is the phenolic hydroxy-group of the tyrosine residue.

4 Methods

Fragmentation.—*Aminopeptidases.* A fairly simple purification procedure yields an aminopeptidase from *Clostridium histolyticum* which is capable of cleaving all types of *N*-terminal amino-acids.[607] So far it has only been used on small peptides, and as pointed out in a report on the properties of a similar enzyme from *Aeromonas proteolytica*,[608] results obtained from proteins are often different. The *Aeromonas* enzyme is interesting in that acidic residues are not cleaved and glycyl residues are only released very slowly. Enzymes such as this with defined specificity ranges are always useful in protein chemistry and in this case an obvious application would be amide determinations.

Unfortunately, neither of these enzymes is commercially available whereas the leucine aminopeptidase from bovine lens is. A further study of the specificity of this enzyme has been made,[609] and a sensitive detection method for contaminating endopeptidases is also included in this paper. A method for the immobilization of the enzyme and the related aminopeptidase M to arylamino-glass has been published.[610] Both enzymes retain activity and may be used for sequencing or, in conjunction with pronase, for total enzymic digestion.

Carboxypeptidase. A method for the large-scale preparation of carboxypeptidase Y, the enzyme from yeast that releases all *C*-terminal amino-acids, has been described.[611] Apart from the useful broad specificity, this carboxypeptidase is also active in 6M-urea and possesses amidase activity. It would appear to be related, at least in specificity, to carboxypeptidase C from citrus leaves (see last year's Report, p. 45) which is now commercially available.

Di- and Tri-peptidases. Following the demonstration by Ovchinnikov and Kiryushkin (see last year's Report, p. 49) that dipeptides liberated by using dipeptidyl aminopeptidase 1 could be identified by a combination of g.l.c. and mass spectrometry, an American group [612] have now investigated the

[606] R. A. Anderson, R. F. Parrish, and D. J. Graves, *Biochemistry*, 1973, **12**, 1901.
[607] E. Kessler and A. Yaron, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 405.
[608] S. H. Wilkes, M. E. Bayliss, and J. M. Prescott, *European J. Biochem.*, 1973, **34**, 459.
[609] B. Wiederanders, J. Lasch, H. Kirschke, P. Bohley, S. Ansorge, and H. Hanson, *European J. Biochem.*, 1973, **36**, 504.
[610] G. P. Roger and J. P. Andrews, *J. Biol. Chem.*, 1973, **248**, 1807.
[611] R. Hayashi, S. Moore, and W. H. Stein, *J. Biol. Chem.*, 1973, **248**, 2296.
[612] R. M. Caprioli, W. E. Seifert, and D. E. Sutherland, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 67.

method using insulin A-chain as substrate. No major problems were encountered but this peptide contains no proline or basic amino-acids, all of which provide difficulties. Prolyl bonds, in particular, are not cleaved but reports on the purification [613] and specificity [614] of an iminodipeptidase from pig kidney may offer a solution to this problem.

An aminotripeptidase which cleaves the *N*-terminal residue from any tripeptide not having aspartic or glutamic acid at the *N*-terminus, or proline in the central position, has been purified from swine kidney.[615] Although allowing a rapid determination of the sequences of tripeptides,[616] the purification procedure is rather tedious and the work involved probably outweighs the advantages of using the enzyme.

Chemical Fragmentation. It has been known for some years that cyanide ion will cause peptide bond cleavage in cysteine-containing proteins but it was only after Degani and Patchornik introduced 2-nitro-5-thiocyanobenzoic acid as the CN^- source that the reaction could be made to go to completion. A recent publication [617] shows that in 6M-guanidinium chloride–0.1M-sodium borate, pH 9.0, cleavage was complete after 12 h at 37 °C. No significant side-reactions were detected. A mechanism is proposed (Figure 5) whereby specific hydroxide-ion catalysis is followed by concerted peptide bond cleavage and ring closure, without the intermediate formation of an acyliminothiazolidine. As can be seen from Figure 5, the new fragments possess an NH_2-terminal iminothiazolidinyl residue, and as yet there is no way of removing this blocking group. This reduces the value of the method considerably as these large fragments would otherwise have been ideal for automated sequence analysis. (This method has now been slightly amended.[618])

Hydroxylamine is coming into more common usage for the cleavage of asparaginylglycine bonds. A mechanism has been proposed [619] for this cleavage whereby the reaction proceeds through the intermediate formation of a succinimide ring.

Fractionation.—Strid has made the interesting observation [620] that in the presence of charged detergents such as dodecylamine or dodecanoic acid, peptides are fractionated on Sephadex not only by size but also according to their charge. Thus, two large fragments (97 and 76 residues) from carbonic anhydrase B are eluted from Sephadex G-50 in the expected order when 1M-acetic acid is used as eluent. However, addition of 2%

[613] H. Mayer and A. Nordwig, *Z. physiol. Chem.*, 1973, **354**, 371.
[614] A. Nordwig and H. Mayer, *Z. physiol. Chem.*, 1973, **354**, 380.
[615] D. Chenoweth, R. E. J. Mitchel, and E. L. Smith, *J. Biol. Chem.*, 1973, **248**, 1672.
[616] D. Chenoweth, D. M. Brown, M. A. Valenzwela, and E. L. Smith, *J. Biol. Chem.*, 1973, **248**, 1684.
[617] G. R. Jacobson, M. H. Schaffer, G. R. Stark, and T. C. Vanaman, *J. Biol. Chem.*, 1973, **248**, 6583.
[618] Y. Degani and A. Patchornik, *Biochemistry*, 1974, **13**, 1.
[619] M. Deselnica, P. M. Lange, and E. Heidemann, *Z. physiol. Chem.*, 1973, **354**, 105.
[620] L. Strid, *F.E.B.S. Letters*, 1973, **33**, 192.

dodecylamine to the solvent reverses this order. Strid proposes that since the critical micelle concentration of the detergent is 1.2×10^{-2} mol l^{-1} (~0.2%) the micelles formed are probably too large to penetrate the gel matrix, which leads to a concentration gradient of dodecylammonium ions. This produces a Donnan effect which would be expected to influence

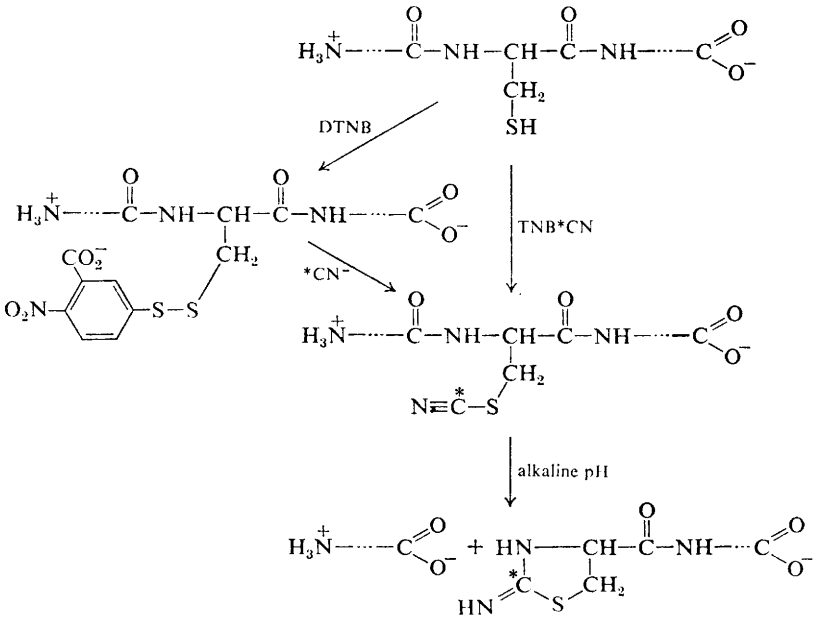


Figure 5 *Flow chart for the modification and cleavage reactions*
(Reproduced by permission from *J. Biol. Chem.*, 1973, **248**, 6583)

the partition of the peptides such that negatively charged molecules would tend to be outside the gel and positively charged molecules would tend to appear inside. The anionic detergent dodecanoic acid would show the reverse effect, and this is in fact the case. The potential usefulness of this method for purification of large fragments is increased by the ease of removal of the detergents. Extraction with heptane–butanol (5 : 1) is all that is required.

A conventional automated peptide analyser has been described [621] and a modified gradient introduced.[622] A preparative, but rather long and cumbersome, method for the isolation of *C*-terminal tryptic peptides from proteins has also been published.[623] Cystine-containing peptides may be detected in column effluents by reduction with $NaBH_4$ in guanidine

[621] K. D. Lin and H. F. Deutsch, *Analyt. Biochem.*, 1973, **56**, 155.
[622] E. Triantaphyllopoulos, *Biochem. J.*, 1973, **133**, 843.
[623] P. A. Hargrave and F. Wold, *Internat. J. Peptide Protein Res.*, 1973, **5**, 85.

followed by destruction of excess reducing agent by acid and acetone.[624] Liberated thiol groups are quantitated using DTNB.

One of the most powerful fractionation techniques available is electrophoresis in polyacrylamide gels containing the detergent sodium dodecylsulphate (SDS). The method involves the fractionation of microgram quantities of material in a denatured state, and although methods are available for both scaling-up the procedure and re-naturation of the proteins, these are not always feasible at the practical level. However, following the demonstration that proteins eluted from these gels may be characterized by their *N*-terminal sequence,[625] it has now been shown [626] that they may also be analysed by iodination with ^{125}I, enzymic digestion, and fingerprinting of the radioactive peptides. The method has been used on actin, tubulin, and paramyosin with both one- and two-dimensional fingerprint methods. The smallest sample was 2 ng (120 fmol) of lysozyme and the method will clearly be useful for characterizing small quantities of protein.

Amino-acid Analysis.—*Sample Preparation.* Two papers advocate the dilution of samples with low pH buffer after hydrolysis rather than evaporating off the hydrolysis acid and redissolving in buffer. The first of these [627] emphasizes the greater accuracy of the method as adsorption of amino-acids onto the glass hydrolysis tube is minimized. The second paper [628] simply states the method to be more convenient. A new internal standard, *S*-β-(4-pyridylethyl)-DL-penicillamine, has been proposed [629] although it appears to have been tested only on the short (basic amino-acids) column system. For the measurement of thiol groups, 3-bromopropionic acid has been proposed [630] as an alternative to iodoacetate (see Chemical Modification section, p. 78).

Buffer Systems. A gradient-elution system using lithium citrate buffers has been employed [631] for the resolution of more than 50 amino-acids. Apparently, the separation is improved over conventional step-wise buffer-change systems. However, the same authors also propose [632] a mixed lithium–sodium gradient system which they say has the advantages of cost, availability, and decreased environmental pollution over all-lithium systems. Finally, an ammonia filtration system using a small column of Dowex 50W-X8, 20–50 mesh, has been published.[633] The filter operates on the first two buffers only. The third buffer contains 10% (v/v)

[624] A. F. S. A. Habeeb, *Analyt. Biochem.*, 1973, **56**, 60.
[625] A. M. Weiner, T. Platt, and K. Weber, *J. Biol. Chem.*, 1972, **247**, 3242.
[626] D. Bray and S. M. Brownlee, *Analyt. Biochem.*, 1973, **55**, 213.
[627] E. Robel, *Analyt. Biochem.*, 1973, **51**, 137.
[628] H. D. Spitz, *Analyt. Biochem.*, 1973, **56**, 66.
[629] M. Friedman, A. T. Noma, and M. S. Mani, *Analyt. Biochem.*, 1973, **51**, 280.
[630] A. F. Bradbury and D. G. Smyth, *Biochem. J.*, 1973, **131**, 635.
[631] J. L. Young and M. Yamamoto, *J. Chromatog.*, 1973, **78**, 221.
[632] J. L. Young and M. Yamamoto, *J. Chromatog.*, 1973, **78**, 349.
[633] L. G. Gürtler, *J. Chromatog.*, 1973, **76**, 255.

methanol to improve resolution of lysine and ammonia but whether the methanol affects the life of the fractionation resin is not discussed.

Unusual Amino-acids. Conditions have been formulated [634] for the separation of 5-hydroxylysine, ornithine, lysine, N^{ϵ}-monomethyl-lysine, N^{ϵ}-dimethyl-lysine, histidine, N^{τ}-methylhistidine, N^{π}-methylhistidine, ammonia, methylamine, $N^{G}N^{G}$-dimethylarginine, $N^{G}N'^{G}$-dimethylarginine, N^{G}-monomethylarginine, and arginine. Essentially a 30 cm column of Durrum DC-2A resin is eluted with 0.35M-citrate, pH 5.8, at 28 °C. After 200 min a new buffer of the same molarity but at pH 4.7 is substituted and the temperature increased to 55.5 °C. Slight adjustment of the pH of the buffers may be used to optimize any required separation. The methods have been applied to myelin basic protein from various species, which is known to contain methylated arginines.

The amino-acids desmosine, isodesmosine, mesodesmosine, and lysinonorleucine, commonly found in elastin hydrolysates, can be resolved using a two-column system. In this method [635] a conventional acidic + neutrals analysis is performed on the long column and this is then eluted with short-column buffer, which effects a good separation of these amino-acids. The same workers have devised a similar method for the resolution of hydrolysis products of collagen and elastin cross-links.[636]

Capillary Analyser. An instrument capable of detecting down to 60 pmol of an amino-acid has been described.[637] Unfortunately, serine and aspartic acid are not resolved and the system, which is designed primarily for physiological fluids, is a two-column one with a long total programme time.

Integration of Results. Surprisingly little has been published on this topic in 1973. A relatively economic method has been presented [638] which uses a digital voltmeter to take a signal direct from the chart recorder. The signal is converted into a binary coded form and punched onto paper tape for computer processing. The program, written in FORTRAN IV, allows the following information to be printed out: peak number, peak start and end, area, area ratio to norleucine standard, peak name, residue %, and residue % amino N_2.

Fluorescamine. Five papers have appeared on the use of this reagent for amino-acid analysis, all of them from either the Roche Institute of Molecular Biology or Hoffmann-La Roche Inc. One of these [639] gives details of a two-column instrument capable of analysing down to 1 μg of bovine serum albumin. However, proline, as secondary amine, is not detected.

[634] G. E. Deibler and R. E. Martenson, *J. Biol. Chem.*, 1973, **248**, 2387.
[635] D. Volpin and G. Michelotto, *J. Chromatog.*, 1973, **79**, 335.
[636] D. Volpin, M. G. Giro, and I. Castelloni, *J. Chromatog.*, 1973, **79**, 337.
[637] A. M. Gressner, *Analyt. Biochem.*, 1973, **56**, 532.
[638] R. Taylor and M. G. Davies, *Analyt. Biochem.*, 1973, **51**, 180.
[639] S. Stein, P. Bohlen, J. Stone, W. Dairman, and S. Udenfriend, *Arch. Biochem. Biophys.*, 1973, **155**, 203.

Following the publication [640] of a method for oxidizing proline to a primary amine with *N*-chlorosuccinimide, details have appeared [641] of a single-column analyser which allows the automated addition of the oxidizing reagent at the point in the programme where proline is expected. Optimum amino-acid concentration for this system would appear to be about 400 pmol although smaller amounts can be detected. An alternative single-column methodology, but without *N*-chlorosuccinimide addition, has also been published from the same laboratory.[642] One problem is that acetone, used for dissolving the fluorescamine, will precipitate salts from buffers of high ionic strength. The inability to use these buffers means that the run-time cannot be shortened much below 4 h with present systems. Some speeding-up has been achieved by using higher flow-rates and a stainless-steel column but in no cases are the operating pressures revealed. Finally, an automated system which will also detect hydroxyproline has been described.[643] Unfortunately, as with proline, the fluorescence yield is considerably less than that for the other amino-acids so that a minimum of 250 pmol is required for analysis. This may introduce difficulties in quantitation of some of the other amino-acids.

Several alternatives to fluorescamine have been proposed. One of these,[644] a mixture of 1-phthalaldehyde and 2-mercaptoethanol, has the advantage of being stable in aqueous solutions so that it may be dissolved directly in borate buffer, pH 10.0, and high-salt buffers may then presumably be used. Good quantitation is claimed down to 0.5 nmol. Systems using pyridoxal–Zn^{2+} have also been described [645, 646] and these are claimed to be one to two orders of magnitude more sensitive than conventional ninhydrin procedures.

Several simpler, more economical but possibly less convenient methods of high-sensitivity amino-acid analysis have been proposed. These are all based on reaction with the highly fluorescent 1-dimethylaminonaphthalene-5-sulphonyl (dansyl) chloride followed by two-dimensional t.l.c. of the dansyl-amino-acids. The differences stem from the method of quantitation used.

Varga and Richards [647] prefer fluorescence scanning, photocopying, and densitometry. As the fluorescence yield depends markedly on pH, humidity, solvent, and the batch of t.l.c. plates, it is essential to use double-sided plates with an internal marker of dansyl-amino-acids on one side and the sample on the other. The concentration of the marker mixture must be accurately known and the spots must both be the same size. Variations in

[640] M. Weigele, S. DeBernardo, and W. Leimgruber, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 352.
[641] A. M. Felix and G. Terkelsen, *Arch. Biochem. Biophys.*, 1973, **157**, 177.
[642] A. G. Georgiadis and J. W. Coffey, *Analyt. Biochem.*, 1973, **56**, 121.
[643] A. M. Felix and G. Terkelsen, *Analyt. Biochem.*, 1973, **56**, 610.
[644] M. Roth and A. Hampi, *J. Chromatog.*, 1973, **83**, 353.
[645] M. Maeda and A. Tsuji, *Analyt. Biochem.*, 1973, **52**, 555.
[646] M. Maeda, A. Tsuji, S. Ganno, and Y. Onishi, *J. Chromatog.*, 1973, **77**, 434.
[647] J. M. Varga and F. F. Richards, *Analyt. Biochem.*, 1973, **53**, 392.

these two parameters are likely to provide the major contributions to error. The solvent systems used for resolving the dansyl-amino-acids are the same as those conventionally used for the dansyl-Edman sequencing procedure but in a different order. Thus, solvent 1 is benzene–acetic acid (9 : 1), solvent 2 is formic acid–water (3 : 200), and solvent 3 is ethyl acetate–methanol–acetic acid (20 : 1 : 1). In this way dansic acid does not interfere. If solvent 3 is run in the same direction as solvent 1, dansyl-His separates from dansyl-Arg and the Asp and Glu derivatives also resolve. To resolve dansyl-Arg as well as the serine and threonine derivatives, solvent 3 is run in the same direction as solvent 2. This method of running the plates is likely to be of interest to people using conventional dansyl-Edman methods of sequencing as well as those trying to quantitate dansyl-amino-acids. Reasonable quantitation (± 15%) is claimed for the 5—20 pmol range and apparently spots can be detected at levels down to 1 fmol. A similar technique has been described[648] for quantitation of amino-acids in biological fluids, except that the samples were eluted from the plate and the fluorescence of the solution was measured. In cases where the amino-acid was radioactive, quantitation was by scintillation counting. A more accurate method is proposed by Brown and Perham.[649] In this case the mixture of unknown amino-acids is added to a known concentration of ^{14}C-labelled amino-acids and the mixture then dansylated with ^{3}H-labelled dansyl chloride. Determination of the ratio $^{3}H/^{14}C$ for each dansyl-amino-acid then allows accurate quantitation of the unlabelled amino-acid originally present. Accuracy is ± 6—8%, and as little as 20 pmol can be measured.

With all these micro-methods the reliability will depend on the operator as much as on the technique. A human fingerprint contains approximately 10 nmol of amino-acids, so that working at levels around 0.1% of this quantity requires clean samples, clean reagents, and a clean atmosphere, a combination not always found in biochemical laboratories.

Gas–Liquid Chromatography. Now that ion-exchange methods permit amino-acid analyses down to 50 pmol of material in less than 1 h, the advantages of a g.l.c. system are essentially those of simplification and economy. However, until a simple, reliable, and reproducible method becomes available for producing volatile derivatives the technique is unlikely to supplant conventional methods.

This year a single-column separation of *N*-trifluoroacetyl n-butyl ester derivatives has been presented,[650] and an investigation of the corresponding methyl esters has also been made.[651] The effect of varying the nature of the substituent group has been discussed theoretically.[652] A short glass

[648] J. Airhart, S. Sibiga, H. Sanders, and E. A. Khairallah, *Analyt. Biochem.*, 1973, **53**, 132.
[649] J. P. Brown and R. N. Perham, *European J. Biochem.*, 1973, **39**, 69.
[650] C. W. Gehrke and H. Takeda, *J. Chromatog.*, 1973, **76**, 63.
[651] A. J. Cliffe, N. J. Berridge, and D. R. Westgarth, *J. Chromatog.*, 1973, **78**, 333.
[652] S.-C. J. Fu and D. S. H. Mak, *J. Chromatog.*, 1973, **78**, 211.

capillary column has been used [653] for the resolution of *N*-heptafluorobutyryl-amino-acid n-propyl esters with promising results but once again only after a complicated derivatization procedure.

One area where g.l.c. does have an obvious application is in the identification of amino-acids which are either not resolved or not detected by conventional methods. A method had been described [654] for the estimation of the n-butyl esters of *N*-trifluoroacetylasparagine, glutamine, pyrrolidonecarboxylic acid, carboxymethylcysteine, homoserine, hydroxyproline, and ε-methyl-lysine. Tryptophan may also be detected by a similar method.[655]

Other Methods. High-resolution, high-frequency ^{1}H n.m.r. has been used [656] to determine the composition of equine cytochrome *c*; 10—20 mg of protein is required and it is not clear whether this is a serious suggestion for amino-acid analysis. The method does only take 10 min. For determination of tryptophan, magnetic c.d. has been proposed.[657] Results from 17 proteins (of known composition) are presented and these are most impressive, but the complexity of the instrumentation and its absence from most biochemical laboratories will preclude general use. Mass fragmentography, a technique where a mass spectrometer is used only to detect certain preselected ions, has been used for the simultaneous quantitation of ten amino-acids from soil.[658] The method is relatively rapid (1 h) and sensitive (0.1 nmol) but the number of amino-acids that can be quantitated is limited by the number of ions that can be monitored.

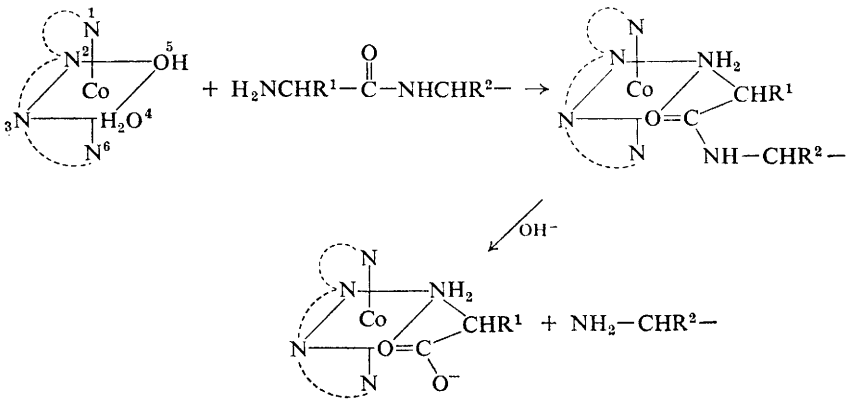
Terminal Group Identification.—Two methods have appeared for identification of *C*-terminal amino-acids in peptides and proteins. Both techniques are based on reduction to an amino-alcohol followed by hydrolysis and identification. In the first case [659] reduction is with sodium dihydrobis-(2-methoxyethoxy)aluminate and identification is by paper chromatography. The second method [660] uses sodium borohydride as reducing agent and the amino-alcohols are analysed using a conventional amino-acid analyser. Neither method appears as simple or convenient as the use of carboxypeptidases.

The thermally induced formation of methyl- and phenyl-thiohydantoin amino-acid derivatives from the corresponding *N*-methyl- and *N*-phenylthiourea derivatives of amino-acids and peptides has been shown [661] to

[653] J. Jönsson, J. Eyem, and J. Sjöquist, *Analyt. Biochem.*, 1973, **51**, 204.
[654] H. Hediger, R. L. Stevens, H. Brandenberger, and K. Schmid, *Biochem. J.*, 1973, **133**, 551.
[655] C. W. Gehrke and H. Takeda, *J. Chromatog.*, 1973, **76**, 77.
[656] G. Bemski and T. Hynes, *Nature*, 1973, **245**, 37.
[657] B. Holmquist and B. L. Vallee, *Biochemistry*, 1973, **22**, 4409.
[658] W. E. Pereira, Y. Hoyano, W. E. Reynolds, R. E. Summons, and A. M. Duffield, *Analyt. Biochem.*, 1974, **55**, 236.
[659] A. K. Saund, B. Prashad, A. K. Koul, J. M. Bachhawat, and N. K. Mathur, *Internat. J. Peptide Protein Res.*, 1973, **5**, 7.
[660] T. Hamada and O. Yonemitsu, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 1081.
[661] T. Fairwell, S. Ellis, and R. E. Lovins, *Analyt. Biochem.*, 1973, **53**, 115.

occur in the mass spectrometer. As methods already exist for the mass-spectrometric identification of MTH and PTH-amino-acids, this could provide a rapid, sensitive method for *N*-terminal amino-acid identification.

An alternative, simpler procedure [662] involves *N*-terminal peptide-bond hydrolysis with the *cis*-β-hydroxyaquo(triethylenetetramine)cobalt(III) ion (Scheme 5). The procedure is rapid and the derivatives may be identified

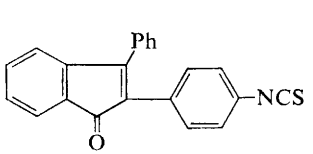


N-Terminal peptide hydrolysis with cobalt(III) tetra-amine chelates

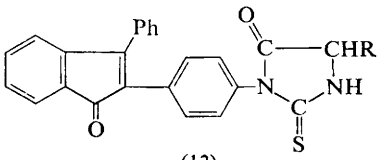
Scheme 5

directly or cleaved with cyanide ion, hydrogen sulphide, or borohydride to give the free amino-acid.

Sequence Analysis.—Two compounds have been proposed as alternatives to the phenyl isothiocyanate (PITC) normally used in the Edman degradation. One of these [663] is the analogue 2-*p*-isothiocyanophenyl-3-phenylindone (12), which interacts with α-amino-groups to give diphenylindonyl thiohydantoins (13). The advantage of this compound is that after t.l.c.



(12)



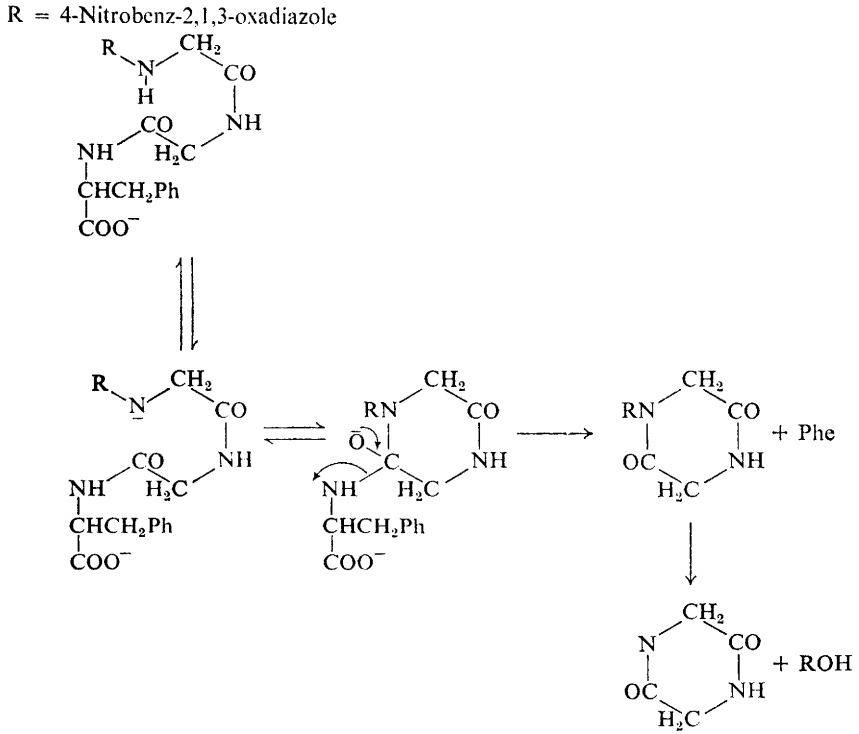
(13)

separation [664] the derivatives are coloured and may be identified directly.

The second method [665] is based on the fact that 4-nitrobenz-2,1,3-oxadiazole (NBD) derivatives of amino-acids can interact *via* a negatively

[662] K. W. Bentley and E. H. Creaser, *Biochem. J.*, 1973, **135**, 507.
[663] C. P. Ivanov and I. N. Mancheva, *Analyt. Biochem.*, 1973, **53**, 420.
[664] C. P. Ivanov and I. N. Mancheva, *J. Chromatog.*, 1973, **75**, 129.
[665] A. A. Aboderin, R. E. K. Semakula, E. Boedefeld, and R. A. Kenner, *F.E.B.S. Letters*, 1973, **34**, 90.

charged nitrogen with the carbonyl carbon atom of the second peptide bond. Whilst the mechanism is unclear the result is the removal of two amino-acids under relatively mild conditions (Scheme 6).

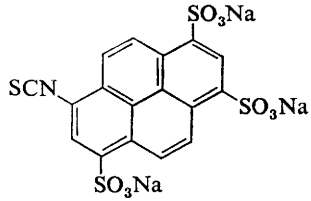


Proposed mechanism for the removal of Gly-Gly from the peptide Gly-Gly-Phe after reaction with NBD chloride

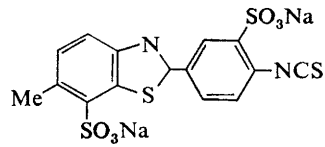
Scheme 6

Sequencer. Braunitzer has introduced[666] two new sulphonylated PITC derivatives primarily for automated sequence analysis. The problem here is that peptides tend to be extracted from the spinning-cup during the organic solvent washes, particularly after the ε-amino-groups of lysine have reacted with the hydrophobic PITC. Prior reaction with PITC derivatives carrying negatively charged sulphonic acid groups alleviate this problem and in some cases allows the complete sequence of the peptide to be obtained. The two new derivatives (14) and (15) have once more allowed the automated sequence analysis of the A and B chains of insulin.

[666] G. Braunitzer, B. Schrank, S. Petersen, and U. Petersen, *Z. physiol. Chem.*, 1973, **354**, 1563.



(14)



(15)

If the peptide contains no lysine residues, an alternative method [667] is to attach 2-aminonaphthalene-1,5-disulphonic acid to carboxy-groups *via* a water-soluble carbodi-imide. Using this method the peptide Gly-Ala-Arg-Gly-Gly-Val-Gly-Val-Gly-Gly-Leu-Pro-Thr-Phe was automatically sequenced, although yields were down to 0.9% at the penultimate residue. One snag is that the derivatized amino-acids are only identifiable by hydrolysis and subsequent amino-acid analysis.

Some impressive results on *N*-terminal ribosomal protein sequences have been obtained [668] using a much-modified Beckman sequencer. Ten proteins were analysed and sequences between 34 and 57 residues were derived. The details of the modifications, which are mostly to the vacuum systems, will not be given here but they include an automatic device for conversion of thiazolinones into thiohydantoins. Details of this unit are to be published separately. Whether, having spent about £20 000, one is prepared to follow this paper and virtually rebuild the instrument in the hope of improving the results must remain an individual decision. A full description of a manually operated peptide sequencer has also been given.[621]

Solid-phase Sequencing. Considerable advances have been made in this field and the technique is beginning to look a possible successor to spinning-cup sequencers as well as to the manual Edman degradation. It is certainly a method which no one involved in large-scale primary structure work can afford to neglect. Major advances include methods for attaching homoserine peptides and a general method for coupling *C*-terminal carboxy-groups to the resin. A new solid support has also been described which looks useful for protein degradation.

For coupling cyanogen bromide fragments [669] the *C*-terminal homoserine is first converted into the lactone by treatment with trifluoroacetic acid. After drying, the peptide is coupled directly onto a triethylenetetramine derivative of polystyrene (Figure 6) simply by stirring in DMF. Excess sites on the resin are then blocked with CH_3NCS and the peptide sequenced either manually, or preferably using the automatic instrument first

[667] J. A. Foster, E. Bruenger, C. L. Hu, K. Albertson, and C. Franzblau, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 70.
[668] B. Wittmann-Liebold, *Z. physiol. Chem.*, 1973, **354**, 1415.
[669] M. J. Horn and R. A. Laursen, *F.E.B.S. Letters*, 1973, **36**, 285.

described by Laursen and now available commercially. The only potential problem arises from the poor solubility of some peptides in DMF. This may be circumvented by adding a small amount of water to the coupling medium or substituting a hydrophilic support which will swell in aqueous media.

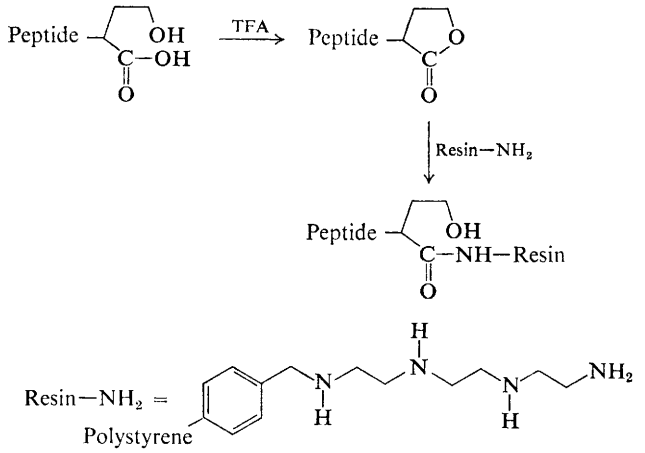


Figure 6 *Attachment of homoserine peptides to triethylenetetraminepolystyrene* (Reproduced by permission from *F.E.B.S. Letters*, 1973, **36**, 285)

For the specific attachment of peptides *via* their *C*-terminal carboxy-groups a water-soluble carbodi-imide is used.[670] After treatment with BOC-azide to block NH_2-groups, the *C*-terminal carboxy-group is activated with *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodi-imide (EDC) for 90 min and then stirred with the triethylenetetramine resin described above. Several peptides containing both aspartic and glutamic acid residues have been coupled in this way and their sequences have been determined with no unusual drops in yield. The success of the method may depend on the ability of the user to keep all nucleophiles absent during the activation reaction and thus minimize side-reactions. This is similar to the procedure previously described by Schellenberger *et al.*,[671] where resin, peptide, and EDC are mixed without prior blocking of the amino-terminus.

Aminopropyl-glass, prepared by treating controlled-pore glass beads with 3-aminopropyltriethoxysilane, has also been used as a solid support.[672] Peptides coupled, whether by carbodi-imide methods or by using *p*-phenylene di-isothiocyanate, have been satisfactorily degraded with a

[670] A. Previero, J. Derancourt, M. A. Coletti-Previero, and R. A. Laursen, *F.E.B.S. Letters*, 1973, **33**, 135.
[671] A. Schellenberger, H. Jeschkeit, H. Graubaum, C. Mech, and G. Sternkopf, *Z. Chem.*, 1972, **12**, 62.
[672] E. Wachter, W. Machleidt, H. Hefner, and J. Otto, *F.E.B.S. Letters*, 1973, **35**, 97.

repetitive yield of about 92%. A second paper [673] from this group reports on the attachment of cytochrome *c* from *Candida krusei* to this support and its degradation over 35 residues with a repetitive yield of 95%. Clearly this will be a very useful method for the attachment of larger molecules.

Identification of Hydantoins.—Whether one uses liquid- or solid-phase, manual or automatic methods, identification of the released thiohydantoins still has to be made. Although systems exist for the use of t.l.c., g.l.c., h.p.l.c., mass spectroscopy, and hydrolysis followed by amino-acid analysis, there is still no method capable of resolving all the derivatives in a single run.

A slightly modified solvent system for t.l.c. resolution of PTH derivatives has been described [674] for use on aluminium-backed silica plates. Apparently this solvent is particularly useful for identification of PTH-CMCys. Another solvent system (xylene–95% ethanol–acetic acid, 50 : 50 : 0.5) is recommended [675] for resolution of PTH-His and PTH-Arg, which are generally identified by specific spot-tests on the t.l.c. plates. Micro-t.l.c. methods have also been published. One of these [676] uses polyamide layers similar to those used for identification of dansyl-amino-acids but with an internal fluorescent indicator. Sensitivity is claimed to be less than 0.5 nmol but only 16 PTH-amino-acids are resolved. Similar sensitivity is claimed [677] for a method which uses silica-gel layers and where detection is by spraying with starch–iodine. A mixture of ninhydrin, collidine, and ethanol has also been used for detection.[678] This staining mixture dehydrates the PTH's to form red hydrindantins. Finally, four new solvent systems have been introduced which permit the separation of all the PTH-amino-acids.[679] The plates are stained with t-butyl hypochlorite (1% solution in dichloroethane containing 10% acetic acid) followed by spraying with equal volumes of 0.05M-KI and a saturated solution of *o*-tolidine in 2% acetic acid. This permits detection at levels down to 1 nmol.

A system has been described [680] for separating both the MTH and PTH derivatives by g.l.c. on capillary columns. The glass columns used have an internal diameter of 0.13 mm and are capable of resolving most thiohydantoins in a single run.

High-pressure liquid chromatography (h.p.l.c.) is a technique likely to become much more common in biochemical laboratories. The advantages of applying the method to thiohydantoin separation are sensitivity (as little as 10 pmol can be detected), speed, and the fact that potentially all the amino-acids can be identified in a single run. The main limitation at

[673] W. Machleidt, E. Wachter, M. Scheulen, and J. Otto, *F.E.B.S. Letters*, 1973, **37**, 217.
[674] A. S. Inglis and P. W. Nicholls, *J. Chromatog.*, 1973, **79**, 344.
[675] T. Inagami, *Analyt. Biochem.*, 1973, **52**, 318.
[676] M. R. Summers, G. W. Smythers, and S. Oroszlan, *Analyt. Biochem.*, 1973, **53**, 624.
[677] M. Salal and J. L. Bernard, *J. Chromatog.*, 1973, **80**, 140.
[678] W. Schäfer and E. Bauer, *Z. physiol. Chem.*, 1973, **354**, 1307.
[679] T. Suzuki, K. Komatsu, and K. Tuzimura, *J. Chromatog.*, 1973, **80**, 199.
[680] J. Eyem and J. Sjöquist, *Analyt. Biochem.*, 1973, **52**, 255.

present is an inadequate resolution but hopefully this will only be a temporary problem. This first communication [681] reports on a partial separation of the PTH derivatives using ~1 μg (~5 nmol) of each and with a run-time of 26 min. Selected steps from a sequencer run on myoglobin are also shown. Better resolution has been obtained using a two-column system.[682]

Mass Spectrometry.—Following their observation (see last year's Report, p. 50) that limited permethylation reaction-time cuts down salt formation in histidyl peptides, Morris and co-workers [683] have applied a similar technique to peptides containing carboxymethylcysteine, methionine, and arginine. In all cases interpretable mass spectra were obtained and it now appears that all of the problematical amino-acids can be dealt with by this technique. For arginyl peptides conversion into pyrimidylornithine is still, of course, necessary and these authors include new, milder conditions for this reaction which are claimed to give very few side-reactions.

Glycopeptides.—This topic has already been comprehensively reviewed this year [684] and only some more recent references will be included here. A sensitive radiochemical method for monitoring glycopeptides in chromatographic eluents has been developed.[685] The method, based on hydrolysis of the glycopeptide followed by reduction with NaB^3H_4, is claimed to be up to 100 times more sensitive than conventional colorimetric methods.

An alternative is thin-layer electrophoresis followed by the application of stains for specific types of carbohydrate.[686] While providing a useful mapping method for small quantities of glycoprotein, the technique is limited in not being preparative. Identification of released carbohydrate units by g.l.c. has also been usefully discussed.[687]

5 Partial Sequences

Even today the elucidation of the complete sequence of a protein is a relatively rare event, especially if the protein is unrelated to other proteins of known sequence. On the other hand, each year sees the publication of an enormous number of sequences of various parts of proteins. These partial sequences are determined for a variety of purposes, although the main one is generally connected with the definition of sequences around particular types of amino-acid residues. However, the increasing availability of automatic sequencers has led to the determination of amino-

[681] C. L. Zimmerman, J. Pisano, and E. Appella, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 1220.
[682] G. Frank and W. Strubert, *Chromatographia*, 1973, **6**, 522.
[683] H. R. Morris, R. J. Dickinson, and D. H. Williams, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 247.
[684] R. G. Spiro, *Adv. Protein Chem.*, 1973, **27**, 350.
[685] E. D. Lehman, B. G. Hudson, and K. E. Ebner, *J. Chromatog.*, 1973, **76**, 345.
[686] E. Moczar, *J. Chromatog.*, 1973, **76**, 417.
[687] R. Chambers, J. R. Clamp, R. Bayard, and J. Montreuil, *Biochimie*, 1973, **55**, 1195.

Table 4 *Amino-terminal sequences* [a]

Protein	*Amino-terminal sequence*	*Ref.*
Rabbit anti-hapten antibody heavy chains	ISLEES SVEES	688
Mouse IgM	EVQLEESGPELVKP	689
Rabbit anti-*p*-azobenzenearsonate light chain	AVVLTETASPVSAPVGGTYTI	690
Anti-Pr cold agglutinin kappa chain	DIVMTQSPDSLAVSLGETATID	691
Bence-Jones protein (Ta)	SYALTQPPSVSVSPGQTA	692
Porcine trypsin	A chain IVGGYTCAANS(I,V)PYQVSLNS GSHFCGGSLI B chain SSGSSYPSLL	693
Bovine pepsin	VSQEPLQNY	694
Rabbit plasminogen	EPLDDYVNTQGA	695
Haemoglobin (*Vipera aspis*)	VLSEDDKNRVRTSVGKNPELPGEYGSETLTRMFA	696
Haemoglobin (Badger)	GLSDGEWQLVLDVWGKVEADLAGHGQEVLIR LFKGHPETLEKFDKFKHLKSEDEM	697
Haemoglobin (Trout)	α-VXSA β-VDXTD	698
Myoglobin (Woolly Monkey)	Major GLSDGEWQ Minor FKGLSDGEWQ	699
Myoglobin (Marmoset)	Major GLSDGEWQ Minor FKGLSDGEWQ	699

[688] B. D. Seon, A. O. Roholt, and D. Pressmann, *Imm. Chem.*, 1973, **10**, 495.
[689] E. A. Robinson, E. Appella, and K. R. McIntire, *J. Biol. Chem.*, 1973, **248**, 7112.
[690] R. G. Mage, E. Appella, G. O. Young-Cooper, J. B. Winfield, J. H. Pincus, and A. Chessi, *J. Immunol.*, 1973, **110**, 227.
[691] A. C. Wang, H. H. Fridenberg, J. V. Wells, and D. Roelcke, *Nature New Biol.*, 1973, **243**, 126.
[692] C. Tonnello, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 1112.
[693] J. E. Walker, N. Zylber, and B. Keil, *F.E.B.S. Letters*, 1973, **32**, 223.
[694] M. Harboe and B. Foltmann, *F.E.B.S. Letters*, 1973, **34**, 311.
[695] F. J. Castellino, G. E. Siefring, J. M. Sodetz, and R. K. Bretthauer, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 845.
[696] F. Le Goffic and N. Moreau, *F.E.B.S. Letters*, 1973, **29**, 289.
[697] D. Tetaert, K. Han, M. Dantrevaux, S. Ducastaing, I. Hombrados, and E. Neuzil, *F.E.B.S. Letters*, 1973, **29**, 38.
[698] D. Barra, F. Bossa, J. Bonaventura, and M. Brunori, *F.E.B.S. Letters*, 1973, **35**, 151.
[699] A. E. Romero-Herrera and H. Lehmann, *F.E.B.S. Letters*, 1973, **31**, 175.

Table 4 (*cont.*)

Protein	*Amino-terminal sequence*	*Ref.*
Alkaline phosphatase (*E. coli*)	RTPEMPVLENRAAEGDITAP	700
Y-Crystallin (Haddock lens)	GKITFYE	701
Hirudin	VV	702
Adrenocorticotrophin (Human)	SYSMEHFRWGKPVGKKRRPVK	703
Neurophysin (Human)	AAPDLDVRKCLPCGPG	704
K-Caseinoglycopeptide (Sheep)	MAIPPKKDQDKTEIPAINTIASAEPT VHSTPTPEAVVNAV	705
Proglucagon (peptide B7)	HSQGTFTSDYSKYLDSRRAQDFV(QWL) MNTKRNNKNIA	706
Co-lipase (Porcine)	GIIINLDEGELC-LNCAQSK	707
Amyloid P-component (Human)	HADLXTKVFVFXXSEVVXSVVSI	708
Coat protein (ϕX174)	MFQTFISRH	709
Serum prealbumin (Rhesus Monkey)	GPTG$\binom{V}{I}$DESKC	710
Lysozyme (Black Swan)	C-type KVYERCELAAAMKRLGLDKY G-type RTDCYGNVNRIDTTGASCKT	727
Alcohol dehydrogenase (*B. sterothermophilus*)	MKAAVVEQFKKPLLQVKEVEK PKISYGEVLVRIKACGVCHTVLG	711

[a] The single-letter notation is that recommended by the IUPAC–IUB Commission; see these Reports, Vol. 5, p. 487.

[700] J. Bridgen and D. Secher, *F.E.B.S. Letters*, 1973, **29**, 55.
[701] L. R. Croft, *Biochim. Biophys. Acta*, 1973, **295**, 174.
[702] L. Graf, A. Patthy, E. B. Barabas, and D. Bagdy, *Biochim. Biophys. Acta*, 1973, **310**, 416.
[703] H. P. J. Bennett, P. J. Lowry, and C. McMartin, *Biochem. J.*, 1973, **133**, 11.
[704] I. Foss, K. Sletten, and O. Trygstad, *F.E.B.S. Letters*, 1973, **30**, 173.
[705] K. G. Welinder, *F.E.B.S. Letters*, 1973, **30**, 243.
[706] H. S. Tager and D. F. Steiner, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2321.
[707] M. F. Maylie, M. Charles, M. Astier, and P. Desnuelle, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 291.
[708] M. Skinner and A. S. Cohen, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 732.
[709] G. M. Air and J. Bridgen, *Nature New Biol.*, 1973, **241**, 40.
[710] P. Van Jaarsveld, W. T. Branch, J. Robbins, F. J. Morgan, Y. Kanda, and R. E. Canfield, *J. Biol. Chem.*, 1973, **248**, 7898.
[711] J. Bridgen, E. Kolb, and J. I. Harris, *F.E.B.S. Letters*, 1973, **33**, 1.

terminal sequences of a large number of proteins. In this section these partial sequences are considered more or less in relation to the general region of the protein from which they are derived rather than the particular purpose for which the sequences were determined.

Amino-terminal Sequences.—A summary of the proteins which have had their amino-terminal sequences determined recently, together with the actual sequences obtained, is shown in Table 4. In some cases the reported sequences are too extensive to be included in the table and are either described in the text below or are not reproduced here.

Immunoglobulins. It is not surprising that a major proportion of the amino-terminal sequences reported belongs to immunoglobulin chains since the variable domains of these are located at the amino-termini of the heavy and light chains. Earlier studies on lambda chains from plasmacytomas of BALB/c mice indicated a much simpler pattern of variability than human lambda or kappa chains, and the BALB/c mice appear to express only one lambda variable-region subgroup, which has been taken as evidence for a single germ-line lambda variable-region gene. A recent report [712] describes more detailed sequencing on the variable regions of two indistinguishable lambda chains and confirms the argument that they are identical (Figure 7). Therefore, of the 12 mouse lambda chains, eight are identical whilst the other four only seem to differ in the hypervariable region.

Specifically purified anti-*p*-azobenzoate antibody heavy chains from several rabbits have been sequenced at the amino-terminus and found to be virtually identical.[688] This contrasts with the high degree of variability in amino-terminal sequences of the light chains. Another study on 12 rabbit antibody light chains from bacterial-polysaccharide-induced antibody has shown that rabbit, mouse, and human kappa chains are homologous.[713] Like the mouse and human chains, rabbit kappa chains could be sub-grouped and a minimum of six subgroups could be identified on the results obtained. The results do not support a suggestion that there are species-specific amino-acid residues. A very interesting observation is that the degree of homology appears to be a function of the breeding relationship of individual rabbits, which would imply inheritance of structural V-region genes for the synthesis of specific anti-polysaccharide antibodies.

Relatively little sequence information has been obtained on IgA, which is the predominant immunoglobulin in external secretion systems and is therefore associated with the local defence mechanism in the body. A 65-residue fragment has now been isolated from human IgA myeloma protein and sequenced.[714] The fragment was isolated by specific cleavage at an Asn-Gly bond in the α-heavy chain with hydroxylamine. The cleavage was reasonably efficient and the final yield of the fragment was over about

[712] I. M. Cesari and M. Weigert, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2112.
[713] D. G. Braun and J. C. Jaton, *Imm. Chem.*, 1973, **10**, 387.
[714] T. Shinoda, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 1246.

J558, xS104, and MOPC-104E	**1**	**5**
	Pca-Ala-Val-Val-Thr-Glx-Glx-Ser-Ala-	
S178		

	10	**15**	**20**	**25**	
	Leu-Thr-Thr-Ser-Pro-Gly-Glx-Thr-Val-Thr-Leu-Thr-Cys-Arg-Ser-Ser-				
					Asn-

J558, xS104, and MOPC-104E		**30**	**35**	**40**	**45**
	Thr-Gly-Ala-Val-Thr-Thr-Ser-Asx-Tyr-Ala-Asx-Trp-Val-Glx-Glx-Lys-Pro-Asp-His-Leu-Phe-Thr-Gly-Leu-Ile-				
S178					

J558, xS104, and MOPC-104E		**55**	**60**	**65**	**70**
	Gly-Gly-Thr-Asx-Asx-Arg-Ala-Pro-Gly-Val-Pro-Ala-Arg-Phe-Ser-Gly-Ser-Leu-Ile-Gly-Asx-Lys-Ala-Ala-Leu-				
S178	—Asn—				

J558, xS104, and MOPC-104E		**80**	**85**	**90**	**95**
	Thr-Ile-Thr-Gly-Ala-Glx-Thr-Glx-Asx-Glx-Ala-Ile-Tyr-Phe-Cys-Ala-Leu-Trp-Tyr-Ser-Asx-His-Trp-Val-Phe-				
S178					Arg—

J558, xS104, and MOPC-104E	**105**	**110**
	Gly-Gly-Gly-Thr-Lys-Leu-Thr-Val-Leu-Gly-Glx-Pro-Lys	
S178		

Figure 7 *The tentative sequence of the variable region of mouse lambda chains*
(Reproduced by permission from *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2112)

20%. It could prove to be of considerable use in the generation of large fragments with other proteins.

The allotypy of the variable region in rabbit heavy chains is a puzzling phenomenon, and resolution of some of the paradoxes implicit in it might provide information on the genetic mechanism underlying antibody diversity. Unfortunately, the presence of a blocked amino-terminal residue in rabbit H chains has made them difficult to study. Recently, however, evidence has been obtained that group *b* allotype markers correlate with certain mixtures of variable region sequences in rabbit light chains.[715] The approach used was to carry out automatic sequence analyses on IgG kappa chains from rabbits homozygous for each of the four group *b* allotypes and see whether the profiles derived from the variable region correlated with the serologically determined allotypic specificities. Such a correlation was obtained and if confirmed would permit more detailed investigation of the phenomenon, since light chains are much more amenable to analysis. Light chains of the *b5* allotype have been obtained from rabbit anti-*p*-azobenzene-arsonate antibody.[690] The first 20 residues were sequenced from light chains and gave a unique sequence when third-course antibody was used. However, limited heterogeneity was evident when second- and fourth-course antibodies were used.

The sequence of the variable region of an anti-Pr cold agglutinin kappa chain has been found to have at least a 34% difference with known subgroups, and was assigned to a new subgroup V_κIV.[691] Cold agglutinins with anti-Pr specificities have variable-region subgroups which clearly differ from those in cold agglutinins with anti-I specificity. As these specificities correspond to chemical differences in the red cell membrane antigens, this feature illustrates a direct correlation between the antibody specificity and the structure of the variable regions of the light and heavy chain.

The kappa-type light chain from mouse myeloma MPC11 has been found to contain an extra 12 residues at its amino-terminus.[716] The first six of these are identical with residues 13—18 (which is the start of the typical kappa-chain sequence) and it suggests that the MPC11 gene was produced by a recent breakage and reunion of two kappa variable-region genes. This hypothesis is preferred to the possibility that the extra residues correspond to the precursor sequence since the Leu-Leu-Leu sequence, which is typical of the MOPC 321 precursor, is not present. It was also observed that the MPC11 protein had some very labile bonds in the region of the extra sequence. These bonds seemed to be very easily broken at pH 4.8—5.5, conditions which do not usually cause peptide bond cleavage.

The genetic markers which are present in various domains of the immunoglobulins have proved extremely useful in studies on the synthesis and inheritance of these molecules. However, only one serologically detectable

[715] M. D. Waterfield, J. E. Morris, L. E. Hood, and C. W. Todd, *J. Immunol.*, 1973, **110**, 227.
[716] G. P. Smith, *Science*, 1973, **181**, 941.

allotypic marker exists for variable region (*cf.* above). Studies on the b9 light chains from rabbit antibodies to streptococcal carbohydrate [717] have now revealed that glutamic acid is present at position N(16) of the light chains of the V_KIII length. This amino-acid has not been found at N(16) in any other light chains, suggesting that it could serve as a marker for genetic studies on the V_L region.

The use of manual sequencing by the Edman method has not been completely superseded by the automatic procedure, as evidenced by a report on the sequence of the first 18 residues of the human lambda Bence-Jones protein.[692] Only 200 nmol of protein were used even though the protein contains 30% of hydroxylated amino-acids in this region. The sequence shows that the protein belongs to group IV.

Proteinases. Porcine α-trypsin contains two polypeptide chains joined together by disulphide bridges. The two chains (A and B) have been isolated and partially sequenced on the automatic sequencer.[693] The A chain sequence contains only three differences from the amino-terminal sequence of bovine trypsin whilst the B chain is very similar to the *C*-terminus of bovine trypsin. The first 62 residues of chymosin have also been sequenced and found to have about 50% homology with bovine and porcine pepsin (Figure 8).[718]

The two major forms of rabbit plasminogen which can be isolated from rabbit plasma appear to be very similar in composition, size, and activation properties. These proteins have now been sequenced through the first twelve residues of the amino-terminus and found to be identical.[695] However, the carbohydrate contents of the two proteins are different and are probably the cause of their differences in chromatographic behaviour.

The amino-terminal sequences of prothrombin and the activation intermediates have been determined and used to demonstrate that the activation proceeds by the sequential cleavage of the amino-terminus of prothrombin and the intermediates.[719] This is illustrated in Scheme 5. One interesting feature is that the total sequence of the prothrombin molecule can be deduced from the sequences of the intermediates.

Although the process of pepsinogen activation has been intensively studied the exact site involved in the conversion into pepsin has not been pin-pointed. Several reports which appeared recently have been concerned with this problem.[694, 720] The most comprehensive approach used has been the elucidation of the sequence of the first 119 residues of swine pepsinogen, which has permitted the unambiguous identification of the activation site.[721] Previous studies had suggested [722] that the activation

[717] A. L. Thumberg, H. Lackland, and T. J. Kindt, *J. Immunol.*, 1973, **111**, 1755.
[718] V. B. Pedersen and B. Foltmann, *F.E.B.S. Letters*, 1973, **35**, 250.
[719] C. M. Heldebrant, C. Noyes, H. I. Kingdon, and K. G. Mann, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 155.
[720] V. B. Pedersen and B. Foltmann, *F.E.B.S. Letters*, 1973, **35**, 255.
[721] V. M. Stepanov, L. A. Baratova, I. B. Pugacheva, L. P. Belyanova, L. P. Revina, and E. A. Timokhina, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 1164.
[722] E. B. Ong and G. E. Perlmann, *J. Biol. Chem.*, 1968, **243**, 6104.

	10	20
Cymosin	Gly-Glu-Val-Ala-Ser -Val -Pro-Leu-Thr-Asn-Tyr-Leu-Asp-Ser -Gln-Tyr-Phe-Gly-Lys-Ile-Tyr-Leu-Gly-Thr-Pro-
Bovine pepsin	Val-Ser -Gln-Glu-Pro-Leu-Gln-Asn-Tyr(Leu.Asx-Thr-Glx-Tyr.Phe.Gly.Thr.Ile.Tyr.Ile .Gly.Thr.Pro.
Porcine pepsin	Ile -Gly-Asp-Glu-Pro-Leu-Glu-Asn-Tyr-Leu-Asn-Thr-Glu-Tyr-Phe.Gly-Thr-Ile-Gly-Ile -Gly-Thr-Pro-
Common	Pro Leu Asn Tyr Leu Tyr Phe Gly Ile Gly Thr Pro

	30	40	50
Cymosin	Pro-Gln-Glu-Phe-Thr-Val-Leu-Phe-Asp-Thr-Gly-Ser-Ser-Asp-Phe-Trp-Val-Pro-Ser-Ile -Tyr-Cys-Lys-Ser-Asn-
Borine pepsin	Ala.Glx.Asx.Phe.Thr)Val-Ile -Phe-Asp-Thr-Gly-Ser-Ser-Asn.Leu-Trp-Val-Pro-Ser.Ile -Tyr-Cys.Ser -Ser-Glu-
Porcine pepsin	Ala-Gln-Asp-Phe.Thr-Val-Ile -Phe.Asp-Thr-Gly-Ser-Ser-Asn.Leu-Trp-Val-Pro-Ser.Val-Tyr-Cys-Ser -Ser-Leu-
Common	Phe Thr Val Phe Asp Thr Gly Ser Ser Trp Val Pro Ser Tyr Cys Ser

	60
Cymosin	Ala-Cys-Lys-Asn-His-Gln-Arg-Phe-Asp-Pro-Arg/
Bovine pepsin	Ala-Cys-Thr-Asn-His-Asn-Arg/
Porcine pepsin	Ala-Cys-Ser -Asp-His-Asn-Gln-Phe-Asn-Pro-Asp/
Common	Ala Cys His

Figure 8 *Comparison between the* N*-terminal amino-acid sequences of gastric proteases*
(Reproduced by permission from *F.E.B.S. Letters*, 1973, **35**, 273)

occurs through the cleavage of a Glu-Ile bond, although there was independent evidence that the Ile residue at the amino-terminus of pepsin was preceded by Leu and not Glu. In the detailed study described recently [721] it was found that residues 40 and 41 are Glu-Ala, and not Ala-Glu as suggested previously. In addition, the sequence that follows Glu_{41} is

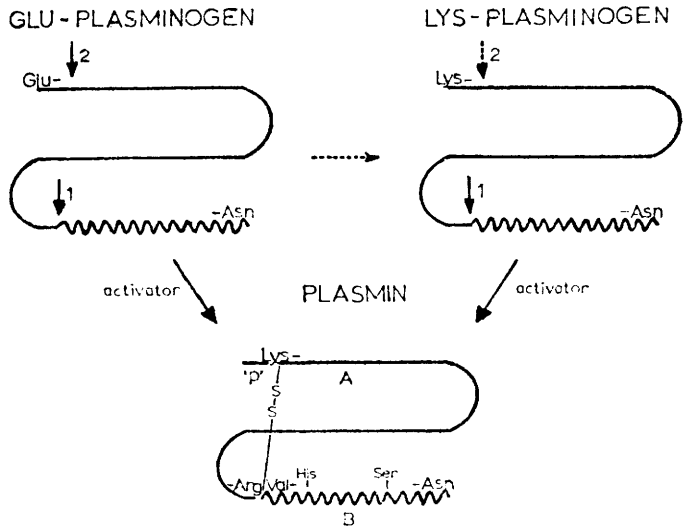


Figure 9 *Molecular events which take place during the activation of Glu- and Lys-plasminogens to give Lys-plasmin. The heavy (A) and light (B) chains, a peptide ('p'), the active-centre histidine (His) and serine (Ser) residues, and the hinge disulphide bond (—S—S—) connecting the chains are noted in the figure*
(Reproduced by permission from *J. Biol. Chem.*, 1973, **248**, 7242)

Ala-Ala-Leu, with Ile at position 45. This confirms the view that activation occurs through a Leu_{44}-Ile_{45} split in pepsinogen. The existence of the Ala-Ala sequence in this region is consistent with the activation of pepsinogen by neutral proteinases.

The amino-terminal sequences of several plasminogens and the corresponding plasmin-derived heavy (A) and light (B) chains have been determined and compared.[723] Residue 9 of the plasminogen sequences has not been identified but another study indicates that it is threonine.[695] As a result of this and other studies it has been considered necessary to re-evaluate the activation of plasminogen to plasmin. In the revised scheme (Figure 9) two specific peptide-bond cleavages are required instead of the one previously reported. The extra split was probably missed because the peptide produced is released from the plasmin.

[723] K. C. Robbins, P. Bernabe, L. Arzadon, and L. Summaria, *J. Biol. Chem.*, 1973, **248**, 7242.

Other Proteins. The amino-terminal sequence of alkaline phosphatase from *E. coli* has been investigated in order to check reports of molecular heterogeneity in the apparently identical subunits.[701] Only a single sequence was obtained, and this in turn suggests that the subunits of the ochre mutants of this enzyme which have been described have been substantially degraded during isolation. The amino-terminal cyanogen bromide fragment from two isozymes of the same enzyme has been sequenced and gives results which are consistent with that obtained in the above study.[724] One isozyme lacks the arginine residue at the amino-terminus.

A careful study of crystalline glucagon has shown that such preparations contain various peptides with glucagon activity as contaminants.[707] One highly basic peptide was isolated and subjected to sequence analysis and found to contain the normal glucagon sequence together with an eight-residue extension at its carboxyl-terminus. This appears to be a fragment of proglucagon and not the prohormone itself because there is strong evidence that the latter is much larger. It is not known whether the proglucagon fragment is bovine or porcine since a mixed preparation was used. Thus it appears that many, if not most, hormones are synthesized as larger precursors and pruned down to the correct size when required. The reason for this is not known but it could reflect special characteristics of the protein synthesizing/secreting mechanism for extracellular proteins.

A useful application of the automatic sequencer to toxinology has been reported in connection with a study of some snake neurotoxins.[725] Venom falsification can be an important problem in toxinology and one clear example exists of the assumption that an Elapidae-type toxin exists in a Viperidae venom. The conclusion was based on pharmacological and immunological results. But compositional and sequence studies showed it to have the structure of the Elapid *Naja nigricollis* toxin. It appears that artifactual mixtures of these venoms can arise owing to the overlap in geographical distribution. Since the task of collecting venom is, understandably, often delegated to non-experts it is comforting that a rapid check on identity can be made in the laboratory.

Studies on the evolutionary relationship between lysozyme and α-lactalbumin have been extended to the determination of the amino-terminal sequence of grey kangaroo α-lactalbumin I.[726] Only 7 of the first 22 residues are identical with corresponding regions in the bovine, human, or guinea-pig proteins but the 23—42 sequence is very similar. It is also quite similar to human and chicken lysozyme. It appears from the comparative studies that the deletion of two residues was a key event in the conversion of a

[724] P. M. Kelley, P. A. Neumann, K. Schriefer, F. Cancedda, M. J. Schlesinger, and R. A. Bradshaw, *Biochemistry*, 1973, **12**, 3499.
[725] C. Kopeyan, J. van Rietschoten, E. Martinez, H. Rochat, F. Miranda, and S. Lissitsky, *European J. Biochem.*, 1973, **35**, 244.
[726] K. Brew, H. M. Steinman, and R. L. Hill, *J. Biol. Chem.*, 1973, **248**, 4739.

Table 5 *Sequences around reactive residues*

Protein	*Type of residue*	*Sequence*	*Ref.*
Shrimp trypsin	Serine	DSCQGSDGGPLACSNTGSTY	727
Yeast enolase A	Cysteine	IGLDCSASEFFK	728
Yeast enolase B	Cysteine	VE(IEL)(DCSA)(SE)FYK	
Bovine kidney rhodanese	Cysteine	KVDLTKPLIATCR	729
Oestradiol dehydrogenase	Cysteine	ALACPPGSLETLQLDVR	730
Horseradish peroxidase	Histidine Cystine	LHFHDCFVNGCDASILLDNTTSFR	731
Isoleucyl-tRNA synthetase	Cysteine	.CPR .AVGCAK .VAEHAEICGR .CWHYTEDVGK .MEHSYPCCWR .CVSDVAGDGEK	732
Antheraea pernyi protease	Serine	DACNGDSGGPVQNAGR	733
Aspartate aminotransferase	Lysine	YFVSEGFGLFCAQSFSKNFGLYNER	734
Flavocytochrome c_{552} (*Chromatium*)	Cysteine	YTCY	735
IgA1 myeloma globulin	Cystine	—	736
Acetylcholinesterase (Eel)	Serine	GGESSEGAAG	737
Rabbit anti-*p*-azophenyltrimethylammonium antibody	Tyrosine	CARDPY (H-chain) SSEY (L-chain)	738
Human plasminogen	Histidine	HFCGGTLISPEWVLSAAHCL	739
Hydroxynicotine oxidase	Histidine	SGGNNPDHY(QP)A	740
Porcine IgG	Cystine	TAPSVYPLAPCGR	741

Rabbit anti-*p*-azobenzoate antibody	Tyrosine	YTGY	742
Human glomerular basement membrane antigen	Hydroxylysine	hydroxylys-GEDG	743
Epidermal growth factor	Cystine	YSGDRC-CN YC-CT NSYPGC-PCV	744
Yeast carboxypeptidase	Serine	HIAGESYAHGYIPVF	745

[727] N. Arnheim, A. Hindenberg, G. S. Begg, and F. J. Morgan, *J. Biol. Chem.*, 1973, **248**, 8036.
[728] S. K. Oh, J. Travis, and J. M. Brewer, *Biochim. Biophys. Acta*, 1973, **310**, 421.
[729] F. Bossa, C. Cannella, G. Federici, D. Barra, and L. Pecci, *F.E.B.S. Letters*, 1973, **29**, 171.
[730] J. Nicolas and J. I. Harris, *F.E.B.S. Letters*, 1973, **29**, 173.
[731] K. G. Welinder, *F.E.B.S. Letters*, 1973, **30**, 243.
[732] M. R. Kula, *F.E.B.S. Letters*, 1973, **35**, 299.
[732a] G. L. E. Koch, Y. Boulanger, and B. S. Hartley, reported by B. S. Hartley, *Int. Congress of Biochemistry, Stockholm*, 1973, **9**, 199.
[733] K. J. Kramer, R. L. Felsted, and J. H. Law, *J. Biol. Chem.*, 1973, **248**, 3021.
[734] Y. Morino and M. Okamoto, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 1061.
[735] W. C. Kenney, D. Edmondson, R. Seng, and T. P. Singer, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 434.
[736] V. Moore and F. W. Putnam, *Biochemistry*, 1973, **12**, 2361.
[737] N. K. Schaffer, H. O. Michel, and A. F. Bridges, *Biochemistry*, 1973, **12**, 2946.
[738] Y. Takeda, B. Friedensen, O. A. Roholt, and D. Pressman, *European J. Biochem.*, 1973, **3**, 461.
[739] K. C. Robbins, P. Bernabe, L. Arzadon, and L. Summaria, *J. Biol. Chem.*, 1973, **248**, 1631.
[740] M. Bruhmullee and K. Decker, *European J. Biochem.*, 1973, **37**, 256.
[741] J. J. Metzger and P. Pesy, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 253.
[742] O. A. Roholt, B. Friedensen, F. Radzimski, and D. Pressman, *J. Immunol.*, 1973, **111**, 1367.
[743] P. M. Mahieu, P. H. Lambert, and G. R. Maghium-Rogister, *European J. Biochem.*, 1973, **40**, 599.
[744] C. R. Savage, J. H. Hash, and S. Cohen, *J. Biol. Chem.*, 1973, **248**, 7669.
[745] R. Hayashi, S. Moore, and W. H. Stein, *J. Biol. Chem.*, 1973, **248**, 8366.

lysozyme-like molecule into the α-lactalbumin type. There are two types of such deletions and the possibility that they occurred after divergence between mammals and marsupials is raised. If so the 21-22 deletion would serve as a marker for marsupials and the 14-15 one for non-marsupial mammals. The available evidence also supports the view that marsupials are more evolutionarily remote from the cow, the guinea pig, and man than the latter are from one another.

Partial sequencing at the amino-terminus of the two types of lysozyme found in the eggs of the black swan has been carried out and clearly shows them to be quite different.[727] One of them is very similar to the lysozyme of the chicken and many other birds and the other resembles the lysozyme found in the egg white of the Embden goose. The differences between the two types of lysozyme are sufficient to show that they are from distinct genetic loci. It is believed that the genetic loci for both species are present in all species, but the phenotype is controlled by a species-specific regulatory mechanism.

Sequences around Reactive Residues.—The amino-acid residues which contain reactive groups in their side-chains have been widely studied because they can yield useful, albeit sometimes limited, information about protein structure and function. This is particularly true when the residues concerned are at or close to the active centres of the proteins. Furthermore, the tendency for such residues such as cysteine to be conserved during evolutionary changes of a protein means that comparisons of the sequences around such residues can prove useful in evolutionary studies. In this section consideration is given to sequences around the reactive residues irrespective of whether or not they originate from the active centres of the proteins concerned. We shall only consider those cases in which the sequence around the reactive residue has been published. Other examples may be found in the section on the chemical modification of proteins. The information relevant to this section is summarized in Table 5.

Cysteine Sequences. In contrast to previous studies it has now been shown that yeast enolase contains one cysteine residue per subunit.[728] These residues are very unreactive towards sulphydryl reagents but they can be made to react after dissociation with urea. The possibility that the cysteines are the result of a recent mutation was excluded. Previous studies also indicated that the two chains of yeast enolase are identical although anti-co-operative binding of Mg^{2+} has been reported. The finding of two different cysteine sequences does not contradict this view since it was shown that they come from different forms of the enzyme, namely enolase A and B, respectively. The sequences obtained were:

A Val-Glu-(Ile, Glu, Leu) (Asp, Cys, Ser, Ala) (Ser, Glu) Phe-Tyr-Lys

B Ile-Gly-Leu-Asp-Cys-Ser-Ala-Ser-Glu-Phe-Phe-Lys

The significant differences in sequence between the two peptides contradict the suggestion that the B form is a deamidation product of the A form.

Bovine kidney rhodanese contains one cysteine residue which is essential for enzymic activity. This residue is rapidly carboxymethylated in the native enzyme and could be involved in the formation of a persulphide group at the active site during catalysis. The tryptic peptide containing this residue has been isolated and sequenced.[729] There appear to be some differences with the analogous peptide from the liver enzyme, indicating that the enzyme might be organ-specific.

The sequence around the reactive cysteine residue in oestradiol dehydrogenase from human placenta has been determined from the corresponding tryptic peptide.[730] The fact that oestradiol does not protect the thiol group against reaction with *N*-ethylmaleimide argues against its involvement in the substrate binding site, but by the same argument the fact that NADP does protect suggests it may be near the coenzyme binding site.

The cysteine peptides from tryptic digests of isoleucyl-tRNA synthetase from *E. coli* have been isolated and sequenced.[731] The available sequences account for only eight of the cysteine residues calculated from the amino-acid composition, suggesting that the polypeptide chain contains large segments of repeated sequence. Although this is indirect evidence it supports the view that many aminoacyl-tRNA synthetases, and probably other types of enzymes as well, have elongated by the addition of sequence repetitions.[732a] Similar evidence has also been obtained with phosphofructokinase from rabbit muscle.[746]

The reactivity of the cysteine residues of aspartate aminotransferase has been studied as a probe for the topological distribution of these residues,[747] following the elucidation of the complete primary structure of the enzyme. Two of the five residues, Cys-45 and Cys-82, are very reactive and are the most exposed in the native enzyme. Alkylation of these two residues has no effect on the activity of the enzyme. Cys-390 has been identified as the residue which undergoes selective alkylation and is the syn-catalytically modified residue. The other two residues Cys-252 and Cys-191 are deeply buried in the protein since they only react upon denaturation of the protein.

The involvement of cysteine in the attachment of flavin to cytochrome c_{552} from *Chromatium* has been confirmed by the isolation of the peptide concerned from peptic and tryptic/chymotryptic digests of the protein.[735] The peptide which has the sequence Tyr-Thr-Cys-Tyr is joined to the 8α carbon of the flavin through the cysteine. There is some doubt about the type of linkage involved but the available evidence favours a thiohemiacetal rather than the thioether, as exists in monamine oxidase.

The disulphide bridges of mouse epidermal growth factor have been placed in the complete amino-acid sequence from studies on the thermolysin peptides (Figure 10).[744] The intact disulphide bridges are necessary for

[746] C. J. Coffee, R. P. Aaronson, and C. Frieden, *J. Biol. Chem.*, 1973, **248**, 1381.
[747] O. L. Polyanovsky, V. V. Nosikov, S. M. Deyev, A. E. Braunstein, E. V. Grishim, and Yu. A. Ovchinnikov, *F.E.B.S. Letters*, 1973, **35**, 322.

biological activity and they can be reversibly reduced with mercaptoethanol. Removal of the C-terminal pentapeptide does not affect the activity of the factor nor does it seem to be important for re-folding of the reduced molecule since its removal has no effect on the reversible reduction.

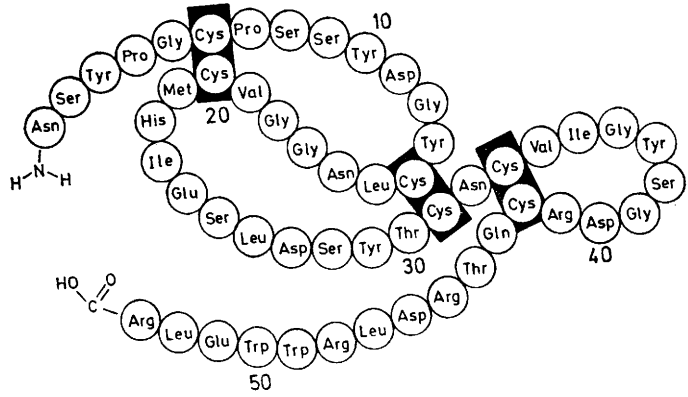


Figure 10 *The amino-acid sequence of EGF, with placement of disulphide bonds*
(Reproduced by permission from *J. Biol. Chem.*, 1973, **248**, 7669)

Serine Sequences. Shrimp trypsin differs from mammalian trypsins in having an acidic isoelectric point, no requirement for calcium for stability, resistance to autodigestion, and it is irreversibly inactivated below pH 5.0. However, the sequence around the reactive serine residue of this enzyme is very similar to that of mammalian trypsins (Figure 11),[727] showing that

Shrimp	Asp-Ser-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Leu-
Dogfish trypsin	Asp-Ser-Cys-Glx-Gly-Asp-Ser-Gly-Gly-Pro-Val-
Pig tryspin	Asp-Ser-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Val-
Cow trypsin	Asp-Ser-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Val-
Cow chymotrypsin	Ser-Ser-Cys-Met-Gly-Asp-Ser-Gly-Gly-Pro-Leu-
Pig elastase	Ser-Gly-Cys-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Leu-

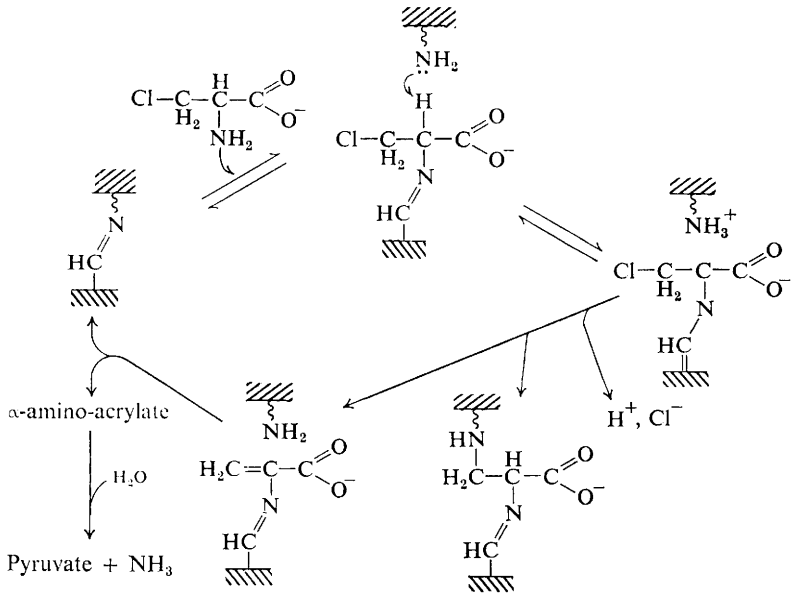
Shrimp	Ala-Cys-Ser-Asn-Thr-Gly-Ser-Thr-Tyr
Dogfish trypsin	Val-Cys
Pig trypsin	Val-Cys-Gly-Gln-Gln-Leu
Cow trypsin	Val-Cys-Ser-Gly-Lys
Cow chymotrypsin	Val-Cys-Lys-Lys-Asn
Pig elastase	His-Cys-Leu-Val-Asn

Figure 11 *Comp rison of active-site peptide from shrimp trypsin with other proteinases*
(Reproduced by permission from *Biochim. Biophys. Acta*, **310**, 137)

the sequences around this residue have been conserved despite the fact that mammals and crustaceans diverged at least 500 million years ago. Another arthropod trypsin-like protease called cocoonase has also been partially

sequenced and found to have an active serine residue which is placed in a sequence which is highly homologous with the corresponding sequence in mammalian trypsins.[733] It is surprising in view of the amount of work that has been carried out on proteases that the above two reports are the first on arthropod enzymes. The reports on the active-site serine residues of eel acetylcholinesterase[737] and yeast carboxypeptidase[745] have also been considered in the section on chemical modification.

Other Residues. It has been shown that the labelling of cytoplasmic aspartate aminotransferase by β-chloro-L-alanine occurs at the ε-amino-group of the lysyl residue involved in the binding of the coenzyme pyridoxal phosphate.[734] The process by which inactivation occurs is shown in Scheme 7. The results support the view that the lysine residue involved



Scheme 7

not only participates in the formation of an aldimine bond with pyridoxal phosphate but also functions as a base to abstract the α-H atom of the bound substrate. C.d. studies on the tryptic peptide from the labelled site suggest that the phosphopyridoxyl chromophore is embedded in a dissymmetric environment unlike in the derivative obtained by borohydride reduction.

Tyrosine residues have been found at the binding sites of several antibody molecules. Use of the paired labelling technique with ^{131}I on anti-*p*-

azophenyltrimethylammonium antibody yielded a peptide with the sequence Cys-Ala-Arg-Asp-Pro-Tyr, which corresponds to positions 97—102 in the heavy chain.[738] Another peptide, Cys-Cys-Gln-Tyr, was assigned to the light chain of the molecule. A peptide containing tyrosine has also been isolated from the active site of a rabbit anti-*p*-azobenzoate antibody.[742] The sequence, which is Tyr-Thr-Gly-Tyr, is from the light chain, and since it did not correspond to any of the known sequences of light chains it was concluded that the sequence came from a hypervariable region. Interestingly, only the *C*-terminal tyrosine is iodinated, emphasizing the specificity of this approach. The iodination of the tyrosine residue does not eliminate hapten binding completely but there is a slight reduction in affinity. This enhances the argument that the tyrosine residue is at the binding site since it implies that there is no gross rearrangement of the antibody molecule during the iodination.

The sequence around the light-chain active-centre histidine residue of human plasminogen has been determined. The histidine residue was labelled with radioactive TLCK and the molecule cleaved with cyanogen bromide. The largest of the fragments contained the labelled histidine residue, which was then found to be at position 18 from the *N*-terminus of the fragment by automatic sequencing. The sequence obtained is homologous with the histidine loop sequence of other serine proteases. Histidine has also been found to be involved in the covalent attachment of flavin to hydroxynicotine oxidase.[740] The peptide concerned was isolated and sequenced.

The glycopeptide carrying the major antigenic determinants of anti-glomerular-basement-membrane autosensitization has been isolated from an enzymatic digest of basement-membrane antigens and sequenced.[743] The disaccharide is attached to the hydroxylysine residue of the sequence Hyl-Gly-Glu-Asp-Gly. The immunodominant antigenic site is formed by the Galactose-Hyl unit, which could explain the cross-reaction with collagen.

6 Complete Sequences

Muscle Proteins.—*Actin.* The first complete amino-acid sequence of a myofibrillar protein, actin from rabbit skeletal muscle, has been published by Elzinga and co-workers.[748] The 374-residue structure has a molecular weight of 41—42 000, which is 5—10% less than that predicted by physicochemical methods. Considerable similarity is found between this structure and partial sequences from trout [749] and amoeba [750] actin, and it appears that the protein is highly conserved throughout nature. It will be of interest to see whether this high degree of homology is maintained in non-muscle actins.

[748] M. Elzinga, J. H. Collins, W. M. Kuehl, and R. S. Adelstein, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2687.
[749] J. Bridgen, *Biochem. J.*, 1971, **126**, 21.
[750] R. R. Weihing and E. D. Korn, *Biochemistry*, 1972, **11**, 1538.

Parvalbumins. Complete sequences have appeared for these proteins from carp,[751] hake,[752, 753] and pike[754] muscle. In the case of the carp parvalbumin a 1.85 Å resolution *X*-ray map has also been deduced[755] (see Part II), enabling the nature of the calcium-binding sites to be deliniated and also indicating a possible gene triplication, a feature not obvious from the primary structure. In all three cases the sequence around the calcium-binding sites has been conserved. Although there is some resemblance in properties to the calcium-binding component of troponin from mammalian muscle, troponin C (see below), the function of these proteins, found only in fish and amphibia, remains to be determined. One interesting point in the determination of the carp sequence was the use of carboxypeptidase A at pH 5.2 to remove acidic residues. Both aspartic and glutamic acid were rapidly removed and the method may clearly be useful for amide assignments, as well as for C-terminal sequence determination.

Troponin C. Troponin C (TN-C) is the calcium-binding component of the troponin complex and is the first element of this complex to have its amino-acid sequence determined.[756] Comparison of this sequence with that of carp parvalbumin (MCBP, Figure 12) shows 35 identities and the structures are particularly similar in the *C*-terminal region. A more detailed comparison, particularly with the three-dimensional structure of the carp protein, has led these authors to speculate that there are four calcium-binding segments in troponin C which not only show homology with each other but also with the calcium-binding segments of the parvalbumins. This correlates with evidence that troponin C binds four Ca^{2+} ions per mole.

Structural Proteins.—*Wool Proteins.* Component 0.62 of merino wool is a 61-residue protein containing more than 50% of glycine and aromatic residues but no lysine, histidine, glutamic acid, isoleucine, or methionine. The amino-acid sequence[757] shows, not surprisingly, glycine-rich regions but there is no obvious homology with the sulphur-rich class of wool proteins, and the role played by component 0.62 is not yet clear. The primary structure of one of these sulphur-rich proteins, protein SCMKB-IIIA3, has appeared[758] and although the methods of sequencing were relatively conventional these authors used three 150 cm columns of Sephadex G-50 or G-25 connected in series for peptide fractionation. This method apparently affords more highly purified peptides than are

[751] C. J. Coffee and R. A. Bradshaw, *J. Biol. Chem.*, 1973, **248**, 3305.
[752] J.-P. Capony and J.-F. Pechère, *European J. Biochem.*, 1973, **32**, 88.
[753] J.-P. Capony, L. Rydèn, J. Demaille, and J.-F. Pechère, *European J. Biochem.*, 1973, **32**, 97.
[754] F. Frankenne, L. Joassin, and Ch. Gerday, *F.E.B.S. Letters*, 1973, **35**, 145.
[755] R. H. Kretsinger and C. E. Nockolds, *J. Biol. Chem.*, 1973, **248**, 3313.
[756] J. H. Collins, J. D. Potter, M. J. Horn, G. Wilshire, and N. Jackman, *F.E.B.S. Letters*, 1973, **36**, 268.
[757] T. A. A. Dopheide, *European J. Biochem.*, 1973, **34**, 120.
[758] L. S. Swart and T. Haylett, *Biochem. J.*, 1973, **133**, 641.

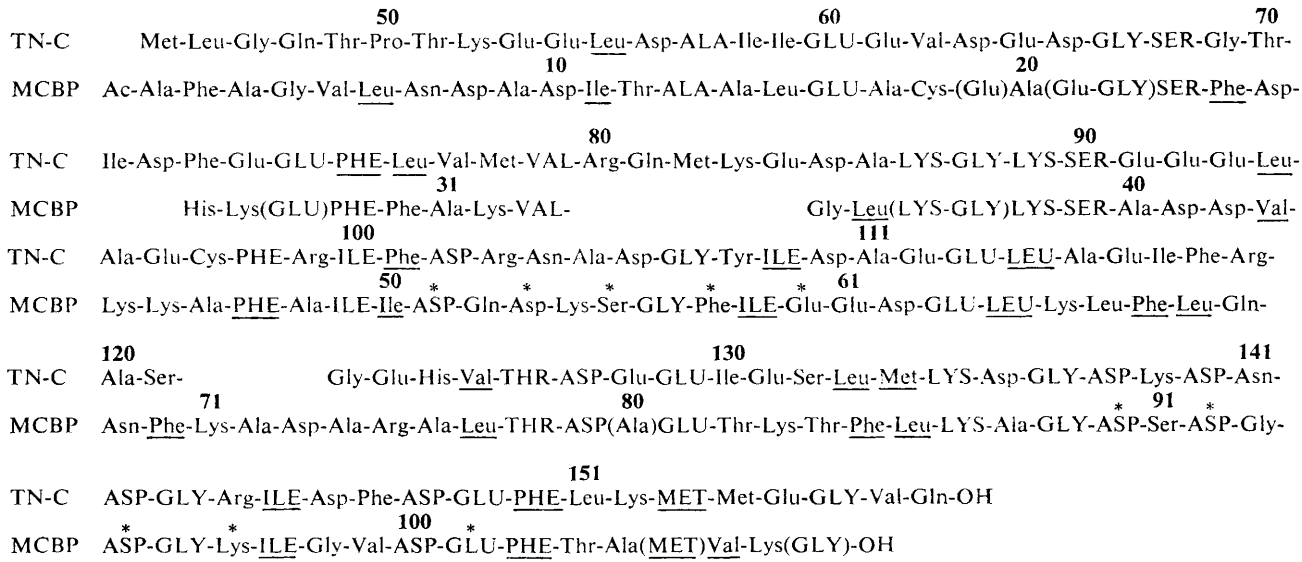


Figure 12 *Comparison of the sequences of MCBP and rabbit TN-C. The MCBP sequence is that of carp component* 3 *except where corresponding residues (given in parentheses) in components* 2 *or* 5 *or in the hake protein give identity or closer similarity with TN-C. Identical residues are capitalized. Hydrophobic core residues in MCBP and corresponding hydrophobic residues in TN-C are underlined. Asterisks (*) are used to indicate residues in MCBP that are involved in calcium-binding. Gaps in the MCBP sequence are those introduced by Kretsinger and Nockolds in their comparison of three segments of the MCBP with each other* [755]

(Reproduced by permission from *F.E.B.S. Letters*, 1973, **36**, 268)

produced by fractionation on paper. Examination of this sequence, along with those from other high-sulphur wool proteins, has led to the proposal [759] that this class of molecules contains a repeating pentapeptide unit and is derived from a single ancestral protein of sequence:

Pro-Ser-Cys-Cys-Gln-Pro-Ile-Cys-Cys-Asp

This proposal is refuted by Elleman *et al.*,[760] who suggest that the repeat unit is a decapeptide.

Proteases and Protease Inhibitors.—*Trypsin.* The primary structure of porcine trypsin has been determined [761] by Neurath's group using the methods of automated sequencer analysis that they described previously (see last year's Report, p. 47). Indeed the prime reason for completing the sequence appears to be a demonstration of the value of these methods. Analysis of the protein and eight fragments generated by autolysis, cyanogen bromide, tryptophan cleavage, hydroxylamine, and trypsin was sufficient to determine the virtually complete sequence, thus considerably simplifying peptide fractionation procedures. The protein was found to possess 82% identity with bovine trypsin.

Trypsin Inhibitor. It is well known that catalytic amounts of trypsin at pH 3.75 will specifically cleave trypsin inhibitors at a reactive Arg-Ile bond. This fact has been used [762] as a starting point for determining the sequence of the soybean inhibitor (Kunitz). The complete sequence of 181 residues [763] shows very little homology with the sequence of the Bowman–Birk inhibitor but does have the internal repetitions previously noted for the latter protein (Figure 13).

Others. Evidence for the previously reported amino-acid sequence of porcine elastase [764] has been published in great detail and a tentative amino-acid sequence of hog pepsin has finally appeared.[765]

Nucleases.—*Ribonuclease.* With the publication of complete primary structures for this enzyme from rats,[766] sheep,[767] Red Deer, and Roe Deer,[768] the total number of sequences available increases to seven. From this information it is clear that the enzyme is evolving fairly rapidly relative to cytochrome *c*, and a value of 1% substitution per 3×10^6 years has been calculated as the rate of molecular evolution. There is some evidence for microheterogeneity in both the sheep and Roe Deer enzymes

[759] L. S. Swart, *Nature New Biol.*, 1973, **243**, 27.
[760] T. C. Elleman, H. Lindley, and R. J. Rowlands, *Nature*, 1973, **246**, 530.
[761] M. A. Hermodson, L. H. Ericsson, H. Neurath, and K. A. Walsh, *Biochemistry*, 1973, **12**, 3146.
[762] T. Koide and T. Ikenaka, *European J. Biochem.*, 1973, **32**, 401.
[763] T. Koide and T. Ikenaka, *European J. Biochem.*, 1973, **32**, 417.
[764] D. M. Shotton and B. S. Hartley, *Biochem. J.*, 1973, **131**, 643.
[765] L. Moravek and V. Kestka, *F.E.B.S. Letters*, 1973, **35**, 276.
[766] J. J. Beintema and M. Gruber, *Biochim. Biophys. Acta*, 1973, **310**, 161.
[767] R. Kobayashi and C. H. W. Hirs, *J. Biol. Chem.*, 1973, **248**, 7833.
[768] H. Zwiers, A. J. Scheffer, and J. J. Beintema, *European J. Biochem.*, 1973, **36**, 569.

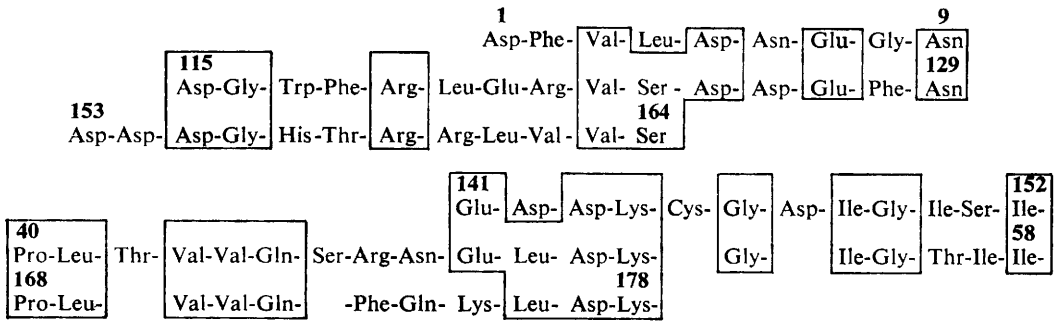


Figure 13 *The repetition in amino-acid sequence of soybean trypsin inhibitor (Kunitz) The regions enclosed in blocks show the same amino-acid residue or sequence*
(Reproduced by permission from *European J. Biochem.*, 1973, **32**, 417)

and the latter has been isolated in a glycosidated and carbohydrate-free form. In Red Deer the sugar attachment position (Asn-34 in other species) has been substituted by lysine, and this enzyme is found in the carbohydrate-free form only.

Deoxyribonuclease. The 257-residue sequence of bovine pancreatic DNase A has been determined [769, 770] and the positions of the two S—S bridges have been elucidated. As expected, there is no homology with bovine ribonuclease. Deoxyribonuclease C has been shown to differ from the A form by a single His → Pro substitution. Fingerprint analysis together with the above sequence data has now shown [771] that this substitution is at histidine-118. Hence, this residue cannot be essential for the enzyme's activity.

Venom Proteins.—In an attempt to investigate possible homology between non-neurotoxic and neurotoxic proteins, Fryklund and Eaker have determined [772] the 61-residue amino-acid sequence of a haemolytic protein from the venom of *Haemachatus haemachates* (African Ringhals Cobra). Gel-filtration fractions of guanidated and chymotryptic peptides were further purified by zone electrophoresis in a column of pyridine-washed cellulose and the sequences then determined by conventional means. There is clear homology between this non-neurotoxin and other non-neurotoxins as well as with neurotoxic venom components. There is also similarity between this sequence and those from toxin $V^{11}1$ of the Banded Egyptian Cobra (*Naja haje annulifera*),[773] toxin B of the Indian Cobra (*Naja naja*),[774] toxin F_{V11} of the Green Mamba (*Dendroaspis angusticeps*),[775] toxins $V_n^{I}I$ and $V_N^{III}I$ from Jameson's Mamba (*Dendroaspis jamesoni kaimosae*),[776] and toxins a and b of the King Cobra (*Ophiophagus hannah*).[777]

Strydom has isolated [778] two less toxic proteins from the venom of the Black Mamba (*Dendroaspis polylepsis polylepsis*) which show no homology to the above neurotoxins or cardiotoxins but appear to be similar in sequence to trypsin inhibitors (Figure 14). Furthermore, one of these (toxin K) has 15% of inhibitory activity against trypsin whereas the other (toxin I) is 5% active against chymotrypsin. As Strydom points out, this yields positive evidence that snake venom glands may be homologous to the enterosecretory system. A similar result has been obtained [779] by

[769] J. Salnikow, T.-H. Liao, S. Moore, and W. H. Stein, *J. Biol. Chem.*, 1973, **248**, 1480.
[770] T-.H. Liao, J. Salnikow, S. Moore, and W. H. Stein, *J. Biol. Chem.*, 1973, **248**, 1489.
[771] J. Salnikow and D. Murphy, *J. Biol. Chem.*, 1973, **248**, 1499.
[772] L. Fryklund and D. Eaker, *Biochemistry*, 1973, **12**, 611.
[773] K. H. K. Weise, F. H. H. Carlsson, F. J. Joubert, and D. J. Strydom, *Z. physiol. Chem.*, 1973, **354**, 1317.
[774] M. Ohta and K. Hayashi, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 431.
[775] C. C. Viljoen and D. P. Botes, *J. Biol. Chem.*, 1973, **248**, 4915.
[776] A. J. C. Strydom, *Biochim. Biophys. Acta*, 1973, **328**, 491.
[777] F. J. Joubert, *Biochim. Biophys. Acta*, 1973, **317**, 85.
[778] D. J. Strydom, *Nature New Biol.*, 1973, **243**, 88.
[779] H. Takahashi, S. Iwanaga, Y. Hokama, T. Suzuki, and T. Kitagawa, *F.E.B.S. Letters*, 1973, **38**, 217.

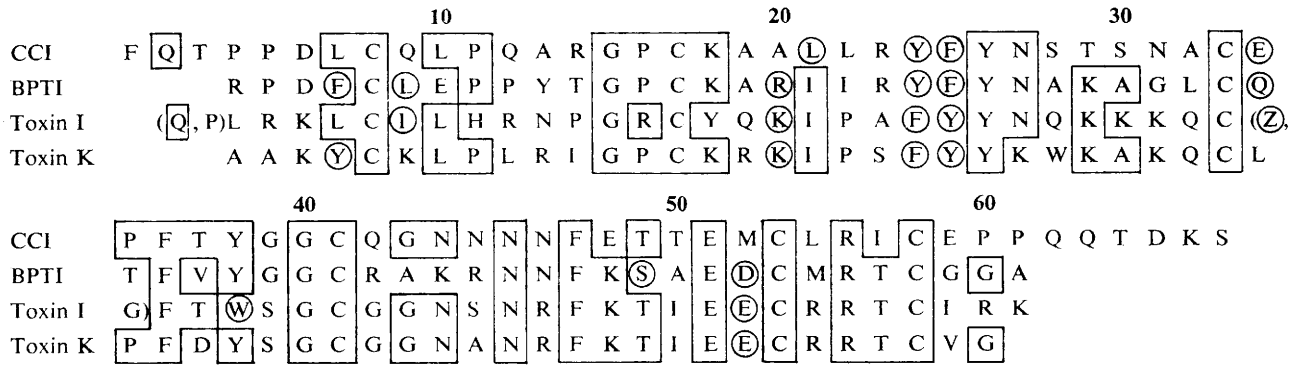


Figure 14 *The alignment of mamba toxins I and K with trypsin inhibitors. CCI = cow colostrum inhibitor. BPTI = bovine pancreatic trypsin inhibitor. Blocks drawn around residues indicate identity between at least one of the toxins and one of the inhibitors. Circles indicate similarity*

(Reproduced by permission from *Nature New Biol.*, 1973, **243**, 88)

a Japanese group working on an inhibitor isolated from the venom of Russell's Viper (*Vipera russelli*).

Finally, an enterprising if somewhat ambitious attempt has been made [780] to construct a three-dimensional structure for the *Naja nigricollis* α-toxin. The primary structures of twenty-seven different toxins along with chemical modification results and other physico-chemical evidence have been correlated and a model has been constructed by application of known rules of protein tertiary structure. The final structure, based on a mixture of intuition and speculation, will not be discussed in detail here but it will be interesting to contrast it with that produced from *X*-ray work reportedly now in progress. Venom proteins, which are both small (60—74 residues) and possess a large number (4 or 5) of disulphide bridges are particularly suited to this approach. If nothing else, the paper provides a model capable of being tested, *e.g.* by cross-linking experiments, and a useful compilation of chemical data for the venom proteins.

Immunoglobulins.—Putnam's group have now reported [781] on the sequence of the heavy chain from human IgM, thus providing a complete structure for this molecule (Figure 15). The location of the 11 disulphide bridges and five carbohydrate units is also described. Inspection of the sequence leads to the rather unexpected conclusion that there is greater similarity between variable (V) region sequences of heavy chains of the same subgroup but different class than between constant (C) regions of different classes. This finding lends support to the hypothesis that two genes code for each heavy chain, one for the V region and one for the C region. Hilschmann's group have also reported [782, 783] on a complete sequence of a human IgM protein. As expected, there are many differences between the variable regions of these two proteins but there are also apparent substitutions in the constant regions. Presumably these are due to sequencing errors which will almost certainly be present when working with molecules of this size.

Hilschmann's group have also determined the sequence of the Bence-Jones protein Rei.[784] Since a 2.8 Å resolution *X*-ray map of this protein already exists it should now be possible to construct a three-dimensional model of this light chain and, together with electron microscope data, to determine the nature of the antibody combining site. Sequences have also appeared for human κ-light chains from protein Ni,[785, 786] amyloid fibres,[787]

[780] L. Rydén, D. Gable, and D. Eaker, *Internat. J. Peptide Protein Res.*, 1973, **5**, 261.
[781] F. W. Putnam, G. Florent, C. Paul, T. Shinoda, and A. Shimizu, *Science*, 1973, **182**, 287.
[782] C. J. Laure, S. Watanabe, and N. Hilschmann, *Z. physiol. Chem.*, 1973, **354**, 1503.
[783] S. Watanabe, H. U. Barnikol, J. Horn, J. Bertram, and N. Hilschmann, *Z. physiol. Chem.*, 1973, **354**, 1505.
[784] W. Palm and N. Hilschmann, *Z. physiol. Chem.*, 1973, **354**, 1651.
[785] T. Shinoda, *J. Biochem.*, 1973, **73**, 417.
[786] T. Shinoda, *J. Biochem.*, 1973, **73**, 423.
[787] F. W. Putnam, E. J. Whitley, C. Paul, and J. N. Davidson, *Biochemistry*, 1973, **12**, 3763.

and protein Car,[788] and for a light chain derived from a rabbit anti-*p*-azobenzoate antibody of restricted heterogeneity.[789]

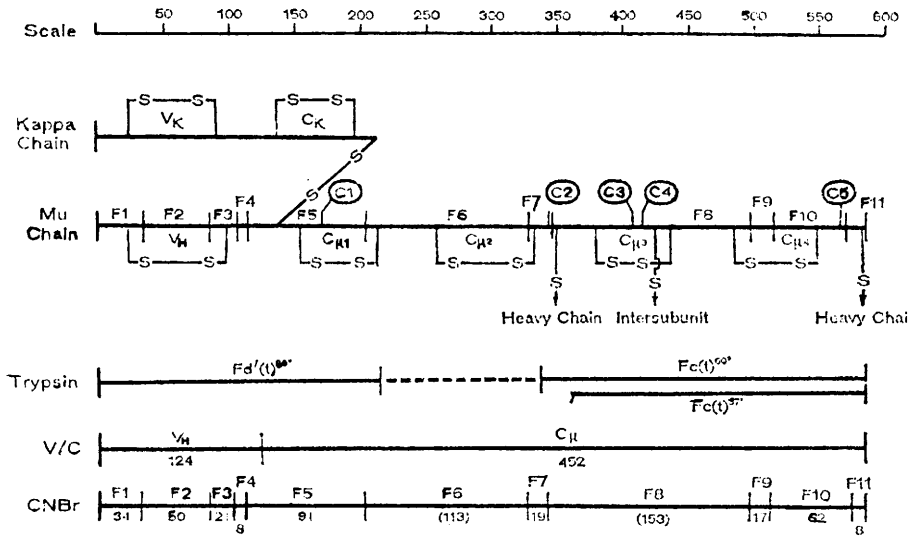


Figure 15 *Schematic structure of the μ heavy chain and the κ light chain of IgM Ou showing (i) the interchain and intrachain disulphide bridges, (ii) the two homology regions of the light chain (V_κ and C_κ), and (iii) the five homology regions of the heavy chain (V_H and $C_\mu 1$ to $C_\mu 4$). The figure for the μ chain also shows the location of the five oligosaccharides (C1 to C5), the points of cleavage by trypsin with the respective fragments (Fd′ and Fc), the sites of cleavage by* CNBr *and the respective fragments (F1 to F11), and the point of division between the variable and constant regions (V_H and C_μ). The scale and the numbers indicate the number of amino-acid residues in each chain and fragment*
(Reproduced by permission from *Science*, 1973, **182**, 287)

Peptides and Hormones.—Somatomammatropin is a hormone possessing both lactogenic and growth-promoting activities. In order to determine possible homologies with either pituitary growth hormone or pituitary lactogenic hormone, the sequence of human chorionic somatomammatropin has been elucidated by Li's group in California. Now, in a further account of this work,[790] with several changes from the previously published sequence, it has been shown that the molecule has very high homology with human growth hormone. Furthermore, a re-examination of one of the different regions in the latter sequence, also determined by Li's group, shows that elimination of an error raises this homology to 96%. This compares with a similarity of about 60% with the lactogenic hormone.

[788] C. P. Milstein, *F.E.B.S. Letters*, 1973, **30**, 41.
[789] E. Appella, D. A. Roholt, A. Chersi, G. Radzimski, and D. Pressman, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 1122.
[790] C. H. Li, J. S. Dixon, and D. Chung, *Arch. Biochem. Biophys.*, 1973, **155**, 95.

A third, considerably revised, structure has appeared for bovine growth hormone.[791] However, comparison of this new sequence with one proposed by Wallis[792] for the same molecule shows that several differences remain. In particular, residues 84—96 are given by the Argentinian group as:

Ser-Gln-Trp-Leu-(Gln, Pro)-Gly-Phe-Leu-Arg-Val

whereas Wallis has derived:

Gln-Ser-Trp-Leu-Gly-Pro-Leu-Gln-Phe-Leu-Ser-Arg-Val

for the corresponding region. It is probably significant that the latter sequence is identical to that found for the corresponding region of the ovine hormone.[793] In fact, comparison of these two sequences in their entirety shows the only differences to be at residues 99 and 130, which change from asparagine and valine in the ovine sequence to aspartic acid and glycine, respectively. The Argentinian group have also published a preliminary version of equine growth hormone.[794]

Also from Li's laboratory comes the primary structure of the β-subunit of human pituitary thyrotropin.[795] By eluting tryptic and chrymotryptic peptides from two-dimensional fingerprints it was possible to determine the 112-residue sequence using less than 2 μmol of protein. The sequence shows high homology with the bovine β-subunit. The α-subunit of this hormone has also been fingerprinted and preliminary evidence suggests that it is identical to human interstitial-cell-stimulating hormone α-subunit.

This similarity of hormone α-subunits is again apparent in a recent publication[796] of the complete sequence of human chorionic gonadotrophin (hCG). The 92-residue α-subunit of this hormone differs by only a two-residue inversion and three-residue deletion from the corresponding subunit in human luteinizing hormone (LH). There is also considerable homology with bovine thyroid-stimulating hormone α-subunit (Figure 16). Thus, a clear pattern emerges of various glycoprotein hormones having very similar α-subunits which can be interchanged without affecting the biological, immunological, or electrophoretic properties of the native hormone. In all these cases, however, the carbohydrate moieties appear to differ and it is possible that this subunit is responsible for the hormonal activity, suggesting a similar mechanism of action, whereas the β-subunit confers the hormonal specificity.

The β-subunit of human chorionic gonadotrophin, revised from a previous account (see last year's Report, p. 108), also shows considerable homology

[791] J. A. Santomé, J. M. Dellacha, A. C. Paladini, C. Peña, M. J. Biscoglio, S. T. Daurat, E. Poskus, and C. E. M. Wolfenstein, *European J. Biochem.*, 1973, **37**, 164.
[792] M. Wallis, *F.E.B.S. Letters*, 1973, **35**, 11.
[793] C. H. Li, D. Gordon, and J. Knorr, *Arch. Biochem. Biophys.*, 1973, **156**, 493.
[794] M. M. Zakin, E. Poskus, J. M. Dellacha, A. C. Paladini, and J. A. Santomé, *F.E.B.S. Letters*, 1973, **34**, 353.
[795] M. R. Sairam and C. H. Li, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 426.
[796] R. Bellisario, R. B. Carlsen, and O. P. Bahl, *J. Biol. Chem.*, 1973, **248**, 6796.

```
                                              10                                           20
hCG-α                H-Ala-Pro-Asx-Val-Glx-Asx-Cys-Pro-Glx-Cys-Thr-Leu-Glx-Glx-Asx-Pro-Phe-Phe-Ser-Glx-Pro-Gly
hLH-α                           H-Val
pLH-α                       H-Thr-Met    Gly                  Lys     Lys         Lys Tyr        Lys Leu
bTSH-α   H-Phe-Pro-Asp-Gly-Glu-Phe-Thr-Met-   Gly                  Lys     Lys         Lys Tyr        Lys     Asx

                                30                                40
hCG-α    Ala-Pro-Ile-Leu-Gln-Cys-Met-Gly-Cys-Cys-Phe-Ser-Arg-Ala-Tyr-Pro-Thr-Pro-Leu-Arg-Ser-Lys-Lys-Thr-Met-Leu-Val-
hLH-α
pLH-α            Tyr                                                  Ala
bTSH-α           Tyr                                                  Ala

         50                                   60                                  70
hCG-α    Gln-Lys-Asn(CHO)-Val-Thr-Ser-Glx-Ser-Thr-Cys-Cys-Val-Ala-Lys-Ser -Tyr-Asn-Arg-Val-Thr-Val-Met-Gly-Gly-Phe-
hLH-α
pLH-α    Pro           Ile          Ala(Thr, Cys, Cys)       Ala Phe Thr Lys Ala                 Asx Ala
bTSH-α   Pro           Ile          Ala                      Ala Phe Thr Lys Ala                 Asn Val

                            80                           90
hCG-α    Lys-Val-Glx-Asn(CHO)-His-Thr-Ala-Cys-His-Cys-Ser-Thr-Cys-Tyr-Tyr-His-Lys-Ser-OH
hLH-α                                         Ser  Cys                           -OH
pLH-α    Arg                 Ser      Glx(Cys, His, Cys, His, Thr, Cys)          -OH
bTSH-α   Arg                          Glu                                        -OH
```

Figure 16 *Comparison of the amino-acid sequences of human and porcine LH-α and bovine TSH-α with hCG-α. Only residues not identical with hCG-α are shown. Bovine and ovine LH-α have identical amino-acid sequences with TSH-α. The portions of the pLH-α sequence enclosed in* parentheses *were determined by composition only*
(Reproduced by permission from *J. Biol. Chem.*, 1973, **248**, 6796)

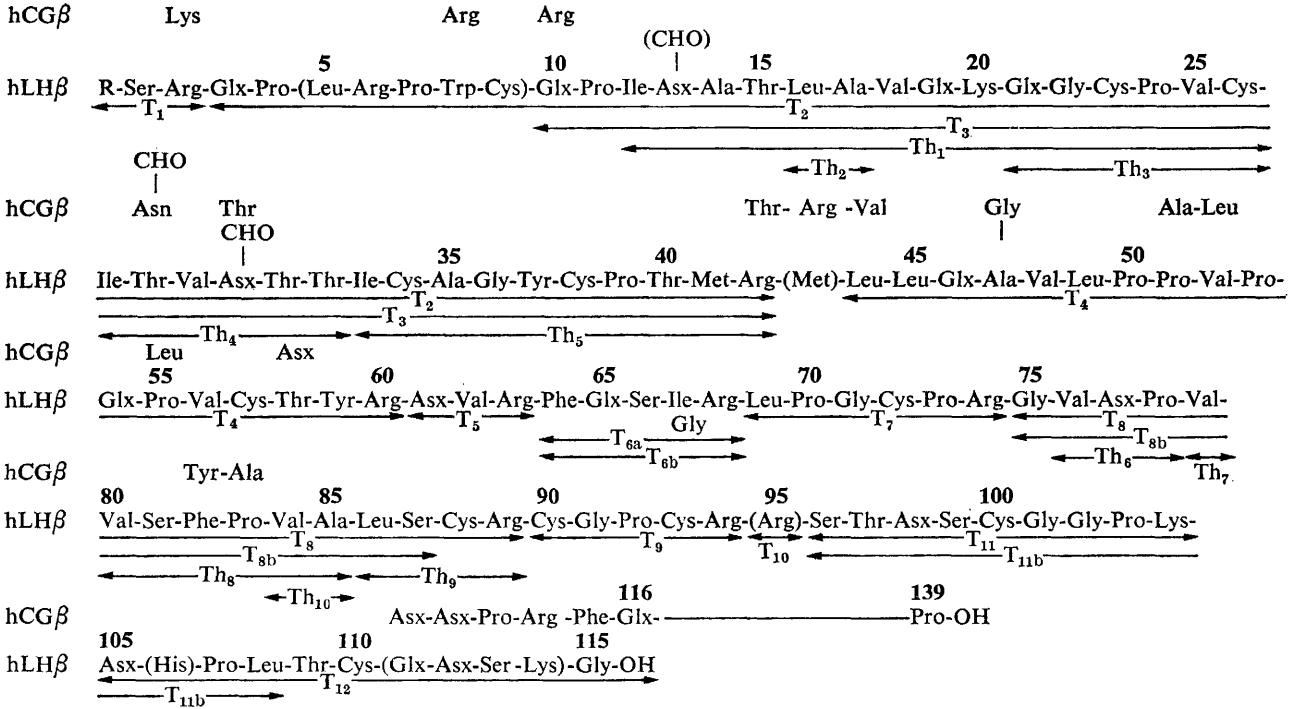


Figure 17 *The amino-acid sequence of human luteinizing hormone β-subunit. T_i indicates tryptic peptides; Th_i, peptides resulting from thermolysin digestion of tryptic peptides.* CHO *symbolizes polysaccharide prosthetic group; R, an unidentified acyl group. Parts of sequences shown in parentheses represent gaps in the elucidation of sequence. Non-identical residues in HCGβ are indicated above the LHβ sequence*

(Reproduced by permission from *F.E.B.S. Letters*, 1973, **29**, 97)

with luteinizing hormones.[797] Thus the structure of the LH-β-subunit from humans[798] shows greater similarity to hCG than to other animal luteinizing hormones (Figure 17), the most striking difference being an additional 24 residues at the *C*-terminus of hCG-β. The sequence variability found previously in bovine LH-α has again been noted for the human hormone β-subunit, both isoleucine and glycine being found at position 67. Heterogeneity has also been observed in porcine LH-β but not in the bovine hormone.[799] Comparison of these structures and also those of the porcine[800] and ovine LH-α subunits is complicated by the fact that the one is obtained by reference to the other. For instance, parts of the porcine LH-β structure were derived from the assumption that tryptic peptides of the same composition will also have the same sequence. This assumption has often been shown to be invalid and these comparisons should therefore be treated as preliminary.

Hypothalamic Peptides. Three papers have appeared from the same group on the structure and mode of action of somatostatin, a peptide purified from ovine hypothalamus which is claimed to inhibit the release of growth hormone.

Firstly, an assay procedure was developed[801] which, by monitoring production of immunoreactive growth hormone, enabled the purification of 8.5 mg of peptide from 500 000 ovine hypothalamic fragments. The structure was then established[802, 803] as

Ala-Gly-Cys-Lys-Asn-Phe-Phe-Trp-Lys-Thr-Phe-Thr-Ser-Cys

by a combination of conventional Edman procedures, with modifications to compensate for the hydrophobicity of the peptide, and mass spectroscopy.

Both this molecule and its synthetic replicate have been shown to be active *in vivo* in rats although the physiological significance of the *in vivo* assay is not yet clear. The difficulty inherent in interpreting this type of experiment is demonstrated by the controversy over which peptides inhibit the release of melanocyte-stimulating hormone from the hypothalamus. Claims and counter-claims have been made for the peptide Pro-Leu-Gly-NH_2 and for tocinoic acid (Cys-Tyr-Ile-Gln-Asn-Cys). A further report[804] now finds that neither of these molecules nor the analogue, des-Cys-tocinoic acid, blocks or stimulate secretions of the hormone.

[797] R. B. Carlsen, O. P. Bahl, and N. Swimanathan, *J. Biol. Chem.*, 1973, **248**, 6810.
[798] J. Closset, G. Hennen, and R. M. Lequin, *F.E.B.S. Letters*, 1973, **29**, 97.
[799] G. Maghuin-Rogister and G. Hennen, *European J. Biochem.*, 1973, **39**, 235.
[800] G. Maghuin-Rogister, Y. Combarnous, and G. Hennen, *European J. Biochem.*, 1973, **39**, 255.
[801] N. Ling, R. Burgus, J. Rivier, W. Vale, and P. Brazeau, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 127.
[802] R. Burgus, N. Ling, M. Butcher, and R. Guillemin, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 684.
[803] P. Brazeau, W. Vale, R. Burgus, N. Ling, M. Butcher, J. Rivier, and R. Guillemin, *Science*, 1973, **179**, 77.
[804] N. H. Grant, D. E. Clark, and E. T. Rosanoff, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 100.

Miscellaneous. Following the ancient belief that honeybee venom is beneficial in certain arthritic and rheumatical conditions, a group of London researchers have assayed various venom fractions for anti-inflammatory activity.[805] Only one active fraction was found and this contained a 22-residue peptide whose structure was determined as

Ile-Lys-Cys-Asn-Cys-Lys-Arg-His-Val-Ile-Lys-Pro-His-Ile-Cys-Arg-Lys-Ile-Cys-Gly-Lys-Asn-NH_2

Activity was lost after reduction and carboxymethylation or by trypsin treatment. Interestingly, this structure is identical to that of a peptide isolated previously which was shown to have powerful mast-cell-degranulating activity, although these authors do not feel that this explains the anti-inflammatory effects. As there are only very small amounts of the peptide in the venom it is unlikely to find use as a therapeutic agent, although synthetic analogues may be of value.

Subtilin. Subtilin is a heterodetic polycyclic peptide isolated from *Bacillus subtilis.* Amino-acid analysis shows[806] that it is 27 residues long and in addition to the common amino-acids contains four residues of 3-methyl-lanthionine, one residue of lanthionine, and one residue of 3-methyl-dehydroalanine. By a combination of tryptic and thermolytic fragmentation[807–809] subtilin has been shown[810] to be a heterodetic pentacyclic peptide with a unique distribution of configurations at α-carbon atoms. The first alanine moiety of lanthionine and all four 2-aminobutyric acid moieties of the 3-methyl-lanthionine residues are of the D-configuration (Figure 18).

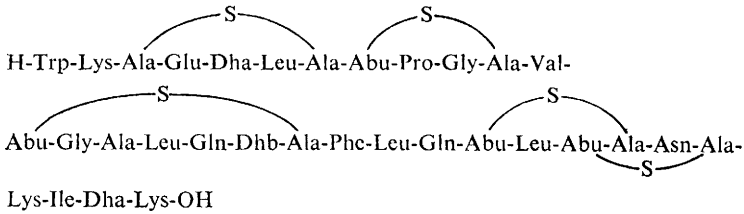


Figure 18 *Structure of subtilin.* (*Dha = dehydroalanine, Abu = α-aminobutyric acid, Dhb = dehydro-α-aminobutyric acid*)

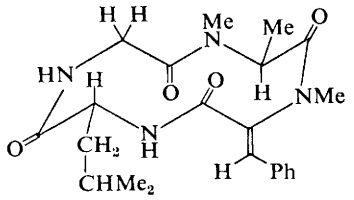
Tentoxin. The fungus *Alternaria tenuis* Auct. synthesizes a cyclic tetrapeptide, tentoxin, which causes chlorosis in germinating seedlings. The observation that treatment of tentoxin with base causes the ring system to

[805] M. E. J. Billingham, J. Morley, J. M. Hanson, R. A. Shipolini, and C. A. Vernon, *Nature*, 1973, **245**, 163.
[806] E. Gross, H. H. Kiltz, and L. C. Craig, *Z. physiol. Chem.*, 1973, **354**, 799.
[807] H. H. Kiltz and E. Gross, *Z. physiol. Chem.*, 1973, **354**, 802.
[808] H. H. Kiltz and E. Gross, *Z. physiol.Chem.*, 1973, **354**, 805.
[809] E. Nebelin and E. Gross, *Z. physiol. Chem.*, 1973, **354**, 807.
[810] E. Gross, H. H. Kiltz, and E. Nebelin, *Z. physiol. Chem.*, 1973, **354**, 810.

open has allowed conventional sequencing methods to be applied to this peptide.[811] The analysis was somewhat complicated by the presence of the unusual amino-acids *N*-methylalanine and *N*-methyldehydrophenylalanine, but a sequence

MeDHPhe-L-Leu-Gly-L-MeAla

was derived by comparison with standards. High-resolution mass spectroscopy confirmed this structure and also indicated the nature of the ring system (16).



(16)

Finally, the structure of the wall peptidoglycan from *Streptomyces* k39 has been determined.[812] Using a series of specific amino- and carboxypeptidases the structure L-alanyl-D-isoglutaminyl-(L)-*meso*-diaminopimelyl-(L)-D-alanine, cross-linked through D-alanyl-(D)-*meso*-diaminopimelic acid linkages, was deduced.

Redox Proteins.—*Cytochrome c.* The amino-acid sequences of nine further cytochromes *c* have appeared this year. Three of these, leek,[813] spinach,[814] and *Nigella*,[815] have come from Boulter's group as part of their investigation into plant phylogeny based on protein sequence studies. The approach here has been to perform a single chymotryptic digest, to purify every peptide, and to align these by homology with known sequences from other plant cytochromes *c*. Although requiring great care to avoid errors, sequences can be obtained using very small quantities of protein; only 200 nmol was used for the *Nigella* structure. One surprising result was that residue 53 in the spinach protein was serine whereas it has been glycine in all eukaryotic mitochondrial cytochrome *c* sequences to date. This reduces the number of 'invariant' residues to 28.

The sequences obtained from three species of *Pseudomonas*[816] were surprising in that, although by bacteriological criteria the organisms were closely related, the differences between the structures varied from 22% to

[811] M. Koncewicz, P. Mathiaparanam, T. F. Uchytil, L. Sparapano, J. Tam, D. H. Rich, and R. D. Durbin, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 653.
[812] J.-M. Ghuysen, M. Leyh-Bouille, J. N. Campbell, R. Moreno, J.-M. Frère, C. Duez, M. Nieto, and H. R. Perkins, *Biochemistry*, 1973, **12**, 1243.
[813] R. H. Brown and D. Boulter, *Biochem. J.*, 1973, **131**, 247.
[814] R. H. Brown, M. Richardson, R. Scogin, and D. Boulter, *Biochem. J.*, 1973, **131**, 253.
[815] R. H. Brown and D. Boulter, *Biochem. J.*, 1973, **133**, 251.
[816] R. P. Ambler and M. Wynn, *Biochem. J.*, 1973, **131**, 485.

39%. Thus, the relationship between these cytochromes *c*-551 and the mitochondrial proteins is unlikely to be determined from primary structure work and may have to await detailed three-dimensional analysis.

Two groups have independently arrived at the same structure for cytochrome *c*-558 from the photosynthetic flagellate *Euglena gracilis*.[817, 818] The structure shows some similarities with that from another flagellate, *Crithidia oncopelti*, in that residue 14 from both species is alanine rather than the usual cysteine. This would imply that the haem is bound through only one covalent link, leaving one vinyl group unsaturated. Also, substitutions which were unique in the original *Crithidia* sequence have also been observed in *Euglena* together with a new substitution, $Tyr_{67} \rightarrow Phe$. These protozoa differ in 48 out of 103 positions and it has not proved possible to establish evolutionary relationships between these primitive eukaryotes and prokaryotic species.

Flavodoxins. Flavodoxins are a low molecular weight class of flavoenzymes whose primary structures are of interest now that tertiary structure information is becoming available.[819] Two sequences have been published this year, one from *Desulphovibrio vulgaris* [820] and one from *Peptostreptococcus elsdenii*.[821] Identity between these two is low, about 24%, but both structures retain residues Ser-10, Thr-12, Thr-15, and Asn-14 (numbered according to the *Desulphovibrio* sequence), which the *X*-ray studies have shown to be important for hydrogen-bonding to the phosphate group of the coenzyme. However, Trp-60, which forms part of a cleft for the isoalloxazine ring, is absent from the *Peptostreptococcus* sequence.

The primary structure of the thermostable ferredoxin from *Clostridium thermosaccharolyticum* is interesting in that the only differences from the mesostable protein from *C. tartarivorum* are the substitutions of Gln-31 and Gln-44 by glutamic acid residues.[822] Thus, the increased stability may arise from hydrogen-bonding of one or both of these glutamic acids to another amino-acid side-chain. This possibility may also be demonstrated by reference to the tertiary structure of the ferredoxin from *Micrococcus aerogenes*.[823]

In spite of a very high difference in redox potential it has previously been shown that the arrangement of the two iron–sulphur clusters in *Clostridium* ferredoxins is almost identical to that in the high-potential iron–sulphur proteins. However, the recently determined primary structure of the latter

[817] G. W. Pettigrew, *Nature*, 1973, **241**, 531.
[818] D. K. Lin, R. L. Niece, and W. M. Fitch, *Nature*, 1973, **241**, 533.
[819] K.-D. Watenpaugh, L. C. Sieker, L. H. Jensen, J. Le Gall, and M. Dubourdieu, *Proc. Nat. Acad. Sci. U.S.A.*, 1972, **69**, 3185.
[820] M. Dubourdieu, J. Le Gall, and J. L. Fox, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 1418.
[821] M. Tanaka, M. Haniu, K. T. Yasunobu, S. Mayhew, and V. Massey, *J. Biol. Chem.*, 1973, **248**, 4354.
[822] M. Tanaka, M. Haniu, K. T. Yasunobu, R. H. Himes, and J. M. Akagi, *J. Biol. Chem.*, 1973, **248**, 5215.
[823] E. T. Adman, L. C. Sieker, and L. H. Jensen, *J. Biol. Chem.*, 1973, **248**, 3987.

protein from *Chromatium vinosum*[824] shows little resemblance to ferredoxins although, as has been claimed previously for the latter molecules, there is a possible internal sequence repetition (Figure 19).

A corrected version of the amino-acid sequence of another iron–sulphur protein, bovine adrenodoxin, has also appeared.[825] There appears to be no homology with other proteins of this class although the sequence Cys-X-Y-Cys-His which is found in cytochromes *c* also appears in this molecule.

Globins.—Complete amino-acid sequences have appeared for the haemoglobin α and β chains of Savannah Monkey (*Cereopithecus aethiops*),[826] the Capuchin Monkey (*Celus apella*),[827] the Spider Monkey (*Ateles geoffroyi*),[828] chicken (β-chain, AII component),[829] the echidna (*Tachyglossus aculeatus aculeatus*, α-chain),[830] mice from strains C_{57}BL, SWR, and NB (β-chains only),[831] and mice of strain BALB/c (2 β-chain components).[832] The two components from the BALB/c mice differed in six positions, were present in ratios of 4 : 1, and possibly arose by a gene-duplication mechanism. The only other interesting fact to emerge from all this work was the finding of alanine at position 116 in the chicken β-chain. In other adult mammalian haemoglobins this has always been arginine or histidine, and as this residue forms important hydrogen bonds in the α,β contact it may imply a functional difference between mammalian and bird haemoglobins.

Hb Mutants. A list of abnormal haemoglobins characterized during 1973 is given in Table 6 (see p. 157). Of these Hb Camden, Hb St. Antoine, Hb J. Nyanza, Hb D Ouled Rabah, Hb D Iran, Hb Sawara, Hb F Kuala Lumpur, and Hb Beograd were little changed or unchanged physiologically. Hb Ube-1 turned out to be identical to Hb Köln. This is all in accordance with the predictions of Perutz that whenever the mutation involves an amino-acid on the surface of the molecule there will be little change in the functional characteristics of the haemoglobin. An exception to this rule was Hb Seattle which, although having an abnormally low oxygen affinity, was reported to have only one surface substitution, β_{76} Ala → Glu. Re-examination of this structure by both crystallographic and[833a] fingerprint techniques[833b] now shows that the original protein chemistry was

[824] K. Dus, S. Tedro, and R. G. Bartsch, *J. Biol. Chem.*, 1973, **248**, 7318.
[825] M. Tanaka, M. Haniu, K. T. Yasunobu, and T. Kimura, *J. Biol. Chem.*, 1973, **248**, 1141,
[826] G. Matsuda, T. Maita, B. Watanabe, A. Araya, K. Morokuma, M. Goodman, and W. Prychodko, *Z. physiol. Chem.*, 1973, **354**, 1153.
[827] G. Matsuda, T. Maita, B. Watanabe, A. Araya, K. Morokuma, Y. Ota, M. Goodman, J. Barnabas, and W. Prychodko, *Z. physiol. Chem.*, 1973, **354**, 1513.
[828] G. Matsuda, T. Maita, Y. Suzuyama, M. Setoguchi, Y. Ota, A. Araya, M. Goodman, J. Barnabas, and W. Prychodko, *Z. physiol. Chem.*, 1973, **354**, 1517.
[829] G. Matsuda, T. Maita, K. Mizuno, and Y. Ota, *Nature New Biol.*, 1973, **244**, 244.
[830] R. G. Whittaker, W. K. Fisher, and E. O. P. Thompson, *Austral. J. Biol. Sci.*, 1973, **26**, 877.
[831] R. A. Popp, *Biochim. Biophys. Acta*, 1973, **303**, 52.
[832] R. A. Popp and E. G. Bailiff, *Biochim. Biophys. Acta*, 1973, **303**, 61.
[833] (*a*) N. L. Anderson, M. F. Perutz, and G. Stamatoyannopoulos, *Nature New Biol.*, 1973, **243**, 274; (*b*) S. Kurachi, M. Hermodson, S. Hornung, and G. Stamatoyannopoulos, *ibid.*, p. 275.

	7																					**23**
	Val-	Ala-	Ala-	Asp-	Asn-	Ala-	Thr-	Ala-	Ile	-Ala-	Leu-	Lys-	Tyr-	Asn-	Gln-	Asp	—					Ala
	29																					**44**
	Val-	Ala-	Ala-	Ala	-Arg-	Pro	—		Gly-	Leu-	Pro-	Pro-	Glu	—		Glu-	Gln-	His	-Cys-	Ala
Base changes	0	0	0	1	2	1			1	1	1	1	1			2	0	1	—	0

Figure 19 *Internal homology of* Chromatium *high-potential iron–sulphur proteins. These two segments are listed as an example because they show the best fit and cover a number of residues sufficient to be significant. Shorter segments homologous to the above sequences or other areas of the molecule were also noted. A dash is used to indicate a gap or deletion*
(Reproduced by permission from *J. Biol. Chem.*, 1974, **248**, 7318)

incorrect and the substitution is in fact, β_{70} Ala → Asp. This residue is a haem contact, hence explaining the low oxygen affinity.

Finally, the structure of Hb P-Nilotic has been shown [834] to constitute an 'anti-Lepore' haemoglobin. These, and the corresponding 'Lepore' haemoglobins, are of great theoretical interest as they illustrate at the molecular level the consequences of gene crossing-over. Fingerprint analysis of this particular mutant has shown that the non-α chain is a reversed or anti-Lepore β–δ chain, with the *N*-terminal region derived from the β-chain and the cross-over occurring between positions 22 and 50. As these workers point out, it will be of interest to study the biosynthesis of the β–δ chains to ascertain whether the cross-over occurs only in the earlier stages of erythroid maturation, as is the case for the δ–β chain of Hb Lepore.

Chromosomal Proteins.—*Histones.* The essentially complete primary structure of histone IIb1 from Rainbow Trout testes has been published by Bailey and Dixon.[835] The sequence shows four substitutions and four deletions relative to that from calf thymus, most of these changes being at the *C*-terminal end of the molecule. Comparison with other histone sequences shows that the evolutionary rate of this protein is intermediate between those of histone I and histone IV. Thus, we have a situation where histone IV is evolving at a rate of 0.013 residue per 100 per 10^7 years, histone III at a rate of 0.07 residue per 100 per 10^7 years, histone IIb1 at 0.16, and histone I at 1.6 residues per 100 per 10^7 years, a rate approaching that of α-haemoglobin. This would substantiate the proposal that histone IV interacts with an equally unvariable region of eukaryotic DNA whereas, at the other extreme, histone I may interact with chromosomal components which vary in a species-specific manner.

A similar approach to that used on histone IIb1, maleylation followed by tryptic cleavage at arginine bonds, has been used for the structure determination of histone III from calf thymus by Emil Smith's group in California.[836] As has been previously noted, not all arginine bonds are completely cleaved under these conditions, possibly as a result of introducing large negatively charged maleyl groups. Treatment with hydroxylamine, which removes maleyl side-chains from serine and threonine hydroxy-groups, may alleviate this problem. The complete 135-residue sequence was assembled by overlapping with chymotryptic [837] and cyanogen bromide fragments.[838] Histone III is the only histone in calf thymus to contain cysteine. The equivalent protein has now been isolated from carp testes and it has been shown [839] that the only sequence difference

[834] F. M. Badr, P. A. Lorkin, and H. Lehmann, *Nature New Biol.*, 1973, **242**, 107.
[835] G. S. Bailey and G. H. Dixon, *J. Biol. Chem.*, 1973, **248**, 5463.
[836] R. J. DeLange and E. L. Smith, *J. Biol. Chem.*, 1973, **248**, 3248.
[837] J. A. Hooper and E. L. Smith, *J. Biol. Chem.*, 1973, **248**, 3255.
[838] R. J. DeLange, J. A. Hooper, and E. L. Smith, *J. Biol. Chem.*, 1973, **248**, 3261.
[839] J. A. Hooper, E. L. Smith, K. R. Sommer, and R. Chalkley, *J. Biol. Chem.*, 1973, **248**, 3275.

is the substitution of one of the cysteines found in the calf protein by serine. This is a similar substitution to that reported last year for histone III from chicken erythrocytes. Thus, the hypothesis that reduction–oxidation of the cysteine residues may play a role in chromosomal structure may be redundant, although it remains to be explained why these thiol groups are unreactive in interphase chromatin.

As has been noted earlier, there is only one amino-acid substitution between the amino-acid sequences of histone IV from calf thymus and pea embryo. The corresponding sequences from histone III have now been shown to differ by four substitutions [840] and this rate of mutation concurs with the scheme discussed above.

Protamines. Following his work on the Y-components of protamine from tuna fish (see last year's Report, p. 125), Bretzel has now published sequences for the two Z-components.[841, 842] The approach used was the same for both proteins. Partial digestion with pepsin in 2M-NaCl cleaved a single Tyr-Arg bond, yielding two fragments. After separation of these on phosphocellulose, thermolysin digestion of each peptide generated enough major and minor fragments to provide the complete sequences:

Z1 Pro-$(Arg)_5$-Ser-Ser-Arg-Pro-Val-$(Arg)_5$-Tyr-$(Arg)_2$-Ser-

Z2 Pro-$(Arg)_5$-Ser-Ser-Arg-Pro-Val-$(Arg)_5$-Tyr-$(Arg)_2$-Ser-

Z1 Thr-Val-Ala-$(Arg)_5$-Val-Val-$(Arg)_4$

Z2 Thr-Ala-Ala-$(Arg)_5$-Val-Val-$(Arg)_4$

Miscellaneous.—*Carbonic Anhydrases.* Human erythrocytes contain at least two distinctly different enzymes, each catalysing the reversible hydration of carbon dioxide. The primary structure of one of these enzymes, carbonic anhydrase C, and a new version of the structure of the second enzyme, carbonic anhydrase B, have recently been published. In the latter case Lin and Deutsch have used the approach of sequencing cyanogen bromide fragments and overlapping these with tryptic peptides.[843, 844] Considerable difficulty was experienced with the tryptic digestion as a large core of undigested material was formed and many of the fragments were insoluble, even in 5.7M-GuCl. Use was made of the semi-automatic sequencing apparatus devised by this group (see Methods, p. 113) to complete the 256-residue sequence. However, comparison of this structure with that from Anderson *et al.* (see last year's Report, p. 79) shows no fewer than 17 differences, including inversions, deletions, insertions, and amide changes. Meanwhile a French group, working on the same protein, have determined the sequence of the first 148 residues and are in agreement with the Swedish workers. Clearly, more work needs to be done to resolve these differences.

[840] L. Patthy, E. L. Smith, and J. Johnson, *J. Biol. Chem.*, 1973, **248**, 6834.
[841] G. Bretzel, *Z. physiol. Chem.*, 1973, **354**, 312.
[842] G. Bretzel, *Z. physiol. Chem.*, 1973, **354**, 543.
[843] K.-T. D. Lin and H. F. Deutsch, *J. Biol. Chem.*, 1973, **248**, 1881.
[844] K.-T. D. Lin and H. F. Deutsch, *J. Biol. Chem.*, 1973, **248**, 1885.

The Swedish group have also reported on the structure of carbonic anhydrase C, again from human erythrocytes.[845] The approach is the same as was used on the 'B' enzyme and is one which appears to work well for these proteins. Tryptic cleavage was restricted to arginyl bonds by blocking lysine side-chains with methyl acetimidate hydrochloride. Insoluble fragments were fractionated on Sephadex in acetic acid, dodecylamine, or guanidine while soluble fragments were purified on SP-Sephadex. Acetimidyl blocking groups were removed by incubation in conc. NH_4OH, pH 11.3, at room temperature for 16—20 h. Other enzymic digestions were used to order the fragments. Comparison of the two sequences shows that 153 out of the 260 residues are identical. Furthermore, those residues identified from the 2 Å resolution *X*-ray map that exists for the 'C' enzyme as being part of the active site have identical or similar counterparts in the 'B' enzyme sequence. Thus, both enzymes may have similar three-dimensional structures in the active-site region.

Apart from the work on the three-dimensional structure, publication of these sequences permits the characterization of various carbonic anhydrase mutants. Thus the electrophoretic mutant CA 1d has now been shown [846] to be due to a Thr → Lys change at residue 100 and the P mutant [847] corresponds to an Asp → Val change at residue 232.

Glutamate Dehydrogenase. A seventh paper has appeared [848] from Emil Smith's group on the structure of this enzyme from bovine liver, this time on 82 peptic peptides. During the course of this study three examples were encountered of cleavage on the *N*-terminal side of prolyl bonds. As pointed out by these authors, this is unlikely to be enzymic cleavage but probably reflects the known lability of prolyl bonds at acid pH. In a final (eighth) paper [849] hydroxylamine cleavage of an Asn-Gly bond and cleavage of a tryptophan bond with the bromine adduct of 2-(2-nitrophenylsulphenyl)-3-methylindole were used to produce overlapping fragments and complete the 500-residue sequence, considerably correcting an earlier version in the process.

A study has also been made of the structure of the same enzyme from chicken liver;[850] the almost complete sequence shows only 30 differences relative to the bovine counterpart.

Tryptophan Synthetase. Primary structures have been proposed for the α-chains of this enzyme from *Salmonella typhimurium* [851] and *Aerobacter aerogenes*.[852] These studies must be regarded as tentative as they rely

[845] L. E. Henderson, D. Henriksson, and P. O. Nyman, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 1388.
[846] R. J. Tanis, R. E. Ferrell, and R. E. Taskian, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 699.
[847] S. Funakoshi and H. F. Deutsch, *J. Biol. Chem.*, 1970, **245**, 4913.
[848] D. Piszkiewicz, M. Landon, and E. L. Smith, *J. Biol. Chem.*, 1973, **248**, 3067.
[849] K. Moon and E. L. Smith, *J. Biol. Chem.*, 1973, **248**, 3082.
[850] K. Moon, D. Piszkiewicz, and E. L. Smith, *J. Biol. Chem.*, 1973, **248**, 3093.
[851] S.-L. Li and C. Yanofsky, *J. Biol. Chem.*, 1973, **248**, 1830.
[852] S.-L. Li and C. Yanofsky, *J. Biol. Chem.*, 1973, **248**, 1837.

heavily on homology considerations with the published sequence from *E. coli* as well as rather circumspect direct evidence. Thus, the author's claim of 79% identity between the three chains must represent a maximum as certain parts of the sequences are constructed by assuming maximum similarity. However, residues at positions where inactivating amino-acid changes have been noted in *E. coli* mutants do appear to have been conserved.

α_1-Acid Glycoprotein. Progress in determining the amino-acid sequence of this human plasma protein has been hampered by its resistance to enzymic digestion and its pronounced tendency to aggregate. Cleavage by cyanogen bromide, however, produces two fragments and following last year's publication of the structure of the *C*-terminal peptide, the sequence of the *N*-terminal fragment has now appeared.[853] Assembly of the complete primary structure shows that α_1-acid glycoprotein contains 181 residues and 5 carbohydrate units, all linked *N*-glycosidically to asparaginyl residues. The carbohydrate moieties account for 40% of the glycoprotein. Inspection of the sequence shows two unusual properties. Firstly, there are 21 amino-acid substitutions along the chain, some of which may represent genetic markers. As these authors point out, this exceptionally high degree of heterogeneity suggests a possible relationship with the immunoglobulins, the only other class of proteins which possess such a high number of amino-acid replacements. Secondly, following the observation that the *C*-terminal region of this protein resembled an IgG heavy chain it is now noted that the *N*-terminal region has 27% homology with an IgG κ-type light chain (Figure 20). It is hypothesized that this molecule may be related to a possible ancestral immunoglobulin.

β_2-Microglobulin. Another human serum protein that has structural similarities to IgG is β_2-microglobulin. A previous report showing homology with certain regions of the myeloma protein Eu has now been confirmed by Cunningham *et al.*,[854] who have completed the 100-residue sequence and placed the single S—S bridge. Peptides were purified by conventional methods and by automatic computer monitoring of column effluents in the cascade process described by Edelman and Gall.[855] The strategy employed was to isolate the easily purified peptides from a large number of different enzymic digests and hope that these cover the entire molecule. Until recently the function of this molecule was unknown but a recent report[856] shows that it is synthesized by human lymphoid cells and is found on the lymphocyte surface. Furthermore, a note added in proof to Cunningham's paper states that β_2-microglobulin constitutes

[853] K. Schmid, H. Kaufmann, S. Isemura, F. Bauer, J. Emura, T. Motoyama, M. Ishiguro, and S. Nanno, *Biochemistry*, 1973, **12**, 2711.

[854] B. A. Cunningham, J. L. Wang, I. Berggard, and P. A. Peterson, *Biochemistry*, 1973, **12**, 4811.

[855] G. M. Edelman and W. E. Gall, *Proc. Nat. Acad. Sci. U.S.A.*, 1971, **68**, 1444.

[856] T. H. Hütteroth, H. Cleve, S. D. Lituin, and M. D. Poulik, *J. Exp. Med.*, 1973, **137**, 838.

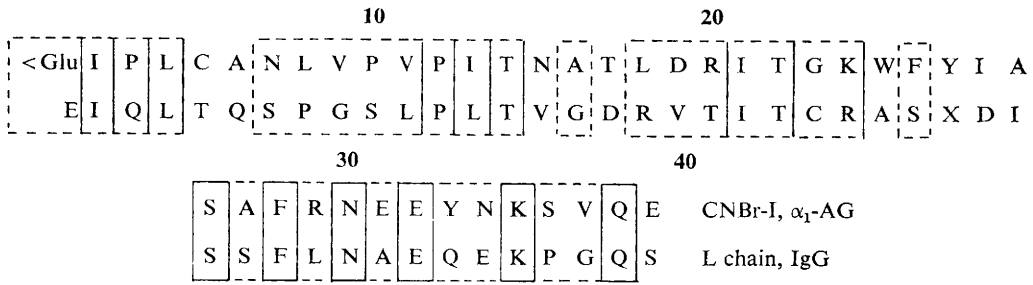


Figure 20 *The homology between the amino-terminal* 43*-residue segment of* α_1*-acid glycoprotein and the amino-terminal segment of a* κ*-type L chain. For this presentation the single-letter code for amino-acids is used. Identical residues are boxed in and the residues which require a single point mutation for identity are in boxes with broken frames*
(Reproduced by permission from *Biochemistry*, 1973, **12**, 2711)

Table 6 *Complete amino-acid sequences*

Protein	*Source*	*No. of residues*	% *automated*	*Notes*	*Ref.*
α_1 Acid glycoprotein	Human plasma	180	—	5 carbohydrate units	853
Actin	Rabbit muscle	374	6		748
Adrenodoxin	Bovine	114	—	Corrected version	825
Anti-inflammatory peptide	Honeybee venom	22	—		805
Aspartate aminotransferase	Pig heart cytoplasm	412	2		857, 858
Carbonic anhydrase B	Human erythrocytes	256	—	17 differences from previous structure	844
Carbonic anhydrase C	Human erythrocytes	259	—	2 Å *X*-ray map exists	845
Carbonic anhydrase C	Human, mutant CA 1d	—	—	One substitution	846
$\kappa\beta$-Casein	Bovine	169	—	2 genetic variants	859
Chromatium Fe/S protein	*Chromatium vinosum*	85	—		824
α-Crystallin	Bovine	173	—	αA_2 chains	860
Cytochrome b_5	Human liver	90	—	Tentative	861
Cytochrome *c*	*Allium porrum* L. (leek)	111	—	Uses 0.4 μmol of protein	813
Cytochrome *c*	*Dromeus novaehollandiae* (Emu)	104	—		862
Cytochrome *c*	*Euglena gracilis*	103	—		817, 818
Cytochrome *c*	Lamprey	104	—	Using 0.2 μmol of protein	863
Cytochrome *c*	*Nigella damascena* L.	111	—		815
Cytochrome *c*	*Pseudomonas fluorescens*	82	—		816

[857] Yu. A. Ovchinnikov, C. A. Egorov, N. A. Aldanova, M. Y. Feigina, V. M. Lipkin, N. G. Abdulaev, E. V. Grishin, A. P. Kiselev, N. N. Modyanov, A. E. Braunstein, O. L. Polyanovsky, and V. V. Nosikov, *F.E.B.S. Letters*, 1973, **29**, 31.

[858] S. Doonan, H. J. Doonan, R. Hanford, C. A. Vernon, J. M. Walker, F. Bossa, D. Barra, M. Carboni, P. Fasella, F. Riva, and P. L. Walton, *F.E.B.S. Letters*, 1973, **38**, 229.

[859] J.-C. Menier, G. Brignon, and B. Ribadeau-Dumas, *European J. Biochem.*, 1973, **35**, 222.

[860] F. J. van der Ouderaa, W. W. de Jong, and H. Bloemendal, *European J. Biochem.*, 1973, **39**, 207.

[861] M. A. Rashid, B. Hagihara, M. Kobayashi, S. Tani, and A. Tsugita, *J. Biochem.*, 1973, **74**, 985.

[862] R. C. Augusteyn, *Biochim. Biophys. Acta*, 1973, **303**, 1.

Table 6 (*cont.*)

Protein	*Source*	*No. of residues*	% *automated*	*Notes*	*Ref.*
Cytochrome *c*	*Pseudomonas mendocina*	82	—		816
Cytochrome *c*	*Pseudomonas stutzeri*	82	—		816
Cytochrome *c*	*Spinacea oleracea* L. (spinach)	111	—	Uses 1.0 μmol of protein	814
Cytochrome *c'*	*Alcaligenes*	127	—		864
Deoxyribonuclease A	Bovine pancreas	257	—	Carbohydrate at Asn-18	770
Deoxyribonuclease C	Bovine pancreas	257	—	One substitution, $His_{118} \rightarrow Pro$	771
Elastase	Porcine pancreas	240	—		764
Ferredoxin	*Clostridium thermosaccharolyticum*	55	60		822
Flavodoxin	*Desulphovibrio vulgaris*	148	26	2.5 Å *X*-ray map exists	820
Flavodoxin	*Peptostreptococcus elsdenii*	137	30		821
Glutamate dehydrogenase	Chicken liver	503	—		850
Glutamate dehydrogenase	Bovine liver	500	—		849
Gonadotrophin	Human	89—92 (α), 147 (β)	—	α- and β-subunits	796
Growth hormone	Bovine	189	—	2 forms – one minus *N*-terminal Ala	791, 792
Growth hormone	Equine	190	—		794
Growth hormone	Ovine	191	—		793
Haemoglobin	Capuchin Monkey	141 (α), 146 (β)	—		827
Haemoglobin	Chicken	146	—	AII component, β-chain	829
Haemoglobin	*Echidna*	—	—	α-chain	830
Haemoglobin	Mouse	146	—	β-chain, strains C_{57}BL, SWR, NB	831
Haemoglobin	Mouse	146	—	β-chain – 2 components, strain BALB/c	832
Haemoglobin	Savannah Monkey	141 (α), 146 (β)	—	α and β chains	826

Haemoglobin	Spider Monkey	141	—	α chain	828
Hb mutants	Beograd	—	—	β_{121} Glu → Val	865
Hb mutants	Camden	—	—	β_{131} Gln → Glu	866
Hb mutants	Daneshgah Tehran	—	—	α_2 72 His → Argβ_2^A	867
Hb mutants	Iran	—	—	β_{22} Glu → Gln	868
Hb mutants	F. Kuala Lumpur	—	—	$\alpha_2\gamma_2$ Asp^{22} → Gly; 136 Ala	869
Hb mutants	Little Rock	—	—	β_{143} His → Gln	870
Hb mutants	p-Nilotic	—	—	β-δ chain	834
Hb mutants	J. Nyanza	—	—	α_2 ^{21}Ala → β_2Asp	871
Hb mutants	Ouled Rabah	—	—	β_{19} Asn → Lys	868
Hb mutants	St. Antoine	—	—	Gly-Leu, β_{74-75} deleted	872
Hb mutants	St. Louis	—	—	β_{28} Leu → Gln	873
Hb mutants	Sawara	—	—	α_6 Asp → Ala	874
Hb mutants	Seattle	—	—	β_{70} Ala → Asp	833*b*
Hb mutants	Tours	—	—	Thr, β_{87} deleted	872
Hb mutants	Ube-1	—	—	β_{98} Val → Met	875
Hb mutants	G. Waimanolo	—	—	α_{64} Asp → Asn	876
Histone IIb_1	Rainbow Trout	125	47		835

[863] C. Nolan, W. M. Fitch, T. Uzzell, L. J. Weiss, and E. Margoliash, *Biochemistry*, 1973, **12**, 4052.
[864] R. P. Ambler, *Biochem. J.*, 1973, **135**, 751.
[865] G. D. Efremov, H. Duma, R. Ruvidii, K. Zolovic, J. B. Wilson, and T. J. H. Huisman, *Biochim. Biophys. Acta*, 1973, **328**, 81.
[866] P. T. W. Cohen, A. Yates, A. J. Bellingham, and E. R. Huehns, *Nature*, 1973, **243**, 467.
[867] S. Rahbar, G. Nowzari, and P. Danèshmand, *Nature New Biol.*, 1973, **245**, 268.
[868] J. Elion, O. Belkhodja, H. Wajeman, and D. Labie, *Biochim. Biophys. Acta*, 1973, **310**, 360.
[869] L.-I. L. Eng, B. G. Wiltshire, and H. Lehmann, *Biochim. Biophys. Acta*, 1973, **322**, 224.
[870] P. A. Bromberg, J. O. Alben, G. H. Bare, S. P. Balcerzak, R. T. Jones, B. Brimhall, and F. Padilla, *Nature New Biol.*, 1973, **243**, 177.
[871] D. G. Kendall, R. D. Bau, A. Lang, and H. Lehmann, *Biochim. Biophys. Acta*, 1973, **310**, 357.
[872] H. Wajeman, D. Labie, and G. Schapira, *Biochim. Biophys. Acta*, 1973, **295**, 495.
[873] M. Cohen-Solal, M. Seligmann, J. Thillet, and J. Rosa, *F.E.B.S. Letters*, 1973, **33**, 37.
[874] I. Sumida, Y. Ohta, T. Imamura, and T. Yanase, *Biochim. Biophys. Acta*, 1973, **322**, 23.
[875] Y. Ohba, T. Miyaji, and S. Shibata, *Nature New Biol.*, 1973, **243**, 205.
[876] R. Q. Blackwell, R. T. S. Jim, T. G. H. Tan, M.-I. Weng, C.-S. Liu, and C.-L. Wang, *Biochim. Biophys. Acta*, 1973, **322**, 27.

Table 6 (*cont.*)

Protein	*Source*	*No. of residues*	% *automated*	*Notes*	*Ref.*
Histone III	Calf thymus	135	—		838
Histone III	Carp testes	135	—	One difference from calf	839
Histone III	Pea embryo	135	—	Four differences from calf	840
Immunoglobulin heavy chains					
IgM-μ	Human	576	—	Protein Gal	783
IgM-μ	Human	576	—	Protein Ou	781
Immunoglobulin light chains					
Bence-Jones κ	Human amyloid	214	—		787
Bence-Jones Ni	Human	218	—		786
Bence-Jones Rei	Human	214	—	2.8 Å *X*-ray map exists	784
κ, Car	Human myeloma	214	—		788
κ	Rabbit	214	—	Anti-*p*-azobenzoate antibody	789
IgM, κ	Human	214	—		782
Insulin	Duck	51	—		877
Insulin C	Duck	26	—		878
β-Lactoglobulin	Bovine	162	100		879
Luteinizing hormone	Bovine	119	—	β-subunit	799
Luteinizing hormone	Human	—	—	β-subunit	798
Luteinizing hormone	Porcine	—	—	α- and β-subunits	799, 800
β_2-Microglobulin	Human serum	100	19	Automated peptide purification	854
Myoglobin	Galago	153	—		880
Myoglobin	Lemur	153	—		880
Myoglobin	Marmoset	153	—		881
Myoglobin	Mollusc	145	—		882
Myoglobin	Humboldt's Woolley Monkey	153	—		881
Myoglobin	Squirrel Monkey	153	—		881
Nerve growth factor	Mouse	118	—	2 forms differing by 8 residues at *N*-terminus	883, 884

Parvalbumin	Hake	108	—		752, 753
Parvalbumin	Pike	109	—		754
Parvalbumin	Carp	108	—	Peptides overlapped from 2 Å *X*-ray map	751
Pepsin	Hog	324	—		765
Peptidoglycan	*Streptomyces* R39	'3'	—	Cross-linked polymeric structure	812
Ribosomal protein A1, A2	*E. coli*	120	98	See last year's Report	885
Ribosomal protein S4	*E. coli*	203	>25		886
Ribonuclease	Ovine	124	—	Possible microheterogeneity	767
Ribonuclease	Rat	127	—		766
Ribonuclease	Red Deer	124	—		768
Ribonuclease	Roe Deer	124	—	Microheterogeneity at residue 64	768
Somatostatin	Ovine	14	—	Part determined by mass spectrometry, synthesized	803
Somatomammotropin	Human	191	—	Revision of previous work	790
Subtilin	*B. subtilis*	27	—	Pentacyclic peptide	810
Tentoxin	*Alternaria tenuis* Auct.	4	—	Part determined by mass spectrometry	811
Thynnin Z1	Tuna	34	—		841

[877] J. Markussen and F. Sundby, *Internat. J. Peptide Protein Res.*, 1973, **5**, 37.
[878] J. Markussen and F. Sundby, *European J. Biochem.*, 1973, **34**, 401.
[879] G. Braunitzer, R. Chen, B. Schrank, and A. Stangl, *Z. physiol. Chem.*, 1973, **354**, 867.
[880] A. E. R. Herrera and H. Lehmann, *Biochim. Biophys. Acta*, 1973, **322**, 10.
[881] A. E. R. Herrera and H. Lehmann, *Biochim. Biophys. Acta*, 1973, **317**, 65.
[882] L. Tentori, G. Vivaldi, S. Carta, M. Marinucci, A. Massa, E. Antonioni, and M. Brunori, *Internat. J. Peptide Protein Res.*, 1973, **5**, 187.
[883] R. H. Angeletti, D. Mercanti, and R. A. Bradshaw, *Biochemistry*, 1973, **12**, 90.
[884] R. H. Angeletti, M. A. Hermodson, and R. A. Bradshaw, *Biochemistry*, 1973, **12**, 100.
[885] C. Terhorst, W. Möller, R. Laursen, and B. Wittmann-Liebold, *European J. Biochem.*, 1973, **34** 138.
[886] J. Rheinbolt and E. Schiltz, *F.E.B.S. Letters*, 1973, **36**, 250.

Table 6 (*cont.*)

Protein	*Source*	*No. of residues*	% *amino-acid*	*Notes*	*Ref.*
Thynnin Z2	Tuna	34	—		842
Thyrotropin	Human	112	—	β-subunit only	795
Triose phosphate isomerase	Rabbit muscle	248	—		887
Troponin C	Rabbit muscle	158	—	Homologies with parvalbumins	756
Trypsin	Pig	223	100		761
Trypsin inhibitor (Kunitz)	Soybean	181	—		763
Tryptophan synthetase	*Aerobacter aerogenes*	269	55	Tentative sequence of α-chain	852
Tryptophan synthetase	*Salmonella typhimurium*	268	56	Tentative sequence of α-chain	851
Venom proteins	*Dendroaspis angusticeps*	61	52	4 S—S bridges	775
Venom proteins	*Dendroaspis jamesoni kaimosae*	60/72	—	Toxins $V_n^{I}I$ and $V_n^{III}I$	776
Venom proteins	*Dendroaspis polylepsis polylepsis*	61/60	—	Toxins 1 and κ	778
Venom proteins	*Haemachatus haemachates*	61	—		772
Venom proteins	*Naja naja*	71	—	5 S—S bridges	774
Venom proteins	*Naja haje annulifera*	60	70	Toxin $V^{II}I$	773
Venom proteins	*Ophiophagus hannah*	73	92	Toxins a and b	777
Venom proteins	*Vipera russelli*	60	—		779
Wool proteins					
$IIIA_3$	Merino sheep	131	37	High sulphur	758
0.62	Merino sheep	61	—	Tyrosine and glycine rich	757
IIIA3 A	Merino sheep	130	—	Minor component of IIIA3	758
Yellow mosaic virus protein	Turnip	189	—		888

[887] P. H. Corran and S. G. Waley, *F.E.B.S. Letters*, 1973, **30**, 97.
[888] D. Stehelin, R. Peter, and H. Duranton, *Biochim. Biophys. Acta*, 1973, **317**, 253.

one of the two polypeptide chains of papain-solubilized human transplantation (HL-A) antigens. Thus, this protein may be of considerable importance and the structure would be very useful for defining similar proteins in species where detailed functional and genetic studies can be carried out.

Finally, a list of complete sequences published during 1973 is given in Table 6.

PART II: *X*-Ray Studies *by C. Chothia*

1 Introduction

In writing this section,* the Reporter has assumed the reader to be moderately familiar with the results of previous years, and also feels bound to declare his interest (prejudice ?) by saying that he found the most significant results this year to be the simple elegance of the immunoglobin structures, the complexity of the cytochromes, and the similarity of the coenzyme binding parts of the glycolytic enzymes.

2 Amino-acids and Peptides

The crystal structure analyses of amino-acids and peptides [1–12] are listed in Table 1. The series of neutron diffraction structure analyses of amino-acids by Koetzle and his colleagues [7] is of particular use because hydrogen atoms are accurately located.

Two structure analyses of γ-aminobutyric acid, $^{+}H_3N-CH_2-CH_2-CH_2-CO_2^-$, have been described. In the zwitterion form the τ(C—C—C—C) torsion angle is *gauche* [2] and in the hydrochloride it is *trans*.[3] The two molecules in the asymmetric unit of L-methionine have different side-chain conformations.[9] In one molecule it is extended (χ_1 = 194, χ_2 = 174.2, χ_3 = 179.7°) and in the other it is folded back (χ_1 = 194.0, χ_2 =

[1] S. K. Arora, M. Sundarlingam, J. S. Danez, R. H. Stanford, and R. E. Marsh, *Acta Cryst.*, 1973, **B29**, 1849.
[2] E. G. Steward, R. B. Player, and D. Warner, *Acta Cryst.*, 1973, **B29**, 2038.
[3] E. G. Steward, R. B. Player, and D. Warner, *Acta Cryst.*, 1973, **B29**, 2825.
[4] S. T. Rao, *Acta Cryst.*, 1973, **B29**, 1718.
[5] K. A. Kerr and J. P. Ashmore, *Acta Cryst.*, 1973, **B29**, 2124.
[6] S. N. Rao, R. Parthasarathy, and F. E. Cole, *Acta Cryst.*, 1973, **B29**, 2373.
[7] T. F. Koetzle, M. S. Lehmann, and W. C. Hamilton, *Acta Cryst.*, 1973, **B29**, 231.
[8] E. Benedetti, C. Pedone, and A. Sirigu, *Acta Cryst.*, 1973, **B29**, 730.
[9] K. Torii and Y. Iitaka, *Acta Cryst.*, 1973, **B29**, 2799.
[10] S. N. Rao and R. Parthasarathy, *Acta Cryst.*, 1973, **B29**, 2379.
[11] R. E. Stenkamp and L. H. Jensen, *Acta Cryst.*, 1973, **B29**, 2872.
[12] I. L. Karle, J. Karle, T. Weiland, W. Burgermaster, H. Faulstick, and B. Witcop, *Proc. Nat. Acad. Sci. U.S.A.* 1973, **70**, 1836.

* The bibliography of this review was compiled from *Amino-acid, Peptide, and Protein Abstracts* and supplemented by a search through all 1973 issues of *Journal of Molecular Biology, Proceedings of the National Academy of Sciences, Biochemistry, Nature, Biopolymers, Acta Crystallographica, Journal of Biological Chemistry, Science, European Journal of Biochemistry, Macromolecules,* and *International Journal of Peptide and Protein Research.*

Table 1 *Amino-acid and peptide crystal structure analyses*, 1973

Amino-acid or peptide	*Unit cell*/Å	*Space group*	*R factor*	σ(C—C)/Å	*Ref.*
3-Aminobenzoic acid hydrochloride $C_7H_7NO_2HCl$	$a = 14.706$, $b = 5.836$, $c = 4.567$, $\alpha = 96.22°$, $\beta = 91.11°$, $\gamma = 98.62°$	$P\bar{1}$	4.3	0.004	1
γ-Aminobutyric acid $C_4H_9NO_2$	$a = 8.285$, $b = 10.224$, $c = 7.203$, $\beta = 110.76°$	$P2_1/a$	6.8	0.005	2
γ-Aminobutyric acid hydrochloride $C_4H_9NO_2HCl$	$a = 9.040$, $b = 6.578$, $c = 5.900$, $\beta = 99.82$	$P2_1$	10.5	0.02	3
DL-Aspartic acid $C_4H_7NO_4$	$a = 18.947$, $b = 7.433$, $c = 9.184$, $\beta = 123.75$	$C2/c$	3.4	0.002	4
L-Cysteine $C_3H_7NO_2S$	$a = 8.116$, $b = 12.185$, $c = 5.426$	$P2_12_12_1$	3.75	0.003	5
L-β,β′-Dimethylcysteine hydrochloride monohydrate $C_5H_{12}NB_2HCl,H_2O$	$a = 12.585$, $b = 6.129$, $c = 6.110$, $\beta = 107.35$	$P2_1$	7.0	0.007	6
4-Hydroxy-L-proline $C_5H_9NO_3$	$a = 4.995$, $b = 8.307$, $c = 14.195$	$P2_12_12_1$	3.2	0.002	7
DL-Isoleucine $C_6H_{13}NO_2$	$a = 14.66$, $b = 5.39$, $c = 5.27$, $\alpha = 109.2$, $\beta = 114.0$, $\gamma = 85.2$	$P\bar{1}$	15.0	—	8
L-Methionine $C_5H_{11}NO_2S$	$a = 9.498$, $b = 5.189$, $c = 15.318$, $\beta = 97.69$	$P2_1$	9.0	0.015	9
L-Norleucine $C_6H_{13}NO_2$	$a = 9.550$, $b = 5.260$, $c = 15.377$, $\beta = 95.60$	$C2$	6.0	0.007	9
Glycylglycine nitrate $C_4H_9N_2O_3,NO_3$	$a = 8.459$, $b = 9.909$, $c = 9.501$, $\beta = 90.46$	$P2_1/c$	4.4	0.002	10
N-Acetyl-L-phenylalanyl-L-tyrosine $C_{20}H_{22}N_2O_5$	$a = 11.530$, $b = 8.589$, $c = 10.635$, $\beta = 114.52$	$P2_1$	4.7	0.003	11
Li^+ antamanide	$a = 11.912$, $b = 23.206$, $c = 13.884$, $\beta = 110.75$	$P2_1$	10.3	—	12
Na^+[Phe⁴,Val⁶]-antamanide	—	$P2_12_12_1$	—	—	12

73.6, χ_3 = 73.6°). The peptide group in crystals of *N*-acetyl-phenylalanyl-tyrosine is twisted from its normal planar configuration [11] so the dihedral angle is 17.7° (ω = 162.3°).

Lithium Antamanide and Sodium-[Phe^4, Val^6]Antamanide Complexes.— Antamanide is a cyclic decapeptide isolated from the poisonous mushroom

Figure 1 *Diagram of the Li–antamanide complex. The phenyl rings of Phe-5 and Phe-10 are not shown*
(Reproduced by permission from *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1836)

Amanita phalloides, and forms complexes with Li^+, Na^+, K^+, and Ca^{2+} ions. Karle *et al.*[12] have described the crystal structures of Li^+-antamanide (1) and Na^+-[Phe^4, Val^6]antamanide (2) complexes.

Pro-Phe-Phe-Val-Pro	Pro-Phe-Phe-Val-Pro
| |	| |
Pro-Phe-Phe-Ala-Pro	Pro-Val-Phe-Phe-Pro
Li^+	Na^+
(1)	(2)

Crystals of the lithium complexes were grown from acetonitrile and those of the sodium complex from ethanol.

Both molecules have the compact 'tennis ball crease' configuration found in some other cyclic peptide metal complexes (Figure 1). This structure is facilitated by *cis* peptide bonds between residues Pro-2 and Pro-3 and between Pro-7 and Pro-8, and is stabilized by two main-chain hydrogen-bonds: Phe-6 NH···OC Pro-3 (3.05 Å) and Val-1 NH···OC Pro-8

(3.00 Å). The ion sits at the saddle point of the molecules where it bonds to four carbonyl oxygens (average bond lengths for Li···O, 2.1 Å). The ions are just above the plane of the oxygens, and are bonded to a fifth ligand, acetonitrile in case of lithium and ethanol in the case of sodium. The square-pyramid co-ordination of lithium has not been observed before.

3 Low-resolution Studies of Proteins

Fibrous Proteins.—The results of *X*-ray diffraction experiments on poly-amino-acids [13–17] are summarized in Table 2. Of particular interest are the studies that show sodium poly-L-glutamate and poly-L-ornithine hydrobromide going from a β-sheet structure at low humidity to an α-helical structure at high humidity.

Collagen. Previously, Miller and Wray [18] have argued that the structural unit of collagen is a five-stranded microfibril. Each strand is a triple helical tropocollagen molecule. These molecules form the microfibril by packing with a 670 Å axial stagger and a 108° azimuthal displacement between neighbours.

From a detailed examination of the *X*-ray diffraction given by rat tail collagen, Miller and Parry [19] have discussed in detail the structure and packing of microfibrils. The microfibril is supercoiled so the 5_1 symmetry of the uncoiled structure is converted into 4_1. The microfibrils pack so that they sit on the two-fold screw axis of a tetragonal cell symmetry *I*4, and are separated by 38 Å from their four neighbours. Neighbouring microfibrils are in azimuthal register with each other and have an axial displacement of zero or a multiple of 670 Å. The microfibrils fill about 39% of the unit-cell volume.

Katz and Li have calculated the intermolecular space in hydrated, purified, and reconstituted steer skin collagen using an *X*-ray diffraction technique.[20] They found the intermolecular space to be 1.14 ml gm^{-1} of collagen. Nemetschek and Hosemann [21] have followed the changes that occur in the *X*-ray pattern when dry collagen is hydrated.

α-Keratin. The proteins that form α-keratin can be divided into three groups on the basis of their amino-acid composition: high glycine and tyrosine, high sulphur, and low sulphur. The last group are believed to form the microfibrils in the structure [22] and to be surrounded by a matrix

[13] A. Del Pra, P. Spadon, and G. Valle, *Biopolymers*, 1973, **12**, 941.
[14] Y. Mitsui, *Biopolymers*, 1973, **12**, 1781.
[15] A. Del Pra, M. Mammi, and G. Peggion, *Biopolyners*, 1973, **12**, 937.
[16] M. Suwalsky and L. der la Hoz, *Biopolymers*, 1973, **12**, 1997.
[17] B. Lotz, F. Heitz, and G. Spach, *Compt. rend.*, 1973, **276**, *C*, 1715.
[18] A. Miller and J. S. Wray, *Nature*, 1971, **230**, 437.
[19] A. Miller and D. A. D. Parry, *J. Mol. Biol.*, 1973, **75**, 441.
[20] E. P. Katz and S. T. Li, *J. Mol. Biol.*, 1973, **73**, 351.
[21] T. Nemetschek and R. Hosemann, *Naturwiss.*, 1973, **60**, 304.
[22] E. Suzuki, W. C. Crewther, R. D. B. Fraser, T. P. MacRae, and N. M. McKern, *J. Mol. Biol.*, 1973, **73**, 275.

Table 2 *Structure analysis of polyamino-acids*, 1973

Polyamino-acid	*Water content*	*Unit cell*/Å	*Structure*	*Ref.*
Poly-N^{γ}-carbobenzoxy-L-$\alpha\gamma$-diaminobutyric acid	Dry	$a = b = 27.5$, $c = 27$, $\gamma = 120°$	α-helix (18 residues in 5 turns)	13
	Almost dry	$a = 7.0$, $b = 14.4$, $c = 4.7$	β-sheet	
Sodium poly-L-glutamate	64% Relative humidity	$a = b = 16.9$, $c = ?$ $\gamma = 120°$	α-helix	14
Poly-*im*-benzyl-L-histidine	Dry	$a \simeq c = 17.4$, $b = 27.0$, $\gamma = 120°$	α-helix (18 residues in 5 turns)	15
	Half to three H_2O per residue (0—76% relative humidity)	$a = 4.60$, $b = 30.2$, $c = 6.64$	β-sheet	
Poly-L-ornithine hydrobromide	Four to nine H_2O per residue (86—~100% relative humidity)	$a = b =$ *ca.* 15.0, $c = ?$ $\gamma = 120°$	α-helix	16
Poly-N^{δ}-carbobenzoxy-L-ornithine	Dry	$a = 23.2$, $b = 22.7$, $c = 16.2$, $\gamma = 119.2$	α-helix (11 residues in 3 turns)	13
Poly-benzyl-DL-glutamate	Dry	$a = 4.76$, $b = 22.5$, $c = 5.4$	α-helix (left- and right-handed)	17

composed of the first two.[23] The *X*-ray diffraction pattern of the low-sulphur protein[22] shows it to have a coiled-coil α-helical conformation, to be about 160 Å long, and to form aggregates >2000 Å long.

Low-angle *X*-Ray Scattering by Proteins in Solution.—The 1973 results for low-angle *X*-ray scattering experiments on proteins in solution[24–27] are summarized in Table 3. Pilz *et al.* examined the diffraction pattern of two anti-poly-D-alanine immunoglobins in the absence and in the presence of antigen.[24] They found the binding of the antigen gave a small but significant decrease in the volume of the antibodies, and they suggest this apparent contraction may be due to the antigens causing an improvement in the packing of less dense regions of the molecule, or decrease in the flexibility of the hinge regions. (The results of the high-resolution crystal structure analysis of Fab′ fragments of immunoglobins are described below.)

The 114 000 molecular weight of lysine tRNA synthetase given by low-angle *X*-ray scattering[26] is less than that (138 000) given by equilibrium sedimentation. The results from experiments on holo-pyruvate decarboxylase are summarized in Table 3. The apo-protein tended to aggregate in solution so its *X*-ray diffraction pattern was of limited use, but it did suggest that the apo structure was similar to the holo.[27]

To examine the relation between the crystal and solution structures of myoglobin, Stuhrmann has calculated from the atomic co-ordinates the small-angle *X*-ray diffraction pattern that would be given if the crystal and resolution structures were the same, and he has compared this calculated pattern with that observed.[28] From the similarity of the two patterns Stuhrmann concluded that the crystal and solution structures are the same, to within a resolution of 5 Å.

The Unit Cell and Symmetry of New Crystals of Globular Proteins.—The recent preliminary *X*-ray data for crystals of globular proteins are listed in Table 4. The crystallographic data show that the two monomers in streptomyces subtilisin inhibitor and in tyrosyl tRNA synthetase are related by a two-fold axis, that the three monomers in 2-keto-3-deoxy-6-phosphogluconic aldolase are related by a three-fold axis, and that the six monomers in canavalin are related by 32 symmetry. The orientation of the molecules in crystals of pig M_4 and H_4 lactate dehydrogenase was determined by the rotation function and their position in the unit cell determined empirically.

Low-resolution Studies of Globular Proteins.—*Aspartate Transcarbamylase* (*ATCase*). ATCase is an allosteric enzyme catalysing the first committed

[23] R. D. B. Fraser, J. M. Gillespie, and T. P. MacRae, *Comp. Biochem. Physiol.*, 1973, **44B**, 943.
[24] I. Pilz, O. Kratky, A. Licht, and M. Sela, *Biochemistry*, 1973, **12**, 4998.
[25] P. Stockel, A. Mayer, and R. Keller, *European J. Biochem.*, 1973, **37**, 193.
[26] R. Osterberg, B. Sjoberg, L. Rymo, and U. Lagerkvist, *J. Mol. Biol.*, 1973, **77**, 153.
[27] I. Pilz and J. Ullrich, *European J. Biochem.*, 1973, **34**, 256.
[28] H. B. Stuhrmann, *J. Mol. Biol.*, 1973, **77**, 363.

Table 3 *Low-angle* X-*ray analysis of proteins in solution*, 1973

Protein	Volume/Å^3	*Radius of gyration*/Å	*Molecular weight*	*Shape and dimensions*	*Specific volume*	*Hydration*/ gm H_2O per gm protein	*Ref.*
Anti-poly(D-alanyl) Immunoglobulin C_1							
(i) no antigen							
a	371 000	60 8	150 000	[T shape]	—	—	24
b	399 000	65.0	150 000	[T shape]	—	—	
(ii) with antigen							
a	359 000	60.0	150 000	[T shape]	—	—	
b	358 000	60.0	150 000	[T shape]	—	—	
Haemoglobin from Tubifin				Hollow cylinder height 156 Å, outer			
(i) oxygenated	5.6×10^6	110	—	radius 260 Å, inner	—	—	25
(ii) deoxygenated	5.8×10^6	112	[3.01×10^6]	radius 18 Å or hollow sphere	—	—	
Lysine tRNA synthetase	295 000	37.7	114 000	Ratio 1.25 : 0.47 $63 \times 50 \times 23$ Å	0.732	0.82	26
Pyruvate decarboxylase (holo)	440 000	41.3	185 000	Ratio 1 : 1 : 2.2 $35 \times 35 \times 78$ Å	0.751	0.55	27

Table 4 *Preliminary X-ray data for globular proteins*

Protein	*Source*	*Molecular weight*	*Unit cell*/Å	*Space group*	*Weight of protein in asymmetric unit*	*Resolution (Å) of observed X-ray pattern*	*Ref.*
Abrin (anti-tumour protein)	*Abrus precatorius*	65 000	$a = 74.9$, $b = 269.8$, $c = 70.3$	$P2_12_12_1$	130 000	6	29
Actinidine (a thiol protein)	*Actinidia chinensis*	26 000	$a = 78.1$, $b = 81.2$, $c = 33.0$	$P2_12_12_1$	36 000	2	30
2-Keto-3-deoxy-6-phospho-gluconic aldolase	*Pseuodomonas putida*	72 000	$a = 103.1$	$P2_13$	24 000	2.5	31
2-Amylase	*B. subtilis*	—	$a = 66$, $b = 85$, $c = 205$	$P22_12_1$ or $P2_12_12_1$	—	13	32
L-Asparaginase	*Proteus vulgaris*	120 000	$a = 69.3$, $b = 125.6$, $c = 101.4$, $\beta = 97°$	$P2_1$ or $P2$	120 000	3.4	33
Canavalin	Jack bean	113 000	$a = 82.4$, $\gamma = 111°$	$R32$	19 500	2.5	34
Carbonic anhydrase	Bovine erythrocyte	30 000	$a = b = 68$, $c = 244$	$P6_122$ or $P6_522$	30 000	3	35
Catalase	Beef liver	250 000	$a = 89$, $b = 140$, $c = 231$	$P2_12_12_1$	250 000	2.8	36
Creatine kinase	Rabbit muscle	80 000 (dimer)	$a = 54$, $b = 114$, $c = 145$, $\beta = 91°$	$P2_1$	160 000	3	37
Cytochrome *c*	Spinach	12 700	$a = 34.9$, $b = 80.7$, $c = 46.1$	$P2_12_12_1$	12 700	1.8	38
Polypeptide chain elongation factor (Tu)-GPD	*E. coli*	40 000	$a = 98.2$, $b = 100.1$, $c = 160.6$	$C222_1$	80 000	2.8	39
'Esteroproteolytic enzyme'	Bovine pancreas	32 000	i, $a = 59.2$, $b = 96.4$, $c = 47.4$	$P2_12_12$	*ca.* 30 000	—	40
			ii, $a = 59.2$, $b = 96.4$, $c = 94.8$	$P2_12_12_1$	*ca.* 60 000	—	

Histidinal dehydrogenase	*S. typhimurium*	80 000 (dimer)	$a = 149.6$, $b = 88.9$, $c = 105.6$, $\beta = 124.5°$	$C2$	80 000	3.5	41
$I_\alpha G$ myeloma protein	Human	150 000	$a = 170 \pm 5$, $\alpha = 109 \pm 2°$	$R32$	150 000	—	42
β-Lactamase	*E. coli*	22 000	$a = 47.2$, $b = 76.9$, $\beta = 98.6$	$C2$	22 000	2.4	43
Lactate dehydrogenase	Mouse testis	—	$a = 84$, $b = 76$, $c = 64$, $\alpha = 109.4$, $\beta = 90.6$, $\gamma = 96.8$	$P1$	—		44
Lactate dehydrogenase M4	Pig	144 000	$a = 86$, $b = 60$, $c = 136.2$	$P22_12_1$	72 000	—	45
Lactate dehydrogenase H4	Pig	144 000	$a = 163.4$, $b = 60.7$, $c = 138.6$, $\beta = 93$	$C2$	144 000	—	45

[29] Chin Hsnan Wei, *J. Biol. Chem.*, 1973, **248**, 3745.
[30] E. N. Baker, *J. Mol. Biol.*, 1973, **74**, 411.
[31] D. Tsernoglou, D. A. Walz, L. E. McCoy, and W. H. Seegers, *J. Biol. Chem.*, 1974, **249**, 999.
[32] A. McPherson and A. Rich, *J. Ultrastruct. Res.*, 1973, **44**, 75.
[33] B. Lee and H. J. Yang, *J. Biol. Chem.*, 1973, **248**, 7620.
[34] A. McPherson and A. Rich, *J. Biochem.*, 1975, **74**, 155.
[35] U. Carlsson, S. Lindskog, E. Andersson, E. Lindqvist, and G. Olsson, *J. Mol. Biol.*, 1973, **80**, 373.
[36] A. McPherson and A. Rich, *Arch. Biochem. Biophys.*, 1973, **157**, 23.
[37] A. McPherson, *J. Mol. Biol.*, 1973, **81**, 83.
[38] Y. Morita, F. Yagi, S. Ida, K. Asada, and T. Takahashi, *F.E.B.S. Letters*, 1973, **31**, 186.
[39] D. Sneden, D. L. Miller, S. H. Kim, and A. Rich, *Nature*, 1973, **241**, 530.
[40] L. M. Amzel, H. P. Avey, L. N. Becka, and R. J. Poljak, *J. Mol. Biol.*, 1973, **81**, 87.
[41] H. J. Yang, B. Lee, and J. L. Haslam, *J. Mol. Biol.*, 1973, **81**, 517.
[42] G. E. Connell, M. H. Freedmann, S. C. Nyburg, R. H. Painter, D. M. Parr, L. Pinterie, and W. Pruzanski, *Canad. J. Biochem.*, 1973, **51**, 1137.
[43] J. R. Knox, P. E. Zorksy, and N. S. Muothy, *J. Mol. Biol.*, 1973, **79**, 597.
[44] A. D. Adams, M. J. Adams, M. G. Rossmann, and E. Goldberg, *J. Mol. Biol.*, 1973, **78**, 721.
[45] M. L. Hackert, G. C. Ford, and M. G. Rossmann, *J. Mol. Biol.*, 1973, **78**, 665.

Table 4 (*cont.*)

Protein	*Source*	*Molecular weight*	*Unit cell*/Å	*Space group*	*Weight of protein in asymmetric unit*	*Resolution* (Å) *of observed X-ray pattern*	*Ref.*
Laticotoxin a	*Laticauda laticaudata*	6970	$a = b = 39.78$, $c = 72.74$	$P4_12_12$	6970	2.5	46
Lysozyme	*T4*	18 700	$a = b = 61.1$, $c = 96.3$	$P3_121$	18 700	2.0	47
6-Phosphogluconate dehydrogenase	Sheep liver	94 000	$a = 72.7$, $b = 149.0$, $c = 103.8$	$C222_1$	47 000	—	48
Ricin (anti-tumour protein)	*Ricinus communis*	65 000	$a = 117.3$, $b = 57.3$, $c = 92.2$, $\beta = 105° 40'$	$C2$	65 000	3	29
Subtilisin inhibitor	*Streptomyces alloguseolus*	27 000	$a = b = 40.8$, $c = 116.7$	$P3_121$	14 000	3.0	49
Subtilisin + subtilisin inhibitor complex	*Streptomyces alloguseolus*	41 000	$a = 77.5$, $b = 186.9$, $c = 69.6$	$I222$ or $I2_12_12_1$	41 000	2.5	49
Thrombin		38 000	i, $a = b = 87$, $c = 101$ ii, $a = b = 88$, $c = 98$	$P42_12$	38 000	2.5	50
Tyrosyl–tRNA synthetase	*B. stearothermophilus*	90 000	$a = 644$, $c = 238$	$P3_121$	45 000	2.7	51

[46] J. E. Seare, W. W. Fullerton, and B. W. Low, *J. Biol. Chem.*, 1973, **248**, 6057.
[47] B. W. Matthews, F. W. Dahlquist, and A. Y. Maynard, *J. Mol. Biol.*, 1973, **78**, 575.
[48] M. J. Adams, quoted by M. Silverberg and K. Dalziel, *European J. Biochem.*, 1973, **38**, 229.
[49] Y. Satow, Y. Mitsui, Y. Iitaka, S. Murao, and S. Sato, *J. Mol. Biol.*, 1973, **75**, 745.
[50] R. L. Vandlen, D. L. Ersfeld, A. Tulinsky, and W. A. Wood, *J. Biol. Chem.*, 1973, **248**, 2251.
[51] B. R. Reid, G. L. E. Koch, Y. Boulanger, B. S. Hartley, and D. M. Blow, *J. Mol. Biol.*, 1973, **80**, 199.

step in pyrimidine biosynthesis. The molecule has twelve sub-units, six catalytic and six regulatory. The regulatory sub-units first associate as dimers and in the molecule three dimers form a ring. The six catalytic sub-units form two trimers one of which sits on top of the ring of regulatory sub-units and the other below. This arrangement gives the molecule 32 symmetry. In two related papers[52, 53] Lipscomb and his colleagues have described some of the details of its structure as seen from a 5.5 Å resolution electron-density map.

The molecular dimensions are 110 × 110 × 90 Å. At its centre, there is an aqueous cavity 50 × 50 × 25 Å. The active site has been labelled with 2-chloromercuri-4-nitrophenol and the position of the mercury atom in the electron-density map shows that the active site is only accessible from the internal cavity. The main channels between the solvent and the cavity occur between the regulatory sub-units and are about 15 Å in diameter. The authors point out that the regulatory sub-units could function by steric control of access to the active site or by changing the environment of the central cavity.

Hexokinase. Hexokinase catalyses the first step in the metabolism of glucose which binds with positive co-operativity. Steitz, Fletterick, and Hwang discuss the structure seen in a 5 Å resolution electron-density map.[54] The crystal asymmetric unit contains a dimer whose molecular weight is *ca.* 102 000. Each sub-unit is folded into two distinct lobes with a narrow connection. The contacts between the two sub-units of the dimer are unusual. One monomer is related to the other by a 180° rotation and a 3.6 Å translation. Using Steitz's terminology the contact areas are not isologous (like with like) but heterologous (like with unlike).

Adenylate Kinase. This enzyme catalyses the reaction ATP + AMP $\rightleftarrows$ 2ADP. An electron-density map at 6.0 Å resolution is described by Schultz *et al.*[55] This shows the molecule is an oblate ellipsoid, 40 × 40 × 30 Å. Appreciable amounts of α-helix are apparent in the structure. There is a large cleft in the molecule, 10 Å deep, in which the active site is probably situated.

Alkaline Phosphatase. This enzyme catalyses the hydrolysis of phosphate esters. Knox and Wyckoff have reported a 7.7 Å resolution crystallographic study.[56] The molecule is a dimer with the sub-units related by a crystallographic two-fold axis. Each monomer contains a zinc ion and its removal by edta is accompanied by a shift of 8 Å by a small region of electron density 10 Å from the zinc.

[52] D. R. Evans, S. G. Warren, B. F. P. Edwards, C. H. McMurry, P. H. Bethge, D. C. Wiley, and W. N. Lipscomb, *Science*, 1973, **179**, 683.
[53] S. G. Warren, B. F. P. Edwards, D. R. Evans, D. C. Wiley, and W. N. Lipscomb, *Proc. Nat. Acad. Sci. U.S.A.* 1973, **70**, 1117.
[54] T. A. Steitz, R. J. Fletterick, and K. J. Hwang, *J. Mol. Biol.*, 1973, **78**, 551.
[55] G. E. Schultz, K. Biedermann, W. Kabsch, and R. H. Schirmer, *J. Mol. Biol.*, 1973, **80**, 857.
[56] J. R. Knox and H. W. Wyckoff, *J. Mol. Biol.*, 1973, **74**, 533.

4 High-resolution Studies of Globular Proteins

Immunoglobins.—One of the most exciting results in protein crystallography is the description of the molecular structure of immunoglobin fragments.[57–59] The simplest immunoglobins consist of two identical light chains (molecular weight 25 000) and two identical heavy chains (molecular weight 50 000) linked by disulphide bridges. On the basis of sequence homologies it was suggested that the chains consist of a series of structurally similar domains, two in the light chains V_L and C_L, and four in the heavy V_H, C_H1, C_H2, C_H3. The *V* domains are distinguished from the *C* by their having peptides whose amino-acid composition varies greatly from one immunoglobin to another. These variable regions are believed to provide the structural basis of antibody specificity.

Immunoglobulins can be fragmented by proteolytic enzymes. One such fragment is Fab′ which consists of the light chains and the first two heavy chain domains V_H and C_H1. Certain multiple myelomas cause dimers of light chains (Bence-Jones proteins) to be excreted in urine. In 1973 Poljak *et al.* described the molecular structure of a Fab′ fragment of a human immunoglobin,[57] Padlan *et al.* described a low-resolution study of the binding of an antigen to another Fab′ fragment,[58] and Schiffer, Girling, Ely, and Edmundson described the molecular structure of a λ-type Bence-Jones protein.[59] This last group has published the description of a heavy-atom derivative in which a mercury atom is inserted into an intersub-unit disulphide bridge so the link becomes sub-unit—S—Hg—S—sub-unit.[59] The reactions which gave the substitution were carried out with the protein in the crystal. This derivative is isomorphous with the native protein and lasted 1.5—2 times as long in the *X*-ray beam as did the other useful derivatives.

A 3.5 Å electron-density map clearly showed the fold of the two light chains in the Bence-Jones protein.[59] The *V* region (111 residues) and the *C* region (105 residues) of each chain appear as distinct domains separated by a short run of extended polypeptide chain. The domains have a double β-sheet structure: one sheet composed of four antiparallel strands covers a second composed of three antiparallel strands (Figure 1a). At the centre of each domain the sheets are linked by a disulphide bridge.

The description of the Fab′ fragment by Poljak *et al.* is based on a 2.8 Å resolution electron-density map.[57] The molecule has the dimensions 80 × 50 × 50 Å and the V_L, C_L, V_H, and C_H1 regions form similar distinct domains *ca.* 40 × 25 × 25 Å. The two domains in each chain are

[57] R. J. Poljak, L. M. Amzel, H. P. Avey, B. L. Chen, R. P. Phizackerley, and F. Said, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3305.

[58] E. A. Padlan, D. M. Segal, T. F. Spande, D. R. Davis, S. Rudikoff, and M. Potter, *Nature New Biol.*, 1973, **245**, 165.

[59] K. R. Ely, R. L. Girling, M. Schiffer, D. E. Cunningham, and A. B. Edmundson, *Biochemistry*, 1973, **12**, 4233; M. Schiffer, R. L. Girling, K. R. Ely, and A. B. Edmundson, *ibid.*, 1973, **12**, 4620.

connected by a short extended polypeptide. The V_H region has extensive contact with V_L as C_L does with C_H1. All four domains have a similar simple structure in which the polypeptide runs back and forth parallel to its long axis (Figure 2). The *V* domains have one run which does not occur in the *C* regions. Different classes of immunoglobins are defined in part by the position of the disulphide bridges that link the heavy (H) and light (L) chains. In human $\gamma1$ the bridge is between 214(L) and 214(H). In human $\gamma2$, $\gamma3$, $\gamma4$, and μ, rabbit γ and $\gamma2a$, mouse $\gamma2a$ and $\gamma2b$, and

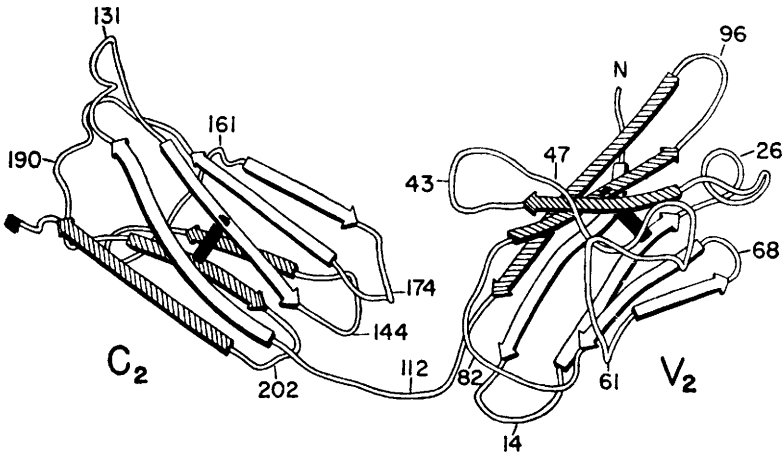


Figure 1a *λ-Type Bence-Jones protein: schematic drawing of monomer* 2. *The arrows indicate chain directions but not chain lengths*
(Reproduced by permission from *Biochemistry*, 1973, **12**, 4620)

guinea-pig $\gamma2a$ the bridge is between 131(H) and 214(L). Intrachain disulphide bridges are found in rabbit IGγ (131—221, H chain) and human $\gamma1$ Daw (35—101, H chain). In this human $\gamma1$ Fab′ fragment only the 214(H)—214(L) bridge is observed, but in all cases residues equivalent to those that form bridges in other classes of immunoglobins are within 6 Å. This implies that the structure of the Fab′ part of these other immunoglobins is very similar to that observed here for $\gamma1$.

The hypervariable peptides occur in some of the loops connecting the parallel runs of the protein chain. The loops from the *V* and the V_L domains are clustered together at one end of the molecule to form a shallow groove 15 × 6 Å in cross-section and 6 Å deep. The 'lower' part of the groove is formed by residues 55—65(H) and 30—33(H), the 'upper' part by 27a—30(L) and 49—52(L), the 'left' side by 90—95(L) and the 'right' side by 102—107(H).

Padlan *et al.* have described the structure of a Fab′-antigen complex as seen in a 4.5 Å resolution electron-density map.[58] This Fab′ fragment of

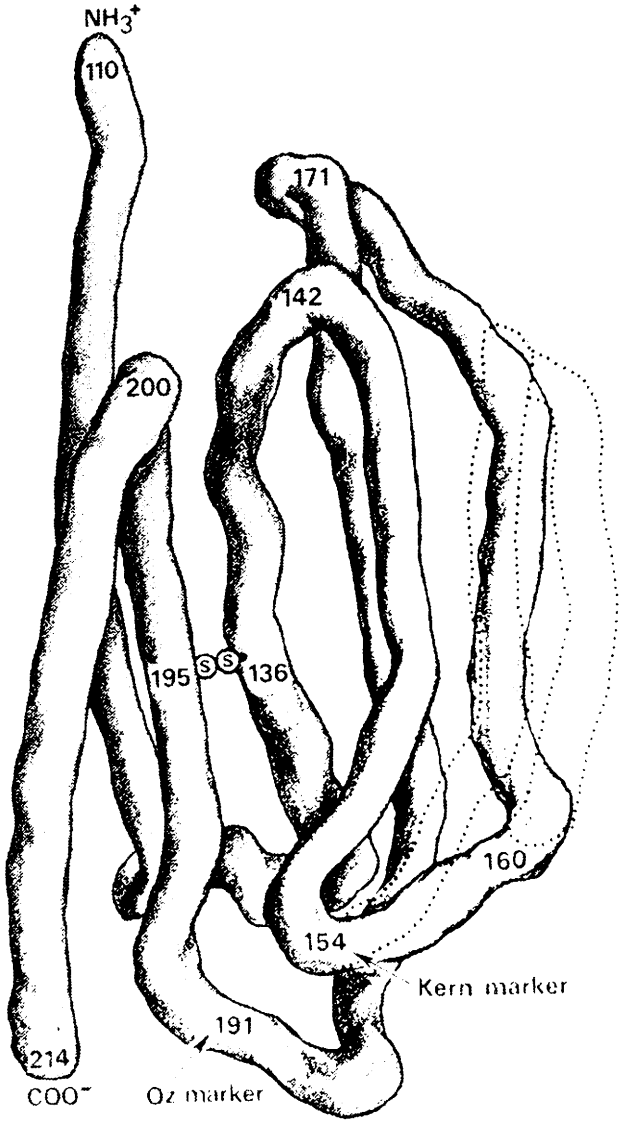
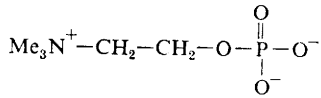


Figure 2 *Diagram of the basic 'immunoglobulin fold'. Solid trace shows the folding of the polypeptide chain in the constant domains* (C_L *and* C_H1). *Broken lines indicate the additional loop of polypeptide characteristic of the* V_L *and* V_H *domains*

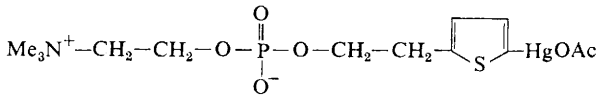
(Reproduced by permission from *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3305)

IgA(K) MOPC 603 binds phosphorylcholine (3) in solution with an association constant of 1.7×10^{-5}. The appearance of this structure at 4.5 Å resolution is very similar to the Fab′ fragment described by Poljak *et al.*



(3)

Its overall dimensions are 40 × 50 × 80 Å. There are two distinct globular units, one composed of the two variable domains (V_H and V_L) and the other composed of the constant domains (C_H1 and C_L). The antibody binding site has been found using 2-(5′-acetoxymercury-2′-thienyl)ethyl phosphorylcholine (4) (AMTEPC). Difference Fourier maps show that



(4)

this analogue binds to two sites in the protein; at one in the *C* region and at the end of the molecule in a crevice between the two domains of the *V* region. *p*-Nitrophenyl phosphorylcholine competes with AMTEPC for this second site (but not for the first) indicating it to be the antigen binding site. Only a small part of the crevice between the variable domains is used to bind AMTEPC. The difference map used to identify the binding site shows no other significant peaks, implying that there are no significant conformational changes by the antibody when the antigen binds. This result can be related to the low-angle scattering studies of immunoglobins in solution [24] described above.

Electron Transport Proteins.—*Structure of Cytochromes c_2 and c_{550}.* Different forms of cytochrome *c* transfer electrons in the respiratory and photosynthetic pathways and in chemical synthesis. Timkovich and Dickerson [60] have pointed out that a great variety of evidence suggests that all cytochrome *c* molecules of the eukaryotic–mitochondrial respiratory pathway have the same molecular structure, but the molecular weights (82—137 residues) and different amino-acid sequences of prokaryotic cytochromes *c* mean that their structural relation to each other and to the eukaryotic form is not clear. The crystal structure of two bacterial cytochromes *c* are now known.

Salemme *et al.* have described, from a 2 Å resolution electron-density map, the atomic structure of cytochrome c_2 from the bacterium *Rhodo-*

[60] R. Timkovich and R. E. Dickerson, *J. Mol. Biol.*, 1973, **79**, 39.

spirillum rubrum.[61–63] This protein is part of the photosynthetic pathway functioning as an electron donor to bacterochlorophyll. It has 112 residues (cytochrome *c* has 104) and an ellipsoidal shape 25 × 33 × 40 Å. The fold of the polypeptide chain and the environment of the haem are very similar to those found for eukaryotic cytochrome *c*. There are four helical

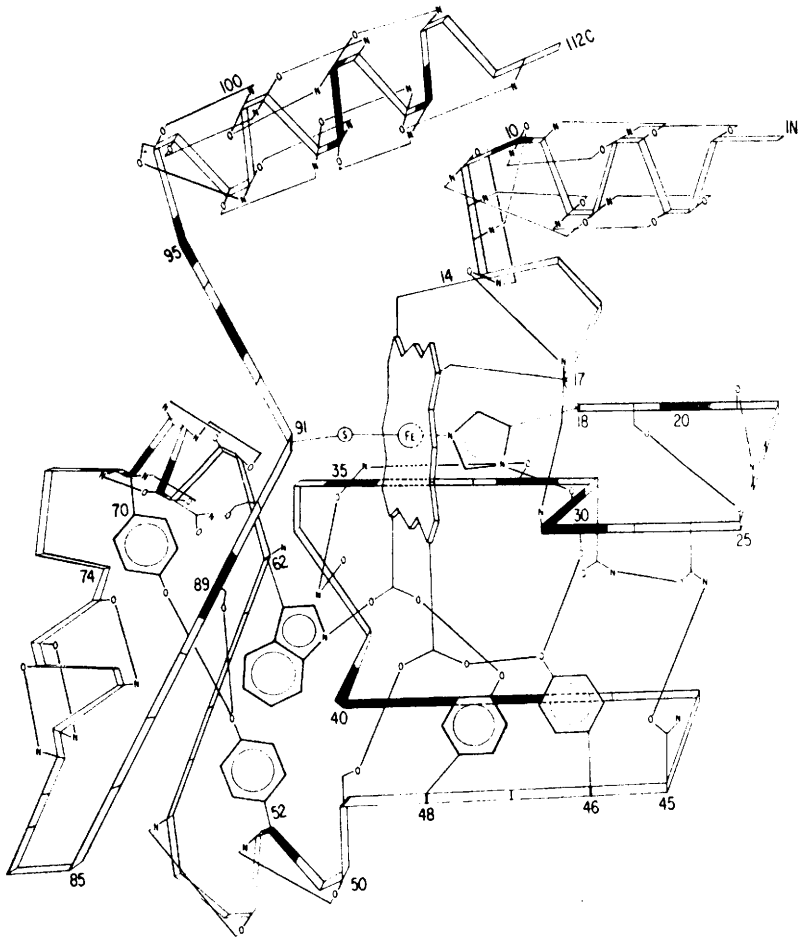


Figure 3 *A schematic drawing of cytochrome c_2 molecule showing hydrogen-bonded interactions. Residues in contact with the haem are shaded*
(Reproduced by permission from *J. Biol. Chem.*, 1973, **248**, 3910)

[61] F. R. Salemme, S. T. Freer, Ng. H. Xuong, R. A. Alden, and J. Kraut, *J. Biol. Chem.*, 1973, **248**, 3910.
[62] F. R. Salemme, S. T. Freer, S. T. Alden, and J. Kraut, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 47.
[63] F. R. Salemme, J. Kraut, and M. D. Kaman, *J. Biol. Chem.*, 1973, **248**, 7701.

regions: residues 2—10, 64—71, 75—80, and 96—110, and single or double β bends occur in the regions 10—18, 21—24, 32—38, and 50—57 (Figure 3).

The primary structure of the c_2 protein differs from that of *c* by the deletion of residues equivalent to 9, 10, 84, 100, and 101 in *c*, and the insertion of one residue at the *N*-terminal, one after 13, three after 52, and eight after 78 (*c* numbering). The eight-residue insertion forms a loop on the surface of the protein. The deletions and insertions do not prevent the c_2 chain from forming contacts with the haem and between residues very similar to those formed by homologous residues in *c*.

Timkovich and Dickerson have described the respiratory cytochrome c_{550} from *Micrococcus denitrificans*.[60] A 4 Å resolution electron-density map allowed many parts of the chain to be traced unambiguously. When distinct chains coalesced, a pathway similar to that in cytochrome *c* could be found; alternative tracings did not make sense of the map.

Though cytochrome c_{550} has 137 residues whereas horse and tuna cytochrome *c* have 104, the fold of the polypeptide chain of all three is essentially the same (Figure 4). At 4.0 Å resolution the 33 residues found in cytochrome c_{550} but not in cytochrome *c* can be assigned to six insertions in the *c* structure: five residues added at the amino-terminal end and five added at the carbonyl end; five inserted after residue 87 (*c* numbering); ten residues added to the loop 18—29; four residues inserted after residue 52; and eight inserted after residue 77. These last two insertions also occur in cytochrome c_{550}.

Timkovich and Dickerson point out that the similar structure of cytochromes *c*, c_2, and c_{550} imply that the respiratory and photosynthetic pathways developed from a common ancestor or the respiratory pathway developed from the photosynthetic.

Conformational Changes on Oxidation or Reduction of Cytochrome c. The reduced and oxidized forms of bonito heart cytochrome *c* and of cytochrome c_2 have the same conformation, but the conformation of the reduced form of tuna heart cytochrome *c* is not the same as oxidized horse heart cytochrome *c*.

Dickerson *et al.* carried out the crystal structure analysis of oxidized horse heart cytochrome *c*. Reduction of the oxidized protein in crystals destroyed the crystal structure and they were not able to crystallize the reduced horse protein, so to find the conformation of reduced cytochrome *c* they used crystals of reduced tuna heart cytochrome *c*. Takano *et al.* have described the structure of this molecule and compared it with that of the oxidized form.[64] The fold of the polypeptide chain is very similar in the two structures. The major conformational differences are:

(i) In the oxidized form there is a narrow channel to the haem bounded by the initial and final helices and the extended piece of chain formed by residues 14—22. In the reduced form the loop 18—28 blocks this channel.

[64] T. Takano, O. B. Kallai, R. Swanson, and R. E. Dickerson, *J. Biol. Chem.*, 1973, **248**, 5234.

(ii) The haem in both the reduced and oxidized forms sits in a crevice with one edge exposed to the solvent. In the oxidized form, part of the crevice is also exposed to the solvent, but on reduction residues 80—85 move towards the haem so the ring of Phe-82, which sticks out into the

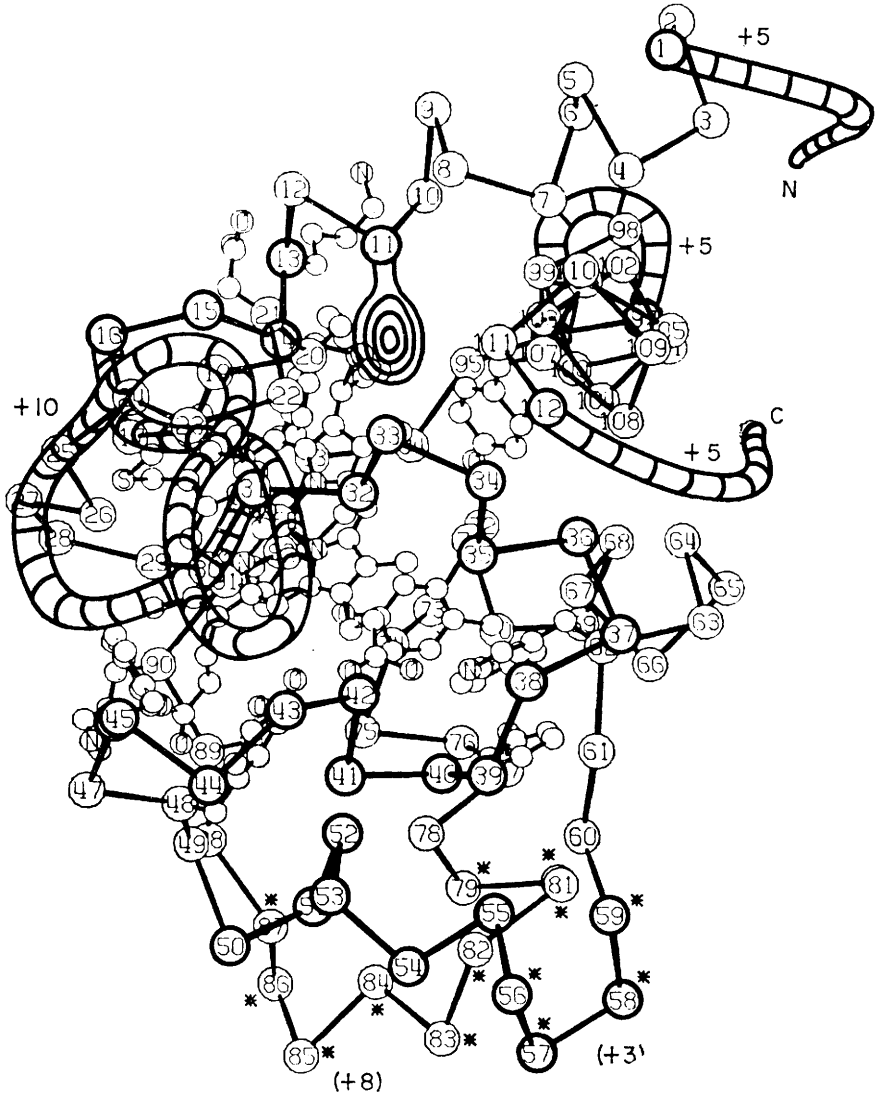


Figure 4 *The α-carbon skeleton of cytochrome c_{550}. Starred residues are present in c_2 but not in c. 'Tubular' residues are only present in c_{550}*
(Reproduced by permission from *J. Mol. Biol.*, 1973, **79**, 39)

solvent in the oxidized form, is parallel and in contact with the haem thus completely blocking the crevice.

(iii) Trp-59 and Tyr-48 are hydrogen-bonded to the buried haem propionate group in both forms. In the reduced form Trp-59 is closer to the haem and Tyr-48 further away than in the oxidized form.

In two preliminary notes Tsukihara, Yamane, Tanaka, Ashida, and Kakudo have described the structure of reduced and oxidized bonito heart cytochrome *c*.[65, 66] The reduced structure is from the interpretation of a 2.3 Å resolution electron-density map and the note implies that it is the same as reduced tuna heart cytochrome *c*. This group was able to oxidize the reduced crystals without any change in the unit cell dimensions. A difference Fourier between the reduced and oxidixed forms showed that the only structural difference was that Lys-99 bound a sulphate ion in the oxidized form. The reason for the different results of the two groups is unclear. Tsukihara *et al.* suggest that their oxidized structure may be a quasi-stable intermediate.[65]

Salemme *et al.* have found that the part of oxidized cytochrome c_2 that is homologous with cytochrome *c* folds in the manner observed for reduced cytochrome *c*.[61, 63] Crystals of the reduced form are isomorphous with those of the oxidized form and crystals of the oxidized form may be reduced with no change in the crystal lattice. A difference Fourier between the reduced and oxidized crystals implies that there are no large conformational differences between the two forms.[63]

These different results for the conformational effect of reduction and oxidation on cytochrome *c* and their mechanistic implications will probably be resolved by refinement of the protein structures. However, Takano *et al.* and Salemme *et al.* have tried to relate their structures to the solution chemistry and have suggested mechanisms for the function of the protein. The first group suggests[64] that reduction takes place *via* the transfer of an electron through the aromatic rings of Tyr-74, Trp-59, and Tyr-67 to the haem and the iron. The second group suggests[63] that the change is transferred to the iron *via* the sulphur of Met-80 which participates in a 'charge relay system' in which hydrogens are transferred between Tyr-70, Tyr-52, and Ser-89.

Bacterial Ferrodoxin. Ferrodoxins transfer electrons in anaerobic fermentative metabolism and in nitrogen fixation. They are small, *ca.* 54 residues, and contain eight iron atoms and eight inorganic sulphur atoms.

The structure of the ferrodoxin from *Peptococcus aerogenes* has been described by Adman, Sieker, and Jensen from the interpretation of a 2.8 Å resolution electron density map.[67] The molecule is a prolate ellipsoid

[65] T. Ashida, N. Tanaka, T. Yamane, T. Tsukihara, and M. Kakudo, *J. Biochem.*, 1973, **73**, 463.

[66] T. Tsukihara, T. Yamane, N. Tanaka, T. Ashida, and M. Kakudo, *J. Biochem.*, 1973, **73**, 1163.

[67] E. T. Adman, L. C. Sieker, and L. H. Jensen, *J. Biol. Chem.*, 1973, **248**, 3987.

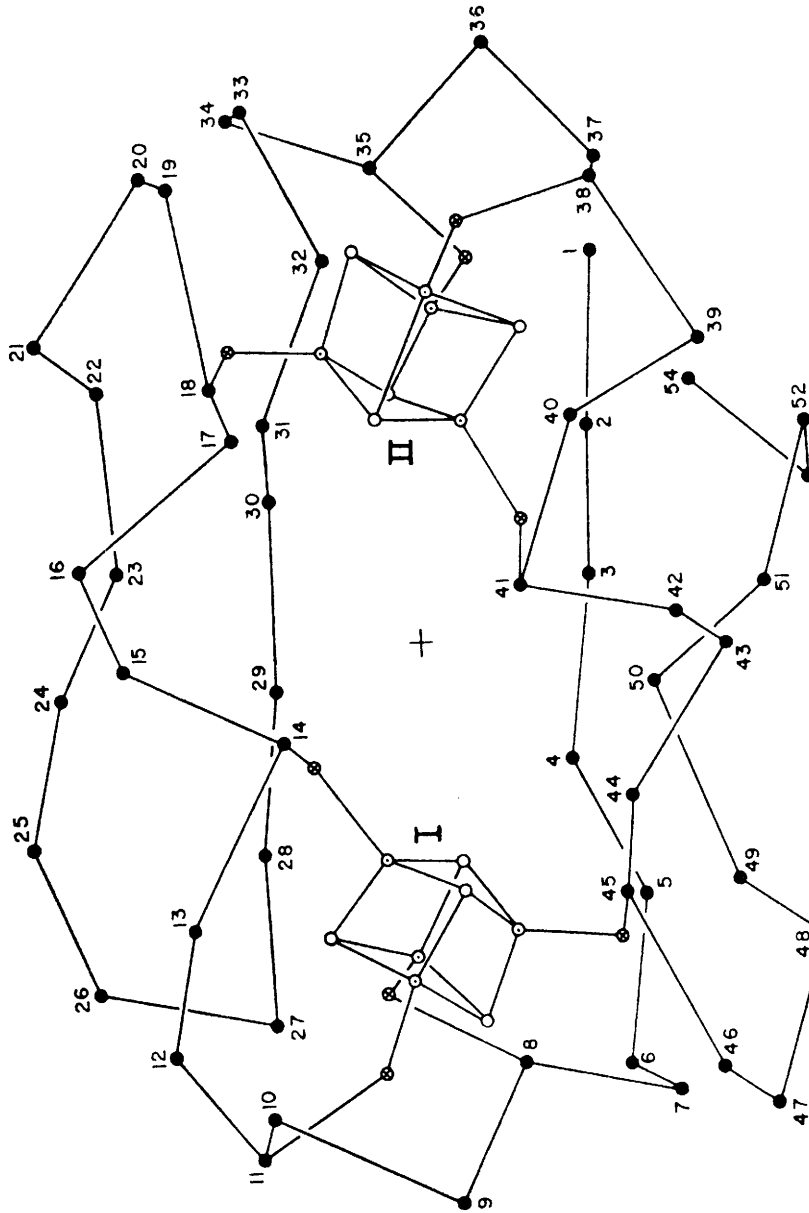


Figure 5 *Ferredoxin: Plot of α-carbon (black circles), iron (dotted circles), and sulphur (open and crossed circles) positions*
(Reproduced by permission from *J. Biol. Chem.*, 1973, **248**, 3987)

22 × 22 × 27 Å. The protein chain runs back and forth across the length of the molecule. There are two iron–sulphur complexes separated by 12 Å. Both are situated in the interior of the molecule and are completely shielded from the solvent by the protein (Figure 5). Each complex is composed of four iron atoms, four inorganic sulphur atoms, and four sulphurs from protein cysteines. The four iron and four inorganic sulphur atoms sit on the corners of a distorted cube so each iron is bonded to three inorganic sulphur atoms and *vice versa*. The two cubes are linked to the protein by each iron bonding to a cysteine sulphur. The molecule has approximately two-fold symmetry (Figure 5) and this is reflected by the cysteines bonding to the clusters, Cys 8, 11, 14, and 45 to one, and Cys 18, 35, 38, and 41 to the other, and by some of the residues forming their environment. The tyrosine in contact with each cluster has its aromatic ring parallel to one of the faces of the iron–sulphur cube and the edge of the ring exposed to the solvent, so could be responsible for the transfer of electrons to and from the buried complex.

Rubredoxin. Two papers have appeared on the crystal structure analysis of this protein. In one, the technique for the unconstrained refinement of the atomic co-ordinates using 1.5 Å resolution *X*-ray data is discussed.[68] The final conventional *R* factor is 12.7%. Two particularly interesting results of this work are the determination of the temperature factors for individual atoms and the location of 130 water molecules. The second paper discusses the identification of residues in electron-density maps of proteins whose sequence is unknown.[69] In the case of rubredoxin, 40 of the 54 residues were correctly identified from the 2.0 Å resolution electron-density map.

Bacterial Flavodoxin. This is an electron transport protein, of 149 residues and with a riboflavin-5′-phosphate (FMN) prosthetic group, that is capable of replacing ferrodoxin in metabolism and nitrogen fixation. The structure of the protein had been described previously by Watenpaugh, Sieker, and Jensen, and they have now discussed the environment of the FMN group.[70] The FMN group is deeply embedded in the protein forming an extensive series of hydrogen-bonds (Figure 6) and having little contact with the solvent. The flavin ring is sandwiched between Trp-60 and Tyr-98. The plane of the tyrosine ring makes a 15° angle, and the plane of the tryptophan a 45° angle, with the flavin ring. Only the methyl groups on the end of the flavin ring are accessible to solvent.

Carp Muscle Calcium-binding Protein (Myogen).—Kretsinger and Nockolds have given a detailed description of the structure of this protein.[71] The

[68] K. D. Watenpaugh, L. C. Sieker, J. R. Herriott, and L. H. Jensen, *Acta Cryst.*, 1973, **B29**, 943.
[69] J. R. Herriott, K. D. Watenpaugh, L. C. Sieker, and L. H. Jensen, *J. Mol. Biol.*, 1973, **80**, 423.
[70] K. D. Watenpaugh, L. C. Sieker, and L. H. Jensen, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3857.
[71] R. H. Kretsinger and C. E. Nockolds, *J. Biol. Chem.*, 1973, **248**, 3313.

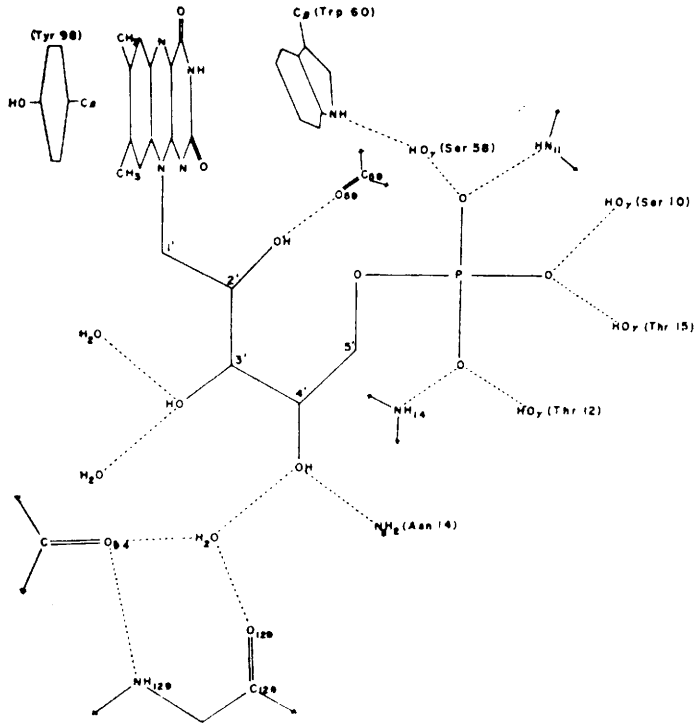


Figure 6a *Flavodoxin: hydrogen bonds between protein and flavin ring* (Reproduced by permission from *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3857)

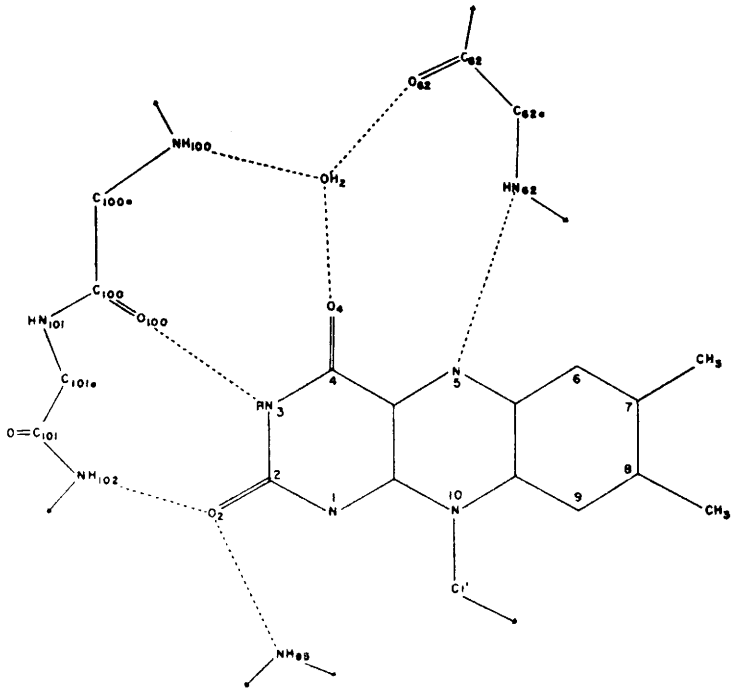


Figure 6b *Hydrogen bonds between protein and ribityl phosphate part of FMN*
(Reproduced by permission from *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3857)

paper contains many interesting details but only a few of them can be mentioned here. The protein has 108 residues and binds two calcium ions. Its approximate shape is an ellipsoid 30 × 30 × 36 Å. The chain folds to form six helices, A (residues 8—18), B (26—33), C (40—51), D (60—70),

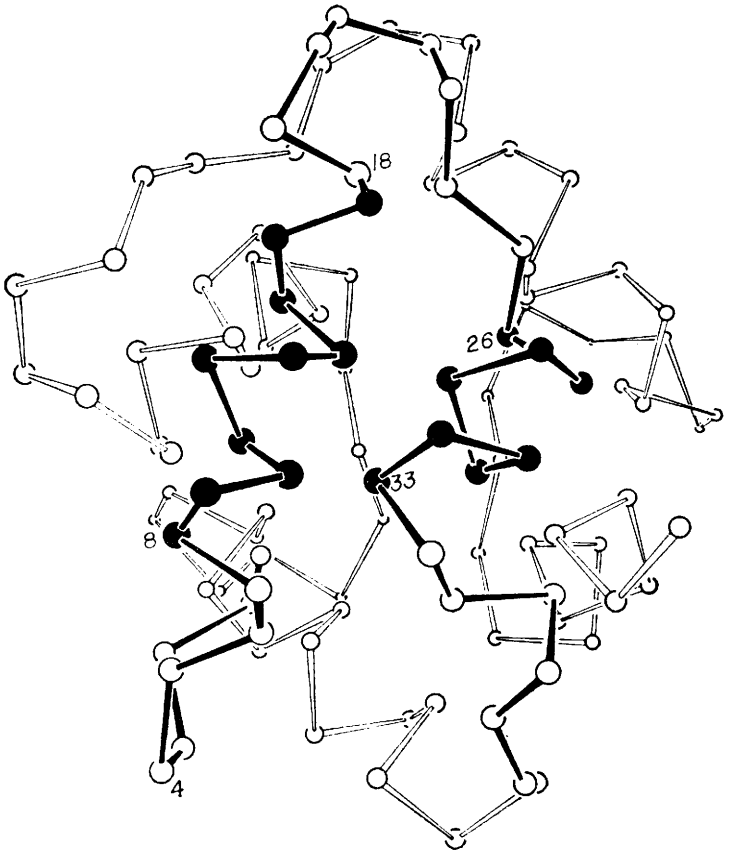


Figure 7a *Calcium binding protein: Plot of α-carbon positions showing the approximate two-fold symmetry (the A helix, AB bend, and B helix are indicated by solid lines)*
(Reproduced by permission from *J. Biol. Chem.*, 1973, **248**, 3313)

E (79—89), and F (99—107). The structures of the regions A–B, C–D, and E–F are similar and suggest that the molecule was formed by gene triplication (Figure 7). The positions of helices C and D and the loop between them, CD, are related by an approximate two-fold axis to helices E and F and the loop EF. The A–AB–B region sits above the rest of the structure with the A helix running antiparallel to the B so the whole

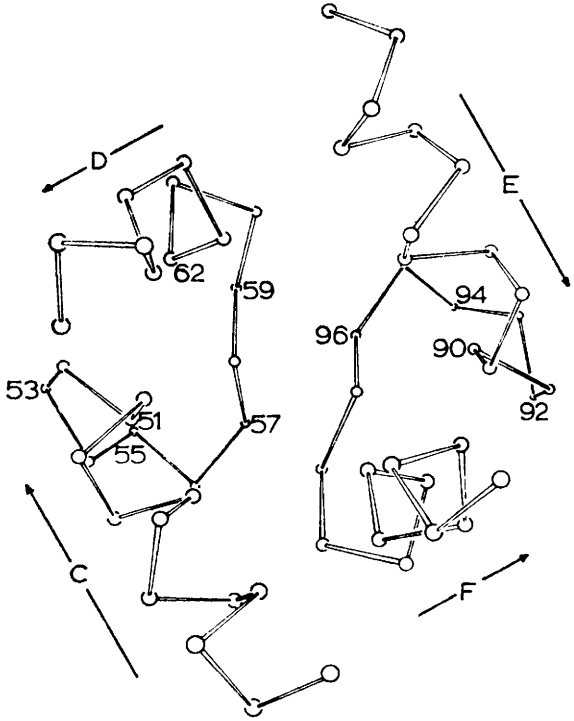


Figure 7b *Calcium binding protein: the two-fold relationship of the C, CD, Dregion to the E, EF, F region*
(Reproduced by permission from *J. Biol. Chem.*, 1973, **248**, 3313)

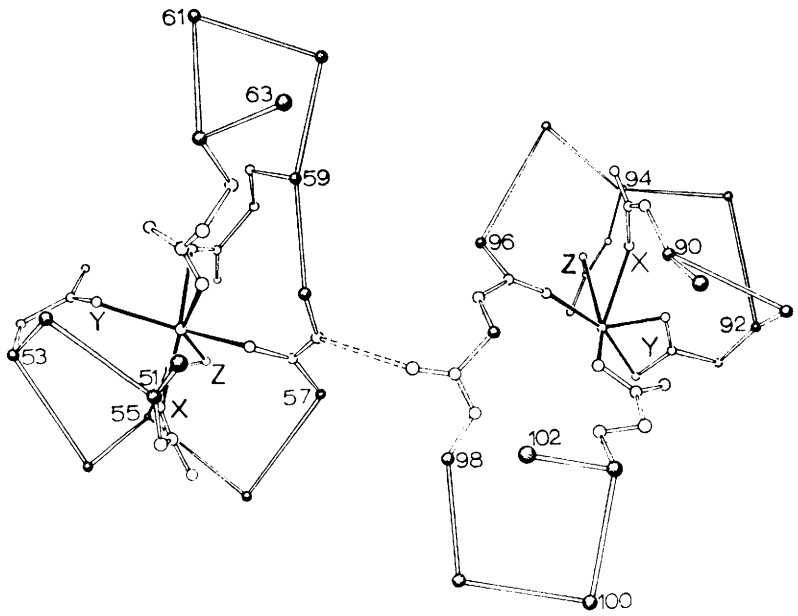


Figure 7c *Calcium binding protein: the co-ordination of the calcium ions by the CD and EF loops*
(Reproduced by permission from *J. Biol. Chem.*, 1973, **248**, 3313)

molecule has an approximate two-fold axis. The CD and EF loops each contain a β-bend and there are antiparallel β-sheet hydrogen-bonds between residues 97 and 58 and between 77 and 64.

Kretsinger and Nockolds define a surface accessible group as one able to make contact with a water molecule radius 1.7 Å. Of the 812 atoms 43% are inaccessible to the solvent; this includes 53% of 324 peptide atoms and 69% of the hydrophobic side-chain atoms. Most polar side-chains are on the surface; those buried include six carbonyl oxygen atoms involved in binding the calcium ions and the side-chains of Glu-81 and Arg-75, which form a buried salt bridge.

The helices pack together so that residues from one side of each form a core to the structure. This core is very hydrophobic, consisting of seven Phe, four Ile, five Leu, three Val, and parts of three other hydrophobic side-chains. The packing of these residues retains the molecule's approximate two-fold symmetry. The volume of the core is $\frac{1}{7}$ of the total molecular volume.

One calcium is located in the CD loop and the other in the EF loop (Figure 7c). Each ion is surmounted by a distorted octahedron formed by four carbonyl and two carboxyl oxygens.

Hendrickson and Karle have described the use of the tangent formula in the structure analysis of carp muscle calcium binding protein.[72]

Haemoglobins.—Since the presentation of a general stereochemical model for the allosteric behaviour of haemoglobin,[73] Perutz and his colleagues have been using it in conjunction with *X*-ray diffraction data to explain the effects of mutations and to answer detailed mechanistic questions. The changes that occur in the tertiary and quaternary structure of haemoglobin on release or uptake of oxygen are difficult to disentangle.

In an elegant paper[74] Anderson describes how he can clamp haemoglobin in the deoxy quaternary structure and then observe the changes that occur in the tertiary structure when oxygen is bound. Crystals of deoxyhaemoglobin are grown in a solution containing acrylamide and bisacrylamide. The acrylamide in the water spaces of the crystals is then polymerized and it is found that this amorphous inert material mechanically clamps the protein so that, on exposure to air, the molecule goes to the met state with tertiary but not quaternary structural changes. Anderson describes these changes as seen in a met acrylamide minus deoxy acrylamide difference Fourier. Only the α-haems are appreciably oxidized suggesting this occurs more readily than the oxidation of the β-haems. The salt bridges between the chain termini are still intact.

The iron moves towards the plane of porphorin ring which tilts, altering the position of the F helix, FG corner, and some other neighbouring residues (Figure 8). These movements change the structure of the sub-unit

[72] W. A. Hendrickson and J. Karle, *J. Biol. Chem.*, 1973, **248**, 3327.
[73] M. F. Perutz, *Nature*, 1970, **228**, 726.
[74] L. Anderson, *J. Mol. Biol.*, 1973, **79**, 495.

$\alpha_1\beta_2$ interface. Tyr-C7(42)α breaks its intrachain hydrogen-bond to Asn-G4(97)α and loosens the bond to Asp-G1(99)β_1. Asn-G4(102)β_2, which is in contact with the haem, breaks the intrasub-unit hydrogen-bond to Asp-G1(99)β_2 and moves towards Asp-G1(94)α which has moved

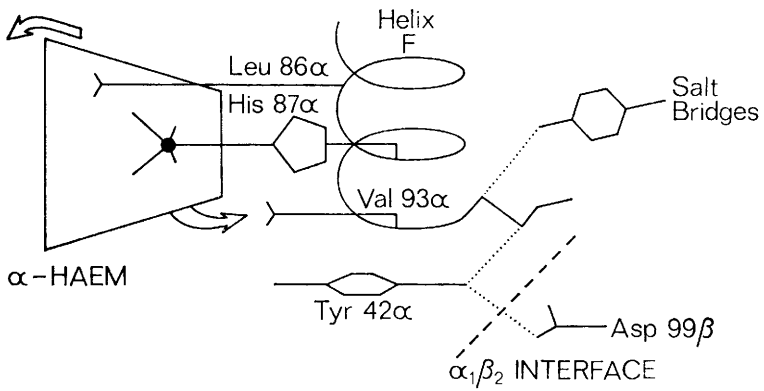


Figure 8 *Haemoglobin A: schematic diagram showing the proposed mechanism of α-chain tertiary structure change on ligation. A decrease in porphyrin–His-F8(87) separation causes helix F to move closer. The iron acts as a ball and socket joint. Leu-F7(86) is rigidly attached to His-87 and pushes strongly on the upper outside haem corner, while Val-FG5(93), less securely fixed to His-87, is displaced by the opposite haem corner, and the nearby main-chain hydrogen-bonding network loosened*
(Reproduced by permission from *J. Mol. Biol.*, 1973, **79**, 495)

in from the internal cavity. These changes tilt the β haem and enlarge its pocket by moving the E and B helices.

The exact stereochemical mechanism of the steps that follow those observed here by Anderson, the breakage of the salt bridges and the change in the quaternary structure, is still not clear.

The structures of the following five mutants of haemoglobin A have been described this year.

Haemoglobin M Boston. In this mutant the histidine distal to the iron atom in the α-chain, E7(58)α, is replaced by a tyrosine, making the oxygen affinity of the β-chain low. A difference Fourier map between deoxyhaemoglobin M Boston and deoxyhaemoglobin A shows that,[75] in the mutant, the iron atom is shifted 0.24 Å to the distal side of the haem plane and is bonded to the phenolate atom of the tyrosine. The bond that normally links His-F8(87)α to the iron is absent in this mutant allowing the F helix to rotate clockwise with the histidine going deep into the haem pocket (Figure 9). The absence of this bond means that the effect of

[75] P. D. Pulsinelli, M. F. Perutz, and R. L. Nagel, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3870.

oxygen binding to the α-sub-unit is not transmitted to the β-sub-unit, and therefore the oxygen affinity of the β-chain is less than that of haemoglobin A.

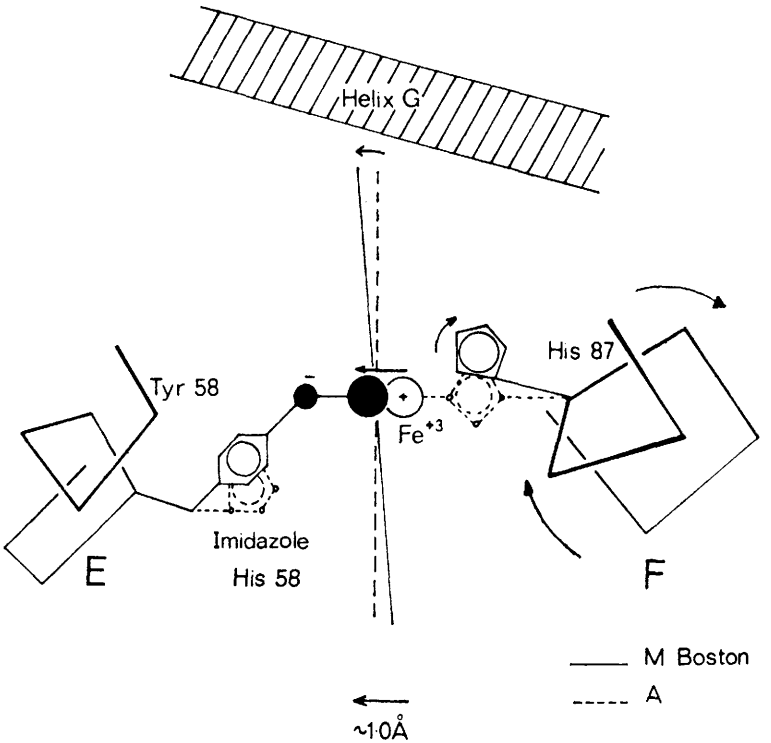


Figure 9 *Haemoglobin M Boston: schematic diagram of the stereochemical changes in the haem pocket of the α subunit on going from deoxyHb A to Hb M Boston*
(Reproduced by permission from *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3870)

Haemoglobin Yakima. This mutant has a high oxygen affinity. The mutation is Asp-G1(99)β → His. In oxyhaemoglobin A, Asp-G1(99)β is in the central cavity of the molecule and on deoxygenation it moves to the $\alpha_1\beta_1$ interface where it hydrogen-bonds to Tyr-C7(42)α. From a difference Fourier map between deoxyhaemoglobin Yakima and deoxyhaemoglobin A, Pulsinelli [76] has shown that the effect of this mutation produces complex structure differences at the $\alpha_1\beta_1$ interface and in the tertiary structure of the sub-units. In general, the structure of the deoxy form of the mutant can be seen as intermediate between the deoxy and oxy forms of the A protein; the salt bridges of the deoxy form are present but the oxy form

[76] P. D. Pulsinelli, *J. Mol. Biol.*, 1973, **74**, 57.

hydrogen-bond between Asp-G1(94)α_1 and Asn-G4(102)β_2 also occurs. The high oxygen affinity of haemoglobin Yakima is due to the mutation destabilizing the deoxy structure.

Haemoglobin Little Rock. This haemoglobin has a high oxygen affinity and the mutation is His-H21(143)$\beta \rightarrow$ Glu. Residue 143 is on the surface in the internal cavity. From the examination of a molecular model,

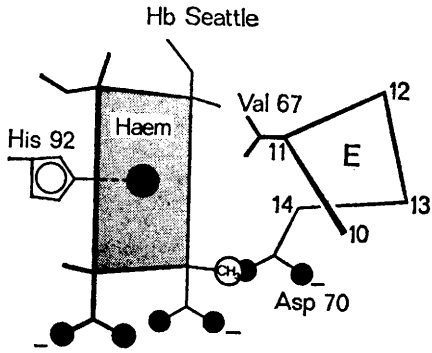


Figure 10 *Haemoglobin Seattle: schematic diagram showing the stereochemistry of the mutation*
(Reproduced by permission from *Nature New Biol.*, 1973, **243**, 274)

Perutz [77] has pointed out that in the oxy structure the glutamic acid of the mutant could form two hydrogen-bonds to Asn-H17(139) in the opposite β-chain. To convert the molecule into the deoxy form, these hydrogen-bonds would have to be broken. Their stabilization of the oxy form therefore explains the high oxygen affinity of the mutant.

Haemoglobin Seattle. This low oxygen affinity haemoglobin has the mutation Ala-E14(70) $\rightarrow$ Asp. Ala-70 is on the surface of the protein near the haem. A difference Fourier map between deoxyhaemoglobin Seattle and deoxyhaemoglobin A shows [78] that the carboxy-group of the mutant Asp is within 3 Å of the methyl group on haem pyrrole ring IV (Figure 10). It is not clear why this contact gives the mutant its low oxygen affinity.

Haemoglobin S. This haemoglobin has the mutation Glu-A3(6)$\beta \rightarrow$ Val and is the cause of sickle-cell anaemia. Oxyhaemoglobin S is very similar to oxyhaemoglobin A in its properties. The deoxy form of S is much less soluble than the deoxy form of A and gives a fibrous precipitate. Fibre formation stabilizes the deoxy structure and so reduces the molecule's oxygen affinity. From the results of an *X*-ray and electron microscope

[77] M. F. Perutz, *Nature New Biol.*, 1973, **243**, 180.
[78] N. L. Anderson, M. F. Perutz, and G. Stamatoyannopoulos, *Nature New Biol.*, 1973, **243**, 274.

examination of deoxyhaemoglobin S, Finch *et al.* have presented a model for the arrangement of the molecules in the fibre.[79]

Thin sections of whole deoxyhaemoglobin S cells show fibres aligned along its length and wrapped around the cell walls. This gives the erythrocytes their characteristic sickle shape. The fibres are packed in hexagonal or tetragonal arrays.

The structure of the fibres is most simply described in terms of stacked rings of six molecules. Adjacent rings are related by a right-handed 7.3° twist so the resulting helical arrangement repeats after 8.2 turns (Figure 11). In part, this description is misleading because the fibres break down to give single filament rings. Thus, a more accurate description of the structure is in terms of six filaments packing in longitudinal register to form a helical fibre. The haemoglobin molecule is 65 × 55 × 50 Å and the rings have a separation of 64 Å suggesting that the long axis of the molecule is nearly parallel to the fibre axis. The position of the other two molecular axes in the fibre is not known.

Haemoglobin from Sea Lamprey. Hendrickson, Love, and Karle have published details of the structure analysis of this protein.[80] Isomorphous phases calculated from three heavy-atom derivatives were partially refined using the tangent formula. A fairly complete account of the structure was published previously. This paper adds a schematic drawing of the intramolecular hydrogen-bonds, stereographic diagrams of segments of the entire atomic model, and details of the contacts formed by the molecule's ten neighbours.

Enzymes of the Glycolytic Pathway.—This year saw the preliminary description of the atomic structure of lobster D-glyceraldehyde-3-phosphate dehydrogenase,[81] liver alcohol dehydrogenase,[82] and yeast phosphoglycerate kinase.[83] The remarkable feature of these results is that the nucleotide binding portions of these three enzymes are very similar in their tertiary structure to that found in lactate dehydrogenase, malate dehydrogenase, and flavodoxin. The availability of the sequence of dogfish LDH has made possible a detailed discussion of the relation between the structure and the function of the enzyme.

D-*Glyceraldehyde-3-phosphate Dehydrogenase.* This enzyme catalyses the reaction

$$\text{D-glyceraldehyde-3-phosphate} + \text{NAD}^+ + \text{phosphate} \rightleftharpoons \text{D-1,3-diphosphoglycerate} + \text{NADH}$$

[79] J. T. Finch, M. F. Perutz, J. F. Bertles, and J. Dobler, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 718.
[80] W. A. Hendrickson, W. E. Love, and J. Karle, *J. Mol. Biol.*, 1973, **74**, 331.
[81] M. Buehner, G. C. Ford, D. Moras, K. W. Olsen, and M. G. Rossmann, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3052.
[82] C. I. Bränden, H. Eklund, B. Nordström, T. Boiwe, G. Soderlund, E. Zeppezauer, I. Ohlsson, and A. Akeson, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2439.
[83] T. N. Bryant, H. C. Watson, and P. L. Wendell, *Nature*, 1974, **247**, 14.

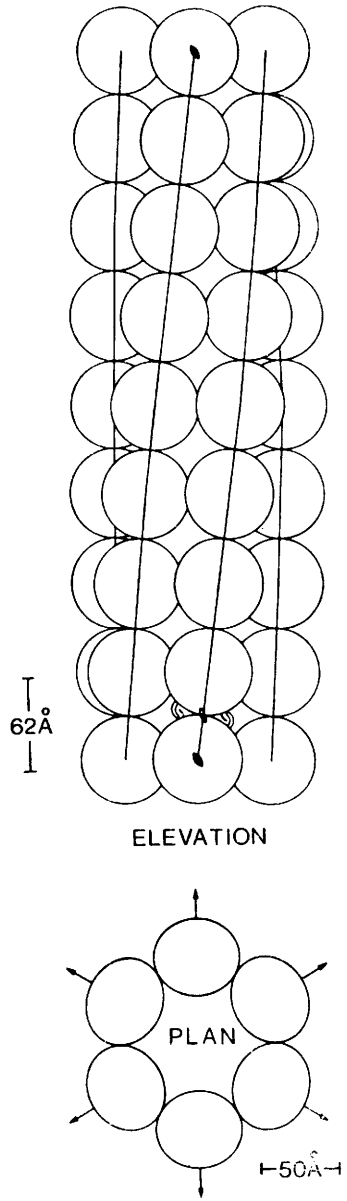


Figure 11 *Haemoglobin S: schematic diagram of the structure of helical tube of deoxyhaemoglobin S. The arrows and signs indicate the probable positions of the molecular dyads normal to the fibre axis. A length of eight rings is shown corresponding to the approximate repeat of the structure. The right-handedness of the helix was established by tilting the specimen about the fibre axis and observing the shift of the superposition pattern*
(Reproduced by permission from *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 718)

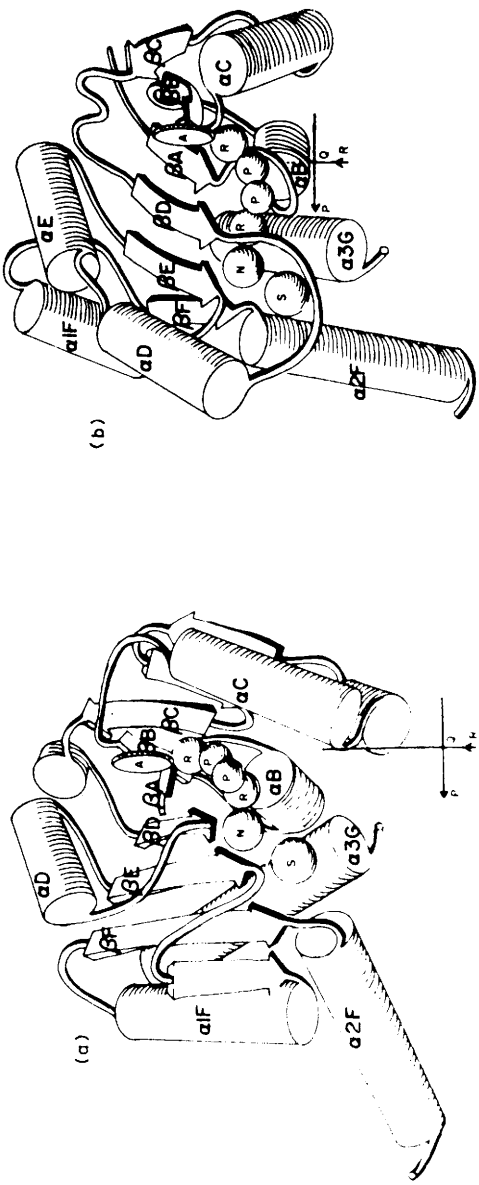


Figure 12 *Schematic diagrams of the coenzyme binding portion of* (a) *glyceraldehyde-3-phosphate dehydrogenase and* (b) *lactate dehydrogenase.* S–N–R–P–R–A *represents the position of the coenzyme. The nomenclature of the secondary structure is based on that of LDH*
(Reproduced by permission from *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3052)

It has four identical sub-units, related by 222 symmetry, and a molecular weight of 143 000. Buehner *et al.*, continuing the work of Watson and Banaszak, have given a description of the nucleotide-binding portion of the enzyme from lobster.[81] This description of the holo enzyme is based on a 3.0 Å electron-density map calculated using one heavy-atom derivative (K_2HgI_4) and averaging the electron density of the four sub-units. The molecule has 333 residues. Residues 22—164 form the coenzyme binding portion and fold to give a structure very similar to that which binds NAD in LDH. The major difference between the two structures (Figure 12) are

(i) In GPD there is an extra run of antiparallel sheet between αC and βC.

(ii) The loop between βD and αD is smaller in GPD.

(iii) Helix α1F between βE and βF in GPD has slipped around the edge of the sheet relative to its position in LDH.

The coenzyme binds in an extended conformation similar to that found in LDH (see below).

In GPD one set of sub-unit contacts occurs between the helices of the coenzyme part of the molecule and is similar to that formed by the Q symmetry axis in LDH (see below). The other sub-unit contacts position the enzyme's active site near a sub-unit interface, so Lys-183 of one sub-unit binds to the pyrophosphate of the coenzyme in the active site of the adjacent sub-unit (Figure 15 below). Buehner *et al.* have suggested that this intersub-unit contact may be involved in the enzyme's co-operative behaviour.

Liver Alcohol Dehydrogenase. This enzyme is a dimer of two identical sub-units of 374 residues and binds four zinc ions. It catalyses the reaction

$$\text{ethanol} + \text{NAD}^+ \rightleftharpoons \text{acetaldehyde} + \text{NADH} + \text{H}^+$$

From a 2.9 Å resolution electron-density map the fold of the polypeptide chain has been described by Bränden *et al.*[82] The molecule has the dimensions 45 × 55 × 110 Å. Each monomer is divided by a wide deep cleft into two unequal parts. The smaller part binds the coenzyme and has a structure very similar to the coenzyme binding part of LDH. Referring to Figures 12 and 13 the major differences between the two structures are

(i) The loop of 20 residues connecting βD and βE in LDH is reduced to three residues in LADH.

(ii) The chain between βE and βF forms a helix in LDH and has an extended conformation in LADH.

(iii) The βF strand is three residues in LDH and seven in LADH.

ADP–ribose binds in the same manner to the coenzyme sites in the two molecules.

The dominant structural features of the larger part of the sub-unit are two regions of pleated sheet, one of five strands and one of six. Two of the four zinc atoms are found at the active sites. The other two occupy a position in the neck between the two parts of the monomers.

The sub-units are related to each other by a two-fold axis. The main chain of βF strand of one sub-unit is hydrogen-bonded to the βF strand of the other so that the β-sheet of the coenzyme part extends right through the molecule.

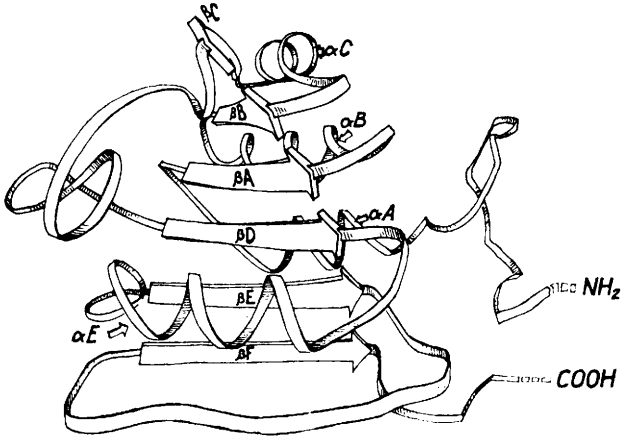


Figure 13 *Schematic diagram of the coenzyme binding portion of liver alcohol dehydrogenase. The nomenclature of the secondary structure is based on that of LDH*
(Reproduced by permission from *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2439)

Yeast Phosphoglycerate Kinase. This enzyme catalyses the reaction

$$\text{1,3-diphosphoglycerate} + \text{ADP} + \text{H}^+ \underset{}{\overset{\text{Mg}^{2+}}{\rightleftharpoons}} \text{3-phosphoglycerate} + \text{ATP}$$

The amino-acid sequence of the enzyme is unknown at present but, from a 3.5 Å resolution electron-density map, Bryant, Watson, and Wendell were tentatively able to trace the entire course of the main chain.[83] The molecule is a monomer. The chain folds to form two lobes connected by a neck; one lobe, A, contains 157 residues and the other B, 181. The chain crosses twice between the lobes and at the crossings forms helices. Five other contacts involving one or two residues occur between the lobes.

The protein binds Mn^{2+}–ADP and a 5 Å resolution difference map showed the position of the binding site in the A lobe. This lobe has six helices and a six-stranded β-sheet associated to form a structure similar to the coenzyme binding lobes of the dehydrogenases. The principal structural feature of the other lobe is a β-sheet with a helix on either side.

Lactate Dehydrogenase. This enzyme catalyses the reaction

$$\text{lactate} + \text{NAD}^+ \rightleftharpoons \text{pyruvate} + \text{NADH} + \text{H}^+$$

Previously, Rossmann and his colleagues have described the structure of the M_4 enzyme from dogfish muscle and with the sequence now available they

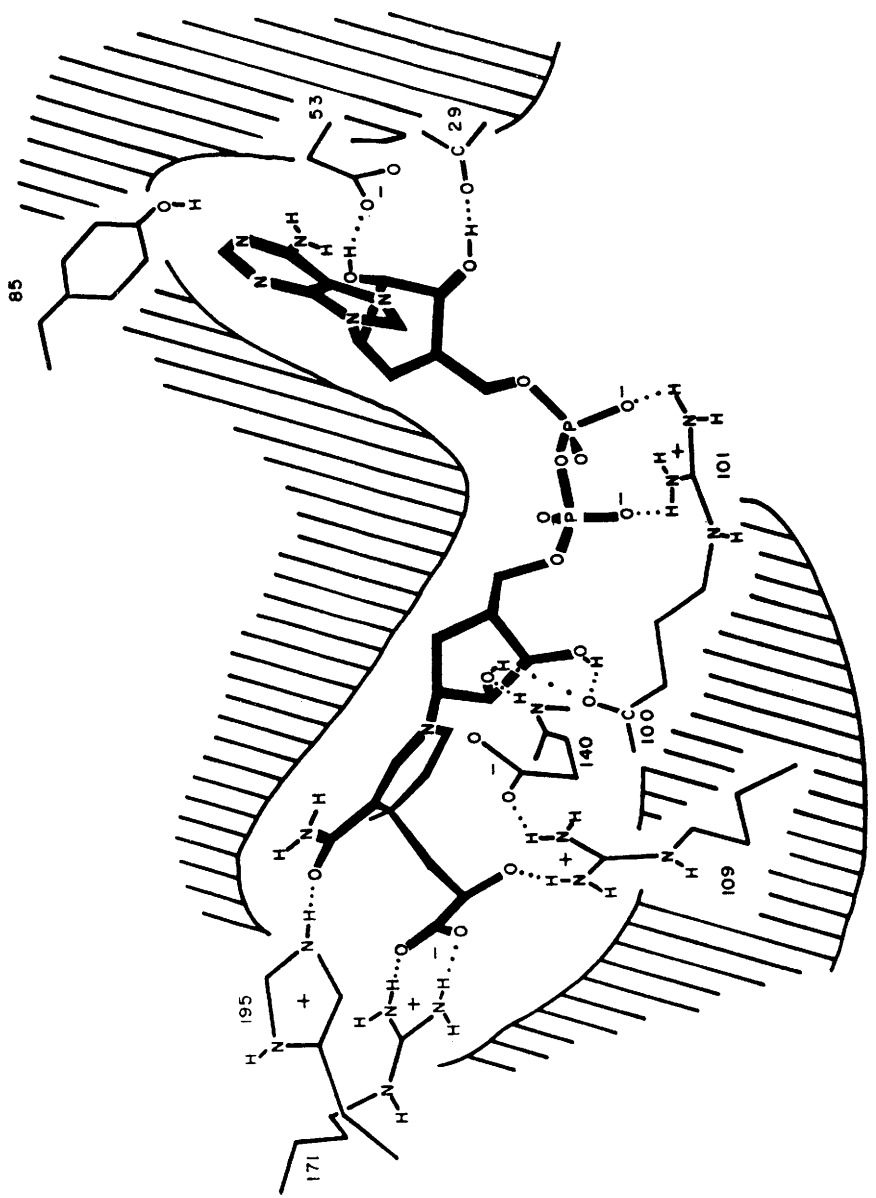


Figure 14 *Lactate dehydrogenase, showing the hydrogen bonds between the protein and coenzyme in the LDH–NAD-pyruvate complex*

have gone on to describe how the protein binds adenosine, 5′-AMP, 5′-ADP,[84] NAD-pyruvate,[85] citrate,[86] and oxamate.[87]

Pyruvate and NAD form an abortive ternary complex with LDH in which the pyruvate is covalently bound to NAD. From a 3.0 Å difference Fourier,[85] Adams *et al.* found that the coenzyme binds in an extended conformation (Figure 14). On formation of the ternary complex the loop formed by residues 98—114 folds down over the active site. The movement

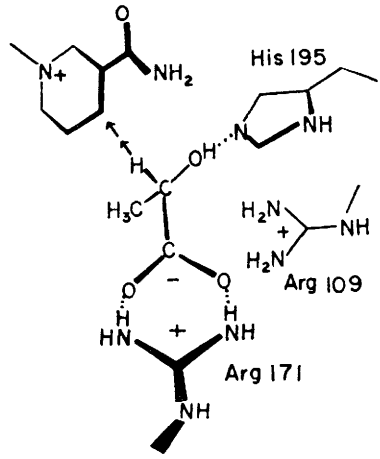


Figure 15 *Lactate dehydrogenase, showing the proposed substrate binding in the active ternary intermediate*
(Reproduced by permission from *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1968)

is quite large: Arg-101 moves 13 Å to cover the coenzyme phosphate and Arg-109 moves 23 Å to hydrogen-bond to the pyruvate carbonylate group. The exact orientation of the coenzyme is due, in the main, to the collapse of the 98—114 loop: the contact between Ala-30 and adenine phosphate, and that between residue 247 and the nicotinamide phosphate. The adenine ring sits in a hydrophobic pocket with the amino-group pointing out into the solvent. The positive charge on the nicotinamide ring is balanced by Glu-140. In the ternary complex His-195 moves 1—2 Å relative to the apo-enzyme positive so it can receive a proton from the hydroxy-group of the lactate, suggesting a mechanism partly shown in Figure 15.

[84] K. Chandrasekhar, A. McPherson, M. J. Adams, and M. G. Rossmann, *J. Mol. Biol.*, 1973, **76**, 503.
[85] M. J. Adams, M. Buehner, K. Chandrasekhar, G. C. Ford, M. L. Hackert, A. Liljas, M. G. Rossmann, I. E. Smiley, W. S. Allison, J. Everse, N. O. Kaplan, and S. S. Taylor, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1968.
[86] M. J. Adams, A. Liljas, and M. G. Rossmann, *J. Mol. Biol.*, 1973, **76**, 519.
[87] A. McPherson, *J. Mol. Biol.*, 1973, **76**, 528.

His-195 is near Cys-165 and the modification of this cysteine, which inactivates LDH, prevents the movement of the histidine. The binding mode of adenine and AMP in the active site is somewhat different from that of the adenine portion of the coenzyme.

There are two anion binding sites in LDH, one at the point in the active site where pyruvate binds in the LDH–NAD–pyruvate complex and one on a sub-unit interface. These sites may be occupied by sulphate, oxalate, and doubly, but not triply, charged citrate. The inhibition of the enzyme by oxamate and citrate is presumably due to this occupation of the active site where they hydrogen-bond to His-195 and Arg-171. The second anion binds at an intersub-unit site which is near the P symmetry axis (see below) where it is in contact with Arg-173 and Trp-190 of one sub-unit and Trp-207 of the other sub-unit. Occupation of this site by ions stabilizes the tetramer and reduces the hybridization by the M and H isoenzymes.

The four sub-units in LDH are related by 222 symmetry. Rossmann *et al.* have labelled these symmetry axes P, Q, and R (Figure 16) and have briefly described the contacts they form:[88] P axis contacts are between the faces of part of the antiparallel β sheet; Q axis contacts are between two parallel triple helices; R axis contacts are formed by the first 22 residues which wrap round the adjacent sub-unit.

Contacts similar to those formed by the Q axis occur in malate dehydrogenase and GPD.

Hydrolytic Enzymes.—*Bovine Trypsin–Bovine Pancreatic Inhibitor Complex.* Bovine pancreatic trypsin inhibitor (PTI) is a small protein that inhibits the action of trypsin by blocking the enzyme's active site. Ruhlmann *et al.* have reported the crystal structure of this complex.[89] The major steps in the catalytic mechanism of serine proteases are the formation of a tetrahedral intermediate in which the reactive serine of the protein covalently bonds to the carbonyl carbon of the substrate, the splitting of this to give an acyl enzyme, and then deacylation. The crystal structure analysis of the complex shows that the inhibitor binds to enzyme like a substrate and that, at the active site, the tetrahedral intermediate is formed in which the side-chain oxygen O_γ of Ser-195 is bonded to the carbonyl carbon of Lys-15 in the inhibitor. Formation of this bond involves a 120° rotation about the C_α—C_β bond of the serine relative to the conformation found in native structure. The position of the active-site His-57 also differs from that found in the acyl and native enzyme. In the acyl enzyme N_τ is hydrogen-bonded to a water molecule; in this complex rotation about C_β—C_γ N_τ moves to within 3.4 Å of O_τ of Ser-195. The tight fit of the enzyme and the inhibitor does not allow room for water molecules in the

[88] M. G. Rossmann, M. J. Adams, M. Buehner, G. C. Ford, M. L. Hackert, A. Liljas, S. T. Rao, L. J. Banaszak, E. Hill, D. Tsernoglou, and L. Webb, *J. Mol. Biol.*, 1973, **76**, 533.

[89] A. Ruhlmann, D. Kukla, P. Schwager, K. Bartels, and R. Huber, *J. Mol. Biol.*, 1973, **77**, 417.

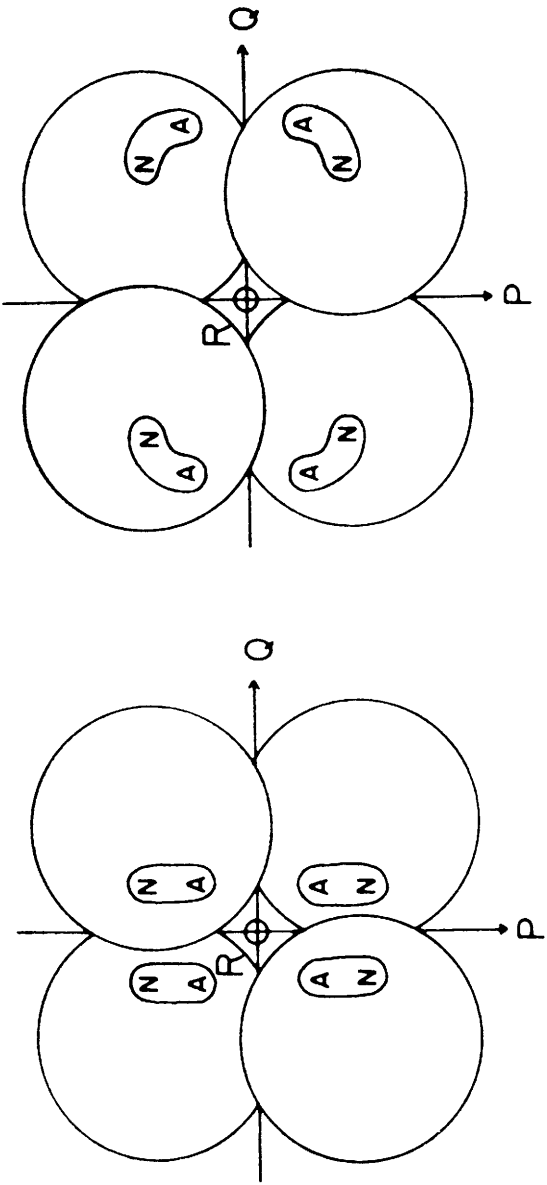


Figure 16 *Diagrammatic comparison of the association of sub-units in glyceraldehyde-3-phosphate dehydrogenase and lactate dehydrogenase.* A N *represents the different position of the active sites in the two enzymes*
Reproduced by permission from *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3052)

contact region. The other contacts between the inhibitor and the enzyme are essentially similar to those previously predicted by model building. Pro-13–Lys-15 of the inhibitor form a hydrogen-bonded antiparallel β-sheet with Lys-214–Gly-216 of the enzyme. The side-chain of Lys-15 sits in the specificity pocket.

Comparison of the Crystal Structures of Chymotrypsinogen A and α-Chymotrypsin. The zymogen chymotrypsinogen A is converted into the active enzyme chymotrypsin by the hydrolytic removal of two dipeptides Ser-14–Arg-15 and Thr-147–Asn-148. To describe the structural basis of the activation process Wright has compared the crystal structures of the zymogen and the enzyme.[90] Though the positions of the important active-site residues, Asp-102, His-57, and Ser-195, and the gross folding of the polypeptide chain are the same in both structures, 13 peptides, 96 residues, are found to have different conformations. Some of these differences can be accounted for by different interpretation of ambiguous sections of the electron-density map or by the fact that the chymotrypsinogen atomic co-ordinates were not refined. But it is quite clear from the comparison of the structures that the major effect of hydrolytic activation is the formation of the specificity pocket at the active site. In chymotrypsinogen only the 'bottom' surface of this pocket exists. On hydrolysis the top part of the pocket is formed by residues Met-192 and Gly-193 moving and the rearrangement of the local hydrogen bonds (Figure 17). In a second paper, Wright argues that the creation of this pocket is sufficient to explain the different catalytic activities of the zymogen and the enzyme.[91]

α-Chymotrypsin. In a series of papers Tulinsky and his colleagues have described their work on the crystal structure of α-chymotrypsin.[92–95] The asymmetric unit of α-chymotrypsin contains two molecules almost exactly related to each other by a two-fold axis. The previous studies on α-chymotrypsin were based essentially on a model built by interpreting the average electron density of this dimer.[96] Tulinsky *et al.* have now looked at the differences between the structures of the two molecules and find that these exist for *ca.* $\frac{1}{6}$ of the molecule. They occur in a 6 Å shell of density around the surface of the molecule and are mostly due to differences in side-chain conformations. The residues showing the most variation are Lys, Glu, and Asn. Asp and Glu tend to be fairly similar. The packing of the residues forming the dimer contact causes some of them, Phe-39, Ser-217, and Ser-218 in particular, to depart by up to 2.5 Å from positions related by

[90] H. T. Wright, *J. Mol. Biol.*, 1973, **79**, 1.
[91] H. T. Wright, *J. Mol. Biol.*, 1973, **79**, 13.
[92] A. Tulinsky, N. V. Mani, C. N. Morimoto, and R. L. Vandlen, *Acta Cryst.*, 1973, **B29**, 1309.
[93] A. Tulinsky, R. L. Vandlen, C. N. Morimoto, N. V. Mani, and L. H. Wright, *Biochemistry*, 1973, **12**, 4185.
[94] A. Tulinsky and L. H. Wright, *J. Mol. Biol.*, 1973, **81**, 47.
[95] R. L. Vandlen and A. Tulinsky, *Biochemistry*, 1973, **12**, 4193.
[96] J. J. Birktoft and D. M. Blow, *J. Mol. Biol.*, 1972, **68**, 187.

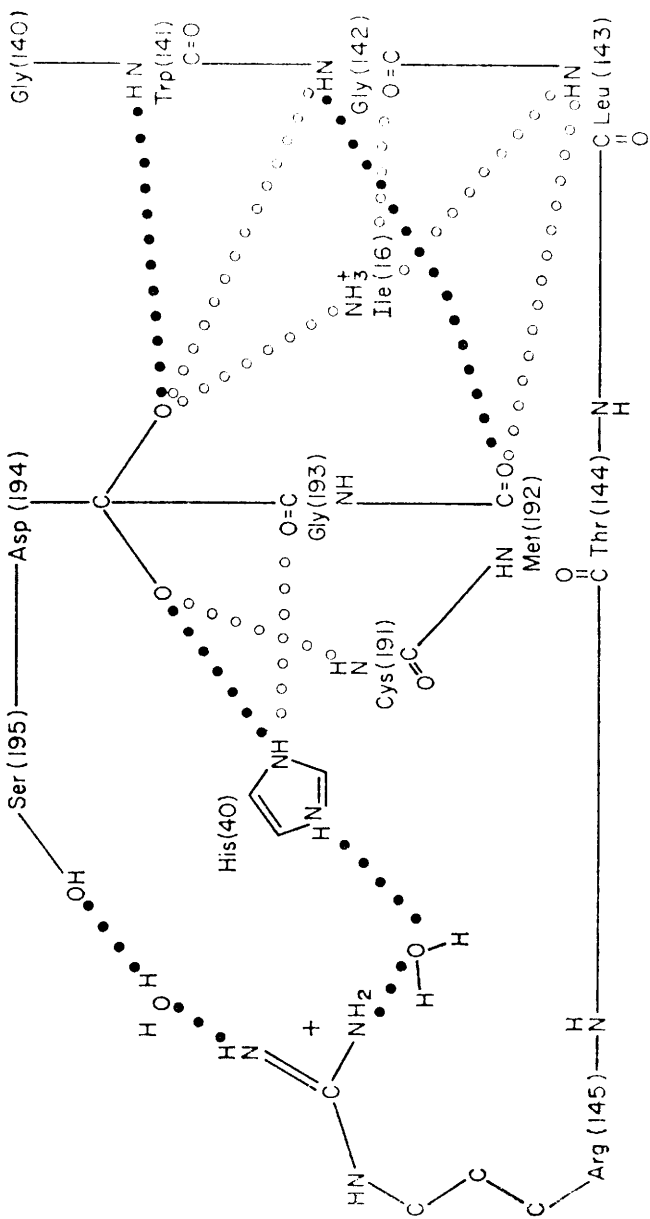


Figure 17 *Chymotrypsinogen: changes in hydrogen-bonding around the Ile-16–Asp-194 ion-pair. Solid circles represent hydrogen bonds which exist in chymotrypsinogen, but are absent in α-chymotrypsin; open circles are hydrogen bonds which exist in α-chymotrypsin but not in chymotrypsinogen. Proton positions are hypothetical for pH 6.3*
Reproduced by permission from *J. Mol. Biol.*, 1973, **79**, 1)

the two-fold axis. The differences between the surfaces of the two molecules are reflected in consistent differences in substitution by the heavy atoms used in the structure analysis.

α-Chymotrypsin crystallizes from ammonium sulphate solutions, and Tulinsky and Wright have described how a difference map between crystals soaked in NH_4SO_4 and NH_4SeO_4 shows the position of seven anions bound to the protein surface.[94] All seven are linked to the protein by hydrogen bonds and five also interact with positively charged groups. The site with the highest occupation occurs on the dimer interface where the ion hydrogen-bonds to O_γ of Ser-195 and NH of Gly-193 of one molecule and OH of Tyr-146 in the next.

In another paper, the structures of crystalline α-chymotrypsin at pH 4.2 and pH 6.7 are compared.[95] At pH 4.2 the enzyme forms dimers and high polymers in solution. Raising the pH of the solution to 6.7 prevents polymerization. Raising the pH of the crystals to 6.7 results in complex changes in the structure of the dimer interface which the authors discuss in some detail.

Ribonuclease. Allewell, Mitsui, and Wycoff[97] have described the structure of N_ε-Lys-41-dinitrophenyl ribonuclease S. From chemical studies, it is found that dinitrophenylation of the ε-NH_3^+ group of Lys-41 substantially reduced the enzymic activity of ribonuclease. Molecular model building suggests that the catalytic role of the ε-NH_3^+ group is to stabilize the five-co-ordinate phosphate intermediate formed during nucleic acid hydrolysis. From a 2.0 Å resolution difference Fourier, Allewell *et al.* have shown that, in the inhibited enzyme, Lys-41 moves away from the active site (Figure 18) so the DNP is partially enclosed by the loop made by residues 33—41. Contacts occur between the DNP group and Asn-34, Asp-38, Arg-39, and the side-chain of Glu-11. The DNP group does not sterically block the active site; pyrimidine inhibitors bind in the same position, and with the same binding constant, to N_ε-Lys-41-dinitrophenyl ribonuclease S as they do to ribonuclease S. The loss of enzymic activity is due to the neutralization and removal from the active site of the ε-amino-group.

Surprisingly, the presence of the dinitrophenyl group seems to cause concerted distortions in the protein structure. The S peptide, helix 50—57, the loop 65—72 and the loop about 114 move 0.5—1.0 Å away from the active site.

Lysozyme. From chemical studies it was found that the irreversible inhibitor 2′,3′-epoxypropyl β-glycoside of *N*-acetyl-D-glucosamine, β(1—4)-linked dimer forms a covalent bond to Asp-48 or Asp-52 of hen egg-white lysozyme. To find which of the two residues was involved, Moult, Eshdat, and Sharon[98] calculated a 2.5 Å resolution difference Fourier in three

[97] N. M. Allewell, Y. Mitsui, and H. W. Wycoff, *J. Biol. Chem.*, 1973, **248**, 5291.
[98] J. Moult, Y. Eshdat, and N. Sharon, *J. Mol. Biol.*, 1973, **75**, 1.

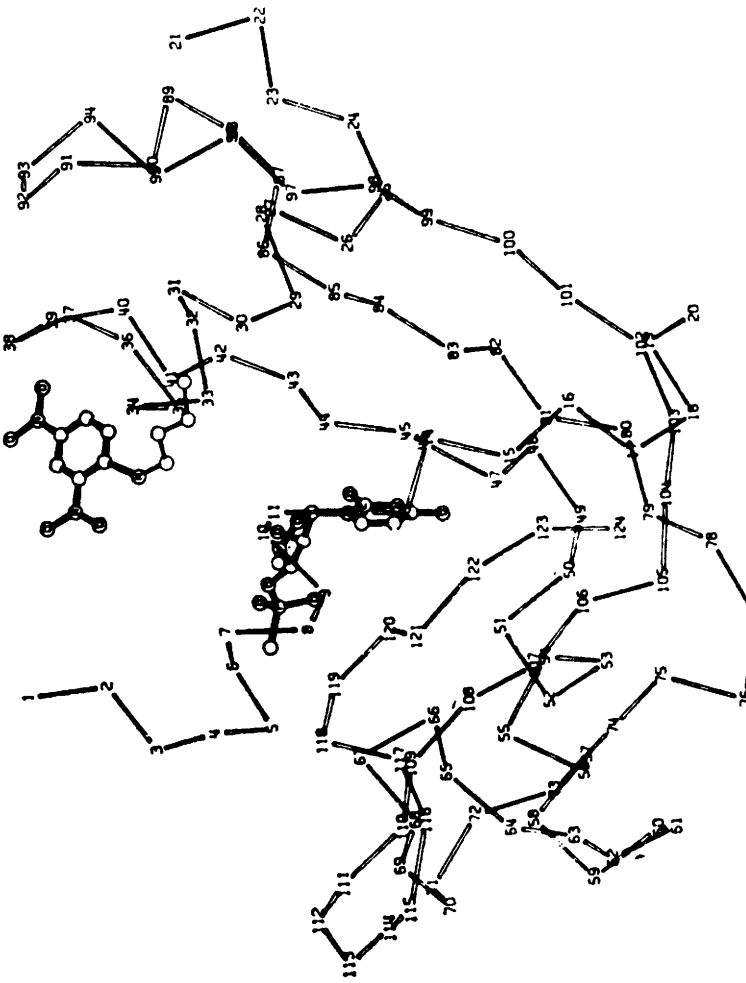


Figure 18 *Lys-41-dinitrophenyl ribonuclease S: diagram showing the position of the dinitrophenyl group in the inhibited enzyme*
(Reproduced by permission from *J. Biol. Chem.*, 1973, **248**, 5291)

projections between the inhibited and native enzyme. Examination of the map showed that the inhibitor occupies the B and C sub-sites of the active site and is bonded to Asp-52.

5 Low-resolution Studies of Virus Structures

Cowpea Chlorotic Mottle Virus.—Crystals of this virus can be grown in 1.5M sodium phosphate, 0.85M sodium citrate, or 1.45M ammonium sulphate.[99] These salt solutions are buffered to give a pH of 4.5—6.5. Crystals grown from citrate are cubic, space group $F432$ and unit cell dimension $a = 560 \pm 15$ Å. The crystals grown from phosphate have a smaller cell dimension, $a = 510 \pm 15$ Å, but are otherwise similar. The virus has a protein coat consisting of 180 identical proteins. From the size and symmetry of the unit cell these proteins must pack so that the coat has at least 23 symmetry. This result is consistent with the previous proposal that the coat has icosahedral symmetry ($T = 3$).

Adenovirus Hexon.—The hexon protein is the major component of the coat of adenovirus. It was crystallized by two groups of workers who reported different crystal densities and therefore different molecular weights for the protein in the crystal asymmetric unit. Cornick, Sigler, and Ginsberg [100] have now reported a new method for determining protein crystal densities based on cross-linking the protein with glutaraldehyde and replacing the mother liquid with pure water. This overcomes the problem of the different contributions of protein-bound pure water and of the bulk salt solution to the crystal density. Application of this technique to crystals of adenovirus gives the molecular weight of the protein in the asymmetry as 102 000 ± 2000. The relation of this mass to that of the hexon is unclear.

Broad Bean Mottle Virus.—This is a spherical RNA virus with a coat of 180 identical protein sub-units packed to give icosahedral symmetry. Small-angle X-ray scattering data have been collected on this virus in solutions whose electron density has been varied by the addition of sucrose and the trisaccharide melezitose.[101] Analysis of the scattering data implies that the virus (i) has a radius of gyration of 117 Å, (ii) has an outer radius of 147 Å and a hollow core radius 60 Å, (iii) has a region of high density between 90 and 130 Å from its centre which contains RNA as well as protein. The region 50—90 Å contains mostly RNA and the central part is solvent possibly with some RNA. Sucrose can penetrate 90% of the regions open to water, indicating that the centre of the virus must be accessible to the external solvent through channels of at least 10 Å.

Bacteriophage fr.—Alkaline degradation of the spherical virus fr produces the intact protein coat free of nucleic acid and with a sedimentation coefficient of 42S, and simple reaggregation of the protein coat sub-units

[99] M. G. Rossmann, I. E. Smiley, and M. A. Wagner, *J. Mol. Biol.*, 1973, **74**, 255.
[100] G. Cornick, P. B. Sigler, and H. S. Ginsberg, *J. Mol. Biol.*, 1973, **73**, 533.
[101] R. A. White and F. A. Fischbach, *J. Mol. Biol.*, 1973, **75**, 549.

produces a range of particles whose average sedimentation coefficient is 45.7S.[102] From low-angle *X*-ray scattering, Zipper, Schubert, and Vogt[102] concluded that the 42S particle has the same dimensions as the protein coat of the intact virus, an outer radius of 132 Å and an inner radius of 106 Å. The reaggregated particles have the same outer radius but a slightly smaller radius of gyration, and from the radial net electron-density distribution it was concluded that, on average, 14.4 additional protein sub-units are located inside the intact protein coat.

Part III: Conformation and Interaction of Peptides and Proteins in Solution

edited by R. H. Pain, with contributions by E. A. Carrey, A. T. Hagler, C. E. Johnson, P. Knowles, G. R. Penzer, H. W. Rattle, B. Robson, and R. M. Stephens

1 Introduction

By R. H. Pain

This year has seen a quickening of interest as to the way in which polypeptides fold up into their native conformation. Firm evidence is appearing for the existence of nucleating states and of stable, intermediate states on the folding pathway. The advances in statistical methods for correlating conformation with sequence in globular proteins, coupled with energy calculations, show that the α-helix and β-structures fulfil the criteria for the initial, nucleating structures which ensure a self-selection by the polypeptide chain of a folding pathway. The general feeling is that the protein folds *via* a series of structures which guide it into a low, free energy state or series of sub-states, hence speeding the folding and preventing the protein from getting trapped in metastable, non-native states. The question remains as to whether the final state is the lowest possible free energy state or not. The expansion of theoretical work on protein conformations is accompanied by a note of optimism that is it now possible to carry out meaningful calculations on tertiary structure or post-nucleation folding interactions.

The still expanding interest in membrane proteins and membrane structure is indicated particularly in the section on spin labels while fluorescence continues to expand as a tool of wide application and, at the same time, is becoming better understood. The increasing availability of Fourier transform spectrometers is extending the usefulness of n.m.r. in biological systems and revealing the great potential of ^{13}C n.m.r. spectra. Mössbauer spectroscopy has shown a fresh lease of life with some interesting studies on proteins new to this technique.

This year it has not been possible to review the literature on circular dichroism or on association–dissociation reactions of proteins. It is planned to cover these next year.

[102] P. Zipper, D. Schubert, and J. Vogt, *European J. Biochem.*, 1973, **36**, 301.

2 Theoretical Aspects of Protein Conformation

Contributed by A. T. Hagler and B. Robson

This section covers theoretical studies of protein conformation published subsequent to, or not covered by, the two previous reviews in this series.[1, 2] The reader is also referred to recent reviews by Brant [3] and Lotan *et al.*,[4] a book by Walton and Blackwell,[5] and the proceedings of a recent Jerusalem Symposium on Quantum Chemistry and Biochemistry.[6] Further, a critical review by Hagler and Lifson [7] is particularly relevant.

Introduction.—The covalent structure of a globular protein is a linear sequence of amino-acid residues which, under suitable conditions, carries all the information for the biologically active conformation of the protein.[8] The mechanism for the translation of this information is the thermodynamically reversible folding of the protein molecule so as to minimize the free energy arising from interactions within the molecule and between the molecule and its environment.[9] Of current interest are those proteins whose biologically active conformation corresponds to the conformation of lowest free energy in aqueous solvent, but there are probably certain proteins such as strongly bound membrane proteins whose biologically active conformations are only stabilized in the non-aqueous environment in which the protein has its biological function.[10]

The prediction of the native conformation of a globular protein therefore depends on evaluation of all the interactions involving the protein molecule, and then minimizing the free energy calculated as arising from these interactions, as a function of conformation. Such a procedure assumes, of course, that the native state corresponds to the global minimum, a hypothesis which is widely accepted as a working rule.[11] It is worth noting, however, that Wetlaufer and Ristow [12] have recently discussed the possibility that the native conformation does not correspond to the lowest minimum overall, but only to the lowest minimum which can be attained during the lifetime of the protein as an experimentally observed entity.

[1] B. Robson, in 'Amino-acids, Peptides, and Proteins', ed. G. T. Young (Specialist Periodical Reports), The Chemical Society, London, 1972, Vol. 4, p. 224.
[2] B. Robson, in 'Amino-acids, Peptides, and Proteins', ed. R. C. Sheppard (Specialist Periodical Reports), The Chemical Society, London, 1974, Vol. 5, p. 180.
[3] D. A. Brant, *Ann. Rev. Biophys. Bioeng.*, 1972, **1**, 369.
[4] N. Lotan, A. Buger, and E. Katchalski, *Ann. Rev. Biochem.*, 1972, **41**, 869.
[5] A. G. Walton and J. Blackwell, 'Biopolymers', Academic Press, New York, 1973.
[6] Proceedings of the Jerusalem Symposium on Quantum Chemistry and Biochemistry V, 'Conformation of Biological Molecules and Polymers', Academic Press, New York, 1973.
[7] A. T. Hagler and S. Lifson in 'The Proteins', Vol. III, 1974, 3rd edn.
[8] C. B. Anfinsen, *Science*, 1973, **181**, 223.
[9] C. Tanford, *Adv. Protein Chem.*, 1968, **23**, 252.
[10] T. L. Steck and C. F. Fox, 'Membrane Molecular Biology', ed. C. F. Fox and A. Keith, Sinauer Associates Inc., 1972, p. 27.
[11] H. A. Scheraga, ref. 6, p. 51.
[12] D. B. Wetlaufer and S. Ristow, *Ann. Rev. Biochem.*, 1973, **42**, 135.

The Problem of Multiple Minima.—Even if the solvent were neglected, a protein molecule has so many conformational degrees of freedom that finding the conformation of lowest free energy is by no means simple. The problem is rendered particularly difficult by the fact that there will invariably be many minima in the conformational energy surface. The lowest energy conformation can only be recognized when it is proven that there is none lower, and this implies an exhaustive search of the conformational energy surface which is well beyond the capabilities even of modern computers. However, some preliminary attempts have been made to develop procedures which can hopefully surmount the multiple minima problem without an exhaustive search, at least when applied to small molecules.[13–17] As pointed out by the authors of these procedures, however, they are not generally applicable and the problem of finding the global minimum in a multidimensional surface with only relatively few minima has not yet been solved. Thus, for example, Crippen and Scheraga [17] have shown that, for tetraglycine, different techniques lead to different sets of minimum energy conformations. One of the conclusions reached is that for larger molecules the choice of starting conformation for the procedure is particularly important.

Heuristic Programs and Nucleation.—In order to overcome the problem of multiple minima, Robson and Pain [18] proposed the use of programs of the 'heuristic' type. Actually, 'heuristic' in its most general sense is no more than a generic name for a varied body of procedures by which information, other than that obtained from the energy surface by more classical minimization routines, is introduced into the overall minimization procedure. In particular, the conclusion of the previous section concerning the importance of the choice of starting conformation confirms the widely held opinion that a choice of starting conformation which is close to, and on a path to, the native conformation can save a great deal of computing time.[11] However, the heuristic approach is by no means limited to suitable starting conformations but can also be applied to directing or biasing the search trajectories of minimization procedures according to the dictates of various types of additional information,[18] which may be provided by n.m.r. and i.r. spectroscopy, c.d., preliminary *X*-ray results, and so on. Indeed, an interesting study on the conformations of actinomycin D has already been carried out in this way by De Santis and his colleagues.[19–21] Further, preliminary results are reported of a test on the search for the global minimum of myoglobin under constraints imposed by low-resolution *X*-ray data.[21]

[13] G. M. Crippen and H. A. Scheraga, *Proc. Nat. Acad. Sci. U.S.A.*, 1969, **64**, 42.
[14] K. D. Gibson and H. A. Scheraga, *Comput. Biomed. Res.*, 1970, **3**, 375.
[15] G. M. Crippen and H. A. Scheraga, *Arch. Biochem. Biophys.*, 1971, **144**, 453.
[16] G. M. Crippen and H. A. Scheraga, *Arch. Biochem. Biophys.*, 1971, **144**, 462.
[17] G. M. Crippen and H. A. Scheraga, *J. Comput. Phys.*, 1973, **12**, 491.
[18] B. Robson and R. H. Pain, ref. 6, p. 279.
[19] P. De Santis, R. Rizzo, and G. Ughetto, *Biopolymers*, 1972, **11**, 279.
[20] P. De Santis, R. Rizzo, and G. Ughetto, *Nature New Biol.*, 1972, **237**, 94.
[21] P. De Santis, ref. 6, p. 493.

Direction of the search trajectory is further justified by the general feeling that a folding protein molecule faces in the real world the same problem as its simulated counterpart, that is, the way in which it folds must be directed in order that it may achieve its native conformation in reasonable time.[22] Since the experiments of Anfinsen [23] show that the protein carries all the information for its own native conformation, a possible mechanism for such direction is that restrictions on the subsequent pathway of folding are imposed by the influence of certain intermediate conformations already attained as a step in the folding process. Such intermediates are generally known as nucleating intermediates and the subsequent direction of the folding process is governed by this nucleation. Clearly, one way of introducing additional information to the computer is to use our current understanding of the way the protein folds in nature in order to simulate the real folding process.

Work continues towards an understanding of nucleation processes. Such locally organized secondary structures as α-helix, β-pleated sheet, and reverse (or β-) turns constitute a large part of the native conformation of most proteins,[24] and continue to be accepted as likely candidates for nucleating structures.[8] From a statistical analysis of proteins of known sequence and conformation, Chou and Fasman [25–27] have proposed a model for the formation of α-helix, β-pleated sheet, and reverse turn regions from a denatured, time-random coil. This model leads to accurate predictions of helix and pleated sheet regions in globular proteins. That at least α-helical regions can form metastable folding intermediates is supported by experimental studies.[18, 28, 29] Further evidence concerning the relation between amino-acid sequence and the tendency to form α-helix, β-pleated sheet, and reverse turns has been obtained by Krigbaum and Knutton.[30]

Theoretical energy calculations supporting and extending these studies have also been carried out, and emphasis in recent years has been on the reverse turn.[31–33] However, a more general kind of investigation has

[22] Nature Molecular Biology Correspondent, *Nature*, 1971, **233**, 523.
[23] C. B. Anfinsen, Brookhaven Symposium on Biology, 1962, **15**, 184.
[24] J. L. Crawford, W. N. Lipscomb, and C. G. Schellman, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 538.
[25] P. Y. Chou and G. D. Fasman, *J. Mol. Biol.*, 1973, **74**, 263.
[26] P. Y. Chou and G. D. Fasman, *Biochemistry*, 1974, **13**, 211.
[27] P. Y. Chou and G. D. Fasman, *Biochemistry*, 1974, **13**, 222.
[28] O. Jardetzky, H. Thielman, Y. Arata, J. L. Markley, and M. N. Williams, *Proc. Cold Spring Harbor Symp. Quant. Biol.*, 1971, **36**, 257.
[29] D. H. Sachs, A. N. Schechter, A. Eastlake, and C. B. Anfinsen, *Proc. Nat. Acad. Sci. U.S.A.*, 1972, **69**, 3790.
[30] W. R. Krigbaum and S. P. Knutton, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2809.
[31] P. N. Lewis, F. A. Momany, and H. A. Scheraga, *Proc. Nat. Acad. Sci. U.S.A.*, 1971, **68**, 2293.
[32] P. N. Lewis, F. A. Momany, and H. A. Scheraga, *Biochem. Biophys. Acta*, 1973, **303**, 211.
[33] R. Chandrasekaran, A. V. Lakshminarayanan, U. V. Pandya, and G. N. Ramachandran, *Biochim. Biophys. Acta*, 1973, **303**, 14.

examined the implications of a stereochemical alphabet in the formation of nucleating structures.[34] In this context, a stereochemical alphabet implies writing the conformation of a polypeptide chain in terms of names for a limited number of allowed regions for the conformation of each residue. These allowed regions are, by definition, determined solely on the basis of interactions within each residue and its associated segment of backbone.

The next stage of folding, the interaction of secondary structure units with the rest of the molecule, has been analysed by Nagano.[35] The role of reverse turns in bringing together α-helix or potential β-pleated sheet regions has also been discussed.[31] The folding of myoglobin as directed by helix stacking interactions has been studied theoretically.[36, 37] Finally, the organization of interacting secondary structure systems into compact and distinct globular units has been discussed by Wetlaufer.[38] These units, containing regions of protein backbone of *ca.* 40—150 residues in length, may constitute the whole protein or parts of a much larger protein, in which case they are referred to as super-secondary structures (a better description might be sub-tertiary).

Energy Refinement of *X*-ray Co-ordinates.—As described above, low-resolution *X*-ray data for a protein may provide additional information for directing the search for the global energy minimum of that protein. Using high-resolution *X*-ray data to construct a starting conformation, the search for the global minimum is simpler, since it should, in principle, be so close that a classical minimization procedure should be applicable without further embellishment. In fact, such energy minimization may ultimately become a standard technique for the refinement of *X*-ray structure, although this has not yet been achieved.

Refinement is generally necessary because the co-ordinates of the atoms in a protein obtained from wire models built to fit high-resolution *X*-ray data contain many large deviations in bond lengths and bond angles from their known standard values in model compounds. Further, there may be severe overlaps between atoms in non-bonded contact. In one of the first procedures designed to overcome this problem Diamond [39–41] has imposed 'standard geometry' on the protein and, starting with the rotation angles of the polypeptide backbone of the protein as obtained from the wire models, he has optimized the overall conformation to fit the electron-density maps. The resultant structure has good stereochemistry (since it was imposed) but may still have bad overlaps.

[34] E. Ralston and J. L. De Coen, *J. Mol. Biol.*, 1974, **83**, 393.
[35] K. Nagano, *J. Mol. Biol.*, 1974, **84**, 337.
[36] O. B. Ptitsyn and A. A. Rashin, *Doklady Akad. Nauk S.S.S.R.*, 1973, **213**, 473.
[37] O. B. Ptitsyn and A. A. Rashin, *Biophys. Chem.*, 1974, in the press.
[38] D. B. Wetlaufer, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 697.
[39] R. Diamond, *Acta Cryst.*, 1966, **21**, 253.
[40] R. Diamond, *Acta Cryst.*, 1971, **A27**, 436.
[41] R. Diamond, *J. Mol. Biol.*, 1974, **82**, 371.

The possibility of refining co-ordinates by minimizing the energy of the protein with respect to those co-ordinates has been discussed by Scheraga,[42] and was first tested by Levitt and Lifson on lysozyme.[43] More recently, a few more attempts to apply the energy refinement technique have been made.[44–46] Scheraga and co-workers developed procedures for energy refinement and applied them to lysozyme[47–49] and actinomycin D.[50] Finally, Levitt[51] has again refined lysozyme, testing various potential functions and calculating the agreement factor (R), as did Birktoft and Blow.[44]

One of the main goals of energy refinement is the refinement of low-resolution X-ray data (approximately 4—6 Å), to obtain atomic co-ordinates of the protein. So far this goal has not been achieved and even the energy refinement of wire-model co-ordinates obtained by high-resolution data has not been as efficient as the Diamond procedure and has not yielded structures in better agreement with X-ray (see discussion by Birktoft and Blow[44] and Levitt[51]).

The problems which must be overcome in order that energy refinement can become a standard tool for the crystallographer include: (i) the need to develop better potential functions; (ii) the lack of a dependable way of treating solvent effects; (iii) the development of more efficient minimization procedures; and (iv) the development of faster computers. The first two problems have been discussed in more detail in the recent review by Hagler and Lifson.[7]

Solvent Calculations.—Proteins are tightly packed structures from the interiors of which solvent molecules are excluded.[52] This emphasizes the importance of hydrophobic interaction in the folding and stability of globular proteins. Experimental studies on model systems continue to throw light on the nature of hydrophobic interactions,[53] but the complexity and uncertainty of the specific role of water in hydrophobic interactions still necessitate an empirical and very approximate approach to including such interactions in conformational energy calculations. Theories of water structure and its effect on macromolecular conformation have recently been reviewed by Hagler *et al.*[54]

[42] H. A. Scheraga, *Adv. Phys. Org. Chem.*, 1968, **6**, 103.
[43] M. Levitt and S. Lifson, *J. Mol. Biol.*, 1969, **46**, 269.
[44] J. J. Birktoft and D. M. Blow, *J. Mol. Biol.*, 1972, **68**, 187.
[45] K. Nishikawa, T. Ooi, Y. Isogai, and N. Saito, *J. Phys. Soc. Japan*, 1972, **32**, 1331.
[46] K. Nishikawa and T. Ooi, *J. Phys. Soc. Japan*, 1972, **32**, 1338.
[47] P. K. Warme, N. Go, and H. A. Scheraga, *J. Comput. Phys.*, 1972, **9**, 303.
[48] P. K. Warme and H. A. Scheraga, *J. Comput. Phys.*, 1973, **12**, 49.
[49] P. K. Warme and H. A. Scheraga, *Biochemistry*, 1974, **13**, 757.
[50] P. K. Ponnuswamy, R. F. McGuire, and H. A. Scheraga, *Internat. J. Peptide Protein Res.*, 1973, **5**, 73.
[51] M. Levitt, *J. Mol. Biol.*, 1974, **82**, 393.
[52] F. M. Richards, *J. Mol. Biol.*, 1974, **82**, 1.
[53] A. Wishnia and S. J. Lappi, *J. Mol. Biol.*, 1974, **82**, 77.
[54] A. T. Hagler, H. A. Scheraga, and G. Nemethy, *Ann. New York Acad. Sci.*, 1973, **204**, 51.

Krimm and Venkatachalam[55, 56] considered only the effects of water molecules strongly hydrogen bonded to the polypeptide. Conformational energy calculations were carried out in which water molecules were attached to the carbonyl groups of polyproline, and were allowed two angular degrees of freedom. Some interesting conclusions were drawn concerning the effect of an aqueous solvent on polyproline conformation, and these studies have been extended to the effect of methanol as a solvent.[56] However, these studies refer to specific binding of solvent to a polar site and thus do not apply to hydrophobic interactions in the currently accepted sense.

A more empirical approach is the solvation shell model, first applied to polypeptides by Gibson and Scheraga.[57] This model depends on the assumption that the free energy of hydration of a polypeptide can be described by the sum of the contributions of the free energies of hydration of individual atoms or groups. The maximum number of water molecules which can hydrate an atom by forming a single shell around the atom was estimated on the basis of space-filling models. The number of water molecules displaced during the formation of a 'hydrophobic bond' depends on the volume of the incoming atom or group and leads to an estimate of the strength of hydrophobic interactions in terms of the partial or total dehydration of the atoms involved. This procedure has recently been adopted[58, 59] to evaluate helical conformation energies for eight homopolypeptides. However, the theory of water on which the original model and parameters were based,[57] namely the Nemethy and Scheraga treatment,[60] has recently been reformulated and yields quite different results.[61–63]

The problems associated with calculations based on dehydration of specific atoms or small groups can to some extent be circumvented by directing attention to the total dehydration of much larger groups. Thus the experimental free energy for transferring an apolar side-chain from water to ethanol may be used as a first approximation to represent the strength of the 'hydrophobic bond' which results from the transfer of the side-chain from water into the fully dehydrated interior of a globular protein.[64] This is the assumption used in the work of Ptitsyn and Rashin,[36, 37] in which hydrophobic interactions of α-helices are calculated in order to simulate the refolding of myoglobin (in this work, the transfer free energies quoted by Tanford[65] in 1962 were used). This work is

[55] S. Krimm and C. M. Venkatachalam, *Proc. Nat. Acad. Sci. U.S.A.*, 1971, **68**, 2468.
[56] C. M. Venkatachalam and S. Krimm, ref. 6, p. 141.
[57] K. D. Gibson and H. A. Scheraga, *Proc. Nat. Acad. Sci. U.S.A.*, 1967, **58**, 420.
[58] A. J. Hopfinger, *Macromolecules*, 1971, **4**, 731.
[59] K. H. Forsythe and A. J. Hopfinger, *Macromolecules*, 1973, **6**, 423.
[60] G. Nemethy and H. A. Scheraga, *J. Chem. Phys.*, 1962, **36**, 3401.
[61] A. T. Hagler, H. A. Scheraga, and G. Nemethy, *J. Phys. Chem.*, 1972, **76**, 3229.
[62] A. T. Hagler, H. A. Scheraga, and G. Nemethy, *Ann. New York Acad. Sci.*, 1973, **204**, 51.
[63] B. Lentz, A. T. Hagler, and H. A. Scheraga, *J. Phys. Chem.*, 1974, in the press.
[64] Y. Nozaki and C. Tanford, *J. Biol. Chem.*, 1971, **246**, 2211.
[65] C. Tanford, *J. Amer. Chem. Soc.*, 1962, **84**, 4240.

important as the first attempt to predict the overall native conformation of a protein given its secondary (locally organized) structure, despite the fact that only hydrophobic interactions were taken into account. Considering the rough nature of the calculations, it is interesting that Ptitsyn and Rashin were able to discover a minimum energy conformation for the myoglobin molecule which coincided approximately with the native tertiary structure. A more complete calculation of the energy of interaction of helices has been presented by Silverman and Scheraga,[66] although this study was confined to two polyalanine α-helices with a more restricted mutual orientation.

Recently, more direct experimental data for interacting α-helices have been presented.[67] From these data it appears to be possible to evaluate the effect of dehydrating side-chains and replacing the water by interactions with another α-helix.

Quantum Mechanical Calculations.—The application of quantum mechanical methods to the study of theoretical aspects of protein structure is of necessity limited by the large amount of computer time needed to carry out *ab initio* calculations even in the case of small model systems.[68] Because of this, most applications to the conformations of model systems have employed approximate quantum mechanical methods such as CNDO/2, EHT, and PCILO.[69–71] One of the first applications of quantum mechanical methods was the calculation of the conformational energy map of the glycine dipeptide by Hoffman and Imamura[72] using the EHT method. Pullman and Maigret[73] have recently carried out PCILO calculations on alanine, glycine, and serine dipeptides, and compared their energy maps with those obtained by other methods. The PCILO method has also been extended to compounds of pharmacological interest.[74] Recently, the more accurate but very time-comsuming *ab initio* calculations have been carried out for the fragment $CHO \cdot NH \cdot CH_2 \cdot CONH_2$ as a model of the polypeptide backbone.[75] *Ab initio* calculations have also been carried out for a few points on the dipeptide map.[76]

In some cases, quantum mechanical calculations may contribute to our understanding of the form of some of the more problematic terms involved in the calculation of conformational energies. Pople and Radom[68] have

[66] D. N. Silverman and H. A. Scheraga, *Arch. Biochem. Biophys.*, 1972.
[67] B. R. Malcolm, *Progr. Membrane Sci.*, 1973. **7**. 183.
[68] J. A. Pople and L. Radom, ref. 6, p. 747.
[69] J. A. Pople and D. L. Beveridge, 'Approximate Molecular Orbital Theory', McGraw-Hill, New York, 1970.
[70] J. N. Murrell and A. J. Harget, 'Semi-empirical Self-consistent-field Molecular Orbital Theory of Molecules', Wiley–Interscience, New York, 1972.
[71] B. Pullman, *Internat. J. Quantum Chem.*, 1971, **4**, 319.
[72] R. Hoffman and A. Imamura, *Biopolymers*, 1969, **7**, 207.
[73] B. Pullman and B. Maigret, ref. 6, p. 13.
[74] B. Pullman and P. Courrière, ref. 6, p. 547.
[75] L. L. Shipman and R. E. Christofferson, *J. Amer. Chem. Soc.*, 1973, **95**, 1408.
[76] B. Pullman and A. Pullman, *Compt. rend.*, in the press.

applied *ab initio* calculations to model compounds in order to study intrinsic rotation potentials in biological molecules. They used both extended and minimal basis sets and compared the results obtained with each (note that the larger model compounds such as $NH_2 \cdot CH_2 \cdot CONH_2$ can only be treated on the basis of the minimal basis set). In another study using the IEHT method, an attempt was made to calculate the electrostatic contribution to the rotational potential functions.[77] Scheraga and co-workers[78–80] have used semi-empirical MO methods (mainly CNDO/2) as a basis for obtaining energy parameters for empirical energy functions. In particular, the partial charges as well as the form and parameters for the hydrogen bond potential were obtained. Ramachandran and co-workers[81] have investigated the effect of deviations from planarity of the peptide unit by carrying out CNDO/2 calculations on the model compound *N*-methylacetamide. The CNDO/2 results indicate that the minimum energy configuration corresponds to a non-planar conformation; it would be interesting to carry out this study using an extended basis set *ab initio* calculation.

Vibrational Free Energy.—The most stable state of a system corresponds to the state of minimum free energy. In the case of a globular protein, the system also includes the solvent. It is well known[3, 82] that the vibrational motion of the protein–solvent system contributes to both the enthalpy and entropy, although as a first approximation such considerations have usually been omitted.

Go and Scheraga[83] and Gibson and Scheraga[84] have considered the vibrational free energy as well as solvent effects with regard to polypeptides and proteins. As formulated by Go and Scheraga,[83] modes of vibration can be separated into hard and soft modes, the soft modes being associated with bond rotations (which require very little energy for their perturbation from equilibrium values) and the hard modes with bond stretching and the distortion of valence angles (which require much higher energies for their perturbation from equilibrium values). This treatment has been reviewed by Brant[3] and Scheraga[82] and the validity of the approximations involved has been discussed by Hagler and Lifson.[7]

Recently, a number of attempts have been made to take account of the vibrational free energy in calculations of polypeptide conformations. The effect of the configurational contribution on the conformations of dipep-

[77] R. Rein, T. J. Swissler, V. R. Krishnan, and G. R. Pack, ref. 6, p. 761.
[78] R. F. McGuire, F. A. Momany, and H. A. Scheraga, *J. Phys. Chem.*, 1972, **76**, 375.
[79] F. A. Momany, R. F. McGuire, J. F. Yan, and H. A. Scheraga, *J. Phys. Chem.*, 1970, **74**, 2424.
[80] J. F. Yan, F. A. Momany, R. Hoffman, and H. A. Scheraga, *J. Phys. Chem.*, 1970, **74**, 420.
[81] G. N. Ramachandran, A. V. Lakshminarayanan, and A. S. Kolaskai, *Biochem. Biophys. Acta*, 1973, **303**, 8.
[82] H. A. Scheraga, *Chem. Rev.*, 1971, **71**, 195.
[83] N. Go and H. A. Scheraga, *J. Chem. Phys.*, 1969, **51**, 4751.
[84] K. D. Gibson and H. A. Scheraga, *J. Physiol. Chem. Phys.*, 1969, **1**, 109.

tides has been calculated.[85] Note that the vibrational free energy actually involves two terms arising from configurational and momentum contributions.[3, 7, 82–84] More recently, new methods have been developed for calculating the conformational entropy of α-helices [86] (see also ref. 87). The vibrational free energy contribution of the soft modes of cyclic polypeptides has been calculated.[88] In a related development, Go and Scheraga have considered the problem of generating the different conformations of cyclic molecules [89] (including cyclic peptides).

Parameters for Conformational Energy Calculations.—Conclusions concerning the relative energies of different conformations are likely to be no more reliable than the least reliable parameter used in that calculation. A general approach to obtaining reliable parameters, the Consistent Force Field (CFF) method, has been discussed by Lifson.[90] Recently, a fast and efficient method of deriving potential parameters for use in conformational calculations from crystal structures of small model compounds has been described.[91]

A problem in the use of such parameters is the question of their transferability from the system from which they were derived to the system to which they are to be applied. It will therefore be necessary to test the parameters obtained by using them to predict the physicochemical properties of polypeptide systems or of small molecule analogues such as dipeptides. Suitable properties might include those measured by i.r. spectroscopy, n.m.r., c.d., and related techniques. For polypeptides, hydrodynamic measurements of the characteristic ratio could be used. Recently, prediction of the experimentally observed Kerr constant (obtained by birefringence studies) has been described as a potentially very sensitive test of the accuracy of conformational energy parameters.[92]

3 Mechanisms of Folding in Globular Proteins

Contributed by E. A. Carrey

Introduction.—In order that the functional state of globular proteins may be attained the genetic information must be further translated using the code relating primary amino-acid sequence to secondary and tertiary structure. Since the molecule must be largely in the native conformation for biological activity to be detected,[93] a great deal of information about its conformation is given by biological criteria such as enzyme activity,

[85] P. N. Lewis, F. A. Momany, and H. A. Scheraga, *Israel J. Chem.*, 1973, **11**, 121.
[86] N. Go, M. Go, and H. A. Scheraga, *Macromolecules*, 1974, **7**, 137.
[87] M. Go, N. Go, and H. A. Scherega, *J. Chem. Phys.* 1970, **52**, 2060.
[88] N. Go, P. Lewis, and H. A. Scheraga, *Macromolecules*, 1970, **3**, 628.
[89] N. Go and H. A. Scheraga, *Macromolecules*, 1973, **6**, 273.
[90] S. Lifson, ref. 6, p. 773.
[91] A. T. Hagler and S. Lifson, *Acta Cryst.*, 1974, in the press.
[92] R. T. Ingwall, E. A. Czurylo, and P. J. Flory, *Biopolymers*, 1974, **12**, 1135.
[93] D. B. Wetlaufer and S. Ristow, *Ann. Rev. Biochem.*, 1973, **42**, 135.

hormone activity (insulin[94]), and serological activity.[95, 96] Physical techniques involving polarimetric or spectrophotometric phenomena are often used in studies of the types of conformational transition undergone by proteins, but these give parameters such as the helix content or overall shape of a protein which need not be related to biological activity.

At present most experimental folding studies investigate the effect of a small number of denaturants on the protein in a simple solution with the ultimate aim of demonstrating the properties of the physiological system.

Much of the work published in 1973 demonstrates the diversity of techniques and interpretations rather than an underlying simplicity in protein folding, but it is apparent that interest in the subject is increasing.

The Folded Protein.—*The Influence of Sequence on Conformation.* Evolutionary forces would be expected to ensure that each part of the primary sequence is essential for the formation of a thermodynamically stable and functionally active structure with the dense packing typical of globular proteins.[97] Even 'conservative' substitutions in the primary sequence[98] may affect the overall stability of the molecule. A comparison of the equilibrium unfolding parameters of homologous myoglobins[99, 100] and haemoglobin subunits[100] demonstrates differences in free energies of stabilization which are difficult to relate to the number of amino-acid substitutions because of the complexity of side-chain interactions. It appears that loftiness in the phylogenetic tree is no guarantee of a more stable protein, although all these proteins have closely similar oxygen affinities.

The folding of derivatives of ribonuclease A which lack some carboxy-terminal residues has been investigated. The des(121—124) derivative is enzymically inactive,[101] and derivatives lacking five or six *C*-terminal residues[102] also have a lower ellipticity and a lower stability to unfolding by guanidinium chloride (GuCl) than the native form, although there is evidence of some folded structure. The S-protein resulting from subtilisin cleavage near the amino-terminus of ribonuclease has a stable conformation and will recombine with free S-peptide (fragment 1—20) forming several interactions, among them a salt bridge link.[103] The thermodynamics of the process have been investigated using free S-protein and S-peptide immobilized on agarose.[104]

[94] A. Cosmatos and P. G. Katsoyannis, *J. Biol. Chem.*, 1973, **248**, 7304.
[95] D. H. Sachs, A. N. Schechter, A. Eastlake, and C. B. Anfinsen, *Proc. Nat. Acad. Sci. U.S.A.*, 1972, **69**, 3790.
[96] C.-L. Lee and M. Z. Atassi, *Biochemistry*, 1972, **12**, 2690.
[97] P. J. Flory, in 'Polymerisation in Biological Systems', CIBA Foundation Symposium, 1972, **7**, 109.
[98] R. E. Dickerson, *Scientific American*, 1972, **226**, 58.
[99] D. Puett, *J. Biol. Chem.*, 1973, **248**, 4623.
[100] D. Puett, E. Friebele, and R. G. Hammonds, *Biochim. Biophys. Acta*, 1973, **328**, 261.
[101] V. G. Sacharovsky, I. I. Chervin, G. I. Yakovlev, S. M. Dudkin, M. Ya. Karpeisky, S. V. Shliapnikov, and V. F. Bystrov, *F.E.B.S. Letters*, 1973, **33**, 323.
[102] D. Puett, *Biochemistry*, 1972, **11**, 4304.
[103] F. M. Finn, J. Dadok, and A. A. Bothner-By, *Biochemistry*, 1972, **11**, 455.
[104] T. H. Gawronski and F. Wold, *Biochemistry*, 1972, **11**, 441, 449.

The work on staphylococcal nuclease T, the complex of large tryptic peptides, has been extended since last year's summary.[105] While certain regions of the sequence must be intact for the complex to be formed, it is shown[106] that analogues of peptide (6—48) with up to five *C*-terminal amino-acids deleted will form a complex (nuclease T′) with peptide (49—149). Nuclease T′ has also been used to investigate dynamic equilibria between folded and unfolded conformations which may apply to intact nuclease.[107] The rate of exchange between radioactively labelled and unlabelled fragment (50—149) is first-order with respect to complex concentration and involves a slow unfolding step followed by a rapid refolding. The conformational equilibrium of polypeptides has also been measured using an immunological approach,[95] since antibodies from anti-nuclease serum will react with 'native format determinants' generated by reversible spontaneous folding of the peptide fragments in solution. A conformational equilibrium constant is derived from first principles.

Structural Regions and Domains. In a review of the structure of several proteins,[108] Wetlaufer points out distinct regions formed by the folded backbone. If the eventual structure of proteins is at all closely related to the intermediately folded structures, then local regions of polypeptide backbone are good candidates for the site of nucleation events necessary for rapid folding of the molecule.[93] Several such events may take place, at least one in each globular intrachain region.

The nucleus, to which other structure is rapidly added, may be as small as eight residues but not larger than about 18 residues, the limits set by stability and time factors, respectively.[108] Evidence from this and other sources[109] suggests that the α-helix has an important role to play in nucleation.

The 'domain hypothesis' for immunoglobulins is a special case of intrachain ordered regions and is supported by genetic evidence, by enzymic cleavage of immunoglobulin IgG light chain into separate globular fragments and by *X*-ray crystallographic analysis.[110] Kinetic and equilibrium studies of the reversible denaturation of immunoglobulin IgG in GuCl[111] were interpreted in terms of independent transitions of two regions of similar size and composition, with similar dependence of equilibrium constant on GuCl concentration but governed by different rate constants for denaturation.[112] The transitions of separate fragments of the immuno-

[105] R. D. Ryder, ref. 2, p. 187.
[106] G. R. Sanchez, I. M. Chaiken, and C. B. Anfinsen, *J. Biol. Chem.*, 1973, **248**, 3653.
[107] H. Taniuchi, *J. Biol. Chem.*, 1973, **248**, 5164.
[108] D. B. Wetlaufer, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 697.
[109] B. Robson and R. H. Pain, 'Conformation of Biological Macromolecules and Polymers', Jerusalem Symposia on Quantum Chemistry and Biochemistry V, ed. E. D. Bergmann and A. Pullman, Israel Academy of Sciences and Humanities, 1973, p. 161.
[110] M. Schiffer, R. L. Girling, K. R. Ely, and A. B. Edmundson, *Biochemistry*, 1973, **12**, 4620.
[111] E. S. Rowe and C. Tanford, *Biochemistry*, 1973, **12**, 4822.
[112] A. Ikai and C. Tanford, *J. Mol. Biol.*, 1973, **73**, 145.

globulin IgE molecule enabled the irreversibly thermolabile region to be located at the *C*-terminus of the heavy chain, in one of the three 'domains' proposed for this portion.[113]

Protein–Protein Interactions.—A substantial amount of the work to date in this field concerning protein aggregation and polymerization, virus assembly, antigen–antibody reactions, and multi-subunit enzymes is summarized in ref. 114.

The importance of subunit interactions in the activity of enzymes can be illustrated by reference to triose phosphate isomerase.[115] Refolding after GuCl denaturation comprises two processes: the dissociated monomer refolds rapidly and reversibly, and this first-order process is followed by slower dimerization to give the active state, using interactions made available by correct folding of the monomers. Concentration-dependent spectral changes in insulin self-aggregation[116] have been distinguished from those resulting from the folding of monomers. The effects of pH and ionic strength on the equilibrium constant for dimerization are also described. A model for the association of casein subunits[117] proposes that electrostatic as well as hydrophobic interactions are important in the aggregation of κ-casein subunits with others forming a micelle-like structure with a hydrophobic centre and polar surface.

Direct calorimetry[118] of the association of heavy and light chains derived from several myeloma immunoglobulin IgG molecules shows not only a high affinity between the two types of chain (see also ref. 113) but also a wide range of enthalpy values for the association, suggesting that the relatively small variations in a limited region of the primary sequence between corresponding chains, which are otherwise identical, cause substantial variation in the non-covalent interactions between pairs of heavy and light chains. Even in the absence of inter-subunit disulphide bonds the IgG molecule can be reversibly unfolded in GuCl,[119] although the transition is broader than in the native molecule. Isolated heavy and light chains are much less stable in denaturant than when subunit interactions act upon them.

Another important subunit-stabilized system is that of haemoglobin. The α- and β-globin chains are known to combine successfully with the haem group (see also cytochrome *c*) and will form a 'half-filled' molecule comprising one species of globin chain and the other species of haem-containing subunit.[120] The interaction with a β-subunit will induce the

[113] K. J. Dorrington and H. Bennick, *J. Biol. Chem.*, 1973, **248**, 8378.
[114] 'Protein–Protein Interactions', ed. R. Jaenicke and E. Helmreich, Proceedings 23rd Mosbach Colloquium, Springer-Verlag, 1972.
[115] S. G. Waley, *Biochem. J.*, 1973, **135**, 165.
[116] R. S. Lord, F. Gubensek, and J. A. Rupley, *Biochemistry*, 1973, **12**, 4385.
[117] C. W. Slattery and R. Evard, *Biochim. Biophys. Acta*, 1973, **317**, 529.
[118] K. J. Dorrington and C. Kortan, *Biochem. Biophys. Res. Comm.*, 1974, **56**, 529.
[119] S. Lapanje and K. J. Dorrington, *Biochim. Biophys. Acta*, 1973, **322**, 45.
[120] M. Waks, Y. K. Yip, and S. Beychok, *J. Biol. Chem.*, 1973, **248**, 6462.

disordered α-chain to refold in the absence of free haem. This 'alloplex' interaction is distinct from the folding events initiated in α-chains by the haem moiety.

Folding Transitions.—Wetlaufer and Ristow [93] have written a critique of work concerning the process of folding (through which the stable conformation of proteins is attained) up to the end of 1972. Some important models for the kinetic processes involved in folding were discussed in this publication last year.[105, 112, 121] The 'two-state debate' as to whether the native and denatured states are the only significant species of single subunit globular proteins in solution will always be with us, but its value as a first approximation to the behaviour of many systems does not halt the search for kinetic and equilibrium intermediates.[122]

Ribonuclease. A sequential model for folding [121] including a rate-limiting nucleation step was able to predict several properties of the unfolding of ribonuclease in pH-jump experiments across the thermal transition. More recently, the authors have shown that both the fast and slow reactions of refolding of thermally denatured ribonuclease A by pH-jump yield native enzyme (as estimated by activity towards inhibitor and substrate molecules).[123] The different reactions appear to be due not to the formation of different products but perhaps to the presence of more than one species in the population of thermally unfolded ribonuclease A (see also ref. 135.)

While comparison of enthalpy values for the thermal unfolding of native and acylated ribonuclease [124] with calorimetrically determined values [125] supports an apparent two-state system, n.m.r. studies of this reaction at low pH indicate at least two regions unfolding at different temperatures.[126]

Cytochrome-c. The denaturation kinetics of cytochrome *c* in GuCl and the mathematical analysis [112, 127] fit most satisfactorily the mechanism

$$N \rightleftharpoons X_1 \rightleftharpoons D \rightleftharpoons X_2$$

where X_2 is a highly ordered 'dead-end' intermediate and X_1 is on the direct renaturation pathway. The major barrier to the conversion of X_2 into the native form is an unfavourable entropy change.[128]

On the other hand, the thermal unfolding of ferricytochrome-*c* has three kinetic phases at acid pH.[129] The time ranges of the two slower phases are comparable with that of the GuCl-induced unfolding but the process

[121] T. Y. Tsong, R. L. Baldwin, P. McPhie, and E. L. Elson, *J. Mol. Biol.*, 1972, **63**, 453.
[122] C. Tanford, ref. 97, p. 125.
[123] J.-R. Garel and R. L. Baldwin, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3347.
[124] S. J. Leach and H. Boyd, *Internat. J. Peptide Protein Res.*, 1973, **5**, 239.
[125] P. L. Privalov, E. I. Tiktopulo, and N. N. Khechinasvili, *Internat. J. Peptide Protein Res.*, 1973, **5**, 229.
[126] D. G. Westmoreland and C. R. Matthews, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 914.
[127] A. Ikai, W. W. Fish, and C. Tanford, *J. Mol. Biol.*, 1973, **73**, 165.
[128] R. W. Henkens and S. R. Turner, *Biochemistry*, 1973, **12**, 1618.
[129] T. Y. Tsong, *Biochemistry*, 1973, **12**, 2209.

is interpreted as before [121] in terms of a nucleation-dependent sequential process.

Thermodynamic studies on the opening of the haem crevice [130] suggest that the major energetic factors are large opposing entropy and enthalpy changes. The stability of the haem crevice is fundamental to that of the molecule; apo-cytochrome-*c* is disordered but can be induced to fold by binding to porphyrin or to a haem-bound peptide fragment.[131]

Carbonic Anhydrase and Penicillinase. In contrast to the usual close correlation between equilibrium parameters determined by different methods, bovine carbonic anhydrase in GuCl has slightly different transition midpoints, depending on whether circular dichroism (midpoint 1.4M-GuCl), difference spectroscopy (1.8M), or optical rotation was used.[132] Preliminary kinetic data indicate that refolding is complex, although the effect of the zinc atom must be taken into account.

Whether or not the distinction between the transitions is significant enough to propose three separate conformational changes in this enzyme, it is certainly reminiscent of a more pronounced effect observed in staphylococcal penicillinase.[109] Here the initial loosening of the structure and exposure of aromatic side-chains is nearly complete before the optical rotation transition (midpoint at 1.4M-GuCl) begins.

Lysozyme.—*Conformational Transitions.* An essentially two-state transition has been confirmed for GuCl denaturation of lysozyme under most conditions,[112, 133] and this is supported by the thermal unfolding in the presence of LiBr.[134] The activation parameters suggest that both thermal and GuCl-mediated unfolding proceed *via* the same activated state (with native conformation but some intramolecular hydrogen bonds broken) even though the denatured states are different. A study of the deuterium exchange properties of lysozyme at different temperatures [135] suggests that there are two denatured forms, differing in the extent and localization of hydrogen-bonding. Four classes of denatured state can be distinguished in lysozyme depending on the denaturant;[136] GuCl and urea give most complete disorder but there is some residual structure after heat treatment alone.

It has been suggested that the lack of 'dead-end' intermediates on the refolding pathway [133] is due to the directing influence of intact disulphide bonds. However, it can be seen [137] that the cross-linked peptide chain in strong denaturants is essentially a random coil, and that while the rate-limiting step in refolding the reduced chain is indeed the formation of

[130] L. S. Kaminsky, V. J. Miller, and A. J. Davison, *Biochemistry*, 1973, **12**, 2215.
[131] W. R. Fisher, H. Taniuchi, and C. B. Anfinsen, *J. Biol. Chem.*, 1973, **248**, 3188.
[132] K.-P. Wong and C. Tanford, *J. Biol. Chem.*, 1973, **248**, 8518.
[133] C. Tanford, K. C. Aune, and A. Ikai, *J. Mol. Biol.*, 1973, **73**, 185.
[134] S.-I. Segawa, Y. Husimei, and A. Wada, *Biopolymers*, 1973, **12**, 2521.
[135] M. Nakanishi, M. Tsuboi, and A. Ikegami, *J. Mol. Biol.*, 1973, **75**, 673.
[136] M. Kugiyama and C. C. Bigelow, *Canad. J. Biochem.*, 1973, **51**, 581.
[137] S. Lapanje and J. A. Rupley, *Biochemistry*, 1973, **12**, 2370.

disulphide links [138] a substantial amount of the chain is correctly folded [139] before this step.

The equilibrium transition zone of lysozyme unfolded by GuCl obtained by optical mixing spectroscopy [140] is similar to that obtained in other studies. The translational diffusion coefficient is 45% larger in the native enzyme than in the denatured form.

A reporter chromophore (whose spectrum is not affected by pH changes) attached to a Tyr residue does not affect the alkaline denaturation of lysozyme [141] and has been used to find the pK for the transition. It is suggested that the ionization of another Tyr residue is closely linked with conformational change (see p. 222).

The primary sequence of α-lactalbumin is known to be similar to that of lysozyme and the structure has been deduced to be so; deuterium exchange experiments [142] show that fewer hydrogen bonds are formed in α-lactalbumin and that the midpoint of the thermal transition is considerably lower than in lysozyme. Similarly, the equilibrium and kinetics of unfolding by GuCl [143] are appropriate to a protein with a less stable native state than lysozyme, with ΔG^0 about half of the latter. The kinetics show a faster (but still apparently two-state) transition than in lysozyme.

Nucleation. A system has been devised [138] in which reduced lysozyme is rapidly reactivated by a mixture of reduced and oxidized thiols in physiological concentrations. The regeneration process was examined by quenching the partly folded products in acid, alkylating free thiol groups and producing two-dimensional 'fingerprints' of the enzymic digests.[139] The large number of disulphide-paired peptides expected from a random search by the chain was not found, and the conclusion was that the restriction of pairing results from nucleation steps early in the refolding process, leading to a much restricted choice of three-dimensional structures for subsequent steps.

The pathways available for the folding of a protein are constrained by the necessity to fold in a biologically feasible time, and such a necessity may prohibit the attainment of the global free energy minimum if that structure is not kinetically accessible. That similar constraints also act on denaturation has been proposed by Wetlaufer's group using their glutathione system for the regeneration of reduced lysozyme,[144] but this time at high temperatures. Although both reduced and oxidized lysozyme are denatured at these temperatures, in the regeneration mixture the thermodynamically unstable native form (with enzyme activity) is a transient intermediate on

[138] V. P. Saxena and D. B. Wetlaufer, *Biochemistry*, 1970, **9**, 5015.
[139] S. S. Ristow and D. B. Wetlaufer, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 544.
[140] S. B. Dubin, G. Feher, and G. B. Benedek, *Biochemistry*, 1973, **12**, 714.
[141] A. A. Aboderin, E. Boedefeld, and P. L. Luisi, *Biochim. Biophys. Acta*, 1973, **328**, 31.
[142] H. Takesada, M. Nakanishi, and M. Tsuboi, *J. Mol. Biol.*, 1973, **77**, 605.
[143] S. Sugai, H. Yashiro, and K. Nitta, *Biochim. Biophys. Acta*, 1973, **328**, 35.
[144] D. B. Wetlaufer, E. Kwok, W. L. Anderson, and E. R. Johnson, *Biochem. Biophys. Res. Comm.*, 1974, **56**, 380.

the pathway to the disordered, cross-linked chain, with a maximum yield after *ca.* 10 min at 66 °C. This 'kinetic determinism' has interesting implications if it can be found in other proteins.

Albumins.—These proteins have been investigated using physical parameters which are more dependent on large conformational changes. The mixture of components resulting from heat and alkali treatment of bovine serum albumin can be separated by gel electrophoresis;[145] one component appeared to result from an intramolecular rearrangement of disulphide bonds and had a slightly altered sedimentation coefficient, while oligomers were formed by intermolecular disulphide exchange.

Heat denaturation of a protein in unbuffered solution results in an abrupt change of pH owing to the normalization in pK values of side-chains not fully exposed to solvent in the native state. A study of 20 proteins, with especial reference to egg albumin,[146] gives melting temperatures in agreement with other methods, and this technique works even if the protein coagulates.

Dilatometry has also been used to investigate the thermal transitions of egg and serum albumins, lysozyme, and β-lactoglobulin.[147] Dilatometric measurements of protein solutions depend upon the difference between the partial specific volumes of native and denatured proteins in solution. Thus, the changes in the molar volume of bovine serum albumin in GuCl [148] have been interpreted as an increase to a maximum in 2M-GuCl as polar surface groups are surrounded by denaturant ions, a decrease during the unfolding transition as solvent penetrates the structure, and a further increase after unfolding is complete. Similar studies have investigated the structure–volume relationships of serum albumin and myoglobin [149] and haemoglobin [150] during acid–base titration in water and denaturing agents.

The Nature of Denaturation.—The change of volume of metmyoglobin obtained from dilatometric measurements (ΔV = 98 ml/mol protein) [149] is similar to that obtained by denaturation under pressure [151] in various conditions of pH and temperature. Other parameters obtained from this work are, however, less reassuring; in particular the heat capacity, which was expected to decrease at higher pressures, does not do so. Other properties are out of line if the interior of the molecule resembles a liquid or solid hydrocarbon, and the thermodynamics of denaturation are assumed to be due solely to the exposure of internal apolar groups to water.

[145] K. Aoki, K. Sato, S. Nagaoka, M. Kamada, and K. Hiramatsu, *Biochim. Biophys. Acta*, 1973, **328**, 323; K. Aoki and S. Nagaoka, *ibid.*, p. 334.
[146] H. B. Bull and K. Breese, *Arch. Biochem. Biophys.*, 1973, **156**, 604; **158**, 681.
[147] H. B. Bull and K. Breese, *Biopolymers*, 1973, **12**, 2351.
[148] S. Lapanje and J. Skerjanc, *Biochem. Biophys. Res. Comm.*, 1974, **56**, 338.
[149] S. Katz, J. K. Crissman, and J. A. Beall, *J. Biol. Chem.*, 1973, **248**, 4840.
[150] S. Katz, J. A. Beall, and J. K. Crissman, *Biochemistry*, 1973, **12**, 4180.
[151] A. Zipp and W. Kauzmann, *Biochemistry*, 1973, **12**, 4217.

The solvation of side-chains transferred to GuCl is complicated by binding of the denaturant to the protein. The enthalpy change is less negative at 30 °C than at 25 °C when lysozyme is transferred to more concentrated solutions of GuCl.[152] C.d. studies on several polypeptides lead to different conclusions about the type of binding. One [153] finds some helical forms of charged homopolymers whose stability is enhanced at higher temperatures in 6M-GuCl, and concludes that it is an effect involving solvent ions and the polyelectrolyte homopolymer. On the other hand,[154] evidence from several proteins and polypeptides, notably poly-L-Pro II, suggests univalent hydrogen-bonding between backbone carbonyl groups and GuCl or urea molecules. Similar bonding is proposed for the interaction of poly-L-Ala and trifluoroacetic acid.[155]

The effect of hydrophobic interactions on the thermal stability of proteins can be seen by the destabilizing effect of univalent alcohols, the greater effect being given by more hydrophobic molecules,[156] and the stabilizing effect of polyvalent alcohols, which may intensify hydrophobic interactions inside the protein.[157]

Sodium dodecyl sulphate (SDS) and its homologues [158] not only mediate a change to a higher order of helix content [159] but also stabilize the protein structure against pH unfolding.[160] Strong binding of SDS to positive side-chains on the protein is suggested,[158] with the helical conformation stabilized by hydrophobic interactions along the alkyl chain of the surfactant.

Amino-acid Residues involved in Denaturation.—Perutz [161] has investigated the crystallographic structure of haemoglobin and pinpointed the amino-acid residues whose titration is responsible for conformational changes in alkali denaturation. He interprets the species variations in stability in terms of substitutions of a few defined amino-acids in the homologous proteins. Similar deductions from the structure of the insulin dimer [116] allow the tyrosine residues which are buried during dimer interactions to be distinguished from other tyrosine residues.

The effect of a change of pH from 4.2 to 6.7 was actually examined in crystals of α-chymotrypsin [162] and the difference map demonstrates that this transition, too, is dependent on a small number of polar side-chains whose interactions are important in the crystal structure.

[152] S. Lapanje, *F.E.B.S. Letters*, 1973, **31**, 67.
[153] M. Cortijo, B. Panijpan, and W. B. Gratzer, *Internat. J. Peptide Protein Res.*, 1973, **5**, 179.
[154] M. L. Tiffany and S. Krimm, *Biopolymers*, 1973, **12**, 575.
[155] P. Combelas, C. Garrigon-Lagrange, and J. Lascombe, *Biopolymers*, 1973, **12**, 611.
[156] R. M. Parodi, E. Bianchi, and A. Ciferri, *J. Biol. Chem.*, 1973, **248**, 4047.
[157] S. Y. Gerlsma and E. R. Stuur, *Internat. J. Peptide Protein Res.*, 1972, **4**, 377.
[158] I. Satake and J. T. Yang, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 930; A. H. Hunt and B. Jirgensons, *Biochemistry*, 1973, **12**, 4435.
[159] B. Jirgensons, *Biochim. Biophys. Acta*, 1973, **317**, 131.
[160] J. Steinhardt and N. Stocker, *Biochemistry*, 1973, **12**, 1789.
[161] M. F. Perutz, *Nature*, 1974, **247**, 341.
[162] R. L. Vandlen and A. Tulinsky, *Biochemistry*, 1973, **12**, 4193.

4 Nuclear Magnetic Resonance

Contributed by H. W. Rattle

Introduction.—From being a relative novelty in the study of peptides and proteins a few years ago, n.m.r. has firmly established itself as a technique which, although often difficult to use in this field, can produce information obtainable in no other way. An introduction to current techniques in the n.m.r. of biological molecules, together with examples on their application to enzymes and enzyme systems, is provided in a recent book.[163] Some work which will be of considerable interest to all whose n.m.r. studies are beset by poor resolution of numerous overlapping resonances is described in ref. 164. By means of the data manipulation and difference spectroscopy which is possible using the new generation of Fourier transform spectrometers with their associated computers, it is possible to improve the resolution of a spectrum by up to a factor of two, which is often sufficient to convert a broad featureless envelope into a series of resolved peaks. In addition to such 'convolution difference' methods, the use of difference spectroscopy in detecting spectral changes brought about by the use of paramagnetic line shift and broadening reagents is discussed.

Peptide Structure.—Among the parameters which are determinable using n.m.r. are the dihedral angle of a C—C bond, and the *cis* or *trans* state of a peptide bond. Theoretical calculations of coupling constants for *cis* and *trans* forms of model compounds have been published,[165] as have experimental data,[166, 167] while a revised equation for proton–proton coupling constants in terms of the dihedral angle of $C_\alpha—C_\beta$ bonds in amino-acids is:[168]

$$J = 11.0 \cos^2 \theta - 1.4 \cos \theta + 1.6 \sin^2 \theta$$

Identification of *cis* and *trans* forms of the peptide bonds of proline derivatives by ^{13}C resonance has been described.[169, 170] The vastly improved effective sensitivity of Fourier transform spectrometers is now enabling natural-abundance ^{13}C studies of proteins to be made; some preliminary work on amino-acids has been described in earlier reviews, and this has now been extended to assignments of tryptophan resonances,[171] to a number of apolar and charged residues in the pentapeptide Gly-Gly-X-

[163] R. A. Dwek, 'N.M.R. in Biochemistry: Application to Enzyme Systems', Clarendon, Oxford, 1973.
[164] I. D. Campbell, C. M. Dobson, R. J. P. Williams, and A. V. Xavier, *J. Magn. Resonance*, 1973, **11**, 172.
[165] V. N. Solkan and V. E. Bystrov, *Tetrahedron Letters*, 1973, 2261.
[166] D. B. Davies and M. A. Khaled, *J.C.S. Perkin II*, 1973, 1651.
[167] D. B. Davies and M. A. Khaled, *Tetrahedron Letters*, 1973, 2829.
[168] K. D. Kopple, G. R. Wiley, and R. Tauki, *Biopolymers*, 1973, **12**, 627.
[169] D. E. Dorman and F. A. Bovey, *J. Org. Chem.*, 1973, **38**, 2379.
[170] R. Deslauriers, R. Walter, and I. C. P. Smith, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 244.
[171] J. H. Bradbury and R. S. Norton, *Biochim. Biophys. Acta*, 1973, **328**, 10.

Gly-Gly,[172, 173] and the titration of histidine.[174] Another nucleus which may in future prove useful is ^{15}N.[175] Sequences of small peptides (up to penta- or hexa-peptides in some cases) have been determined using ^1H n.m.r.[176, 177] An application of lanthanide shift reagents to a dipeptide is found in a ^1H n.m.r. study of dipeptides representing parts of actinomycin.[178]

The possibility of determining dihedral angles and peptide bond isomerism makes cyclic peptides an attractive group of compounds for n.m.r. study. Of the synthetic cyclic peptides investigated in 1973 we may mention *cyclo*-(Gly-L-Pro-Gly)$_2$ which has been the subject of two separate studies by ^1H and ^{13}C resonance,[179, 180] the cyclic hexapeptide *cyclo*-(L-Ala)$_5$-D-Ala,[181] which is in a rapidly exchanging conformational equilibrium in solution, some peptides containing the dipeptide D-Phe-L-Pro [182] which behaves rather as it does in gramicidin-S, and an application of shift reagents to *cyclo*-(L-Pro-L-Pro) and (L-Pro-D-Pro).[183]

Polypeptide Conformation.—In the field of synthetic polypeptides of higher molecular weight, the 'two-peak controversy' regarding the origin of the separate resonance peaks observed simultaneously in the spectra of partly α-helical polypeptides grinds wearily on. The field of polypeptide n.m.r. has been extensively covered, though not strictly reviewed,[184] and the authors conclude that the origin of two peaks in the spectrum of a partly helical polypeptide results from a segregation of the molecules into largely random-coil and largely helical populations by virtue of polydispersity in their molecular weight. Considerable weight is added to their argument by a paper [185] describing fractionation of poly-γ-benzyl-L-glutamate into fractions of low polydispersity and subsequent loss of the two-peak effect. A similar view has been expressed,[186] but an entirely different conclusion

[172] P. Keim, R. A. Vigna, R. C. Marshall, and F. R. N. Gurd, *J. Biol. Chem.*, 1973, **248**, 6104.
[173] P. Keim, R. A. Vigna, J. S. Morrow, R. C. Marshall, and F. R. N. Gurd, *J. Biol. Chem.*, 1973, **248**, 7811.
[174] W. F. Reynolds, I. R. Peat, M. H. Freedman, and J. R. Lyerla, *J. Amer. Chem. Soc.*, 1973, **95**, 328.
[175] J. A. Sogn, W. A. Gibbons, and E. W. Randall, *Biochemistry*, 1973, **12**, 2100.
[176] T. Kasai and S. Sakamura, *Agric. and Biol. Chem. (Japan)*, 1973, **37**, 2155.
[177] M. Antennis and J. Gelan, *J. Amer. Chem. Soc.*, 1973, **95**, 6502.
[178] A. B. Mauger, W. J. Rzeszotaski, and R. A. Ford, *Org. Magn. Resonance*, 1973, **5**, 231.
[179] Ch. Grathwohl, R. Schwyzer, A. Tun-Kyi, and K. Wuthrich, *F.E.B.S. Letters*, 1973, **29**, 271.
[180] L. G. Pease, C. M. Deber, and E. R. Blout, *J. Amer. Chem. Soc.*, 1973, **95**, 258.
[181] A. E. Tonelli and A. I. R. Brewster, *Biopolymers*, 1973, **12**, 193.
[182] K. D. Kopple, A. Go, T. J. Schamper, and C. S. Wilcox, *J. Amer. Chem. Soc.*, 1973, **95**, 6090.
[183] P. E. Young, V. Madison, and E. R. Blout, *J. Amer. Chem. Soc.*, 1973, **95**, 6142.
[184] E. M. Bradbury, P. D. Cary, C. Crane-Robinson, and P. G. Hartman, *Rev. Pure Appl. Chem.*, 1973, **36**, 53.
[185] K. Nagayama and A. Wada, *Biopolymers*, 1973, **12**, 2443.
[186] E. M. Bradbury, C. Crane-Robinson, and P. G. Hartman, *Polymer*, 1973, **14**, 543.

is reached on the basis of apparently quite similar experiments.[187, 188] The different chemical shifts for backbone protons in the random and helical forms, which have hitherto been explained on a basis of protonation of one of the forms by the solvent, is shown [189–191] not to be due to protonation, as it is produced in non-protonating solvents such as DMSO and DMF. The polypeptide used in most of these studies, polybenzylglutamate, has also been the subject of relaxation and nuclear Overhauser studies [192, 193] and of studies in the liquid crystal form [194] and by 'magic angle' rotation.[195] The solvent dependence of linewidths of various polypeptides has been reported.[196] Other polypeptide studies reported in 1973 have concerned the ^{13}C resonance spectrum of polylysine,[197] and polyalanine,[198] polyproline,[199] and a determination of ΔG for *cis–trans* isomerism using poly-(*N*-alkylglycine).[200]

The study of all these model compounds is, of course, only a preliminary to the study of molecules of biological significance. The application of n.m.r. to the study of peptide hormone structure and receptor binding, including ^{2}D, ^{13}C, and ^{15}N studies, has been reviewed.[201] A very 'biological' study of prodigiosin, a red metabolite of *Serrata marcescens*,[202] follows the incorporation patterns of ^{13}C-labelled Ala, Pro, Ser, and Gly by Fourier transform spectroscopy. A new subject for n.m.r. study has been angiotensin II, where several previously proposed structures are excluded by studies of amide NH and α-CH coupling constants [203, 204] and for which a hydrogen–deuterium exchange study [205] reveals two slowly exchanging amides belonging to Val(3) and Val(5). Structural studies of oxytocin reveal three energetically favourable conformations [206] and give assignments

[187] J. B. Milstien and J. A. Ferretti, *Biopolymers*, 1973, **12**, 2335.
[188] J. A. Ferretti and R. L. Jernigan, *Macromolecules*, 1973, **6**, 687.
[189] E. M. Bradbury, C. Crane-Robinson, L. Paolillo, and P. Temussi, *Polymer*, 1973, **14**, 303.
[190] L. Paolillo, P. Temussi, E. Trivellone, E. M. Bradbury, and C. Crane-Robinson, *Macromolecules*, 1973, **6**, 831.
[191] E. M. Bradbury, C. Crane-Robinson, L. Paolillo, and P. Temussi, *J. Amer. Chem. Soc.*, 1973, **95**, 1683.
[192] A. Allerhand and E. Oldfield, *Biochemistry*, 1973, **12**, 3428.
[193] M. Chien, E. T. Samulcki, and C. G. Wade, *Macromolecules*, 1973, **6**, 638.
[194] E. Iizuka, *J. Phys. Soc. Japan*, 1973, **34**, 1054.
[195] D. Doskocilova and B. Schneider, *Macromolecules*, 1973, **6**, 76.
[196] J. H. Rai, W. G. Miller, and R. G. Bryant, *Macromolecules*, 1973, **6**, 262.
[197] H. Saito and I. C. P. Smith, *Arch. Biochem. Biophys.*, 1973, **158**, 154.
[198] M. Goodman, F. Toda, and N. Veyama, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 331.
[199] D. E. Dorman, D. A. Torchia, and F. A. Bovey, *Macromolecules*, 1973, **6**, 80.
[200] M. Sisido, Y. Imanishi, and T. Higashimura, *Biopolymers*, 1973, **12**, 2375.
[201] P. Fromageot, Tran Dinh Son, and S. Fermandjian, *Bull. Inform. Sci. Tech. Commis. Energ. At.*, 1973, No. 183, p. 17.
[202] H. H. Wasserman, R. J. Sykes, P. Peverada, C. K. Shaw, R. I. Cushley, and S. R. Lipsky, *J. Amer. Chem. Soc.*, 1973, **95**, 6874.
[203] G. R. Marshall, H. E. Bosshard, W. H. Vine, and J. D. Glickson, *Nature New Biol.*, 1973, **245**, 125.
[204] J. D. Glickson, J. Dadok, and G. R. Marshall, *Biochemistry*, 1974, **13**, 11.
[205] H. E. Bleich, R. E. Galardy, M. P. Printz, and L. C. Craig, *Biochemistry*, 1973, **12**, 4950.
[206] A. Brewster, V. J. Hruby, J. A. Glasel, and A. E. Tonelli, *Biochemistry*, 1973, **12**, 5294.

of all 43 resonances in the spectrum.[207] Conformational studies of valinomycin[208] and side-chain motion studies of gramicidin-S[209] have also been reported.

Non-haem Proteins.—In the area of non-haem proteins, the freedom of the histidine C-2 and C-4 proton resonances from the rest of the proton spectrum and the association of His residues with active sites in enzymes has led to further results. Histidine titration studies[210] of RNase A in the presence and absence of phosphate support the hypothesis that titration curves may be perturbed by the interaction of the imidazole groups with nearby carboxy-groups in the molecule; histidine titration studies on the same molecule[211] and on soybean trypsin inhibitor[212] have also been reported. A particularly interesting study[213] of the single histidine of a serine protease shows that, over the pH range 4—6.7, both the histidine and its neighbouring aspartic acid are un-ionized. From this and other considerations the authors were able to propose a new catalytic mechanism for the molecule in which the histidine has a double role, firstly to ensure a hydrophobic environment for the carboxylate group of the Asp and then to provide a relay for proton transfer from the serine —OH to the aspartic acid carboxylate group.

The catalytic subunit of *E. coli* aspartate transcarbamylase is inhibited by the binding of succinate, and studies[214] of the transverse relaxation rate T_2 and its variation with temperature over a range of pH values have now led to a proposal for the binding mechanism involving two carboxy-groups having separate roles. Glutamine binding to a glutamine binding protein from *E. coli*[215] and AMP binding to a modified phosphorylase *b*[216] have also proved to be susceptible to n.m.r. investigation. The ready availability of small molecules labelled with ^{13}C has facilitated studies of the binding of ^{13}CN[217] and $^{13}CO_2$ and $H^{13}CO_3$[218] to carbonic anhydrase. ^{13}C enrichment has been used in conjunction with ^{2}D labelling to find further details of the mobility of the His residues of tryptophan synthetase α subunit from *E. coli*,[219] while the much more difficult method of natural-

[207] A. Brewster, V. J. Hruby, A. F. Spatola, and F. A. Bovey, *Biochemistry*, 1973, **12**, 1643.
[208] E. Grell, T. Funck, and H. Sauter, *J. Biochem. (Japan)*, 1973, **34**, 415.
[209] A. Allerhand and R. A. Komoroski, *J. Amer. Chem. Soc.*, 1973, **95**, 8228.
[210] J. S. Cohen, J. H. Griffin, and A. N. Schechter, *J. Biol. Chem.*, 1973, **248**, 4305.
[211] H. Rueterjans, *Deut. Gesellschaft Chem. Apparatew. Monogr.*, 1973, **71**, 219.
[212] J. L. Markley, *Biochemistry*, 1973, **12**, 2245.
[213] M. W. Hunkapiller, S. H. Smallcombe, D. R. Whitaker, and J. H. Richards, *Biochemistry*, 1973, **12**, 4732.
[214] C. B. Beard and P. G. Schmidt, *Biochemistry*, 1973, **12**, 2255.
[215] G. Kreishman, D. Robertson, and C. Ho, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 18.
[216] R. Gaspar and S. Damjanovich, *Biochim. Biophys. Acta*, 1973, **315**, 191.
[217] J. Feeney, A. S. V. Burgen, and E. Grell, *European J. Biochem.*, 1973, **34**, 107.
[218] S. Koenig, R. Brown, T. Needham, and N. Matwiyoff, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 624.
[219] D. T. Browne, G. L. Kenyon, E. L. Packer, H. Sternlicht, and D. M. Wilson, *J. Amer. Chem. Soc.*, 1973, **95**, 1316.

abundance ^{13}C resonance has again been applied to lysozyme, yielding partial assignments [220, 221] and some titration curves.[221]

Relaxation enhancement, either of solvent water protons or of protein protons, by the addition of a paramagnetic metal ion to the protein solution or complex has produced some interesting results. For example, it has been used [222] to detect an apparent competition for inhibition of a threonine-sensitive aspartokinase/homoserine dehydrogenase complex between threonine and a Mn–ATP complex. Concanavalin A has been subjected to several paramagnetic studies, including proton relaxation enhancement which reveals [223, 224] a pH-dependent rate of structural rearrangement at the manganese binding site following the addition of manganese. In the complex of the Mn and Zn derivatives of concavalin A with a ^{13}C-enriched sugar, the selective effect of the paramagnetic ion on the longitudinal relaxation rate of sugar resonances has enabled the orientation of the sugar to be calculated.[225] Other paramagnetic relaxation enhancement studies include those on pyruvate kinase,[226] on the ^{31}P relaxation in a Mn–alkaline phosphatase–phosphate complex,[227] and on the location of a histidine residue relative to a copper atom in human and bovine superoxide dismutase.[228]

The use of halide atoms, particularly ^{35}Cl and ^{19}F as heteronuclear labels in protein n.m.r., has been discussed [229] with reference to some studies of a synthetic copolypeptide of glutamic acid and cysteine in which the relaxation of ^{35}Cl in solvent sodium chloride was observed. Quadrupole relaxation has been used to discriminate between metal-co-ordinative and general anion binding of ^{35}Cl to proteins [230] and ^{19}F resonance to investigate the binding of *N*-trifluoroacetyl derivatives of phenylalanine to chymotrypsin,[231] while both ^{81}Br and ^{79}Br resonances have been employed in studies of the —SH groups of horse methaemoglobin.[232] Other studies of non-haem proteins include bovine serum albumin/surfactant interactions [233] and the states of water in lysozyme crystals.[234]

[220] A. Allerhand, R. F. Childres, and E. Oldfield, *Biochemistry*, 1973, **12**, 1335.
[221] M. H. Freedman, J. R. Lyerla, I. M. Chaiken, and J. S. Cohen, *European J. Biochem.*, 1973, **32**, 215.
[222] R. S. Erlich and M. Takahashi, *Biochemistry*, 1973, **12**, 4309.
[223] B. H. Barber and J. P. Carver, *J. Biol. Chem.*, 1973, **248**, 3353.
[224] E. Meirovitch and A. J. Kalb, *Biochim. Biophys. Acta*, 1973, **303**, 258.
[225] C. F. Brewer, H. Sternlicht, D. M. Marcus, and A. P. Grollman, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1007.
[226] M. Flashner, I. Tamir, A. Mildvan, H. Meloche, and M. Coon, *J. Biol. Chem.*, 1973, **248**, 3419.
[227] R. Zukin, D. Hollis, and G. Gray, *Biochim. Biophys. Res. Comm.*, 1973, **53**, 238.
[228] A. M. Stokes, H. A. O. Hill, W. H. Bannister, and J. V. Bannister, *F.E.B.S. Letters*, 1973, **32**, 119.
[229] D. E. Carlstrom, W. G. Miller, and R. G. Bryant, *J. Phys. Chem.*, 1973, **77**, 2759.
[230] J. R. Norne, T. E. Bull, R. Einasson, B. Lindman, and M. Zepperzauer, *Chem. Scripta*, 1973, **3**, 142.
[231] B. C. Nicholson and T. M. Spotswood, *Austral. J. Chem.*, 1973, **26**, 135.
[232] T. R. Collins, Z. Starcuk, A. H. Burr, and E. J. Wells, *J. Amer. Chem. Soc.*, 1973, **95**, 1649.
[233] J. Oakes, *European J. Biochem.*, 1973, **36**, 553.
[234] J. E. Jentoft and R. G. Bryant, *J. Amer. Chem. Soc.*, 1974, **96**, 297.

Haem Proteins.—In some ways, haem proteins present easier problems to the n.m.r. spectroscopist than others because the presence of the porphyrin ring and iron atom spreads out the resonances. The binding of carbon monoxide to haemoglobin trifluoroacetonylated at cysteine $\beta 93$ has been shown [235] to be preferential for the α-chains, and differences between the ^{13}C resonances of ^{13}CO bound to the α- and β-chains have been demonstrated.[236, 237] The latter study also revealed the presence of abnormal haemoglobins in some samples from rabbit by the presence of a third resonance; it should be possible to study the redox behaviour of α-chains, β-chains, and abnormal chains separately by this method. The dynamic states of human ferrihaemoglobin from relaxation studies have been discussed.[238] One of the problems facing biochemists is the determination of intracellular pH in such cells as erythrocytes, and a method using ^{31}P resonance of intracellular phosphates has been described.[239]

A series of papers [240–242] on cyanoferrimyoglobins gives a large amount of assignment and relaxation data, including the effects of such changes as removing the haem group and changing the pH. Assignments published for the low-field spectrum of horse ferrocytochrome *c* have been used to show that inorganic phosphate, ADP, and ATP all bind in the region of His 26.[243]

Water.—The state of water in protein systems has always been susceptible to n.m.r. investigation, and the method is still attracting a surprising amount of attention. No less than four studies on hydrated collagen have been published [244–247] and wool keratin has also come in for attention.[248] The state of water in whole tissues has also been reported,[249–251] although it is doubtful whether any but the most general conclusions may be drawn from the results.

[235] W. H. Huestis and M. A. Raftery, *Biochemistry*, 1973, **12**, 2531.
[236] P. J. Vergamini, N. A. Matwiyoff, R. C. Wohl, and T. Bradley, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 453.
[237] N. A. Matwiyoff, P. J. Vergamini, T. E. Needham, C. T. Gregg, J. A. Volpe, and W. S. Caughey, *J. Amer. Chem. Soc.*, 1973, **95**, 4429.
[238] G. Pifat, S. Maricic, and S. Grundja, *Biopolymers*, 1973, **12**, 905.
[239] R. B. Moon and J. H. Richards, *J. Biol. Chem.*, 1973, **248**, 7276.
[240] A. M. Nigen, P. Keim, R. C. Marshall, J. S. Morrow, R. A. Vigna, and F. R. N. Gurd, *J. Biol. Chem.*, 1973, **248**, 3724.
[241] A. M. Nigen and F. R. N. Gurd, *J. Biol. Chem.*, 1973, **248**, 3708.
[242] A. M. Nigen, P. Keim, R. C. Marshall, V. Glushko, P. J. Lawson, and F. R. N. Gurd, *J. Biol. Chem.*, 1973, **248**, 3716.
[243] E. Stellwagen and R. G. Shulman, *J. Mol. Biol.*, 1973, **75**, 683.
[244] B. M. Fung and S. C. Wei, *Biopolymers*, 1973, **12**, 1053.
[245] C. Migchelsen and H. J. C. Berendsen, *J. Chem. Phys.*, 1973, **59**, 296.
[246] R. E. Dehl, *Biopolymers*, 1973, **12**, 2329.
[247] P. Lindner and E. Forslind, *Chem. Scripta*, 1973, **3**, 57.
[248] L. J. Lynch and K. H. Marsden, *J. Colloid Interface Sci.*, 1973, **42**, 209.
[249] R. K. Outhred and E. P. George, *Biophys. J.*, 1973, **13**, 97.
[250] P. S. Bolton, K. J. Packer, and T. C. Selwood, *Biochim. Biophys. Acta*, 1973, **304**, 56.
[251] R. E. Block, *F.E.B.S. Letters*, 1973, **34**, 109.

5 Spin Labels

Contributed by P. Knowles

The main conclusion to emerge from the review of papers appearing during 1973 is that careful studies of real biochemical significance are being attempted. Many papers of doubtful value emerged during the 'band-wagon' days of spin labelling (1966—1972). Perhaps this was a reflection of the awe with which most biochemists regard new biophysical techniques. Most of the papers discussed here focus attention on the important questions about protein structure and action which need answering whilst at the same time pointing out the limitations of the spin label method. An excellent textbook on the application of n.m.r. to enzymes, which includes a chapter on spin labelling, has been written by Dwek.[252]

Haemoglobin.—Further spin label electron paramagnetic resonance (e.p.r.) and nuclear magnetic resonance (n.m.r.) studies have been reported on haemoglobin Kempsey.[253] This mutant form has the aspartic acid residue at β-99 replaced by asparagine and shows low co-operativity for oxygen binding. The cysteines of β-93 were labelled with *N*-(1-oxyl-2,2,6,6-tetramethyl-4-piperidinyl)iodoacetamide (IAM 6). When the protein was titrated with carbon monoxide (a homotropic effector like oxygen) the spin label sensed just two conformational states as evidenced by clear isosbestic points in the e.p.r. spectrum. The addition of inositol hexaphosphate (an allosteric effector) caused the Hill coefficient to rise from 1.1 to 1.6 and also eliminated the isosbestic points. This clearly suggests that there are now more than two conformational states as is the case with haemoglobin A. The most plausible explanation for these observations is that inositol hexaphosphate stabilizes the T state[254] of haemoglobin Kempsey in which the $\alpha_1\beta_2$ subunit interface is responsive to molecular strain. Confirmation for this conclusion comes from high-resolution n.m.r. studies: the n.m.r. spectra of haemoglobin Kempsey as a function of carbon monoxide concentration show that the β-haems have a higher ligand affinity and that the β-haem pocket in the haemoglobin Kempsey–carbon monoxide complex has been altered. Thus e.p.r. and n.m.r. are providing structural information on haemoglobin in solution.

Ribonuclease.—Earlier work[255] had shown that conformational changes in ribonuclease could be monitored by e.p.r. following reaction with *N*-(1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidinyl)bromoacetamide. In this study, no attempt was made to determine whether one or several amino-acid residues had been labelled. Hiskey and co-workers[256] have identified two main

[252] R. A. Dwek, 'N.M.R. in Biochemistry: Application to Enzyme Systems', Clarendon Press, Oxford, 1973.
[253] T. R. Lindstrom, J. J. Baldassare, H. F. Bunn, and C. Ho, *Biochemistry*, 1973, **12**, 4212.
[254] J. Monod, J. Wyman, and J.-P. Changeux, *J. Mol. Biol.*, 1965, **12**, 88.
[255] I. C. P. Smith, *Biochemistry*, 1968, **7**, 745.
[256] W. E. Daniel, jun., J. D. Morrisett, J. H. Harrison, H. H. Dearman, and R. G. Hiskey, *Biochemistry*, 1973, **12**, 4918.

components in the reaction mixture; ribonuclease spin-labelled at a methionine and ribonuclease spin-labelled on the imidazole-3-nitrogen of His-105. For reasons of stability, the enzyme with modified methionine was not further characterized. The enzyme spin-labelled at His-105 shows lowered enzymic activity even though the histidines located in the active site have been shown by n.m.r.[257] and *X*-ray diffraction analysis[258] to be His-12 and His-119. The His-105 spin-labelled enzyme behaves similarly to the native enzyme during affinity chromatography and towards iodoacetate inactivation, suggesting that any results obtained from e.p.r. studies will provide valid information on the native enzyme. In the present paper, the only result reported from e.p.r. studies was that the spin label at His-105 is not totally immobilized relative to the protein molecule.

Phosphofructokinase.—The studies by Jones *et al.*[259] on ATP binding to rabbit skeletal muscle phosphofructokinase have been extended.[260] Spin-labelling of a single reactive thiol group per 90 000 subunit in the enzyme with IAM 6 (see p. 229) has been shown to have only a minimal effect on the catalytic activity and allosteric behaviour of the enzyme. From titration studies using e.p.r. detection, it could be concluded that Mg–ATP was bound at the active site and produced positive homotropic interactions[254] between at least two active sites (Hill coefficient of 2.03). Similar titration experiments using Mn–ATP showed that this effector bound both at the active site and at a second site which is probably the inhibitory site. Binding at this second site was non-co-operative, which perhaps indicates that the spin label had modified at least some of the allosteric properties. The distance of the Mn–ATP (bound probably at the active site) from the spin label site has been estimated by a method using data derived from the observed quenching of the spin label e.p.r. signal and from measurements of water proton relaxation rates. Certain assumptions need to be made during this analysis and it is reassuring that the distance is consistent with the value estimated from high-resolution n.m.r. studies where broadening by the spin label of various proton resonances on the bound Mg–ATP is observed.

Citrate Synthase.—E.p.r. and pulsed n.m.r. have been used in a study of the reaction between oxaloacetate and a spin-labelled analogue of acetyl coenzyme A (RCoA).[261] The number of binding sites and the dissociation constant for the binding of RCoA to the enzyme can be determined by titration using e.p.r. detection. Bound RCoA can be displaced from the enzyme by acetyl CoA and CoASH; this displacement can be followed by e.p.r. and allows an estimate of the dissociation constants for these two

[257] G. C. K. Roberts, D. H. Meadows, and O. Jardetzky, *Biochemistry*, 1969, **8**, 2053.
[258] H. W. Wyckoff, K. D. Hardman, N. M. Allewell, T. Inagami, L. N. Johnson, and F. M. Richards, *J. Biol. Chem.*, 1967, **242**, 3984.
[259] R. Jones, R. A. Dwek, and I. O. Walker, *F.E.B.S. Letters*, 1972, **26**, 92.
[260] R. Jones, R. A. Dwek, and I. O. Walker, *European J. Biochem.*, 1973, **34**, 28.
[261] S. W. Weidman, G. R. Drysdale, and A. S. Mildvan, *Biochemistry*, 1973, **12**, 1874.

ligands. By contrast, addition of oxaloacetate did not displace bound RCoA though the e.p.r. spectrum indicated that the spin label had become immobilized as the ternary complex was formed. There was an enhancement of water proton relaxation rate during the addition of oxaloacetate to the enzyme–RCoA complex; this was used as a method for determining the number of oxaloacetate molecules bound (two per enzyme molecule).

Trypsin.—The active-site serine in trypsin can be selectively labelled with *N*-(1-oxyl-2,2,6,6-tetramethyl-4-piperidinyl) methyl phosphofluoridinate.[262] The e.p.r. signal, which indicated intermediate immobilization, did not change when the trypsin was bound to a porous glass matrix.[263] This supports the conclusion that the active-site conformation in trypsin is similar for the bound and unbound forms of the enzyme. More subtle probing of changes in the active-site conformation is planned using a series of spin labels of differing geometry. The stoicheiometry of trypsin binding to a matrix was also determined by a spin count method which should have applicability to other matrix bound enzymes.

ATPases.—Dicyclohexylcarbodi-imide (DCCD) is a potent inhibitor of mitochondrial ATPase to which it binds irreversibly. Studies on ATPase action have been made[264] using a spin-labelled analogue of DCCD [*N*-(1-oxyl-2,2,6,6-tetramethyl-4-piperidinyl)-*N*-cyclohexylcarbodi-imide] which is abbreviated NCCD. NCCD becomes highly immobilized ($\tau_c \sim 10^{-8}$ s) on binding to the ATPase present in sonicated mitochondria. Addition of Mn–ATP to the preparation broadens the spin-label e.p.r. spectrum and allows a calculation to be made of the manganese–spin-label distance. The value obtained (20 Å) indicates that the sites for ATP and NCCD binding are relatively close. It is interesting to note that the addition of succinate to NCCD-labelled mitochondria leads to a reduction of the e.p.r. signal. This suggests that electrons can be transferred from the respiratory chain to the ATPase. Further work is in progress to determine the location of the link point between the ATPase and the respiratory chain.

It has previously been shown[265] that thiol groups in muscle myosin can be labelled with IAM 6 (see p. 229). The strongly immobilized e.p.r. signal is further immobilized following reaction with actin and slightly mobilized by the addition of nucleotides. Other biochemical studies[266] have indicated that interaction occurs between the actin and nucleotide binding sites on myosin. Stone[267] has attempted to quantify these effects through application of the spin-labelling technique. Serious efforts have been made to

[262] L. J. Berliner and S. S. Wong, *J. Biol. Chem.*, 1973, **248**, 1118.
[263] L. J. Berliner, S. T. Miller, R. Uy, and G. P. Boyer, *Biochim. Biophys. Acta*, 1973, **315**, 195.
[264] A. Azzi, M. A. Bragadin, A. M. Tanburro, and M. Santano, *J. Biol. Chem.*, 1973, **248**, 5520.
[265] J. Quinlivan, H. M. McConnell, L. Stowring, R. Cooke, and M. F. Moralis, *Biochemistry*, 1969, **8**, 3645.
[266] B. Kiely and A. Martinosi, *Biochim. Biophys. Acta*, 1969, **172**, 158.
[267] D. B. Stone, *Biochemistry*, 1973, **12**, 3672.

restrict the labelling just to the two rapidly reacting thiol groups in heavy meromyosin (heavy meromyosin is the 'head' end of the myosin molecule and possesses both the ATPase- and actin-binding sites). The effect of addition of actin to the spin-labelled heavy meromyosin–ADP complex is dependent on both the concentration of actin and of ADP. At low ADP concentrations, the immobilization resulting from actin addition can be explained by displacement of ADP. At high ADP concentrations, a ternary actin–heavy meromyosin—ADP complex is probably formed. The ATP binding sites for spin-labelled heavy meromyosin can be completely blocked using the affinity label 6-mercapto-9-β-D-ribofuranosylpurine-5′-triphosphate; a large increase in spin-label mobility results. Addition of actin caused little change in the spectrum, suggesting that no interaction between sites had occurred. This was confirmed by parallel viscosity measurements. The author believes that the affinity label has caused some non-physiological modification of the meromyosin. The use of other synthetic nucleoside triphosphates which bind to meromyosin but which are only slowly hydrolysed shows a similar nucleotide concentration dependence to that observed with ADP. It can be concluded that, although the sites for nucleotide and actin binding are separate, attachment of each ligand is dependent on the presence of the other, supporting the biochemical evidence [266] on interaction between the sites.

Spin Labels as Enzyme Substrates.—The studies on mitochondrial ATPase reported above use the spin label as an electron accepter. Stier and Sackmann [268] have employed a similar approach to investigate another membrane-bound system, namely the cytochrome P450–cytochrome P450 reductase system. The temperature dependence of the reduction of (1-oxyl-2,2,6,6-tetramethyl-4-piperidinyl)phosphate and *N*-(oxyl-4,4-dimethyloxazolidine) derivatives of stearic acid supports the conclusion that the phospholipids are distributed heterogeneously in the membrane, the protein complex being surrounded by a quasi-crystalline lipid 'halo' at temperatures below 32 °C.

General Studies on Proteins.—It has long been recognized [269] that the polarity of the environment in which a spin label resides is reflected in the e.p.r. spectrum. This has found application in studies on membranes.[270] Lassmann *et al.*[271] have carried out experiments in which the hydrophobic regions of proteins are probed by spin labels. The thiol groups of leucine amino-peptidase were *non-selectively* spin labelled with IAM 6 (see p. 229) which limits the value which can be placed on the findings. However, the method has potential in determining the depths of hydrophobic pockets

[268] A. Stier and E. Sackmann, *Biochim. Biophys. Acta*, 1973, **311**, 400.
[269] R. Briere, H. Lemaire, and A. Rassat, *Bull. Soc. chim. France*, 1965, 3273.
[270] J. Hsia, H. Schneider, and I. C. P. Smith, *Chem. Phys. Lipids*, 1970, **4**, 120.
[271] G. Lassmann, B. Ebert, A. N. Kuznetsov, and W. Damerau, *Biochim. Biophys. Acta*, 1973, **310**, 298.

in enzymes; a similar approach has been successfully applied to antibody-binding sites by Hsia and Piette.[272]

Price[273] has described the use of the radioactive spin label *N*-(1-oxyl-2,2,6,6-tetramethyl-4-piperidinyl)iodo[2-^{14}C]acetamide to facilitate the determination of the number and location within the polypeptide chain of reactive thiol groups. The reaction conditions needed to give selective spin labelling could thereby be ascertained; some results obtained with phosphorylase, bovine serum albumen, and glyceraldehyde-3-phosphate dehydrogenase were presented.

Membranes.—The number of spin-labelling papers on membranes published during 1973 exceeded that on proteins. Whilst emphasis will be given in this report to spin labelling of membrane proteins, coverage of the more important papers dealing with lipid spin labelling will be attempted.

A good introductory discussion of magnetic resonance methods (both n.m.r. and e.p.r.) in membrane research is available.[274] In addition, Schreier-Muccillo and Smith[275] have written an essentially qualitative but exhaustive review on the use of spin labels as probes of organization in biological and model membranes, and a critical summary of this field has been given by Keith *et al.*[276]

Nature and Role of Proteins in Membranes. The most interesting studies in this area are those in which the perturbations of a membrane labelled with a lipid spin label are monitored following addition of purified proteins. Hong and Hubbell[277] reported on their studies with the lecithin–rhodopsin system: investigation of the interactions between rhodopsin and membranes has contributed towards an understanding of the elementary steps in the visual process. The ordering effect of rhodopsin on egg lecithin vesicles, which resembles the effect of cholesterol, clearly indicated that an appreciable fraction of the protein molecule is located within the hydrophobic interior of the membrane; this result has been substantiated by freeze-etch electron microscopic studies. The work has been extended[278] to a study of the conditions leading to regeneration of rhodopsin–phospholipid systems with high photochemical sensitivity. Traüble and Sackmann[279] have used the value for the lateral diffusion coefficient in membranes, calculated from their elegant spin-labelling studies, to estimate how rapidly rhodopsin rotates about an axis perpendicular to the membrane surface. Their value for the relaxation time of between 18 and 50 μs agrees reasonably with that calculated independently by Cone.[280]

[272] J. Hsia and L. Piette, *Arch. Biochem. Biophys.*, 1969, **129**, 296.
[273] N. C. Price, *F.E.B.S. Letters*, 1973, **36**, 351.
[274] J. Seelig, *Experientia*, 1973, **29**, 509.
[275] S. Schreier-Muccillo and I. C. P. Smith, *Prog. Surface Membrane Sci.*, 1974, **9**, in the press.
[276] A. D. Keith, M. Sharnoff, and G. E. Cohn, *Biochim. Biophys. Acta*, 1973, **300**, 379.
[277] K. Hong and W. L. Hubbell, *Proc. Nat. Acad. Sci. U.S.A.*, 1972, **69**, 2617.
[278] K. Hong and W. L. Hubbell, *Biochemistry*, 1973, **12**, 4517.
[279] H. Traüble and E. Sackmann, *Nature*, 1973, **245**, 210.
[280] R. A. Cone, *Nature New Biol.*, 1972, **236**, 39.

Griffith and co-workers [281-283] have carried out careful studies using stearic acid spin labels on the vesicular structure formed when delipidated cytochrome oxidase, isolated from beef heart mitochondria, is titrated with phospholipid. The e.p.r. spectra revealed mobile and immobilized components and were consistent with the protein assembly being surrounded by a monolayer of phospholipid. Computational methods for spectral analysis are used to good effect by the authors.

Kirkpatrick and Sandberg [284, 285] have spin labelled the thiol groups of proteins in erythrocyte ghosts with *N*-(1-oxyl-2,2,6,6-tetramethyl-4-piperidinyl)maleimide. Qualitative information on the conformational changes produced by detergents and by neutral salts has been obtained. The effect of various proteins of known amino-acid composition on the organization of lipid spin-labelled multibilayers has been studied.[286] None of the proteins employed had an *in vivo* location in a membrane; the main conclusion drawn was that net positive charge on the protein leads to an ordering effect on the membrane, an effect which can be observed also with relatively high salt concentrations.

Lateral Diffusion and Lateral Phase Separation. The method used to determine lateral diffusion coefficients of phospholipid molecules in dipalmitoyl lecithin vesicles [287] has been applied to cytoplasmic membranes isolated from fatty acid auxotrophs of *E. coli*.[288] The value determined (3.25×10^{-8} cm^2 s^{-1}) was in close agreement with that found for the model system. Estimates have been made of the lateral diffusion coefficients for membrane proteins (*ca.* 3×10^{-10} cm^2 s^{-1}). In the same paper [288] it was shown that the values for the phase-transition temperatures in both dipalmitoyl lecithin vesicles and membranes from *E. coli* determined using spin-labelled stearic acid probes were dependent upon the distance of the paramagnetic centre from the membrane surface. The stearic acid probe bearing the nitroxide near the methyl terminus gave phase-transition temperatures in agreement with those determined by other physical techniques. The biological significance of possible 'pretransitions' as reported by the other spin-labelled stearic acid probes is at present unclear.

Independent studies from the laboratories of McConnell and Fox [289, 290] support the concept of 'pretransition' temperatures in *E. coli* membranes.

[281] P. C. Jost, O. H. Griffith, R. A. Capaldi, and G. Vanderkooi, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 480.
[282] P. C. Jost, O. H. Griffith, R. A. Capaldi, and G. Vanderkooi, *Biochim. Biophys. Acta*, 1973, **311**, 141.
[283] O. H. Griffith, *J. Supramolecular Structure*, 1973, **1**, 269.
[284] F. H. Kirkpatrick and H. E. Sandberg, *Biochim. Biophys. Acta*, 1973, **298**, 209.
[285] F. H. Kirkpatrick and H. E. Sandberg, *Arch. Biochem. Biophys.*, 1973, **156**, 653.
[286] K. W. Butler, A. W. Hansen, I. C. P. Smith, and H. Schneider, *Canad. J. Biochem.*, 1973, **51**, 980.
[287] H. Traüble and E. Sackmann, *J. Amer. Chem. Soc.*, 1972, **94**, 4499.
[288] E. Sackmann, H. Traüble, H.-J. Galla, and P. Overath, *Biochemistry*, 1973, **12**, 5360.
[289] C. D. Linden, K. L. Wright, H. M. McConnell, and C. F. Fox, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2271.
[290] H. M. McConnell, *J. Supramolecular Structure*, 1973, **1**, 285.

'Pretransition' temperatures were discussed in terms of lateral phase separations of lipids in the membranes which might facilitate membrane transport processes. Shimshick and McConnell [291] have investigated lateral phase separations in model membrane vesicles of known phospholipid composition. A possible way by which lateral phase separations might participate in transport processes is proposed. In this context, it is interesting to note that the half-time for phospholipid transverse motion ('flip-flop') across the bilayer is *ca.* 4 min in excitable membrane vesicles prepared from the electroplax cells of the electric eel, and possibly < 1 min in *Mycoplasma* membranes.[292] These half-times are short enough to suggest that phospholipid 'flip-flop' might be a mechanism for the transport of small molecules or ions across membranes.

Spin–Spin Interaction Effects in Phospholipid Membranes. Mention has already been made [287] of lateral diffusion coefficient measurements from the spin–spin broadened e.p.r. spectra of spin-labelled membranes. Levine *et al.*[294] have observed that addition of various spin-labelled lipids to dipalmitoyl lecithin vesicles causes a reduction in the T_1 values of protons in the lecithin-NMe_3 groupings and fatty acid chains.

Marsh and Smith [295] have interpreted the e.p.r. spectra from high concentrations of cholestane spin label incorporated into multi-bilayers composed of lecithin plus cholesterol to give distance information. A model describing how cholesterol and lecithin might pack together was proposed. Another paper from Smith's laboratory [296] used the spin label technique to distinguish between the effects of cholesterol on dipalmitoyl lecithin and egg lecithin bilayer systems.

Chemical Studies on Spin Labels.—Whilst scarcely appropriate to a literature survey for 1973, the attention of workers in this research area is drawn to an excellent monograph giving detailed syntheses of a wide variety of spin labels.[297]

The synthesis and properties of nitronyl nitroxides, some of which have unusually large nitrogen hyperfine coupling constants, have been reported.[298] The same property has been observed [299] for certain bicyclic nitroxyl radicals. Such spin labels would be valuable for observing and quantifying conformational changes in biological macromolecules.

[291] E. J. Shimshick and H. M. McConnell, *Biochemistry*, 1973, **12**, 2351.
[292] M. G. McNamee and H. M. McConnell, *Biochemistry*, 1973, **12**, 2951.
[293] C. W. M. Grant and H. M. McConnell, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1238.
[294] Y. K. Levine, N. J. M. Birdsall, A. G. Lee, and J. C. Metcalfe, *Biochim. Biophys. Acta*, 1973, **291**, 592.
[295] D. Marsh and I. C. P. Smith, *Biochim. Biophys. Acta*, 1973, **298**, 133.
[296] S. Schreier-Muccillo, D. Marsh, H. Dugas, H. Schneider, and I. C. P. Smith, *Chem. Phys. Lipids*, 1973, **10**, 11.
[297] E. G. Rosantsev, 'Free Nitroxyl Radicals', Plenum Press, New York and London, 1970.
[298] E. F. Ullman, L. Call, and S. S. Tseng, *J. Amer. Chem. Soc.*, 1973, **95**, 1677.
[299] I. Morishima, K. Yoshikawa, K. Bekki, M. Kohno, and K. Arita, *J. Amer. Chem. Soc.*, 1973, **95**, 5815.

Detailed studies on di-t-butyl nitroxide have been reported by two groups.[300, 301]

General References.—There have been several studies made on the effect of anaesthetics on membranes using lipid spin-label probes.[302–304] The converse approach of using spin-labelled anaesthetics to study their action on membranes is currently under investigation.[305] The authors draw attention to the need for careful controls to ensure that the pharmacological activity of the anaesthetic has not been altered by attachment of the spin label. More studies using spin-labelled biological effector molecules, *e.g.* drugs, anaesthetics, and hormones, may confidently be expected.

Zavrie *et al.*[306] have used 1-[*N*-oxyl-2,2,6,6-tetramethyl]-4-[β-*N*-ethyl-iminopropionyl]-oxypiperidine to label both DNA and RNA. Temperature variation studies on the spin-labelled RNA derivative have shown that a conformational change probably occurs at a temperature *ca.* 20 °C below the melting temperature.

6 Fluorescence

Contributed by G. R. Penzer

A wide range of observations potentially interesting to those wishing to learn about protein behaviour by use of fluorescence spectroscopy was reported in 1973. The spectrum extends from topics like quenching mechanisms and expressions for the concentration dependence of fluorescence polarization (found in journals of physics) to ideas for new fluorescent probes suited to a particular application (in, perhaps, a histology publication). In preparing this year's Report the obvious industry of the growing number of those whose work is relevant has been striking (and daunting). The aim has again been to outline the scope of what has been covered without filling in all the details. This means that much has been knowingly omitted (unknowingly, probably, there has been a good deal more because 'fluorescence' is not a key-word in the abstracts of many relevant papers). Like last year,[307] two extensive areas, immunofluorescence and the study of proteins in membranes, have been largely ignored.

Theory, Methods, and Techniques.—There is more to making satisfactory fluorescence measurements than using a commercial instrument to record

[300] G. B. Birrell, S. P. Van, and O. H. Griffith, *J. Amer. Chem. Soc.*, 1973, **95**, 2451.
[301] M. S. Davis and R. W. Kreilick, *J. Amer. Chem. Soc.*, 1973, **95**, 5514.
[302] J. R. Trudell, W. L. Hubbell, and E. N. Cohen, *Biochem. Biophys. Acta*, 1973, **291**, 321.
[303] F. Letterrier, F. Rieger, and J. F. Mariand, *J. Pharmacol. and Experimental Therapeutics*, 1973, **186**, 609.
[304] K. W. Butler, H. Schneider, and I. C. P. Smith, *Arch. Biochem. Biophys.*, 1973, 154, 548.
[305] R. J. Garguilo, G. J. Giotta, and H. W. Wang, *J. Medicin. Chem.*, 1973, **16**, 707.
[306] S. Zavrie, G. Grigoryan, I. Krilova, and R. Karmanov, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 123.
[307] G. R. Penzer, ref. 2, p. 203.

the excitation and emission spectra of a sample. It is sometimes hard to know whether an attempt to repeat published work has succeeded because the original report failed to quote corrected spectra or accurate quantum yields. Articles on absolute fluorometry,[308] and suggestions for fluorescence standards both for quantum yields[309] and lifetimes[310] have appeared. A thick volume on practical fluorescence measurements[311] and an article on newer methods for fluorescence analysis of biologically important compounds[312] have been published, and more is promised.[313]

Fluorescence Lifetimes. All kinds of new, improved, or slightly modified instruments and attachments have been reported. Probably most attention has been focused on the rapid measurement of the decay of fluorescence following pulsed excitation, and devices ranging from the highly sophisticated to the comparatively simple have been described.[314] There have also been discussions of the best way to analyse the data obtained in such studies,[315–318] and the likely errors.[319, 320] There are two general approaches to handling the data. One may either assume that the observed decay comprises contributions from several species all of which emit with an exponential decay,[315, 316] or simply try to find a function which accurately describes the measured decay curve and then attempt to understand the function in terms of contributions from the various emitters.[317, 318] The second approach is essential when individual emission decays are not exponential (*e.g.* when there are environment changes caused by solvent relaxation during the decay period). Two of the papers in particular[315, 317] include discussion and evaluation of the methods. Treatments of non-random instrumental errors (by moment index displacement)[319] and of statistical errors[320] have also been given. Clearly it is now possible to resolve with precision the lifetimes of several species contributing to a single decay, with resulting improvement in the sophistication of the

[308] W. H. Melhuish, *J. Res. Nat. Bur. Stand.* (*A*), 1972, **76**, 547; G. A. Crosby, J. N. Demas, and J. B. Callis, *ibid.*, p. 561; R. F. Chen, *ibid.*, p. 593.
[309] R. Reisfeld, *J. Res. Nat. Bur. Stand.* (*A*), 1972, **76**, 613; R. A. Velapoldi, *ibid.*, p. 641.
[310] R. F. Chen, *Analyt. Biochem.*, 1974, **160**, 106.
[311] G. G. Guilbault, 'Practical Fluorescence Theory, Methods and Techniques', Marcel Dekker Inc., New York, 1973,
[312] G. G. Guilbault, *J. Res. Nat. Bur. Stand.* (*A*), 1972, **76**, 607.
[313] R. F. Chen and H. Edelhoch, 'Concepts in Biochemical Fluorescence', Marcel Dekker Inc., New York, 1974.
[314] F. Rodier and M. J. Icole, *Photochem. Photobiol.*, 1973, **17**, 473; K. Osada, *Rev. Sci. Instr.*, 1973, **44**, 656; V. U. Arsen'ev, V. A. Gavanin, V. Z. Pashchenko, S. P. Protasov, L. B. Rubin, and A. B. Rubin, *Zhur. priklad. Spectroskopii*, 1973, **18**, 1093; M. Siebert, R. R. Alfano, and S. L. Shapiro, *Biochim. Biophys. Acta*, 1973, **292**, 493; C. Lewis, W. R. Ware, L. J. Doemeny, and T. L. Nemzck, *Rev. Sci. Instr.*, 1973, **44**, 107; H. C. Beall and A. Haug, *Analyt. Biochem.*, 1973, **53**, 98.
[315] I. Isenberg, R. D. Dyson, and R. Hanson, *Biophys. J.*, 1973, **13**, 1090.
[316] B. Valeur and J. Moirez, *J. Chim. phys. physicochim. biol.*, 1973, **70**, 500.
[317] W. R. Ware, L. J. Doemeny, and T. L. Nemzck, *J. Phys. Chem.*, 1973, **77**, 2038.
[318] M. Almgren, *Chem. Scripta*, 1973, **3**, 145.
[319] I. Isenberg, *J. Chem. Phys.*, 1973, **59**, 5696.
[320] I. Isenberg, *J. Chem. Phys.*, 1973, **59**, 5708.

possible molecular interpretations. But it is important to avoid over-simplified assumptions, *e.g.* that the decay from any complex system can be properly described by a single exponential. Quite good fits to the data can be obtained with this assumption, even when decay is truly composed of contributions from two species with different lifetimes.[321]

The wavelength dependence of emission decay has also been studied,[322, 323] though the approaches to interpretation of the results have varied. In one case [322] emission was from tryptophan residues in human serum albumin. The variation in decay time with wavelength was used to distinguish different residues, assuming that each caused exponential decay. In the second example [323] a red shift during decay in the emission maximum of 2-*p*-toluidinyl-6-naphthalenesulphonate (TNS) adsorbed by lipid bilayers was observed. It was attributed either to an excited state interaction between TNS and lipid or to solvent relaxation during the decay.

Many aspects of relaxation fluorometry (both theoretical and practical) have been discussed in two useful general articles.[324]

Steady-state Fluorescence Measurements. Other instrumental improvements have involved read-out and data-handling systems,[325, 326] low-temperature measurements,[327] flow-cell design [328] and application in an automated polarization fluorometer,[329] photon-counting detection,[330] and a device for attachment to commercial instruments giving automatic compensation for Rayleigh scatter.[331] The advantages of measuring the absorbance as well as the fluorescence of a sample if there are computational facilities to make full use of all the data have been demonstrated.[325] The design of cells to give a sensitivity higher by an order of magnitude or more by utilizing multiple internal reflection has been described,[332] and the theory of emission between mirrors (relevant also to fibre optics) has been considered.[333] The opportunity afforded by fluorescence spectroscopy for the study of molecules in their natural (cellular) environments by increasingly sophisticated microfluorometry has been developed further [334] but as yet no detailed study of a particular protein species under intracellular conditions has been reported. At a rather simpler level, however, it has been shown

[321] F. Rodier, *Biochimie*, 1973, **55**, 647.
[322] P. Wahl and J. C. Auchet, *Biochim. Biophys. Acta*, 1972, **285**, 99.
[323] J. H. Easter and L. Brand, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 1086.
[324] J. Yguerabide, *Methods Enzymol.*, 1972, **26C**, 498; R. Rigler and M. Ehrenberg, *Quart. Rev. Biophys.*, 1973, **6**, 139.
[325] J. F. Holland, R. E. Teets, and A. Timnick, *Analyt. Chem.*, 1973, **45**, 145.
[326] T. M. Sheperd, *Chem. and Ind.*, 1973, 332.
[327] K. Kawai, *Analyt. Biochem.*, 1974, **57**, 190.
[328] D. C. Street, R. C. Denney, J. Mendham, and B. Bush, *Chem. and Ind.*, 1973, 191.
[329] R. D. Spencer, F. B. Toledo, B. T. Williams, and N. L. Yoss, *Clinical Chem.*, 1973, **19**, 838.
[330] R. E. Curry, H. L. Pardue, G. E. Mieling, and R. E. Santini, *Clinical Chem.*, 1973, **19**, 1259.
[331] B. Armitage and D. E. Ryan, *Canad. J. Spectroscopy*, 1973, **18**, 52.
[332] N. J. Marrick and G. I. Loeb, *Analyt. Chem.*, 1973, **45**, 687.
[333] M. R. Philpott, *Chem. Phys. Letters*, 1973, **19**, 435.
[334] E. Kohen, C. Kohen, B. Thorell, and J. M. Salmon, *Rev. Sci. Instr.*, 1973, **44**, 1784.

that fluorescein-labelled histones and RNA polymerase only bind to restricted regions of polytene chromosomes of *Drosophila*, leading to the suggestion that large regions of the DNA are readily accessible to extra-chromosomal proteins.[335]

Fluorescence intensity is not directly proportional to emitter concentration because of various optical screening and reabsorption effects. Detailed consideration has been given to such factors for several instrument geometries, and ways to optimize conditions so as to achieve minimum deviations from linearity for a given detection sensitivity have also been described.[336] Fluorometers usually have large bandwidths (1—20 nm) for both excitation and emission. A method for calculating quasimonochromatic spectra from measurements using significant bandwidths has been given. In some cases differences between experimental and quasimonochromatic spectra are greater than normal experimental error.[337]

Quenching of Fluorescence. In 1973 there was renewed activity in the application of fluorescence quenching to the study of proteins. The traditional first choice quencher, KI, has been used in some cases,[338–340] but there have also been studies involving $NaNO_3$ and/or CsCl,[341–344] $AgNO_3$,[345] and O_2.[346, 347] In most of these studies quenching efficiency is taken to measure the degree of exposure to solvent of an emitting group (usually, when intrinsic protein fluorescence is studied, Trp). Semi-quantitative discussions of the effect of quenchers on Trp emission, and attempts to categorize Trp residues according to their degrees of exposure, have been given.[342, 343]

The behaviour of $AgNO_3$ has been carefully documented.[345] Both Ag^+ and NO_3^- are efficient collisional quenchers, but additionally Ag^+ forms a complex with Trp and mercaptide bonds by reaction with thiol groups. Mercaptide formation causes absorption at wavelengths out to 340 nm, and it has been calculated that quenching of Trp emission by energy transfer is probable. This offers a way to measure Trp–SH separations. The activity of O_2 as a quencher[346] contrasts in some respects with the properties typical of ions. Only emitters with lifetimes >100 ns are much quenched by O_2 at atmospheric pressure but, at pressures up to 100 atm,

[335] R. B. Khesin, *Mol. Biol.*, 1973, **7**, 776.
[336] R. Van-Slageren, G. Den-Boff, and W. E. Van der Linden, *Talanta*, 1973, **20**, 501.
[337] L. Vize and L. Szalay, *Acta Phys.*, 1973, **33**, 33.
[338] G. Jori and N. Genov, *Internat. J. Peptide Protein Res.*, 1973, **5**, 171.
[339] H. Onishi, E. Ohtsuka, M. Ikehara, and Y. Tonomura, *J. Biochem. (Japan)*, 1973, **74**, 435.
[340] K. O. Honikel and N. B. Madsen, *Canad. J. Biochem.*, 1973, **51**, 344.
[341] R. Kaplanas, T. G. Bukolova, and E. A. Burshtein, *Mol. Biol.*, 1973, **7**, 753.
[342] I. U. Ostashevskii, V. A. Velling, and A. G. Bezrukova, *Mol. Biol.*, 1973, **7**, 307.
[343] E. A. Burshtein, N. S. Vedenkina, and M. N. Ivkova, *Photochem. and Photobiol.*, 1973, **18**, 263.
[344] D. R. Sellers and C. A. Ghiron, *Photochem. and Photobiol.*, 1973, **18**, 393.
[345] R. F. Chen, *Arch. Biochem. Biophys.*, 1973, **158**, 605.
[346] J. R. Lacowicz and G. Weber, *Biochemistry*, 1973, **12**, 4161.
[347] J. R. Lacowicz and G. Weber, *Biochemistry*, 1973, **12**, 4171.

$[O_2]$ can be increased to *ca.* 100 mM in aqueous solutions and emissions with lifetimes < 10 ns are quenched. Quenching efficiency is almost diffusion controlled for several chromophores. It is insensitive to electrostatic effects (*cf.* I^- and other ions) but sensitive to the fluorophore being bound to a macromolecule and hence, presumably, protected from quencher.[347]

The sensitivity of fluorescence to the molecular environment of the emitter has been widely exploited in fluorescence titrations of many kinds. A review letter illustrating aspects of this work in the study of lactate dehydrogenase has been published.[348] There have been several discussions of aspects of environmental sensitivity – in particular the physical bases of fluorescence quenching,[349] concentration depolarization of fluorescence,[350] and energy transfer between different donor and acceptor species.[351, 352] There is a warning to those who use energy transfer efficiency as a way to calculate the separation of chromophores in the observation that the efficiency of Förster dipole–dipole transfer is greater than that predicted by the usual theory when solvent viscosities are low.[352]

Energy transfer can be exploited in ways which are (so far as the energy transfer is concerned) either qualitative or quantitative. An example of the former is measurement of hydrolase activity by observing the change in transfer efficiency between donor and acceptor groups on a single substrate during hydrolysis.[353] Quantitative studies usually involve distance calculations based on the Förster formula for transfer efficiency. The sixth-power dependence on distance of this effect limits its use to a comparatively small range of values for a particular donor–acceptor pair unless either very high or very low transfer efficiencies can be determined accurately. This is often attempted by fluorescence yield measurements, but study of fluorescence lifetime changes is also promising. For example, an antibody covalently labelled with two fluorophores (dansyl and fluorescein) showed a fall in emission lifetime of *ca.* 1 ns (from 23.6 ns) when energy transfer was allowed. This corresponds to a transfer efficiency of 3.7% and led to calculation of a donor–acceptor separation of 8.1 nm.[354]

[348] J. J. Holbrook and H. Gutfreund, *F.E.B.S. Letters*, 1973, **31**, 157.

[349] V. I. Tomin, A. N. Rubinov, and V. F. Voronin, *Optics and Spectroscopy*, 1973, **34**, 643; H. Lami, *Chem. Phys. Letters*, 1973, **21**, 140; T. C. Werner and R. M. Hoffman, *J. Phys. Chem.*, 1973, **77**, 1611; I. B. Berlman, *J. Phys. Chem.*, 1973, **77**, 562; H. Dreeskamp and M. Zander, *Z. Naturforsch.*, 1973, **28a**, 1743.

[350] A. N. Sebchenko, V. I. Popechits, and A. M. Sarzhevskii, *Doklady Akad. Nauk Beloruss. S.S.R.*, 1973, **17**, 405; A. Kawski and J. Kaminski, *Izvest. Akad. Nauk S.S.S.R., Ser. fiz.*, 1973, **37**, 761; V. I. Popechii, I. N. Kozlov, A. M. Sarzhevskii, and A. H. Sevchenko, *Zhur. priklad. Spectroskopii*, 1973, **18**, 653; I. N. Kozlov and A. M. Sarzhevskii, *ibid.*, 1973, **19**, 1048; C. Bojarski, *Acta Phys. Polon.* (*A*), 1973, **43**, 391; R. P. Hemenger and R. M. Pearlstein, *J. Chem. Phys.*, 1973, **59**, 4064.

[351] C. Bojarski, F. Burak, E. Grabowska, and L. Kaczynska, *Z. Naturforsch.*, 1973, **28a**, 1697.

[352] A. Kawski, E. Kuten, and J. Kaminski, *J. Phys.* (*B*), 1973. **6**, 1907.

[353] A. Carmel, M. Zur, A. Yaron, and E. Katchalski, *F.E.B.S. Letters*, 1973, **30**, 11.

[354] J. R. Bunting and R. E. Cathou, *J. Mol. Biol.*, 1973, **77**, 223.

The interpretation of fluorescence titration curves has received some attention.[355, 356] In particular the error in assuming that when several ligands bind to the same protein they all have similar effects on fluorescence emission has been illustrated. The use of measurement of the pH-dependence of fluorescence to obtain excited-state pK_a values has also been critically discussed and the method has been shown to break down in some instances.[357]

General. A trend, noted last year,[307] was again apparent in 1973, as more workers applied several physical techniques in studying a particular problem. Usually, in addition to fluorescence, one or more of n.m.r., e.p.r., c.d., and u.v. absorption were used, but less common techniques have also been applied. An example is a study of molecules resembling the Schiff base intermediates of pyridoxal phosphate dependent enzymes.[358] Two conformationally distinct transition-state analogues have been synthesized and studied by fluorescence, n.m.r., c.d., and their behaviour as haptens.

Fluorescence Probes.—Probe, for the purposes of this section, refers to any non-protein substrate, prosthetic group, analogue, or totally unrelated species whose fluorescence can be measured and interpreted to yield information about a protein.

Model Studies. Two sorts of probe study filled many printed pages in 1973. The first involves model studies, the second application of probes to the solution of real biochemical problems. The model work aims to discover the molecular and spectroscopic properties of probes in defined systems so that better interpretation of results for situations of more immediate biological relevance can be achieved. Table 1 illustrates the scope of the reports, which fall into three groups. Some (A) simply involve studying the physical and chemical properties of pure probes; some (B) characterize

Table 1 *Model studies with fluorescent probes*

Probe	*Study (category)*	*Ref.*
2-Aminopurinetriphosphate	Enzyme kinetic parameters (B)	389
1-Anilinonaphthalene-8-sulphonate (ANS)	ANS/bovine serum albumin (C)	*a*
ANS and *N*-phenyl-1-naphthylamine	Electrostatic interactions with membranes (C)	*b*
ANS, dansyl, anthroyl and some others	Orientation and rotation in bilayer membranes (C)	*c*
Dansyl amino-acids	Enhanced emission in water (A)	*d*
$1,N^6$-EthenoAMP	Conformation and emission properties (A)	360

[355] J. A. Bishop, *Analyt. Chim. Acta*, 1973, **63**, 305.
[356] B. G. Archer, T. C. McGuire, and H. Krakauer, *Biochemistry*, 1973, **12**, 2151.
[357] N. Lasser and J. Feitelson, *J. Phys. Chem.*, 1973, **77**, 1011.
[358] V. Raso and B. D. Stollar, *J. Amer. Chem. Soc.*, 1973, **95**, 1621.

Table 1 (*cont.*)

Probe	*Study (category)*	*Ref.*
1,N^6-Ethenoadenine derivatives of various coenzymes	Conformation and emission properties (A)	361
1,N^6-Ethenoadenine polyribonucleotides	Physical properties (A)	366
1,N^6-Ethenoadenine DNA	Physical properties (A)	365
1,N^6-EthenoATP	Enzyme studies with myosin and heavy meromyosin (B)	363
1,N^6-EthenoATP and 1,N^6-ethenoCTP	Enzyme activities (B)	374
1,N^6-EthenoATP and 1,N^6-etheno tRNA	Activity with Phe:tRNA and Ser:tRNA ligases (B)	362
1,N^6-EthenoFAD	Conformation and emission properties (A)	359
NADH	Temperature dependence of spectroscopic properties (A)	*e*
Pyrene	Emission lifetimes in different solvents (A)	*f*
2-*p*-Toluidinylnaphthalene-6-sulphonate	Time-resolved emission when bound to lecithin vesicles (C)	323
Retinol	Polarized emission spectra (A)	376, *g*

(*a*) E. Daniel and J. T. Yang, *Biochemistry*, 1973, **12**, 508; D. Layton and W. Symmons, *F.E.B.S. Letters*, 1973, **30**, 325. (*b*) M. T. Flanagan and T. R. Hesketh, *Biochim. Biophys. Acta*, 1973, **298**, 535. (*c*) R. A. Badley, W. G. Martin, and H. Schneider, *Biochemistry*. 1973, **12**, 268; R. A. Badley, H. Schneider, and W. G. Martin, *Biochem. Biophys. Res. Comm.*, 1972, **49**, 1292. (*d*) T. Kinoshita, F. Iinuma, and A. Tsuji, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 666. (*e*) A. D. B. Malcolm, *Analyt. Biochem.*, 1973, **55**, 278. (*f*) A. Nakajima, *Bull. Chem. Soc. Japan*, 1973, **46**, 2602. (*g*) T. A. Moore and P. S. Song, *Chem. Phys. Letters*, 1973, **19**, 128.

the biochemical properties of the probes; and some (C) attempt to provide understanding of the causes of the environmental sensitivity of emission spectra.

Undoubtedly the most reported class of probes in 1973 was the 1,N^6-ethenoadenosine derivatives, and amongst the publications are various model studies.[359-367] Often [362-364] but not invariably [368, 369] the derivatives

[359] J. R. Barrio, G. L. Tolman, N. J. Leonard, R. D. Spencer, and G. Weber, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 941.
[360] G. R. Penzer, *European J. Biochem.*, 1973, **34**, 297.
[361] C.-Y. Lee and J. Everse, *Arch. Biochem. Biophys.*, 1973, **157**, 83.
[362] H. S. Hertz and H. G. Zachau, *European J. Biochem.*, 1973, **37**, 203.
[363] W. D. McCubbin, G. E. Willick, and C. M. Kay, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 926.
[364] G. E. Willick, K. Oikawa, W. D. McCubbin, and C. M. Kay, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 923.
[365] C. M. Lee and J. G. Wetmur, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 879.
[366] R. F. Steiner, W. Kinnier, A. Lunasin, and J. Delac, *Biochim. Biophys. Acta*, 1973, **294**, 24.
[367] G. R. Penzer and K. D. Robertson, *Biochim. Biophys. Acta*, 1974, **336**, 1.
[368] Y. H. Chien and G. Weber, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 538.
[369] M. DeLuca, N. J. Leonard, B. J. Gates, and W. D. McElroy, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1664.

have biochemical properties similar to those of the parent compounds, and n.m.r. studies suggest that in aqueous solutions molecular conformations of adenine nucleotides are little affected by derivative formation.[360, 361] However, the quantum yields and energies of the emissions are not very sensitive to environment except where intramolecular quenching effects (as with FAD [359, 361] and NAD [361]) are possible. If binding to a protein is to be detected by fluorescence, measurements of polarization are often best [368, 370] because wavelength shifts and yield changes on binding are small, though not always too small to measure.[371] The causes of the low environmental sensitivity in the fluorescence of etheno derivatives and the high quantum yields in comparison with free nucleotides have been discussed.[360] Rigidity may be the key factor here as in some other systems.[307, 372]

It is interesting to note that the fluorescent 'Y' bases found in some tRNA's, whilst not strictly ethenoadenosine derivatives, contain a closely related modification based on guanine.[373] 3,N^4-Ethenocytosine derivatives (including ethenoCTP) have been synthesized, and it has been shown that ethenoCTP replaces ATP in the reaction with glyceraldehyde-3-phosphate dehydrogenase.[374] There are clearly some respects in which ethenoCTP resembles ATP more closely than does ethenoATP itself. The preparation of γ[^{32}P]-labelled ethenotriphosphate derivatives has also been described.[374] The conversion of ethenoadenosine derivatives into fluorescent 2-aza-1,N^6-ethenoadenosine compounds has been reported.[375]

Space precludes detailed comment on other model studies listed in Table 1, except to record the least contentious statement of 1973: 'It thus appears that we must reconcile ourselves to the possibility that nature will not in every instance abide by the dictates of approximate molecular orbital computations'.[376]

Probe Design. The ideal fluorescent probe combines the convenient biochemical properties of an etheno-derivative with an environmentally sensitive fluorescence emission when interacting with the system under study. The design and application of such molecules have developed considerably since the time, way back at the turn of the decade, when nearly all the work was with anilino- and toluidino-naphthalenesulphonates (non-covalent probes) and with dansyl and fluorescein (covalent labels) (see the early reports in this series). Table 2 indicates the extent to which the situation has changed.

[370] J. R. Barrio, J. A. Secrist, Y. Chien, P. J. Taylor, J. L. Robinson, and N. J. Leonard, *F.E.B.S. Letters*, 1973, **29**, 215.
[371] D. A. Hilborn and G. G. Hammes, *Biochemistry*, 1973, **12**, 983; C. Tondre and G. G. Hammes, *Biochim. Biophys. Acta*, 1973, **314**, 245.
[372] M. Nishikimi and M. Yoshino, *J. Biochem.* (*Japan*), 1972, **72**, 1237.
[373] S. H. Blobstein, D. Grunberger, I. B. Weinstein, and K. Nakanishi, *Biochemistry*, 1973, **12**, 188.
[374] J. R. Barrio, L. G. Dammann, L. H. Kirkegaard, R. L. Switzer, and N. J. Leonard, *J. Amer. Chem. Soc.*, 1973, **95**, 961.
[375] K. F. Yip and K. C. Tsou, *Tetrahedron Letters*, 1973, 3087.
[376] B. S. Hudson and B. E. Kohler, *Chem. Phys. Letters*, 1973, **23**, 139.

Table 2 *Fluorescence probes and their applications*

Probe	*Application*	*Ref.*
N-(9-Acridinyl) maleimide	SH reagent	378
1-Anilinonaphthalene-8-sulphonate	Human serum albumin	406
	Phosphofructokinase	*a*
	Human luteinizing hormones	411
	Pyridine nucleotide transhydrogenase	*b*
N-(1-Anilinonaphthyl)-4-maleimide	SH reagent	377
Aurovertin	Mitochondrial ATPase	*c*
2-Aza-1,N^6-ethenoadenine derivatives	Fluorescent adenine analogues	375
Berberine	Liver alcohol dehydrogenase	395
Bis-(3-aminopyridinium)-1,10-decane	Cholinergic receptor protein	*d*
Bis-(1-anilinonaphthalene-8-sulphonate)	Ribosomal proteins	408
7-Chloro-4-nitrobenzo-2-oxa-1,3-diazole	Papain	392
Cytidine	Fluorescent modifications	396
Dansyl	Bovine trypsin	321
	Immunoglobulins IgG and IgE	414
	Poly-L-lysine	407
Dansyl and anthraniloyl	Attachment to 5′ end of tRNA	391
Dansyl and fluorescein	Immunoglobulin IgG	354
1-(Dansylamido)-3-*NN*-dimethylaminopropane	Horse serum cholinesterase	386
1-Dansyl-3-trimethylamino-propane	Acetylcholine receptor protein	*d*
DNA/9-aminoacridine	Poly-L-lysine	*e*
5-Dibutylaminonaphthalene-1-sulphonyl chloride	NH_2 reagent	383
1,N^6-Ethenoadenine nucleotides	Pyruvate kinase	370
1,N^6-EthenoADP	Mitochondrial ATPase	371
	Myosin, heavy meromyosin and subfragment I	364
1,N^6-EthenoAMP and 1,N^6-ethenoATP	Luciferase	369
1,N^6-EthenoATP	Glutamine synthetase	393
	Heavy meromyosin	339
	Aspartate transcarbamylase	368
Fluorescein	Concanavalin A	417
	Glutamic aspartic transaminase	*f*
	Prothrombin	*g*
	Lactate dehydrogenase	*h*
	Glyceraldehyde-3-phosphate dehydrogenase	*i*
	Sorbitol dehydrogenase	*j*
2-(6′-Hydroxy-2-benzothiazolyl)-4-hydroxymethylthiazole	Luciferase	390
N-(Iodoacetylaminoethyl)-5-naphthylamine-1-sulphonate and the 1,8-isomer	SH reagent	381
L-Kynurenine	Serum albumins	394

Table 2 (*cont.*)

Probe	*Application*	*Ref.*
2-Methoxy-2,4-diphenyl-3(2*H*)-furanone	NH_2 reagent	382
N-Methyl-2-anilinonaphthalene-6-sulphonyl peptides	Pepsin	387
4-Methylumbelliferyl-α-D-manno-pyranoside	Concanavalin A	399
Nitrosonaphthol	Tyr reagent	384
N-(1-Oxyl-2,2,6,6-tetramethyl-4-piperidinyl)-5-dimethylamino-naphthalene sulphonamide	Spin-labelled fluorescent probe	*k*
N-(3-Pyrene)-maleimide	SH reagent	379
Sulphonphthalein dyes	Various proteins	372
Tetracyclines	Ribosomal subunits	*l*
	Bovine serum albumin	409
2-*p*-Toluidinylnaphthalene-6-sulphonate	Placental 17β-hydroxy-steroid dehydrogenase	*m*
	Poly-L-lysine	*n*
	Ile:tRNA ligase	412
	Phe:tRNA ligase	413
Y-base($tRNA^{Phe}$)	Phe:tRNA ligase	*o*

(*a*) D. P. Bloxham, *Biochemistry*, 1973, **12**, 1596. (*b*) R. R. Fisher and N. O. Kaplan, *Biochemistry*, 1973, **12**, 1182. (*c*) D. Layton, A. Azzi, and P. Graziotti, *F.E.B.S. Letters*, 1973, **36**, 87; T. M. Chang and H. S. Penefsky, *J. Biol. Chem.*, 1973, **248**, 2746. (*d*) J. B. Cohen and J. P. Changeux, *Compt. rend.*, 1973, **277**, *D*, 603. (*e*) H. Hasumi, K. Akasaka, H. Hatano, and K. Hiromi, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 992. (*f*) C. G. Dimitropoulos, N. G. Oikonomakos, I. A. Karni-Katsadina, T. G. Kalogerakos, and A. E. Evangelopoulos, *European J. Biochem.*, 1973, **38**, 537. (*g*) D. N. Fass and K. J. Mann, *J. Biol. Chem.*, 1973, **248**, 3280. (*h*) D. Dimitrov and G. Detchev, *Stud. Biophys.*, 1972, **34**, 231. (*i*) D. S. Markovich and G. B. Krapivinskii, *Biokhimiya*, 1973, **38**, 1106. (*j*) J. R. Heitz, *J. Biol. Chem.*, 1973, **248**, 5790. (*k*) R. A. Long and J. C. Hsia, *Canad. J. Biochem.*, 1973, **51**, 876. (*l*) G. Fey, M. Reiss and H. Kersten,, *Biochemistry*, 1973, **12**, 1160. (*m*) J. C. Mani, J. Dornand, M. Mousseron-Canet, and F. Vial-Reveillon, *Biochimie*, 1973, **55**, 851. (*n*) G. Witz and B. L. Van Duuren, *J. Phys. Chem.*, 1973, **77**, 648. (*o*) G. Krauss, R. Romer, D. Riesner, and G. Mass, *F.E.B.S. Letters*, 1973, **30**, 6.

It is possible to distinguish a few common strands. The ideal covalent label is chemically specific for a particular functional group, and preferably it has a large change in emission properties on reaction with a protein. It is sometimes possible to modify a group-specific reagent of known reactivity (*e.g.* maleimide) by attachment of a fluorophore of known environmental sensitivity (*e.g.* anilinonaphthyl [377] or acridinyl [378]) or with a useful property for a particular application (*e.g.* pyrene, with its long fluorescence lifetime [379]). Linking the fluorophore to maleimide causes strong fluorescence quenching [377, 378] which produces the ideal situation of a non-fluorescent reactant forming a highly fluorescent product. An older

[377] Y. Kanaoka, M. Machida, M. Machida, and T. Sekine, *Biochim. Biophys. Acta*, 1973, **317**, 563.

[378] Y. Nara and K. Tuzimura, *Bunseki Kagaku*, 1973, **22**, 451.

[379] J. K. Weltman, R. P. Szaro, A. R. Franckelton, R. M. Dowben, J. R. Bunting, and R. E. Cathou, *J. Biol. Chem.*, 1973, **248**, 3173.

—SH specific reagent with these properties is 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole, which was used this year in a study of papain.[380] Slightly different effects are found with —SH specific *N*-(iodoacetylaminoethyl)-naphthalenesulphonates, which are themselves readily susceptible to photodegradation, but which form a stable product after reaction with a protein.[381]

A promising suggestion for an $-NH_2$ specific fluorogenic reagent is 2-methoxy-2,4-diphenyl-3(2*H*)-fluoranone (a close relative of the popular analytical reagent, fluorescamine).[382] The disadvantage of fluorescamine itself is the instability of the fluorescent product in aqueous solution. The new reagent forms a product which is stable over a wide pH range. In this case (as with fluorescamine) excess reagent reacts rapidly with water to produce a non-fluorescent product. Another $-NH_2$ reagent, 5-dibutylaminonaphthalene-1-sulphonyl chloride, resembles dansyl chloride though the reaction products are slightly more soluble in non-aqueous solvents.[383] A Tyr-specific reagent with potential is nitrosonaphthol which gives a fluorescent product (perhaps a benzphenoxazinone) with Tyr.[384]

Another common principle is to prepare a fluorescent substrate or inhibitor analogue, so that interactions with a protein, though non-covalent, are likely to be specific for a given site. The ethenonucleoside derivatives fall into this class, but there are many other examples (Table 2).[353, 385–392] Sometimes it is even possible to attach a fluorescent group to a protein covalently by way of a suitable specific substrate, *e.g.* the etheno-adenylation of glutamine synthetase with ethenoATP.[393] Some specific fluorescent probes are produced naturally, *e.g.* L-kynurenine is an analogue of Trp.[394]

Harder to categorize at all are the comparatively unspecific probes, mostly non-covalent, many of which have advantages for particular applications. For some, emission properties show a strong dependence on the environment. A common (and perhaps significant) feature is that they contain an easily twisted link between two conjugated systems (*e.g.* anilino- and toluidino-naphthalenesulphonates,[307] kynurenine,[394]

[380] G. Allen and G. Lowe, *Biochem. J.*, 1973, **133**, 679.
[381] E. N. Hudson and G. Weber, *Biochemistry*, 1973, **12**, 4154.
[382] M. Weigele, S. De Bernardo, W. Leimgruber, R. Cleeland, and E. Grunberg, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 899.
[383] N. Seiler, T. Schmidt-Glenewin, and H. H. Schneider, *J. Chromatogr.*, 1973, **84**, 95.
[384] R. Hakanson, A. L. Roennberg, and K. Sjoelund, *Analyt. Biochem.*, 1973, **51**, 523.
[385] C. H. Yang and D. Soell, *J. Biochem. (Japan)*, 1973, **73**, 1243.
[386] L. M. Chau, C. M. Himel, and A. R. Main, *Biochemistry*, 1974, **13**, 86.
[387] G. P. Sachdev, A. D. Brownstein, and J. S. Fruton, *J. Biol. Chem.*, 1973, **248**, 6292.
[388] B. R. Dean and R. B. Homer, *Biochim. Biophys. Acta*, 1973, **322**, 141.
[389] W. R. McClure and K. H. Scheit, *F.E.B.S. Letters*, 1973, **32**, 267.
[390] L. J. Bowie, V. Horak, and M. DeLuca, *Biochemistry*, 1973, **12**, 1845.
[391] C. H. Yang and D. Soell, *Arch. Biochem. Biophys.*, 1973, **155**, 70.
[392] A. Bienvenue and J. Tournon, *Biochimie*, 1973, **55**, 1167.
[393] P. B. Chock, C. Y. Huang, R. B. Timmons, and E. R. Stadtman, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3134.
[394] J. E. Churchich, *Biochim. Biophys. Acta*, 1972, **285**, 91.

berberine,[395] some sulphonaphthalein dyes,[372] and some cytidine modifications [396]).

Applications. Having devised and obtained a fluorescent probe which interacts in an observable way with a protein, different types of information can be obtained. Sometimes the stoicheiometry, strength and kinetics of binding, and the properties of the bound probe itself are interesting, particularly if it is a substrate,[397] nucleotide,[348, 398–403] prosthetic group,[404] or some analogue.[339, 371, 388] Sometimes changes in binding following chemical modification of the protein can be studied, as in an investigation of the interaction between glucose oxidase and FAD.[405] In some systems spectral changes can suggest things about the nature of the binding site; for example, the interaction of pepsin with a number of labelled peptides was studied and it was found that 'It is difficult to accommodate the findings with a relatively rigid extended active site that can be mapped in terms of subsites; instead pepsin appears to exhibit considerable conformational flexibility at the active site in response to enzyme substrate (or enzyme inhibitor) interaction'.[387] Energy transfer measurements can be interpreted to give the distances between fluorophores in the probe–protein complex. Sometimes both donor and acceptor are extrinsic probes (as in the immunoglobulin case already discussed [354]), but in other cases transfer from protein Trp to probe is studied, *e.g.* the interaction of two probes at the anionic subsite of horse serum cholinesterase.[386]

Probe fluorescence can also be used to monitor conformation [340, 406, 407] or aggregation changes (for molecules ranging from small peptides to ribosomes) [408–411] or the binding of non-fluorescent substrates.[412, 413] The extreme simplicity with which such measurements can often be performed makes the technique suitable for the collection of large quantities of data (using different physical conditions, on a large variety of ligands) and also

[395] J. Kovar and L. Skursky, *European J. Biochem.*, 1973, **40**, 233.
[396] J. R. Barrio and N. J. Leonard, *J. Amer. Chem. Soc.*, 1973, **95**, 1323.
[397] G. Krauss, R. Roemer, D. Riesner, and G. Maass, *F.E.B.S. Letters*, 1973, **30**, 6.
[398] J. H. Yuan and B. M. Anderson, *Arch. Biochem. Biophys.*, 1973, **156**, 328.
[399] M. Cassman, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 666.
[400] P. L. Coleman and H. Weiner, *Biochemistry*, 1973, **12**, 1705.
[401] K. E. Lind, *European J. Biochem.*, 1973, **33**, 67.
[402] P. L. Luisi, A. Olomucki, A. Baici, and D. Karlovic, *Biochemistry*, 1973, **12**, 4100.
[403] R. B. Wallis and J. J. Holbrook, *Biochem. J.*, 1973, **133**, 173.
[404] C. Vernotte and I. Moxa, *Photochem. and Photobiol.*, 1973, **17**, 245.
[405] H. Tsuge and H. Mitsuda, *J. Biochem. (Japan)*, 1973, **73**, 199.
[406] G. Sudlow, D. J. Birkett, and D. N. Wade, *Mol. Pharmacol.*, 1973, **9**, 649.
[407] T. Ito, *Bull. Chem. Soc. Japan*, 1973, **46**, 335.
[408] F. Pochon and B. Ekert, *European J. Biochem.*, 1973, **36**, 311.
[409] J. K. H. Ma, H. W. Jun, and L. A. Luzzi, *J. Pharm. Sci.*, 1973, **62**, 1261.
[410] B. I. Kurganov, N. P. Lisovskaya, N. B. Livanova, and T. B. Eronina, *Biochemistry (U.S.S.R.)*, 1973, **38**, 199.
[411] S. M. Aloj, K. C. Ingham, and H. Edelhoch, *Arch. Biochem. Biophys.*, 1973, **155**, 478; K. C. Ingham, S. M. Aloj, and H. Edelhoch, *ibid.*, 1973, **159**, 596.
[412] E. Holler, *Biochemistry*, 1973, **12**, 1142; E. Holler, P. Rainey, A. Orme, E. L. Bennett, and M. Calvin, *ibid.*, p. 1150.
[413] H. M. Kosakowski and E. Holler, *European J. Biochem.*, 1973, **38**, 274.

for fast reaction studies. A study illustrating these points involves the emission of TNS bound to isoleucine–tRNA ligase.[412]

Another approach uses the anisotropy of probe emission to yield information about rigidity and molecular motion. Usually such studies use covalently labelled proteins, *e.g.* dansyl-labelled immunoglobulins IgG and IgE. Although neither molecule is rigid, IgE is much less mobile than IgG.[414] Polarization studies of dansyl derivatives of human serum albumin [415] and plasminogen [416] have also been reported. The interaction with different cells (normal and malignant) of the fluorescein conjugate of concanavalin A has been investigated. It appears that, whereas malignant transformation of cells in suspension *in vivo* reduces the mobility of concanavalin binding sites, the reverse effect is found for cells in solid tissue.[417] The binding of fluorescent substrates can be studied in a similar fashion, and fluorescence polarization studies have been used to show that, when ethenoADP and ethenoATP bind to pyruvate kinase, the fluorescent parts of the nucleotides remain more mobile than the protein as a whole.[370]

Intrinsic Fluorescence.—As with probes, there have been model studies using intrinsic fluorescence as well as investigations of conformation, aggregation, and binding effects. Dealing briefly with the first of these categories, there have been a number of studies of the environmental sensitivity of Trp [418–420] and Tyr [418] emissions, and reports on the determination of the fluorescence of His [421] and iodotyrosines.[422] The formation of complexes between aromatic amino-acids and nucleotides has been studied.[423] The emission characteristics of tryptophanyl-tRNA^{Trp} and some modified derivatives have been reported.[424]

Protein Conformation. Various intrinsic emission characteristics of proteins have been used to learn about molecular conformation. The wavelength dependence of fluorescence decay curves has been used to distinguish Trp residues in human serum albumin,[322] and quenching effects have been interpreted to give similar information for a number of species.[342, 343] Particularly interesting is some work on the quenching effects of O_2 at high pressures.[347] The fundamental observation was that at pressures which have no effect on enzyme activity, the fluorescence emissions of all

[414] R. S. Nezlin, Y. A. Zagyansky, and A. I. Kaeivaeraeinen, *Immunochemistry*, 1973, **10**, 681.
[415] B. N. Preston, B. Obrink, and T. C. Laurent, *European J. Biochem.*, 1973, **33**, 401.
[416] E. J. Castellino, W. J. Brockway, J. K. Thomas, H. Liao, and B. Rawitch, *Biochemistry*, 1973, **12**, 2787.
[417] M. Inbar, M. Shinitzky, and L. Sachs, *J. Mol. Biol.*, 1973, **81**, 245.
[418] R. McGuire and I. Feldman, *Photochem. and Photobiol.*, 1973, **18**, 119.
[419] J. Chrysochoos, *Mol. Photochem.*, 1973, **5**, 1.
[420] J. Bello and H. R. Bello, *European J. Biochem.*, 1973, **34**, 535.
[421] I. Tatischeff, P. Vigny, R. Klein, and M. Duquesne, *Compt. rend.*, 1973, **276**, *D*, 1217.
[422] E. E. Gussakovskii, T. Saatov, and T. A. Badaev, *Mol. Biol.*, 1973, **7**, 753.
[423] C. Hélène, *Photochem. and Photobiol.*, 1973, **18**, 255; T. Montenay-Garesti and C. Hélène, *J. Agric. Food Chem.*, 1973, **21**, 11; T. Montenay-Garesti and C. Hélène, *J. Chim. phys. physicochem. biol.*, 1973, **70**, 1391.
[424] R. H. Buckingham and P. Danchin, *F.E.B.S. Letters*, 1973, **30**, 236.

the Trp residues in a number of proteins were quenched. An energy-transfer mechanism was ruled out, and it was concluded that O_2 (unlike I^- and other charged quenchers) can approach close to all Trp residues, whether 'exposed' or 'buried' by other criteria. Thus 'the functional properties of protein molecules are not properly represented by rigid models that do not include the rapid structural fluctuations necessary to explain . . . O_2 quenching'.

The shapes of excitation and emission spectra can suggest both the polarity of the environment of the emitting groups and the extent of energy transfer between residues (especially Tyr → Trp).[338] The pH-dependence of energy transfer has been studied for glucagon.[425] The efficiencies of transfer from Tyr and Tyr^- are different, and it has been possible to correlate the spectra found at several pH values with a molecular model of glucagon coiled into a compact sphere. Energy transfer has also been invoked to explain differences between the fluorescence spectra of α-lactalbumins isolated from four species: guinea-pig, goat, ox, and human.[426] The fluorescence yield from Trp residues in the guinea-pig protein is about twice that in any of the others, and it is suggested that this is related to the absence of Trp-60. Using the lysozyme analogy model, this can be rationalized by assuming that in the presence of Trp-60 there is an efficient quenching mechanism linking Trp-104 → Trp-60 → —S—S—. Thus in the absence of Trp-60 the yield from Trp-104 increases.

Other energy-transfer effects which have been studied include Trp → probe (*e.g.* in a study of the interaction of ethenoATP with heavy meromyosin[339]) and energy transfer (including triplet–triplet and also non-resonance transfer) of proteins in glasses (*e.g.* carboxypeptidase[427] and α-trypsin[428]). The use of both intrinsic fluorescence and probe emission in the same study has also been employed in a study of thermally induced structural transitions in Bence-Jones proteins.[429]

Emission intensity and polarization of fluorescence can be used to monitor conformation changes. Systems studied in this way include bovine plasma albumin,[430] ribosomal proteins S-4 and S-7 (from *E. coli*),[431] and human thyroglobulin under normal conditions and in nodular goitre.[432] The anisotropic u.v. fluorescence of collagen fibres has been studied,[433]

[425] R. M. Epand, *Photochem. and Photobiol.*, 1973, **18**, 245.
[426] P. B. Sommers, M. J. Kronman, and K. Brew, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 98.
[427] N. Shaklai, N. Zisapel, and M. Sokolovsky, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2025.
[428] C. A. Ghiron, J. W. Longworth, and N. Ramachandran, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3703.
[429] A. C. Ghose, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 240.
[430] M. Sogami, S. Nagaoka, K. B. Itoh, and S. Sakata, *Biochem. Biophys. Acta*, 1973, **310**, 118; M. Sogami, M. Uyeda, and S. Ogura, *ibid.*, p. 487.
[431] G. Lemieux and D. Gerard, *F.E.B.S. Letters*, 1973, **37**, 234.
[432] Y. K. Turakulov, T. Saatov, P. A. Khakimov, and M. Yangibaev, *Biochemistry (U.S.S.R.)*, 1973, **37**, 930.
[433] U. S. Borovikov, L. V. Kukhareva, U. M. Rozanov, and V. F. Vorob'ev, *Tsitolgiya*, 1973, **15**, 1097.

and the polarization of lysozyme emission has been measured as a preliminary to a binding study.[434] Fluorescence has been used in the detection of differences between proteins from two sources, *e.g.* normal and pathological thyroglobulins,[435] and to monitor the extent of chemical modification of Trp residues, *e.g.* work on ovine pituitary lactogenic hormone,[436] papain,[437] and azobacter flavodoxin.[438]

Table 3 *Binding studied by intrinsic protein fluorescence*

Protein	*Ligand(s)*	*Ref.*
Acetylcholinesterase	*N*-Methyl acridine and *N*-methylpyridine	*a*
	Bis-quaternary ammonium ligands	*b*
Azobacter flavodoxin	FMN	438, 440
Bovine α_{s1}-casein	Ca^{2+}	*c*
Concanavalin A	Mn^{2+} and rare-earth ions	*d*
Dihydrofolate reductase	NADH, dihydrofolate	*e*
Dihydropteridine reductase	Quinoid dihydropteridine	401
Egg-white flavoprotein	FMN	*f*
Glyceraldehyde-3-phosphate dehydrogenase	NAD	*g*
Immunoglobulin IgG	Univalent haptens	356
Lysozyme	Oligosaccharides	434, 444, 445
Melittin	Lecithin	446
Met:tRNA ligase	$tRNA^{Met}$	443
Myeloma proteins	Phosphorylcholine	*h*
Myosin	ATP	*i*
Octopine dehydrogenase	NADH	402
Pyrocatechase	Fe^{III}	*j*
Ser:tRNA ligase	$tRNA^{Ser}$ and Ser:RNA^{Ser}	*k*
Streptomyces R61 DD-carboxy-peptidase transpeptidase	Penicillin	442
Wheat-germ agglutinin	Saccharides	*l*

(*a*) M. Shinitzky, Y. Dudai, and I. Silman, *F.E.B.S. Letters*, 1973, **30**, 125. (*b*) P. Taylor and N. M. Jacobs, *Mol. Pharmacol.*, 1974, **10**, 93. (*c*) D. G. Dalgleish, *European J. Biochem.*, 1973, **40**, 375. (*d*) A. D. Sherry and G. L. Cottam, *Arch. Biochem. Biophys.*, 1973, **156**, 665. (*e*) J. S. Erickson and C. K. Mathews, *Biochemistry*, 1973, **12**, 372. (*f*) M. Nishikimi and Y. Kyogoku, *J. Biochem.* (*Japan*), 1973, **73**, 1233. (*g*) D. L. Sloan and S. F. Velick, *J. Biol. Chem.*, 1973, **248**, 5419. (*h*) R. Pollet and H. Edelhoch, *J. Biol. Chem.*, 1973, **248**, 5443. (*i*) E. M. Mandelkow and E. Mandelkow, *F.E.B.S. Letters*, 1973, **33**, 161. (*j*) K. Nagami, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 364. (*k*) A. Pingoud, D. Reisner, D. Boehme, and G. Mass, *F.E.B.S. Letters*, 1973, **30**, 1. (*l*) R. Lotan and N. Sharon, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 1340.

[434] V. I. Teichberg and M. Shinitzky, *J. Mol. Biol.*, 1973, **74**, 519.
[435] Y. K. Turakulov, T. A. Babaev, E. E. Gussakovskii, T. Saatov, and U. K. Makhmudov, *Mol. Biol.*, 1973, **7**, 732.
[436] A. Kawauchi, T. A. Bewley, and C. H. Li, *Biochemistry*, 1973, **12**, 2124.
[437] J. E. Mole and H. R. Horton, *Biochemistry*, 1973, **12**, 5278.
[438] J. Ryan and G. Tollin, *Biochemistry*, 1973, **12**, 4550.

Binding Interactions. Intrinsic fluorescence measurements are also widely used in investigation of the interactions between proteins and other molecules. A number of such studies have already been mentioned in other contexts: these, and some other reports, are listed in Table 3. Any one or more of the commonly studied fluorescence parameters may be sensitive to an intermolecular interaction, and although intensity is generally the simplest to measure, both fluorescence polarization (especially for interactions with other macromolecules)[434, 439] and fluorescence lifetimes[440] can be used. It is less straightforward, but still useful in some instances where other techniques fail, to monitor the rate at which fluorescence changes when the protein undergoes a reaction. The effects of ligands on these rates (usually to inhibit them) can then be interpreted to give dissociation constants. The effects of trypsin on phosphorylase[441] and urea on dihydropteridine reductase have been treated in this way.[401]

The simplicity of using fluorescence titrations of some description to study binding interactions often permits a large number of separate measurements in a short time using only a small amount of material. For example the binding of penicillin quenches the fluorescence of the *Streptomyces* R61 DD carboxypeptidase transpeptidase, and this effect has been used to measure the effects of different salt concentrations on the penicillin binding.[442] A second, time-dependent fluorescence change (attributed to a conformational effect) associated with binding could not be detected by c.d., and so presumably it reflects only a small alteration in three-dimensional structure. Another example showing the advantages offered by the simplicity of fluorescence titrations is a study of the binding of various ligands (tRNA, Met, *etc.*) to methionine:tRNA ligase.[443] Many different combinations were observed under a variety of physical conditions.

There has been the usual crop of studies with lysozyme,[434, 444, 445] as well as investigation of less common proteins like melittin from bee venom.[446] Human lysozyme has been compared with the hen egg-white enzyme in the pH-dependence of its emission properties.[444] Although the two enzymes are similar in the absence of substrate, this is not the case in the presence of oligomers of *N*-acetylglucosamine. The differences under these conditions are attributed to the substitution of Trp-62 (egg-white) by Tyr-63 (62) (human enzyme). A very detailed study of the egg-white enzyme under many different conditions illustrates again the value of the

[439] J. A. D'Anna and I. Isenberg, *Biochemistry*, 1973, **12**, 1035.
[440] L. J. Andrews, M. L. MacKnight, J. Ryan, and G. Tollin, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 1165.
[441] N. B. Livanova, B. I. Kurganov, and N. P. Lissovskaja, in 'Mechanisms and Control Properties of Phosphotransferases', Joint Biochemical Symposium USSR–GDR, Akademie-Verlag, Berlin, 1973, p. 307.
[442] M. Nieto, H. R. Perkins, J. M. Frere, and J. M. Ghuysen, *Biochem. J.*, 1973, **135**, 493.
[443] S. Blanquet, M. Iwatsubo, and J. P. Waller, *European J. Biochem.*, 1973, **36**, 213; S. Blanquet, M. Iwatsubo, and J. P. Waller, *ibid.*, p. 227.
[444] R. S. Mulvey, R. J. Gualtieri, and S. Beychok, *Biochemistry*, 1973, **12**, 2683.
[445] S. K. Bannerjee and J. A. Rupley, *J. Biol. Chem.*, 1973, **248**, 2117.
[446] C. Mollay and G. Kreil, *Biochim. Biophys. Acta*, 1973, **316**, 196.

simplicity of fluorometric measurements of binding.[445] The work with melittin has shown that this lytic peptide (26 residues with Trp-19) has an emission which is sensitive to the binding of lecithin but not cholesterol.[446]

7 Infrared Spectroscopic Studies of Proteins

Contributed by R. M. Stephens

Model Compounds.—New i.r. overtone and combination bands of several primary and secondary amides have been assigned to amide group vibrational modes. Three bands in the 6700—7000 cm^{-1} region from $RCONH_2$ (R = H, Me, or Ph) have been attributed to the $2\nu_{asym}$, $\nu_{asym} + \nu_{sym}$, and $2\nu_{sym}$ modes of the NH group, where ν_{asym} and ν_{sym} are the antisymmetrical and symmetrical stretching vibrations of the group. Bands resulting from the combination of ν_{asym} and ν_{sym} with the amide II and amide III vibrations were detected in the region 4700—5100 cm^{-1}. The 2ν(NH) overtone, the [ν(NH) + amide II] and the [ν(NH) + amide III] combination bands in the near i.r. of five secondary amides were corroborated. A [2ν(C=O) + amide III] combination band for the primary amides was confirmed for the secondary amides. Anharmonicity constants for the N—H stretching modes are of the same magnitude as those for primary and secondary amines. The Beer Law studies indicated that the near i.r. bands can be used for quantitative analysis of amides.[447] Hydrogen-bonding studies from *N*-methylacetamide–*NN'*-dimethylamide complexes and *N*-methylacetamide–*NN'*-diphenylamide complexes have shown that there is a stronger association between the former pair than the latter. I.r. frequencies of free and bonded N—H bands, absorbances of free N—H bands, and equilibrium constants for the H-bonded complexes are given and the results are explained on the basis of varying contributions from dipolar resonance structures.[448] Analysis of the i.r. and n.m.r. spectra of *N*-ethylacetamide in the liquid state in polar and non-polar media over a broad temperature and concentration range revealed the presence of a single pair of equivalent mirror image conformers, with the conformational angle ψ in the vicinity of $\pm 75°$. *N*-Isopropylacetamide exists in a similar single conformer or conformer pair.[449] The i.r. dichroism of MeCONHCOMe, $CD_3CONHCOCD_3$, and their *N*-deuteriated compounds in the stable crystalline form, where the molecules have a *trans-cis* conformation, has been examined between 400 and 4000 cm^{-1}. Vibrational assignments were made on the basis of isotopic shifts and i.r. dichroism. A normal-coordinate treatment was made for the out-of-phase vibrations of the *trans-cis* planar model, and from the torsional constants about the C—N bonds the potential barrier for internal rotation about the C—N bond was found to be approximately 110 kJ mol^{-1}.[450]

[447] S. E. Krikorian and M. Mahpour, *Spectrochim. Acta* (*A*), 1972, **29**, 1233.
[448] K. Ravindranath and K. Ramiah, *Current Sci.*, 1973, **42**, 706.
[449] P. Schmidt, D. Doskocilova, and B. Schneider, *J. Mol. Structure*, 1973, **15**, 383.
[450] Y. Kuroda, Y. Saito, K. Machida, and T. Uno, *Spectrochim. Acta* (*A*), 1973, **29**, 411.

The i.r. spectra of 13 N-monosubstituted acetamides and four *N*-alkylamides showed that, in the non-associated state at least, two ν(NH) absorption bands were present from each amide. These were assigned to various conformations in terms of rotation about the C—N bond.[451]

The potassium bromide disc technique has been used to obtain the i.r. spectra of many amino-acids and dipeptides in the hydrogenated and deuteriated forms. These studies, coupled with those in aqueous media over a wide range of pH or pD, have allowed assignments of various group frequencies to be given.[452]

Polypeptide Conformation.—The intensities and other spectral parameters of the main components of i.r. amide bands for the random and antiparallel pleated sheet forms were detected for poly-*S*-carbobenzoxymethylcysteine, polylysine, and silk fibroin of *Bombyx mori* in heavy-water solutions and in organic solvents. Assuming that the optical spectra of these two types of structure are additive, a method was proposed for determining relative contents and molar absorption coefficients of the amide I band. The integral intensity and maximum frequency values of the main components of the amide I band were specific for all the samples in the β-form, whereas the integral intensity of the amide I band in the random coil form varied within certain limits. The maximum frequency for all the samples remained the same.[453] The i.r. spectra of copoly-γ-benzyl DL-glutamates and sequential copolymers of γ-benzyl D- and L-glutamates in the region 700—200 cm^{-1} showed that each chain of the meso copolymers of γ-benzyl D- and L-glutamates had a random coil portion and an α-helical portion, and that the latter consists of the L and D residues with the ratio of D : L either 5 : 1 or 1 : 5.[454] The sequential copolymers glycine-L-alanine, L-leucine-L-alanine, L-phenylalanine-L-alanine, and some containing L-proline residues have been synthesized and the i.r. spectra in the region 700—200 cm^{-1} measured, a comparison of structures being made. Several bands in the region 600—200 cm^{-1} clearly reflect the backbone conformation and also the local conformation of component amino-acid residues of polypeptides with the α-helical, β-form, and polyglycine structures.[455] The copolypeptide γ-methyl D-glutamate–γ-methyl L-glutamate was synthesized in the presence of Et_3N as an initiator and the copolypeptide molecules were shown to exist in a conformation very similar to that of the α-helix of either poly-γ-methyl D-glutamate or poly-γ-methyl L-glutamate. The sum of right- and left-handed helix contents was determined from the amide bands in the i.r. spectrum. O.r.d. studies of the copolypeptide in helicogenic solvents was used to ascertain the helical character.[456]

[451] J. Smolikova, A. Vitek, and K. Blaha, *Coll. Czech. Chem. Comm.*, 1973, **38**, 548.
[452] J. F. Pearson and M. A. Slifkin, *Spectrochim Acta* (*A*), 1972, **28**, 2403.
[453] Y. Chirgadze, B. V. Shestopalov, and S. Y. Ven'yaminov, *Biopolymers*, 1973, **12**, 1337.
[454] K. Itoh, T. Ozaki, K. Nagayama, A. Wada, and M. Tsuboi, *Macromolecules*, 1973, **6**, 654.
[455] K. Itoh and H. Katabuchi, *Biopolymers*, 1973, **12**, 921.
[456] A. Nakajima, T. Hayashi, K. Itoh, and T. Fujiwara, *Polymer. J.*, 1973, **4**, 10.

Both the content of the hydrogen-bonded β-form of poly-γ-benzyl L-glutamate (PBLG) and the extent of association have been examined in the following solvents: dimethylformamide, trifluoroethanol, trimethylphosphate, chloroform, dioxan, ethylene dichloride, and ethylene dibromide. The i.r. spectra suggested that the conformation of PBLG in dichloroacetic acid differed from either the isolated random chain or β-conformation found in other solutions, and an approximate method was developed for estimating the content of β-structure from a single spectrum of dissolved PBLG. The reduced viscosities of some solutions were scarcely dependent on the PBLG concentration when a single conformation predominated over a given concentration range, whilst in other solutions the reduced viscosity showed a strong concentration dependence or an anomaly. The observed viscosity behaviour was attributed to changes in size and shape of the aggregates, which was determined by the number of hydrogen bonds in the aggregate.[457]

The nature of the peptide–acid specific interactions responsible for the helix–random-coil transition of polymer chains has been studied by comparing the i.r. spectra of poly-L-alanine–trifluoroacetic acid mixtures with those of the model amide *N*-methylacetamide. The model compound was studied to determine the spectral characteristics of the different complexes or species formed between amide and acid. At a low concentration hydrogen-bonded complexes were observed but no association between amide N—H and acid C—O groups was detected. Amidium and carboxylate bonds were localized at 1680, 1705 and 620, 1625 cm^{-1}, respectively. If the cation band was observed, the anion band appeared only for the most acidic solutions. For the polymer, proton transfer similar to that noted with a model amide was never observed. Thus the helix–random-coil transition in polypeptides may not be due to protonation of the peptide group but to a strong hydrogen-bond interaction between the polymer and acid molecules.[458]

Using Wilson's G-F matrix method as modified by Higgs for infinite helical polymers, dispersion curves and frequency distribution functions have been calculated for poly-L-proline II. The i.r. spectrum was recorded and a Urey–Bradley force field was evaluated to provide the best fit for the observed frequencies. The results were discussed in relation to the conformation characteristics of the two crystal modifications of poly-L-proline.[459] A satisfactory interpretation of the splitting observed from the amide I modes can be obtained by including in the frequency calculation the D_{11} term. The physical origin of this term is explained by transition dipole coupling. This mechanism also explains the splitting of the C=O stretching vibration in H-bonded carboxylic acid dimers. The D_{10} interaction constant is *ca.* 0 rather than the large value required by previous

[457] T. Imae and S. Ikeda, *Biopolymers*, 1973, **12**, 1203.
[458] P. Combelas, C. Garrigou-Lagrange, and J. Lascombe, *Biopolymers*, 1973, **12**, 611.
[459] V. D. Gupta, R. D. Singh, and A. M. Dwivedi, *Biopolymers*, 1973, **12**, 1377.

applications of perturbation treatment.[460] The influence of the substituents upon the amide I and ν(NH) frequencies has been analysed for ten amides substituted by aliphatic side-chains. By considering the aliphatic chain field effect some data have been obtained as to the conformation of the N—C_α bond, and the consequences for polypeptide conformation, especially poly-L-valine, discussed.[461]

Proteins.—Small conformational changes in a molecule of sperm whale myoglobin in its native state for different pH values at room temperature, as well as during heat denaturation in alkaline media at different stages of unfolding of the globular protein, were observed by using far-i.r. spectroscopy in the region 30—600 cm^{-1}. The changes appeared in the bands near 420 and 470 cm^{-1}, and these were ascribed to the side-chain vibrations of helical segments of the myoglobin molecule.[462]

I.r. and c.d. spectra have been obtained for myoglobin, lysozyme, RNase and serum albumin in D_2O–Tris buffer (pH 7.5). The protein concentration was 30 mg ml^{-1} and the two spectra were recorded at 5825—6250 μm and 210—240 nm, respectively. A mathematical procedure has been developed for the computation of the basic c.d. spectra and of the three main polypeptide conformations. The basic i.r. spectra were also used to compute the independent conformational analysis of the same proteins. These determinations showed a closer fit to the protein conformational determination by *X*-ray structure analysis than did the basic c.d. spectra of poly-L-lysine. For the conformational analysis of the membrane-bound protein the modified spectra allowed computation which eliminated the optical artifacts arising from the particulate nature of the materials studied. The mathematical analysis of the i.r. spectra of proteins was valuable as an independent method for the determination of the percentage β-structure in proteins, especially if the analysis of the c.d. was obscured by optical artifacts.[463]

Analysis of i.r. spectra from solutions of human serum albumin in D_2O within the range 1550—1800 cm^{-1} at different pD's and temperatures has shown that the thermostability of the serum albumin solutions as a function of pD has a maximum at the isoelectric point (pD = 5.3: T_m = 71 °C) and decreased for pD < 5.2 and pD > 7. The decrease in thermostability of serum albumin with an increase of total charge is in agreement with the statistical theory of helix–coil transitions in synthetic polypeptides.[464] The unpolarized i.r. spectrum of partially deuteriated collagen from bovine achilles tendon with the primary hydration sheath intact showed an unexpected variation in the intensity of the N—D stretching

[460] S. Krimm and Y. Abe, *Proc. Nat. Acad. Sci. U.S.A.*, 1972, **69**, 2788.
[461] F. Fillaux and C. de Loze, *Biopolymers*, 1972, **11**, 2063.
[462] Y. Chirgadze and A. Ovsepyan, *Biopolymers*, 1973, **12**, 637.
[463] R. Grosse, J. Malur, W. Meiske, J. G. Seich, and K. Tepka, *Acta Biol. Med.*, 1972, **29**, 777.
[464] M. A. Semenov and V. Maleev, *Dopovidi Akad. Nauk Ukrain. R.S.R.* (*B*), 1972, **34**, 1095.

band upon rotation of the fibre direction. This suggests that the intrachain hydrogen bonds in the triple helical structure of collagen are of unequal length, with those of similar strength lying in the same plane.[465]

In order to examine the binding of water to proteins using i.r. spectroscopy a new type of cell has been developed, enabling better sampling methods to be used. The amount of protein and bound water (measured in gm cm^{-2}) in the optical path can be varied over a considerable range and the amount of each can be determined independently, by appropriate weighings. When the method was applied to a film consisting of 18.4 mg of gelatin per cm^2, bands due to the presence of bound water were found near 1.45, 1.6, 1.8, 1.95, 2.1, 2.25, and 2.5 μm. A comparison was made of these with other reported results, and the dependence of the resolution of these bands upon the percentage water present, including the possible forms of water which some of these bands represent, were discussed.[466]

A new fibre-mounting technique for i.r. spectroscopy and applied to Nylon 66 has been developed. A single filament is wound between two rock-salt windows which are separated by a spacer. The arrangement resulted in a layer of adjacent parallel fibres and good orientation was observed in the fibres. A high level of *A* conformation orientation was present in the fibres, and structural changes that occur when the nylon fibres were shrunk was also observed.[467] Stretched and non-stretched samples of deuteriated and non-deuteriated Nylon 6 films were studied between 1200 and 500 cm^{-1} using polarized i.r. light. The spectra of three films crystallized in the α-structure were given. Measuring the intensity of the 930 cm^{-1} band was insufficient for the determination of the crystallinity. No indication of β-structure was found.[468]

8 Mössbauer Spectroscopy

Contributed by C. E. Johnson

Mössbauer spectroscopy continues to provide valuable data on biological molecules. Recent applications include the study of iron in haem proteins, iron–sulphur proteins, and iron storage and transport proteins, and of iodine in hormones.

The main source of information in searching the literature for this report has been 'The Index of Publications in Mössbauer Spectroscopy of Biological Materials', by Dr. Leopold May, Department of Chemistry, The Catholic University of America, Washington D.C., 20017, U.S.A. Dr. May kindly made available an advance copy of the latest additions to the Index.

Two particularly interesting developments in 1973 have been (i) the study of an enzyme system, putidaredoxin and cytochrome P450,[469] and

[465] N. Roberts, *J.C.S. Faraday II*, 1973, **69**, 1084.
[466] N. Ressler, *Biochim. Biophys. Acta*, 1973, **295**, 30.
[467] G. A. Tirpak and J. P. Sibilia, *J. Appl. Polymer. Sci.*, 1973, **17**, 643.
[468] P. Simak, *Angew. Makromol. Chem.*, 1973, **28**, 75.
[469] P. G. Debrunner, in 'Perspectives in Mössbauer Spectroscopy', ed. S. G. Cohen and and M. Pasternak. Plenum, New York, 1973, p. 89.

(ii) the determination of different iron sites in conalbumin by selectively binding iron to each of them in turn.[470]

The measurements on the putidaredoxin–cytochrome P450 enzyme system have been described by Debrunner.[469] Cytochrome P450 is a haem protein which binds molecular oxygen, and inserts one atom of it into a specific enzyme, putidaredoxin, a two-iron iron–sulphur protein. This particular reaction mechanism occurs in the hydroxylation of camphor in the bacterium *Pseudomonas putida*, and it can be considered as a model for similar enzyme systems which are found in mammalian cells. The reaction was studied by taking Mössbauer spectra of the reaction intermediates which were frozen at various stages of the cycle.

Haem Proteins.—A review of the interpretation of Mössbauer isomer shifts and quadrupole splittings in ferrous haemoglobins has been given by Trautwein.[471] The study of the ferrous haem compounds, deoxyhaemoglobin (Hb), oxyhaemoglobin (HbO_2), and haemoglobin carbon monoxide (HbCO), was one of the first applications of Mössbauer spectroscopy to biological molecules, and experimental and theoretical work on them is continuing.

The axis of the electric field gradient in deoxygenated myoglobin (Mb) has been deduced[472] from measurements of the relative intensities of the two quadrupole-split lines in the Mössbauer spectra of single crystals as a function of the γ-ray direction. The single crystals were enriched in ^{57}Fe and had a diameter of about 0.5 mm. V_{zz} was found to be positive in sign and is oriented along one of the four Fe—N directions in the haem group.

Mössbauer spectra of dehydrated α- and β-sub-units of deoxyhaemoglobin have been measured.[473] The data for each sub-unit show a superposition of a ferrous high-spin ($S = 2$) and a ferrous low-spin ($S = 0$) quadrupole doublet. However, the spectrum of anhydrous haemoglobin is not a direct sum of the separate α- and β-spectra, and this is interpreted as showing that the interaction between different sub-units in the Hb tetramer ($\alpha_2\beta_2$) results in structural rearrangements near the haem groups which change the relative stability of the low- and high-spin spectra upon dehydration. Since each of the isolated chains shows considerable spin-state mixing, the model previously proposed[474] for anhydrous haemoglobin with the α- and β-chains in different spin states is suggested to be unlikely.

Sickle-cell haemoglobin (Hb-S) when deoxygenated and anhydrous (lyophilized) has been compared with normal haemoglobin (Hb-A) using

[470] P. Aisen, G. Lang, and R. C. Woodworth, *J. Biol. Chem.*, 1973, **248**, 649.
[471] A. Trautwein, in ref. 469, p. 101.
[472] A. Trautwein, Y. Maeda, V. Gonser, F. Parak, and H. Formanek, Proceedings of the 5th International Conference on Mössbauer Spectroscopy, Czechoslovak Academy of Science, Bratislava, Czechoslavakia, 1973.
[473] G. Papaefthymiou, B. H. Huynh, Y. W. Chow, J. L. Groves, C. S. Yen, and C. S. Wu, *Bull. Amer. Phys. Soc.*, 1973, **18**, 671.
[474] A. Trautwein, H. Eicher, and A. Mayer, *J. Chem. Phys.*, 1970, **52**, 2473.

Mössbauer spectroscopy.[475] The main features of the spectrum are the same in the two cases.

Native catalase prepared from the bacterium *Micrococcus lysodeikicus* grown in an ^{57}Fe-enriched culture medium has been studied by Maeda *et al.*[476] Catalase is a haem protein, important in the catalysis of reactions of peroxide compounds. At low temperatures the spectrum consists of a quadrupole doublet (ΔE_Q = 0.9 mm s^{-1}) and a broad magnetic hyperfine structure pattern; these are both characteristic of high-spin ferric irons. By applying a small magnetic field, the magnetic pattern exhibits six well-defined lines with an effective field of 293 kG, and a quadrupole perturbation of (−)0.45 mm s^{-1}, while the quadrupole doublet is not influenced. The saturated value of the hyperfine field is 488 kG, using the values D = 12 cm^{-1} and E = 0.3 cm^{-1} from magnetic susceptibility data. At high temperatures the magnetic spectrum collapses. The areas of the magnetic and quadrupole spectra were equal within experimental error. The results suggest that the four haem groups in bacterial catalase show pair-wise different interactions, probably owing to different Fe—Fe interatomic distances.

The haem protein cytochrome P450$_{cam}$ from *Pseudomonas putida* was investigated by Mössbauer spectroscopy in both oxidized and reduced states by Sharrock *et al.*[477] The oxidized enzyme in the presence of the substrate, camphor, contains a mixture of high-spin (S = $\frac{5}{2}$) and low-spin (S = $\frac{1}{2}$) ferric haem sites. The high-spin fraction increases as temperature is raised. Removal of camphor results in a conversion from high spin to low spin. The components of the hyperfine field tensor for high-spin iron are −448, −224 and −101 kG along the x, y, and z directions, respectively, of the electric field gradient tensor. For low-spin iron the hyperfine fields are 120, 54, and −275 kG. Anaerobic reduction of P450$_{cam}$ in camphor solution produces a high-spin ferrous (S = 2) state. Exposure of this preparation to oxygen results in a new complex whose Mössbauer spectra are similar to those observed for oxygenated haemoglobin. Both reduced proteins show large quadrupole splittings and only moderate isomer shifts relative to iron metal; no paramagnetic effects are observed even in large applied magnetic fields. Such spectra appear to be characteristic of the haem group with an O_2 molecule as one axial ligand. P450$_{cam}$ also forms a stable adduct with carbon monoxide. The Mössbauer spectra of this complex are very similar to those of haemoglobin carbon monoxide.

The haem protein chloroperoxidase, which is isolated from the mould *Caldariomyces fumago*, has been investigated by Champion *et al.*[478] using

[475] C. S. Yen, J. L. Groves, G. Papaefthymiou, B. H. Huynh, P. H. Swerdlow, and C. S. Wu, *Bull. Amer. Phys. Soc.*, 1973, **18**, 670.

[476] Y. Maeda, A. Trautwein, U. Gonser, K. Yoshida, K. Kichuchi-Torii, T. Homma, and Y. Ogura, *Biochim. Biophys. Acta*, 1973, **303**, 230.

[477] M. Sharrock, E. Münck, P. G. Debrunner, V. Marshall, J. D. Lipscomb, and I. C. Gunsalus, *Biochemistry*, 1973, **12**, 258.

[478] P. M. Champion, E. Münck, P. G. Debrunner, P. F. Hollenberg, and L. P. Hager, *Biochemistry*, 1973, **12**, 426.

Mössbauer spectroscopy. The haem iron of native chloroperoxidase is in a low-spin ferric state at low temperatures and undergoes a temperature-dependent spin transition to high-spin ferric around 200 K. The low-temperature Mössbauer spectra could be described by a model where the haem iron resides in a ligand field potential of orthorhombic symmetry. The low-temperature Mössbauer spectra of the chloroperoxidase–Cl complex are quite similar to those of the native enzyme, suggesting that chloride does not bind as an axial ligand to the haem iron. The complexes of chloroperoxidase with iodide and fluoride are high-spin ferric at all temperatures; both complexes reveal unusually large rhombic distortions at the haem iron. The ferrous form of chloroperoxidase is a high-spin species with an optical absorption spectrum and Mössbauer parameters which are almost identical with those of the reduced form of cytochrome $P450_{cam}$. Like the cytochromes of the P450 type, reduced chloroperoxidase forms a stable complex with carbon monoxide, characterized by a Soret band at an unusually long wavelength (443 nm). The haem iron of this complex is found to be in a low-spin ferrous state. This investigation suggests close structural similarities between the active sites of chloroperoxidase and $P450_{cam}$.

Iron–Sulphur Proteins.—Previous work has shown that the active centre in reduced two-iron ferredoxins contains one Fe^{3+} (spin $S_1 = \frac{5}{2}$) and one Fe^{2+} (spin $S_2 = 2$) atom coupled antiferromagnetically together to give a total spin $S = \frac{1}{2}$. The effective fields at the ^{57}Fe nuclei are *ca.* -185 and $+120$ kG, respectively. These rather small fields have been shown [479, 480] to be explained as a result of the antiferromagnetic coupling. The effective fields at the nuclei in each atom are proportional to the projection of the spins of each atom on the resultant spin S, *i.e.*

$$B_{\text{eff}}(Fe^{3+}) = \frac{\langle \mathbf{S}_1 . \mathbf{S} \rangle}{S(S+1)} B_{\text{hf}}(Fe^{3+})$$

$$B_{\text{eff}}(Fe^{2+}) = \frac{\langle \mathbf{S}_2 . \mathbf{S} \rangle}{S(S+1)} B_{\text{hf}}(Fe^{2+})$$

where the hyperfine fields B_{hf} were assumed to be the same as those found in oxidized (Fe^{3+}) and reduced (Fe^{2+}) forms of rubredoxin. In the latter proteins there is only one iron atom in the molecule and it is tetrahedrally co-ordinated to four sulphurs, which is believed to be the same environment as in ferredoxins.

Transferrin.—Transferrin, the iron transport protein of the extracellular fluid of vertebrates, contains two metal-binding sites per protein molecule. Each site will reversibly bind one Fe^{3+} atom. Despite numerous investigations (e.p.r., ^1H n.m.r., optical spectroscopy, *etc.*) the chemical nature and

[479] C. E. Johnson, in ref. 469, p. 79.
[480] C. E. Johnson, *J. de Phys.*, 1974, **35**, C1-57.

electronic character of the sites remain obscure. There is disagreement as to whether the two sites are identical or not, and whether there is any interaction between the metal ions bound at the two sites. Several groups are now working on these proteins using Mössbauer spectroscopy.

Mössbauer spectra of ^{57}Fe in transferrin from humans and from rabbits have been studied by Tsang *et al.*,[481] at several magnetic fields and temperatures. The two iron sites appear to be closely similar. Well-resolved magnetic hyperfine spectra were observed at 4.2 K and 77 K in zero applied magnetic field. Thus the electric relaxation times are long and spin–spin interaction is negligible, so it is suggested that the two iron atoms of that molecule are at least 20 Å apart. The high-spin ferric iron ($S = \frac{5}{2}$) is shown to be in a rhombic environment, and can be described by a spin Hamiltonian

$$\mathscr{H} = D\{[S_z^2 - \tfrac{1}{3}S(S+1)] + \lambda(S_x^2 - S_y^2)\}$$

where the data give $D = 0.12\ \text{cm}^{-1}$ and $\lambda = 0.25$. The nuclear quadrupole interaction $\frac{1}{4}QV_{zz}$ was found to be $0.09\ \text{mm s}^{-1}$. No difference was found between spectra of transferrin from humans and rabbits, or with different levels of saturation with iron.

Spartalian and Oosterhuis [482] have measured the ^{57}Fe Mössbauer spectra of transferrin from human serum and, like Tsang *et al.*, detected no difference between the iron sites. The spectra were analysed in terms of a spin Hamiltonian with zero field splitting of the electronic ground-state given by $D = 0.15\ \text{cm}^{-1}$ and $E/D\ (= \lambda) = 0.31$, which are close to the values found by Tsang *et al.* The magnetic hyperfine coupling was observed to be anisotropic and measured to be $A_x : A_y : A_z = 221 : 225 : 213$ kG per unit spin, respectively. The quadrupole splitting $\Delta E_Q = 0.75\ \text{mm s}^{-1}$. The effect of ligand (probably ^{14}N) nuclear moments was observed in the zero field spectra. Mössbauer spectra were also taken with ^{151}Eu bound to the protein, but no hyperfine splitting was observed, because the ground state of the Eu^{3+} ion is a singlet ($J = 0$).

Conalbumin from hen egg-white (ovotransferrin) and transferrin from humans have been studied by Aisen *et al.*[470] using Mössbauer and e.p.r. spectroscopy. They used the technique of isoelectric focusing which enabled them to isolate the protein molecules where only one site has a bound iron atom and hence to determine the contribution of each site to the spectra of the protein. Occupancy of the first (or 'inner') binding site by Fe^{3+} produces spectra which are readily distinguishable from those shown by the second ('outer') site. Summation of the spectra from the two sites gave the spectrum observed from diferric conalbumin. It is concluded that the specific binding sites of conalbumin are not identical, and that the protein does not bind metal irons in a simple random fashion.

[481] C. P. Tsang, A. J. F. Boyle, and E. H. Morgan, *Biochim. Biophys. Acta*, 1973, **328**, 84.
[482] K. Spartalian and W. T. Oosterhuis, *J. Chem. Phys.*, 1973, **59**, 617.

Ferritin.—Ferritins are iron storage proteins. The ferritin from the single-cell plant-like fungus *Phycomyces blakesleanii* has been studied by Spartalian *et al.*[483] Mössbauer spectra from different parts of the cells have been observed, and indicated that the iron forms superparamagnetic clusters within the ferritin molecules. Horse ferritin has been similarly studied. The size of the iron clusters may be determined from the temperature at which a six-line hyperfine structure pattern becomes resolved. The clusters were found to be smaller in *Phycomyces* ferritin than in horse ferritin, and to increase as the cell matures. The results also suggest that *Phycomyces* ferritin near the top of the cell where the spores are located seems to be more fully saturated with iron than the ferritin near the bottom.

Iron storage and transport proteins have been reviewed.[484]

Other Proteins.—Atkin *et al.*[485] have measured the Mössbauer effect of ^{57}Fe in protein B2, one of the sub-units of ribonucleotide reductase from *E. coli*. This protein contains two iron atoms per molecule, and the optical spectra are in many ways similar to those of oxy- and met-haemerythrin. Below 195 K the iron was completely diamagnetic. The Mössbauer data suggest that protein B2 contains two non-identical high-spin Fe^{3+} irons in an antiferromagnetically coupled dinuclear complex that resembles both oxyhaemerythrin and methydroxohaemerythrin. The results throw light on the function of iron in protein B2.

Hormones.—Groves *et al.*[486] have used the ^{129}I Mössbauer effect to study iodine bonding in some iodine-containing hormones. They have made measurements on L-thyroxine (T_4) which is synthesized by the thyroid gland and contains four iodine atoms, and on L-3,5-di-iodotyrosine which is believed to be a precursor for T_4 and contains two iodine atoms. Identical Mössbauer spectra were found for the two compounds, and the iodine sites in each compound were indistinguishable by the Mössbauer effect. The observed quadrupole splittings and isomer shifts can be explained by a small amount of *s* hybridization in the 5*p* orbitals.

[483] K. Spartalian, N. Smarra, and W. T. Oosterhuis, 19th Annual Conference on Magnetism and Magnetic Materials, Boston, 1973.
[484] K. Spartalian, W. T. Oosterhuis, and B. Window, *Mössbauer Effect Methodol.*, 1973, **8**, 137.
[485] C. L. Atkin, L. Thalander, P. Reichard, and G. Lang, *J. Biol. Chem.*, 1973, **248**, 7464.
[486] J. L. Groves, M. J. Potasek, and G. Depasquali, *Phys. Letters* (*A*), 1973, **42**, 493.

3
Peptide Synthesis

BY B. RIDGE AND J. H. JONES

PART I: Methods *by B. Ridge*

1 Introduction

The general arrangement of the Peptide Synthesis section of this Report follows that used in previous volumes, except that the sections on 'Methods' and 'Results' have been separated. Thus Part I deals with methods of peptide synthesis and Part II contains the results of peptide synthesis and also the valuable compilations of 'Syntheses Reported during 1973' (Appendix A), and 'Some Useful New Synthetic Intermediates Described during 1973' (Appendix B).

The year under review has seen the appearance of the second edition of Jakubke and Jeschkeit's excellent book 'Aminosäuren, Peptide, Proteine – Eine Einführung'.[1] This work contains a useful hundred-page account of the methods of peptide synthesis, similar to the first edition, but updated with relevant new material. A plethora of conference proceedings [2–5] has been published; the editors of certain of these should be congratulated on the rapidity with which they steered their works through the publication maze. A companion [6] to these Reports has appeared, which looks at the literature in two-year quanta rather than the quanta of one year used here. In some ways the publications are complementary. A number of general

[1] H.-D. Jakubke and H. Jeschkeit, 'Aminosäuren, Peptide, Proteine – Eine Einführung', 2., verbesserte Auflage, Akademie-Verlag, Berlin, 1973.

[2] Proceedings of the Second American Peptide Symposium held at the Cleveland Clinic, 1970: 'Progress in Peptide Research', Vol. II, ed. S. Lande, Gordon and Breach, New York, 1972.

[3] Proceedings of the Eleventh European Peptide Symposium held at Vienna, Austria, 1971: 'Peptides 1971', ed. H. Nesvadba, North-Holland Publishing Co., Amsterdam, 1973.

[4] Proceedings of the Third American Peptide Symposium held at Boston, Massachusetts, 1972: 'Chemistry and Biology of Peptides', ed. J. Meienhofer, Ann Arbor Science Publishers, Ann Arbor, Michigan, 1972.

[5] Proceedings of the Twelfth European Peptide Symposium held at Schloss Reinhardsbrunn, German Democratic Republic, 1972: 'Peptides 1972', ed. H. Hanson and H.-D. Jakubke, North-Holland Publishing Co., Amsterdam and London, 1973.

[6] MTP International Review of Science, Organic Chemistry Series One, Volume 6, 'Amino Acids, Peptides and Related Compounds', ed. D. H. Hey and D. I. John, Butterworth, London, and University Park Press, Baltimore, 1973, (*a*) C. H. Stammer, p. 135; (*b*) R. Wade, p. 161.

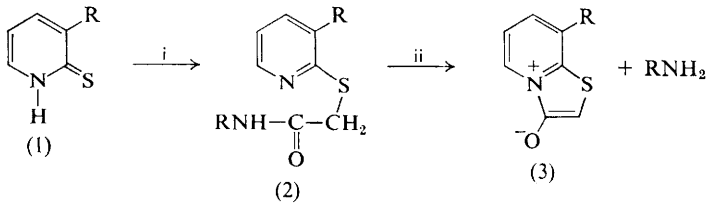
reviews [7–11] of the field have been published, together with a specific review [12] of some of the problems encountered in the synthesis of biologically active peptides, and some stimulating 'Guesses and Pious Hopes'.[13, 14]

2 Protective Groups

A book has been published concerning the chemistry of protecting groups [15] which contains a chapter on the protection of N—H bonds.[15a]

Established Methods of Amino-group Protection.—A review has been printed giving the fascinating story of the turn of events which led to the development of the t-butoxycarbonyl amino-protective group, base-sensitive protective groups, and solvent-sensitive protective groups.[16]

The *N*-chloroacetyl group can be removed by a novel two-stage process involving treatment of the chloroacetyl derivative with pyridine-2-thione (1; R = H) to yield an *N*-(2-pyridylthio)acetyl derivative (2; R = H), which can be cyclized in cold trifluoroacetic acid to a thiazolo[3,2-*a*]-pyridinium-3-olate (3; R = H), with concomitant liberation of the free amine (Scheme 1).[17] The most effective cleavage reagent of a series of



Reagents: i, $ClCH_2CONHR$, aq. $NaHCO_3$ at 40 °C; ii, CF_3CO_2H at 20 °C, **1—4 h**

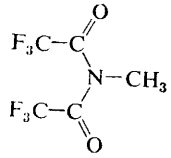
Scheme 1

thiolactams is 3-nitropyridine-2-thione (1; R = NO_2), partly on account of the crystallinity of the corresponding thioacetyl derivatives (2; R = NO_2), and partly on account of the convenient cyclization time. The procedure has been applied to a number of amino-acid and simple

[7] M. Oya, *Sen'i To Kogyo*, 1972, **5**, 489 (*Chem. Abs.*, 1973, **78**, 84 772).
[8] N. Yanaihara, *Farumashia*, 1972, **8**, 721 (*Chem. Abs.*, 1973, **78**, 136 632).
[9] J. Meienhofer, *Chem. Technol.*, 1973, **3**, 242 (*Chem. Abs.*, 1973, **79**, 5538).
[10] N. Yanaihara, *Seibutsu Butsuri*, 1972, **12**, 195 (*Chem. Abs.*, 1973, **79**, 53 752).
[11] M. Waki and N. Izumiya, *Kobunshi*, 1973, **22**, 12 (*Chem. Abs.*, 1974, **80**, 3776).
[12] Th. Wieland, ref. 5, p. 38.
[13] R. Hirschmann, ref. 4, p. 721.
[14] J. Rudinger, ref. 4, p. 729.
[15] 'Protecting Groups in Organic Chemistry', ed. J. F. W. McOmie, Plenum Press, London, 1973, (*a*) J. W. Barton, p. 43; (*b*) C. B. Reese, p. 95; (*c*) E. Haslam, p. 145; (*d*) E. Haslam, p. 183; (*e*) R. G. Hiskey, V. R. Rao, and W. G. Rhodes, p. 235; (*f*) J. F. W. McOmie, p. 403.
[16] L. A. Carpino, *Accounts Chem. Res.*, 1973, **6**, 191.
[17] K. Undheim and P. E. Fjeldstad, *J.C.S. Perkin I*, 1973, 829.

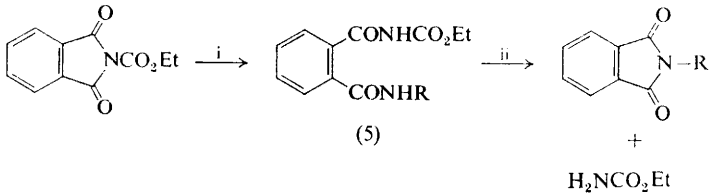
dipeptide derivatives, without racemization. A kinetic study of the cyclization of a series of *para*-substituted *N*-(3-nitro-2-pyridylthioacetyl)anilines showed that protonation of the amide group is the rate-determining factor. The method should prove of value for the selective removal of the *N*-chloroacetyl group only in the presence of groups which are stable to prolonged treatment with trifluoroacetic acid.

Trifluoroacetylation of amino-, hydroxy-, and thiol groups can be accomplished under mild non-acidic conditions using the reagent (4).[18]



(4)

N-Phthaloylation of 4-chloro- and 4-hydroxy-2-amino-esters with *N*-ethoxycarbonylphthalimide, under aqueous conditions, proceeds with difficulty, owing to the intervention of displacement and amide–hydroxy neighbouring-group reactions. These problems can be circumvented by using an aprotic solvent, such as dimethyl sulphoxide, containing triethylamine, as the reaction medium.[19] Confirmation of the mechanism proposed by Nefkens and colleagues[20] for the aminolysis of *N*-ethoxycarbonylphthalimide has been obtained by isolation of crystalline intermediate diamides (5) which cyclize readily in the presence of triethylamine (Scheme 2).[19] Further details have appeared concerning the use of



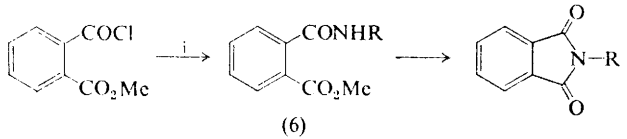
(5)

Reagents: i, RNH_2 [R = Bu^n, CH_2Ph, or $CH(CO_2Me)CH_2CH_2Cl$]; ii, NEt_3

Scheme 2

o-methoxycarbonylbenzoyl chloride in anhydrous medium for the introduction of *N*-phthaloyl groups into amino-acids and amino-esters.[21] In the case of some sterically hindered amino-acids, mixtures of *o*-methoxycarbonylbenzamides (6) and phthaloyl derivatives are produced (Scheme 3).

[18] M. Donike, *J. Chromatog.*, 1973, **78**, 273.
[19] S. Clarke, R. C. Hider, and D. I. John, *J.C.S. Perkin I*, 1973, 230.
[20] G. H. Nefkens, G. I. Tesser, and R. J. Nivard, *Rec. Trav. chim.*, 1960, **79**, 688.
[21] D. N. Reinhoudt, T. S. Lie, J. J. Gunneweg, and H. C. Beyerman, *Rec. Trav. chim.*, 1973, **92**, 819.



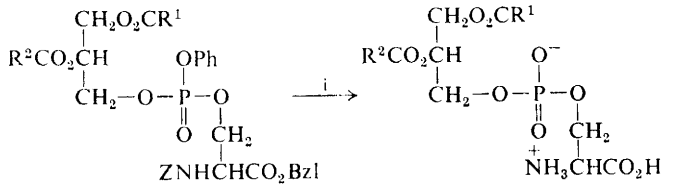
(6)

Reagents: i, RNH_2 [R = $CH(CO_2Bzl)CHMeEt$, or $CMe_2(CO_2Bzl)$] [in the latter case (6) is the sole product]

Scheme 3

The constant pH, controlled-potential electrolytic detosylation of *N*-tosylamino-acids and peptides has been reported in full [22] (see Vol. 4 of these Reports, p. 316). The electrolytic cell, described in detail in the paper, consists of a lead cathode, a carbon anode, and a cellular membrane separating anode and cathode compartments; the electrolysis is carried out at a current density of 3.2 A dm^{-2}. The catholyte, which consists of the substrate in 20% aqueous methanol containing two equivalents of sodium or potassium chloride, is maintained at pH 11 by addition of caustic soda solution from an autotitrator; the anolyte (20% hydrochloric acid) is replaced continuously. The method was tested using a range of *N*-tosylamino-acids (including *N*-tosyl-L-serine) and small peptides (*e.g.* *N*-tosyl-*S*-benzylglutathione); very good yields of optically pure materials were regenerated. In addition it was shown that selective cleavage of *N*-tosyl groups in the presence of *N*-benzyloxycarbonyl, t-amyloxycarbonyl, t-butoxycarbonyl, and *S*-benzyl groups is possible, although *N*-benzoyl groups are attacked.

In quite a different context, *viz.* the synthesis of unsaturated phosphatidylserines, *N*-benzyloxycarbonyl and benzyl ester groups have been removed by electrochemical reduction at a mercury cathode (graphite anode) using 95% dimethylformamide containing tetraethylammonium iodide as the electrolyte (Scheme 4).[23] The progress of the reduction was



Reagents: i, electrolysis in 95% Me_2NCHO containing $\overset{+}{N}Et_4\ \bar{I}$ at −2.87 V relative to the standard calomel electrode ($R^1 = C_{15}H_{31}$, $R^2 = C_{17}H_{33}$)

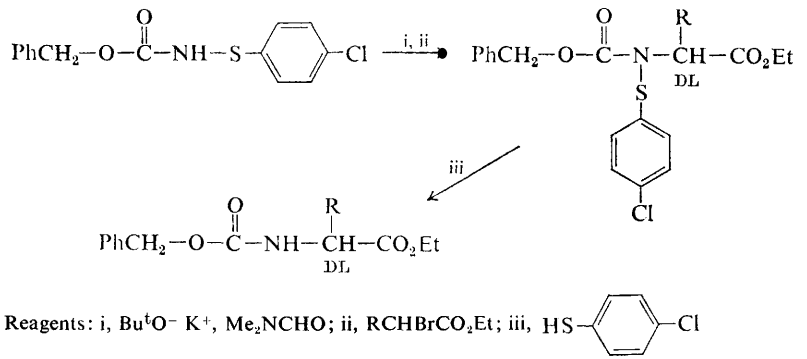
Scheme 4

[22] T. Iwasaki, K. Matsumoto, M. Matsuoka, T. Takahashi, and K. Okumura, *Bull. Chem. Soc. Japan*, 1973, **46**, 852.

[23] M. A. Kabanova, N. E. Suleimanyan, N. F. Loginova, V. G. Mairanovskii, V. I. Shvets, and R. P. Evstigneeva, *J. Gen. Chem.* (*U.S.S.R.*), 1973, **43**, 189.

conveniently followed by examining the behaviour of aliquots of reaction mixture in an analytical polarograph.

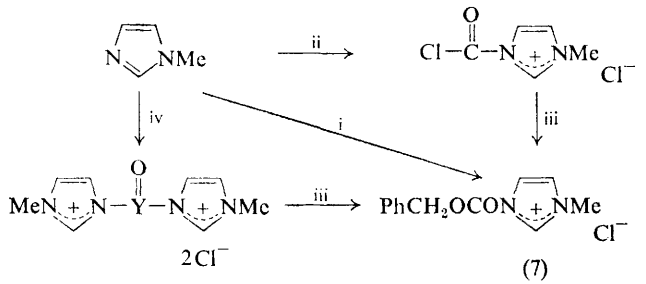
The chemistry of phosgene has been reviewed.[24] Racemic *N*-benzyloxycarbonyl-amino-acids have been prepared in moderate yield by the reaction of the potassium salt of benzyl *N*-(*p*-chlorophenylthio)carbamate with an α-halogeno-ester (Scheme 5).[25]



Reagents: i, Bu^tO^- K^+, Me_2NCHO; ii, $RCHBrCO_2Et$; iii, $HS-C_6H_4-Cl$

Scheme 5

N^1-Benzyloxycarbonyl-N^3-methylimidazolium chloride (7) is a stable water-soluble reagent which can be used for the introduction of the *N*-benzyloxycarbonyl group.[26] The reagent (7) can be prepared by the methods outlined in Scheme 6. The acylation reaction occurs rapidly in



Reagents: i, $PhCH_2OCOCl$; ii, $COCl_2$ (1 equivalent); iii, $PhCH_2OH$; iv, $YOCl_2$ (half an equivalent, where Y = S or C)

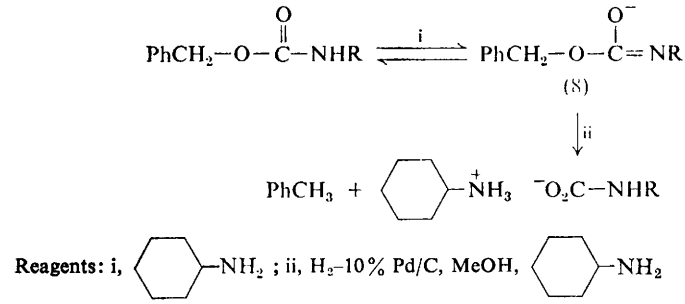
Scheme 6

neutral aqueous solution at room temperature and shows selectivity (see later). The method is likely to find application only for the benzyloxycarbonylation of substances sensitive to the conditions of the conventional

[24] H. Babad and A. G. Zeiler, *Chem. Rev.*, 1973, **73**, 75.
[25] T. Taguchi and T. Mukaiyama, *Chem. Letters*, 1973, 1 (*Chem. Abs.*, 1973, **78**, 97 987)
[26] E. Guibé-Jampel, G. Bram, and M. Vilkas, *Bull. Soc. chim. France*, 1973, 1021.

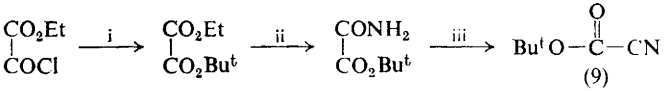
Schotten–Baumann acylation and for selective acylation required in partial synthesis.

Catalytic hydrogenolysis of *N*-benzyloxycarbonyl-amino-acids occurs four to five times faster in a basic medium [27] (methanol containing excess cyclohexylamine or triethylamine) than in the corresponding neutral solvent or acidic medium. This effect is attributed to the probable presence of appreciable amounts of the deprotonated form (8) of the substrate, which undergoes hydrogenation more readily than the uncharged form (Scheme 7).



Scheme 7

Acylation of amino-acids with t-butoxycarbonyl chloride proceeds readily in dioxan–water mixtures at 0 °C at a carefully controlled pH corresponding to the p*K* of the amino-group.[28] Carpino [29] had shown that t-butyl cyanoformate (9) (prepared as shown in Scheme 8) is a more reactive

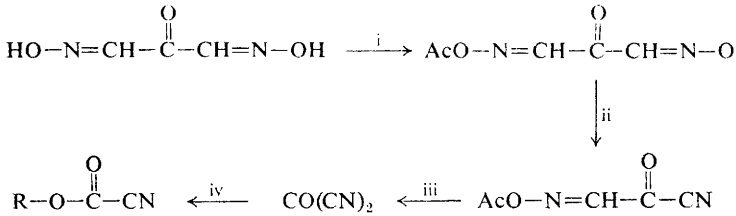


Reagents: i, Bu^tOH, C_5H_5N; ii, NH_3; iii, $(CF_3CO)_2O$, C_5H_5N

Scheme 8

acylating agent (towards simple amines) than t-butyl azidoformate, but that the crude acylated products are less pure. The cyanoformate (9) can also be prepared from t-butyl alcohol and carbonyl cyanide [30] (see Scheme 9; R = Bu^t), although the preparation of carbonyl cyanide is laborious. The reagent (9) can be used for the acylation of amino-acid benzyltrimethylammonium salts in dimethylformamide, very good yields being obtained. t-Amyl cyanoformate (10) can be prepared similarly (Scheme 9; R = CMe_2Et) and used for preparing t-amyloxycarbonyl-amino-acid derivatives.[30]

[27] H. Medzihradszky-Schweiger, *Acta Chim. Acad. Sci. Hung.*, 1973, **76**, 437.
[28] H. Klengel, K. J. Schumacher, and G. Losse, *Z. Chem.*, 1973, **13**, 221.
[29] L. A. Carpino, *J. Amer. Chem. Soc.*, 1960, **82**, 2725.
[30] M. T. Leplawy and J. Zabrocki, ref. 5, p. 112.

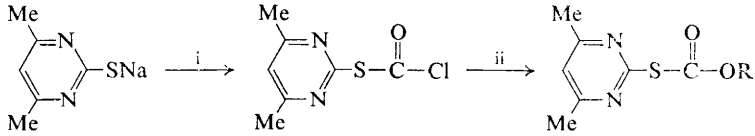


(9) R = Bu^t
(10) R = CMe_2Et
(11) R = *p*-$MeOC_6H_4CH_2$

Reagents: i, heat, Ac_2O; ii, heat, 60—130 °C, 12 mmHg; iii, heat, 150—180 °C, 650 mm Hg; iv, ROH

Scheme 9

A new stable thiocarbonate has been used for the introduction of *N*-t-butoxycarbonyl groups, *viz.* t-butyl *S*-(4,6-dimethylpyrimid-2-yl)-thiocarbonate (12; R = Bu^t) prepared according to Scheme 10. Amino-



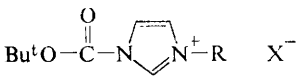
(12)

Reagents: i, $COCl_2$; ii, ROH, pyridine (R = Bu^t or *p*-$MeOC_6H_4CH_2$)

Scheme 10

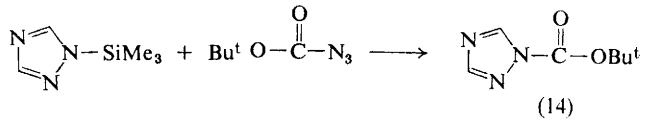
acids react smoothly with (12) in aqueous dioxan or aqueous dimethylformamide, containing triethylamine, to give excellent yields of acylated derivatives; a wide range of derivatives are reported.[31]

Other highly reactive t-butoxycarbonylating reagents are water-soluble 1-t-butoxycarbonyl-3-alkylimidazolium salts (13)[32] and N^1-t-butoxycarbonyl-1,2,4-triazole (14)[33] prepared as shown in Scheme 11. An observation of considerable interest is that acylation of amino-acid



(13) a; R = Et, X = BF_4
b; R = Me, X = ClO_4

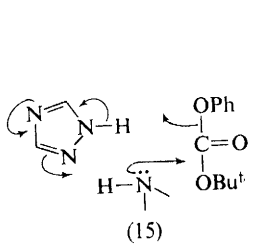
[31] T. Nagasawa, K. Kuroiwa, K. Narita, and Y. Isowa, *Bull. Chem. Soc. Japan*, 1973, **46**, 1269.
[32] E. Guibé-Jampel, G. Bram, M. Wakselman, and M. Vilkas, *Synthetic Comm.*, 1973, **3**, 111.
[33] G. Bram, *Tetrahedron Letters*, 1973, 469.



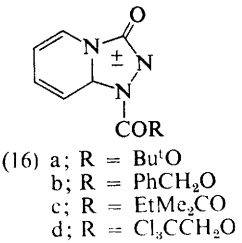
(14)

Scheme 11

benzyltrimethylammonium salts in DMF or DMSO with t-butyl phenyl carbonate proceeds in higher yield when the stoicheiometric quantity of 1,2,4-triazole is added to the reaction mixture [33] (apparently benzotriazole and α-pyridone do not show this effect). Two possible explanations have been put forward to account for the behaviour; either (14) is formed *in situ* or the triazole acts as a bifunctional catalyst [see (15)]. Mesoionic anhydro-1-acyl-3-hydroxy-*sym*-triazolo[4,5-*a*]pyridinium hydroxides (16) have been



(15)



(16) a; R = Bu^tO
b; R = $PhCH_2O$
c; R = $EtMe_2CO$
d; R = Cl_3CCH_2O

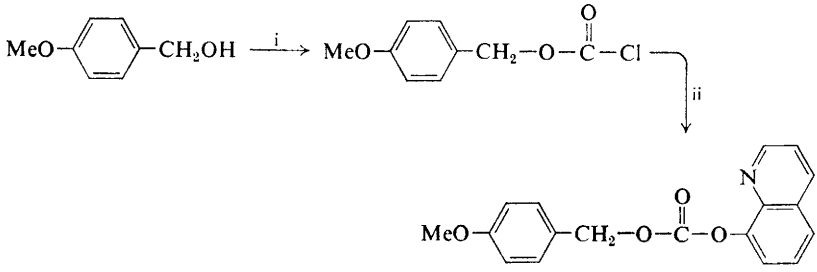
used to introduce *N*-t-butoxycarbonyl (also *N*-benzyloxycarbonyl, *N*-t-amyloxycarbonyl, and *N*-trichloroethoxycarbonyl) groups into amino-acids, penicillins, and cephalosporins.[34]

Aqueous trifluoroacetic acid (70%) is more selective than trifluoroacetic acid in dichloromethane for the cleavage of *N*-t-butoxycarbonyl groups in the presence of *N*-benzyloxycarbonyl substituents [35] (see Vol. 4 of these Reports, p. 312). However, the presence of cation scavengers during acidolytic deprotection can reduce such selectivity.[36] The rate of removal of the benzyloxycarbonyl group (largely an S_N2 process) is increased by the nucleophilic scavenger, while the rate of removal of the t-butoxycarbonyl group (S_N1 process) is decreased, presumably owing to the reduced solvent polarity. Boron trifluoride etherate in acetic acid is a particularly suitable reagent for the cleavage of *N*-t-butoxycarbonyl groups in the presence of *S*-trityl, except in the presence of hydroxy side-chains, when partial acetylation can occur [35] (see also ref. 114).

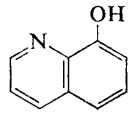
p-Methoxybenzyl cyanoformate (11), prepared as indicated in Scheme 9, can be used (in the crude state) to acylate amino-acid sodium salts in aqueous THF with good yields.[30] A series of *N*-*p*-methoxybenzyloxycarbonyl

[34] B. Shimizu and A. Saito, Jap. Kokai 73 62 702 (Cl. 16 A6) (*Chem. Abs.*, 1974, **80**, 3799).
[35] E. Schnabel, H. Klostermeyer, and H. Berndt, ref. 3, p. 69.
[36] D. F. Veber, S. F. Brady, and R. Hirschmann, ref. 4, p. 315.

derivatives of side-chain-protected trifunctional amino-acids have been prepared [37] in very good yield using *p*-methoxybenzyl 8-quinolyl carbonate (17) (Scheme 12) as acylating reagent.[38] *p*-Methoxybenzyl *S*-(4,6-di-



(17)

Reagents: i, $COCl_2$, Et_2O; ii,  , EtOAc, NEt_3

Scheme 12

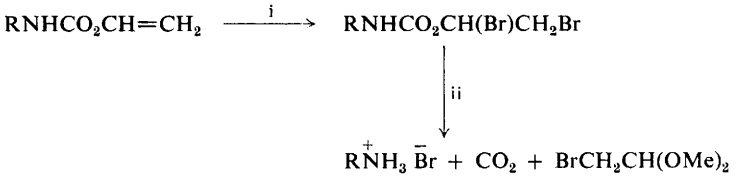
methylpyrimid-2-yl)thiocarbonate (12; R = *p*-$MeOC_6H_4CH_2$), prepared according to Scheme 10, is also a convenient reagent for preparing a wide range of these derivatives.[31]

The preparations of a range of *N*-1-(*p*-biphenylyl)-1-methylethoxycarbonylamino-acid cyclohexylammonium or dicyclohexylammonium salts (suitably side-chain-protected, where necessary, for solid-phase peptide synthesis) have been described.[39] Conversion of these derivatives into the free acids, using an extraction procedure, and removal of the *N*-protecting groups with 10% trifluoroacetic acid in dichloromethane, enabled the optical purity of the cycled amino-acids to be verified.[39]

New Methods of Amino-group Protection.—An interesting development in the field of solvent-sensitive protecting groups is the introduction of the *N*-vinyloxycarbonyl group.[40] Vinyl chloroformate, which is readily available, reacts smoothly with amino-acids in aqueous dioxan in the presence of magnesium oxide to yield *N*-vinyloxycarbonyl derivatives, some of which are oils and must be characterized as their dicyclohexylammonium salts. The only scant indication of the use of these derivatives in peptide synthesis comes from a commercial brochure.[41] The preferred

[37] H. Yajima, F. Tamura, Y. Kiso, and M. Kurobe, *Chem. and Pharm. Bull.* (*Japan*), 1973, **21**, 1380.
[38] H. Yajima, F. Tamura, and Y. Kiso, *Chem. and Pharm. Bull.* (*Japan*), 1970, **18**, 2574.
[39] R. S. Feinberg and R. B. Merrifield, *Tetrahedron*, 1972, **28**, 5865.
[40] R. A. Olofson and Y. S. Yamamoto, U.S.P. 3 711 458 (Cl. 260/112.5; C 07*cg*, C 08*h*) (*Chem. Abs.*, 1973, **78**, 98 024).
[41] PCR, Incorporated, Report October 1973.

method of cleavage of this group is treatment with bromine in dichloromethane, followed by methanol, when carbon dioxide is evolved [40] (Scheme 13). This result can be rationalized by a neighbouring-group process similar



Reagents: i, Br_2 in CH_2Cl_2; ii, MeOH

Scheme 13

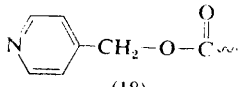
to that formulated for the 'self cleavage' of 2-bromo-1,1-dimethylethoxycarbonyl derivatives in alcoholic solution (see Vol. 4 of these Reports, p. 317). The t-butoxycarbonyl group is stable under these conditions. The vinyloxycarbonyl group can also be removed with hydrogen chloride in methanol, or by hydrolysis induced by mercury(II) ion.

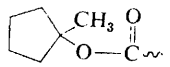
A pair of 'chemically' selective protecting groups has been devised,[36] in order to overcome the difficulties due to the loss of *N*-t-butoxycarbonyl groups during isolation procedures involving 50% aqueous acetic acid as solvent and due to the lack of selectivity in the removal of t-butoxycarbonyl groups in the presence of benzyloxycarbonyl.

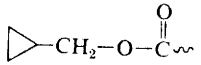
The 4-picolyloxycarbonyl (isonicotinyloxycarbonyl) group (18) (used for the protection of the ε-amino-group of lysine – see later) is quite stable in liquid hydrogen fluoride or trifluoroacetic acid (owing to protonation) and only slowly removed by hydrogen bromide in acetic acid. The group is, however, smoothly removed by zinc dust in 50% aqueous acetic acid and by catalytic hydrogenation.[36] This group confers increased solubility, both in aqueous and organic phases, of peptides containing it; furthermore, the basic 'handle' should aid purification.

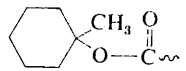
Of a series of cycloalkoxycarbonyl derivatives (19), (20), (22), and (23) [36, 42] (prepared *via* the corresponding crude chloroformates), the 1-methylcyclobutoxycarbonyl group (20) was found to be stable in 50% aqueous acetic acid and yet to be removed by trifluoroacetic acid.[36] The group is in fact removed completely from 1-methylcyclobutoxycarbonylphenylalanine in less than 30 min at 20 °C, but it is stable in 50% aqueous acetic acid for 48 h (under these latter conditions t-butoxycarbonylphenylalanine loses 10—15% of its protecting group) (note: for a given protecting group the rate of acid-catalysed cleavage from an amino-acid is about twice that from a peptide, due, presumably, to the reduced basicity of the urethane group in the protected peptide [36]). In the case of the groups (21) and (22), which were also introduced *via* the corresponding azide (24)

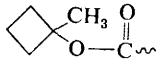
[42] H. Otsuka and K. Inouye, Ger. Offen. 2 234 797 (Cl. C 07*c*) (*Chem. Abs.*, 1973, **78**, 124 891).

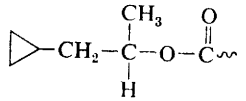

(18)


(21)

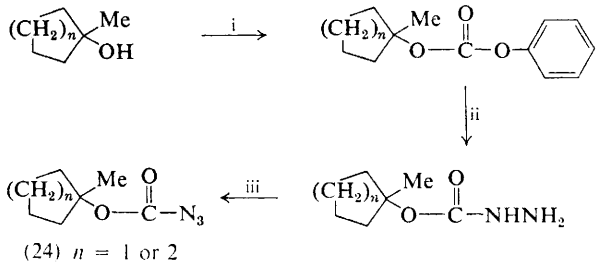

(19)


(22)


(20)


(23)

(see Scheme 14; several mixed carbonates are also reported), derivatives of (21) are more labile to hydrogen chloride in acetic acid than derivatives of (22), which have a similar lability to the t-butoxycarbonyl group. The application of these principles in the synthesis of a large peptide using N^{α}-1-methylcyclobutoxycarbonyl, N^{ϵ}-4-picolyloxycarbonyl, and *S*-acetamidomethyl protection is awaited with interest.


(24) n = 1 or 2

Reagents: i, PhO_2CCl; ii, N_2H_4; iii, HNO_2

Scheme 14

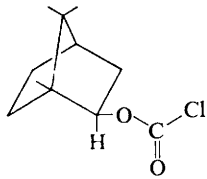
The *N*-piperidino-oxycarbonyl group can be removed selectively in the presence of t-butoxycarbonyl- or benzyloxycarbonyl-amino-groups, 4-picolyl esters, or nitro-arginine, when sodium dithionite in aqueous acetic acid or aqueous DMF is used as cleavage reagent.[43]

Further conference discussions of the versatile αα-dimethyl-3,5-dimethoxybenzyloxycarbonyl group (see Vol. 5 of these Reports, p. 264) have appeared.[44, 45] The group is easily introduced and can be removed by acidolysis (5% trifluoroacetic acid in dichloromethane) or by photochemically induced heterolysis (the *m*-methoxy-groups acting as electron

[43] G. T. Young, D. S. Stevenson, and T. Pinker, ref. 5, p. 131.
[44] Ch. Birr, F. Flor, P. Fleckenstein, and Th. Weiland, ref. 3, p. 175.
[45] Ch. Birr, ref. 5, p. 72.

donors; the oxycarbonyl group as the electron acceptor). The acid lability of this group is intermediate between that of t-butoxycarbonyl and 1-(*p*-biphenylyl)-1-methylethoxycarbonyl group.[45] The applicability of the new protecting group has been tested in a conventional stepwise synthesis (mixed anhydride method) of a linear decapeptide of [6-tyrosine]-antanamide,[45] and in solid-phase syntheses of a linear decapeptide of antanamide,[44] the linear decapeptide of [6-tyrosine]-antanamide,[45] and the *C*-terminal decapeptide of the insulin A-chain.[45] Interesting details of the continuous synthesis of peptides have been described.[45] A cold dichloromethane solution of the *N*-methylmorpholinium salt of the carboxy-component and a cold dichloromethane solution of anisoyl chloride are pumped through tubing to mix at a Y-junction. The mixture then passes through a 10 min reaction coil at 0 °C. Meanwhile the αα-dimethyl-3,5-dimethoxybenzyloxycarbonyl peptide in THF solution is continuously deprotected photochemically in a quartz spiral. The solutions of both components mix at another Y-junction and pass through a 10 min reaction coil at 20 °C. In the solid-phase variant of this method, coupling is accomplished with dicyclohexylcarbodi-imide, and deprotection with trifluoroacetic acid in dichloromethane; the absorbance at 276 or 282 nm of the filtrates from the deprotection cycle gives a quantitative measure of the amount of 3,5-dimethoxy-α-methylstyrene produced and hence of the efficiency of deprotection.[45]

Another group has independently investigated the isobornyloxycarbonyl moiety as a protective function for amino-groups (see Vol. 5 of these Reports, p. 262). In addition to the stable chloroformate (25),[46–48] a number of mixed carbonates (26)—(28) can be used to introduce the



(25) Bornoc-Cl

Bornoc-ONSu
(26)

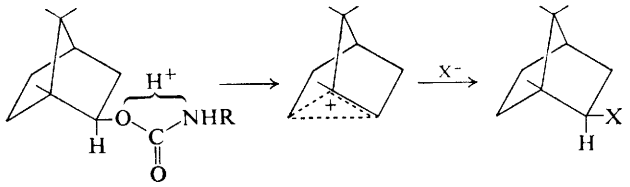
Bornoc-ONp
(27)

Bornoc-OTcp
(28)

masking group. The ready acidolysis of isobornyloxycarbonylamino-derivatives with trifluoroacetic acid or hydrogen chloride in acetic acid has been confirmed.[47, 48] This lability is a consequence not only of the stability of the incipient non-classical carbonium ion, but of the *exo*-orientation of the leaving group (since bornyloxycarbonylamino-derivatives undergo

[46] M. Fujino, S. Shinagawa, O. Nishimura, and T. Fukuda, *Chem. and Pharm. Bull. (Japan)*, 1972, **20**, 1017.
[47] G. Jäger and R. Geiger, ref. 3, p. 78.
[48] G. Jäger and R. Geiger, *Annalen*, 1973, 1535.

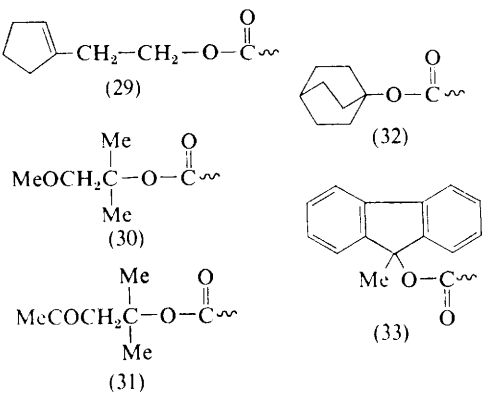
acidolysis less readily than derivatives in the iso-series)[47, 48] (Scheme 15). The solubility of amino-acids and peptides is enhanced more by this new protecting moiety than by others, particularly if it is used in the racemic form, but then of course purification of synthetic intermediates can become



Scheme 15

more difficult on account of their lack of crystallinity.[48] The new protecting group has been used in the synthesis of a number of peptides,[48] including gonadotropin-releasing hormone[49] and thyrotropin-releasing hormone.[50]

Use of the 2-(cyclopenten-1-yl)ethoxycarbonyl group (29),[51] the 2-methoxy-1,1-dimethylethoxycarbonyl group (30),[52] the 2-methylpentan-4-on-2-yloxycarbonyl group (31),[52] the bicyclo[2,2,2]oct-1-yloxycarbonyl group (32),[53] and of the 9-methyl-9-fluorenyloxycarbonyl group (33)[54] has been reported.



[49] M. Fujino, T. Fukuda, S. Kobayashi, and M. Obayashi, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 87.

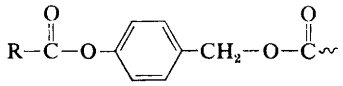
[50] M. Fujino, T. Fukuda, and C. Kitada, *Takeda Kenkyusho Ho*, 1973, **32**, 12 (*Chem. Abs.*, 1973, **79**, 32 291).

[51] E. Vargha, A. Balog, F. Gonczy, D. Breazu, and L. Beu, *Rev. Roumaine Chim.*, 1973, **18**, 1627.

[52] I. Kitsukawa, Jap. Kokai 73 54 002 (Cl. 16 A6, 16 B65, 16 B663) (*Chem. Abs.*, 1974, **80**, 15 210).

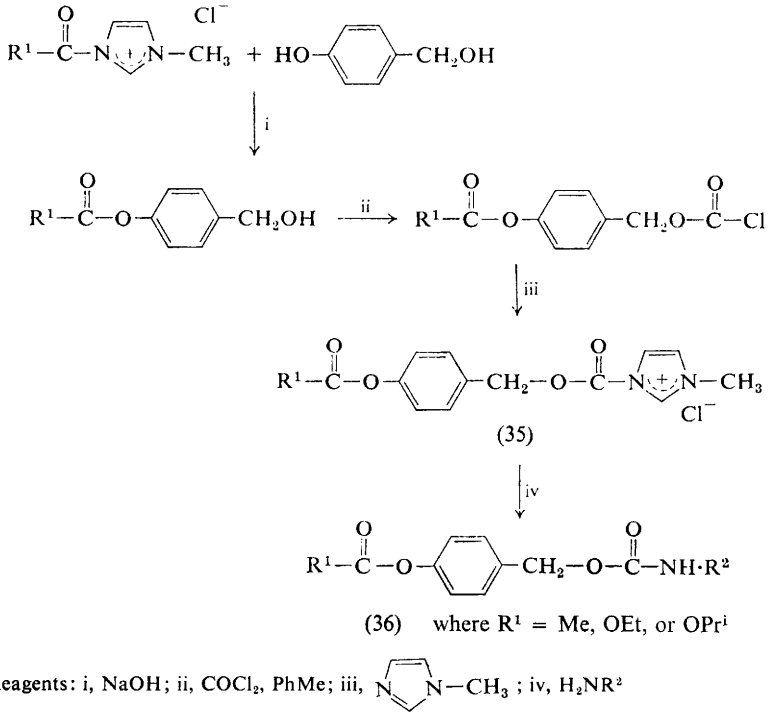
[53] I. Kitsukawa, Jap. Kokai 73 56 603 (Cl. 16 A6) (*Chem. Abs.*, 1973, **79**, 146 859).

[54] H. Otsuka and K. Inouye, Ger. Offen. 2 234 798 (Cl. C 07c) (*Chem. Abs.*, 1973, **78**, 124 886).



(34) R = Me, OEt, or OPr^i

A series of alkali-labile substituted benzyloxycarbonyl amino-protective groups has been investigated,[55] *viz.* 4-acetoxybenzyloxycarbonyl (34; R = Me), 4-ethoxycarbonyloxybenzyloxycarbonyl (34; R = OEt), and 4-isopropoxycarbonyloxybenzyloxycarbonyl (34; R = OPr^i). They are introduced into amino-acids (Scheme 16) using the crystalline water-



(35)

(36) where R^1 = Me, OEt, or OPr^i

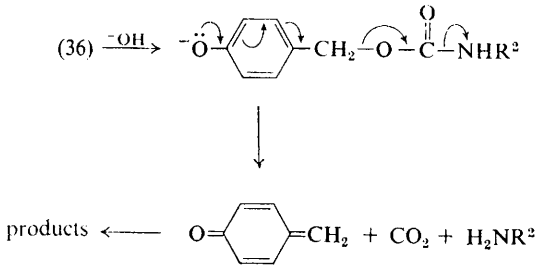
Reagents: i, NaOH; ii, $COCl_2$, PhMe; iii, ; iv, H_2NR^2

Scheme 16

soluble reagents (35; R^1 = Me, OEt, or OPr^i), whose synthesis is also shown in Scheme 16. The glycine derivatives (36; R^1 = Me, OEt, or OPr^i; R^2 = $CH_2 \cdot CO_2H$) undergo quantitative cleavage in the presence of 40% hydrogen bromide in acetic acid or on catalytic hydrogenation. The compound (36; R^1 = OPr^i, R^2 = $CH_2 \cdot CO_2H$) is stable in trifluoroacetic acid in dichloromethane but is readily cleaved in 5% aqueous potassium

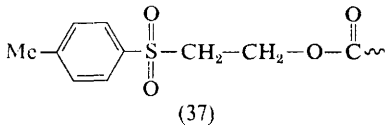
[55] M. Wakselman and E. Guibé-Jampel, *J.C.S. Chem. Comm.*, 1973, 593.

carbonate containing one equivalent of hydrazine (98% cleavage after 2 h) and completely cleaved with 0.1M sodium hydroxide solution. The group suffers fission, presumably *via* a 1,6-elimination involving a quinone methide intermediate (Scheme 17).



Scheme 17

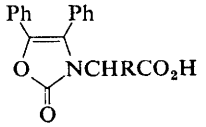
It would seem that what is required is an alkali-labile group which will undergo fission with non-hydrolytic basic reagents, *e.g.* the 9-fluorenylmethoxycarbonyl amino-protective group which is removable by treatment with liquid ammonia, ethanolamine, piperidine, or morpholine (see Vol. 5 of these Reports, p. 263). Similar deblocking reagents bring about the removal[56] of the 2-(toluene-*p*-sulphonyl)ethoxycarbonyl amino-protector (37).[57] In an attempt to investigate systems more labile in the presence of anhydrous base the alcohol (38) has been prepared;[58] however, under the conditions required to convert the alcohol into a urethane, dehydration occurs.[16]



(37)

$(PhSO_2)_2CHCH_2OH$

(38)

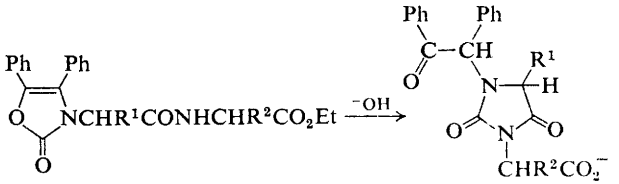


(39)

The ingenious method of protection of amino-groups by incorporating them into the stable unreactive 4,5-diphenyl-4-oxazolin-2-one ring system (39)[59] has been evaluated in detail[60] (see Vol. 5 of these Reports, p. 268). The most convenient procedure for the incorporation involves treating

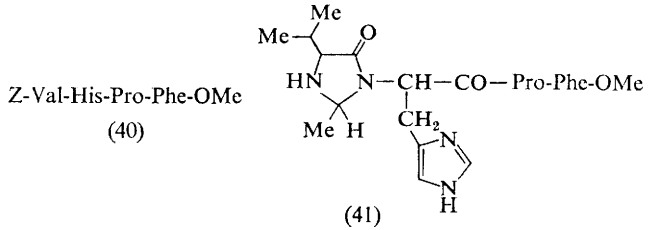
[56] L. A. Carpino and G. Y. Han, *J. Org. Chem.*, 1972, **37**, 3404.
[57] A. T. Kader and C. J. M. Stirling, *J. Chem. Soc.*, 1964, 258.
[58] L. A. Carpino, *J. Org. Chem.*, 1973, **38**, 2600.
[59] F. S. Guziec, jun., and J. C. Sheehan ref. 4, p. 321.

an amino-acid tetramethylammonium salt in DMF with benzoin cyclic carbonate; the resulting hydroxyoxazolidinone mixture is subsequently dehydrated with trifluoroacetic acid. A wide range of amino-acid derivatives was prepared in very good yield. The reductive removal of the group by catalytic hydrogenolysis is best effected in ethanol or DMF containing an equivalent of aqueous mineral acid. The oxidative removal using *m*-chloroperbenzoic acid is too vigorous for general use with sensitive amino-acids. In coupling situations no racemization could be detected at 1% sensitivity. Hydantoin formation is a troublesome side-reaction during the alkaline hydrolysis of protected dipeptide esters (Scheme 18).[60]



Scheme 18

The tendency of 4-imidazolidinones to arise as by-products which interfere with normal peptide synthesis has been discussed,[61] *e.g.* hydrogenolysis of (40) in ethanol in the presence of triethylamine and palladium-on-charcoal as the catalyst yielded varying amounts of the epimeric

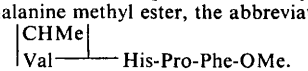


(40) (41)

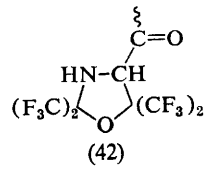
mixture (41), the acetaldehyde that was responsible arising from ethanol and triethylamine by oxidation processes.[62] Clearly, inadvertent formation and splitting of imidazolidinone peptides (*NN'*-alkylidene-peptides)* during synthesis is a source of uncertainty with regard to sequence homogeneity. Such an *NN'*-ethylidene-peptide is relatively stable below pH 3.5 but readily loses acetaldehyde when warmed in neutral aqueous solution.

[60] J. C. Sheehan and F. S. Guziec, jun., *J. Org. Chem.*, 1973, **38**, 3034.
[61] F. Cardinaux and M. Brenner, ref. 3, p. 65.
[62] F. Cardinaux and M. Brenner, *Helv. Chim. Acta*, 1973, **56**, 339.

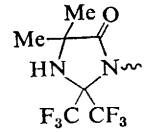
* The suggested nomenclature[62] for compounds of the type (41) is (*NN'*-ethylidene-valyl-histidyl)-prolyl-phenylalanine methyl ester, the abbreviation being



Two equivalents of hexafluoroacetone in DMSO react with *N*-terminal glycyl-peptides and glycine ester to yield fluorinated products containing an oxazolidine ring (42).[63] When, however, the *N*-terminal residue is α-methylalanyl, equivalent amounts of the reactants combine to yield a polyfluorinated imidazolidinyl peptide (43).[63]

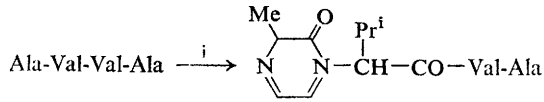


(42)



(43)

Certain free *N*-terminal peptides react with glyoxal to yield 3-alkylpyrazin-2-on-1-yl acylpeptides (Scheme 19).[64] Mixed anhydrides of 1-ethoxycarbonylpropen-2-yl or 2-ethoxycarbonylcyclopenten-1-yl amino-acids (generated from the corresponding dicyclohexylammonium salts)

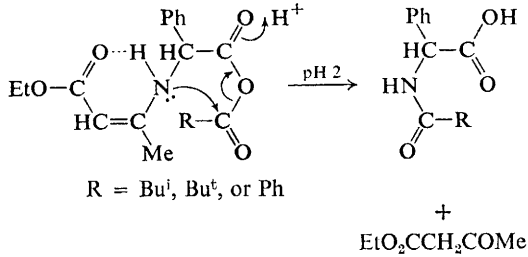


Reagents: i, $(CHO)_2$, pH 5, 100 °C

Scheme 19

undergo an acyl migration at pH 2 with concomitant cleavage of the β-dicarbonyl component (see for example Scheme 20).[65] Racemization of the amino-acid residue occurs only if a catalytic amount of *N*-methylmorpholine is present; the rearrangement is not general.

Heretofore the allyloxycarbonyl group (44) has not been employed in peptide synthesis; this situation may change in view of the report that the

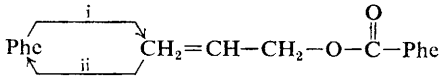


R = Bu^i, Bu^t, or Ph

Scheme 20

[63] C. A. Panetta, T. G. Casanova, and Chia-chi Chu, *J. Org. Chem.*, 1973, **38**, 128.
[64] E. Vargha, F. Gonczy, and A. Balog, *Studia Univ. Babes-Bolyai, Ser. Chem.*, 1972, **17**, 125 (*Chem. Abs.*, 1973, **78**, 136 654).
[65] A. Balog, D. Breazu, V. Voinescu, M. Herman, E. Vargha, and E. Ramontian, *Rev. Roumaine Chim.*, 1973, **18**, 123.

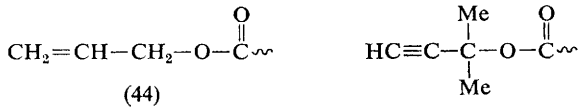
group can be removed readily both from nitrogen and oxygen functions with the aid of nickel carbonyl.[66] Scheme 21 gives the conditions for acylation and deacylation of racemic phenylalanine. This important observation points the way to the need for a systematic study of the



Reagents: i, CH_2:$CHCH_2OCOCl$, C_5H_5N, in THF or Et_2O; ii, $Ni(CO)_4$ (5 equivalents), $Me_2NCH_2CH_2NMe_2$ (3 equivalents), Me_2NCHO–H_2O (95 : 5), 55 °C, N_2 atmosphere, 4 h

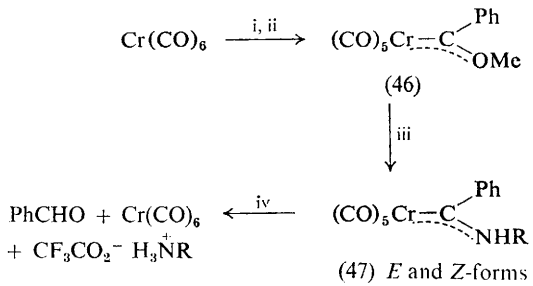
Scheme 21

behaviour of peptide derivatives towards various metal carbonyls [*e.g.* it would be interesting to see if this new reagent will also cleave the substituted propargyl group in the recently described 1,1-dimethyl-2-propynyloxycarbonyl group (45)[67]].



(44) (45)

Pentacarbonyl(phenylcarbene)chromium has been evaluated as an amino-protective group in peptide synthesis.[68] Aminolysis of pentacarbonyl[methoxy(phenyl)carbene]chromium (46), prepared[69] as shown in Scheme 22, with amino-esters yields the corresponding intensely yellow



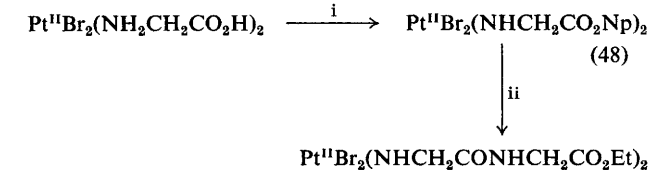
Reagents: i, LiPh in Et_2O; ii, $[Me_3\overset{+}{O}]\overset{-}{B}F_4$ in H_2O; iii, H_2NR in Et_2O; iv, CF_3CO_2H, 10 min at 20 °C or 80% HOAc, 30 min at 80 °C

Scheme 22

[66] E. J. Corey and J. W. Suggs, *J. Org. Chem.*, 1973, **38**, 3223.
[67] G. L. Southard, B. R. Zaborowsky, and J. M. Pettee, *J. Amer. Chem. Soc.*, 1971, **93**, 3302.
[68] K. Weiss and E. O. Fischer, *Chem. Ber.*, 1973, **106**, 1277.
[69] E. O. Fischer, B. Heckl, K. H. Dötz, J. Müller, and H. Werner, *J. Organometallic Chem.*, 1969, **16**, P29.

[amino(phenyl)carbene]pentacarbonylchromium complex (47). Such complexes are stable to heat and to dilute aqueous acids and bases; they are very soluble in ether, dioxan, benzene, dichloromethane, and alcohol, and sparingly soluble in water and carbon tetrachloride. The conditions for the removal of the protecting group with acetic acid or trifluoroacetic acid are also shown in Scheme 22. The amino-ester complexes (47; R = $CHR^2 \cdot CO_2Me$) can be saponified under normal conditions and coupled with amino-esters, using *N*-hydroxysuccinimide–dicyclohexylcarbodi-imide, to give good yields of yellow crystalline protected peptides.[68] We await with interest a report on the optical purity of the peptides which were synthesized.

Ligand-substitution reactions of *trans*-dichlorobis(amino-active ester)-platinum(II) complexes have been studied.[70] *trans*-Dibromo- and -dichloro-bis(glycine)platinum(II) complexes (see Vol. 5 of these Reports, p. 271) can be converted into active ester complexes (48) by the action of dicyclohexylcarbodi-imide and the substituted phenol. Such *N*-co-ordinated active ester ligands can react with amino-esters and peptide esters to yield N-protected peptides (see for example Scheme 23). It is



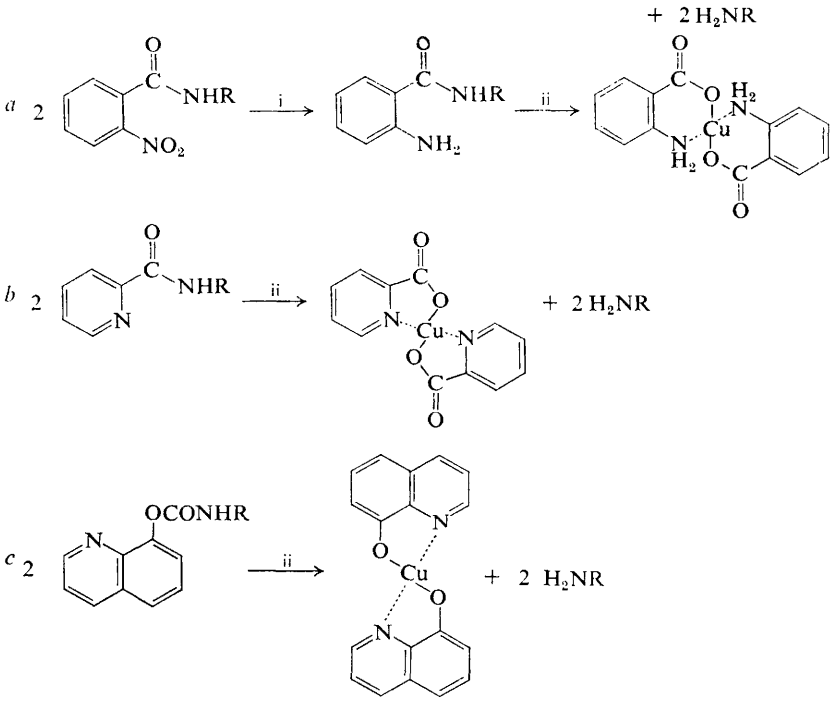
Reagents: i, $C_6H_{11}N{:}C{:}NC_6H_{11}$–HONp–$Me_2NCHO$; ii, H-Gly-OEt, Me_2NCHO

Scheme 23

time that the thorny problem of the removal of these peptides from the platinum(II) complex was studied.

The synthesis of some simple peptides has been accomplished using *N*-(*o*-nitrobenzoyl)amino-acids, but racemization is to be expected. The N-protecting group can be removed, after hydrogenation to the corresponding *N*-anthranilylpeptide, by copper(II)-catalysed hydrolysis (Scheme 24a).[71] *N*-(α-Picolinoyl)amino-acids have been used in peptide synthesis. The *N*-α-picolinoyl group can also be removed by hydrolysis that is promoted by copper acetate (Scheme 24b).[72] The removal of these two protecting groups is highly reminiscent of the copper(II)-catalysed hydrolysis of the 8-quinoloxycarbonyl amino-protective group [73] (Scheme 24c).

[70] W. Beck, B. Purucker, and E. Strissel, *Chem. Ber.*, 1973, **106**, 1781.
[71] A. K. Koul, J. M. Bachhawat, B. Prashad, N. S. Ramegourda, A. K. Mathur, and N. K. Mathur, *Tetrahedron*, 1973, **29**, 625.
[72] A. K. Koul, N. K. Mathur, N. S. Ramegourda, B. Prashad, and J. M. Bachhawat, *Synthetic Comm.*, 1972, **2**, 383.
[73] E. J. Corey and R. L. Dawson, *J. Amer. Chem. Soc.*, 1962, **84**, 4899.



Reagents: i, PtO_2–H_2; ii, aq: $Cu(OAc)_2$

Scheme 24

Protection of Carboxy-groups.—A general review has appeared concerning the protection of carboxy-groups in synthesis.[15d] It contains a useful tabular summary of twenty-one carboxy-protective groups, including the methods for their introduction and removal.

The use of phenyl esters for carboxy-protection has been investigated with more complex peptides. The group is nevertheless still removed, rapidly, by hydrolysis catalysed by peroxide anion (up to three equivalents of peroxide are used) at pH 10.5, without racemization. However, peptide phenyl esters tend to undergo solvolysis in the presence of methanol and so the use of isopropyl alcohol is recommended for their chromatography.[74]

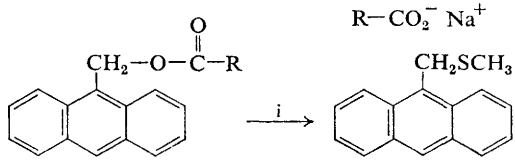
Two conference communications concerning the picolyl ester method of peptide synthesis indicate that the method can now be used routinely to synthesize peptides of up to twelve amino-acid residues in length.[75, 76] The preferred procedures are discussed.

[74] D. Hudson, G. W. Kenner, B. Mason, B. Morgan, R. Ramage, B. Singh, and R. Tyson, ref. 5, p. 70.
[75] G. A. Fletcher and G. T. Young, ref. 3, p. 82.
[76] G. A. Fletcher, A. Gosden, P. P. Nash, and G. T. Young, ref. 5, p. 65.

2,2,2-Trichloroethyl esters of *N*-benzyloxycarbonyl-amino-acids can be prepared by the acid chloride method. Selective removal of the N-protecting group with hydrogen bromide in acetic acid enables the synthesis of peptides, using dicyclohexylcarbodi-imide in acetonitrile for coupling. Selective removal of the carboxy-protective group from benzyloxycarbonyl peptide esters is achieved with zinc dust in 90% acetic acid.[77] The *p*-nitrobenzyl ester carboxy-protective group can be cleaved with zinc dust and hydrochloric acid.[78]

Methylthiomethyl esters can be prepared by heating carboxylic acid triethylammonium salts with chloromethyl methyl sulphide. The ester is stable to mild reducing agents such as sodium borohydride or zinc dust in methanol, but cleaved by trifluoroacetic acid (15 min treatment) or by heating under reflux with methyl iodide in aqueous acetone (17 h treatment).[79] The method has not so far been applied to amino-acids or peptides.

9-Anthrylmethyl esters (49), prepared by treating carboxylic acid triethylammonium salts with 9-chloromethylanthracene, can be cleaved in very high yield by treatment with the sodium salt of methanethiol in DMF (the reaction is extraordinarily fast in hexamethylphosphoramide) (Scheme 25).[80] The method has not been applied to peptides.



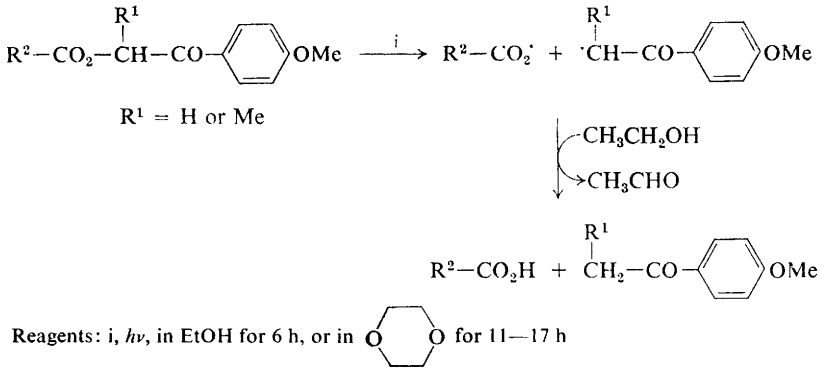
(49)

Reagents: i, $MeS^- Na^+$ in Me_2NCHO at −20 °C for 1 h, or in $(Me_2N)_3PO$ at 25 °C for 30 s

Scheme 25

The *p*-methoxyphenacyl group and α-methylphenacyl group are useful photolabile groups for protecting carboxy-functions.[81] The photocleavage is carried out in ethanol or dioxan (hydrogen donors) (at substrate concentrations of 5×10^{-3}—10^{-2} mol l^{-1}) using a mercury immersion lamp fitted with a Pyrex absorption sleeve (exciting wavelengths above 313 nm). The mechanism of cleavage is thought to be a radical scission of the carbon–oxygen bond (Scheme 26), since the reaction occurs in the presence of cumene (a hydrogen radical donor), and involves a long-lived triplet

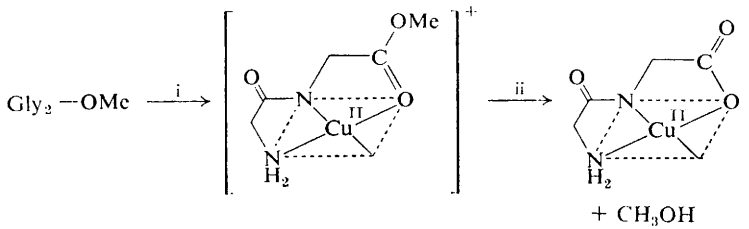
[77] B. Marinier, Y. C. Kim, and J.-M. Navarre, *Canad. J. Chem.*, 1973, **51**, 208.
[78] E. L. Smithwick, jun., and J. E. Shields, Ger. Offen. 4 256 573 (Cl. C 07*c*) (*Chem. Abs.*, 1973, **79**, 92 602).
[79] T.-L. Ho and C. M. Wong, *J.C.S. Chem. Comm.*, 1973, 224.
[80] N. Kornblum and A. Scott, *J. Amer. Chem. Soc.*, 1974, **96**, 590.
[81] J. C. Sheehan and K. Umezawa, *J. Org. Chem.*, 1973, **38**, 3771.



Scheme 26

state (evidence from quenching experiments). *N*-Benzyloxycarbonyl, *N*-t-butoxycarbonyl groups, and peptide bonds are stable under the conditions described; the phthaloyl group suffers partial decarboxylation and the trityl group is unstable; within these limitations, however, cleavage of the esters occurs in good yield.

The rate of hydrolysis of the methyl ester of glycylglycine is increased by three orders of magnitude in the presence of copper(II).[82] Dissociation of the amide proton is of major importance in the catalysed process, since co-ordination of this deprotonated centre moves the ester group into position to bind with the metal ion, which promotes the hydrolysis of the ester (Scheme 27). The copper(II) acts as a 'template' catalyst preventing



Reagents: i, Cu^{II}, pH 6.5—8.0; ii, $\bar{O}H$

Scheme 27

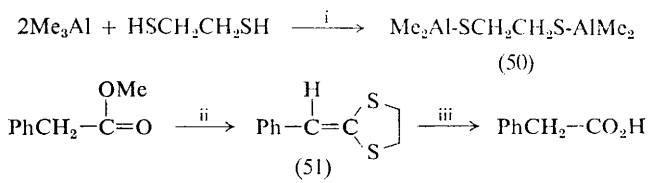
side-reactions such as 2,5-dioxopiperazine formation, which is observed in the uncatalysed hydrolysis. Naturally, the hydrolysis of glycylsarcosine methyl ester is not promoted by copper ions.

A few β-(dimethylamino)ethyl esters of amino-acids and peptides have been prepared.[83] The removal of a benzyl ester by electrolytic reduction

[82] R. Nakon and R. J. Angelici, *J. Amer. Chem. Soc.*, 1973, **95**, 3170.
[83] O. L. Mndzhoyan and S. A. Kazaryan, *Armyan. khim. Zhur.*, 1973, **26**, 395 (*Chem. Abs.*, 1973, **79**, 126 767). See also ref. 274.

has already been mentioned.[23] The partial hydrolysis of the bis-benzyl ester of diaminopimelic acid derivatives has been described.[84] An improved synthesis of L-histidine methyl ester has appeared.[85] The stability of the 1-butanol–hydrogen chloride Fischer esterification reagent has been studied.[86] Methods for the chemical synthesis of amino-acid and peptide esters of nucleosides, nucleotides, and oligonucleotides have been reviewed.[87]

An ingenious technique has been developed for the protection of esters (and lactones) against nucleophilic attack. Esters possessing an α-hydrogen atom react with bis(dimethylaluminium)1,2-ethanedithiolate (50) to yield keten thioacetals (51) (Scheme 28) which are fairly stable to reagents such



Reagents: i, $PhCH_3$–CH_2Cl_2 at −78 °C, gradually being raised to 25 °C; ii, (50) in CH_2Cl_2, 25 °C, 12 h; iii, HgO–BF_3,Et_2O in aq. THF at 60 °C, 4 h

Scheme 28

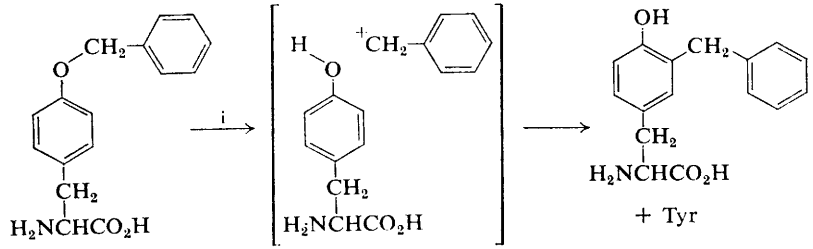
as methanolic potash, lithium aluminium hydride in ether, or aqueous acetic acid. Keten thioacetals yield carboxylic acids when treated with mercuric oxide–boron trifluoride.[88] Unfortunately, the stereochemistry of α-chiral centres is lost in this process.

Protection of Hydroxy-groups.—General reviews have appeared concerning the protection of alcoholic [15b] and phenolic [15c] hydroxy-groups.

Protection of the phenolic group of tyrosine as a benzyl ether (for the preparation of *O*-benzyltyrosine see ref. 89) is unsatisfactory on two counts.[90–92] Under the conditions of *N*-α-deprotection (50% trifluoroacetic acid in dichloromethane for the t-butoxycarbonyl group, extensively used in solid-phase synthesis), *O*-benzyltyrosine suffers partial deprotection,[90–93] and the product mixture consists of tyrosine and 3-benzyl-

[84] A. Arendt, A. Kolodziejczyk, and T. Sokolowska, *Roczniki Chem.*, 1973, **47**, 1087 (*Chem. Abs.*, 1973, **79**, 115 878).
[85] N. A. Brusentov, V. S. Gruzdev, V. A. Novoselov, and Yu. P. Shvachkin, Russ. P. 320 487 (Cl. C 07*c*, C 07*d*) from *Otkrytiya, Izobret., Prom. Obraztsy, Tovarnye Znaki*, 1972, **49**, 186 (*Chem. Abs.*, 1973, **78**, 124 889).
[86] J. P. Hardy, S. L. Kerrin, and S. L. Manatt, *J. Org. Chem.*, 1973, **38**, 4196.
[87] T. L. Tsilevich, A. A. Kraevskii, and B. P. Gottikh, *Russ. Chem. Rev.*, 1972, **41**, 822.
[88] E. J. Corey and D. J. Beames, *J. Amer. Chem. Soc.*, 1973, **95**, 5829.
[89] H. Eckstein, R. E. Sievers, and E. Bayer, *Annalen*, 1973, 1467.
[90] B. W. Erickson and R. B. Merrifield, ref. 4, p. 191.
[91] D. Yamashiro, R. L. Noble, and C. H. Li, ref. 4, p. 197.
[92] B. W. Erickson and R. B. Merrifield, *J. Amer. Chem. Soc.*, 1973, **95**, 3750.
[93] D. Yamashiro and C. H. Li, *J. Org. Chem.*, 1973, **38**, 591.

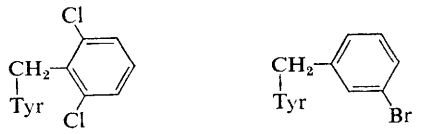
tyrosine.[90, 91, 93] Since essentially similar results are obtained when a carbonium-ion scavenger is present,[90, 92] the 3-benzyltyrosine is probably formed by an essentially intramolecular rearrangement (Scheme 29). In



Reagents: i, CF_3CO_2H–CH_2Cl_2

Scheme 29

an attempt to overcome these difficulties the *m*-bromobenzyl[91, 93] and 2,6-dichlorobenzyl groups[90–92] have been investigated [2,6-dichlorobenzyltyrosine (52) is prepared[92, 94] by the action of 2,6-dichlorobenzyl bromide on tyrosine–copper complex]. Both derivatives (52) and (53)



(52) (53)

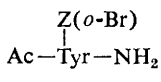
show an enhanced stability towards trifluoroacetic acid in dichloromethane (stable for at least 350 h at 20 °C[92]) but are, nevertheless, readily cleaved under the conditions of final deprotection (hydrogen fluoride, 10 min at 0 °C), yielding a mixture of tyrosine (60%) and the substituted 3-benzyl derivative (40%).[92] However, in the presence of anisole the amount of isomerized product is reduced to 5%, thus indicating that ring benzylation in hydrogen fluoride in the absence of anisole occurs mainly by an intermolecular route. Thus, use of the *m*-bromobenzyl or 2,6-dichlorobenzyl ether represents an improvement over the benzyl group, but does not entirely eliminate the problem.[90–92] A successful solid-phase synthesis of human adrenocorticotropin has been accomplished using 2,6-dichlorobenzyl protection for tyrosine.[94]

O-Benzyloxycarbonyltyrosine is not sufficiently stable to trifluoroacetic acid in dichloromethane for the *O*-benzyloxycarbonyl group to be a useful alternative protection to the substituted benzyl groups mentioned above, although the undesired intramolecular benzylation no longer takes place.[92] Furthermore, *O*-benzyloxycarbonyltyrosine loses its protection in the

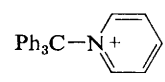
[94] D. Yamashiro and C. H. Li, *J. Amer. Chem. Soc.*, 1973, **95**, 1310.

presence of nucleophiles such as di-isopropylethylamine in dichloromethane (the nucleophile here is apparently the liberated amino-group), and thiols.[92] In order to overcome these difficulties the use of *N*-α-t-butoxycarbonyl-*O*-(*o*-bromobenzyloxycarbonyl)tyrosine has been investigated in a successful solid-phase synthesis of an octapeptide which occurs in the human growth hormone sequence (residues 138—145).[93] This protecting group is stable under the conditions of repeated N-α-deprotection [*N*-α-acetyl-*O*-(*o*-bromobenzyloxycarbonyl)tyrosinamide (54) shows a 1% loss in O-protection after 24 h in 50% trifluoroacetic acid in dichloromethane], and yet is completely removed in hydrogen fluoride (10 min at 0 °C), only a single product being obtained. However, the group still shows slight instability to nucleophiles [compound (54) loses *ca.* 5% of its O-protection during a 24 h treatment with 10% di-isopropylethylamine in DMF].[93]

Brief mention has been made of the use of the *O*-4-picolyl group for tyrosine protection. Deblocking is effected by electrolytic reduction.[95]



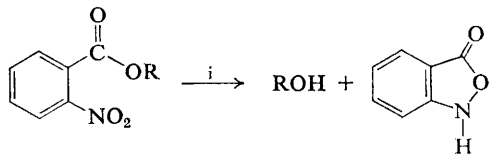
(54)



(55)

The 9-anthrylmethyl group is suitable for the protection of phenolic hydroxy-groups.[80] 9-Anthrylmethyl ethers are prepared by allowing a phenate salt to react with 9-chloromethylanthracene. Deblocking is accomplished with the sodium salt of methanethiol in DMF. This new method has not yet been applied to the tyrosine problem.

Tritylpyridinium fluoroborate (55) is a useful reagent for the tritylation of primary alcoholic functions.[96] The *o*-nitrobenzoyl group is potentially useful for protecting alcohols and phenols since it can be removed in high yield by reduction with zinc dust and ammonium chloride (Scheme 30).[97]



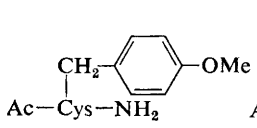
Reagents: i, Zn dust–NH_4Cl, aq. THF

Scheme 30

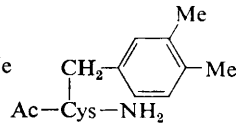
Protection of Thiol Groups and the Synthesis of Cystine Peptides.—A detailed critical review of the protection of thiols has appeared.[15e]

[95] G. T. Young and A. Kotai, ref. 5, p. 128.
[96] S. Hanessian and A. P. A. Staub, *Tetrahedron Letters*, 1973, 3555.
[97] D. H. R. Barton, I. H. Coates, and P. G. Sammes, *J.C.S. Perkin I*, 1973, 599.

Deprotection of *S*-*p*-methoxybenzyl-L-cysteine with anhydrous hydrogen fluoride, using the literature conditions,[98, 99] does not proceed to completion even in the presence of anisole and tyrosine (subsequent oxidation giving only an 83% yield of cystine).[100] Deprotection of the tetra-*S*-*p*-methoxybenzyl-A-chain of insulin under similar conditions is incomplete; only 70% of the free thiol groups are liberated during a 30 min reaction time.[100] On the other hand, under the conditions for removal of *N*-t-butoxycarbonyl groups in solid-phase synthesis, the *S*-*p*-methoxybenzyl group may not provide sufficient protection during the synthesis of large peptides[90, 92] [(56) loses 27% of its S-protection during a 23 h treatment with 50% trifluoroacetic acid in dichloromethane[91]]. It is interesting to note that in the synthesis of arginine-vasopressin using *S*-*p*-methoxybenzyl protection, final deprotection was accomplished with anhydrous hydrogen



(56)



(57)

fluoride containing trifluoroacetic acid and anisole.[98] In an effort to overcome the problem of the lability of the *S*-*p*-methoxybenzyl group in solid-phase synthesis, the use of *S*-3,4-dimethylbenzyl protection has been investigated [(57) loses only 0.2% of its S-protection during a 23 h treatment with 50% trifluoroacetic acid in dichloromethane, and yet is completely removed under conditions for final deprotection (anhydrous hydrogen fluoride, 15 min at 0 °C)].[91] The suitability of this protecting group has been demonstrated with a solid-phase synthesis of the *C*-terminal cyclic dodecapeptide of human growth hormone.[101] The 4-methylbenzyl group has also been proposed as an alternative thioether protection; *S*-4-methylbenzyl-L-cysteine is stable to trifluoroacetic acid in dichloromethane but the group is removable with hydrogen fluoride–anisole at 0 °C within 60 min.[92]

The acetamidomethyl protecting group has become one of the most widely used S-protecting groups. Improvements in the isolation of *S*-acetamidomethyl-L-cysteine (58) have been reported.[102] The compound is prepared by treating cysteine with acetamidomethanol, over several

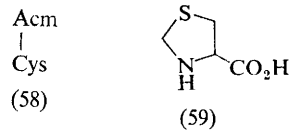
[98] S. Sakakibara, Y. Shimonishi, M. Okada, and Y. Kishida, Proceedings of the Eighth European Peptide Symposium held at Noordwijk, The Netherlands, 1966: 'Peptides', ed. H. C. Beyerman, A. van de Linde, and W. Maassen van den Brink, North-Holland Publishing Co., Amsterdam, 1967, p. 44.

[99] S. Sakakibara, Y. Shimonishi, M. Okada, and H. Sugihara, *Bull. Chem. Soc. Japan*, 1967, **40**, 2164.

[100] K. Lübke and E. Schröder, ref. 3, p. 89.

[101] D. Yamashiro, R. L. Noble, and C. H. Li, *J. Org. Chem.*, 1973, **38**, 3561.

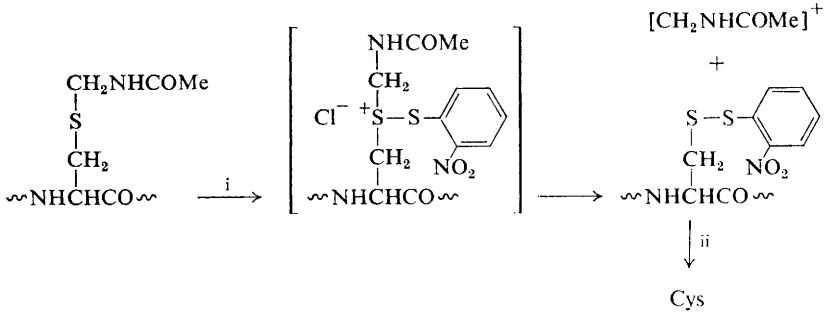
[102] P. Hermann and E. Schreier, ref. 5, p. 126.



(58) (59)

days, in hydrochloric acid at pH 0.5. Under these conditions the formaldehyde released from the amido-alcohol leads to the formation of L-thiazolidine-4-carboxylic acid (59). The mixture of (58) and (59) isolated from the reaction product by ion-exchange chromatography is treated with a mild oxidizing agent (such as periodate, potassium tri-iodide, or iodine), which converts (59) selectively into cystine. Repetition of the chromatography (strong cation exchanger, eluting with water then aqueous pyridine) enables the hydrate of (58) to be obtained pure in 75% yield, free of cystine and (59).[102]

The acetamidomethyl group is usually removed by mercury(II) salts at pH 4 in aqueous acetic acid, or by oxidation with iodine (if disulphides are required). The group can also be removed with the aid of *o*-nitrophenylsulphenyl chloride (Scheme 31).[103] In practice the acetamidomethylated



Reagents: i, Nps–Cl; ii, RSH

Scheme 31

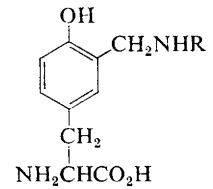
peptide is treated with the sulphenyl chloride in acetic acid containing 1M-hydrochloric acid to effect the cleavage.[104]

Veber *et al.*[105] had reported that acetamidomethyl-peptides on acidic hydrolysis gave rise to a mixture of cysteine and cystine in various proportions, unless the mixture was subsequently oxidized to cystine. Other workers have observed the practically quantitative conversion of acetamidomethyl-cysteine residues into thiazolidine-4-carboxylic acid as a result of this hydrolysis;[104] however, if an *o*-nitrophenylsulphenyl-amino-acid is

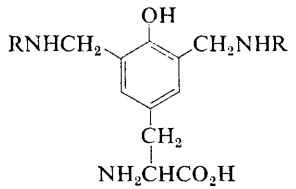
[103] A. Fontana, ref. 5, p. 122.
[104] L. Moroder, F. Marchiori, G. Borin, and E. Scoffone, *Biopolymers*, 1973, **12**, 493.
[105] D. F. Veber, J. D. Milkowski, S. L. Varga, R. G. Denkewalter, and R. G. Hirschmann, *J. Amer. Chem. Soc.*, 1972, **94**, 5456.

added to the hydrolysis mixture, as an internal standard, then cystine is released quantitatively.[104]

A warning has been issued that the acetamidomethylation of proteins in anhydrous hydrogen fluoride can lead to irreversible modification of tyrosine residues.[131] Model experiments in which cysteine–tyrosine mixtures (or tyrosine alone) were exposed to acetamidomethanol in hydrogen fluoride showed that tyrosine had been modified (even when a deficiency of the amido-alcohol was used), basic products being detected on an amino-acid analyser after acidic hydrolysis of the reaction mixture. It is assumed that these products arise from the direct acetamidomethylation of tyrosine and possess structures (60; R = Ac) and (61; R = Ac) before hydrolysis, and after hydrolysis the structures (60; R = H) and (61; R = H).

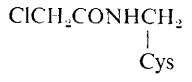


(60) R = H or Ac

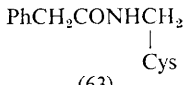


(61) R = H or Ac

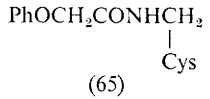
(62)



(64)



(63)

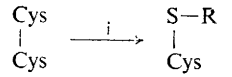


(65)

Modified acetamidomethyl derivatives (62)—(65) of cysteine have been prepared.[106] The corresponding amido-alcohols are more readily freed from formaldehyde than is the parent compound. The derivatives are prepared by treating cysteine with the amido-alcohol in anhydrous medium (DMF or glacial acetic acid, in the presence of hydrogen chloride in acetic acid or boron trifluoride). Compounds (62) and (63) crystallize very readily. It appears that *S*-acylamidomethyl groups can be removed with thiocyanogen; the method can be used for the construction of unsymmetrical disulphide bridges.[106]

The preparation of a further series of *S*-(alkylthio)cysteine derivatives has been reported (Scheme 32). With the exception of *S*-(cyclohexylthio)-cysteine and *S*-(*p*-tolylthio)cysteine the derivatives underwent rapid thiol–disulphide interchange with glutathione.[107]

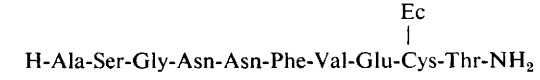
[106] H. Arold and M. Eule, ref. 5, p. 78.
[107] P. Hartter and U. Weber, *Z. physiol. Chem.*, 1973, **354**, 365.



Reagents: i, $PhCO_3H$ in $CHCl_3$ plus RSH in MeOH (R = CHMeEt, C_6H_{11}, Bu^n, CMe_2Et, CH_2Bu^t, n-C_8H_{17}, or *p*-MeC_6H_4)

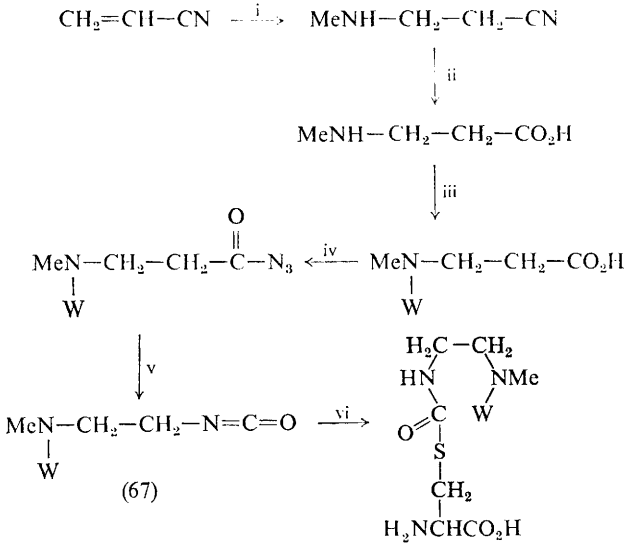
Scheme 32

The synthesis of the cysteine-containing tetracosapeptide amide (which corresponds to positions 81—104; subfragment G) of the proposed primary sequence of ribonuclease T_1 has been reported. The proposed stepwise tactics of synthesis of subfragment G_3 (66) had to be modified in favour



(66)

of a fragment approach, presumably because of difficulties arising in part from the presence of the alkali-labile *S*-ethylcarbamoyl moiety.[108] A modified *S*-ethylcarbamoyl protecting group has now been devised which uses the 'safety catch' principle.[109] The β-(*N*-acyl-*N*-methyl-aminoethyl)-



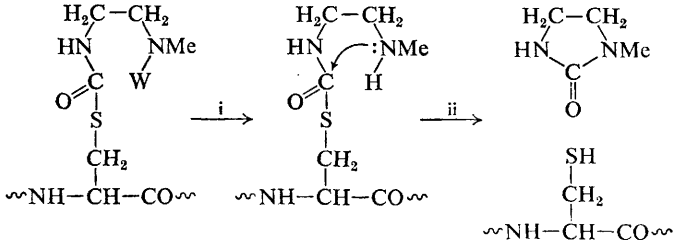
Reagents: i, $MeNH_2$; ii, $Ba(OH)_2$; iii, WX (where WX = ZCl, Boc-N_3, or Adoc-Cl); iv, mixed anhydride, HN_3; v, heat; vi, Cys, HCl

Scheme 33

[108] K. Kawasaki, R. Camble, G. Dupuis, H. Romovacek, H. T. Storey, C. Yanaihara, and K. Hofmann, *J. Amer. Chem. Soc.*, 1973, **95**, 6815.

[109] G. Jäger and R. Geiger, ref. 5, p. 90.

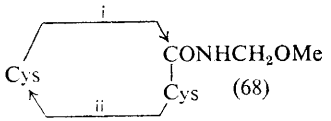
carbamoyl group is introduced into cysteine using the corresponding isocyanate (67) prepared according to Scheme 33, and removed according to Scheme 34. Deamino-oxytocin was prepared using this group. The



Reagents: i, acid; ii, tertiary base, 40—50 °C, 2—3 h

Scheme 34

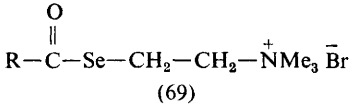
SH-group of cysteine can be blocked by reaction with methoxymethyl isocyanate. The resulting *N*-methoxymethyl-*S*-carbamoyl derivative (68) is stable at pH 6 but can be deblocked at alkaline pH (Scheme 35).[110]



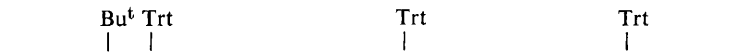
Reagents: i, $MeOCH_2N{:}C{:}O$, aq. solution, pH 4—5; ii, aq. solution, pH 9.6, 30—60 min

Scheme 35

Selective *S*-benzyloxycarbonylation of cysteine can be accomplished using (7) in neutral aqueous solution.[26] The kinetics of alkaline hydrolysis of some *S*-acylcysteine peptides have been studied.[111] The reactivity of *S*-acetylcysteine but not *S*-benzoylcysteine derivatives towards hydrolysis is enhanced by the presence of an adjacent hydroxy-amino-acid residue.[111] Acylselenol choline derivatives (69) can be used for the acylation of thiols.[112] On a small scale, selective removal of the *o*-nitrophenylsulphenyl group



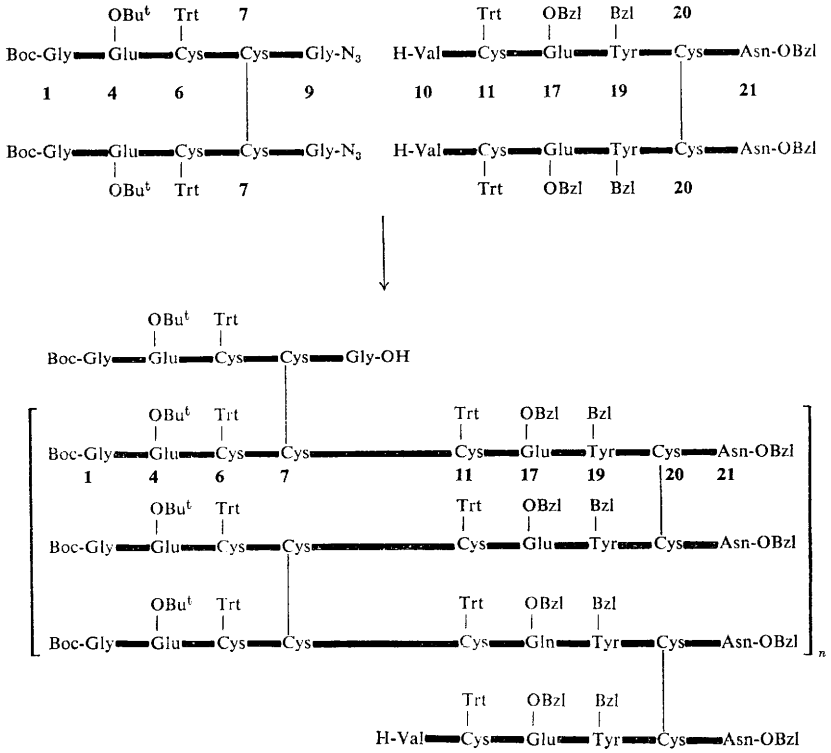
(69)



(70) (71)

[110] H. Tschesche and H. Jering, *Angew. Chem. Internat. Edn.*, 1973, **12**, 756.
[111] D. G. Clark and E. H. Cordes, *J. Org. Chem.*, 1973, **38**, 270.
[112] A. Makriyannis, W. H. H. Günther, and H. G. Mautner, *J. Amer. Chem. Soc.*, 1973, **95**, 8403.

from the *S*-tritylcysteine peptide (70) was accomplished using *o*-nitrophenylsulphenyl chloride in the presence of mercaptoethanol, the hydrochloride salt of the tetrapeptide being produced. However, on a larger scale, side-reactions intervened (loss of *S*-trityl ?) and coloured impurities were produced.[113] A high-yield synthesis of arginine vasopressin has been accomplished using *S*-trityl protection, cyclization of the intermediate (71) being accomplished in 70% yield with iodine in acetic acid.[114] Syntheses of the insulin B-chain using symmetrical cystine peptides[115, 116] have been discussed further. A similar approach has been applied to the insulin A-chain (Scheme 36).[117]

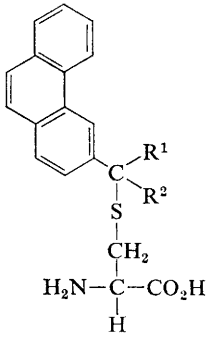


A-Chain polymer

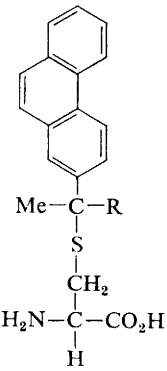
Scheme 36

[113] R. G. Hiskey, E. T. Wolters, G. Ülkü, and V. Ranga Rao, *J. Org. Chem.*, 1972, **37**, 2478.
[114] D. A. Jones, jun., R. A. Mikulec, and R. H. Mazur, *J. Org. Chem.*, 1973, **38**, 2865.
[115] H. Zahn, H. Berndt, P. Fehrenbach, D. von Wachtendonk, and H. Klostermeyer, ref. 3, p. 101.
[116] M. I. Titov, Z. A. Adremasova, Zh. D. Bespalova, and L. I. Leont'eva, *Doklady Akad. Nauk S.S.S.R.*, 1973, **209**, 227 (*Chem. Abs.*, 1973, **79**, 5577).
[117] W. O. Dano and E. Engels, ref. 4, p. 335.

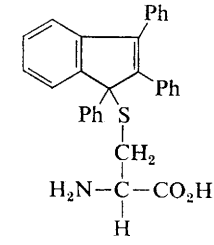
Thiol groups can be protected as 9-anthrylmethyl thioethers,[80] prepared by reaction of the corresponding thiolate anion with 9-chloromethylanthracene in DMF at 25 °C. Deblocking is most readily accomplished with 10 equivalents of sodium methanethiolate in hexamethylphosphoramide (2 h at 25 °C). The method has only been explored with hexane-1-thiol and toluene-ω-thiol. A series of substituted *S*-phenanthrylmethyl-L-cysteines (72) and (73) has been prepared[118] from the corresponding



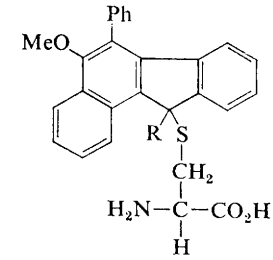
(72) a; $R^1 = R^2 = Me$
b; $R^1 = Me$, $R^2 = Ph$



(73) a; R = Me
b; R = Ph



(74)



(75) a; R = Me
b; R = Ph

phenanthrylcarbinols and L-cysteine hydrochloride in the presence of boron trifluoride etherate. A series of *S*-indenyl-L-cysteines[119] (74) and of *S*-benzofluorenyl-L-cysteines (75)[120] has been prepared similarly.

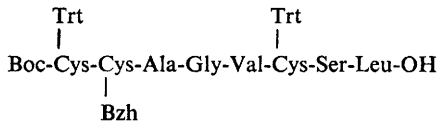
The autoxidation of a series of dicysteine oligopeptides has been studied.[121] Kinetic studies of the oxidation of cysteine with hydrogen peroxide in the presence of copper(II) ions are consistent with the view that the catalysed autoxidation of cysteine proceeds in two steps.[122] Hydrogen peroxide is generated as an intermediate and is subsequently used to oxidize cysteine directly.[122] The formation of disulphides of

[118] L. Mladenova-Orlinova and M. A. Tsvetkova, *Doklady Bolg. Akad. Nauk*, 1973, **26**, 493 (*Chem. Abs.*, 1973, **79**, 79 149).
[119] L. Mladenova-Orlinova, *Doklady Bolg. Akad. Nauk*, 1973, **26**, 497 (*Chem. Abs.*, 1973, **79**, 79 163).
[120] L. Mladenova-Orlinova, *Doklady Bolg. Akad. Nauk*, 1973, **26**, 643 (*Chem. Abs.*, 1973, **79**, 105 563).
[121] P. M. Hardy, B. Ridge, H. N. Rydon, and F. O. dos S. P. Serrão, ref. 3, p. 384.
[122] A. Hanaki and H. Kamide, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 1421.

cysteine derivatives, mediated by hexacyanoferrate(III) ions, is catalysed by small amounts of copper salts, present as impurities.[123] The observed rate law is consistent with a kinetic scheme involving several competing paths, and thiol complexes of copper(I—III) and an intermediate, thought to be a dithio-radical anion or its protonated form. Kinetic measurements have been carried out in the presence of added edta.[124]

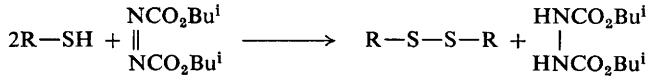
The amount of parallel and antiparallel dimers formed (as side-products) during the oxidation of oxytoceine to oxytocin with potassium hexacyanoferrate(III) can be reduced by a factor of three by excluding air from the system. The implication is that most of the side-products are formed by aerial oxidation of the dithiol solution prior to addition of hexacyanoferrate.[125]

The formation of an unsymmetrical disulphide corresponding to the A-(6—13)-octapeptide of insulin using the action of thiocyanogen on peptide (76) has been discussed further.[126] *S*-Tetrahydropyranyl derivatives



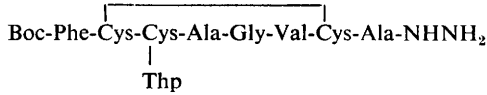
(76)

are insufficiently reactive, in the presence of thiocyanogen, to be used for the formation of disulphides.[106] Thiols can be rapidly oxidized to disulphides quantitatively, simply, and elegantly, using di-isobutyl azodicarboxylate (Scheme 37).[106] The sulphydryl derivative in aqueous alcohol



Scheme 37

or dioxan is titrated with a standard alcoholic solution of the reagent to a pale yellow end-point. The co-product di-isobutyl hydrazinodicarboxylate is easily removed from the reaction mixture. The cyclic disulphide (77)

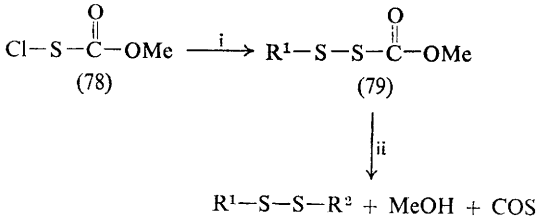


(77)

[123] G. J. Bridgart, M. W. Fuller, and I. R. Wilson, *J.C.S. Dalton*, 1973, 1274.
[124] G. J. Bridgart and I. R. Wilson, *J.C.S. Dalton*, 1973, 1281.
[125] M. Wälti and D. B. Hope, *Experientia*, 1973, **29**, 389.
[126] R. G. Hiskey, L. M. Beacham, tert., V. G. Matl, V. Ranga Rao, G. Ülkü, and E. T. Wolters, ref. 3, p. 107.

has been prepared by this method using high-dilution conditions. *S*-Alkylthioisothiuronium derivatives are suitable reagents for the synthesis of unsymmetrical disulphides, including *S*-(alkylthio)cysteines.[127]

The formation of unsymmetrical disulphides by the thiol-induced fragmentation of sulphenylthiocarbonates (see Vol. 4 of these Reports, p. 331) (Scheme 38) has now been applied to peptides,[128] with the synthesis



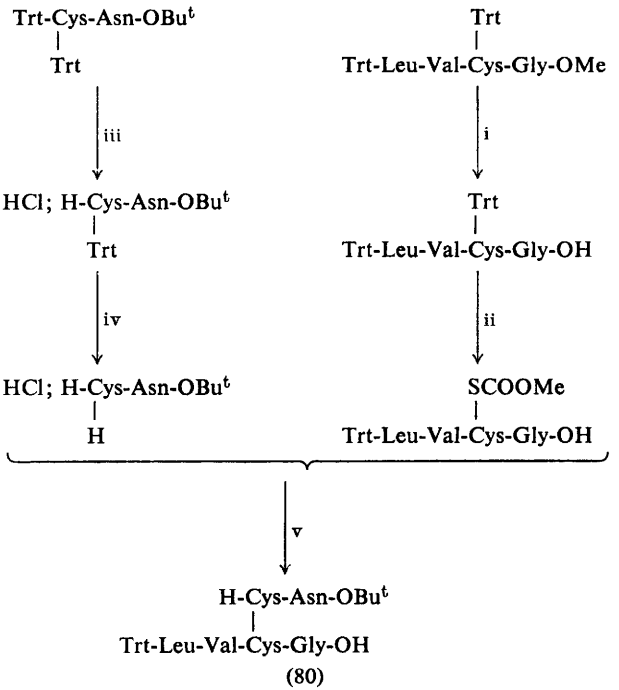
Reagents: i, R^1SH; ii, R^2SH

Scheme 38

of open-chain unsymmetrical derivatives of insulin fragments containing the A^{20}—B^{19} bridge. The reagent (78) is not stable in basic milieu; in methanol containing triethylamine, dimethyl carbonate and elemental sulphur are produced, while in an inert solvent such as THF containing triethylamine, polymerization occurs. In spite of this, however, sulphenylthiocarbonate derivatives of peptides (79) can be prepared readily, either by the action of the reagent (78) on a protected thiol-peptide, or by the action of the reagent on an S-protected (*S*-acetamidomethyl or *S*-trityl)-peptide. Naturally, for the formation of (79) the *N*-terminus of the substrate must be protected, but its carboxy-terminus can remain as a diethylammonium salt. These thiocarbonates are obtained in very good yield and are crystalline, stable both in the solid state and in solution, stable in the presence of trifluoroacetic acid (1 h at 0 °C), but labile in the presence of basic reagents such as hydroxide anion and hydrazine. The formation of peptide disulphides from a thiol-peptide and a sulphenylthiocarbonate occurs cleanly in high yield at room temperature in a variety of solvents. The most demanding disulphide synthesis reported is shown in Scheme 39, where the thiol component possesses a free amino-group and the sulphenylthiocarbonate a free carboxy-group. The resulting unsymmetrical cystine peptide (80) has been subjected to an elegant series of selective transformations (Scheme 40).[128] In a parallel study, where disulphides were produced by iodine oxidation of the corresponding *S*-tritylcysteine peptides, surprisingly high yields of unsymmetrical cystine derivatives were produced [*i.e.* higher than that expected on a purely statistical basis – *e.g.* compound

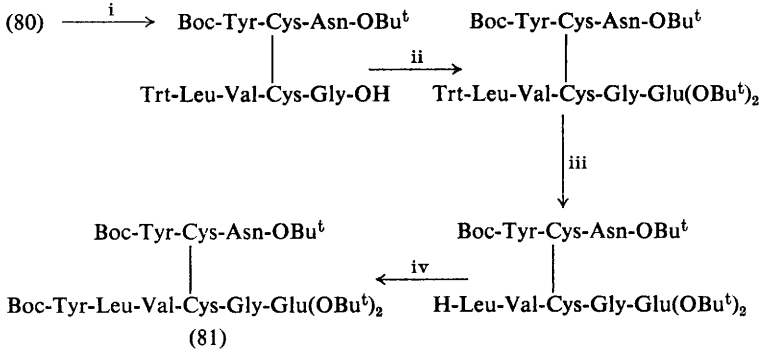
[127] O. Aki, M. Numata, and K. Shirakawa, *Takeda Kenkyusho Ho*, 1973, **32**, 1 (*Chem. Abs.*, 1973, **79**, 32 280).

[128] B. Kamber, *Helv. Chim. Acta*, 1973, **56**, 1370.



Reagents: i, NaOH, aq. dioxan; ii, $HNEt_2$ in $CHCl_3$–MeOH followed by MeOCOSCl (78); iii, aq. HOAc; iv, $Hg^{II}(OAc)_2$ in EtOAc–MeOH; v, $CHCl_3$–MeOH

Scheme 39



Reagents: i, Boc-Tyr-OPcp, NEt_3; ii, HCl; H-Glu(OBu^t)$_2$, NEt_3, $C_6H_{11}N{:}C{:}NC_6H_{11}$; iii, aq. HOAc; iv, Boc-Tyr-OPcp

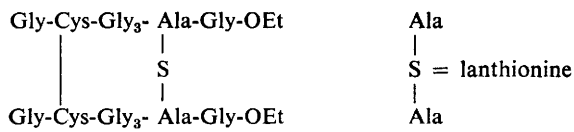
Scheme 40

(81) was obtained in 96% yield].[128, 129] This would seem to indicate that either the two *S*-trityl groups possess different reactivities towards iodine or the two derived sulphenyl iodides possess different intrinsic reactivities.

S-Picolyl derivatives [130] of the insulin chains have been prepared from tributylphosphine-reduced insulin.[131] Disulphides can be reduced quantitatively by polymer-bound dihydrolipoic acid.[132]

Zervas [133] suggested that the synthesis of interchain disulphides could be facilitated using a bifunctional blocking group. This idea has been applied to a model situation using cross-linked insulin, prepared by the action of the bis-succinimidyl esters of succinic and suberic acid, on intact insulin (this yields derivatives cross-linked between the α-amino-group of glycine-A1 and the ε-amino-group of lysine-B29). Conversion of either of these derivatives into the corresponding hexa-*S*-sulphonate, followed by reduction to the hexathiol and subsequent copper-catalysed autoxidation, regenerated the activity of cross-linked insulin *ca.* three times more efficiently than activity is regenerated by the re-oxidation of the isolated reduced chains [134] (see, however, Section 6).

The synthesis of a heterodetic cyclopeptide (82) containing a lanthionine residue has been reported.[135] Cyclic lanthionine derivatives (83) and the



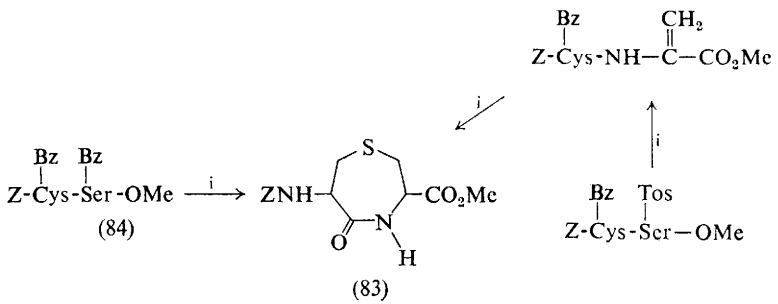
(82)

14-membered-ring dimer can arise during attempted transesterification of the dipeptide (84). These compounds were also obtained *via* the corresponding dehydroalanine peptides (Scheme 41).[136]

Methionine residues can be converted stereospecifically into methionine sulphoxide residues using a mild oxidative procedure involving tetrachloroauric(III) acid. L-Methionine gives rise to L-methionine-(*S*)-sulphoxide.[137]

Protection of Basic Side-chains.—Selective N^{im}-benzyloxycarbonylation of histidine can be accomplished in neutral aqueous solution using (7).[26] An indication has been obtained that the N^{im}-tosyl protecting group may not

[129] B. Kamber, *Helv. Chim. Acta*, 1971, **54**, 398.
[130] G. T. Young, A. Gosden, and D. Stevenson, ref. 5, p. 128.
[131] J. H. Seely, U. Ruegg, and J. Rudinger, ref. 5, p. 86.
[132] M. Gorecki and A. Patchornik, *Biochim. Biophys. Acta*, 1973, **303**, 36.
[133] L. Zervas, *Coll. Czech. Chem. Comm.*, 1962, **27**, 2242,
[134] S. M. L. Robinson, I. Beetz, O. Loge, D. G. Lindsay, and K. Lubke, *Tetrahedron Letters*, 1973, 985.
[135] A. Schöberl, M. Rimpler, and E. Graf, *Annalen*, 1973, 1379.
[136] L. Zervas and N. Ferderigos, *Experientia*, 1973, **29**, 262.
[137] E. Bordignon, L. Cattalini, G. Natile, and A. Scatturin, *J.C.S. Chem. Comm.*, 1973, 878.



Reagents: i, NaOMe, MeOH

Scheme 41

be as stable to trifluoroacetic acid as had been reported; nevertheless, it has been used successfully in a synthesis of the heptapeptide acid (residues 2 through 8) of insulin A-chain, using trifluoroacetic acid to deblock the intermediate N^{α}-t-butoxycarbonyl peptide esters.[138] N^{im}-Benzenesulphonyl and N^{im}-4-nitrobenzenesulphonyl derivatives of histidine have been prepared.[139] The N^{im}-2,4-dinitrophenyl group* can be removed from histidine bound to a polymeric carrier using thiophenol in DMF.[140, 141] The N^{im}-isobornyloxycarbonyl moiety has been used in a synthesis of thyrotropin-releasing hormone; it is removed by trifluoroacetic acid treatment.[142] A comparison of the ^{13}C chemical shift – pH profiles for imidazole, histidine, and 1-methyl- and 3-methyl-histidine shows that the predominant form of the imidazole ring of histidine, in basic solution, is the 1*H*-tautomer [143] (the *tele*-tautomer).

The synthesis of some protected L-2,4-diaminobutyric acid derivatives has been described.[144] The preparation of the four stereoisomeric forms of 2,3-diaminobutyric acid, and of some derivatives suitable for peptide

[138] G. P. Schwartz and P. G. Katsoyannis, *J.C.S. Perkin I*, 1973, 2894.
[139] S. Sakakibara and T. Fujii, Jap. Kokai, 73 07 436 (Cl. C 07*d*) (*Chem. Abs.*, 1973, **79**, 19 118).
[140] J. M. Stewart, M. Knight, A. C. M. Paiva, and T. Paiva, ref. 2, p. 59.
[141] G. R. Marshall, N. Eilers, and W. Vine, ref. 2, p. 15.
[142] M. Fujino, T. Fukuda, and C. Kitada, *Takeda Kenkyusho Ho*, 1973, **32**, 12 (*Chem. Abs.*, 1973, **79**, 32 291).
[143] W. F. Reynolds, I. R. Peat, M. H. Freedman, and J. R. Lyerla, jun., *J. Amer. Chem. Soc.*, 1973, **95**, 328.
[144] N. Nakamizo and M. Naruse, Jap. Kokai, 73 29 724 (Cl. 16 B81) (*Chem. Abs.*, 1973, **79**, 79 177).

* Volume 4 of these Reports (p. 446, footnote 5) underlines the ambiguity in the chemist's and biochemist's approach to the nomenclature of N^{im}-substituted histidines. A good example of this pitfall is to be found on p. 337 of the same volume, where this Reporter considered the paper of Henkart (*J. Biol. Chem.*, 1971, **246**, 2711) in terms of chemist's nomenclature. More careful consideration of the paper makes it likely that it is the biochemist's system of enumeration which was used, although it is not explicitly stated or clarified by formula. If this is the case then N^{α}-acetyl-N^{im}-2,4-dinitrophenylhistidine should be formulated as the *tele*-isomer [Ac-His(τDnp)-OH] and not the *pros*-isomer. This means that Scheme 77 is in error and should be amended accordingly.

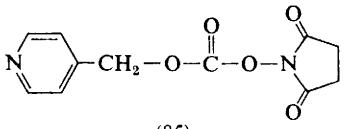
synthesis, have been described; they were prepared by the action of ammonia in methanol on the corresponding *N,O*-ditosylthreonine (or allothreonine) methyl esters.[145]

The synthesis and resolution of DL-6-nitro-2-aminocaproic acid have been described. 6-Nitro-2-aminocaproic acid derivatives can be used in peptide synthesis for the introduction of masked lysine residues.[146] A number of N^{ϵ}-protected derivatives have been prepared using edta disodium salt to destroy the intermediate copper complexes.[89] Selective N^{ϵ}-benzyloxycarbonylation of lysine can be accomplished using (7) in neutral aqueous solution without need of a copper complex.[26]

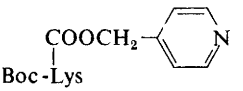
The protection of ε-amino-groups of lysine with benzyloxycarbonyl causes difficulties in solid-phase synthesis because acidolytic deprotection of N^{α}-t-butoxycarbonyl groups brings about some cleavage of the ε-protector (ε-benzyloxycarbonyl-lysine in 50% trifluoroacetic acid in dichloromethane loses 1.4% of its protecting group per hour [90] or 42% over a 20 h period;[91] the corresponding experiment using polymer-bound ε-benzyloxycarbonyl-lysine leads to a loss of 0.8% per hour). This generates a proportion of free ε-amino-groups which can form loci for subsequent chain-branching.[90, 91, 147] That this is indeed a serious problem was shown by the solid-phase synthesis of decalysylvaline (using extreme deprotection conditions – 50% trifluoroacetic acid in dichloromethane for 1 h per cycle). Chromatography revealed that the crude undecapeptide mixture contains peptides with 11—19 lysine residues.[90, 147] Accordingly, a search has been made for modified benzyloxycarbonyl groups, substituted with electron-withdrawing groups, which are more stable to deprotection cycles, but which are labile under the conditions of final deprotection.[90, 147] 3-Chlorobenzyloxycarbonyl and 2,6-dichlorobenzyloxycarbonyl groups are quite stable to deprotection cycles but are resistant to the conditions of final deprotection (in hydrogen fluoride at 0 °C for 60 min, 15% of N^{ϵ}-2,6-dichlorobenzyloxycarbonyl-lysine remains unattacked). 2-Chloro-, 2,4-dichloro-, and 3,4-dichloro-benzyloxycarbonyl groups are suitable for the side-chain protection of lysine during solid-phase synthesis, since they are stable in trifluoroacetic acid in dichloromethane, and yet completely deprotected by hydrogen fluoride after 1 h at 0 °C. Resynthesis of decalysylvaline using 2,4-dichlorobenzyloxycarbonyl side-chain protection gave a crude undecapeptide free of branched material.[147] 4-Bromo- and 2-bromobenzyloxycarbonyl groups [91] have also been advocated for lysine protection (4-bromobenzyloxycarbonyl-lysine loses 12% of its protection during a 20 h treatment with 50% trifluoroacetic acid, while the 2-bromo-derivative loses only 0.7%). A successful solid-phase synthesis of human adrenocorticotropin has been carried out using the latter protection for lysine.[93]

[145] E. Atherton and J. Meienhofer, *Z. physiol. Chem.*, 1973, **354**, 689.
[146] E. Bayer and K. Schmidt, *Tetrahedron Letters*, 1973, 2051.
[147] B. W. Erickson and R. B. Merrifield, *J. Amer. Chem. Soc.*, 1973, **95**, 3757.

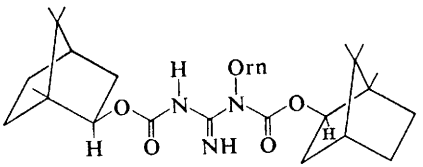
Another solution to the problem of adequately protecting the lysine side-chain is use of the 4-picolyloxycarbonyl group (18) mentioned earlier.[36] ε-4-Picolyloxycarbonyl-lysine is prepared by the reaction of lysine–copper complex with 4-picolylsuccinimidyl carbonate (85). The derivative (86) has been used for peptide synthesis by classical and solid-phase techniques.



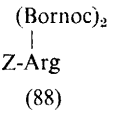
(85)



(86)



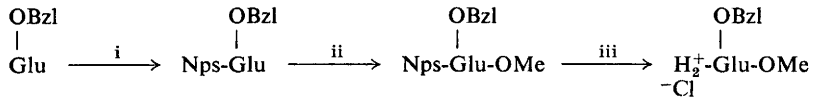
(87)



(88)

The isobornyloxycarbonyl group has been recommended for the protection of the guanidino-function of arginine. N^{δ},N^{ω}-bis-isobornyloxycarbonylarginine (87) is prepared by the action of isobornyl chloroformate on arginine in basic solution, and the corresponding N^{α}-benzyloxycarbonyl derivative (88) has been used in synthesis; although the solubility of the arginine derivative is increased, its use is complicated by marked steric hindrance.[47, 48] Arginine side-chains (protected by protonation for coupling steps) can be used as 'basic handles' to facilitate the subsequent isolation of synthetic peptides.[148]

Miscellaneous Matters relating to Protective Groups.—The γ-(*p*-bromobenzyl) ester of glutamic acid has been used in a solid-phase synthesis of the carboxy-terminal dodecapeptide of human pituitary growth hormone.[101] A simple preparation of γ-benzyl-α-methyl-L-glutamate (Scheme 42) has been reported.[149] Of a series of carboxamide protecting groups for asparagine and glutamine, the diphenylmethyl group was found to be most



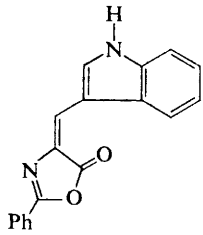
Reagents: i, NpsCl and NaOH; ii, CH_2N_2; iii, HCl in Et_2O

Scheme 42

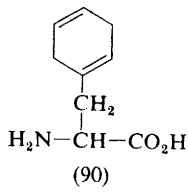
[148] D. J. Schafer and A. D. Black, *Tetrahedron Letters*, 1973, 4071.
[149] C. Daicoviciu, *Rev. Chim.* (*Roumania*), 1972, **23**, 735 (*Chem. Abs.*, 1973, **78**, 160 083).

suitable.[150] The rates of deamidation of peptides containing glutamine and asparagine in phosphate buffer at 37 °C have been studied.[151]

Hydrogenation of some antamanide analogues over platinum gave the corresponding β-cyclohexylalanine derivatives.[152] In the synthesis of luteinizing hormone-releasing hormone, during hydrogenolytic removal of the nitro-group from the arginine residue, the saturation of the tryptophan residue was observed. A control experiment showed that perhydrotryptophan could be produced from tryptophan by hydrogenation in acetic acid over palladium–charcoal for 12 h.[153] Synthesis of some dehydrotryptophan dipeptides has been accomplished using the unsaturated oxazolinone (89).[154] The use of the N^{in}-formyl group to protect the indole moiety of tryptophan from destruction during acidic deprotection cycles (using either 0.4M-hydrogen chloride in formic acid or 50% trifluoroacetic acid in dichloromethane) in solid-phase synthesis has been further evaluated.[155, 156] Parallel syntheses of peptides with and without N^{in}-formyl protection confirmed its superiority. The N^{in}-formyl group is stable under acidic conditions (including hydrogen fluoride) but can be removed by hydrazine hydrate in DMF (although it is apparently stable to



(89)



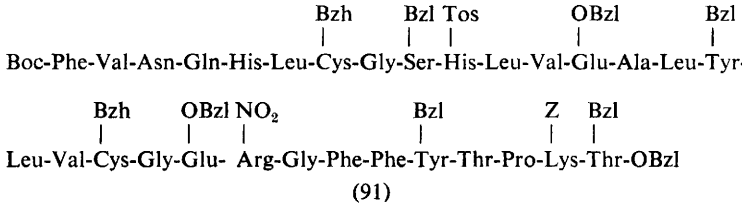
(90)

anhydrous triethylamine in DMF),[155] 0.01M-ammonium bicarbonate buffer of pH 9,[156] and liquid ammonia.[156] Under the latter two sets of conditions the formyl group undergoes an $N^{in} \rightarrow N^{\alpha}$-migration to give a contaminated product;[156] this can be partly circumvented, in the first case by using 1M-buffer, and in the second by using hydroxylamine hydrochloride as an additive.

The feasibility of using L-cyclohexa-1,4-diene-1-alanine (90) as a phenylalanine analogue in peptide synthesis has been demonstrated.[157]

[150] V. J. Hruby, F. A. Muscio, W. Brown, and P. M. Gitu, ref. 4, p. 331.
[151] A. B. Robinson, J. W. Scotchler, and J. H. McKerrow, *J. Amer. Chem. Soc.*, 1973, **95**, 8156.
[152] A. I. Miroshnikov, K. Kh. Khalilulina, N. N. Uvarova, V. T. Ivanov, and Yu. A. Ovchinnikov, *Khim. prirod. Soedinenii*, 1973, 214 (*Chem. Abs.*, 1973, **79**, 42 831).
[153] S. Bajusz, A. Turán, I. Fauszt, and A. Juhász, ref. 5, p. 93.
[154] M. Bakhra, G. S. Katrukha, and A. B. Silaev, *Khim. prirod. Soedinenii*, 1973, 280 (*Chem. Abs.*, 1973, **79**, 19 087).
[155] M. Ohno, S. Tsukamoto, S. Sato, and N. Izumiya, *Bull. Chem. Soc. Japan*, 1973, **46**, 3280.
[156] D. Yamashiro and C. H. Li, *J. Org. Chem.*, 1973, **38**, 2594.
[157] G. R. Nagarajan, L. Diamond, and C. Ressler, *J. Org. Chem.*, 1973, **38**, 621.

Final Deprotection.—Although many examples in the literature record apparently successful final deprotection using anhydrous hydrogen fluoride in the presence of anisole [see for instance the demanding example of the deprotection of the insulin B-chain (91) [138]], a number of difficulties have



(91)

come to light. Problems due to the ring migration [90, 91, 93] of the benzyl group of *O*-protected tyrosine and due to the acetamidomethylation of tyrosine [131] have already been mentioned. A series of α-aspartyl peptides was shown to have undergone rearrangement to β-aspartyl peptides during removal of protecting groups with hydrogen fluoride–anisole (1 h at 0 or −25 °C, followed by treatment for 24 h with 1% ammonium hydrogen carbonate at 40 °C). Of the peptides, aspartylhistidine gave mostly β-peptide, aspartylglycine and aspartylserine gave some β-peptide, and aspartylalanine gave only α-peptide.[158] This work is difficult to evaluate since insufficient control experiments were performed. In another case the deprotected peptide apparently incorporated the anisole scavenger.[159]

Boron tris(trifluoroacetate) (92) in trifluoracetic acid or dichloromethane can be used for final deprotection;[160] it is similar in its action to hydrogen fluoride but does not require a special apparatus or special precautions for its use. The reagent is readily formed (Scheme 43) and can be easily

$$3CF_3CO_2H + BBr_3 \longrightarrow \underset{(92)}{B(O_2CCF_3)_3} + 3HBr$$

Scheme 43

isolated. Using a moderate excess of the reagent (3 equivalents) the following protective groups can be removed within 1 hour at room temperature: *N*-benzyloxycarbonyl, *N*-t-butoxycarbonyl, *N*-*o*-nitrophenylsulphenyl, *N*-trityl, and other acid-labile protecting groups; benzyl esters (including peptidyl-polymer), benzyl-based side-chain protection of serine, threonine, and tyrosine. With more reagent (5—8 equivalents) the following groups can be removed: *N*-ω-nitro and *N*-ω-tosyl groups from arginine, and *S*-*p*-methoxybenzyl groups from cysteine. Methyl and ethyl esters, methyl ethers, and of course peptide bonds are stable. After deprotection the

[158] T. Baba, H. Sugiyama, and S. Seto, *Chem. and Pharm. Bull.* (*Japan*), 1973, **21**, 207.
[159] S. Sano and S. Kawanishi, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 46.
[160] J. Pless and W. Bauer, *Angew. Chem. Internat. Edn.*, 1973, **12**, 147.

reaction solution is evaporated at room temperature, and volatile boron salts are removed by repeated evaporation with methanol or by chromatography.

Trifluoromethanesulphonic acid (93) is another reagent which resembles hydrogen fluoride and can be used for final deprotection.[161] The reagent

$$CF_3SO_3H$$

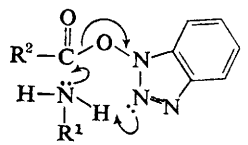
(93)

consists of (93) dissolved in trifluoroacetic acid or dichloromethane with anisole. Using 5—10 equivalents of reagent at 20 °C, *N*-t-butoxycarbonyl, *N*-*p*-methoxybenzyloxycarbonyl, and *N*-*o*-nitrophenylsulphenyl groups are cleaved quantitatively within 3—5 min; the *S*-*p*-methoxybenzyl group of cysteine within 15 min; t-butyl-based and benzyl-based side-chain protection (of serine, threonine, and aspartic and glutamic acids), *N*-benzyloxycarbonyl (including tryptophan and tyrosine) within 30 min; N_{im}-tosyl group of histidine in 60 min; at 40 °C the ω-tosyl group of arginine is cleaved in 60 min. Deprotection of *O*-benzyl-tyrosine is unsuccessful because the rearrangement product, 3-benzyltyrosine, predominates. The recovery of methionine from benzyloxycarbonyl-methionine is poor (improvement is to be expected if the additives anisole, methyl ethyl sulphide, or dithiothreitol are used). Under the most drastic conditions employed, ω-nitro-arginine proved to be partially resistant to the reagent. The new reagents (92) and (93) appear to hold considerable potential.

Pyridinium salts (chloride, bromide, or perchlorate) can be used to cleave *N*-trityl, *N*-(1-biphenylyl-1-methyl)ethoxycarbonyl, and *N*-*o*-nitrophenylsulphenyl protecting groups selectively in the presence of benzyl- and t-butyl-based protection.[162]

3 Formation of the Peptide Bond

Activated Esters.—It is now firmly established that 1-hydroxybenzotriazole catalyses dicyclohexylcarbodi-imide-mediated peptide condensation reactions and suppresses racemization. This behaviour is explained by the formulation of a complex between the intermediate benzotriazolyl ester and the incoming nucleophile (94).[163] The striking observation that



(94)

[161] H. Yajima, N. Fujii, H. Ogawa, and H. Kawatani, *J.C.S. Chem. Comm.*, 1974, 107.
[162] H. Klostermeyer and E. Schwertner, ref. 5, p. 108.
[163] W. König and R. Geiger, ref. 4, p. 343.

hydroxybenzotriazole markedly catalyses the aminolysis of certain active esters (*e.g.* Scheme 44) (principally *p*-nitrophenyl and 2,4,5-trichlorophenyl

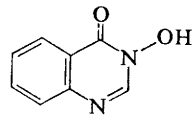
Boc-Leu-Phe-OX + HCl; H-Val-OBu^t $\xrightarrow{i}$ Boc-Leu-Phe-Val-OBu^t

—OX	*Catalyst*	*Reaction time*	D-Phe *Content*
—ONp	—	70 h	<2%
—ONp	HOBt	5 min	<2%
—OPcp	—	70 h	40%
—OPcp	HOBt	5 min	27%
—OTcp	—	70 h	<2%
—OTcp	HOBt	10 min	<2%

Reagents: i, Et—N⟨ ⟩O, Me_2NCHO

Scheme 44

esters)[163, 164] in polar solvents, such as dimethylformamide or dimethylacetamide, promises to increase the utility of these esters. *N*-Hydroxy-additives with a pK_a close to that of acetic acid are effective, the most effective being 1-hydroxybenzotriazole (pK_a = 4.00), 1-hydroxy-2-pyridone (pK_a = 4.08), and 3-hydroxy-4-oxo-3,4-dihydroquinazoline (95) (pK_a = 4.10). [*N*-hydroxysuccinimide (pK_α = 4.04) has, in general, a much weaker



(95)

effect, as do additives which are much more acidic, or less acidic, than acetic acid; furthermore the effect of 1-hydroxybenzotriazole on *N*-succinimidyl ester couplings is only marginal.] The nature of the solvent plays a crucial role, since in THF the additive, in general, inhibits aminolysis (this represents a reversal with respect to the behaviour of the additives acetic acid, imidazole, pyrazole derivatives, 2-hydroxypyridine, and 1,2,4-triazole, which are more effective in non-polar solvents[165]). Rate studies and physical measurements indicate that in the catalysed process an equilibrium exists between the starting active ester and the intermediate benzotriazolyl ester, which favours the former ester in acidic medium and the latter ester in basic medium.[164] In some cases the catalysed reaction is complete in the time it takes to dissolve the components! Application of the Weygand gas chromatographic racemization test (Scheme 44) shows that

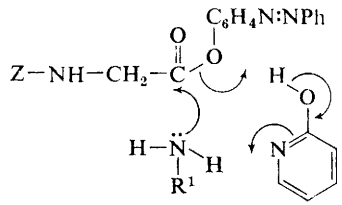
[164] W. König and R. Geiger, *Chem. Ber.*, 1973, **106**, 3626.
[165] J. Rudinger, quoted in refs. 163 and 164.

the additive does not cause racemization. Salts of *N*-hydroxy-compounds can be used as additives, thus enabling the free base of the amino-component to be liberated from its salt directly. (1-Hydroxybenzotriazole, however, can cleave *o*-nitrophenylsulphenamides.[165]) The successful use of the additive (95) in the solid-phase synthesis of the peptide (96) using 2,4,5-trichlorophenyl esters has been reported.[166]

It has been claimed that, in the coupling between *p*-nitrophenyl *N*-benzyloxycarbonylglycinate and glycine t-butyl ester in benzene, the starting materials, the dipeptide product, and the salt formed from added *p*-nitrophenol with the amino-component all catalyse the process.[167] In the 2-hydroxypyridine-catalysed coupling of *N*-benzyloxycarbonylglycine *p*-(phenylazo)phenyl ester with glycine t-butyl ester in benzene, the observed

H-Gly-Ile-Val-Glu-Gln-OH

(96)



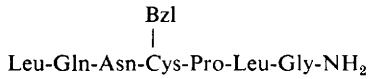
(97)

second-order rate coefficient increases with time, probably owing to product autocatalysis. A bifunctional mechanism (97) is postulated for the catalysis, since pyridine and 1-methyl-2(1*H*)-pyridone are catalytically inactive.[168] The kinetics of some model couplings of phenyl, *p*-nitrophenyl, *p*-bromophenyl, and 3,4-dinitrophenyl esters in non-polar solvents have been interpreted in terms of a highly polar tetrahedral transition state, whose decomposition is the rate-controlling step.[169] The kinetics of the acetic acid-catalysed aminolysis of phenyl esters, and of the intramolecular catalysis of aminolysis in succinic acid monophenyl ester, have been studied.[170]

Active esters would seem to be ideal acylating agents for solid-phase peptide synthesis, using minimal side-chain protection; however, they have only been used generally for the introduction of asparagine and glutamine residues (as their N-protected *p*-nitrophenyl esters). Active esters have not

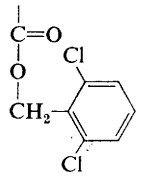
[166] H. Wissmann, W. König, and R. Geiger, ref. 5, p. 158.
[167] V. V. Kosmynin, L. M. Litvinenko, and Yu. A. Sharanin, *Zhur. org. Khim.*, 1973, **9**, 745 (*Chem. Abs.*, 1973, **79**, 5579).
[168] L. M. Litvinenko, Yu. A. Sharanin, L. P. Drizhd, and V. A. Savielova, *J. Gen. Chem. (U.S.S.R.)*, 1973, **43**, 1808.
[169] V. V. Kosmynin, Yu. A. Sharanin, L. M. Litvinenko, and V. A. Savielova, *Reakts. spos. org. Soedinenii*, 1972, **9**, 977 (*Chem. Abs.*, 1973, **79**, 79 160).
[170] F. Muzalewski, ref. 3, p. 39.

been used more widely because they are insufficiently reactive in the heterogeneous system. In a re-examination of the problem it has been shown that although *o*-nitrophenyl esters are only marginally more reactive than *p*-nitrophenyl esters in solution, they are more efficient in solid-phase synthesis [171] (using the solid-phase technique with 0.02 mol l^{-1} concentration of the reactants in ethyl acetate, the various *N*-t-butoxycarbonyl-amino-acid *o*-nitrophenyl esters had half-reaction times of 150—700 min) and their rates of aminolysis are only slightly solvent-dependent.[172] The heptapeptide amide (98) has been prepared on a benzhydrylamine resin both by the *o*-nitrophenyl ester method and by dicyclohexylcarbodi-imide. The peptide prepared by the former method possesses superior purity to that from the latter, but each coupling cycle requires *ca.* 2—15 h.[171] The



(98)

Leu-Gln-Lys-Leu-Leu-Gln-Gly-Leu-Val



(99)

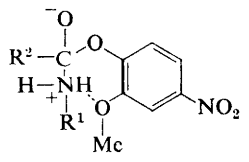
new active esters are best prepared from the carboxylic acid, *o*-nitrophenol, and dicyclohexylcarbodi-imide in pyridine solution.[173] They have also been applied to the stepwise synthesis of the *C*-terminal nonapeptide amide of secretin (99) in solution using an '*in situ*' technique. In this technique the whole chain-lengthening process is carried out in a centrifuge tube; after each operation (coupling, and deprotection) the solvent is evaporated and replaced by a 'non-solvent' [*i.e.* a solvent in which the required peptide is insoluble but in which the co-products and side-products are soluble (ethyl acetate)], which enables the peptide to be recovered by centrifugation.[173]

It has been shown that pentafluorophenyl activated esters are sufficiently reactive to be useful in solid-phase synthesis.[174] Optically pure pentafluorophenyl esters of N-protected amino-acids can be prepared by the dicyclohexylcarbodi-imide method,[175] and of N-protected peptides by the crystal-

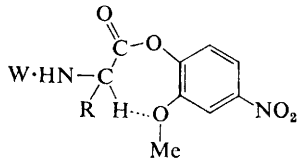
[171] M. Bodanszky and K. W. Funk, *J. Org. Chem.*, 1973, **38**, 1296.
[172] M. Bodanszky, R. J. Bath, A. Chang, M. L. Fink, K. W. Funk, S. M. Greenwald, and Y. S. Klausner, ref. 4, p. 203.
[173] M. Bodanszky, K. W. Funk, and M. L. Fink, *J. Org. Chem.*, 1973, **38**, 3565.
[174] K. Kovács and B. Penke, ref. 5, p. 187.
[175] L. Kisfaludy, M. Löw, O. Nyéki, T. Szirtes, and I. Schön, *Annalen*, 1973, 1421.

line adduct of pentafluorophenol (3 moles) and dicyclohexylcarbodi-imide (1 mole).[175] They are stable and possess a high reactivity towards aminolysis in solution; they are polarimetrically stable in dioxan, ethyl acetate, dichloromethane, and chloroform, but not in THF and DMF. Providing that excess base and the latter two solvents are avoided, they can be used in coupling situations without racemization (Bodanszky racemization test).[175]

The preparation and use of 2-pyridyl esters,[176] and some problems concerning the stability of *O*-benzyloxycarbonyltyrosine-containing peptide *o*-hydroxyphenyl esters in the presence of triethylamine,[177] have received further mention. 4-Nitroguaiacyl esters are slightly more reactive to aminolysis than the corresponding *p*-nitrophenyl esters [this is ascribed to intramolecular base catalysis; see (100)], and yet undergo racemization less readily [this is ascribed to hydrogen bonding; see (101)].[178] A study of the aminolysis of esters of *N*-hydroxysulphon-amides and -imides (102) and



(100)



(101)

$$ZNHCH_2COONHSO_2C_6H_4X(p)$$

(102)

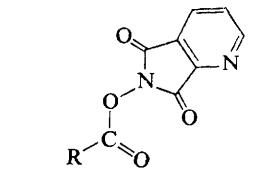
$$ZNHCH_2COON[SO_2C_6H_4X(p)]_2$$

(103)

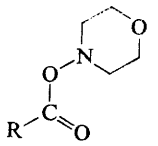
(103) revealed that the former compounds are suitable for peptide synthesis, but the latter derivatives, although more reactive, undergo nucleophilic attack at sulphur as well as at the carbonyl group, so that mixtures of amide and sulphonamide derivatives are obtained.[179] Further use of *p*-(acetylsulphamoyl)phenyl esters in peptide synthesis has been reported.[180]

Other relevant papers have been concerned with active esters of '*N*-hydroxyquinolinimide'[181] [presumably this refers to esters of structure (104)], of *N*-hydroxymorpholine (105),[182] of *N*-hydroxy-5-norbornene-2,3-dicarboximide (106),[183] of ketoximes of dicarbonyl compounds (107),

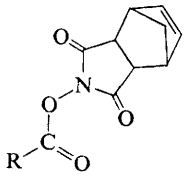
[176] A. S. Dutta and J. S. Morley, ref. 3, p. 21.
[177] Y. Trudelle and G. Spach, ref. 3, p. 36.
[178] K. Bankowski and S. Drabarek, ref. 3, p. 31.
[179] G. Kupryszewski and J. Przybylski, ref. 5, p. 114.
[180] B. Tomicka, G. Kupryszewski, and F. Karczynski, *Roczniki Chem.*, 1973, **47**, 185 (*Chem. Abs.*, 1973, **79**, 5578).
[181] M. Dzieduszycka, L. Lubiewska-Nakonieczna, and E. Taschner, ref. 3, p. 28.
[182] K. Okamoto and S. Shimamura, *Yakugaku Zasshi*, 1973, **93**, 333 (*Chem. Abs.*, 1973, **79**, 19 081).
[183] M. Fujino, S. Kobayashi, M. Obayashi, S. Shinagawa, and T. Fukuda, Ger. Offen. 2 311 786 (Cl. C 07*c*) (*Chem. Abs.*, 1973, **79**, 146 854).



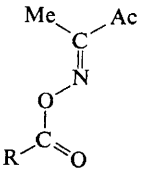
(104)



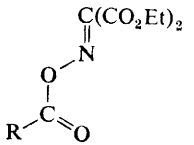
(105)



(106)



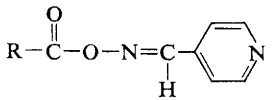
(107)



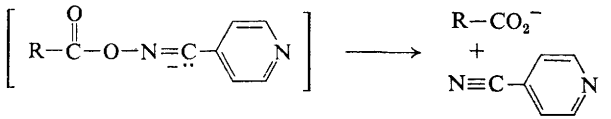
(108)

(108),[184] and of 4-pyridine-aldoxime (109).[185] The latter esters possess a convenient reactivity to aminolysis in the presence of twelve equivalents of acetic acid (protonation of the ring nitrogen atom), and the product of coupling can be separated from reaction co-products and excess acylating agent, either by extracting them into an acidic phase or by treating the reaction mixture with a base, when the decomposition products can be separated by acid–base extractions (Scheme 45).

4-Iminoxy activated esters of 5-(benzyloxycarbonylglycyl)imino-3-methyl-4-oximino-1-phenylpyrazoline (110) (see Vol. 4 of these Reports, p. 341) have been used to prepare lysozyme-(2—11)-decapeptide (111) and



(109)



Reagent: i, Et_3N

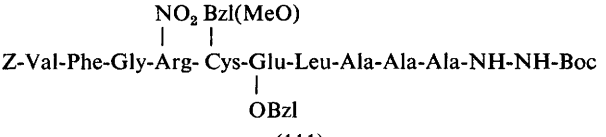
Scheme 45

[184] M. Fujino, C. Hatanaka, and T. Nishimura, Jap. Kokai 73 21 922 (Cl. C 07*cb*) (*Chem. Abs.*, 1973, **79**, 92 593).

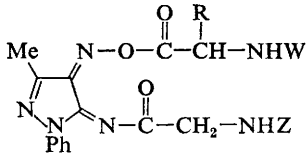
[185] P. Ya. Romanovski, V. E. Muizhnieks, and G. I. Chipens, ref. 5, p. 97.

bradykinin.[186] The pyrazole derivative (112) has been used for the acylation of amino- and hydroxy-groups.[187]

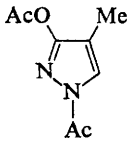
Further 2-pyrimidinyl thiolesters (113; R^1 = H) and also 4,6-dimethyl 2-pyrimidinyl thiolesters (113; R^1 = Me) have been used for acylation.[188]



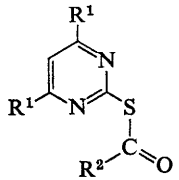
(111)



(110) W = Z, Boc, or Nps

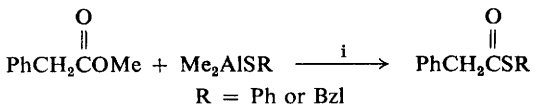


(112)



(113)

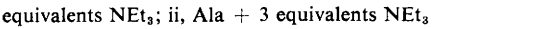
Cholineselenol esters (69) have been mentioned in an earlier section.[112] Mention has been made of a simple one-step procedure for activating the carboxy-group of an ester with reagents derived from monothiols and trimethylaluminium (Scheme 46).[88]



Reagents: i, CH_2Cl_2, 2.5 h, 25 °C

Scheme 46

Enol esters derived from α-chlorobenzylideneaniline (Scheme 47) have been mentioned.[189] The reaction of *N*-t-butyl-5-methylisoxazolium perchlorate (114) (a stable reagent) with carboxylate anion yields an *N*-t-butyl-β-acyloxycrotonamide (116). These enol ester acylating agents are,



Reagents: i, 3 equivalents NEt_3; ii, Ala + 3 equivalents NEt_3

Scheme 47

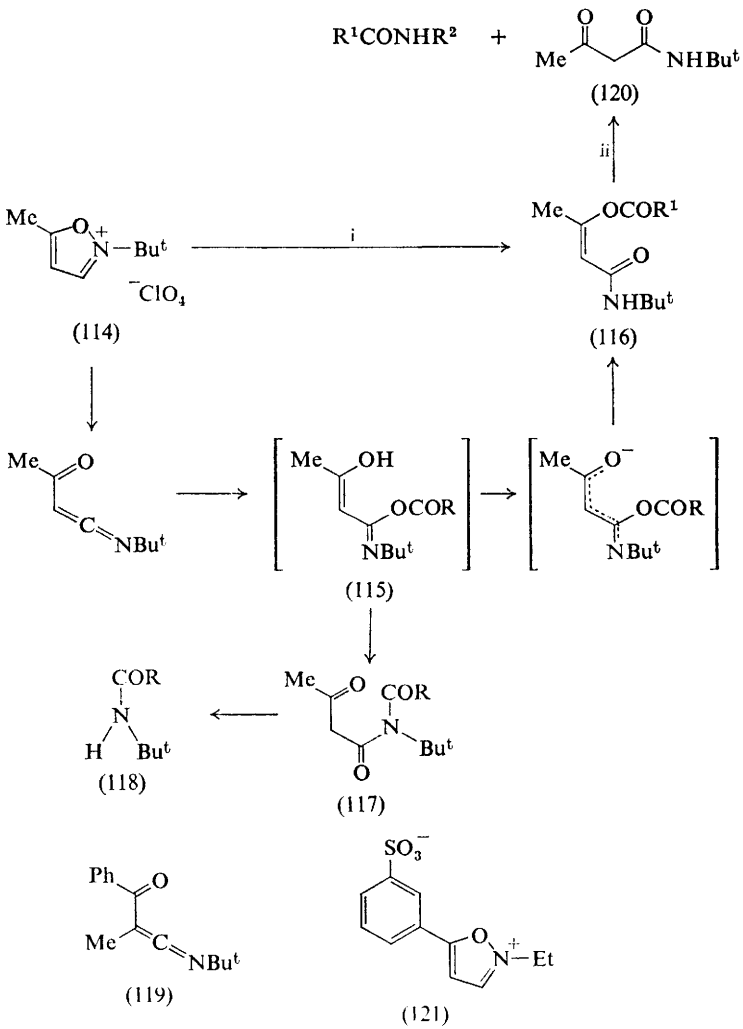
[186] M. Guarneri, P. Giori, R. Ferroni, A. Guggi, E. Menegatti, R. Tomatis, and C. A. Benassi, ref. 3, p. 22.

[187] K. Arakawa, Jap. Kokai 73 16 881 (Cl. C 07*cd*) (*Chem. Abs.*, 1973, **79**, 66 818).

[188] T. Nagasawa, K. Kuroiwa, and K. Narita, Ger. Offen. 2 246 334 (Cl. C 07*d*) (*Chem. Abs.*, 1973, **78**, 160 109).

[189] J. S. P. Schwarz, U.S.P. 3 732 199 (Cl. 260—112.5; C 07*c*) (*Chem. Abs.*, 1973, **79**, 19 125).

in general, stable and crystalline. However, their formation can be accompanied by the formation of *N*-acyl t-butylamides (118),[190, 191] which appear to arise *via* intermediate diacylamides (117), which cannot themselves arise from the enol esters and must, therefore, arise directly by a



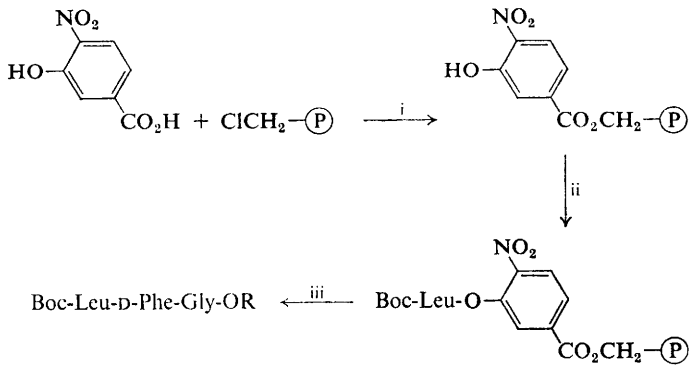
Reagents: i, $R^1CO_2^-$ $H\overset{+}{N}Et_3$, MeCN; ii, H_2NR^2

Scheme 48

[190] D. J. Woodman and A. I. Davidson, *J. Org. Chem.*, 1970, **35**, 83.
[191] D. J. Woodman and A. I. Davidson, *J. Org. Chem.*, 1973, **38**, 4288.

competing pathway from the adduct (115)* *via* an imino-anhydride rearrangement.[191] (These pathways are summarized in Scheme 48.) However, this side-reaction can be diminished by the use of excess (114) in wet acetonitrile, thereby providing a generally useful method for preparing these active esters.[190] The reaction succeeds with acids containing an unprotected hydroxy-group (serine and tyrosine) and an unprotected carboxamide group (glutamine), but fails with asparagine (side-chain dehydration being observed) and yields a mixture of the oxazolinone and enol ester with benzoyl-leucine. No significant differences are observed when coupling reactions are carried out with purified enol esters or esters generated *in situ*. Reaction of the amino-component with the coupling by-product (120) to form an enamine is only a problem in the case of highly hindered carboxy-components. It would appear that the reagent (114) possesses no great advantages over the original zwitterionic reagent (121).[191]

A few polymeric active esters have been described. A polymeric reagent, similar to the previously described poly(4-hydroxy-3-nitrostyrene), has been produced by forming an ester linkage between 3-hydroxy-4-nitrobenzoic acid and chloromethylated polystyrene.[192] The derived polymeric active ester has been used for limited peptide synthesis (Scheme 49).



Reagents: i, NEt_3 in EtOH; ii, Boc-Leu-OH, $C_6H_{11}N{:}C{:}NC_6H_{11}$, CH_2Cl_2; iii, HCl; H-D-Phe-Gly-OR, NEt_3, CH_2Cl_2 [Ⓟ = polystyrene–2% divinylbenzene copolymer]

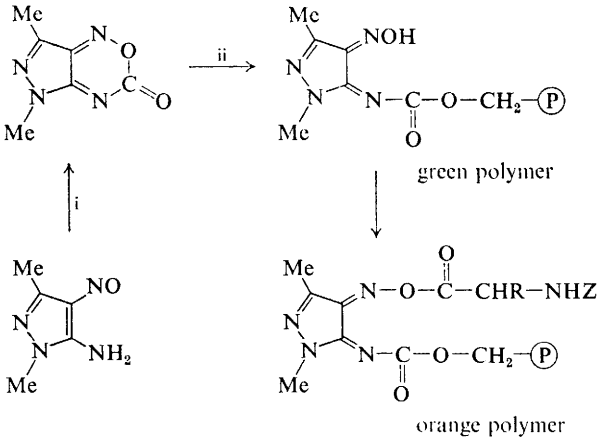
Scheme 49

Another polymeric reagent incorporates the 'safety catch' principle. *p*-Mercaptophenol is incorporated into polystyrene, an N-protected amino-acid is anchored to this matrix, and this is used for successive deprotection and chain-elongation cycles. At the end of the synthesis the anchor ester

[192] G. T. Panse and D. A. Laufer, ref. 2, p. 3.

* Evidence for the existence of this postulated reactive intermediate comes from the appearance of acid anhydrides, by intermolecular trapping of the intermediate derived from the disubstituted ketenimine (119) with unconsumed carboxylic acid.

bond can be cleaved with ammonia in methanol or hydrazine in DMF, or it is further activated to nucleophilic attack by oxidation to the sulphone, thus enabling cyclization or chain elongation to be carried out.[193] Further mention of the preparation and use of activated esters of copoly(ethylene-*N*-hydroxymaleimide) has been made.[194, 195] Polymeric active esters based on 1,3-dimethyl-4-nitroso-5-aminopyrazole have now been prepared (Scheme 50). These active esters couple readily with amino-components

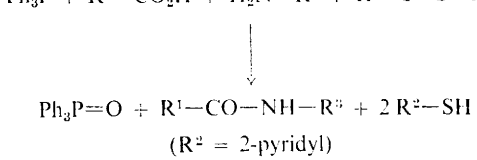


Reagents: i, $COCl_2$, ii, Ⓟ—CH_2OH; iii, Z-AA-OH or Boc-AA-OH and $C_6H_{11}N{:}C{:}N_6CH_{11}$. Available reagent sites 0.2—1.0 mmol g^{-1}

Scheme 50

(reaction times 15—60 min), the coupling being followed visually as the orange active polymer reverts to the green exhausted polymer.[196]

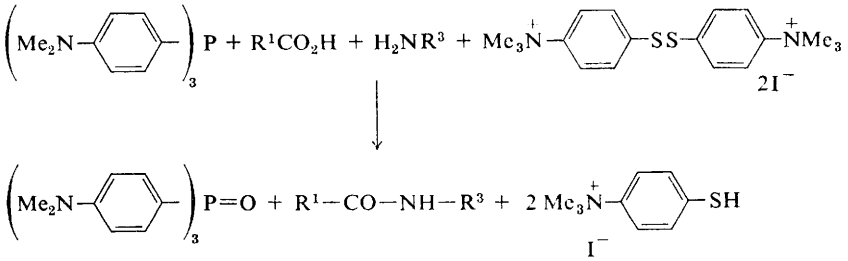
Coupling Methods involving Phosphorus Derivatives.—A general review of new organic reactions based on organosulphur compounds discusses the 'oxidation–reduction' method of peptide bond formation.[197] This method, summarized in Scheme 51, has been used for a synthesis of luteinizing-



Scheme 51

[193] E. Flanigan and G. R. Marshall, ref. 2, p. 7.
[194] M. Fridkin, A. Patchornik, and E. Katchalski, ref. 2, p. 19.
[195] M. Narita, T. Teramoto, and M. Okawara, *Bull. Chem. Soc. Japan*, 1972, **45**, 3149.
[196] M. Guarneri, R. Ferroni, P. Giori, and C. A. Benassi, ref. 4, p. 213.
[197] T. Mukaiyama, *Internat. J. Sulfur. Chem., Part B*, 1972, **7**, 173.

hormone-releasing hormone by a fragment-condensation strategy.[198] A study has been conducted of the effect on peptide synthesis of substituting the phenyl rings of the triphenyl phosphite component by electron-attracting or electron-releasing groups, using the 'oxidation–reduction' method. The most satisfactory results were obtained using tri(*p*-bromophenyl) phosphite,[199] which enabled the Young test peptide to be obtained optically pure in very high yield. Recent patents involve the use of tri-n-butylphosphine instead of triphenylphosphine, and *N*-(phenylthio)-succinimide, t-butyl 2,4-dinitrophenylsulphenate, *N*-(phenylthio)phthalimide, 2,4-dinitrophenyl disulphide, or 2-nitrophenyl disulphide with dihydropyran, n-butyl vinyl ether, and ethyl acrylate, instead of 2,2′-dipyridyl disulphide.[200–202] The reagents used in the 'oxidation–reduction' method have been modified so that co-products of coupling can be removed readily by washing with dilute aqueous acid (Scheme 52).[203] The modified



Scheme 52

method may be used without protecting the side-chain functional groups of serine, tyrosine, asparagine, and glutamine, and it is reported to be free of racemization.

The method of peptide synthesis using triphenyl phosphite and imidazole has been reviewed.[204] The Izumiya test peptide was obtained with an optical purity of 95.8% with respect to the alanyl residue when acetonitrile was used as solvent, 95.4% with dioxan, 92.6% with DMF, and 88.6% with ethyl acetate.[205] The method is applicable to the majority of amino-acid residues (suitably side-chain-protected where necessary), but gives low yields with proline, asparagine, and glutamine residues, and fails in the

[198] R. Matsueda, H. Maruyama, and E. Kitazawa, Jap. Kokai 73 40 769 (Cl. 16 E362) (*Chem. Abs.*, 1973, **79**, 79 194).
[199] M. Ueki, S. Takashi, A. Hayashida, and T. Mukaiyama, *Chem. Letters*, 1973, 733.
[200] T. Mukaiyama, M. Ueki, R. Matsueda, and H. Maruyama, Jap. Kokai 73 23 042 (Cl. C 07*c*, B 01*j*) (*Chem. Abs.*, 1973, **79**, 105 579).
[201] T. Mukaiyama, M. Ueki, R. Matsueda, and H. Maruyama, Jap. Kokai 73 23 041 (Cl. C 07*c*, B 01*j*) (*Chem. Abs.*, 1973, **79**, 105 580).
[202] T. Mukaiyama, M. Ueki, R. Matsueda, and H. Maruyama, Jap. Kokai 73 23 043 (Cl. C 07*c*, B 01*j*) (*Chem. Abs.*, 1973, **79**, 105 581).
[203] G. I. Shelykh, G. P. Vlasov, and Yu. V. Mitin, *J. Gen. Chem.* (*U.S.S.R.*), 1973, **43**, 367.
[204] Yu. V. Mitin, A. T. Gudkov, N. P. Zapevalova, and E. E. Maximov, ref. 5, p. 57.
[205] Yu. V. Mitin and E. E. Maximov, *J. Gen. Chem.* (*U.S.S.R.*), 1973, **43**, 199.

cases of cystine and ω-nitro-arginine (side-reactions intervene). Successful syntheses of (122) and (123) of good optical purity have been reported.[204]

Amide bond formation with a series of trisdimethylaminophosphonium salts (124a—g) has been evaluated.[206] With benzoic acid, compounds

Z-Glu-His-Phe-Arg-Trp-Gly-OH (122)

H-Trp-Met-Asp-Phe-NH_2 (123)

$(NMe)_3\overset{+}{P}—X,\ PF_6^-$

(124) (a) X = Br
(b) X = CN
(c) X = NCS
(d) X = N_3
(e) X = CCl_3
(f) X = OPh
(g) X = O—C$_6$H$_4$—NO_2

(124a—c) give rise to benzoic anhydride, while (124d) yields benzoyl azide. In the Young test, reagent (124b) causes complete racemization, while reagent (124d) yields benzoyl-L-leucylglycine ethyl ester with an optical purity of 98.5%. This reagent (124d) can be prepared according to Scheme 53.[206]

$$P(NMe_2)_3 \xrightarrow{i} (NMe_2)_3\overset{+}{P}—Br,\ PF_6^- \xrightarrow{ii} (NMe_2)_3\overset{+}{P}—N_3,\ PF_6^-$$

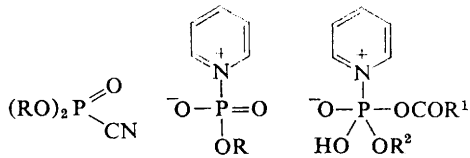
(124a) (124d)

Reagents: i, Br_2–Et_2O at 0 °C, followed by H_2O–KPF_6; ii, NaN_3

Scheme 53

Further mention has been made of the use of diphenylphosphoryl azide and also of diphenylphosphoryl cyanide (125; R = Ph) in peptide bond formation.[207] Diethylphosphoryl cyanide (125; R = Et) is also a useful reagent for racemization-free amide bond formation.[208] It is prepared by the reaction of triethyl phosphite with cyanogen bromide; coupling is accomplished by mixing the carboxy- and amino-components with the reagent in the presence of two equivalents of triethylamine (apparently it is necessary to generate the carboxylate anion). In the Young test, benzoyl-L-leucylglycine ethyl ester was obtained in very high yield, 96% optically pure. A detailed study of the mechanism of formation of simple amides, from carboxylic acids and amines, mediated by the oxidation mixture prepared from phosphorous acid (or its mono-, di-, or tri-esters) and mercury(II) chloride in pyridine solution, is consistent with the proposed intermediacy of *N*-phosphonium salts of pyridine (126), and *N*-acyloxyphosphonium salts of pyridine (127).[209, 210]

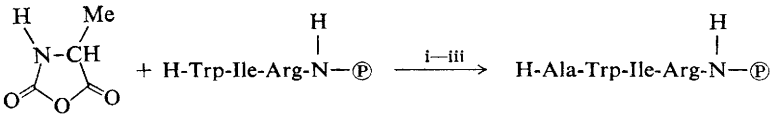
[206] B. Castro and J. R. Dormoy, *Tetrahedron Letters*, 1973, 3243.
[207] S. Yamada and T. Shiori, Jap. Kokai 73 67 202 (Cl. 16 A6) (*Chem. Abs.*, 1974, **80**, 27 475).
[208] S. Yamada, Y. Kasai, and T. Shiori, *Tetrahedron Letters*, 1973, 1595.
[209] N. Yamazaki and F. Higashi, *Bull. Chem. Soc. Japan*, 1973, **46**, 1235.
[210] N. Yamazaki and F. Higashi, *Bull. Chem. Soc. Japan*, 1973, **46**, 1239.



(125) (126) (127)

It has been postulated that a number of coupling procedures involving phosphorus compounds (in various oxidation states) may proceed *via* acyloxyphosphonium salts as intermediates. Such species may be detectable by ^{13}C–^{31}P couplings using heteronuclear magnetic resonance,[211] but were not observed in the reaction examined.

N-Carboxy-anhydrides.—A study of the controlled stepwise synthesis of peptides in aqueous medium using *N*-carboxy-anhydrides (and a volatile organic base to establish and maintain the pH) has concentrated on the influence of the stability of the *N*-carboxy-anhydride, the pK_a of the amine, and the stability of the resulting carbamate.[212] Stepwise controlled synthesis of a model tetrapeptide on a water-soluble polyethylenimine support by the *N*-carboxy-anhydride method has been investigated (Scheme 54).[213]



Reagents: i, pH 10.2, 0 °C; ii, pH 3; iii, pH 7.0, diafiltration
[Ⓟ = polyethylenimine of M = 30 000—40 000; arginine loading 2.25 mmol g^{-1} (18 mol%)]

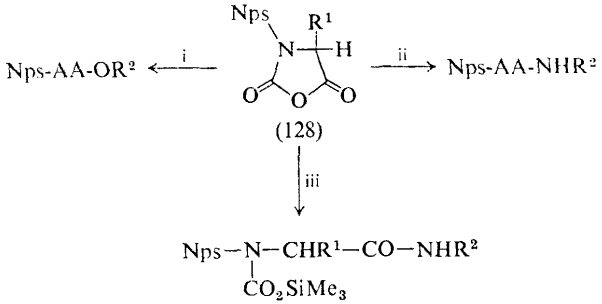
Scheme 54

The finished peptide was removed from the polymer enzymically. *N,O*-Bis(trimethylsilyl)serine trimethylsilyl ester (and the corresponding threonine derivative) react with phosgene in THF to yield the corresponding *N,O*-bis(trimethylsilyl)-amino-acid-*N*-carboxy-anhydrides. In the controlled synthesis of dipeptides the trimethylsilyl groups of these derivatives are hydrolysed simultaneously with the formation of the peptide bond.[214]

In order to overcome the general instability of *N*-carboxy-anhydrides in repetitive synthesis a number of N-protected Leuch's anhydrides have been investigated. *N*-Carboxy-anhydrides can be quantitatively acylated with *o*-nitrophenylsulphenyl chloride in anhydrous solvents containing triethylamine or *N*-methylmorpholine, yielding *N*-*o*-nitrophenylsulphenyl-*N*-carboxy-anhydrides (128), which are crystalline, more stable than the

[211] A. J. Bates, G. W. Kenner, R. Ramage, and R. C. Sheppard, ref. 5, p. 124.
[212] E. M. Grovestine, J. R. Langlois, and R. E. Williams, *Canad. J. Chem.*, 1973, **51**, 1284.
[213] H. Blecher and P. Pfaender, *Annalen*, 1973, 1263.
[214] R. Wies and P. Pfaender, *Annalen*, 1973, 1269.

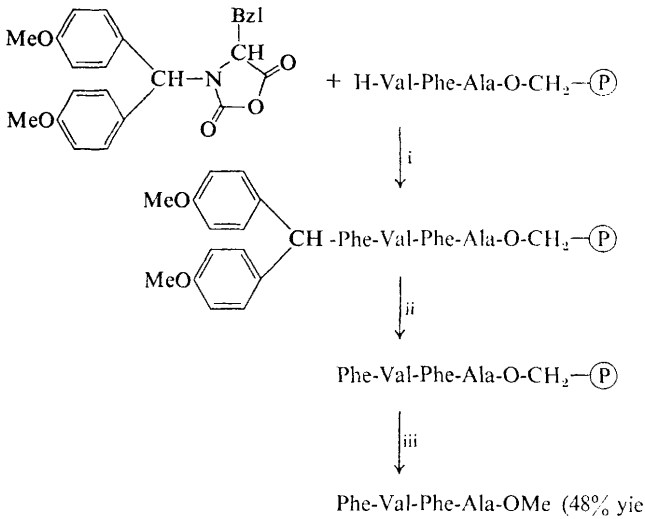
parent *N*-carboxy-anhydride, and do not polymerize.[215] The derivatives (128) react with alcohols, amines, and *N*-silylated amines as shown in Scheme 55.



Reagents: i, R^2OH; ii, R^2NH_2; iii, $R^2NHSiMe_3$

Scheme 55

N-Benzyl-, *N*-2,4-dimethoxybenzyl-, *N*-2,4,6-trimethoxybenzyl-, *N*-benzhydryl-, *N*-9-xanthenyl-, and *N*-4,4′-dimethoxybenzhydryl-amino-acids can be converted into the corresponding crystalline stable *N*-alkyl-*N*-carboxy-anhydrides using phosgene in THF.[216] Limited solid-phase peptide synthesis has been carried out with *N*-4,4′-dimethoxybenzhydryl-*N*-



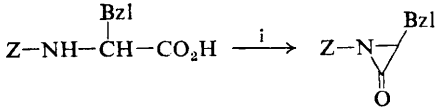
Reagents: i, 2—4 h, 55 °C in benzene or chloroform; ii, anhydrous CF_3CO_2H, 30 min; iii, NEt_3 in MeOH

Scheme 56

[215] H. R. Kricheldorf, *Angew. Chem. Internat. Edn.*, 1973, **12**, 73.
[216] J. Halstrøm and K. Kovács, ref. 5, p. 173.

carboxy-anhydrides according to Scheme 56.[216] The *N*-carboxy-anhydrides are stable to polymerization, but they can only be prepared in low yield (30—40%), nevertheless their application in solid-phase work appears promising. The synthesis of a mixture of erratic sequence peptides by the *N*-carboxy-anhydride method has been described.[217]

Coupling Methods involving Other Cyclic Amino-acid Derivatives.— Optically active 3-substituted 1-benzyloxycarbonylaziridin-2-ones, some of which are crystalline, can be prepared by the action of phosgene (the preferred procedure, although thionyl chloride or phosphoryl chloride can be used) on the corresponding *N*-benzyloxycarbonyl-amino-acid under very carefully controlled conditions (Scheme 57).[218] The nature of the



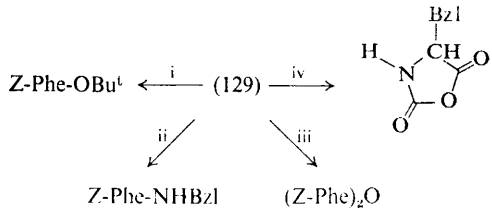
(129)

Reagents: i, THF, $COCl_2$ (1 equivalent), Et_2O, NEt_3 (2 equivalents), 90 min, −20 °C (crude yield 73%, quantitative!)

Scheme 57

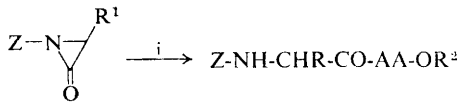
N-protecting group exercises a marked influence on the ability to form the aziridinone: in the case of benzyloxycarbonyl and *p*-bromo- and *p*-chlorobenzyloxycarbonyl, aziridinones are produced cleanly; with tosyl protection, less cleanly; with *p*-methoxybenzyloxycarbonyl, t-butoxycarbonyl, and t-amyloxycarbonyl protection they cannot be detected at all; with benzoyl and acetyl protection, oxazolinones result; and with benzyl protection the *N*-benzyl-*N*-carboxy-anhydride is produced. The aziridinones were shown to be optically intact by conversion into the corresponding amino-acids (although the paper does not quote actual data), and the evidence confirming their structures is discussed.[218] The reactions of (129) are summarized in Scheme 58; they proceed with retention of configuration.[219] 3-Substituted-*N*-1-benzyloxycarbonyl-aziridinones can be used to prepare optically pure peptides in good yield (Scheme 59), either using the isolated aziridinone or the aziridinone generated *in situ*;[219] the method appears to be particularly effective with sterically hindered components. *N*-1-Benzyloxycarbonylaziridinone itself is too unstable to be isolated, and the derivative of proline cannot be formed. Coupling can also be accomplished in aqueous solution using the sodium salt of the amino-component; in this case the symmetrical anhydride may be the acylating species. This new method of synthesis depends for its success

[217] P. Pfaender, E. Kuhnle, B. Krahl, A. Backmansson, G. Gnauck, and H. Blecher, *Z. physiol. Chem.*, 1973, **354**, 267.
[218] M. Miyoshi, *Bull. Chem. Soc. Japan*, 1973, **46**, 212.
[219] M. Miyoshi, *Bull. Chem. Soc. Japan*, 1973, **46**, 1489.



Reagents: i, Bu^{t}OH or Bu^{t}O$^-$; ii, BzlNH$_2$; iii, H$_2$O; iv, H$_2$–Pd/C

Scheme 58



Reagents: i, THF, HCl; H—AA—OR^2 in $CHCl_3$ plus NEt_3 added at −20 °C, 1 h

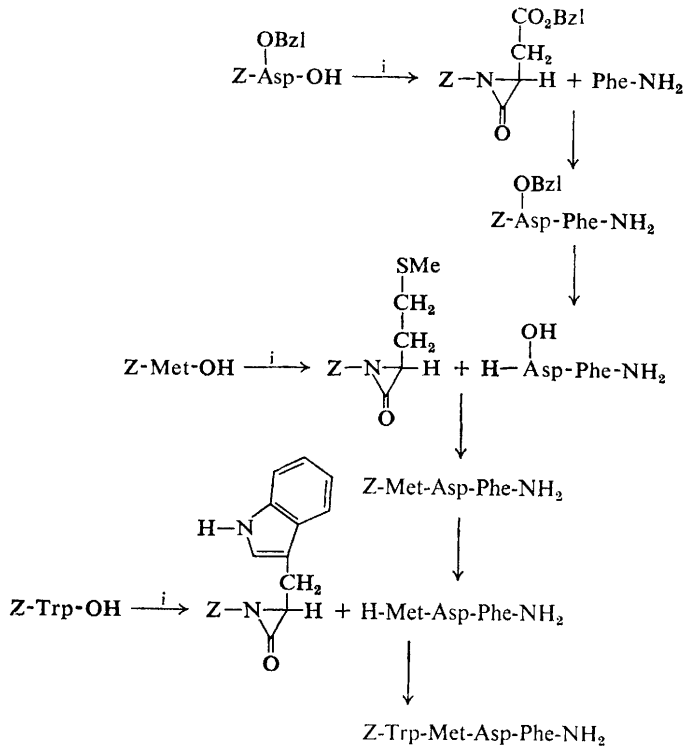
Scheme 59

on the lowered electron density of the strained three-membered ring (MO calculations) so that coupling proceeds *via* addition with concomitant ring opening. The preparation of the gastrin *C*-terminal tetrapeptide amide by this method has been described (Scheme 60).[220] The delineation of the scope and limitations of this promising new method of peptide synthesis is awaited with interest.

3-Isobutyl-7-nitro-2,3-dihydro-benzo[f]-1,4-oxazepin-2-one (130) can be prepared by the carbodi-imide method from the Schiff's base derived from leucine and 5-nitrosalicylaldehyde. Lactones such as (130) readily undergo aminolysis (Scheme 61).[221] Analogous activated lactones can be prepared from 3-formyl-*N*-hydroxyphthalimide, α-formyl-*N*-hydroxysuccinimide, 4-acetyl-*N*-hydroxypiperidine, and pentane-2,4-dione.

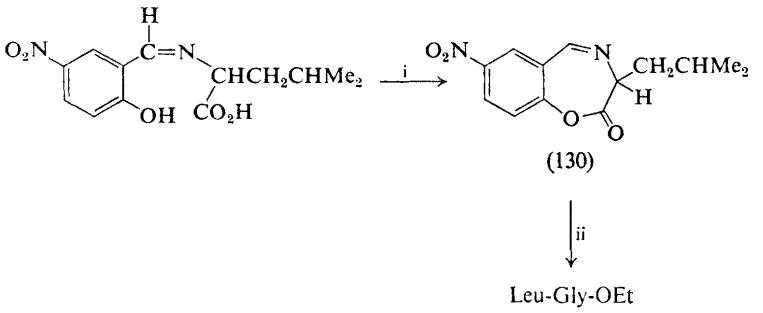
Coupling Methods involving Carbodi-imides.—Addition of dicyclohexylcarbodi-imide to solutions of t-butoxycarbonyl- or t-amyloxycarbonyl-amino-acids in dichloromethane or DMF results in the formation of ninhydrin-positive products, including the free amino-acid. This decomposition warns against the indiscriminate use of carbodi-imide.[222] The use of dicyclohexylcarbodi-imide to form peptide bonds in the presence of trifluoroacetate anion can cause substantial trifluoroacetylation of the amino-component.[223] Trifluoroacetylation is increased the greater the excess of carboxy-component and coupling reagent above that of the amino-component. It is also encountered in couplings involving highly reactive active esters. The difficulty can be circumvented by replacing trifluoroacetate by the less nucleophilic pivalate anion, but is best greatly reduced or

[220] M. Miyoshi, H. Tamura, K. Higaki, and K. Niwa, Ger. Offen, 2 245 459 (Cl. C 07*c*, A 61*k*) (*Chem. Abs.*, 1973, **78**, 148 241).
[221] M. Bodanszky, U.S.P. 3 704 246 (Cl. 260—333; C 07*d*) (*Chem. Abs.*, 1973, **78**, 58 801).
[222] M. Bodanszky and J. C. Sheehan, quoted in ref. 172.
[223] G. A. Fletcher, M. Löw, and G. T. Young, *J.C.S. Perkin I*, 1973, 1162.



Reagent: i, $COCl_2$

Scheme 60



Reagents: i, $C_6H_{11}N{:}C{:}NC_6H_{11}$; ii, Gly-OEt, followed by dil. HCl

Scheme 61

avoided by the addition of 1-hydroxybenzotriazole. Active ester couplings in the presence of triethylammonium formate can lead to ninhydrin-negative products of the amino-component, which probably arise by formylation.[224]

Amide bond formation mediated by ethyl(3-dimethylamino)propyl-carbodi-imide in aqueous solution gives maximum yield at pH 1.[225] In an attempt to prepare 1-nonanoyl-2,2-dimethylaziridine from nonanoic acid, dicyclohexylcarbodi-imide, and 2,2-dimethylaziridine, a substantial quantity of *N*-cyclohexylnonamide was produced, which presumably resulted from nucleophilic attack of the *N*-acylurea.[226] Determination of the product distributions for the reaction of three amines with the symmetrical anhydride of benzyloxycarbonylglycine, and with a mixture of benzyloxycarbonylglycine and dicyclohexylcarbodi-imide under identical conditions, indicates that acylation by the anhydride cannot entirely account for the initial acylation in the latter case.[227]

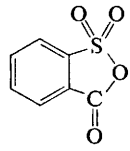
Coupling Methods involving Mixed Anhydrides.—The superiority of the base *N*-methylmorpholine for generating mixed carboxylic–carbonic anhydrides (using isobutyl chloroformate) has been reiterated.[228]

A simple, convenient, rapid, repetitive procedure for the synthesis of peptides utilizes excess mixed anhydride.[229, 230] In the procedure a 1.4-fold excess of mixed anhydride is generated using 1.5 equivalents of carboxy-component and 1.4 equivalents of isobutyl chloroformate and *N*-methyl-morpholine. The reaction is carried out in THF or DMF at -15 °C to avoid disproportionation of the anhydride. The amino-component is added and coupling is stopped after 2 h at -15 °C by decomposing the excess of mixed anhydride with aqueous potassium hydrogen carbonate.[229] The crude peptide can be recovered readily from the water-soluble side-products and, after deprotection, is ready for another coupling cycle. The crude products are obtained in excellent yield, and possess high purity and are free of racemization. The method is only limited in that N-protected asparagine requires side-chain protection to avoid imide formation and that N-protected glycine can give complications due to diacylimide formation; when there is a *C*-terminal proline in the amino-component, coupling can occur at the wrong carbonyl groups of the anhydride, and the activation of N-protected nitro-arginine can give minor by-products. Human growth hormone-(1—10)-decapeptide[231] and secretin-(12—27)-hexadecapeptide[232] have been prepared by this method.

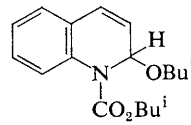
[224] G. P. Schwartz and P. G. Katsoyannis, *J.C.S. Perkin I*, 1973, 2890.
[225] M. E. Addy, G. Steinman, and M. F. Mallete, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 1034.
[226] L. S. Levitt, *Chem. and Ind.*, 1973, 637.
[227] J. Rebek and D. Feitler, *J. Amer. Chem. Soc.*, 1973, **95**, 4052.
[228] G. W. Anderson, ref. 2, p. 343.
[229] M. A. Tilak, M. L. Hendricks, and D. S. Wedel, ref. 2, p. 351.
[230] H. C. Beyerman, ref. 4, p. 351.
[231] H. C. Beyerman, E. W. B. DeLeer, and J. Floor, *Rev. Trav. chim.*, 1973, **92**, 481.
[232] A. van Zon and H. C. Beyerman, *Helv. Chim. Acta*, 1973, **56**, 1729.

An ion-exchange technique has been developed for the isolation of products of the repetitive excess mixed anhydride method.[233] After decomposition of the excess mixed anhydride the doubly protected product is passed selectively, as the only neutral non-volatile component of the reaction mixture, through a mixed-bed column of ion-exchange resins in the triethylammonium and acetate forms; evaporation of the solvent and of the released triethylammonium acetate leaves the required product. In the case of a deprotection cycle, the amino-peptide ester is captured, as the only non-volatile cationic component of the reaction mixture, on an acidic resin in the Na^+ cycle, and subsequently displaced with sodium acetate, the product acetate salt being selectively dissolved from the residue after evaporation.

o-Sulphobenzoic anhydride (131) reacts with the potassium salt of N-protected amino-acids at elevated temperatures; the resulting mixed carboxylic–carboxylic anhydride can be used for the synthesis of peptides.[234]



(131)

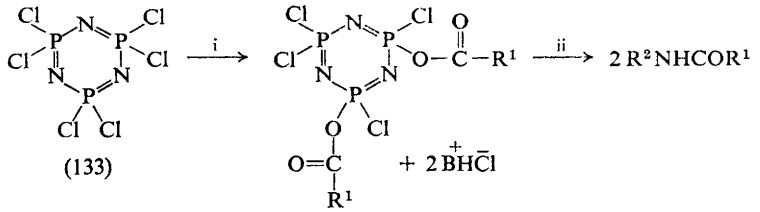


(132)

The full paper [235] concerning preparation and use of the modified 'EEDQ' reagents 2-isobutoxy-1-isobutoxycarbonyl-1,2-dihydroquinoline (132) and the corresponding 2-methoxy- and 2-ethoxy-compounds has appeared. Reagent (132) gave significantly higher yields than the other pseudo-bases. Further mention of polymeric 'EEDQ' has been made.[236] When aqueous bromine or chlorine is added to a mixture of glycine and alanine-3-sulphinic acid at pH 3.5 the major product (45% yield) is glycylcysteic acid; however, at pH 8.5 the major product (45% yield) is cysteylglycine.[237] The results are explained by the generation of a sulphonic–carboxylic mixed anhydride, which is attacked by an amino-group to give the respective dipeptides.

Other Methods.—It has been shown that hexachlorocyclotriphosphatriazene (133) can mediate amide-bond formation (Scheme 62); half an equivalent of the reagent in the presence of one equivalent of triethylamine or *N*-methylmorpholine is effective. The method is simple, rapid, and free of racemization in the Anderson test, but showed significant racemization in the Young test.[238]

[233] D. J. Woodman, L. C. Butler, and J. Stewart, *Tetrahedron Letters*, 1973, 1557.
[234] A. K. Koul, N. S. Ramegowda, M. N. Modi, and N. K. Mathur, *Indian J. Chem.*, 1973, **11**, 612.
[235] Y. Kiso, Y. Kai, and H. Yajima, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 2507.
[236] J. Brown, D. R. Lauren, and R. E. Williams, ref. 4, p. 217.
[237] P. G. Gordon, *Austral. J. Chem.*, 1973, **26**, 1771.
[238] J. Martinez and F. Winternitz, *Bull. Soc. chim. France*, 1972, 4707.

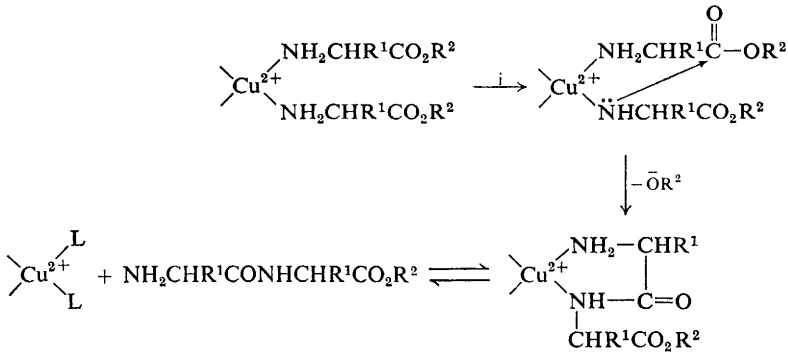


(133)

Reagents: i, $2R^1CO_2H$, 2B; ii, $2R^2NH_2$

Scheme 62

Amino-esters, in general, in the presence of copper(II) chloride in anhydrous methanol give rise to dipeptide esters.[239] The reactions proceed in low yield [1 mole of copper(II) chloride to 12 moles of ester] without racemization. In the case of glycine ester a complex mixture of di-, tri-, and tetra-glycine esters is obtained.[240] Isoleucine ester does not give a peptide product (presumably owing to a steric effect), neither do cysteine or histidine esters (presumably owing to the additional ligands). Lysine and ornithine esters give rise to the corresponding amino-caprolactam and -valerolactam. A mixture of the four possible dipeptide esters is obtained when a mixture of dichlorobis(amino-ester)copper(II) and another amino-ester is used, presumably owing to ligand exchange. When dichlorobis-(amino-ester)copper(II) is treated with a dipeptide ester, a tripeptide ester (amino-acyldipeptide ester only) results.[239] The reaction has also been applied to the formation of simple secondary amides.[241] A combination of chemical and spectroscopic evidence indicates that peptide bond formation



Reagents: i, $Cu^{II}Cl_2$ (1 mole), $NH_2CHR^1CO_2R^2$ (= L, 12 mole) in anhydrous alcohol

Scheme 63

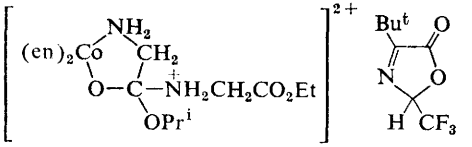
[239] S. Terashima, M. Wagatsuma, and S. Yamada, *Tetrahedron*, 1973, **29**, 1487.

[240] S. Yamada, M. Wagatsuma, Y. Takeuchi, and S. Terashima, *Chem. and Pharm. Bull. (Japan)*, 1971, **19**, 2380.

[241] M. Wagatsuma, S. Terashima, and S. Yamada, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 422.

occurs by N-co-ordination of the amino-ester, activating the carbonyl group of the ester to nucleophilic attack from the anion of a bound amino-group (Scheme 63).[242]

Miscellaneous Matters concerning Coupling Reactions.—The kinetics of the aminolysis reaction between isopropyl glycinate bis(ethylenediamine)-cobalt(III) perchlorate and glycine ethyl ester in DMSO to form ethyl glycylglycinate bis(ethylenediamine)cobalt(III) perchlorate and isopropyl alcohol reveal that the reaction occurs in two stages *via* an addition–elimination process involving a stabilized chelated tetrahedral intermediate (134).[243]



(134)

(135)

An asymmetric synthesis of sterically hindered *N*-pivaloyl-L-t-leucyl-L-valine has been accomplished using the high asymmetric induction afforded by the reaction of the racemic 3-oxazolin-5-one (135) with chiral valine.[244] A full paper has appeared concerning the silylation of amide bonds in peptides and their subsequent acylation to form *N*-acylaminoacyl peptides. Deprotection of the derivatives of α- or β-aminoacyl peptides so formed is followed by rearrangement, with incorporation of the aminoacyl residue into the peptide chain. The inserted optically active amino-acid residues are not racemized.[245]

Simple carboxylic acids and amines give amides in poor yield when treated with iodine and pyridine in liquid sulphur dioxide.[246] Use of liquid hydrogen cyanide as a solvent for peptide synthesis has been alluded to.[247] A simple rule of thumb has been developed concerning fragment condensations mediated by dicyclohexylcarbodi-imide and an additive. Complete couplings are attained by using the activated carboxy-component at concentrations in excess, such that their levels do not fall below 0.05 mol l^{-1} at completion.[138]

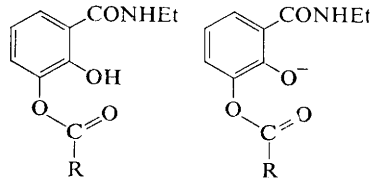
4 Racemization

The results of racemization tests of the various individual coupling methods have been mentioned, where appropriate, in the previous section. A detailed discussion, in mechanistic terms, of the coupling and racemization

[242] M. Wagatsuma, S. Terashima, and S. Yamada, *Tetrahedron*, 1973, **29**, 1497.
[243] D. A. Buckingham, J. Dekkers, and A. M. Sargeson, *J. Amer. Chem. Soc.*, 1973, **95**, 4173.
[244] E. Frauendorfer, W. Steglich, and F. Weygand, *Chem. Ber.*, 1973, **106**, 1019.
[245] J. S. Davies, C. H. Hassall, and K. H. Hopkins, *J.C.S. Perkin I*, 1973, 2614.
[246] M. Nojima, S. Hasegawa, and N. Tokura, *Bull. Chem. Soc. Japan*, 1973, **46**, 1254.
[247] M. Bodanszky, R. Rak, and M. T. Birns, ref. 2, p. 347.

behaviour of *p*-nitrophenyl esters, 3-acyloxy-2-hydroxy-*N*-ethylbenzamides, and *N*-succinimidyl esters has been published.[248]

In the case of *p*-nitrophenyl esters the kinetic data for aminolysis in polar solvents can best be interpreted in terms of solvent participation by general base catalysis. The kinetics of racemization are best explained by specific base catalysis. With 3-acyloxy-2-hydroxy-*N*-ethylbenzamides (136) under neutral coupling conditions, a small amount of the conjugate base is in equilibrium with the ester, the oxy-anion (137) facilitating aminolysis



(136) (137)

(and hydrolysis) by internal general base catalysis. Racemization appears to occur by specific base catalysis; however, oxazolinone formation under strongly basic conditions is inhibited by internal buffering of the oxy-anion. Evidence from labelling experiments indicates that an intermediate *N*-succinimidyl ester is involved in the dicyclohexylcarbodi-imide–*N*-hydroxysuccinimide method of coupling. *N*-Succinimidyl esters show relatively low levels of racemization in THF as solvent compared to DMF.

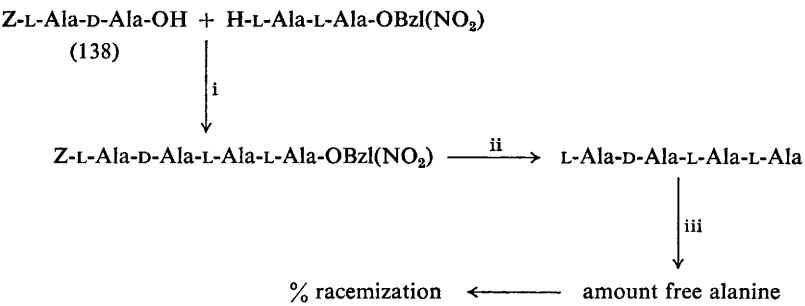
The rate coefficients for racemization (using triethylamine) and for coupling (using valine methyl ester) have been compared for a series of active esters of *S*-benzyl-*N*-benzyloxycarbonyl-L-cysteine,[249] *N*-benzyloxycarbonyl-L-serine,[250] β-methyl *N*-benzyloxycarbonyl-L-aspartate,[249, 251] *N*-benzyloxycarbonyl-L-phenylalanine,[250] γ-methyl *N*-benzyloxycarbonyl-L-glutamate,[249, 251] *N*-benzyloxycarbonyl-L-tryptophan,[250] and *N*-benzyloxycarbonyl-L-alanine in THF solution. Decreasing rates of racemization by α-hydrogen abstraction follow the order of substrates given above; the decreasing rates of racemization for a given N-protected amino-acid active ester follow the order: *N*-succinimidyl > pentafluorophenyl > 2,4,5-trichlorophenyl > *p*-nitrophenyl > pentachlorophenyl. The side-chains of the amino-acids do not appear to exert a significant effect on the rate coefficient for coupling with the valine derivative. For a given amino-acid active ester, however, the decreasing rate coefficient for coupling follows the order: pentafluorophenyl > *N*-succinimidyl > pentachlorophenyl > 2,4,5-trichlorophenyl > *p*-nitrophenyl. The results show that the ability

[248] D. S. Kemp, ref. 3, p. 1.
[249] J. Kovacs, G. L. Mayers, R. H. Johnson, R. Giannotti, H. Cortegiano, and J. Roberts, ref. 2, p. 185.
[250] J. Kovacs, E. J. Davis, R. H. Johnson, H. Cortegiano, and J. Roberts, ref. 4, p. 359.
[251] J. Kovacs, R. E. Cover, R. H. Johnson, T. J. Kalas, G. L. Mayers, and J. Roberts, *J. Org. Chem.*, 1973, **38**, 2518.

of an active ester to couple is not strictly parallel with its ability to racemize by α-hydrogen abstraction, so that a fast-coupling active ester which racemizes relatively slowly is probably the best choice for the synthesis of peptides.[251] The choice of active ester is not so critical for those amino-acid derivatives mentioned at the end of the sequence above (*e.g.* the tryptophan derivative) as for those at the beginning (*e.g.* the cysteine derivative).

Racemization tests which utilize glycine ester as amino-component are limited, in that glycine couples much faster than more hindered amino-acids, so that more extensive racemization is to be anticipated when other amino-components are used.[250]

A new generally applicable method for detecting racemization during peptide synthesis utilizes the stereoselective hydrolysis of diastereomeric peptides by leucine aminopeptidase.[252] A suitably protected dipeptide (Scheme 64) of configuration L-D is coupled with an all-L-alanine peptide.



Reagents: i, coupling method; ii, H_2–Pd/C in aq. HOAc; iii, leucine aminopeptidase, pH 8.6

Scheme 64

The blocking groups are removed and the free peptide is hydrolysed with leucine aminopeptidase. Since this enzyme shows absolute L-specificity, it cleaves only L-peptide formed through racemization. The amount of free alanine determined by amino-acid analysis gives a multiple of the degree of racemization. Under the reported conditions as little as 0.2% racemization could be detected. In the synthesis of the model peptide with dicyclohexylcarbodi-imide in DMF (1.1 equivalents DCCI and 1 equivalent of triethylammonium bromide) the penultimate alanine residue possessed an optical purity of only 20% (*i.e.* 80% racemization); however, in the presence of 2 equivalents of 1-hydroxybenzotriazole (and an extra equivalent of triethylamine) the optical purity was 99.8% (*i.e.* 0.2% racemization). In the synthesis of the *N*-succinimidyl ester of (138) using dicyclohexylcarbodi-imide in dioxan–ethyl acetate (2 equivalents of

[252] H. R. Bosshard, I. Schechter, and A. Berger, *Helv. Chim. Acta*, 1973, **56**, 717.

N-hydroxysuccinimide) and subsequent coupling in DMF in the presence of 1 equivalent of *N*-methylmorpholinium bromide and 1 equivalent of *N*-methylmorpholine, the penultimate alanine residue possessed an optical purity of 99.0% (*i.e.* 1.0% racemization).

The model coupling of *N*-benzyloxycarbonyl-L-prolyl-L-valine (139) with proline methyl ester has been used to evaluate a number of coupling methods.[253, 254] Racemization is detected, after Weygand, using the gas chromatographic resolution of the resulting *N*-trifluoroacetyl-L-prolyl-valyl-L-proline methyl ester diastereomers. Almost no racemization is observed in couplings mediated by the *N*-succinimidyl ester, the 8-hydroxyquinoline ester, and the *p*-nitrophenyl ester of (139), providing each ester

Z-L-Pro-L-Val

(139)

is prepared by a 'backing off' procedure. With the esters of *N*-hydroxypiperidine and thiophenol, large amounts of racemization are observed, even though the 'backing off' procedure is used.[253] In the azide and mixed anhydride methods no racemization is found, but a high degree of racemization is observed when dicyclohexylcarbodi-imide, oxidative coupling of the corresponding hydrazide, or the phosphorazo method is used.[254]

The degree of racemization of an amino-acid can be determined by the reaction of that amino-acid with an optically active amino-acid *N*-carboxy-anhydride, followed by resolution of the resulting dipeptide diastereoisomers by ion-exchange chromatography on an amino-acid analyser. The method can detect 0.01% of a stereochemical impurity. For analysis of the stereochemical homogeneity of the residues in a synthetic peptide, hydrolysis with tritiated hydrochloric acid can be used for determination of the amount of racemization during hydrolysis (for many of the common amino-acids). These two techniques together can be used to establish, to better than 1%, the configurations of most amino-acid residues in either synthetic or natural peptides.[255]

Another method of detecting racemization in peptide synthesis involves preparing a phenylthiocarbamyl tripeptide from a phenylthiocarbamyl dipeptide and an amino-component; racemization of the central residue in the tripeptide is assessed by chromatographic resolution of the dipeptide stereoisomers resulting from one cycle of the Edman degradation on the tripeptide.[256]

Peptide pentachlorophenyl esters can be prepared without racemization using the complex of pentachlorophenol and dicyclohexylcarbodi-imide, if short reaction times and solvents such as dichloromethane–hexane or

[253] I. Tomida, H. Kayahara, and R. Iriye, *J. Agric. Biol. Chem. Japan*, 1973, **37**, 2549.
[254] I. Tomida, H. Kayahara, and R. Iriye, *J. Agric. Biol. Chem. Japan*, 1973, **37**, 2557.
[255] J. M. Manning, A. Marglin, and S. Moore, ref. 2, p. 173.
[256] M. Smulkowski, Ł. Lubiewska-Nakonieczna, and E. Taschner, ref. 3, p. 45.

ethyl acetate are used.[257] In the coupling of *N*-benzyloxycarbonylglycyl-L-alanine pentachlorophenyl ester with leucine benzyl ester, racemization could be best avoided using the base 1-diethylaminopropan-2-ol in dioxan.[257]

Further studies on the dependence of the extent of racemization of peptide azides on the amount and nature of the added base have been reported.[205, 257, 258] From a study of the model coupling of β-t-butyl *N*-benzyloxycarbonyl-L-aspartyl-L-phenylalanine azide with glycine or valine methyl ester, it is clear that triethylamine should be avoided, that *N*-ethylmorpholine is satisfactory if used in not too great an excess, and that for slow couplings *NN*-di-isopropylethylamine should be used.[258] 1-Diethylaminopropan-2-ol has also received further recommendation.[257] Clearly the triethylamine bottle is due to be banished to the farthest corner of other laboratories too![258] Difficulties in synthesis with a hexapeptide azide, and racemization during coupling of the corresponding acid by the dicyclohexylcarbodi-imide–*N*-hydroxysuccinimide method have been reported.[259]

A method has been developed for investigating possible racemization of the amino-acid residue penultimate to the *C*-terminal one, upon activation of an N-protected peptide. The results would appear to indicate that not insignificant amounts of racemization can occur in this position, although it is not clear to what extent the background racemization, inherent in the method, plays a role.[260]

It was mentioned last year (see these Reports, Vol. 5, p. 303) that racemization can attend the deprotection of *N*-methyl-amino-acids and peptide derivatives. This matter has now been discussed in full.[261, 262] Since saponification of alkyl esters of *N*-methyl derivatives (or treatment of the corresponding t-butyl esters with base) can result in appreciable racemization, it should not be used. Racemization which attends the removal, with hydrogen bromide in acetic acid, of benzyloxycarbonyl groups from *N*-methyl derivatives can be suppressed by adding water to the reagent although, clearly, hydrogenolytic deprotection is to be preferred. The full paper has appeared concerning the racemization of *N*-methyl-amino-acid and peptide derivatives during peptide bond formation.[263] A study of the model couplings of *N*-benzyloxycarbonyl-L-alanyl-L-*N*-methyl-leucine (and of the corresponding *N*-t-butoxycarbonyl derivative) with glycine benzyl ester revealed that *N*-methyl-amino-acid residues racemize as readily as amino-acid residues, except in the presence

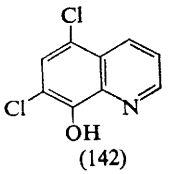
[257] L. Kisfaludy, O. Nyeki, and T. Szirtes, ref. 3, p. 54.
[258] P. Sieber and B. Riniker, ref. 3, p. 49.
[259] O. A. Kaurov, V. F. Martynov, and O. A. Popernatskii, *J. Gen. Chem.* (*U.S.S.R.*), 1973, **43**, 907.
[260] M. Dzieduszycka, M. Smulkowski, and E. Taschner, ref. 5, p. 103.
[261] J. R. McDermott and N. L. Benoiton, *Canad. J. Chem.*, 1973, **51**, 1915.
[262] J. R. McDermott and N. L. Benoiton, *Canad. J. Chem.*, 1973, **51**, 2555.
[263] J. R. McDermott and N. L. Benoiton, *Canad. J. Chem.*, 1973, **51**, 2562.

of salt (triethylammonium toluene-*p*-sulphonate), when their racemization is more pronounced. In the presence of salt a stereochemically pure product was only obtained using the corresponding *N*-succinimidyl ester, while in the absence of salt, coupling with 'EEDQ' and dicyclohexylcarbodi-imide–*N*-hydroxysuccinimide gave essentially optically pure products.

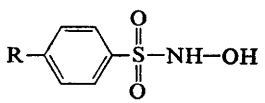
Ethyl 2-hydroximino-2-cyanoacetate (140)[264, 265] and ethyl 2-hydroximino-2-cyanoacetamide (141)[265] have been recommended as racemization-suppressing additives for dicyclohexylcarbodi-imide-mediated peptide

$HON{:}C(CN)CO_2Et$ (140) $HON{:}C(CN)CONH_2$ (141)

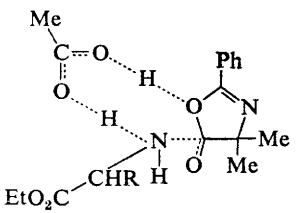
couplings. Either of the two additives gave no detectable racemization in the Anderson test (coupling solvent, THF), the Young test gave ethyl benzoyl-L-leucylglycinate with an optical purity of 93.6%, and the use of (140) in the Bodanszky test gave ethyl acetyl-L-isoleucylglycinate with an optical purity of 98.9% with respect to the α-centre (the figure being corrected for racemization occurring during acidic hydrolysis). 5,7-Dichloro-8-hydroxyquinoline (142) is also an effective racemization depressant in carbodi-imide-mediated condensations, the Bodanszky peptide being obtained with an optical purity of 94.7%.[266] Further depressants have been mentioned;[267] *p*-chloro- and *p*-nitro-benzenesulphohydroxamic acids (143) yield the Bodanszky peptide in 91.0 and 91.9% optical purity, respectively.



(142)



(143) R = Cl or NO_2



(144)

[264] M. Itoh, ref. 4, p. 365.
[265] M. Itoh, *Bull. Chem. Soc. Japan*, 1973, **46**, 2219.
[266] H. Yajima, M. Kurobe, and K. Koyama, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 1612.
[267] H. Yajima, K. Kitagawa, and M. Kurobe, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 2566.

The kinetics of the ring opening of 2-phenyl-4,4-dimethyl-2-oxazolin-5-one with amino-esters, catalysed by acetic acid, have been reported.[268] A cyclic intermediate [see (144)] has been proposed to explain the function of the catalyst.

5 Repetitive Methods of Peptide Synthesis

Solid-phase Synthesis.—A critical assessment of the current position includes consideration of the limitations imposed by the polymer support, the benzyl ester polymer linkage, and the occurrence of failure sequences.[269] Some reflections on solid-phase synthesis,[270] a masterful Harvey Lecture account of the development of the solid-phase strategy culminating in the synthesis of ribonuclease A,[271] and some reviews have appeared.[272, 273] The large number of peptides which have been synthesized during the year under review, utilizing the generalized Merrifield technique, are not discussed here but are listed in Appendix I.

General Improvements to the Merrifield Technique. Wieland has classified [12] the adherents of conventional peptide synthesis into 'fragment condensers' and 'stepwise prolongers'. This classification can now begin to be applied to solid-phase strategists, since the year under review has seen the appearance of several papers concerned with fragment condensations on a polymeric support. The rationale of such studies is that at the end of a synthesis the final peptide will be contaminated only with peptides which have widely dissimilar properties, so that the eventual purification will be facilitated. Furthermore, the smaller number of N-deprotection cycles required will cut down the premature removal of side-chain-protecting groups and of the growing peptide from the resin. However, coupling of fragments in the heterogeneous phase becomes a more formidable problem. Examples of this approach are the synthesis of some small peptides,[274] luteinizing-hormone-releasing hormone,[275] and bovine pancreatic basic trypsin inhibitor,[276, 277] of which the last example represents the most complex and demanding synthetic situation.

In the Merrifield synthesis of (145) it was shown that peptide–resin steric interactions could be minimized by using a resin with the lowest

[268] C. Chuaqui, S. Atala, A. Márquez, and H. Rodríguez, *Tetrahedron*, 1973, **29**, 1197.
[269] R. C. Sheppard, ref. 3, p. 111.
[270] R. B. Merrifield, ref. 2, p. 89.
[271] R. B. Merrifield, *Harvey Lectures*, Ser. 67, p. 143.
[272] J. M. Stewart, *Ann. Reports Medicin. Chem.*, 1972, **7**, 289 (*Chem. Abs.*, 1973, **78**, 43 952).
[273] C. H. Gaozza, *Soc. Argent. Farm. Bioquim. Ind.*, 1973, **12**, 692 (*Chem. Abs.*, 1973, **79**, 42 812).
[274] M. A. Barton, R. U. Lemieux, and J. Y. Savoie, *J. Amer. Chem. Soc.*, 1973, **95**, 4501.
[275] R. Matsueda, H. Maruyama, E. Kitazawa, H. Takahagi, and T. Mukaiyama, *Bull. Chem. Soc. Japan*, 1973, **46**, 3240.
[276] N. Izumiya, T. Kato, M. Waki, N. Mitsyasu, K. Noda, S. Terada, and O. Abe, ref. 2, p. 69.
[277] H. Yajima, Y. Kiso, Y. Okada, and H. Watanabe, *J.C.S. Chem. Comm.*, 1974, 106.

Pro-Pro-Thr(Bzl)-Ile-Val-Val-His(Tos)-Gly
(145)

degree of cross-linking consistent with stability, by using only a low degree of chloromethylation (to avoid additional cross-linking), by esterifying the first amino-acid to the resin in a solvent which does not swell the resin extensively (this limits substitution to the most accessible sites), and by performing the rest of the synthesis with the resin maximally swollen.[278]

Syntheses of a polypeptide with some hen egg-white lysozyme activity have been reported.[279, 280] This work highlights some of the major technical problems associated with application of the stepwise solid-phase technique to the synthesis of protein molecules, namely: the cumulative effects of poor coupling efficiency, premature removal of side-chain blocking groups, premature removal of the growing peptide chain from the resin, and incomplete final deprotection.

Poor coupling efficiency (in some cycles it was as low as 90%) is improved by using an excess of the protected amino-acid (to promote symmetrical anhydride formation), the incremental addition of dicyclohexylcarbodiimide, and repetition of the coupling cycle as dictated by ninhydrin analysis. Furthermore, chains recalcitrant to coupling are terminated by acetylation. The N-deprotection reagent (trifluoroacetic acid, dichloromethane, mercaptoethanol, and anisole) causes loss of N^{ε}-benzyloxycarbonyl protecting groups of lysine, and of peptide from the resin (in one case 93% of the amino-acid originally esterified to the resin was lost during synthesis). The deprotection cycle is improved using hydrogen chloride in dioxan (containing ethane-1,2-dithiol to protect tryptophan), which enables N^{α}-t-butoxycarbonyl groups to be removed without marked loss of side-chain protecting groups or without the benzyl ester linkage suffering too much cleavage (for a 30 min cycle, 1.1% of the total peptide was removed from the resin). Final deprotection with hydrogen fluoride–anisole is far from satisfactory (it fails to remove *S*-benzyl groups completely). The reagent exerts a pronounced destructive effect on the reduced and unfolded native protein.

The attempted synthesis [280] of lysozyme, incorporating the improvements mentioned, took six months to complete (residues being incorporated at the rate of one per day). It yielded a product which after chromatographic purification had 2—3% of the specific activity of lysozyme. Although the product exhibited the expected specificity towards the hexamer of *N*-acetylglucosamine, identity of the active component with natural hen egg-white lysozyme has not been satisfactorily demonstrated.

Further major developments in solid-phase procedures are clearly necessary for the method to be generally applicable to the synthesis of

[278] J. M. Stewart and R. Matsueda, ref. 4, p. 221.
[279] L. E. Barstow, D. A. Cornelius, V. J. Hruby, T. Shimoda, J. A. Rupley, J. J. Sharp, A. B. Robinson, and M. D. Kamen, ref. 4, p. 231.
[280] J. J. Sharp, A. B. Robinson, and M. D. Kamen, *J. Amer. Chem. Soc.*, 1973, **95**, 6097.

proteins. There is a dire need for more suitable solid supports, and milder conditions are required for N- and final deprotection.

Polymeric Carriers. We should perhaps mention in passing that oligonucleotide synthesis has been carried out on a polypeptide support [281] (namely on poly-L-lysine)!

Polystyrene-coated glass beads have been used for solid-phase synthesis of some simple peptides using a column process.[282] The more efficient mass transfer in the thin film of resin enables the time for the coupling cycles to be reduced. Polystyrene chains can be insolubilized with minimal cross-linking by anchoring them at one end to an inert insoluble core resin. Graft copolymers of this type can be prepared by irradiating Kel F beads with γ- or *X*-rays and exposing the irradiated polymer to styrene monomer.[283] These polymers are dense, sink rapidly in most solvents, and do not swell appreciably. Comparative peptide synthesis on chloromethylated graft copolymer and on conventional polystyrene shows both polymers to possess similar reaction characteristics; however, the former polymer possesses considerable design flexibility.

Porous, highly cross-linked polymethacrylate gels have been used for peptide synthesis, using *o*-nitrophenylsulphenyl N-protection and deprotection with hydrogen chloride in dichloromethane.[284] Benzhydrylamine supports [285] and pellicular silicone resins [286] have received further mention.

Peptide synthesis on hydroxypropylated dextran (Sephadex LH 20) has been studied.[287] The anchor ester link is forged with a t-butoxycarbonyl-amino-acid and carbonyldi-imidazole.[288] With different amino-acid derivatives loadings of up to 5 mmol g^{-1} could be achieved. The most suitable N-deprotection reagent is toluene-*p*-sulphonic acid in acetic acid. The amino-groups liberated can be determined spectrophotometrically from the amount of toluene-*p*-sulphonic acid released in the neutralization cycle. Bradykinin has been prepared by this method in 60% yield.[289]

A careful study [290] of the use of a soluble linear monodisperse polystyrene carrier for peptide synthesis utilizes gel-filtration on Bio-beads to separate the growing polymer from reaction components of low molecular weight.

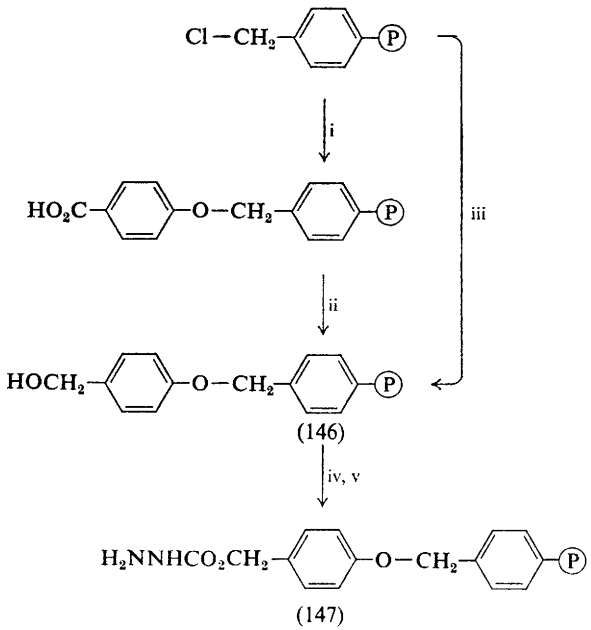
Linkage to the Polymeric Carrier. N-Protected amino-acids can be linked *via* an ester bond to polystyrene by allowing the chloromethylated polymer

[281] T. M. Chapman and D. G. Kleid, *J.C.S. Chem. Comm.*, 1973, 193.
[282] R. P. W. Scott, K. K. Chan, P. Kucera, and S. Zolty, *J. Chromatog. Sci.*, 1971, **9**, 577.
[283] G. W. Tregear, ref. 4, p. 175.
[284] V. Gut, ref. 5, p. 185.
[285] P. G. Pietta, P. F. Cavallo, and G. R. Marshall, ref. 3, p. 172.
[286] K. Grohmann, *Diss. Abs., B*, 1973, **97**, 236.
[287] G. P. Vlasov, A. Yu. Bilibin, N. Yu. Kuznetsova, I. Ditkovskaya, and V. N. Lashkov, *Chem.-Ztg.*, 1973, **97**, 236 (*Chem. Abs.*, 1973, **79**, 42 833).
[288] A. Yu. Bilibin and G. P. Vlasov, *J. Gen. Chem.* (*U.S.S.R.*), 1973, **43**, 1828.
[289] A. Yu. Bilibin, N. Yu. Kozhevnikova, and G. P. Vlasov, *Zhur. obshchei Khim.*, 1973, **43**, 2046 (*Chem. Abs.*, 1974, **80**, 15 176).
[290] R. H. Andreatta and H. Rink, *Helv. Chim. Acta*, 1973, **56**, 1205.

to react with the benzyltrimethylammonium salt of the acid.[290] t-Butoxycarbonylglutamine resin esters are best prepared[291] by the procedure of Dorman and Love.

The reaction of chloromethylated polystyrene–1% divinylbenzene copolymers with caesium salts of N-protected amino-acids, using DMF, proceeds rapidly, without side-reactions, to give N-protected amino-acyl resin esters free of quaternary ammonium sites or of reactive chloride. The method is superior to other esterification techniques because it proceeds quantitatively, enabling the degree of substitution to be controlled.[292]

Two resins based on derivatives of t-butyl alcohol have been used for the synthesis of oligopeptides, which were subsequently used in fragment condensations in homogeneous solution. A similar pair of resins based on polymeric derivatives of *p*-alkoxybenzyl alcohol has now been described,[293, 294] and prepared according to Scheme 65. N^1-(*p*-Biphenylyl)-1-methylethoxycarbonyl-amino-acids can be esterified to the *p*-alkoxybenzyl

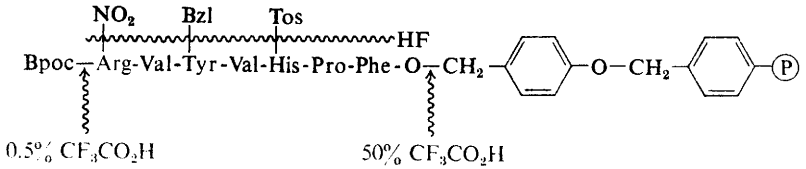


Reagents: i, *p*-$HOC_6H_4CO_2H$, NaOMe; ii, $LiAlH_4$; iii, *p*-$HOC_6H_4CH_2OH$, NaOMe; iv, PhOCOCl, C_5H_5N; v, H_2NNH_2

Scheme 65

[291] L. C. Dorman, D. A. Nelson, and R. C. L. Chow, ref. 2, p. 65.
[292] B. F. Gisin, *Helv. Chim. Acta*, 1973, **56**, 1476.
[293] S.-S. Wang, *J. Amer. Chem. Soc.*, 1973, **95**, 1328.
[294] S.-S. Wang, ref. 4, p. 179.

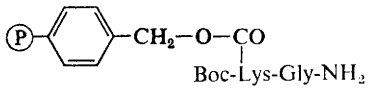
alcohol resin (146) using dicyclohexylcarbodi-imide and pyridine, or using the corresponding *p*-nitrophenyl ester. This resin is suitable for the synthesis of peptide fragments possessing a free carboxy-group; the three-stage selectivity to acid depicted in Scheme 66 is used. N-Protected amino-acids



Scheme 66

can be linked to the *p*-alkoxybenzyloxycarbonyl hydrazide resin (147) with carbodi-imide. This resin is suitable for the synthesis of peptide hydrazides, using similar deprotection conditions to those outlined for the alcohol resin (146) in Scheme 66.

It has already been mentioned that *N*-t-butoxycarbonyl-amino-acids can be esterified to hydroxypropylated dextran (Sephadex LH20) using carbonyldi-imidazole.[288] Lysine has been attached to carboxymethyl-dextran *via* its N^{ε}-amino-group by coupling the copper chelate of lysine with the polymer using dicyclohexylcarbodi-imide.[295] N^{α}-t-Butoxycarbonyl-L-lysylglycine amide has also been attached by its side-chain amino-group to polystyrene–2% divinylbenzene copolymer *via* a urethane linkage. The resulting derivative (148) has been used for the synthesis of lysine-vasopressin by successive N^{α}-deprotection, coupling, and final deprotection



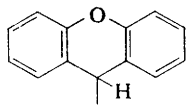
(148)

using hydrogen bromide in dioxan or trifluoracetic acid.[296] Further mention has been made of the bidirectional extension of peptide chains from a histidine residue anchored to the resin *via* its side-chain.[297]

Side-chain Protection. The following solid-phase syntheses have been reported using N_{in}-formyl protection of tryptophan: somatostatin,[298] some oligopeptides,[155, 156] and ribonuclease T_1 [299] (this topic has been discussed in a previous section, p. 301). The hydroxy-group of threonine has been

[295] A. Livshits and A. E. Vasil'ev, *Zhur. obshchei Khim.*, 1973, **43**, 219 (*Chem. Abs.*, 1973, **78**, 111 711).
[296] J. Meienhofer and A. Trzeciak, ref. 3, p. 199.
[297] J. D. Glass, R. Walter, and I. L. Schwartz, ref. 5, p. 135.
[298] D. Yamashiro and C. H. Li, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 882.
[299] N. Izumiya, M. Waki, T. Kato, M. Ohno, H. Aoyagi, and N. Mitzuyasu, ref. 4, p. 269.

protected as the *N*-benzyloxycarbonylamino-2,2,2-trifluoroethyl derivative in the synthesis of human growth hormone-(1—10)-decapeptide.[300] The N^{δ}-xanthenyl group cannot be used as a permanent amide-protective group for glutamine, when used in conjunction with the N^{α}-t-butoxycarbonyl protecting group, on account of the lack of selectivity to acidolysis. However, the compound (149) may still be useful to introduce glutamine



Boc-Gln-OH

(149)

residues by the carbodi-imide method; subsequent simultaneous N^{α}-N^{δ}-deprotection with hydrogen chloride in dioxan proceeds without formation of pyroglutamyl derivatives.[291]

It is important in solid-phase synthesis to accept that, although a side-chain-protecting group may be quite stable during an N^{α}-deprotection cycle, it may nevertheless not survive cumulative multiple deprotection cycles,[301] to the extent required in the stepwise synthesis of protein molecules. This problem has already been alluded to in the case of lysozyme,[280] and descriptions of the modified N^{ε}-benzyloxycarbonyl, *O*-benzyl, and *S*-benzyl protecting groups which have been devised to overcome this problem have been given (pp. 284, 287, and 299). In this section it is necessary to concentrate on these matters from a different angle.

A study of the stability of a series of N^{α}-acetyl-amino-acid amide derivatives (bearing side-chain-protecting groups) to prolonged treatment with 50% trifluoroacetic acid in dichloromethane (mainly using thin-layer chromatography) enabled a rough estimate to be made of the likely survival of those protecting groups in multiple N-deprotection cycles.[91]

The acid stability of a series of amino-acids bearing benzylic side-chain-protecting groups was determined quantitatively using ion-exchange chromatography.[90, 92, 147] The pseudo-first-order rate coefficients for loss of the benzylic groups were measured in 50% trifluoroacetic acid in dichloromethane at 20 °C. Assuming that the relative stabilities of side-chain-protected amino-acids in solution can be used to predict their stabilities under solid-phase conditions, the extent of side-chain deprotection can be estimated (see Table). [The average rate coefficient for the loss of the benzyloxycarbonyl group from N^{ε}-benzyloxycarbonyl-lysine in the solid-phase synthesis of decalysylvaline is about half the value of the rate coefficient for the loss of the benzyloxycarbonyl group from lysine in solution (*i.e.* $1.9 \times 10^{-6}\ s^{-1}$ and $3.96 \times 10^{-6}\ s^{-1}$, respectively).] Com-

[300] E. W. B. de Leer and H. C. Beyerman, *Rec. Trav. chim.*, 1973, **92**, 174.
[301] E. Schaich and F. Schneider, ref. 5, p. 180.

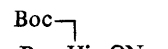
Table *Deprotection of benzylic amino-acid derivatives in 50% trifluoroacetic acid in dichloromethane*

Benzylic amino-acid derivative	k_{rel} [a]	*% loss per cycle* [b]
Tyr (Bzl)	94	0.76
Lys (Z)	59	0.473
Lys (Z[4-Cl])	21	0.166
Tyr (Z)	13	0.11
Cys (Bzl[4-MeO])	12	0.098
Ser (Bzl)	1.6	0.013
Asp (OBzl)	1	0.011
Glu (OBzl)	1	0.009
Thr (Bzl)	[1.0]	0.0083
Lys (Z[2-Cl])	0.94	0.0076
Lys (Z[2,4-Cl_2])	0.73	0.0059
Lys (Z[3,4-Cl_2])	0.36	0.003
Lys (Z[3-Cl])	0.072	0.00062
Lys (Z[2,6-Cl_2])	0.058	0.0005
Tyr (Bzl[2,6-Cl_2])	< 0.018	< 0.00014
Cys (Bzl[4-Me])	< 0.005	< 0.00004

[a] Pseudo-first-order rate coefficient for the loss of protecting group from *O*-benzyl-threonine normalized to unity, and the other rate coefficients quoted relative to this value. *O*-Benzyl-threonine loses 1% of its protection in 40 h (*i.e.* equivalent to 120 deprotection cycles) ($k_\alpha/k_{sc} = 1.0 \times 10^6$, thus fulfilling the criterion laid down); [b] A deprotection cycle involves treatment with the above reagent for 20 min.

parison of the rate of deprotection of side-chain-protecting groups with that of N^α-t-butoxycarbonyl groups enables an estimate to be made of the probable loss of side-chain-protecting groups during an N^α-deprotection cycle. It is suggested that no more than 0.05% cleavage of a side-chain group should occur during the removal of more than 99.95% of the N^α-protecting group, if it is to be acceptable for side-chain masking (*i.e.* $k_\alpha/k_{sc} = 1.5 \times 10^4$). It is possible to calculate the extent of side-chain deprotection to be expected in a given solid-phase synthesis using the data of the Table. These studies were used to assess the merits of the substituted benzyl protecting groups discussed earlier,[90, 92, 147] and represent an important general approach to the study of protective-group stability in solid-phase synthesis.

A study of the suitability of a range of protecting groups for histidine [302, 303] and lysine [303] side-chains has been made. $N^\alpha N^{im}$-Bis-t-butoxycarbonyl-L-histidine *p*-nitrophenyl ester (150) is a useful derivative for introducing histidine residues in solid-phase synthesis.[302]



(150)

[302] E. Schaich and F. Schneider, ref. 3, p. 127.
[303] E. Schaich, A. M. Fretzdorff, and F. Schneider, *Z. physiol. Chem.*, 1973, **354**, 897.

The Coupling Step. The relationships between the expected purity of the product, the chain length of the target peptide, and the average coupling efficiency are illustrated in the Figure. It is apparent that high efficiency in each step of every cycle is required if a large polypeptide is to be synthesized.[294]

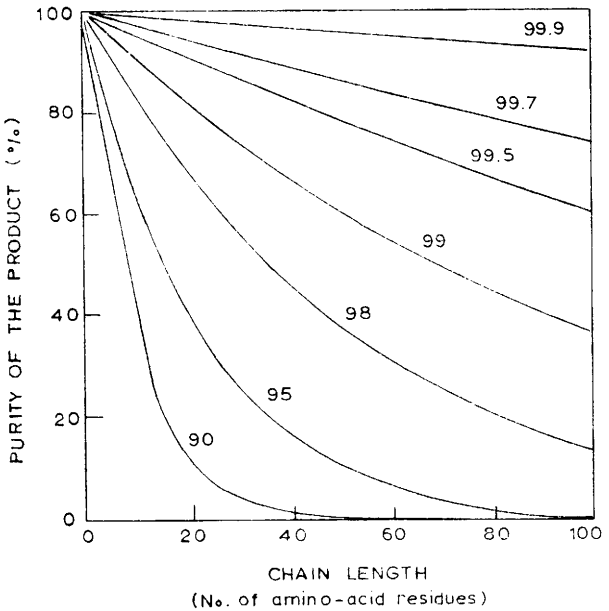


Figure *Relationship between average coupling efficiency and purity of products prepared by the solid-phase method*
(Reproduced by permission from 'Chemistry and Biology of Peptides; Proceedings of the Third American Peptide Symposium', ed. J. Meienhofer, Ann Arbor Science Publishers Inc., 1972, p. 180)

A warning has been given concerning the indiscriminate use of dicyclohexylcarbodi-imide.[222] Use of ionic binding to conserve the carboxy-component in carbodi-imide-mediated solid-phase synthesis has been mentioned.[304] 2,4,5-Trichlorophenyl esters of N-protected amino-acids can be used in solid-phase synthesis with the aid of catalysis from 3-hydroxy-4-oxo-3,4-dihydroquinazoline.[165] *o*-Nitrophenyl esters [171] possess a higher reactivity than *p*-nitrophenyl esters in solid-phase synthesis, but pentafluorophenyl esters are more reactive still.[174] Pentachlorophenyl esters have also been employed for this purpose as *N*-t-butoxycarbonyl [305] or *N*-2-hydroxy-1-naphthal [306] derivatives. In the latter case removal of the

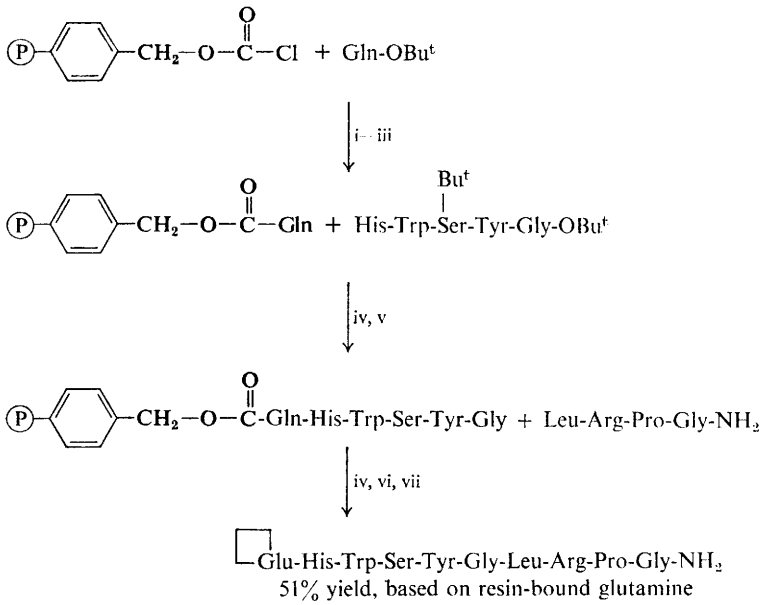
[304] D. F. Elliott, P. Moritz, and R. Wade, ref. 3, p. 168.
[305] B. J. Johnson, *J. Pharm. Sci.*, 1973, **62**, 1019.
[306] M. C. Khosla, R. R. Smeby, and F. M. Bumpus, ref. 2, p. 41.

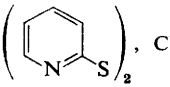
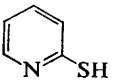
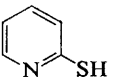
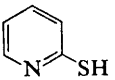
N-2-hydroxy-1-naphthal protecting group was accomplished with benzylamine in dichloromethane. Solid-phase synthesis using 'EEDQ' has been investigated.[307] The use of symmetrical anhydrides has also been mentioned.[44] A comparative study has been made of degree of coupling attained in the solid-phase synthesis of a model peptide using *p*-nitrophenyl esters, 5-chloro-8-hydroxyquinoline esters, selenophenyl esters, *N*-hydroxyurethane esters, the carbodi-imide method, the mixed anhydride method, and Woodward's reagent K.[308] A model system has been devised[309] for studying the coupling yields attained in solid-phase syntheses, using ion-exchange chromatography of the liberated peptides. Some quantitative studies on the formation of peptide bonds in Merrifield synthesis, when sterically hindered components are involved, or when the resin-bound free amino-groups are located close to the matrix or separated from it by a hydrocarbon chain, have been reported.[310] A quantitative comparison of the degree of coupling on a Merrifield support and on a macroreticular support shows that the latter support is suitable for solid-phase synthesis but possesses no great advantage over the former resin.[311]

Solid-phase synthesis using fragment conjunction was attempted in the synthesis of human growth hormone-(1—10)-decapeptide, where the chain was elongated using some dipeptide carboxy-components.[300] A model system involving the coupling of *N*-t-butoxycarbonylglycyl-L-alanine with L-leucyl-polymer has been developed to evaluate coupling methods.[274] 'EEDQ' (1.5 equivalents) in dioxan proved to be the most effective coupling agent, giving an 89% incorporation on to the leucine residue, the optical purity of the alanine residue in the Izumiya peptide being 99.1% (0.9% racemization). The dicyclohexylcarbodi-imide and 'activated hexamethylphosphoramide' methods proved to be much inferior as regards coupling yield and degree of racemization. The versatile oxidation–reduction method of coupling has been applied to the fragment solid-phase synthesis of optically pure luteinizing-hormone-releasing hormone, by a conventional *C*-terminal elongation process (the *N*-terminal pyroglutamyl residue being introduced subsequently by a solution coupling using pentachlorophenyl L-pyroglutamate) and by the less conventional *N*-terminal elongation (Scheme 67).[275, 312]

The synthesis of bovine pancreatic basic trypsin inhibitor represents the most demanding example of the use of fragment condensations on the solid phase.[276, 277] The synthesis is outlined in Scheme 68. However, racemization is avoided by the use of convenient glycine residues. The product is obtained in 7.5% yield, starting with blocked peptide resin, with a specific activity against tosylarginine methyl ester of 82%. The

[307] F. Sipos and D. W. Gaston, ref. 3, p. 165.
[308] H.-D. Jakubke and A. Baumert, ref. 3, p. 132.
[309] H. Hagenmaier and F. Hartmut, *J. Chromatog. Sci.*, 1972, **10**, 663.
[310] G. Losse and R. Ulbrich, ref. 5, p. 152.
[311] A. Losse, *Tetrahedron*, 1973, **29**, 1203.
[312] T. Mukaiyama, M. Ueki, and R. Matsueda, ref. 4, p. 209.



Reagents: i, $CHCl_3$, NEt_3; ii, excess chloroformyl groups destroyed with Et_2NH; iii, CF_3CO_2H, CH_2Cl_2; iv, Ph_3P,  , CH_2Cl_2, Me_2NCHO; v, CF_3CO_2H, CH_2Cl_2,  ; vi, HF, PhOMe,  ; vii, HOAc, PhOMe,  [Ⓟ = polystyrene–2% divinylbenzene copolymer]

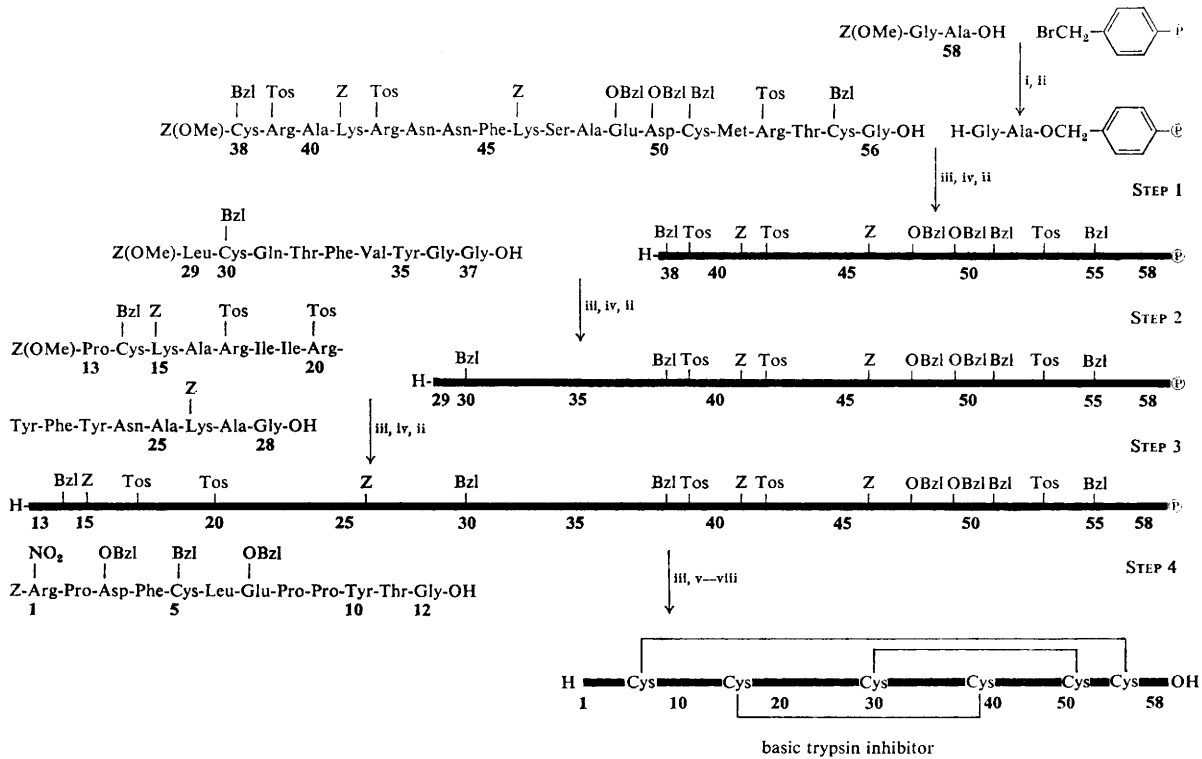
Scheme 67

apparent yields in the four coupling steps indicated in Scheme 68 were 75% in step 1, 82% in step 2, 78% in step 3, and nearly 100% in step 4. Identity of synthetic and natural material appears to have been satisfactorily demonstrated for a preliminary communication. It would seem therefore that the use of a fragment strategy in solid-phase synthesis will become more widespread.

Removal of the Object Peptide from the Resin. Removal of the finished peptide from the Merrifield resin using base-catalysed transesterification (tertiary-base-catalysed methanolysis) has been discussed further.[313]

Self-catalysed transesterification with 2-dimethylaminoethanol in DMF at room temperature is effective for the removal of protected peptides

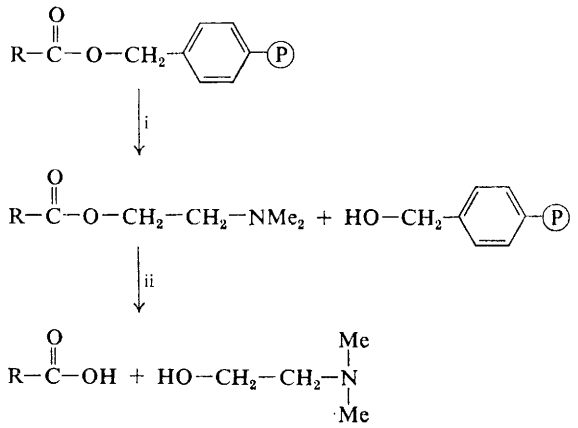
[313] H. C. Beyerman, ref. 2, p. 25.



Reagents: i, $(C_6H_{11})_2NH$; ii, CF_3CO_2H, CH_2Cl_2; iii, $C_6H_{11}N{:}C{:}NC_6C_{11}$–HOSu; iv, $C_6H_{11}N{:}C{:}NC_6H_{11}$–HOAc; v, HF, PhOMe, Met; vi, purification on Sephadex G25; vii, autoxidation; viii, affinity chromatography using trypsin–Sepharose

Scheme 68

from the Merrifield resin in the absence of side-chain benzyl esters.[274] The 2-dimethylaminoethyl function is removed from the resulting protected peptide esters by treatment with DMF–water at room temperature (the hydrolysis of dimethylaminoethyl esters, which have a low solubility, can be accelerated by addition of imidazole, or alternatively aqueous sodium bicarbonate can be used), to yield the corresponding peptide acids (Scheme 69). The series of reactions comprising transesterification and



Reagents: i, $Me_2NCH_2CH_2OH$, Me_2NCHO (1 : 1 v/v); ii, aq. Me_2NCHO

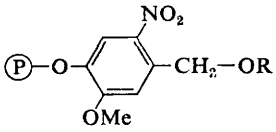
Scheme 69

ester hydrolysis proceed with 0.3 ± 0.1% racemization of the *C*-terminal amino-acid. The transesterification is very slow when the peptide contains a *C*-terminal proline residue but can be expedited by addition of thallium(I) 2-dimethylaminoethoxide. Benzyloxycarbonyl and t-butoxycarbonyl groups in the object peptide are not disturbed. This mild, virtually racemization-free procedure appears most promising, particularly so for 'fragment condensers' and for peptides not containing aspartic or glutamic acids.

Boron tris(trifluoroacetate) (92) in trifluoroacetic acid or dichloromethane can be used for final deprotection [160] and release of peptides from the Merrifield resin. It is to be expected that trifluoromethanesulphonic acid [161] (93) will also find favour for this purpose. Both of these reagents were discussed earlier (pp. 302 and 303).

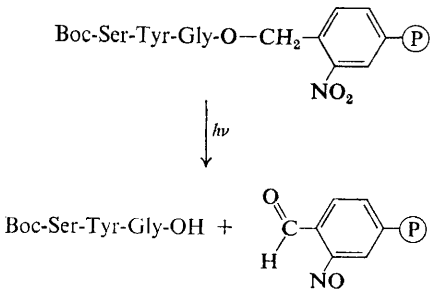
Protected peptides can be removed from an *o*-nitrobenzyl resin by photolysis.[314] The peptidyl-polymer is suspended in methanol, in the absence of air, and irradiated for 12—17 h at 350 nm (wavelengths below 320 nm being filtered out). The yields of released peptides are good; the method does not cause racemization, nor decompose aromatic amino-

[314] D. H. Rich and S. K. Gurwara, *J.C.S. Chem. Comm.*, 1973, 610.



(151) R = saccharide units

acids (Scheme 70). The polymer (151) is light-sensitive[315] and has been used as a support for oligosaccharide synthesis. Final deprotection is achieved by irradiation at >320 nm in dioxan suspension.



Scheme 70

Monitoring the Progress of Solid-phase Synthesis. Because of the multiple uncertainties associated with the solid-phase method of peptide synthesis, rapid and reliable analytical control, both of coupling and deprotection reactions, is increasingly desirable. The polymer support may distort the analytical results, either by masking a population of the functional groups being determined (*e.g.* amino-groups), or by trapping by-products of the coupling reaction, which react with the assay reagent but in a time-dependent fashion. One of the major difficulties is the lack of analytical standards, which necessitates the use of several analytical techniques in parallel. A comparison of the accuracy of the various monitoring techniques has been made,[316] using the synthesis of hexaleucine as a model.

The Dorman method measures the amount of chloride bound to the resin by conversion of free amino-groups into their corresponding hydrochlorides, on treatment of the resin with pyridinium chloride. The chloride ion is displaced from the resin with triethylamine and measured titrimetrically. In the case of the hexaleucine model, however, the method is not sensitive to loss of peptide from the resin. Perhaps a compensating re-esterification occurs.[316] The important fact emerged that repeated titration of peptidyl-resin led to a significant drop in yield in the synthesis

[315] U. Zehavi and A. Patchornik, *J. Amer. Chem. Soc.*, 1973, **95**, 5673.
[316] P. Fankhauser, B. Schilling, P. Fries, and M. Brenner, ref. 3, p. 153.

of acyl-carrier protein-(63—74)-dodecapeptide.[317] The speed and sensitivity of the Dorman method have been improved by the use of pyridinium chloride-36, which enables released chloride to be determined by measurement of the radioactivity, and allows the method to be automated.[317] However, its use is complicated by the variation in measured background chloride; this presumably occurs by dynamic solvation changes of the resin matrix affecting the accessibility of quaternary ammonium groups[318] (arising from residual chloromethyl groups). However, there is evidence that this difficulty can be overcome by using a different resin.[317]

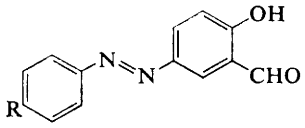
A method has been developed for the determination of residual amino-groups on the polymer, using *N*-[^{14}C]acetylimidazole.[316] Combustion of the acetylated resin enables the incorporated radioactivity to be measured as liberated $^{14}CO_2$. However, discrepancies between this method and those from the use of [^{14}C]acetic anhydride probably lie in the ability of the former reagent to react more readily with hydroxymethyl groups on the polymer.

t-Butoxycarbonyl groups which incorporate a radioactive label can be used for monitoring solid-phase synthesis.[318] *N*-2-[^{14}C]-t-Butoxycarbonyl-amino-acid is used in the coupling step, the progress of the reaction being followed by measuring the decrease in radioactivity of aliquots of the supernatant liquor (but *care* ionic binding to the resin). The possibility of determining the peptide content of a resin with [^{35}S]sulphuric acid has been investigated.[318]

Resin-bound free amino-groups can be determined using one cycle of the Edman degradation employing [^{14}C]phenyl isothiocyanate and measuring the radioactivity released in the phenylthiohydantoin,[319] or determining the nature and amount of released phenylthiohydantoin by gas–liquid chromatography.[320] This method can be adapted to measuring the extent of coupling and of deprotection.

The Esko–Karlsson method determines the residual free amino-groups by converting them into a Schiff's base with 2-hydroxy-1-naphthaldehyde. The Schiff's base is decomposed with benzylamine and the released *N*-(2-hydroxy-1-naphthylidene)benzylamine is determined spectrophotometrically. The method gave low results in monitoring the hexaleucine synthesis; it would appear that the concentration of the reagent used is critical and that the formation of Schiff's base is incomplete.[316] 5-Phenylazosalicylaldehyde (152) and 5-*p*-nitrophenyl azosalicylaldehyde (153) can be used to monitor polymer-bound free amino-groups.[321] When free amino-groups are present the resin turns yellow with (152) and red with (153).

[317] W. S. Hancock, D. J. Prescott, P. R. Vagelos, and G. R. Marshall, *J. Org. Chem.*, 1973, **38**, 774.
[318] H. C. Beyerman, P. R. M. van der Kamp, E. W. B. de Leer, W. Maassen van den Brink, J. H. Parmentier, and J. Westerling, ref. 3, p. 138.
[319] W. Geising and S. Hornle, ref. 3, p. 146.
[320] H. D. Niall, G. W. Tregear, and J. Jacobs, ref. 4, p. 695.
[321] G. Losse and R. Ulbrich, East Ger. P. 83 529 (Cl. G 01*n*) (*Chem. Abs.*, 1973, **78**, 84 822).

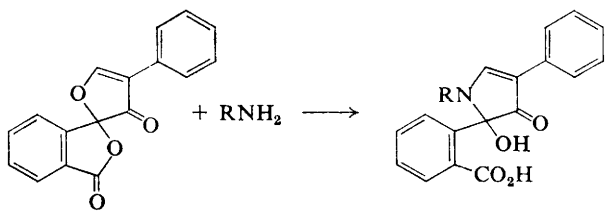


(152) R = H
(153) R = NO_2

The scope and limitations of the automatic monitoring of solid-phase peptide synthesis using potentiometric titration with perchloric acid have been further discussed.[322, 323]

Peptides with or without N-protecting groups can be quantitatively determined using a modification of the Rydon–Smith technique.[324] The peptide is chlorinated with aqueous sodium hypochlorite, the excess being destroyed with hydrogen peroxide, and the solution is treated with *NN'*-dimethyl-4,4'-diaminodiphenylmethane. The blue colour which appears immediately is replaced by an indigo-blue (λ_{max} = 720 nm) and remains stable for several hours. The method has been used for monitoring chromatographic effluents.

Fluorescamine (154)[325] reacts with primary amines giving highly fluorescent products that are detectable in the picomole range (Scheme 71),



(154)

Scheme 71

and it has been applied to monitoring the completeness of solid-phase coupling reactions.[326] Resin-bound amino-groups react rapidly with fluorescamine, giving rise to resin-bound fluorophores which are stable at room temperature, are easily excited by long-wavelength u.v. radiation, and can be estimated visually. The new reagent can detect smaller amounts of resin-bound amino-groups more simply than has been possible heretofore.

An assay procedure has been devised for determining the availability of dicyclohexylcarbodi-imide-activated amino-acids, as a function of pre-

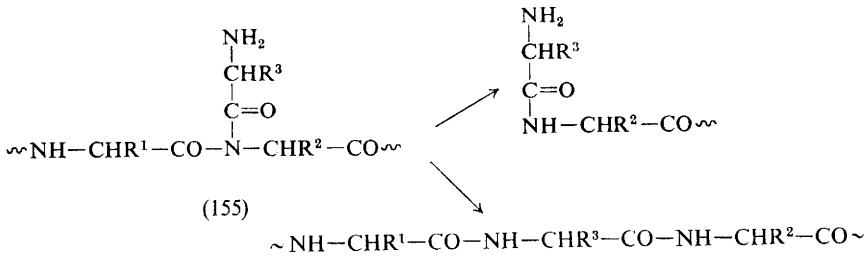
[322] K. Brunfeldt, D. Bucher, T. Christensen, P. Roepstorff, O. Schou, and P. Villemoes, ref. 4, p. 183.
[323] K. Brunfeldt, ref. 5, p. 141.
[324] A. Loffet, J. Gobert, A. Bouteille, and C. Density, ref. 3, p. 249.
[325] S. Udenfriend, S. Stein, P. Bohlen, and W. Dairman, ref. 4, p. 655.
[326] A. M. Felix and M. H. Jimenez, *Analyt. Biochem.*, 1973, **52**, 377.

incubation time, for solid-phase synthesis.[327] The inactivation of activated *N*-t-butoxycarbonyl-L-leucine is significant; in fact in this case there is little to be gained in prolonging a coupling cycle above two hours, unless fresh activated component is added.

Side-reactions in Solid-phase Synthesis. The possibility that the formation of pyroglutamyl derivatives is a source of chain termination in solid-phase synthesis has been verified.[328] t-Butoxycarbonyl-L-glutaminyl-peptide resin can be deprotected cleanly using 4M hydrogen chloride in dioxan;[291, 328] however, the use of trifluoroacetic acid in dichloromethane leads to the formation of some pyroglutamyl derivative. Glutaminyl-peptide resin is stable to DMF, and to DMF containing triethylamine; it is, however, unstable in the presence of acetic acid.

The formation of failure sequences due to intra-chain aminolysis liberating dioxopiperazines,[322, 329–331] inter-chain aminolysis,[330–332] racemization of N_{im}-benzylhistidine derivatives,[329] and chain termination by acetylation with acetic acid leaching out of Teflon components of the reaction system into the coupling mixture [322] have received further mention.

It has been pointed out that under the conditions of solid-phase coupling, where an excess of acylating agent is used, acylation of the peptide backbone is a possibility.[333] After the removal of protecting groups, in addition to the diacylimide (155) yielding fragment peptides, an amino-acyl rearrangement is possible (Scheme 72). A preliminary study using the dicyclohexyl-



Scheme 72

carbodi-imide-mediated coupling of t-butoxycarbonylphenylalanine to benzoyl-oligoleucyl-resins reveals that phenylalanine is indeed incorporated. In the case of benzoyl-leucine resin, after coupling, final deprotection, and esterification, the amino-acyl rearrangement product, benzoyl-

[327] A. M. Tometsko, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 886.
[328] H. C. Beyerman, T. S. Lie, and C. J. van Veldhuizen, ref. 3, p. 162.
[329] M. C. Khosla, R. R. Smeby, and F. M. Bumpus, ref. 4, p. 227.
[330] M. Rothe and J. Mazánek, ref. 4, p. 89.
[331] J. Mazánek, H. Rott, and M. Rothe, ref. 5, p. 186.
[332] H. C. Beyerman, J. Hirt, E. W. B. de Leer, W. van Vossen, and H. A. Billiet, ref. 5, p. 170.
[333] P. Frankhauser, B. Schilling, and M. Brenner, ref. 5, p. 162.

phenylalanyl-leucine methyl ester, was identified.[333] Full studies of the scope of these undesirable side-reactions are awaited so that their general influence on solid-phase synthesis can be evaluated.

An analytical study of the synthesis of acyl carrier protein-(63—74)-dodecapeptide has led to the suggestion that the growing polymer suffers dynamic solvation changes at different stages of the synthesis, which gives rise to alterations in the accessibility of reactive groups.[317] These changes of the polymer matrix lead to the sequence-dependent problems of solid-phase synthesis, such as truncation and deletion sequences, or inability to terminate unreacted chains. These considerations led to an improved synthesis of the dodecapeptide. The problem of change in solvation as the hydrophilicity/lipophilicity of the growing polymer alters could in principle be attacked by using a polymer support whose solvation properties would be similar to those of the polypeptide being synthesized.[269]

Apparatus and Instrumentation. Two systems for manual solid-phase synthesis [334, 335] and a reaction vessel suitable for continuous-flow spectrophotometric monitoring [313] have been illustrated. A further mention of the centrifugal porous Teflon solid-phase reactor ('laundering machine') has been made,[44] and an interesting review of automation in solid-phase synthesis has appeared.[323]

To complete this section on solid-phase synthesis it seems appropriate to refer to the art and virtue of understatement and over-characterization.[336]

Other Repetitive Methods.—The repetitive method of synthesis,[173] using '*in situ*' coupling of *o*-nitrophenyl esters and a 'non-solvent' to wash away co-products and side-products, has been discussed under 'active esters'. The controlled stepwise synthesis of peptides using *N*-carboxy-anhydrides has been discussed earlier.[212, 213, 216] The repetitive excess mixed anhydride method of coupling has also been discussed earlier.[229–233]

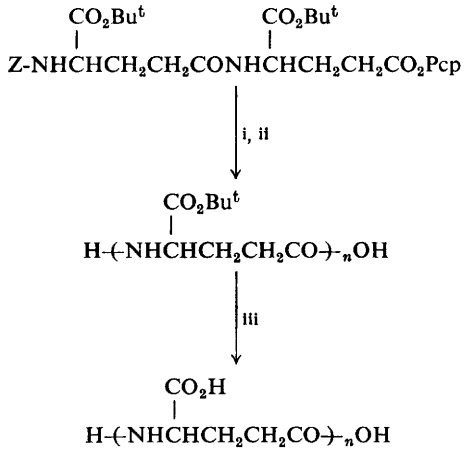
6 Synthesis of Polymeric Models for Studies in Protein Chemistry

Poly(amino-acids).—Poly-L-histidine has been prepared by polymerizing the dihydrobromide of histidine azide in the presence of triethylamine.[337] Poly-DL-lysine has been prepared from strictly racemic N^{ϵ}-benzyloxycarbonyl-lysine *N*-carboxy-anhydride. The extent of isotactic L-blocks of residues occurring in the polymer was found to be 3% using enzymic digestion with pronase and taka-diastase.[338] Poly-L-thialysine (M_w = 5000—10 000) has also been prepared by polymerization of the *N*-carboxy-anhydride of N^{ϵ}-benzyloxycarbonyl-L-thialysine. The protected polymer

[334] J. D. Young, W. Voelter, M. Shimizu, C. Y. Leung, W. J. Peterson, and E. Benjamini, ref. 2, p. 35.
[335] W. K. Park, J. Asselin, and L. Berlinguet, ref. 2, p. 49.
[336] Symposium discussion summarized by J. Meienhofer, ref. 4, p. 261.
[337] W. J. Bailey and H. J. Booth, *Nuova Chim.*, 1973, **49**, 81 (*Chem. Abs.*, 1973, **79**, 137 483).
[338] W. Darge, R. Sass, and W. Thiemann, *Z. Naturforsch.*, 1973, **28c**, 116.

was deprotected with hydrogen bromide in a mixture of acetic acid and dichloroacetic acid, since it is insoluble in the normal reagent.[339]

Poly-(γ-L-glutamic acid) has been prepared as shown in Scheme 73; it has properties in satisfactory agreement with those previously reported,



Reagents: i, H_2–Pd/C, HCl in MeOH; ii, Me_2NCHO, NEt_3; iii, CF_3CO_2H

Scheme 73

and with the enantiomeric natural product from *Bacillus anthracis* and *Bacillus subtilis*.[340]

Sequential Polypeptides.—Collagen models continue to be a fruitful area of study. The practically monodisperse sequential polypeptides (156; n = 5 or 10) have been prepared by the Merrifield technique, by repetitive stepwise coupling of the tripeptide fragment shown in Scheme 74.[341] These oligomers, like (157; n = 5 or 10),[342, 343] form a triple helical structure analogous to that of collagen. The hydroxyproline residues of the former polymers appear to stabilize this structure. A similar repetitive stepwise coupling of a tripeptide active ester fragment, but using solution techniques, has been applied to the synthesis of the collagen-like oligomers (158), for studies of enzymic hydroxylation with protocollagen-proline hydroxylase,[344] and (159), for antigenicity studies.[345] Other sequential collagen models

[339] Š. Štokrová, M. Havránek, P. Hermann, and K. Bláha, *Coll. Czech. Chem. Comm.*, 1973, **38**, 902.

[340] D. Balasubramanian, C. C. Kalita, and J. Kovacs, *Biopolymers*, 1973, **12**, 1089.

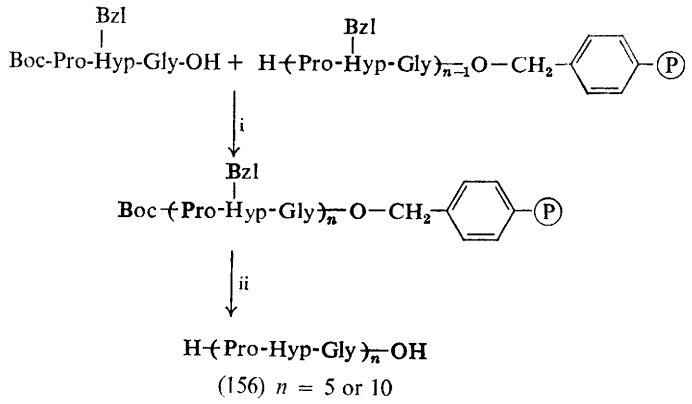
[341] S. Sakakibara, K. Inouye, K. Shudo, Y. Kishida, Y. Kobayashi, and D. J. Prockop, *Biochim. Biophys. Acta*, 1973, **303**, 198.

[342] H. Okuyama, N. Tanaka, T. Ashida, M. Kakudo, S. Sakakibara, and Y. Kishida, *J. Mol. Biol.*, 1972, **72**, 571.

[343] S. Sakakibara, Y. Kishida, and S. Aimoto, ref. 4, p. 293.

[344] K. Okada, Y. Kikuchi, Y. Kawashiri, and M. Hiramoto, *F.E.B.S. Letters*, 1972, **28**, 226.

[345] R. Fairweather and J. H. Jones, *Immunology*, 1973, **25**, 241.



Reagents: i, $C_6H_{11}N{:}C{:}NC_6H_{11}$; ii, HF

Scheme 74

which have been prepared by polymerizing tripeptide 2,4,5-trichlorophenyl esters with triethylamine in DMSO are (160) and (161).[346] Tripeptide *p*-nitrophenyl ester polymerizations, carried out with triethylamine or *N*-methylmorpholine in DMSO, have been used to prepare (162), (163), and (164), whereas a pentachlorophenyl ester was used to prepare (165).[345] The sarcosine-containing tripeptide pentachlorophenyl esters (166) and (167), however, did not yield high molecular weight products, possibly owing to competing cyclization.[345]

H-(-Pro-Pro-Gly-)$_n$OH
(157; n = 5 or 10)

H-(-Glu-Pro-Gly-)$_n$OH
(158; n = 2—4)

H-(-Ala-Gly-Pro-)$_n$OH
(159; n = 1—4)

H-[-Gly-Glu(OMe)-Pro-]$_n$OH
(160; M = 4300)

H-[-Pro-His(Z)-Gly-]$_n$OH
(161; M = 2700)

H-(-Gly-Pro-Gly-)$_n$OH
(162)

H-(-Gly-Pro-Pro-)$_n$OH
(163)

H-(-β-Ala-Pro-Gly-)$_n$OH
(164)

H-(-Ala-Gly-Pro-)$_n$OH
(165; M_w of fractions 2000, 3000, 6000, 11 000, and 15 000)

H-Sar-Ala-Gly-OPcp
(166)

H-Ala-Gly-Sar-OPcp
(167)

H-(-Ala$_x$-Gly$_y$-)$_n$OH
(168)

[346] V. A. Shibnev, Sh. Kh. Kalikov, M. I. Ismailov, and K. T. Poroshin, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1973, 1874 (*Chem. Abs.*, 1973, **79**, 146 837).

The sequential polypeptides (168) have been prepared by polymerization of monomer pentachlorophenyl esters, as models of silks from *Bombyx mori*.[347]

A whole series of histone and protamine sequential analogues (Scheme 75) has been prepared by polymerizing 2,4,5-trichlorophenyl peptide

H-(-Ala-Ala-Lys-)$_n$OH	H-(-Ala-Ala-Ala-Lys-)$_n$OH
H-(-Lys-Lys-Ala-)$_n$OH	H-(-Lys-Lys-Lys-Ala-)$_n$OH
H-(-Lys-Lys-Gly-)$_n$OH	H-(-Lys-Lys-Ala-Ala-)$_n$OH
H-(-Lys-Arg-Ala-)$_n$OH	H-(-Lys-Lys-Gly-Gly-)$_n$OH
H-(-Lys-Arg-Gly-)$_n$OH	H-(-Orn-Orn-Ala-Ala-)$_n$OH
H-(-Gly-Ala-Arg-)$_n$OH	H-(-Orn-Orn-Gly-Gly-)$_n$OH

Scheme 75

active ester monomers in DMF containing triethylamine.[348–351] Some of these sequential polymers possess bactericidal properties against *Staphylococcus aureus*.[351] The preparation of the polypeptide antigens (169) and (170) has been described.[352]

The problem of racemization in the formation of sequential polypeptides has been discussed.[249] A number of sequential polypeptides (171)—(173) containing L-tyrosine and L-glutamic acid have been prepared using racemization-free 2-hydroxyphenyl esters for the polymerization step.[353] During synthesis of the monomers the hydroxy-group of these esters is protected as a phenacyl ether. The rate of polycondensation in benzene increases markedly on addition of increasing amounts of triethylamine, probably owing to ionization of the hydroxy-group of catechol, and consequent intramolecular general base catalysis. The polymerization gives the highest molecular weight polymers when it is carried out in DMF containing triethylamine, the concentration being one molar. The polymer (171) has also been prepared by the 2,4,5-trichlorophenyl ester method using a dipeptide monomer.[354]

H-(-Phe-Glu-Val-Gly-)$_n$Gly-OMe (169)

H-(-Phe-Asp-Val-Gly-)$_n$Gly-OMe (170)

[347] A. Brack, A. Caille, and G. Spach, *Monatsh.*, 1972, **103**, 1604.
[348] V. A. Shibnev, L. I. Mar'yash, and K. T. Poroshin, *Doklady Akad. Nauk S.S.S.R.*, 1972, **207**, 625 (*Chem. Abs.*, 1973, **78**, 84 805).
[349] V. A. Shibnev and L. I. Mar'yash, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1973, 435 (*Chem. Abs.*, 1973, **79**, 5585).
[350] V. K. Burichenko, K. T. Poroshin, G. F. Kasymova, A. K. Mirzoev, and V. A. Shibnev, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1972, 2597 (*Chem. Abs.*, 1973, **78**, 124 935).
[351] G. F. Kasymova, V. K. Burichenko, E. E. Meitus, and K. T. Poroshin, *Khim. prirod. Soedinenii*, 1973, **9**, 88 (*Chem. Abs.*, 1973, **79**, 53 809).
[352] B. J. Johnson, *J. Pharm Sci.*, 1973, **62**, 1564.
[353] Y. Trudelle, *J.C.S. Perkin I*, 1973, 1001.
[354] M. I. Ismailov and V. A. Shibnev, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1973, 1872 (*Chem. Abs.*, 1973, **79**, 146 835).

H-(-L-Tyr-L-Glu-L-Tyr-L-Glu-)$_n$OH
(171)

H-(-L-Glu-L-Tyr-L-Glu-)$_n$OH
(172; M_w = 19 500)

H-(-L-Glu-L-Glu-L-Tyr-L-Glu-)$_n$OH
(173; M_w = 25 000)

H-(-L-Glu-L-Lys-L-Ala-Gly-)$_n$OH
(174; $n \approx 32$)

H-(-L-Ala-D-Glu-L-Lys-D-Ala-Gly-)$_n$OH
(175; $n \approx 32$)

Sequential polypeptides (174) and (175), with a high degree of polymerization, have been obtained by polymerizing pentachlorophenyl ester peptide monomers as a dispersion in benzene in the presence of triethylamine. A high monomer-to-dispersant ratio is not necessary to obtain material of high molecular weight.[355]

The preparations of the sequential polymers (176) and (177) by the *p*-nitrophenyl ester method,[356] of (178) by the succinimidyl ester method,[357] and of (179) by way of the pentachlorophenyl ester [358] have been reported.

OEt
|
H-(-Glu-Gly-Gly-)$_n$OH
(176; M_n = 6800)

H-(-Glu-Glu-Gly-)$_n$OH
| |
OEt OEt
(177)

H-(-Lys-Ala-Ala-)$_n$OH
(178; M_w = 13 000)

H-(-L-Tyr-L-Asp-L-Ala-Gly-)$_n$Gly-OH
(179)

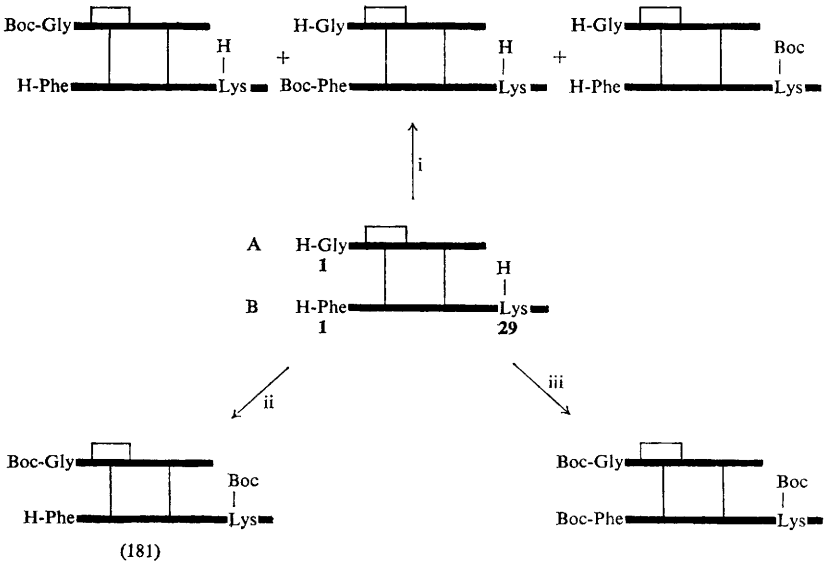
Ala-Lys-Ala
(180)

7 Synthetic Operations with Peptides of Biological Origin

A review of some aspects of partial synthesis with peptides of biological origin has been presented.[359] It includes discussion of fragment couplings of peptides derived from lysozyme, *C*-terminal carboxy-groups being differentiated from side-chain carboxy-groups by trypsin cleavage of material which had been totally esterified with phenyldiazomethane or 2,2-dimethyl-1-diazopropane. Structure–function studies with chemically modified insulins have been discussed,[360] and a review has appeared concerning partially synthetic analogues of staphylococcal nuclease-T′ and bovine pancreatic ribonuclease-S′. Specific N^{ϵ}-acylation with t-butoxycarbonyl azide of the model lysine-containing peptide (180) and of β-melanocyte-stimulating hormone has been studied.[361]

[355] A. R. Zeiger, A. Lange, and P. H. Maurer, *Biopolymers*, 1973, **12**, 2135.
[356] W. Barton Rippon, H. H. Chen, J. M. Anderson, and A. G. Walton, *Biopolymers*, 1972, **11**, 1411.
[357] A. Yaron, N. Tal (Turkeltaub), and A. Berger, *Biopolymers*, 1972, **11**, 2461.
[358] B. J. Johnson and C. Cheng, *J. Medicin. Chem.*, 1973, **16**, 415.
[359] R. E. Offord, ref. 5, p. 52.
[360] D. Brandenberg, W.-D. Busse, H.-G. Gattner, H. Zahn, A. Wollmer, J. Glieman, and W. Puls, ref. 5, p. 270.
[361] S. Lande, ref. 2, p. 363.

N^{α}(A1)-N^{ε}(B29)-Bis-t-butoxycarbonyl-insulin (181) is prepared directly by the action of t-butoxycarbonyl azide on insulin in DMF at pH 9.5—10.[362] Reinvestigation[363] of this reaction shows that the extent of reaction of the acylating reagent with the three nucleophilic centres of insulin is dependent on the excess of reagent used, on the reaction time, and on the nature of the base. Under the appropriate conditions the three possible mono-t-butoxycarbonyl derivatives, one of three possible bis-t-butoxycarbonyl derivatives, and the tris-derivative can be prepared and separately isolated (Scheme 76). The stage is thus set for even more elaborate syn-



Reagents: i, Me_2NCHO, NEt_3 (30 equivalents), Boc—N_3 (120 equivalents) at 38 °C for 2 h; ii, as for i, but with Boc—N_3 (750 equivalents); iii, as for ii, but for 6 h

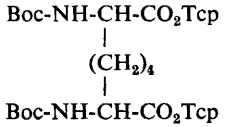
Scheme 76

thetic manipulations of the insulin molecule. A number of N^{α}(A1)-acyl derivatives[364] and a number of N^{α}(A1)-aminoacyl derivatives[365] of bis-t-butoxycarbonyl-insulin (181) have been prepared, the latter acylations being accomplished *via* the corresponding azides.

Several studies central to the issue of regeneration of insulin activity have concentrated attention on facilitating re-formation of the disulphide bridges from chains which are covalently moored alongside each

[362] R. Geiger and D. Langner, Ger. Offen. 2 162 164 (Cl. C 07*c*, A 61*k*) (*Chem. Abs.*, 1973, **78**, 136 665).
[363] D. Levy, *Biochim. Biophys. Acta*, 1973, **328**, 107.
[364] R. Geiger, Ger. Offen. 2 120 269 (Cl. C 07*c*) (*Chem. Abs.*, 1973, **78**, 30 197).
[365] R. Geiger, Ger. Offen. 2 120 268 (Cl. C 07*c*) (*Chem. Abs.*, 1973, **78**, 30 196).

other.[134, 366–369] After demonstrating that N^{α}(B1)-trifluoroacetyl-insulin could be cross-linked between its N^{α}(A1) and N^{ϵ}(B29) amino-groups with the bifunctional reagent (182), the separate insulin A-chain tetrasulphonate and N_{α}(B1)-trifluoroacetyl-insulin B-chain disulphonate were successively made to react with (182) to establish a cross-link (see Scheme 4, p. 265).

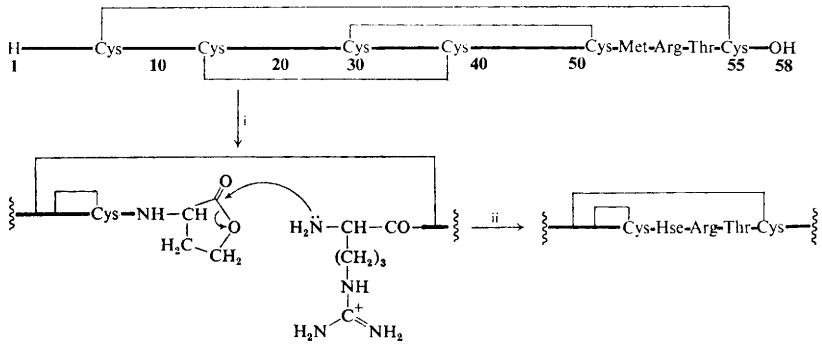


(182)

Reduction of this derivative followed by re-oxidation gave a high yield of the original cross-linked insulin derivative, which is converted into N^{α}(B1)-trifluoroacetyl-insulin by removal of the t-butoxycarbonyl groups and application of the Edman degradation to remove the cross-link. The yield of mono-trifluoroacetyl-insulin from the isolated chains is 20—40%.[366]

This important study will clearly pave the way to improved insulin syntheses. Similar studies in which the N^{α}(A1) and N^{ϵ}(B29) amino-groups are cross-linked with either an adipic,[367, 368] succinic,[134] or suberic acid residue [134, 369] have been reported.

[52-Homoserine]-basic pancreatic trypsin inhibitor has been prepared by chain cleavage and resynthesis in a novel fashion,[370] making use of the intrinsic reactivity of homoserine lactone towards amine nucleophiles. The lactone was produced by cyanogen bromide cleavage of the methionyl-arginine (52–53) bond in the intact protein (Scheme 77). The success of this



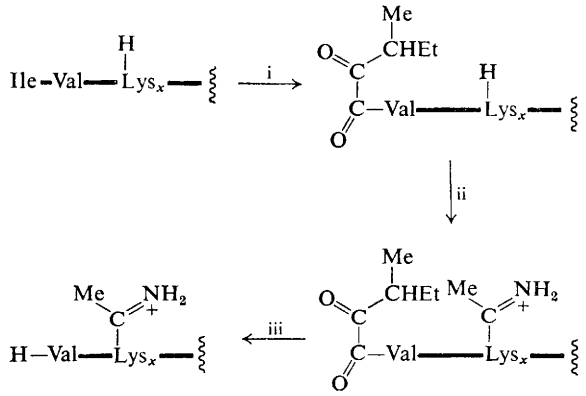
Reagents: i, CNBr; ii, pH 6.3, $t_{\frac{1}{2}} \simeq 30$ h

Scheme 77

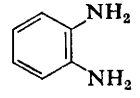
[366] R. Geiger and R. Obermeier, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 60.
[367] D. Brandenburg and A. Wollmer, *Z. physiol. Chem.*, 1973, **354**, 613.
[368] D. Brandenburg and W. Puls, Ger. Offen. 2 209 836 (Cl. C 07*c*) (*Chem. Abs.*, 1973, **79**, 146 856).
[369] D. Brandenburg, W. Schermutzki, and H. Zahn, *Z. physiol. Chem.*, 1973, **354**, 1521.
[370] D. F. Dyckes, T. Creighton, and R. C. Sheppard, *Nature*, 1974, **247**, 202.

resynthesis is due partly to the existence of neighbouring disulphide bonds, which hold the two fragment chains together, but the reaction also has potential for bimolecular combination of natural and synthetic fragments.

Removal of the *N*-terminal residue of a protein and subsequent selective acylation of N^{ε}-amino-groups provides an amino-component suitable for partial synthesis. This has been achieved in the case of trypsin,[359, 371] the removal of the *N*-terminal residue being accomplished by transmination. The product contains an *N*-terminal oxo-acyl group which enables N^{ε}-amino-groups to be selectively acylated. The *N*-oxo-acyl group is removed by quinoxaline formation (Scheme 78), thus providing a protected amino-



Reagents: i, C_5H_5N, HOAc, $Ni(OAc)_2$, NaO_2CCHO, $CaCl_2$; ii, $MeC(:\overset{+}{N}H_2)OEt$; iii,



Scheme 78

component suitable for coupling. This procedure commends itself on account of its mildness and will be particularly useful for proteins which are labile in the presence of anhydrous trifluoroacetic acid, for then the alternative Edman degradation cannot be applied.

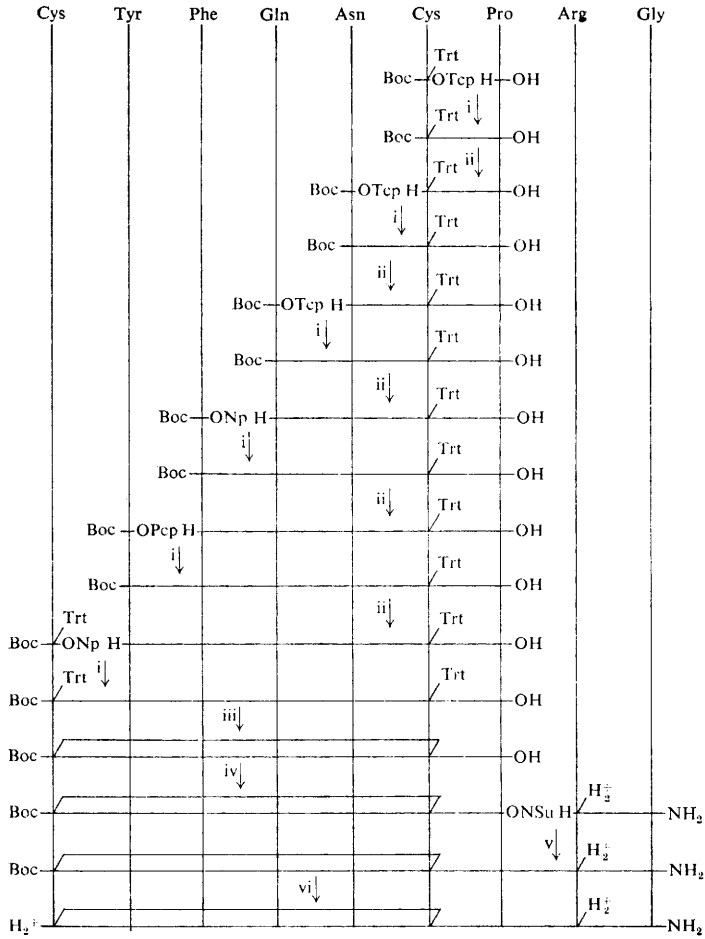
PART II: Syntheses Achieved *by J. H. Jones*

As in the corresponding sections of previous volumes, the discussion here is highly selective, and the reader is referred to the list which appears as an Appendix to this chapter for a comprehensive set of references to the year's literature.

Arginine Vasopressin.—Many syntheses of arginine vasopressin have been described, most of them using *S*-benzyl protection with oxidative disul-

[371] D. Webster and R. E. Offord, *Biochem. J.*, 1972, **130**, 315.

phide bridge formation after vigorous deprotection as the final step. A new synthesis [1] (Scheme 1) has now been reported which depends on a fragment condensation, with the major fragment containing the disulphide bridge of the final hormone (*cf.* recent oxytocin [2] and calcitonin [3] syn-



Reagents: i, active ester–DMF–*N*-methylmorpholine; ii, 2M-HCl in AcOH–dioxan; iii, I_2–AcOH; iv, DCCI–HONSu; v, DMF–*N*-methylmorpholine; vi, 90% CF_3CO_2H

Scheme 1

[1] D. A. Jones, jun., R. A. Mikulec, and R. H. Mazur, *J. Org. Chem.*, 1973, **38**, 2865.
[2] M. Mühlemann, M. I. Titov, R. Schwyzer, and J. Rudinger, *Helv. Chim. Acta*, 1972, **55**, 2854.
[3] *e.g.* B. Riniker, M. Brugger, B. Kamber, P. Sieber, and W. Rittel, *Helv. Chim. Acta*, 1969, **52**, 1059.

theses), the disulphide bridge having been formed under very mild conditions by oxidation of a bis-*S*-trityl peptide with iodine in acetic acid (*cf.* work in the calcitonin [4] field). As well as making the conditions of the late stages of the synthesis mild, this strategy enabled problems attending the arginine residue to be evaded by introducing it at the penultimate stage. The overall yield of the synthesis after purification of the final product was 11% based on the t-butoxycarbonyl-*S*-tritylcysteine trichlorophenyl ester starting material.

The success of a sequence involving repetitive t-butoxycarbonyl removal steps with the preservation of *S*-trityl protection may be of some general interest.* Trifluoroacetic acid (90%) could be used for such stages with only slight net *S*-detritylation, since re-*S*-tritylation occurred on evaporation of the acid during work-up. Since, however, 90% trifluoroacetic acid caused the formation of a pyroglutamyl peptide at the tetrapeptide level in Scheme 1, other acidic conditions were explored: 2M-HCl in 1 : 2 dioxan–acetic acid for 5 min at room temperature was used to remove t-butoxycarbonyl groups from the *S*-trityl peptides to give the corresponding hydrochlorides in nearly quantitative yields.

Corticotropic Peptides containing Radioactive Labels.—Highly radioactive corticotropic peptides with specifically located labels are required for a range of biological experiments. Two preparations of such materials have been published this year.

[(3,5-$^3H_2$)-*Tyr*^{23}]-*β-Corticotropin*-(1—24)-*Tetracosapeptide.* In this case,[6,7] radioactivity was specifically introduced by iodination of a partial sequence containing only one susceptible residue, followed by completion of the peptide chain and catalytic tritium–iodine exchange as outlined in Scheme 2: the product had a specific activity of *ca.* 45 Ci mmol^{-1} in several runs.

[*Phe*^{2}, (4,5-$^3H_2$)-*Nva*^{4}]-*β-Corticotropin*-(1—24)-*Tetracosapeptide.* Here [8] the radioactivity was introduced by catalytic tritiation of a fully protected complete sequence containing an unsaturated side-chain (Scheme 3): after deprotection *etc.*, peptide with specific activity of 7.42 Ci mmol^{-1} was obtained. However, although most of the tritium was shown to be specifically located in the norvaline residue, some 5% had entered histidine-6.

[4] B. Kamber, H. Brückner, B. Riniker, P. Sieber, and W. Rittel, *Helv. Chim. Acta*, 1970, **53**, 556; see also R. Geiger and W. König, Ger. Offen. 1 917 939 (1970) (*Chem. Abs.*, 1970, **73**, 131 307k).

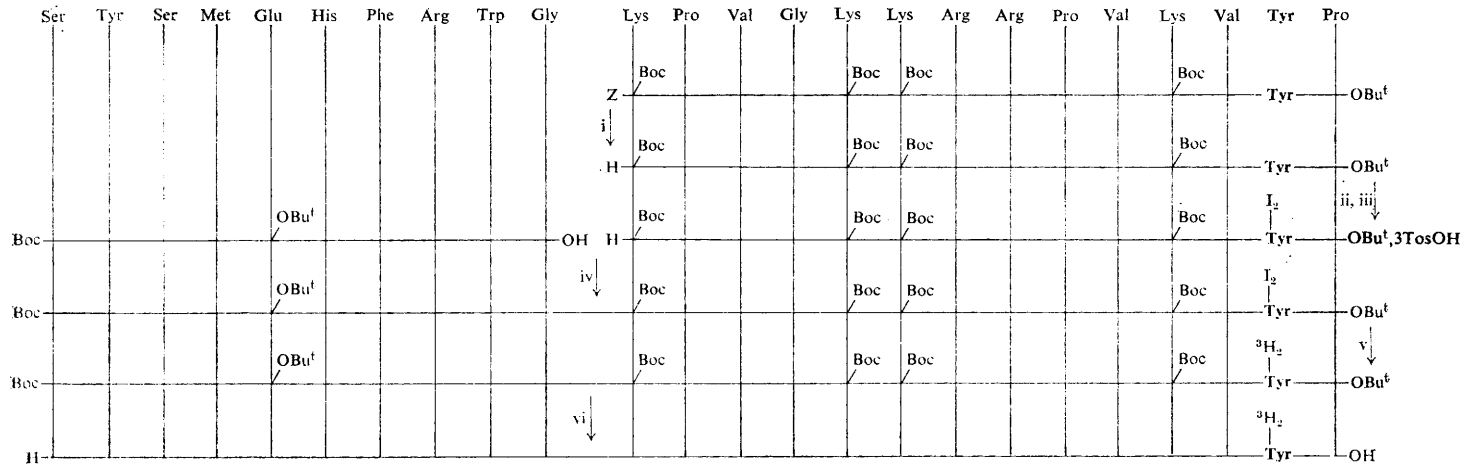
[5] K. Hammerström, W. Lunkenheimer, and H. Zahn, *Makromol. Chem.*, 1970, **133**, 41.

[6] 'Peptides 1972', Proceedings of the 12th European Peptide Symposium, Reinhardsbrunn, German Democratic Republic, September 1972, ed. H. Hanson and H.-D. Jakubke, North-Holland Publishing Co., Amsterdam and London, 1973.

[7] D. E. Brundish and R. Wade, ref. 6, p. 227; *J.C.S. Perkin I*, 1973, 2875.

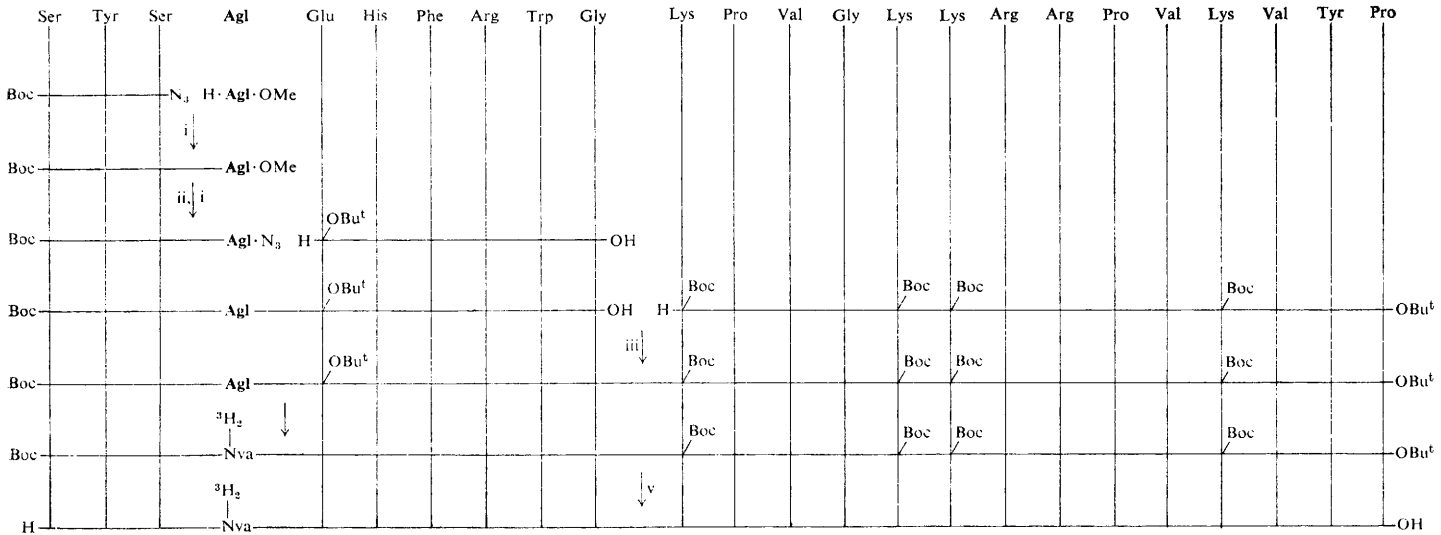
[8] R. Schwyzer and G. Karlaganis, *Annalen*, 1973, 1298.

* The α-Boc, *S*-Trt combination has been shown [5] to be unsuitable for use in Merrifield solid-phase work.



Reagents: i, H_2–Pd/C–AcOH; ii, ICl–AcOH; iii, TosOH–MeOH; iv, DCCI–HONSu–DMF; v, $^3H_2$–Pd/C–Rh/$CaCO_3$–DMF; vi, 90% CF_3CO_2H

Scheme 2



Reagents: i, Honzl–Rudinger procedure; ii, N_2H_4–MeOH; iii, DCCI–HOBt; iv, $^3H_2$–Pd/C; v, HF

Scheme 3 (Agl = allylglycine)

[Har[8]]-β-Corticotropin-(1—24)-Tetracosapeptide.*—This analogue (1) has been synthesized [9] by an overall approach which is similar to that used by Schwyzer [10] for peptides of this class. The 'delicate' intermediate (2) proved

H-Ser-Tyr-Ser-Met-Glu-His-Phe-**Har**-Trp-Gly-Lys-Pro-Val-Lys-Lys-Arg-Arg-Pro-Val-Lys-Val-Tyr-Pro-OH

(1)

Boc-Phe-Har(NO_2)-Trp-Gly-OMe

(2)

very troublesome: the best method of preparation was a 2 + 2 condensation using Mitin's [11] triphenyl phosphite method (the possibility of racemization was ruled out by comparison with a sample more laboriously obtained by a fully stepwise approach). The final complete removal of t-butyl-based protective groups required an unusually extended exposure to acid – it appears to be general that acidolysis of t-butyl ethers near *N*-terminal residues is retarded because of the adjacent protonated amino-group (*cf.* experience [12] in the calcitonin field).

Insulin and Proinsulin.—Katsoyannis and his colleagues have described in detail syntheses of an [Ala[6,11]]-insulin A chain,[13] a [Hcy[6,11]]-A chain,[14]† and of a B chain lacking the four *C*-terminal residues.[15] Fragment condensation approaches were used with sodium–ammonia for final deprotection. In each case combination with a complementary chain from a natural source gave an active analogue, indicating that neither the 6–11 disulphide bridge in the A chain nor the B chain *C*-terminal tetrapeptide sequence is critical for biological acitivity. Among the practical points worthy of note is that benzyloxycarbonyl-ω-tosylarginine was coupled [15] by means of Woodward's reagent K to a tetrapeptide ester (full details are given) to give an 81% yield of pure pentapeptide: this particular carboxyl component is often very troublesome because of lactam formation.

Katsoyannis has also reported a new synthesis of human insulin B chain:[16, 17] in this synthesis *S*-benzhydryl, ω-nitro, *im*-tosyl (which proved precarious), benzyl ether, and benzyl ester side-chain protection were used, with liquid hydrogen fluoride for the final deprotection.

[9] G. I. Tesser, A. W. J. Pleumekers, W. Bassie, and I. C. Balvert-Geers, *Rec. Trav. chim.*, 1973, **92**, 1210.
[10] R. Schwyzer and H. Kappeler, *Helv. Chim. Acta*, 1963, **46**, 1550.
[11] Yu. V. Mitin and O. V. Glinskaya, *Tetrahedron Letters*, 1969, 5267.
[12] B. Riniker, M. Brugger, B. Kamber, P. Sieber, and W. Rittel, *Helv. Chim. Acta*, 1959, **52**, 1058.
[13] P. G. Katsoyannis, Y. Okada, and C. Zalut, *Biochemistry*, 1973, **12**, 2516.
[14] A. Cosmatos and P. G. Katsoyannis, *J. Biol. Chem.*, 1973, **248**, 7304.
[15] P. G. Katsoyannis, J. Ginos, A. Cosmatos, and G. P. Schwartz, *J. Amer. Chem. Soc.*, 1973, **95**, 6427.
[16] G. P. Schwartz and P. G. Katsoyannis, *J.C.S. Perkin I*, 1973, 2890.
[17] G. P. Schwartz and P. G. Katsoyannis, *J.C.S. Perkin I*, 1973, 2894.

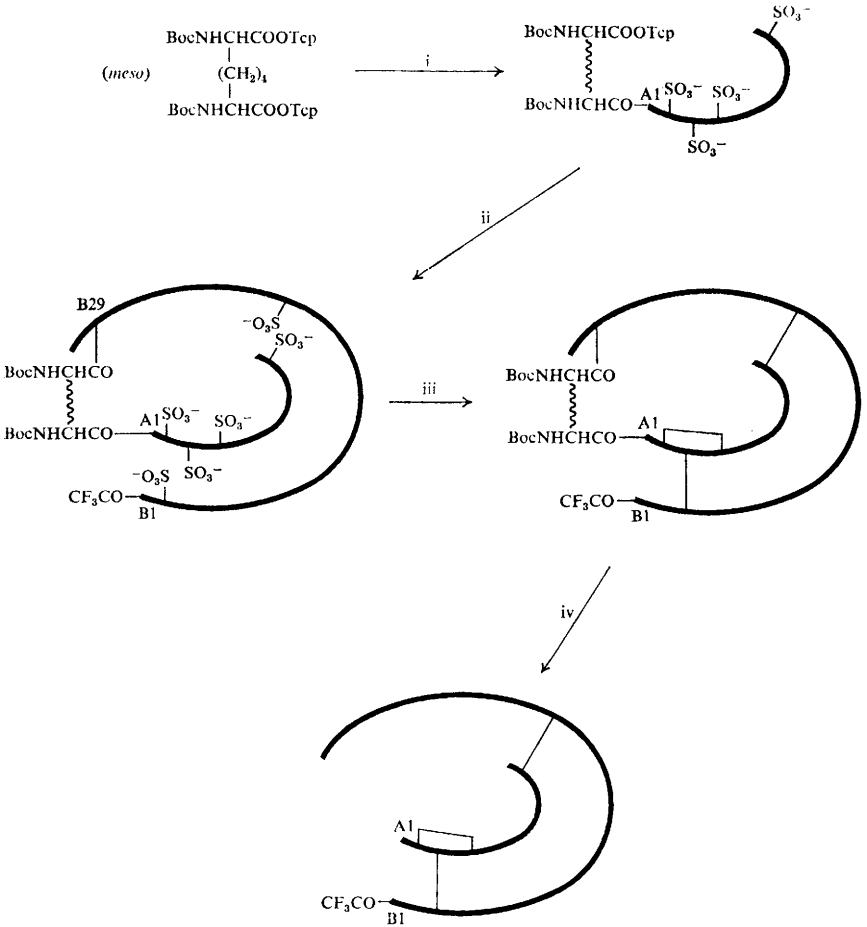
* Har = homoarginine.
† Hcy = homocysteine.

Two syntheses of the human proinsulin connecting peptide have appeared: each has many points of interest, of which we select for special mention the use of amide protection (dimethoxybenzhydryl) in a large peptide,[18] difficulties arising from the use of hydroxybenzotriazole in conjunction with *o*-nitrophenylsulphenyl protection,[18] the use of ammonium thiocyanate for cleavage of *o*-nitrophenylsulphenyl protection,[18] the extensive use of Sephadex LH-20, and the use of N^{α}-trityl protection (which seems to be edging back towards popularity after many years of underemployment) to permit acidolytic N^{α}-deprotection in the presence of t-butyl-based groups.[19] The solubility problems encountered with these syntheses of the human connecting peptide do not seem to have been so great as was the case with some of the approaches to the porcine sequence – in one example [20] connected with the latter, resort to desperate measures (dissolution of amino-components in cresol or molten phenol for coupling) was necessary.

When proinsulin was first discovered, many jumped to the apparently obvious conclusion that the role of the A–B connecting sequence of proinsulin was to ensure that the single chain adopted the conformation which would yield disulphide bridges in the correct positions on oxidation. However, much less complex A–B connections can do this: even a very simple dicarboxylic acid residue [21, 22] inserted between the insulin A and B chains suffices to enable a cycle of reduction and re-oxidation to be performed in good yield. Thus $N^{\alpha A1},N^{\varepsilon B29}$-bisadipoylinsulin [22] can be reduced with mercaptoethanol and then re-oxidized by air to give a 75% recovery of original material: this is a very much higher yield than is obtained with insulin itself. Progression from these discoveries to the design of a reversible cross-linking group which introduces control over the final stages of insulin synthesis has been rapid. Two groups have described their independent preparations and investigations of αA1–εB29-linked α,α′-diaminosuberoylinsulin.[23, 24] In this case the cross-link can be removed by Edman degradation – with the simultaneous removal of phenylalanine-B1 unless this is blocked by *N*-trifluoroacetylation.[24] Scheme 4 [24] shows how this connecting group has been used in what is the first example [24] of a controlled combination of insulin A and B chains to give biologically active material in good yield.

Parathyroid Hormone-(1—34)-Tetratriacontapeptide.—The *N*-terminal sequence of human parathyroid hormone (3) has been synthesized [25] by a

[18] R. Geiger, G. Jager, and W. König, *Chem. Ber.*, 1973, **106**, 2347.
[19] V. K. Naithani, *Z. physiol. Chem.*, 1973, **354**, 659.
[20] R. Geiger, G. Jager, W. König, and A. Volk, *Z. Naturforsch.*, 1969, **24b**, 999.
[21] S. M. L. Robinson, I. Beetz, O. Loge, D. G. Lindsay, and K. Lübke, *Tetrahedron Letters*, 1973, 985.
[22] D. Brandenburg and A. Wollmer, *Z. physiol. Chem.*, 1973, **354**, 613.
[23] D. Brandenburg, W. Schermutzki, and H. Zahn, *Z. physiol. Chem.*, 1973, **354**, 1521.
[24] R. Geiger and R. Obermeier, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 60.
[25] R. H. Andreatta, A. Hartmann, A. Jöhl, B. Kamber, R. Maier, B. Riniker, W. Rittel, and P. Sieber, *Helv. Chim. Acta*, 1973, **56**, 470.



Reagents: i, A-tetra-*S*-sulphonate–*N*-methylmorpholine–DMSO; ii, Tfa-B-di-*S*-sulphonate; iii, mercaptoethanol reduction and air oxidation; iv, removal of Boc groups and one Edman cycle

Scheme 4

fragment condensation approach using t-butyl-based side-chain protection but without substituents on arginine or histidine residues. Selective exposure of N^{α}-amino-groups was performed by mild acidolysis of biphenylyloxycarbonyl or trityl groups using hydrogen chloride in trifluorethanol.

H-Ser-Val-Ser-Glu-Ile-Gln-Leu-Met-His-Asn-Leu-Gly-Lys-His-Leu-Asn-Ser-Met-Glu-Arg-Val-Gln-Trp-Leu-Arg-Lys-Lys-Lys-Gln-Leu-Val-His-Asn-Phe-OH

(3)

Fragments were assembled by dicyclohexylcarbodi-imide–hydroxybenzotriazole or Honzl–Rudinger azide techniques and purified by countercurrent distribution. The free peptide had high hypercalcaemic activity in the thyroparathyroidectomized rat, indicating that the *N*-terminal region of human parathyroid hormone contains the hormonally active region: in this respect the human hormone is like the bovine and porcine materials, despite sequence differences.

Solid-phase Synthesis of High Molecular Weight Polypeptides.—*Acyl Carrier Protein Analogues.* Marshall has now reported in detail [26] on the extension of his group's work on the solid-phase synthesis of acyl carrier protein analogues (see Vol. 4 of these Reports, p. 376): they have been able to identify some apparently inessential parts of this small protein.

Clupeine Z. A brief preliminary outline of a solid-phase synthesis of the Pacific herring protamine clupeine Z (4) is of some interest as the first

H-Ala-Arg_4-Ser-Arg_2-Ala-Ser-Arg-Pro-Val-Arg_4-Pro-Arg_2-Val-Ser-Arg_4-Ala-Arg_4-OH

(4)

recorded attempt at synthesis of a basic nuclear protein or peptide.[27] The material obtained appeared homogenous and identical with the natural protamine on gel electrophoresis, but detailed characterization has not been reported. The overwhelming proportion of arginine in (4) would make a classical synthesis a very difficult undertaking: some simple arginine oligomers have been prepared [28] but no long peptide with such a high proportion of arginine as in (4) has been synthesized.

Lysozyme. A long paper [29] recounting the tribulations of two attempts at solid-phase syntheses of lysozyme has appeared. Notwithstanding serious difficulties and disappointments with the first synthesis – including the loss during the synthesis of 93% of the chains originally attached to the resin, the fact that the greater part of the product had a lower molecular weight than lysozyme, and the ultimate isolation of material with a specific activity of only *ca.* 0.05% – a second attempt was undertaken. Modification of some of the conditions – notably the reagents and procedures used for deprotection – gave somewhat better experience on the second attempt, with isolation after chromatography on chitin of material with a specific activity of 2—3% (or 9—25% of that of native lysozyme which had been similarly abused by subjection to the procedures which were used in the final deprotection and purification of the synthetic material). As an exercise to examine the problems and prospects of solid-phase protein synthesis, this

[26] W. S. Hancock, G. R. Marshall, and P. R. Vagelos, *J. Biol. Chem.*, 1973, **248**, 2424.
[27] H. Yonezawa, S. Terada, H. Aoyagi, and N. Izumiya, *Mem. Fac. Sci., Kyushu Univ., Ser. C*, 1973, **8**, 191 (*Chem. Abs.*, 1973, **79**, 19 093k).
[28] H. Ito, I. Ichikizaki, and T. Ando, *Internat. J. Protein Res.*, 1970, **1**, 59.
[29] J. J. Sharp, A. B. Robinson, and M. D. Kamen, *J. Amer. Chem. Soc.*, 1973, **95**, 6097.

work is a valuable contribution, but it seems unlikely at present that solid-phase work with lysozyme analogues [30] will be able to make a significant or reliable contribution to lysozyme studies.

Ribonuclease T_1. A symposium account [31, 32] of a solid-phase synthesis of ribonuclease T_1 has been published: a small amount of material with 40—50% (depending on the assay) specific activity was isolated.

Staphylococcal Nuclease Systems. Anfinsen's exploration of structure–activity relations in his nuclease-T system (see Vol. 2, p. 177, Vol. 3, p. 257, and Vol. 4, p. 381 of these Reports) with solid-phase preparations continues,[33] and is summarized in the Table.

Table.[a] *Synthetic analogues of fragment P_2 of staphylococcal nuclease*

Peptide	*Does the peptide bind to* P_3?	*Does the peptide activate* P_3?
[Ala[46]]-(6—47)-P_2	Yes	Yes
(6—44)-P_2	Yes	Yes
(6—43)-P_2	Yes	Very slightly?
(6—42)-P_2	Yes	No
[Ala[19]]-(6—47)-P_2	Yes	No
[Asp[43]]-(6—43)-P_2	Yes	No
(9—44)-P_2	Yes	Yes
[Ala[44]]-(9—44)-P_2	Yes	Yes

[a] This Table is a continuation of Table 3 in Vol. 3, p. 259 and the Table, Vol. 4, p.381 of these Reports.

Somatostatin.—Ovine hypothalamic extracts contain a peptide which inhibits the release of growth hormone. This peptide – 'somatostatin' – has been isolated and shown [34] to have the tetradecapeptide structure (5).

H-Ala-Gly-Cys-Lys-Asn-Phe-Phe-Trp-Lys-Thr-Phe-Ser-Cys-OH

(5)

Three solid-phase syntheses [35–37] (one [36] of which is notable for its use of *S*-dimethoxybenzylcysteine and N^{in}-formyltryptophan) and one classical [38]

[30] L. E. Barstow, D. A. Cornelius, V. J. Hruby, T. Shimoda, J. A. Rupley, J. J. Sharp, A. B. Robinson, and M. D. Kamen, ref. 31, p. 231.
[31] 'Chemistry and Biology of Peptides', Proceedings of the 3rd American Peptide Symposium, Boston, June 1972, ed. J. Meienhofer, Ann Arbor Science Publishers, Ann Arbor, Michigan, 1972.
[32] N. Izumiya, M. Waki, T. Kato, M. Ohno, H. Aoyagi, and V. Mitsuyasu, ref. 31, p. 269.
[33] G. R. Sanchez, I. M. Chaiken, and C. B. Anfinsen, *J. Biol. Chem.*, 1973, **248**, 3653.
[34] R. Burgus, N. Ling, M. Butcher, and R. Guillemin, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 684.
[35] J. Rivier, P. Brazeau, W. Vale, N. Ling, R. Burgus, C. Gilon, J. Yardley, and R. Guillemin, *Compt. rend.*, 1973, **276**, *D*, 2737.
[36] D. Yamashiro and C. H. Li, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 882.
[37] D. H. Coy, E. J. Coy, A. Arimura, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 1267.
[38] D. Sarantakis and W. A. McKinley, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 234.

synthesis have been described: the [Ala^{6,11}]-analogue, which has very low but possibly significant activity, has also been prepared [39] by a solid-phase route. There can be little doubt that somatostatin will soon, like the other hypothalamic factors, show that its biological actions include the ability to stimulate peptide chemists into frenzied activity.

[Nle^{26,32}]-Staphylococcal Nuclease-(21—47)-Fully Protected Heptacosapeptide.—As a major step towards a classical synthesis of a staphylococcal nuclease fragment P_2 analogue, Anfinsen has reported [40] the synthesis of a protected fragment comprising more than half the sequence. The use of dicyclohexylcarbodi-imide–*N*-hydroxysuccinimide for fragment condensation in the presence of histidine which was unprotected in the side-chain gave products with lower than expected histidine contents. The chemistry of this side-reaction is obscure, but fortunately the addition of imidazole enticed it away from the peptide. A protected octadecapeptide and the final heptacosapeptide were purified by chromatography on Sephadex LH-20 and LH-60. The Sephadex LH-60 gel was prepared by Pharmacia on an experimental basis only:[41] it is not generally or commercially available at the time of writing, which some may find frustrating in view of Anfinsen's statement that 'Indeed, Sephadex LH-60 appears to offer much promise as a gel for separating high molecular weight, fully protected peptides'.

Substance P.—The first classical synthesis [42] of bovine substance P, which is somewhat unusual in its extensive use of *p*-methoxybenzyloxycarbonyl-N^{α}-protection and the activation at one point by pentachlorophenyl trichloroacetate, is outlined in Scheme 5.

Tuftsin.—A solid-phase preparation of the phagocytosis-stimulating peptide 'tuftsin' (6) has been described.[43]

H-Thr-Lys-Pro-Arg-OH

(6)

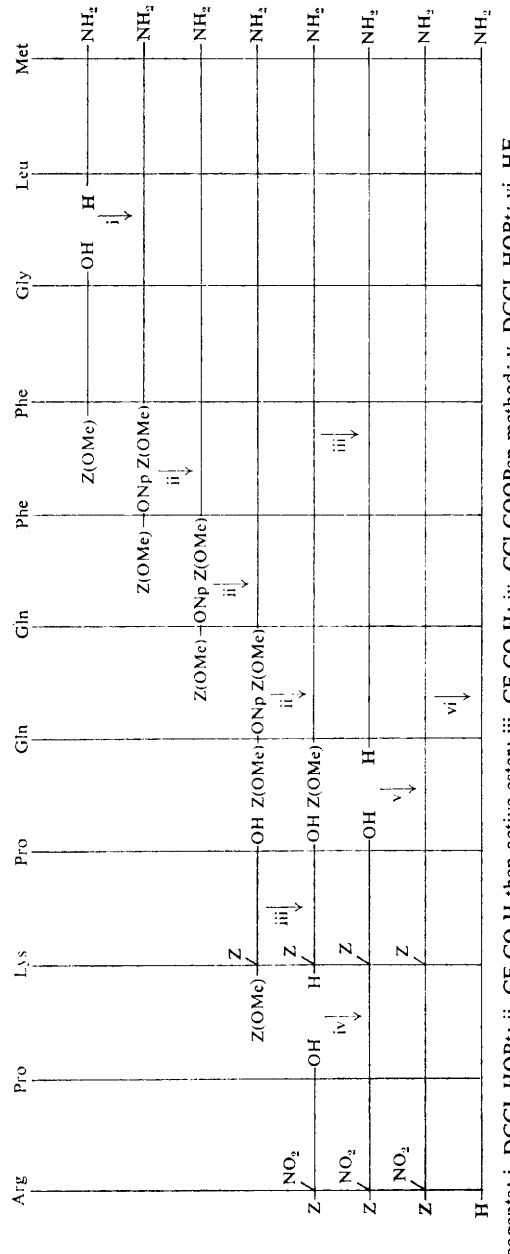
Vasoactive Intestinal Peptide (VIP).—The porcine hormone VIP (7) has been synthesized by a {1—6} + {[7—13] + [(14—17) + (18—28)]} frag-

H-His-Ser-Asp-Ala-Val-Phe-Thr-Asp-Asn-Tyr-Thr-Arg-Leu-Arg-Lys-Gln-Met-Ala-Val-Lys-Lys-Tyr-Leu-Asn-Ser-Ile-Leu-Asn-NH_2

(7)

ment condensation strategy using Honzl–Rudinger azide techniques and minimal side-chain protection (ε-benzyloxycarbonyl for lysine) at the

[39] D. Sarantakis, W. A. McKinley, and H. H. Grant, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 538.
[40] A. R. Zieger and C. B. Anfinsen, *J. Amer. Chem. Soc.*, 1973, **95**, 880.
[41] Personal communication to the Reporter from Pharmacia, 7th June 1973.
[42] H. Yajima and K. Kitagawa, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 682.
[43] K. Nishioka, P. S. Satoh, A. Constantopoulos, and V. A. Najjar, *Biochem. Biophys. Acta*, 1973, **310**, 230.



Reagents: i, DCCI–HOBt; ii, CF_3CO_2H then active ester; iii, CF_3CO_2H; iv, $CCl_3COOPcp$ method; v, DCCI–HOBt; vi, HF

Scheme 5

fragment conjunction stages. The completion of the synthesis has only been outlined,[44] but details[45–48] of the preparations of all the principal fragments have been published. The methods used for these fragment preparations were essentially conventional, but details[47] for an 'EEDQ' coupling [Boc-Thr + Arg(NO_2)-Leu-OMe] are instructive. The principal problem encountered was insolubility, and suspensions had to be used at several deprotection and coupling stages.

Progress towards total synthesis of the related gastric inhibitory peptide has been described.[49–51]

Appendix I: A List of Syntheses Reported during 1973

by J. H. Jones

Naturally Occurring Peptides, Proteins, Analogues, and Partial Sequences.—All syntheses are listed under the natural substance to which they are related.

Peptide	*Ref.*
Acyl carrier protein	
A series of analogues of *E. coli* acyl carrier apoprotein	26
Adrenocorticotropins	
Human ACTH	52, 53
β-Corticotropin-(5—10)-protected hexapeptide	54
[Tyr(3,5-I_2)^{23}]- and [Tyr(3,5-$^3H_2$)^{23}]-β-corticotropin-(1—24)-tetracosapeptide	7
[Phe^2,Nva(4,5-$^3H_2$)^4]-β-corticotropin-(1—24)-tetracosapeptide	8
[Har^8,Lys^{17,18}]-β-corticotropin-(1—18)-octadecapeptide amide	9
[Har^8]-β-corticotropin-(1—24)-tetracosapeptide amide	9
[β-Ala^1,Arg^{11}]-, [β-Ala^1,Gly^{11}]-, and [β-Ala^1,Nle^{11}]-β-corticotropin-(1—17)-heptadecapeptide-4-amino-n-butylamides	55
Angiotensin	
[Val^5]-angiotensin II	56
[Val^8]-, [Ile^8]-, and chlorambucil-des-1-aspartic acid, 8-valine angiotensins I	57

[44] M. Bodanszky, Y. S. Klausner, and S. I. Said, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 382.
[45] M. Bodanszky, Y. S. Klausner, and V. Mutt, *Bio-organic Chem.*, 1972, **2**, 30.
[46] Y. S. Klausner, V. Mutt, and M. Bodanszky, *Bio-organic Chem.*, 1972, **2**, 87.
[47] Y. S. Klausner, C. Y. Lin, V. Mutt, and M. Bodanszky, *Bio-organic Chem.*, 1973, **2**, 345.
[48] Y. S. Klausner and M. Bodanszky, *Bio-organic Chem.*, 1973, **2**, 354.
[49] R. Camble, R. Cotton, A. S. Dutta, J. J. Gormley, C. F. Hayward, J. S. Morley, and M. J. Smithers, ref. 6, p. 200.
[50] K. Kovacs, J. Kovacs-Petres, G. Wendlberger, and E. Wünsch, *Z. physiol. Chem.*, 1973, **354**, 890.
[51] J. Kovacs-Petres, K. Kovacs, K.-H. Diemer, and E. Wünsch, *Z. physiol. Chem.*, 1973, **354**, 894.
[52] L. Kisfaludy, M. Löw, T. Szirtes, I. Schön, M. Sárközi, S. Bajusz, A. Turan, A. Juhász, R. Beke, L. Gráf, and K. Medzihradsky, ref. 31, p. 299.
[53] D. Yamashiro and C. H. Li, *J. Amer. Chem. Soc.*, 1973, **95**, 1310.
[54] Yu. V. Mitin, A. T. Gudkov, N. P. Zapelova, and E. E. Maximov, ref. 6, p. 57.
[55] R. Geiger and H.-G. Schröder, *Z. physiol. Chem.*, 1973, **354**, 156.
[56] S. S. Wang, *J. Amer. Chem. Soc.*, 1973, **95**, 1328.
[57] A. C. M. Paiva, V. L. A. Nouailhetas, M. E. Miyamoto, G. B. Mendes, and T. B. Paiva, *J. Medicin. Chem.*, 1973, **16**, 6.

Peptide	*Ref.*
[Asn^1,Thr^5]- and [Asn^1,Thr(Me)^5]-angiotensins II	58
[Suc^1,Ala^8]-angiotensin II	59
[Suc^1,D-His^6,Ala^8]-angiotensin II	59
[Ala^4,Ile^8]-angiotensin II	59
[Ala^4,D-His^6,Ile^8]-angiotensin II	59
[D-*a*Ile^8]-angiotensin II	59
[D-His^6,D-Ile^8]-angiotensin II	59
[D-His^6]-angiotensin II	59
[β-(2-thienyl)alanine^6], [Phe^6]-, [Met^6]-, [Leu^6]-, [Orn^6]-, [Arg^6]-, and [Glu^6]-angiotensins II	60
[Asp^1,Phe(*p*-F)^4]-angiotensin II	61
[Phe(*p*-F)^4]-angiotensin II	61
[Phe(*p*-F)^8]-angiotensin II	61
Some fragments of [Val^5]-angiotensinamide	62
[Asn^1,Pro^3,Val^5]-angiotensin II	63
[His(πMe)^6]- and [His(τMe)^6]-angiotensins II	64
[Pro^3,Ile^8]-, [Sar^1,Pro^3,Ile^8]-, and [Phe^4,Ile^8]-angiotensins II	65
[Asn^1,Val^5]-angiotensins I and II	66
Angiotensin II *C*-terminal hexapeptide	66
Miscellaneous related peptides and analogues	67—70
Angiotensin-converting enzyme	
A series of inhibitors	71, 72
Apolipoprotein-alanine*	
The -(41—79)-, -(48—79)-, -(55—79)-, and -(61—79)-peptides	73
Bombesin	
A series of analogues and partial sequences	74

[58] E. C. Jorgensen, G. C. Windridge, and T. C. Lee, *J. Medicin. Chem.*, 1973, **16**, 467.
[59] M. C. Khosla, M. M. Hall, R. R. Smeby, and F. M. Bumpus, *J. Medicin. Chem.*, 1973, **16**, 829.
[60] R. J. Freer and J. M. Stewart, *J. Medicin. Chem.*, 1973, **16**, 733.
[61] W. H. Vine, D. A. Brueckner, P. Needleman, and G. R. Marshall, *Biochemistry*, 1973, **12**, 1630.
[62] G. Cipens, A. Pavars, J. Ancans, and V. K. Kibirev, *Khim. prirod. Soedinenii*, 1973, 77 (*Chem. Abs.*, 1973, **79**, 5573e).
[63] R. Vegners and G. Cipens, *Zhur. obshchei Khim.*, 1972, **42**, 2334 (*Chem. Abs.*, 1973, **78**, 58 776j).
[64] P. Needleman, G. R. Marshall, and J. Rivier, *J. Medicin. Chem.*, 1973, **16**, 968.
[65] M. C. Khosla, M. M. Hall, R. R. Smeby, and F. M. Bumpus, *J. Medicin. Chem.*, 1973, **16**, 1184.
[66] H. E. Bleich, R. E. Galardy, M. P. Printz, and L. C. Craig, *Biochemistry*, 1973, **12**, 4950.
[67] D. Regoli, F. Rioux, and W. K. Park, ref. 31, p. 495.
[68] P. Needleman, E. M. Johnson jun., W. Vine, E. Flanagan, and G. R. Marshall, ref. 31, p. 501.
[69] E. C. Jorgensen, G. C. Windridge, K.-H. Hsieh, and T. C. Lee, ref. 31, p. 513.
[70] R. Paruszewski, *Roczniki Chem.*, 1973, **47**, 735 (*Chem. Abs.*, 1973, **79**, 79 165v).
[71] M. A. Ondetti, J. Pluščec, E. R. Weaver, N. Williams, E. F. Sabo, and O. Kocy, ref. 31, p. 525.
[72] J. Pluščec, E. R. Weaver, E. F. Sabo, O. Kocy, and M. A. Ondetti, ref. 6, p. 403.
[73] J. T. Sparrow, A. M. Gotto, jun., and J. D. Morrisett, *Proc. Nat. Acad. Sci., U.S.A.*, 1973, **70**, 2124.
[74] R. de Castiglione, F. Angelucci, V. Erspamer, G. F. Erspamer, and L. Negri, ref. 6, p. 463.

* A major protein constituent of the very-low-density lipoproteins of human plasma: 79 residues.

Peptide	*Ref.*
Bradykinin	75, 76
Some analogues with cyclic imino-acid replacements	77, 78
Bradykinyl-Val-Ala-Pro-Ala-Ser-OH*	79
2-L- and 3-D-pipecolic acid bradykinins	80
Bradykinyl-Gly-Lys-Phe-His-OH†	81
Calcitonin	
Salmon calcitonins-2 and -3	82
Clupeine Z	27
Cytochromes	
A series of protected fragments spanning the entire sequence of bakers' yeast iso-1-cytochrome *c*	83
Cytochrome b_5-(62—93)-ditriacontapeptide	84
Cytochrome b_5-(74—93)-eicosapeptide	85
A series of sequences from the *C*-terminal region of horse heart cytochrome *c*	86
Eledoisin	
[Lys^6,Phe^8]-eledoisin-(6—11)-hexapeptide amide	56
[Asn^5,Thr(Me)^8]-eledoisin-(4—11)-octapeptide amide	87
[Asn^5,8-DL-thiaisoleucine]-eledoisin-(4—11)-octapeptide amide	87
Some depsipeptide analogues	88, 89
Analogues with replacements by *N*-methylamino-acid residues	88, 90
Aza-analogues	91—93
Some analogues of eledoisin partial sequences	94
Encephalitogenic peptides	95

[75] R. Tomatis, R. Ferroni, M. Guarneri, and C. A. Benassi, *J. Medicin. Chem.*, 1973, **16**, 1053.
[76] G. P. Vlasov, A. Yu. Bilibin, N. Yu. Kusnetzova, I. Ditkoskaya, and V. N. Lashkov, *Chem.-Ztg.*, 1973, **97**, 236 (*Chem. Abs.*, 1973, **79**, 42 833k).
[77] J. O'D. McGee, M. H. Jimenez, A. M. Felix, G. J. Cardinale, and S. Udenfriend, *Arch. Biochem. Biophys.*, 1973, **154**, 483.
[78] A. M. Felix, M. H. Jimenez, R. Vergona, and M. R. Cohen, *Internat. J. Peptide Protein Res.*, 1973, **5**, 201.
[79] N. Yanaihara, C. Yanaihara, M. Sakagami, T. Nakayama, and K. Matsumoto, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 616.
[80] K. Neubert, L. Baláspiri, and G. Losse, *Monatsh.*, 1972, **103**, 1575.
[81] T. Yasuhara, M. Hira, T. Nakajima, N. Yanaihara, C. Yanaihara, T. Hasimoto, N. Sakura, S. Tachibana, K. Araki, M. Bessho, and T. Tamanaka, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 1388.
[82] J. Pless, W. Bauer, H. Bossert, K. Zehnder, and S. Guttmann, *Nature New Biol.*, 1972, **240**, 62.
[83] L. Moroder, G. Borin, F. Marchiori, and E. Scoffone, *Biopolymers*, 1973, **12**, (*a*) 477; (*b*) 493; (*c*) 507; (*d*) 521; (*e*) 693; (*f*) 701; (*g*) 721; (*h*) 729.
[84] V. M. Kozhukhovskaya, S. D. L'vova, and R. P. Evstigneeva, *Khim. prirod Soedinenii*, 1972, 514 (*Chem. Abs.*, 1972, **77**, 152 558r).
[85] V. M. Kozhukovskaya, N. N. Monapova, N. L. Alarkon, S. D. L'vova, and R. P. Evstigneeva, *Khim. prirod Soedinenii*, 1972, 510 (*Chem. Abs.*, 1973, **78** 4509f).
[86] S. Kawanishi and S. Sano, *Biochemistry*, 1973, **12**, 3166.
[87] P. Hermann, I. Willhardt, K. Lemke, Š. Štokrova, M. Havránek, and K. Bláha, ref. 6, p. 214.
[88] H. Sugano, *Bull. Chem. Soc. Japan*, 1973, **46**, 2168.
[89] H. Sugano, K. Higaki, and M. Miyoshi, *Bull. Chem. Soc. Japan*, 1973, **46**, 226.
[90] H. Sugano, K. Higaki, and M. Miyoshi, *Bull. Chem. Soc. Japan*, 1973, **46**, 231.
[91] H. Niedrich and J. Pirrwitz, *J. prakt. Chem.*, 1972, **314**, 735.
[92] R. Grupe, B. Baeck, and H. Niedrich, *J. prakt. Chem.*, 1972, **314**, 751.
[93] H. Niedrich and C. Oehme, *J. prakt. Chem.*, 1972, **314**, 759.
[94] H. Immer, V. Nelson, W. Robinson, and M. Gotz, *Annalen*, 1973, 1789.
[95] R. Shapira, F. C. Chou, and R. F. Kibler, ref. 31, p. 225.

* A peptide isolated from the skin of *Rana nigromaculata*.
† A peptide isolated from the skin of *Bombina orientalis* Boulenger.

Peptide	*Ref.*
Ferredoxin	
Clostridium pasteurianum ferredoxin-(40—50)-undecapeptide ethyl ester	96, 97
Gastric inhibitory peptide	
Protected peptides corresponding to positions (1—16) and (17—43) of the proposed sequence of porcine gastric inhibitory peptide	49
Protected peptides corresponding to positions (35—38) and (38—43)	50, 51
Gastrin *C*-terminal tetrapeptide	54
Z-Trp-Met-Asp-Phe-NH_2	98
Z-Trp-Leu-Asp-Phe-NH_2	98
Boc-Ser-, Boc-D-Ser-, and Boc-β-Ala-Trp-Met-Asp-Phe-NH_2	99
Growth hormone	
Human growth hormone-(138—145)-octapeptide	100
Human growth hormone protected *N*-terminal decapeptides	101, 102
Ac-(95—136)-Human growth hormone peptide	103
Human growth hormone *C*-terminal dodecapeptide and its bis-*S*-carbamidomethyl derivative	104
Growth hormone releasing factor	
Two protected pentapeptide fragments	105
Haemoglobin	
A decapeptide which had been isolated from a peptide hydrolysate, 'probably from the α-chain analogous to positions 129—138 in human haemoglobin'	106
Histones	
Protected peptides corresponding to positions -(9—14)-, -(4—9)-, and -(8—14)- of calf thymus histone F2A1	107
Immunoglobulins	
A heptacosapeptide fragment of human immunoglobulin IgG1 γ-chain	108
Insulin	109—111
[$Ala^{A6,A11}$]-ovine insulin	13
[$Hcy^{A6,A11}$]-ovine insulin	14
Des-(B27-B30)-human (porcine) insulin	15

[96] A. Schöberl, M. Rimpler, and U. Dethlefsen, *Naturwiss.*, 1973, **60**, 154.
[97] A. Schöberl, M. Rimpler, and U. Dethlefson, *Annalen*, 1973, 1612.
[98] M. Dabrowska, G. Kupryszewski, B. Lammek, and F. Muzalewski, *Roczniki Chem.*, 1972, **46**, 1295 (*Chem. Abs.*, 1972, **77**, 140 501d).
[99] M. Portelli and G. Renzi, *Farmaco, Ed. Sci.*, 1973, **28**, 316 (*Chem. Abs.*, 1973, **79**, 5574f).
[100] D. Yamashiro and C. H. Li, *J. Org. Chem.*, 1973, **38**, 591.
[101] H. C. Beyerman, E. W. B. de Leer, and J. Floor, *Rec. Trav. chim.*, 1973, **92**, 481.
[102] E. W. B. de Leer and H. C. Beyerman, *Rev. Trav. chim.*, 1973, **92**, 174.
[103] J. Blake and C. H. Li, *Internat. J. Peptide Protein Res.*, 1973, **5**, 123.
[104] D. Yamashiro, R. L. Noble, and C. H. Li, *J. Org. Chem.*, 1973, **38**, 3561.
[105] A. Loffet, ref. 6, p. 242.
[106] J. Brown, H. Brown, and A. Trzeciak, ref. 31, p. 251.
[107] V. K. Burichenko, N. I. Koryakina, L. V. Tashenkova, and V. A. Shibnev, *Khim. prirod. Soedinenii*, 1972, 370 (*Chem. Abs.*, 1972, **77**, 152 562n).
[108] B. J. Johnson, *J. Pharm. Sci.*, 1973, **62**, 1019.
[109] N. A. Yudaev, Yu. P. Shvachkin, M. G. Poznyak, V. P. Fedotov, R. G. Vdovina, E. N. Voluiskaya, M. N. Ryabtsev, V. F. Krivtsov, A. K. Gracheva, *et al.*, *Biokhimiya*, 1973, **38**, 221 (*Chem. Abs.*, 1973, **79**, 5 566e).
[110] Yu. P. Shvachkin, R. G. Vdovina, M. G. Poznyak, E. N. Voluiskaya, M. N. Ryabtsev, V. F. Krivtsov, A. K. Gracheva, S. P. Krasnoshschenkova, V. A. Novoselov, *et al.*, *Zhur. obshchei Khim.*, 1973, **43**, 216 (*Chem. Abs.*, 1973, **78**, 124 863p).
[111] N. A. Yudaev, Yu. P. Shvachkin, M. G. Poznyak, V. P. Fedotov, R. G. Vdovina, E. N. Voluiskaya, M. N. Ryabtsev, V. F. Krivtsov, A. K. Gracheva, *et al.*, *Probl. Endokrinol*, 1972, **18**, 86 (*Chem. Abs.*, 1972, **77**, 165 068f).

Peptide	*Ref.*
Some unsymmetrical disulphide fragments incorporating the A20–B19 bridge	112
Insulin A chain *C*-terminal decapeptide	113, 114
Insulin A chain *N*-terminal pentapeptide	115
Insulin A chain -(7—11)-fully protected pentapeptide	116
Ovine insulin A chain polymeric disulphide	117
Human insulin B chain	16, 17, 118
Miscellaneous related peptides	119—124
Luteinizing hormone releasing factor (LH-RF)	125—135
LH-RF tritiated in the Glp residue	132
[Phe^6]-LH-RF	136, 137
LH-RF-(2—10)-nonapeptide amide and -(3—10)-octapeptide amide	125
A series of shortened peptide amide analogues converging to the *N*-terminus	138

[112] B. Kamber, *Helv. Chim. Acta*, 1973, **56**, 1370.
[113] C. Birr, ref. 6, p. 72.
[114] C. Birr, *Annalen*, 1973, 1652.
[115] H. Wissmann, W. König, and R. Geiger, ref. 6, p. 158.
[116] F. Flor, C. Birr, and T. Wieland, *Annalen*, 1973, 1601.
[117] W. O. Danho and E. Engels, ref. 31, p. 335.
[118] M. I. Titov, Z. A. Adremasova, Zh. D. Bespalova, and L. I. Leont'eva, *Doklady Akad. Nauk S.S.S.R.*, 1973, **209**, 227 (*Chem. Abs.*, 1973, **79**, 5577j).
[119] G. A. Korshunova, G. P. Mishin, and Yu. P. Shvachkin, *Khim. prirod. Soedinenii*, 1973, **9**, 98 (*Chem. Abs.*, 1973, **79**, 19 094m).
[120] G. A. Korshunova, G. P. Mishin, O. M. Vol'pina, and Yu. P. Shvachkin, *Khim. prirod. Soedinenii*, 1973, **9**, 263 (*Chem. Abs.*, 1973, **79**, 42 822f).
[121] G. Weitzel, K. Eisele, V. Schulz, and W. Stock, *Z. physiol. Chem.*, 1973, **354**, 321; G. Weitzel, K. Eisele, H. Gugliemi, W. Stock, and V. Schulz, ref. 6, p. 292.
[122] N. A. Voskova and Yu. P. Shvachkin, *Vestnik. Moskov. Univ., Khim.*, 1972, **13**, 732 (*Chem. Abs.*, 1973, **78**, 84 799p).
[123] G. P. Mishin, L. A. Aleksandrova, and Yu. P. Shvachkin, *Vestnik. Moskov. Univ., Khim.*, 1972, **13**, 612 (*Chem. Abs.*, 1973, **78**, 58 793n).
[124] S. K. Girin, V. F. Krivtsov, and Yu. P. Shvachkin, *Zhur. obshchei Khim.*, 1973, **43**, 1411 (*Chem. Abs.*, 1973, **79**, 66 789b).
[125] W. H. Arnold, W. Hite, and G. Flouret, *J. Medicin. Chem.*, 1973, **16**, 1054.
[126] W. F. White, M. T. Hedlund, R. H. Rippel, W. Arnold, and F. R. Flouret, *Endocrinology*, 1973, **93**, 96.
[127] M. Manning, E. Coy, and A. V. Schally, ref. 31, p. 235.
[128] G. Flouret, W. Arnold, W. Cole, R. Morgan, W. White, M. Hedlund, and R. Rippel, ref. 31, p. 305.
[129] S. Bajus, A. Turán, and I. Fauszt, ref. 31, p. 325.
[130] M. Fujino, T. Fukuda, S. Kobayashi, and M. Obayashi, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 87.
[131] T. Mukaiyama, M. Ueki, and R. Matsueda, ref. 31, p. 209.
[132] D. H. Coy, E. J. Coy, T. W. Redding, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 866.
[133] S. Bajusz, A. Turán, I. Fauszt, and A. Juhász, ref. 6, p. 93.
[134] G. R. Flouret, W. H. Arnold, J. W. Cole, R. L. Morgan, W. F. White, M. T. Hedlund, and R. H. Rippel, *J. Medicin. Chem.*, 1973, **16**, 369.
[135] Y. Yanaihara, C. Yanaihara, M. Sakagami, K. Tsuji, T. Hashimoto, T. Kaneko, H. Oka, A. V. Schally, A. Arimura, and T. W. Redding, *J. Medicin. Chem.*, 1973, **16**, 373.
[136] N. Yanaihara, T. Hashimoto, C. Yanaihara, K. Tsuji, Y. Kenmochi, F. Ashizawa, T. Kaneko, H. Oka, S. Saito, A. Arimura, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 64.
[137] D. H. Coy, E. J. Coy, and A. V. Schally, *J. Medicin. Chem.*, 1973, **16**, 83.
[138] J. Rivier, W. Vale, R. Burgus, N. Ling, M. Amoss, R. Blackwell, and R. Guillemin, *J. Medicin. Chem.*, 1973, **16**, 545.

Peptide	*Ref.*
[Tyr(Me)^5]-LH-RF	139
[Phe(p-NO_2)^5]-LH-RF	139
[Phe(p-NH_2)^5]-LH-RF	139
LH-RF-(1—9)-nonapeptide ethylamide	140
[Leu^3]-LH-RF	136
[Phe^3]-LH-RF	136
[Trp^2,His^3]-LH-RF	136
Des-Trp^3-LH-RF	136
Des-His^2-[Phe^5]-LH-RF	136
[Ala^4]-LH-RF	136, 141, 142
[Ala^4,Phe^5]-LH-RF	136
[Pro^1]-LH-RF	142, 143
1-Orotic acid-LH-RF	143
[Glu^1]-LH-RF	143
[Ser^2]-LH-RF	143
[Leu^2]-LH-RF	143
[Gln^2]-LH-RF	143
[Phe^2]-LH-RF	143
1-(2-oxo-oxazolidine-4-carboxylic acid)- and 1-(2-oxo-5-methyl-oxazolidine-4-carboxylic acid)-LH-RF	142
[Phe^2]-, [His(τMe)^2]-, [Lys^2]-, and [Arg^2]-LH-RF	142
[Thr^4]-LH-RF	142
[Gln^5]- and [Tyr(3-Cl)^5]-LH-RF	142
[Tyr(3,5-Cl_2)^7]-, [Gly^7]-, [Ala^7]-, [Val^7]-, [Ile^7]-, and [Nle^7]-LH-RF	142
[Lys^8]- and [Orn^8]-LH-RF	142
[Ala^{10}]-LH-RF	142
[Gln^8]-, [Leu^8]-, and [Pro^8,Arg^9]-LH-RF	144
[Gly^1]-, HCO-[Gly^1]-, Ac-[Gly^1]-, propionyl-[Gly^1]-, palmitoyl-[Gly^1]-, and Ac-[Gly^1]-LH-RF	145
[Leu^1]-, [Tyr^3]-, [Thr^4]-, [Ile^6]-, [Ala^9]-, and [Gly-NMe_2^{10}]-LH-RF	146
[Ala^6]-, [D-Ala^6]-,* [Val^6]-, [D-Val^6]-, [Pro^6]-, [D-Pro^6]-, and Des-His^2-[D-Ala^6]-LH-RF†	147
Miscellaneous related peptides and analogues	148

139 D. H. Coy, E. J. Coy, and A. V. Schally, *J. Medicin. Chem.*, 1973, **16**, 827.

140 M. Fujino, S. Shinagawa, I. Yamazaki, S. Kobayashi, M. Obayashi, T. Fukuda, R. Nakayama, W. F. White, and R. H. Rippel, *Arch. Biochem. Biophys.*, 1973, **154**, 488.

141 R. Geiger, H. Wissmann, W. König, J. Sandow, A. V. Schally, T. W. Redding, L. Debeljuk, and A. Arimura, *Biochem. Biophys. Res. Comm.*, 1972, **49**, 1467.

142 M. Fujino, S. Kobayashi, M. Obayashi, T. Fukuda, S. Shinagawa, I. Yamazaki, R. Nakayama, W. F. White, and R. H. Rippel, *Biochem. Biophys. Res. Comm.*, 1972, **49**, 698.

143 N. Yanaihara, K. Tsuji, C. Yanaihara, T. Hashimoto, T. Kaneko, H. Oka, A. Arimura, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 165.

144 N. Yanaihara, C. Yanaihara, T. Hashimoto, Y. Kenmochi, T. Kaneko, H. Oka, S. Saito, A. V. Schally, and A. Arimura, *Biochem. Biophys. Res. Comm.*, 1972, **49**, 1280.

145 Y. Okada, K. Kitamura, Y. Baba, A. Arimura, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 1180.

146 D. H. Coy, E. J. Coy, and A. V. Schally, *J. Medicin. Chem.*, 1973, **16**, 1140.

147 M. W. Monahan, M. S. Amoss, H. A. Anderson, and W. Vale, *Biochemistry*, 1973, **12**, 4616.

148 M. W. Monahan, J. Rivier, W. Vale, N. Ling, G. Grant, M. Amoss, R. Guillemin, R. Burgus, E. Nicolaides, and M. Rebstock, ref. 31, p. 601.

* 350—450% of the biological potency of LH-RF itself.

† Potent antagonist.

Peptide	*Ref.*
Lysozyme	29, 30
Some protected fragments	149, 150
A decapeptide (Glu-Phe-Ala-Ala-Glu-Glu-Ala-Ala-Ser-Phe) with slight glycosidase activity	151
Melanocyte-stimulating hormones	
Human and monkey β-MSH	152
[Leu^4]- and [Ile^4]-α-MSH	153
Oxytocin	154, 155
[Ile^2,Leu^4]-Oxytocin	156
[Phe^2,Leu^4]-Oxytocin	156
A series of partially deuteriated oxytocin analogues and partial sequences prepared for n.m.r. studies	157
Deamino-oxytocin	158
Aza-oxytocin	159
[Phe(F_5)^2]-, [D-Phe^2]-, and [D-Phe(F_5)^2]-oxytocins	160
[Nva^1]-Oxytocin	155
Z-Pro-Ile-Gly-NH_2 (mesotocin *C*-terminal protected tripeptide)	161
Parathyroid hormone	
Human parathyroid hormone-(1—34)-tetratriacontapeptide	25
Various partial sequences	162
Physalaemin	
The *C*-terminal hexapeptide amide and other peptides	163, 164
Proinsulin	
Human proinsulin connecting peptide	18, 19
Human proinsulin connecting peptide protected peptides	165—169
Porcine proinsulin-(28—34)-heptapeptide	170
Porcine proinsulin-(33—40)-, -(47—58)-, -(47—59)-, -(47—60)-, -(46—61)-, and -(41—61)-peptides	171

[149] M. Guarneri, R. Ferroni, P. Giori, and C. A. Benassi, ref. 31, p. 213.
[150] K. Noda and E. Gross, ref. 31, p. 241.
[151] P. K. Chakravarty, K. B. Mathur, and M. M. Dhar, *Experientia*, 1973, **29**, 786.
[152] K.-T. Wang, J. Blake, and C. H. Li, *Internat. J. Peptide Protein Res.*, 1973, **5**, 33.
[153] O. Nishimura and M. Fujino, *Takeda Kenkyusho Ho*, 1972, **31**, 293 (*Chem. Abs.*, 1972, **77**, 165 060x).
[154] S. Hase and R. Walter, *Internat. J. Peptide Protein Res.*, 1973, **5**, 283.
[155] O. A. Kaurov, V. F. Martynov, and O. A. Popernatskii, *Zhur. obshchei Khim.*, 1973, **43**, 909 (*Chem. Abs.*, 1973, **79**, 42 824h).
[156] V. J. Hruby, F. Muscio, C. M. Groginsky, P. M. Gitu, D. Saba, and W. Y. Chan, *J. Medicin. Chem.*, 1973, **16**, 624.
[157] A. I. R. Brewster, V. J. Hruby, A. F. Spatola, and F. A. Bovey, *Biochemistry*, 1973, **12**, 1643.
[158] G. Jager and R. Geiger, ref. 6, p. 90.
[159] H. Niedrich, *J. prakt. Chem.*, 1972, **314**, 769.
[160] O. A. Kaurov, V. F. Martynov, and M. P. Smirnova, *Zhur. obshchei Khim.*, 1973, **43**, 217 (*Chem. Abs.*, 1973, **78**, 111 738b).
[161] W. König and R. Geiger, *Annalen*, 1973, **106**, 3626.
[162] J. T. Potts, G. W. Tregear, J. van Rietschoten, H. D. Niall, and H. T. Keutmann, ref. 6, p. 191.
[163] W. Voelter, K. Zech, G. Jung, and K. F. Sewing, *Tetrahedron*, 1972, **28**, 5963.
[164] B. Mehlis, H. Apelt, J. Bergmann, and H. Niedrich, *J. prakt. Chem.*, 1972, **314**, 390.
[165] R. Geiger, G. Treuth, and F. Burow, *Chem. Ber.*, 1973, **106**, 2339.
[166] R. Geiger, G. Jager, W. König, and G. Treuth, *Chem. Ber.*, 1973, **106**, 188.
[167] W. König, *Chem. Ber.*, 1973, **106**, 193.
[168] R. Geiger and A. Volk, *Chem. Ber.*, 1973, **106**, 199.
[169] G. Jager, *Chem. Ber.*, 1973, **106**, 206.
[170] K. Neubert and H.-D. Jakubke, ref. 6, p. 235.
[171] V. K. Naithani, *Z. physiol. Chem.*, 1973, **354**, 67.

Peptide	*Ref.*
Porcine proinsulin-(40—52)-, -(41—52)-, -(43—61)-, -(46—52)-, and -(53—58)-peptides	172
[Glu^{9},Gln^{11}]-Human proinsulin connecting peptide	18
Ribonuclease A	
Bovine pancreatic ribonuclease A *C*-terminal nona-, dodeca-, and tetradeca-peptides	173
Ribonuclease T_1	32
Ribonuclease T_1-(89—94)- and -(95—100)-protected hexapeptides	174
Ribonuclease T_1-(81—104)-protected peptide amide	175
Rubredoxin	
A chelated core peptide related to rubredoxin	176
'Scotophobin', and related peptides	177, 178
Secretin	
Secretin *C*-terminal protected hexadecapeptide	179
Leu-Gln-Lys-Leu-Gln-Gly-Leu-Val (*C*-terminal sequence of a secretin analogue)	180
Secretin *C*-terminal-protected tetrapeptide	161
Somatostatin	34—38
[Ala^{3,14}]-Somatostatin	39
Staphylococcal nuclease	
[Nle^{26,32}]-Staphylococcal nuclease-(21—47)-protected heptacosapeptide	40
A series of analogues and shortened peptides related to the (6—47)-peptide	33
Streptococcal protease	
A 27-residue sequence	181
Substance P	42
Trypsin inhibitor	
Peptides related to the active site of bovine basic trypsin inhibitor	182
Thyroid-stimulating hormone releasing factor	129, 133, 183—186
The DDD, DDL, LDL, and LLD stereoisomers	187

[172] V. K. Naithani, *Z. physiol. Chem.*, 1973, **354**, 141.
[173] R. Hayashi, S. Moore, and R. B. Merrifield, *J. Biol. Chem.*, 1973, 248.
[174] S. Terada, N. Mitsuyasu, M. Waki, T. Kato, and N. Izumiya, *Mem. Fac. Sci. Kyushu Univ., Ser. C.*, 1973, **8**, 179 (*Chem. Abs.*, 1973, **79**, 19 092j).
[175] K. Kawasaki, R. Camble, G. Dupuis, H. Romovacek, H. T. Storey, C. Yanaihara, and K. Hofmann, *J. Amer. Chem. Soc.*, 1973, **95**, 6815.
[176] A. Ali, F. Fahrenholz, J. C. Garing, and B. Weinstein, *Internat. J. Peptide Protein Res.*, 1973, **5**, 91.
[177] A. Ali, J. H. R. Faesel, D. Sarantakis, D. Stevenson, and B. Weinstein, *Internat. J. Peptide Protein Res.*, 1972, **4**, 395.
[178] W. Parr and G. Holzer, ref. 31, p. 397.
[179] A. van Zon and H. C. Beyerman, *Helv. Chim. Acta*, 1973, **56**, 1729.
[180] M. Bodanszky, K. W. Funk, and M. K. Fink, *J. Org. Chem.*, 1973, **38**, 3565.
[181] E. Shaich and F. Schneider, ref. 6, p. 180.
[182] I. Z. Siemion, D. Konopińksa, S. Wiejak, B. Rzeszotarska, and Z. Najbar, ref. 6, p. 210.
[183] S. Bajusz and I. Fauszt, *Acta Chim. Acad. Sci. Hung.*, 1973, **75**, 419 (*Chem. Abs.*, 1973, **78**, 148 224g).
[184] M. Fujino, T. Fukuda, and C. Kitada, *Takeda Kenkyusho Ho*, 1973, **32**, 12 (*Chem. Abs.*, 1973, **79**, 32 291t).
[185] J. D. Glass, R. Walter, and I. L. Schwartz, ref. 6, p. 135.
[186] P. Kurath and A. M. Thomas, *Helv. Chim. Acta*, 1973, **56**, 1656.
[187] G. Flouret, R. Morgan, R. Gendrich, J. Wilber, and M. Seibel, *J. Medicin. Chem.*, 1973, **16**, 1137.

Peptide	*Ref.*
Miscellaneous related peptides and analogues too numerous to list individually	148, 186, 188
Tobacco mosaic virus protein	
TMV protein-(150—158)-protected nonapeptide	189
Tuftsin	43
Thr-Lys-Pro-Pro-Arg	43
Thr-Lys-Pro	43
Lys-Pro-Arg	43
Vasoactive intestinal peptide (VIP)	44
The *C*-terminal hendeca-, tetradeca-, pentadeca-, and docosa-peptides	44
The (7—13) protected heptapeptide	47
The (14—28) protected pentadecapeptide	48, 230
Vasopressin	
1-Deamino, 9-ethylenediamine-lysine vasopressin	190
Arginine vasopressin	1, 191
Arginine vasopressin tritiated in the Phe residue	192
[Thr^4]-arginine vasopressin	191
[Thr^4]-lysine vasopressin	191
[Ile^3,Thr^4]-arginine vasopressin (*i.e.* [Thr^4]-arginine vasotocin)	191
[Ile^3]-arginine vasopressin (*i.e.* arginine vasotocin)	191
1-Deamino-[Thr^4]-arginine vasopressin	193
1-Deamino-[Ile^3,Thr^4]-arginine vasopressin (*i.e.* 1-deamino-[Thr^4]-arginine vasotocin)	193
1-Deamino-[Thr^4]-lysine vasopressin	193
1-Deamino-[Ile^3]-arginine vasopressin (*i.e.* 1-deamino-arginine vasotocin)	193
[Leu^4]-lysine vasopressin	194
1-Deamino-[Leu^4]-lysine vasopressin	194
[Ile^3,Leu^4]-arginine vasopressin (*i.e.* [Leu^4]-arginine vasotocin)	195
1-Deamino-lysine vasopressin	196
1-Deamino-[$Lys(Tos)^8$]-vasopressin	196
1-Deamino-[Val^4,D-Arg^8]-vasopressin	197
[β-Ala^1,Ala^6]-arginine vasopressin	198
Sequential polypeptides	
H-(-Pro-Hyp-Gly)$_n$-OH (n = 5 and 10)	199
H-(-Pro-Pro-Gly-)$_{10}$-OH (crystallized)	200

[188] Yu. P. Shvachkin and A. P. Smirnova, *Zhur. obshchei Khim.*, 1973, **43**, 686 (*Chem. Abs.*, 1973, **79**, 19 086k).
[189] G. R. Pettit, S. K. Gupta, and R. H. Ode, *J.C.S. Perkin I*, 1973, 950.
[190] J. D. Glass and V. du Vigneaud, *J. Medicin. Chem.*, 1973, **16**, 160.
[191] M. Manning, E. J. Coy, W. H. Sawyer, and M. Acosta, *J. Medicin. Chem.*, 1973, **16**, 463.
[192] H. H. Holton, L. A. Branda, and B. M. Ferrier, *Canad. J. Chem.*, 1973, **51**, 1910.
[193] M. Manning, E. J. Coy, M. Acosta, and W. H. Sawyer, *J. Medicin. Chem.*, 1973, **16**, 836.
[194] D. F. Dyckes, M. F. Ferger, V. du Vigneaud, and W. Y. Chan, *J. Medicin. Chem.*, 1973, **16**, 843.
[195] D. Gillessen, R. O. Studer, and J. Rudinger, *Experientia*, 1973, **29**, 170.
[196] R. T. Havran, *Experientia*, 1973, **29**, 520.
[197] M. Manning, L. Balaspiri, M. Acosta, and W. H. Sawyer, *J. Medicin. Chem.*, 1973, **16**, 975.
[198] G. Jager and R. Geiger, *Annalen*, 1973, 1928.
[199] S. Sakakibara, K. Inouye, K. Shudo, Y. Kishida, Y. Kobayashi, and D. J. Prockop, *Biochem. Biophys. Acta*, 1973, **303**, 198.
[200] S. Sakakibara, Y. Kishida, and S. Aimoto, ref. 31, p. 293.

Peptide	*Ref.*
H-Pro(4-Br)-Pro-Gly-(Pro-Pro-Gly)$_9$-OH (crystallized)	200
Poly-(Tyr-Glu-Tyr-Glu)	201
Poly-(Glu-Tyr-Glu)	201
Poly-(Glu-Glu-Tyr-Glu)	201
Poly-(Ala-Ala-Lys)	202
Poly(Lys-Lys-Ala)	202
Poly-(Lys-Lys-Gly)	202
Poly-(Ala-Ala-Ala-Lys)	202
Poly-(Lys-Lys-Lys-Ala)	202
Poly-(Lys-Lys-Ala-Ala)	202
Poly-(Lys-Lys-Gly-Gly)	202
Poly-(Orn-Orn-Ala-Ala)	202
Poly-(Orn-Orn-Gly-Gly)	202
Poly-(Ala-Arg)	203
Poly-(Ala-Ala-Arg)	203
Poly-(Ala-Arg-Arg)	203
Poly-(Gly-Arg-Arg)	203
Poly-(Lys-Arg-Ala)	204
Poly-(Lys-Arg-Gly)	204
Poly-(Ala-D-Ala-Ala-Ala)	205
Poly-[Tyr-Glu(Gly-OBu^t)-Ala-Gly]-Gly-OMe	206
Poly-(Pro-Ala-Gly)	207
Poly-(Pro-Gly-Pro)	207
Poly-(Pro-Ser-Gly)	207
A series of sequential polypeptides (with protected side chains) containing Ala, Lys, and Orn in various proportions	208
Poly-(Glu-Lys-Ala-Gly)	209
Poly-(Ala-D-Glu-Lys-D-Ala-Gly)	209
Miscellaneous sequential polypeptides related to collagen	210

Miscellaneous Peptides

Some synthetic peptide precursors of norphalloine containing, *inter alia*, Nva, Hyp, or *a*Hyp	211
A series of nineteen oligomers (Pro)$_n$, n = 2—20	212
Asp-Arg-Val-Tyr-Val-His-Pro-Phe-His-Leu-Leu (pepsitensin) and the [Ile^5]-analogue	213

[201] Y. Trudelle, *J.C.S. Perkin I*, 1973, 1001.
[202] V. A. Shibnev, L. I. Mar'yash, and K. T. Poroshin, *Doklady Akad. Nauk S.S.S.R.*, 1972, **207**, 625 (*Chem. Abs.*, 1973, **78**, 84 805n).
[203] V. K. Burichenko, K. T. Poroshin, G. F. Kasymova, and V. A. Shibnev, *Izvest. Akad. Nauk S.S.S.R,. Ser. khim.*, 1972, 2596 (*Chem. Abs.*, 1973, **78**, 124 934n).
[204] V. K. Burichenko, K. T. Poroshin, G. F. Kasymova, A. K. Mirzoev, and V. A. Shibnev, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1972, 2597 (*Chem. Abs.*, 1973, **78**, 124 935p).
[205] N. Lotan, F. T. Hesselink, H. Benderly, J. F. Yan, I. Schecter, A. Berger, and H. A. Scheraga, *Macromolecules*, 1973, **6**, 447.
[206] B. J. Johnson and C. Cheng, *J. Pharm. Sci.*, 1973, **62**, 917.
[207] E. R. Heidemann, B. S. Harrap, and H. D. Schiele, *Biochemistry*, 1973, **12**, 2958.
[208] V. A. Shibnev and L. I. Mar'yash, *Izvest. Akad. Nauk, S.S.S.R., Ser. khim.*, 1973, 435 (*Chem. Abs.*, 1973, **79**, 5 585k).
[209] A. R. Zeiger, A. Lange, and P. H. Maurer, *Biopolymers*, 1973, **12**, 2135.
[210] R. Fairweather and J. H. Jones, *Immunology*, 1973, **25**, 241.
[211] E. Nebelin, H. Faulstich, and T. Wieland, *Annalen*, 1973, 45.
[212] M. Rothe and J. Mazanek, *Tetrahedron Letters*, 1972, 3795.
[213] A. C. M. Paiva, T. B. Paiva, and M. E. Miyamoto, *J. Medicin. Chem.*, 1973, **16**, 280.

Peptide	*Ref.*
Decalysylvaline	214
Leu-Gln-Asn-Cys(Bzl)-Pro-Leu-Gly-NH_2	215
An extensive series of peptides containing alanine and proline prepared as substrates for elastase	216
Some simple ^{15}N-enriched dipeptides	217
Arg-DL-*a*Thr-Phe (the L-D-L isomer has antifungal activity) and Arg-Thr-Phe	218
Some derivatives of t-leucylvaline	219
Some cysteine and cystine peptides	220
Various simple peptides and mixtures obtained by the *N*-carboxy-anhydride method	221, 222
Miscellaneous simple peptides obtained by the solid-phase method	223, 224
Peptides containing α-benzylphenylalanine	225
Some thrombin substrates	226
Possible peptide renin inhibitors	227, 228
Some linear peptides related to gramicidin S	229, 230
Some protected oligolysine derivatives	231, 232
Aza-peptides	91—93, 159
Some peptides prepared as protocollagen proline hydroxylase substrates	233
A series of possible renin substrates related to the region around the renin-sensitive Phe–Met bond of κ-casein	234
A series of microbial proteinase substrates	235
A series of possible renin inhibitors with *C*-terminal carbinol groups	236
Gly-Ala-Arg-Gly-Ala-Trp-Gly	237

[214] B. W. Erickson and R. B. Merrifield, *J. Amer. Chem. Soc.*, 1973, **95**, 3757.
[215] M. Bodanszky and K. W. Funk, *J. Org. Chem.*, 1973, **38**, 1296.
[216] R. C. Thompson and E. R. Blout, *Biochemistry*, 1973, **12**, 51, 57, 66.
[217] J. A. Sogn, W. A. Gibbons, and E. W. Randall, *Biochemistry*, 1973, **12**, 2100.
[218] W. A. König, W. Loeffler, W. H. Meyer, and R. Uhmann, *Chem. Ber.*, 1973, **106**, 816.
[219] E. Frauendorder, W. Steglich, and F. Weygand, *Chem. Ber.*, 1973, **106**, 1019.
[220] H. Arold and M. Eule, ref. 6, p. 78.
[221] P. Pfaender, E. Kuhnle, B. Krahl, A. Backmansson, G. Gnauck, and H. Blecher, *Z. physiol. Chem.*, 1973, **354**, 267.
[222] P. Pfaender, E. Kuhnle, B. Krahl, A. Backmansson, and H. Blecher, *Z. physiol. Chem.*, 1973, **354**, 286.
[223] S. Wang, ref. 31, p. 179.
[224] W. Parr and K. Grohmann, ref. 31, p. 169.
[225] H. L. Maia, B. Ridge, and H. N. Rydon, *J.C.S. Perkin I*, 1973, 98.
[226] L. C. Dorman, R. C. Cheng, and F. N. Marshall, ref. 31, p. 455.
[227] K. Poulsen, J. Burton, and E. Haber, *Biochemistry*, 1973, **12**, 3877.
[228] M. J. Parry, A. B. Russell, and M. Szelke, ref. 31, p. 415.
[229] S. Sofuku, *Bull. Chem. Soc. Japan*, 1973, **46**, 968.
[230] S. Matsura, M. Waki, T. Kato, and N. Izumiya, *Bull. Chem. Soc. Japan*, 1973, **46**, 977.
[231] K. T. Poroshin, V. V. Bondar, and V. A. Shibnev, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1972, 2305 (*Chem. Abs.*, 1973, **78**, 58 779n).
[232] L. I. Mar'yash and V. A. Shibnev, *Izvest. Akad Nauk. S.S.S.R., Ser. khim.*, 1973, 1858 (*Chem. Abs.*, 1972, **77**, 152 546k).
[233] K. Okada, Y. Kikuchi, Y. Kawashiri, and M. Hiramoto, *F.E.B.S. Letters*, 1972, **28** 226.
[234] C. Schattenkerk, I. Voskuyl-Holtkamp, and R. Bokhorst, *Rec. Trav. chim.*, 1973, **92**, 92.
[235] K. Morihara and T. Oka, *Arch. Biochem. Biophys.*, 1973, **157**, 561.
[236] K. Shigezane and T. Mizoguchi, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 972.
[237] D. Yamashiro and C. H. Li, *J. Org. Chem.*, 1973, **38**, 2594.

Peptide	*Ref.*
Asp-Phe-OMe	238, 239
Ala-Trp-Ile	240
Some lanthionine derivatives and peptides	241
Ornithine-containing peptides	242—244
Ala-γ-D-Glu-Lys-D-Ala-D-Ala	245
A series of pentapeptides of structure H-Gly_2-X-Gly_2-OH, where X = Gly, Ala, Val, Ile, or Leu, prepared for n.m.r. studies	246
A set of diastereoisomers of Phe-Leu-Phe	247
Some peptides containing β-aminoisovaleric acid	248
A series of aspartic acid dipeptide esters	249
Some substrate analogue inhibitors of human renin	250

Appendix II: A List of Some Useful New Synthetic Intermediates Described during 1973 *by J. H. Jones*

The preambles to the first two of these lists (Vol. 2 of these Reports, p. 188 and Vol. 3, p. 270) still apply.

Compound	*Ref.*
Alanine	
Boc-Ala-ONp(*o*)	180
Boc-Ala-OPfp	251
(*l*)Bornoc-Ala*	252
Bpoc-Ala,Dcha (significantly higher m.p.)	253
Ddz-Ala,Dcha	254
Z-Ala-NHNHBoc	78
β-Alanine	
(*l*)Bornoc-β-Ala,Dcha	252
2-Allylglycine	
2-Allylglycine methyl ester hydrochloride	8

[238] Y. Ariyoshi, T. Yamatani, and Y. Adachi, *Bull. Chem. Soc. Japan*, 1973, **46**, 2611.
[239] Y. Ariyoshi, T. Yamatani, N. Uchiyama, T. Adachi, and N. Sato, *Bull. Chem. Soc. Japan*, 1973, **46**, 1893.
[240] H. Blecher and P. Pfaender, *Annalen*, 1973, 1263.
[241] A. Shöberl, M. Rimpler, and E. Graf, *Annalen*, 1973, 1379.
[242] N. A. Poddubnaya and N. Ya. Krasnobrizhii, *Zhur. obshchei Khim.*, 1973, **43**, 199 (*Chem. Abs.*, 1973, **78**, 111 731u).
[243] N. A. Poddubnaya and N. Ya. Krasnobrizhii, *Zhur. obshchei Khim.*, 1973, **43**, 917 (*Chem. Abs.*, 1973, **79**, 42 820d).
[244] N. A. Poddubnaya, L. S. Zhigis, and N. Ya. Krasnobrizhii, *Vestsi Akad. Navuk Belarusk. S.S.R., Ser. khim.*, 1973, 101 (*Chem. Abs.*, 1973, **79**, 5 567f).
[245] A. R. Zeiger and P. H. Maurer, *Biochemistry*, 1973, **12**, 3387.
[246] P. Keim, R. A. Vigna, R. C. Marshall, and F. R. N. Gurd, *J. Biol. Chem.*, 1973, **248**, 6104.
[247] B. Tomicka, G. Kupryszewski, and F. Karczynski, *Roczniki Chem.*, 1973, **47**, 185 (*Chem. Abs.*, 1973, **79**, 5578k).
[248] C. N. C. Drey, J. Lowbridge, and R. J. Ridge, *J.C.S. Perkin I*, 1973, 2001.
[249] R. H. Mazur, J. A. Reuter, K. A. Swiatek, and J. M. Schlatter, *J. Medicin. Chem.*, 1973, **16**, 1284.
[250] I. Parikh and P. Cuatrecasas, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 1356.
[251] L. Kisfaludy, M. Löw, O. Nyeki, T. Szirtes, and I. Schön, *Annalen*, 1973, 1421.
[252] G. Jager and R. Geiger, *Annalen*, 1973, 1535.

* (*l*)Bornoc = L-isobornyloxycarbonyl *etc.*: in Vol. 5 the less precise symbol Bornoc was used for derivatives which were in fact of the D-terpenoid-series, *i.e.* (*d*)Bornoc derivatives.

Compound	*Ref.*
Arginine	
Boc-Arg(NO_2)-OQ(5-Cl)	170
Bpoc-Arg(Tos),Cha	253
Z(OMe)-Arg(Tos)	255
Z(OMe)-Arg(Tos),Cha	255
Asparagine	
Boc-Asn-ONp(*o*)	180, 256
Boc-Asn-OPcp	108
Boc-Asn-OPfp	251
Boc-Asn(Bzh)-OBzl	156
Boc-Asn(Bzh)	156
Bpoc-Asn-ONSu	257
Bpoc-Asn[Bzh(OMe_2)],Cha	253
Ddz-Asn-[Bzh(OMe_2)]	114
Z-Asn-NHNHBoc	83*f*
Z-Asn-ONp(*o*)	258
Z-Asn-OPfp	52, 259
Aspartic acid	
Boc-Asp-OBzl	156
Boc-Asp(OBu^t)-OPfp	251
Bpoc-Asp(OBzl),Cha	253
Boc-Asp(OBzl)-OPcp	108
Boc-Asp(OBzl)-ONp(*o*)	180
Z-Asp(OMe)-OPcp	260
Z-Asp(OMe)-OPfp	260
Z-Asp(OMe)-OTcp	260
Azetidine-2-carboxylic acid (Aze)	
Z-Aze	261
Aze-OMe,bis(toluene-*p*-sulphonyl)amine salt	261
Boc-Aze	81
Cysteine	
Cys(Acm) (alternative procedure)	262
Cys[Bzl(3,4-Me_2)]	104
Cys(Bu^t)-OBu^t,AcOH	263
Cys(Bu^t)-OBu^t,HCl	263
Boc-Cys(CH_2CONH_2)	104
Boc-Cys(CH_2CONH_2)-ONp	104
Boc-Cys(Bzl)-ONp(*o*)*	256

[253] R. S. Feinberg and R. B. Merrifield, *Tetrahedron*, 1972, **28**, 5865.
[254] C. Birr, W. Lochinger, H. Stahnke, and P. Lang, *Annalen*, 1972, **763**, 162.
[255] H. Yajima, F. Tamura, Y. Kiso, and M. Kurobe, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 1380.
[256] M. Bodanszky and K. W. Funk, *J. Org. Chem.*, 1973, **38**, 1296.
[257] D. Hudson, G. W. Kenner, B. Mason, B. Morgan, R. Ramage, B. Singh, and R. Tyson, ref. 6, p. 70.
[258] M. Bodanszky, R. J. Bath, A. Chang, M. L. Fink, K. W. Funk, S. M. Greenwald, and Y. S. Klausner, ref. 31, p. 203.
[259] L. Kisfaludy, ref. 6, p. 129.
[260] J. Kovacs, R. E. Cover, R. H. Johnson, T. J. Kalas, G. L. Mayers, and J. E. Roberts, *J. Org. Chem.*, 1973, **38**, 2518.
[261] J. Vičar, J. Smolikova, and K. Blaha, *Coll. Czech. Chem. Comm.*, 1973, **38**, 1957.
[262] P. Hermann and E. Schreier, ref. 6, p. 126.
[263] A. Chimiak, T. Kolasa, and J. F. Biernat, *Z. Chem.*, 1972, **12**, 264.

* Named in the experimental part of ref. 256 as a derivative of *S*-benzoyl-cysteine, but this is clearly a misprint. See also ref. 180.

Compound	*Ref.*
Boc-Cys(Bzl)-OPcp	108
Boc-Cys(Bzl)-OPfp	251
Boc-Cys(Acm)-OPfp	251
Boc-Cys[Bzl(3,4-Me_2)],Dcha	104
Boc-Cys[Bzl(OMe)] (purification *via* Dcha salt)	264
Boc-Cys(SBu^t)	264
Boc-Cys(Trt)-OTcp	1
Boc-Cys(Trt)-ONp (higher m.p. than before)	1
(*l*)Bornoc-Cys(Bzl)	252
Bpoc-Cys[Bzl(OMe)],Cha	253
Ddz-Cys(SBu^t),Dcha	254
Ddz-Cys[Bzl(OMe)],Dcha	114
Cystine	
Ddz_2-Cystine	254
$\alpha\beta$-*threo*-Diaminobutyric acid	
Several derivatives suitable for peptide synthesis	265
2,5-Dihydrophenylalanine	
A range of derivatives (Z-, Boc-, *etc.*) for use in peptide synthesis	266
Glutamic acid	
Boc-Glu(OBzl)-OQ(5-Cl)	170
Boc-Glu(OBzl)-OPcp	108
Boc-Glu(OBzl)-ONSu	267
Boc-Glu(OBu^t)-OPfp	251
Boc-Glu(OBzl)-ONp(*o*)	180
Boc-Glu[Bzl(*p*-Br)]	104
(*l*)Bornoc-Glu	252
(*l*)Bornoc-Glu(OBu^t)	252
Ddz-Glu(OBzl) (not crystalline)	114
Z-Glu(OBu^t)-NHNHBoc (oil)	83*e*
Z-Glu(OMe)-OPcp	260
Z-Glu(OMe)-OPfp	260
Glutamine	
Gln-OBu^t,HCl	166
Boc-Gln-OPcp	108
Boc-Gln-OPfp	251
Boc-Gln-ONp(*o*)	180, 256
(*l*)Bornoc-Gln	252
(*l*)Bornoc-Gln-OTcp	252
(*l*)Bornoc-Gln-OPcp	252
(*l*)Bornoc-Gln-ONp	252
Bpoc-Gln[Bzh(OMe_2)],Cha	253
Ddz-Gln[Bzh(OMe_2)]	114
Z-Gln[Bzh(OMe_2)]-ONSu	169
Z-Gln[Bzh(OMe_2)]-NHNHBoc	83*d*
Glycine	
Gly-NHNHZ,CF_3CO_2H	177
Gly-OBu^t,AcOH	263
Boc-Gly-NHNHZ (amorphous)	177
Boc-Gly-ONp(*o*)	180, 258

[264] T. Wieland, F. Flor, and C. Birr, *Annalen*, 1973, 1595.
[265] E. Atherton and J. Meienhofer, *Z. physiol. Chem.*, 1973, **354**, 689.
[266] G. R. Nagarajan, L. Diamond, and C. Ressler, *J. Org. Chem.*, 1973, **38**, 621.
[267] K. Nakajima and K. Okawa, *Bull. Chem. Soc. Japan*, 1973, **46**, 1811.

Compound	*Ref.*
Boc-Gly-OPfp	251
(*l*)Bornoc-Gly	252
Ddz-Gly,Dcha	254
Z-Gly-NHNHBoc	78
Histidine	
His-NHNHBoc	83*b*
Boc-His(Dnp)-ONSu	267
Boc-His(Dnp)-OPcp	108
Boc-His(Boc)-OBzl(NO_2)	268
Boc-His(Boc) (alternative procedures)	268
Boc-His(Boc)-ONp	268
Boc-His(Tos)-OTcp	17
Boc-His(Z) (amorphous)	268
Boc-His(Z)-ONp	268
Bpoc-His,Cha	253
Bpoc-His(Tos),Cha	253
Z-His-NHNHBoc	83*b*
Z(OMe)-His(Bzl)	255
Z(OMe)-His(Tos),Dcha	255
Z(OMe)-His-$NHNH_2$	255
Z(OMe)-His	268
Z(OMe)-His[Z(OMe)]-ONp	268
Z(OMe)-His[Z(OMe)] (oil)	268
Hydroxyproline	
Boc-Hyp(Bzl)	199
Hyp(Bzl)	199
allo-Hydroxyproline	
*a*Hyp-OMe,HCl	211
Boc-*a*Hyp(Bu^t)	269
Isoleucine	
Boc-Ile-OPfp	251
Boc-Ile-ONp(*o*) (oil)	180
(*l*)Bornoc-Ile,Dcha	257
Bpoc-Ile,Cha	253
Ddz-Ile,Dcha	254
Lanthionine (Lan)	
Z_2Lan	241
Z_2-Lan$(ONp)_2$	241
Leucine	
Leu-OBu^t,AcOH	263
Boc-Leu-ONp(*o*)	180, 256
Boc-Leu-OPfp	251
Bpoc-Leu,Cha	253
Ddz-Leu,Dcha	114
Nps-Leu-OSu (oil)	83*b*
Lysine	
Lys[Z(2-Br)]	53
Lys[Z(2,4-Cl_2)]	214
Lys(Boc)-NHNHZ,AcOH	270

[268] E. Schaich, A.-M. Fretzdorff, and F. Schneider, *Z. physiol. Chem.*, 1973, **354**, 897.
[269] J. C. Sheehan and S. L. Ledis, *J. Amer. Chem. Soc.*, 1973, **95**, 875.
[270] D. Blanot and E. Bricas, ref. 6, p. 420.

Compound	*Ref.*
Lys(Boc)-NHNHZ,HBr	270
Boc-Lys[Z(2-Br)],Dcha	53
Boc-Lys[Z(2,4-Cl_2)] (oil)	214
Boc-Lys(Z)-OQ(5-Cl)	170
Boc-Lys[Z(NO_2)]-OQ(5-Cl)	170
Boc-Lys(Z)-OPfp	251
Boc-Lys[Z(2,6-Cl_2)]-ONp(*o*)	180
Bpoc-Lys(Boc)-NHNHZ	270
Bpoc-Lys(Z),Dcha	253
Nps-Lys(CF_3CO),Dcha	83*g*
Nps-Lys(Boc)-NHNHZ	270
Nps-Lys(Boc),Dcha (rotation discrepancy)	271
Z-Lys(CF_3CO)-NHNHBoc	83*a*
Z-Lys(CF_3CO)-ONp	83*a*
Z-Lys(Boc)-OPfp	251
Z-Lys[(*l*)Bornoc]	252
Z-Lys[(*d*)Bornoc]	252
Z(OMe)-Lys(Z) (formerly described as an oil)	255
Z(OMe)-Lys(CHO)	255
Z(OMe)-Lys(Tos),Dcha	255
Methionine	
Boc-Met-OPcp	108
Boc-Met-OPfp	251
Boc-Met-ONp(*o*)	180
Nps-Met-ONp	83*j*
N-Methylamino-acids	
MeIle-OMe,HCl (more details than previously published)	90
MeLeu-OMe,HCl (more details than previously published)	90
MePhe-OBzl,TosOH	90
Sar-OBzl,HCl	90
MeMet-OEt,HCl	88
O-Methylthreonine	
Boc-Thr(Me),Dcha	58
Phenylalanine	
Boc-Phe-ONp(*o*)	180, 256
Boc-Phe-OPfp	251
Ddz-Phe,Dcha	254
Z-Phe-ONp(*o*)	256
Phenylalanine, substituted in the ring	
Boc-Phe(*p*-F)	61
Boc-Phe(*p*-I)-OTcp	272
Proline	
D-Pro-NH_2,HCl	187
Boc-Pro-ONp(*o*)	180, 256
Boc-Pro-OPfp	251
Boc-Pro-OPcp	108
Bpoc-Pro,Cha	253
Ddz-Pro	254
Proline analogues	
Boc-derivatives of various proline analogues	81

[271] I. Barral and J. Savrda, *Synthesis*, 1973, 795.
[272] G Krail, D. Brandenburg, and H. Zahn, *Z. physiol. Chem.*, 1973, **354**, 1497.

Compound	*Ref.*
Pyroglutamic acid	
Z-Glp-OSu	135, 186
D-Glp-OPcp	187
Serine	
Ser(Tms) *N*-carboxyanhydride (oil)	273
Boc-Ser-OPcp	108
Boc-Ser(Bzl)-OPcp	108
Boc-Ser-OPfp	251
Boc-Ser(Bzl)-ONp(*o*)	180
(*l*)Bornoc-Ser,Dcha	252
Ddz-Ser(Bu^t),Dcha	254
Thiazolidine-4-carboxylic acid (Thz)	
Boc-Thz	81
Threonine	
Thr(Tms) *N*-carboxyanhydride (oil)	273
Bpoc-Thr(Bzl),Cha	159
Boc-Thr-OPcp	108
Boc-Thr(Bzl)-OPcp	108
Boc-Thr-OPfp	251
Bpoc-Thr(Bzl),Cha	253
Nps-Thr-ONSu	83*b*
Z-Thr(Bu^t)-NHNHBoc	83*h*
Tryptophan	
Boc-Trp-OPfp	251
Boc-Trp-OPcp	108
Boc-Trp-ONp(*o*)	180
Boc-Trp(CHO),Dcha	237
Bpoc-Trp,Dcha	253
Ddz-Trp	254
Z(OMe)-Trp	255
Tyrosine	
Tyr[Bzl(2,6-Cl_2)]	53, 274
Tyr[Z(2-Br)]	100
Tyr(Bu^t)-NH_2	177
Boc-Tyr[Z(2-Br)],Dcha	100
Boc-Tyr[Bzl(2,6-Cl_2)]	53
Boc-Tyr(Bzl)-ONp(*o*)	190
(*l*)Bornoc-Tyr(Bzl)	252
(*l*)Bornoc-Tyr(Bzl)-ONp	252
(*l*)Bornoc-Tyr(Bzl)-OTcp	252
(*l*)Bornoc-Tyr(Bzl)-OPcp	253
Bpoc-Tyr(Bzl),Cha	253
Bpoc-Tyr[Bzl(2,6-Cl_2)],Cha	253
Ddz-Tyr(Bzl),Dcha	254
Z-Tyr(Bu^t)-NH_2	177
Z(OMe)-Tyr,Dcha	255
Valine	
Val-OBu^t,AcOH	263
Boc-Val-OPfp	251
Boc-Val-ONp(*o*)	180
Bpoc-Val,Cha (higher m.p. than before)	253

[273] R. Wies and P. Pfaender, *Annalen*, 1973, 1269.
[274] B. W. Erickson and R. B. Merrifield, *J. Amer. Chem. Soc.*, 1973, **95**, 3750.

4

Peptides with Structural Features not Typical of Proteins

BY B. W. BYCROFT

1 Introduction

The past twelve months have witnessed significant developments in a number of areas. New microbial peptides with novel features continue to be isolated and the increasing complexity of these molecules reflects the growing sophistication of the methods available for the structural elucidation of peptides containing unusual structural features.

The molecular sizes of, for example, enduracidin, bleomycin, and subtilin move closer to the borderline between classical organic molecules and bio-macromolecules. The nisin and subtilin group are particularly noteworthy in this respect. In addition there is evidence to suggest that the initial biosynthesis of these antibiotics, unlike the majority of small microbial peptides, follows that of the proteins. The possibility that similar highly modified peptides produced under ribosomal control will be isolated and characterized in the next few years is very real.

The application of physicochemical methods for defining the finer points of the spatial structure of peptide systems has concentrated on the biologically active cyclic peptides. The results, though often elegant, have been achieved mainly by the application of the principles and methods reviewed in last year's Report rather than by the introduction of new techniques. Several exploratory studies with a variety of new methods which may be of value in the long term have also been reported.

The major synthetic achievements this year are undoubtedly the elegant syntheses of the epidithiodioxopiperazines, anhydrogliotoxin and sporidesmin G, and the cephalosporin derivative, cephalothin. However, the novel aspects of the syntheses are only of indirect interest to peptide chemists. In fact, synthetic endeavour has been restricted mainly to analogues of cyclic peptides with biological activity. The only significant new development is the use of dehydroalanine-containing peptides in a simple solid-phase procedure for the synthesis of peptide amides.

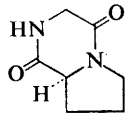
The penicillin–cephalosporin field has been dominated by synthetic approaches to the 6(7)-methoxy-derivatives, together with a considerable resurgence of interest in the biosynthesis. The structural and biosynthetic aspects of peptidoglycan from bacterial cell wall continue to be an area of considerable activity.

2 Cyclic Peptides (Homodetic Peptides)

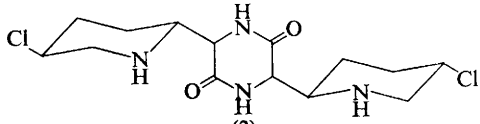
Cyclic peptides continue to be the subject of numerous investigations and in order to reflect adequately the emphasis of this work, headings of a more general nature for this section have been used in this year's Report.

An excellent review on cyclic peptides covering recent advances in synthesis and conformational studies has been published.[1]

2,5-Dioxopiperazines (Cyclic Dipeptides).—Various new naturally occurring cyclic dipeptides have been isolated from a diverse range of organisms. As part of a general survey of marine organisms, the ethanolic extracts of the star-fish *Luidia clathrata* have been shown to exhibit significant activity against lymphocytic leukemia. The simple dioxopiperazine (1) has been obtained from this extract and the structure established[2] by an



(1)



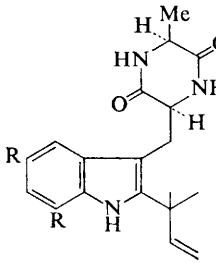
(2)

X-ray crystallographic analysis. Although (1) does not possess antitumour activity, the chlorine-containing metabolite (2) from *Streptomyces griseolutus* is active. The structural assignment and relative stereochemistry (2) are based on spectroscopic evidence.[3]

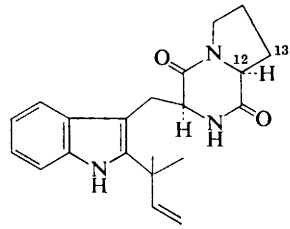
In last year's Report the isolation of a partially purified enzyme which catalyses the isoprenylation of *cyclo*-L-alanyl-L-tryptophan to the mono-isoprenylated compound (3) was described. An extension of this work has demonstrated[4] that (3) is a precursor of echinulin (4) in *Aspergillus amstelodami*. A series of similar metabolites (5)—(9), related to the brevianamide group, are considered to be the toxic principles of the fungus *Aspergillus ustus*.[5] The proline residue possesses the L-configuration in all except (7). The tryptophan moiety has the L-configuration in (5) and (6), but neither the chirality at C-9 in (7)—(9) nor at C-2 in (8) and (9) has been established.

Simple 2,5-dihydroxypyrazine derivatives have been shown to undergo 1,4-cycloadditions, and the implications of these reactions to the final stages of the biosynthesis of brevianamide A and related compounds were discussed.[6] It is probable that the relatively large number of naturally occurring pyrazine derivatives with amino-acid side-chain substituents in

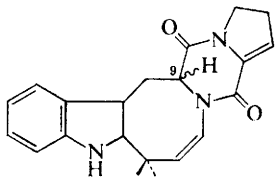
[1] P. M. Hardy and B. Ridge, *Progr. Org. Chem.*, 1973, **8**, 129.
[2] G. R. Pettit, R. B. von Dreele, G. Bollinger, P. M. Traxler, and P. Brown, *Experientia*, 1973, **29**, 521.
[3] B. H. Arison and J. L. Beck, *Tetrahedron*, 1973, **29**, 2743.
[4] C. M. Allen, *J. Amer. Chem. Soc.*, 1973, **95**, 2386.
[5] P. S. Steyn, *Tetrahedron*, 1973, **29**, 107.
[6] P. J. Machin, A. E. A. Porter, and P. G. Sammes, *J.C.S. Perkin I*, 1973, 404.



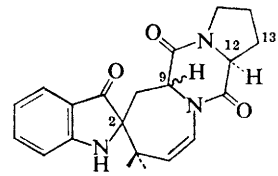
(3) R = H
(4) R = $CH_2CH{=}CMe_2$



(5)
(6) 12,13-Dehydro

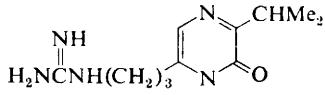


(7)



(8)
(9) 12,13-Dehydro

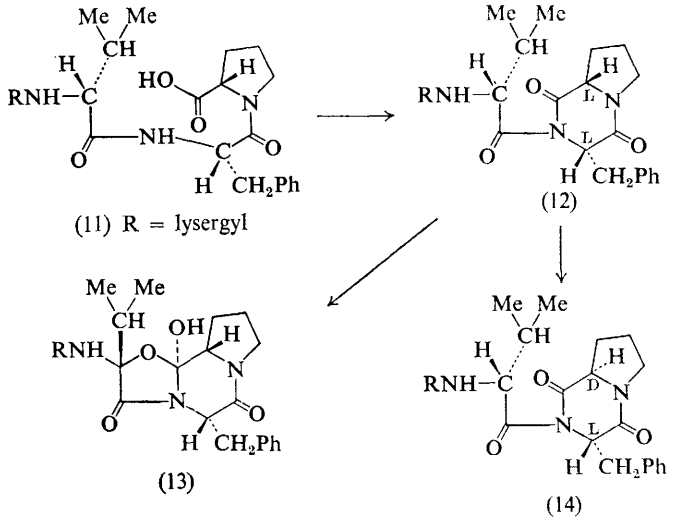
positions 3 and 6 are derived by oxidative processes on the corresponding cyclic dipeptides. Argvalin (10), isolated from a *Streptomyces*, is probably a further example of this class of compounds.[7]



(10)

A possible biosynthetic sequence for the ergot peptides has been proposed and is outlined in Scheme 1. It is suggested that the lysergyl tripeptide (11) cyclizes to (12) which is then oxidized to ergotoxin (13). Circumstantial evidence for this sequence is provided by the isolation and characterization[8] of the acyl dioxopiperazine (14) from the mycelium of a *Claviceps purpurea* strain. This is claimed to be a shunt product formed by the epimerization of the key intermediate (12). The possibility that (12) is involved in the biosynthesis was predicted[9] some time ago, together with the proposal that the oxidative cyclization to (13) is achieved *via* an acylimine intermediate. It is of interest to note that an acylimine intermediate has been implicated in the biogenesis of the 7-methoxycephalosporins (see Section 5).

[7] K. Tatsuta, K. Fujimoto, M. Yamashita, T. Tsuchiya, and S. Umezawa, *J. Antibiotics*, 1973, **26**, 606.
[8] P. Stutz, R. Brunner, and P. A. Stadler, *Experientia*, 1973, **29**, 936.
[9] B. W. Bycroft, *Nature*, 1969, **224**, 595.



Scheme 1

The epithiodioxopiperazines belonging to the sporidesmin and gliotoxin group of metabolites also represent cyclic dipeptides at a higher oxidation level. The considerable interest in this important class of compounds, no doubt because of their antibacterial and cytotoxic activity, is still apparent. The results of an *X*-ray analysis on sporidesmin G, a minor metabolite of *Pithomyces chartarum*, have revealed[10] that the stereochemistry of the tetrathio-bridge is very similar to that observed[11] for *NN'*-dimethyl-3,6-epitetrapiperazine-2,5-dione (Figure 1).

Full details of the chemical and physicochemical evidence for the structures of the dimeric antibiotics verticillin A (15) and B (16) have been presented.[12] Verticillin C is considered to be the epitrithio-analogue of verticillin B. A metabolite, designated hyalodendrin, has been isolated from a species of *Hyalodendron* and the structure (17) assigned on the basis of spectroscopic and chemical data.[13] This metabolite was isolated[14] from an unidentified fungus, together with (18) and (19). These compounds are of interest since it was proposed previously that they might be late intermediates in the biosynthesis of the gliotoxin and aranotin antibiotics.

[10] M. Przybylska, E. M. Gopalakrishna, A. Taylor, and S. Safe, *J.C.S. Chem. Comm.*, 1973, 554.
[11] B. R. Davis and I. Bernal, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 279.
[12] H. Minato, M. Matsumoto, and T. Katayama, *J.C.S. Perkin I*, 1973, 1819.
[13] G. M. Strunz, M. Kakushima, M. A. Stillwell, and C. J. Heissner, *J.C.S. Perkin I*, 1973, 2600.
[14] R. L. DeVault and W. Rosenbrook, *J. Antibiotics*, 1973, **26**, 532.

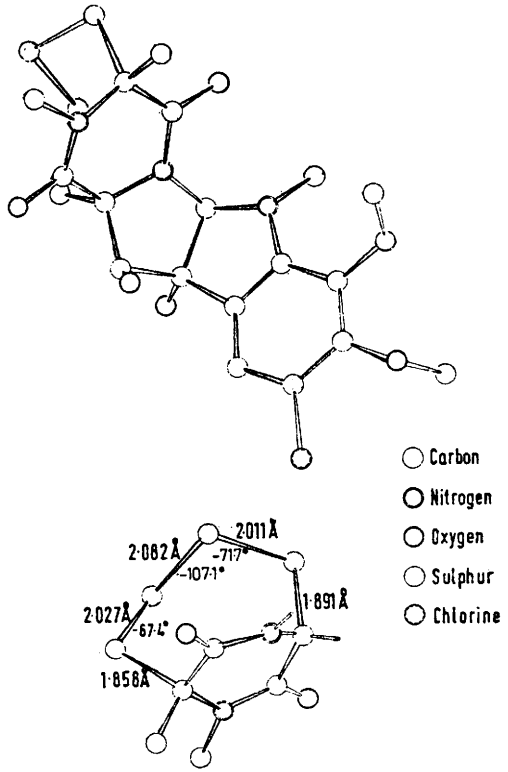
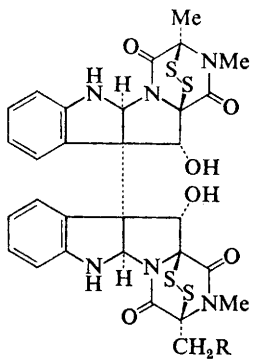
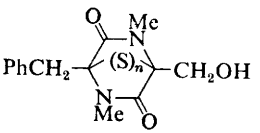


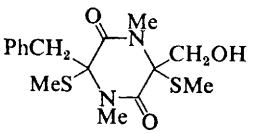
Figure 1 *Perspective drawing of the molecule of sporidesmin G projected down the* c *axis and a view of the tetrathio-piperazinedione portion down the* b *axis*
(Reproduced from *J.C.S. Chem. Comm.*, 1973, 554)



(15) R = H
(16) R = OH

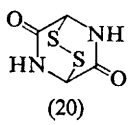


(17) $n = 2$
(18) $n = 3$

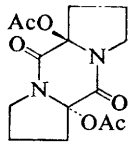


(19)

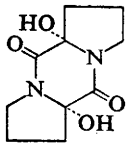
The pronounced antiviral activity of even simple epidithiodioxopiperazine derivatives continues to make this system an attractive challenge for synthetic chemists. The parent ring system (20) has been synthesized [15] by standard procedures from *N*-benzyloxycarbonylthiobenzylglycine. Oxidation of *cyclo*(L-prolyl-L-proline) with lead tetra-acetate affords (21), which is optically active. On acid hydrolysis (21) yields the racemic *cis*-diol (22) which, on treatment with hydrogen sulphide followed by mild oxidation, is converted into the epidithio-derivative (23).[16] Alternative routes for the synthesis of (23) have also been investigated.[17]



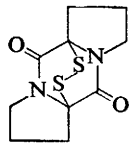
(20)



(21)

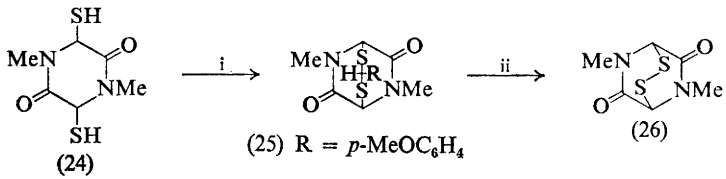


(22)



(23)

A new general route to this important epidithiodioxopiperazine system, outlined in Scheme 2, has been described.[18] Either the *cis*- or the *trans*-dithiol (24) can be converted into the thioacetal (25) which, on subsequent oxidation followed by acid hydrolysis, can be cleaved to give the epidithio-



(24) (25) R = *p*-$MeOC_6H_4$ (26)

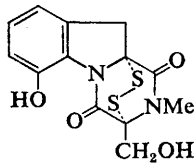
Reagents: i, *p*-$MeOC_6H_4CHO$; ii, *m*-$ClC_6H_4CO_3H/H^+$

Scheme 2

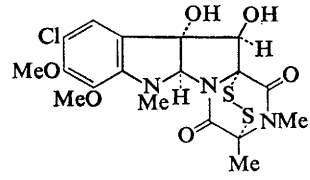
derivative (26). The intermediate thioacetals are stable to acidic, basic, and reductive conditions and therefore can be employed as a very effective protecting group for the epidithio bridge. Their synthetic utility has been admirably demonstrated by the very elegant syntheses of dehydroglio-

[15] T. Petrzilka and Ch. Fehr, *Helv. Chim. Acta*, 1973, **56**, 1218.
[16] E. Oehler, F. Tataruch, and U. Schmidt, *Chem. Ber.*, 1973, **106**, 396.
[17] E. Oehler, F. Tataruch, and U. Schmidt, *Chem. Ber.*, 1973, **106**, 165.
[18] Y. Kishi, T. Fukuyama, and S. Nakatsuka, *J. Amer. Chem. Soc.*, 1973, **95**, 6490.

toxin [19] (27) and sporidesmin A (28).[20] A route to analogues of anhydrogliotoxin involving the addition of α-halogeno-α-aminoacyl chlorides to ethyl indolenine-2-carboxylates, followed by reaction with sulphur nucleophiles and final ring closure, has been developed.[21]



(27)

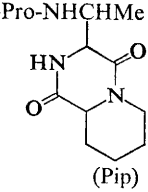


(28)

The structure of the clinically useful antibiotic amphomycin has been studied mainly by standard analytical methods both on the intact molecule and on partial acid hydrolysis products. Amphomycin is a member of a closely related family of antibiotics and is considered to be identical with glumamycin. The structures previously assigned [22] to the latter are disputed [23] and the investigation points to (29) as the general structure for the

FA-Asp-Asp(3-Me)-Asp-Gly-Asp-Gly-*erythro*-αβA_2bu-Val-Pro-NHCHMe (*threo*)

(29) FA = fatty acid

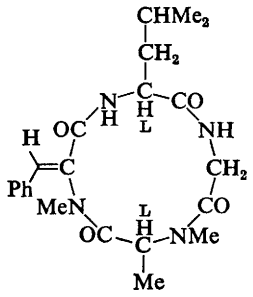


(Pip)

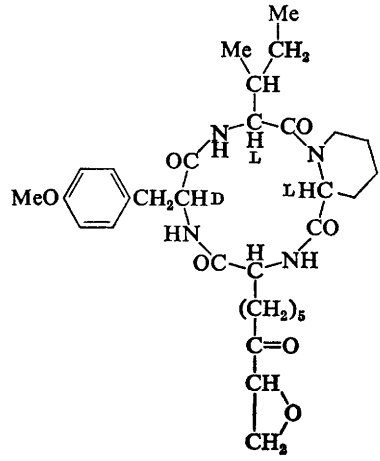
group. Amphomycin is itself a mixture, the major components of which contain the fatty acids (+)-3-anteisotridecenoic acid and (+)-3-isododecnoic acid. Considerable care was taken to establish that the dioxopiperazine system is present in the intact antibiotic and not an artefact.

New Naturally Occurring Cyclic Peptides.—The fungus *Alternaria tenius* produces a cyclic tetrapeptide which causes chlorosis in germinating seedlings. The structure (30) has been proposed [24] for this phytotoxic agent on the basis of standard sequencing techniques and mass spectrometric analysis. The presence of the dehydrophenylalanine residue in the

[19] Y. Kishi, T. Fukuyama, and S. Nakatsuka, *J. Amer. Chem. Soc.*, 1973, **95**, 6492.
[20] Y. Kishi, S. Nakatsuka, T. Fukuyama, and M. Havel, *J. Amer. Chem. Soc.*, 1973, **95**, 6493.
[21] H. C. J. Ottenheym, T. F. Spande, and B. Witkop, *J. Amer. Chem. Soc.*, 1973, **95**, 1989.
[22] M. Fujino, M. Inoue, J. Ueyanagi, and A. Miyake, *Bull. Chem. Soc. Japan*, 1965, **38**, 515.
[23] M. Bodanszky, G. F. Sigler, and A. Bodanszky, *J. Amer. Chem. Soc.*, 1973, **95**, 2352.
[24] M. Koncewicz, P. Mathiaparanam, T. F. Uchytil, L. Sparapano, J. Tam, D. H. Rich, and R. D. Durbin, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 653.



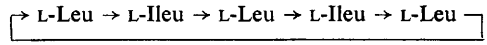
(30)



(31)

toxin is noteworthy. Similar methods were employed in determining the structure (31) of the novel cyclotetrapeptide Cyl-2 from the pathogenic fungus *Cylindrocladium scoparium*.[25, 26] The structure of the unusual 2-amino-8-oxo-9,10-epoxydecanoic acid was established by spectroscopic methods and chemical degradation to aminosuberic acid. The chirality of this residue remains to be determined.

Viscunamide, isolated from mistletoe, contains only L-leucine and L-isoleucine. Sequence analysis of peptide fragments obtained on partial hydrolysis together with end-group determination has allowed the assignment[27] of structure (32) to this cyclopentapeptide. The antibiotic poly-



(32)

myxin M, a further member of the general polymyxin family, and differing from polymyxin E only in the replacement of a leucine residue for threonine, has been reported.[28]

Conformational Studies on Synthetic Homodetic Peptides.—N.m.r. methods, because of their potentially high information content, still dominate as the most generally useful technique for the determination of solution

[25] H. Hirota, A. Suzuki, K. Aizawa, and S. Tamura, *Agric. and Biol. Chem.* (*Japan*), 1973, **37**, 955.
[26] H. Hirota, A. Suzuki, and S. Tamura, *Agric. and Biol. Chem.* (*Japan*), 1973, **37**, 1185.
[27] Y. Okumura and A. Sakurai, *Bull. Chem. Soc. Japan*, 1973, **46**, 2190.
[28] A. B. Silaev, Zh. P. Trifonova, S. N. Maevskaya, N. M. Vasileva, and G. S. Katrukha, *Khim. prirod. Soedinenii*, 1973, 283.

conformation.[29] Considerable effort, both theoretical[30] and experimental,[31, 32] has been made to refine the curve for the dependence of the peptide vicinal NH—C_αH coupling constant on the corresponding dihedral angle. This parameter, which is readily available from n.m.r. data, provides the principal torsion angle ϕ for each amino-acid residue. A further coupling constant–dihedral angle correlation for the HC_α—C_β—H system of amino-acid residues in peptides has been derived experimentally.[33] However, it is pointed out that the rotational freedom of most peptide side-chains limits the information available from this parameter.

Exploratory studies on ^{15}N-enriched amino-acids and peptides have shown[34] that the one-bond N—H coupling constants vary significantly with *cis–trans* isomerism about the peptide bond, suggesting its use in the study of peptide conformation for amino-acids other than proline. The vicinal couplings to side-chain β-protons from nitrogen appear to be a sensitive indicator of the side-chain torsion angle. Unfortunately the vicinal coupling from nitrogen to the next α-proton is too small to be measured accurately at the present time. This is disappointing since it had been anticipated that there would be a direct relationship between this coupling and the peptide torsion angle ψ.

^{1}H N.m.r. spectra of N-substituted peptides exhibit a small five-bond long-range coupling between groups which are anti-periplanar to each other across the peptide bond. The observation of this coupling has been used for *cis* and *trans* amide assignments in peptides containing N-methylated amino-acids.[35, 36]

Various shift-inducing lanthanide salts have been employed in ^{13}C and ^{1}H n.m.r. studies on linear peptides for sequence[37, 38] and conformational[39] analysis and should be of value in the study of cyclic peptides. A detailed investigation[40] of the solution conformations of *cyclo*(L-prolyl-L-proline) and the L-D-isomer, using europium complexes which bind to the carbonyl oxygen of the amides, revealed that the two europium-binding sites are related by C_2 symmetry in the *cis*-isomer and S_2 symmetry in the *trans*. Conformational energy calculations show that there is a flattening of the dioxopiperazine ring from a boat in *cyclo*(L-prolyl-L-proline) to a nearly planar form in cyclo(L-prolyl-D-proline). The calculated pyrrolidine ring

[29] V. F. Bystrov, S. L. Portnova, T. A. Balashova, S. A. Koz'min, Yu D. Gavrilov, and V. A. Afanas'ev, *Pure Appl. Chem.*, 1973, **36**, 19.
[30] V. N. Solkan and V. F. Bystrov, *Tetrahedron Letters*, 1973, 2261.
[31] M. Barfield and H. L. Gearhart, *J. Amer. Chem. Soc.*, 1973, **95**, 641.
[32] V. F. Bystrov, V. T. Ivanov, S. L. Portnova, T. A. Balashova, and Yu. A. Ovchinnikov, *Tetrahedron*, 1973, **29**, 873.
[33] K. D. Kopple, G. R. Wiley, and R. Tauke, *Biopolymers*, 1973, **12**, 627.
[34] J. A. Sogn, N. A. Gibbons, and E. W. Randall, *Biochemistry*, 1973, **12**, 2100.
[35] D. B. Davies and M. A. Khaled, *Tetrahedron Letters*, 1973, 2829,
[36] D. B. Davies and M. A. Khaled, *J.C.S. Perkin II*, 1973, 1651.
[37] M. Anteunis and J. Gelan, *J. Amer. Chem. Soc.*, 1973, **95**, 6502.
[38] E. Bayer and K. Beyer, *Tetrahedron Letters*, 1973, 1209.
[39] L. C. Martinelli, I. L. Honigberg, and A. Sternson, *Tetrahedron*, 1973, **29**, 1671.
[40] P. E. Young, V. Madison, and E. R. Blout, *J. Amer. Chem. Soc.*, 1973, **95**, 6142.

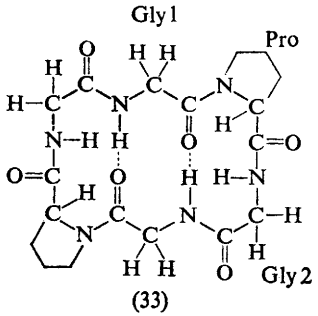
conformers are similar in both cases and correspond well to the data from an *X*-ray analysis on *cyclo*(L-prolyl-L-leucine).[41] In relation to the conformation of 2,5-dioxopiperazines containing a proline residue, it is perhaps noteworthy that the *X*-ray analysis of (1) showed that the N—C_α proline peptide bond is twisted out of planarity by 7.0°. This observation is significant in connection with the general interest in the concept of a non-planar peptide unit.[42]

The crystal structures of *cyclo*(glycyl-L-tyrosine) and *cyclo*(L-seryl-L-tyrosine) have been determined[43] by *X*-ray diffraction techniques. Both possess planar amide units and are folded in the now familar conformation with the aromatic ring of the tyrosine side-chain facing the dioxopiperazine ring.

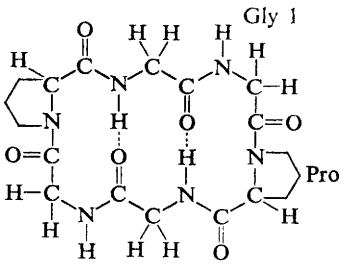
An extensive ^{1}H n.m.r. study on a considerable number of cyclic dipeptides containing pipecolic acid, proline, or 2-azetidine carboxylic acid has been extended and the original conclusions concerning the effect of the ring size of the annealed rings on the conformation were reinforced.[44, 45] Conformational energy calculations on cyclic pentapeptides[46] and cyclic hexapeptides[47, 48] have provided simplified potential energy functions which yield a qualitative conformational distribution. This gives guidance in deriving preferred conformations from experimental data. For example, low-energy conformers of *cyclo*(penta-L-alanyl-D-alanine) whose individual amino-acid residues are in conformations consistent with the observed NH—C_αH coupling constants were searched for in an approximate theoretical treatment. The two lowest energy all-*trans* peptide conformations were then distinguished between by the presence or absence of a single intramolecular hydrogen-bond using the temperature dependence of the chemical shift of the amide protons and deuterium exchange rates.[49]

On the basis of a ^{1}H n.m.r. study, two C_2 symmetric intramolecularly hydrogen-bonded conformations (33) and (34) were proposed for *cyclo*-(glycyl-L-prolyl-glycine)$_2$. Related peptides had already been shown to possess favoured conformations in aqueous solution with the type II β-turns,[50] as in (33). In contrast, gramicidin S assumes solution conformations with type II′ β-turns, analogous to (34). In two further independent[51, 52] investigations on *cyclo*(glycyl-L-prolyl-glycine)$_2$, ^{13}C n.m.r. has

[41] I. L. Karle, *J. Amer. Chem. Soc.*, 1972, **94**, 81.
[42] G. N. Ramachandran, *Biochim. Biophys. Acta*, 1973, **303**, 8.
[43] C. F. Lin and L. E. Webb, *J. Amer. Chem. Soc.*, 1973, **95**, 6803.
[44] J. Vičar, M. Buděšinský, and K. Bláha, *Coll. Czech. Chem. Comm.*, 1973, **38**, 1940.
[45] J. Vičar, J. Smoliková, and K. Bláha, *Coll. Czech. Chem. Comm.*, 1973, **38**, 1957.
[46] G. C. C. Niu, N. Go, and H. A. Scheraga, *Macromolecules*, 1973, **6**, 91.
[47] N. Go and H. A. Scheraga, *Macromolecules*, 1973, **6**, 523.
[48] V. Madison, *Biopolymers*, 1973, **12**, 1837.
[49] A. E. Tonelli and A. I. R. Brewster, *Biopolymers*, 1973, **12**, 193.
[50] C. M. Venkatachalam, *Biopolymers*, 1968, **5**, 1425.
[51] C. Grathwohl, A. Tun-Kyi, R. Schwyzer, and K. Wuthrich, *F.E.B.S. Letters*, 1973, **29**, 271.
[52] L. G. Pease, C. M. Deber, and E. R. Blout, *J. Amer. Chem. Soc.*, 1973, **95**, 258.



(33)



(34)

been used to identify the resonances associated with each of the Gly residues and establish that (33), with the glycine preceding the proline intramolecularly hydrogen-bonded, is the preferred conformer. The related cyclic hexapeptides, *cyclo*(L-X-D-Phe-L-Pro-)$_2$ where X = alanine,

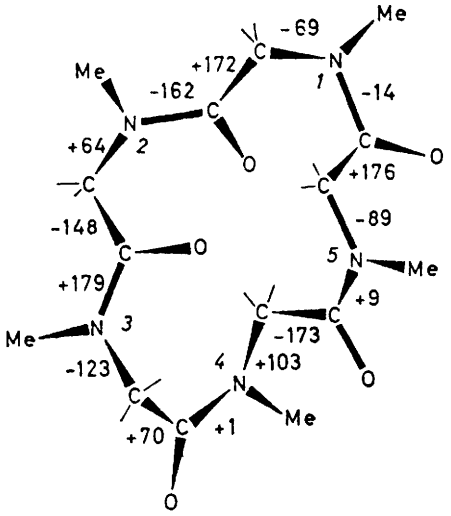


Figure 2 *Crystal structure of cyclopentasarcosyl*
(Reproduced from *J.C.S. Chem. Comm.*, 1973, 646)

histidine, or ornithine, all exhibit C_2 symmetry, but n.m.r. evidence indicates that they all possess solution conformations with the type II′ β-turn structures, *i.e.* (34) with D-phenylalanine and X instead of glycine-1 and glycine-2, respectively.[53]

[53] K. D. Kopple, A. Go, T. J. Schamper, and C. S. Wilcox, *J. Amer. Chem. Soc.*, 1973, **95**, 6090.

An important and interesting correlation has been made between the crystal and solution conformations of cyclopentasarcosyl [54] and cyclo-octasarcosyl.[55] The crystal structures of both molecules have been determined (Figures 2 and 3) and n.m.r. studies in each case have led to the conclusion that one dominant conformer exists in solution which corresponds to that observed in the crystal lattice.

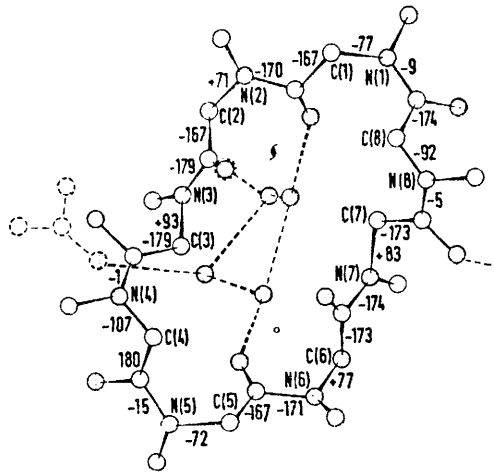


Figure 3 *Crystal structure of cyclo-octasarcosyl tetrahydrate. Open circles are carbon atoms, filled circles nitrogen (marked) or oxygen atoms. Dihedral angles refer to the ring skeleton*
(Reproduced from *J.C.S. Chem. Comm.*, 1973, 346)

Conformational Studies on Naturally Occurring Cyclic Peptides.—The assault by the now considerable battery of physical methods on the conformation of naturally occurring peptides continues unabated. Impressive results have already been achieved, when it is considered that only a few years ago dioxopiperazine conformation was the acme of the investigations in this area. For cyclic dipeptides, the *X*-ray structures and conformations of (1) and sporidesmin G (Figure 1) have been described [2, 10] and an extensive study on the c.d. of gliotoxin and related antibiotics reported.[56] The apparent violation of the skewed diene rule by gliotoxin is accounted for by the coupling of the diene $\pi\pi^*$ transition with the disulphide $n\sigma^*$ transition. In the c.d. spectra of sporidesmin and chaetocin, the sign of the long-wavelength disulphide transition does not obey the usual disulphide chirality rule and it is proposed that this is due to coupling with the peptide $\pi\pi^*$ transitions and perhaps the indolinyl chromophore.

[54] K. Titlestad, P. Groth, and J. Dale, *J.C.S. Chem. Comm.*, 1973, 646.
[55] K. Titlestad, P. Groth, J. Dale, and M. Y. Ali, *J.C.S. Chem. Comm.*, 1973, 346.
[56] R. Nagarajan and R. W. Woody, *J. Amer. Chem. Soc.*, 1973, **95**, 7212.

The solution conformation of malformin A, a cyclic pentapeptide with an intramolecular disulphide bridge, has been studied by n.m.r. and c.d. methods.[57] *P* Chirality for the disulphide linkage is proposed on the basis of the $HC_\beta—C_\alpha H$ coupling constant and the negative band at 250 nm in the c.d. spectrum. The disulphide bridge limits the conformational mobility

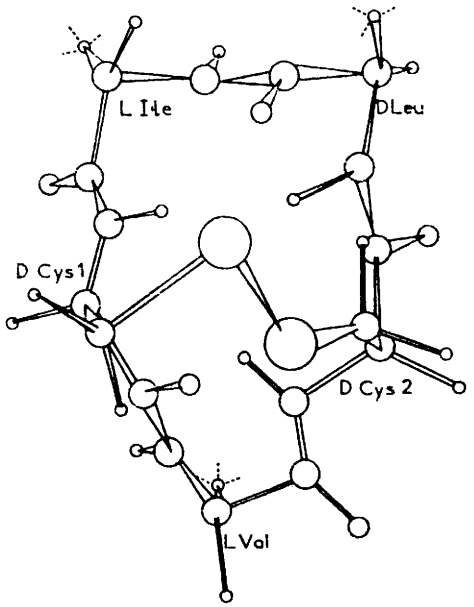


Figure 4 *Proposed conformation of malformin A*
(Reproduced by permission from *Biopolymers*, 1973, **12**, 1575)

of the molecule and the proposed conformation of the peptide backbone is shown in Figure 4.

Several conformations, generated from approximate potential energy calculations and which are consistent with the information obtained from ^1H n.m.r. studies, have been presented for the bicyclic heptapeptide phalloidin.[58] In all the conformers considered the cysteine amide proton is intramolecularly hydrogen-bonded and the tryptophan amide proton internally buried. It is concluded that in solution phalloidin is a relatively rigid molecule.

A more complex situation pertains with respect to the antitoxic cyclododecapeptide antamanide and its Na^+ complex. Detailed studies indicate that similar conformational changes are induced in antamanide by interaction with cations and polar solvents, although the interpretation of these

[57] M. Ptak, *Biopolymers*, 1973, **12**, 1575.
[58] D. J. Patel, A. E. Tonelli, P. Pfaender, H. Faulstich, and Th. Wieland, *J. Mol. Biol.*, 1973, **79**, 185.

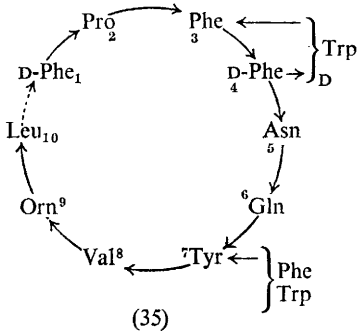
changes is controversial. On the basis of an investigation on antamanide and a number of analogues, the Russian group [59] have proposed an all-*trans* peptide conformer for free antamanide in non-polar solvents, whereas American workers [60] favour a conformation containing *cis* peptide bonds at Val[1]-Pro[2] and Phe[6]-Pro[7]. Similar discrepancies, which perhaps serve to illustrate the present limits of the techniques available for conformational studies, exist concerning the conformation of the sodium complex.[61–63]

The solution conformation of gramicidin S is now well established. Information concerning the rotational motion of the molecular backbone and internal rotation in the side-chains is available from the ^{13}C spin–lattice relaxation times of protonated carbons.[64] A study of the c.d. of copper complexes of gramicidin S has been interpreted in terms of the molecular conformation.[65] Substitution of Ga^{3+} and Al^{3+} for Fe^{3+} in the ferrichromes has little effect on the solution conformation of these peptides but markedly affects the rate of hydrogen–deuterium exchange for the amide protons. It was concluded that it is insufficient to describe the solution conformation of a peptide without reference to its stability in the defined conformer.[66, 67] Tentative assignments for each of the carbon resonances in the ^{13}C n.m.r. spectra of bacitracin have been made and initial conclusions concerning the conformation in aqueous solution drawn.[68] Further inconclusive conformational studies on alamethicin have been reported.[69]

Known Naturally Occurring Cyclic Peptides.—In previous years it has been customary to deal with the well-known naturally occurring cyclic peptides under separate headings. However, in view of the shift in emphasis in this section, headings of a more general nature have been used. The majority of references relate to conformational and synthetic studies and these are now to be found under the appropriate heading.

It is now fairly well established that a ribosomal-type synthesis can be excluded for the majority of microbial peptide antibiotics. The principal exceptions are probably members of the nisin and subtilin families (see Section 2). Further cell-free work on the biosynthesis of tyrocidin (35) has resulted in the isolation of three complementary enzyme fractions.[70] These

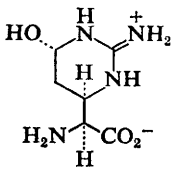
[59] V. T. Ivanov, A. I. Miroshnikov, S. A. Koz'min, E. N. Meshcheryakova, L. B. Senyavina, N. N. Uvarova, K. Kh. Khalilulina, V. A. Zabrodin, and V. F. Bystrov, *Khim. prirod. Soedinenii*, 1973, **9**, 378.
[60] D. J. Patel, *Biochemistry*, 1973, **12**, 667.
[61] D. J. Patel, *Biochemistry*, 1973, **12**, 677.
[62] A. E. Tonelli, *Biochemistry*, 1973, **12**, 689.
[63] V. F. Bystrov, V. T. Ivanov, S. A. Koz'min, I. I. Mihaleva, K. Kh. Khalilulina, Yu. A. Ovchinnikov, E. I. Fedin, and P. V. Petrovski, *F.E.B.S. Letters*, 1972, **21**, 24.
[64] A. Allerhand and R. A. Komoroski, *J. Amer. Chem. Soc.*, 1973, **95**, 8228.
[65] P. DeSantis, L. D'Ilario, G. Lamanna, and S. Morosetti, *Biopolymers*, 1973, **12**, 423.
[66] M. Llinás, P. Klein, and J. B. Neilands, *J. Biol. Chem.*, 1973, **248**, 915.
[67] M. Llinás, P. Klein, and J. B. Neilands, *J. Biol. Chem.*, 1973, **248**, 924.
[68] J. R. Lyerla and M. H. Freedman, *J. Biol. Chem.*, 1972, **247**, 8183.
[69] A. W. Burgess and S. J. Leach, *Biopolymers*, 1973, **12**, 2691.
[70] S. G. Lee, R. Roskoski, K. Bauer, and F. Lipmann, *Biochemistry*, 1973, **13**, 398.



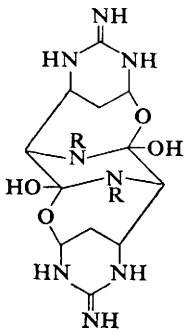
(35)

fractions activate the component amino-acids so that they become thio-ester linked to the enzymes. On combination they polymerize sequentially, beginning from the *N*-terminal phenylalanine. Addition of the last amino-acid, leucine, causes the release of the cyclic decapeptide. The general sequence is similar to that already established for gramicidin S.[71]

Interest still continues in the antitubercular antibiotics viomycin and tuberactinomycin and the relationship between the antimicrobial activity and chemical structure has been investigated.[72] Acidic hydrolysis of tuberactinomycin affords the guanidino amino-acid (36) together with the



(36)



(37) R = CHO

novel dimer (37), which is an artefact formed under the conditions of isolation.[73] The carbinolamine system of (37) is reminiscent of structures of tetrodotoxin and the ergot peptides.

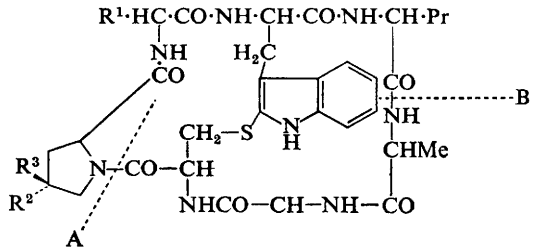
Synthesis of Homodetic Cyclic Peptides.—There appears to have been less activity than usual in this area over the past year. The main emphasis has

[71] F. Lipmann, *Science*, 1971, **173**, 875l
[72] T. Kitagawa, T. Miura, S. Tanaka, and H. Taniyama, *J. Antibiotics*, 1973, **26**, 528.
[73] T. Nakamiya, T. Shiba, T. Kaneko, H. Sakakibara, T. Noda, and T. Take, *Bull. Chem. Soc. Japan*, 1973, **48**, 949.

been centred on the synthesis of specific cyclic peptides for conformational studies, but in a number of cases full details of the methods have not been reported. In addition, analogues of biologically active peptides still present an important synthetic challenge. Reviews which cover the strategy and methodology for the synthesis of cyclic peptides have been published.[1, 74]

[MeLeu^3]-Gramicidin and [MeLeu^3,MeLeu$^{3'}$]-gramicidin, which lack one and two, respectively, of the four intramolecular hydrogen bonds of gramicidin, have been synthesized.[75] The linear precursors were prepared by standard procedures and cyclized using carbodi-imide and *N*-hydroxysuccinimide. This method is claimed to give better yields than the *p*-nitrophenyl ester method previously used for gramicidin synthesis. The carbodi-imide-*N*-hydroxysuccinimide procedure has also been employed for the synthesis of sesqui- and di-gramicidin S.[76] An o.r.d. study suggests that the *N*-methyl analogues have similar conformations to gramicidin, indicating that the hydrogen bonds do not contribute substantially to conformational stabilization. The macro-ring analogues have conformations and antibacterial activity similar to those of the corresponding linear analogues. A further synthesis of [4,5-δ-aminovaleric acid]-gramicidin using the *p*-nitrophenyl ester method has been described.[77]

The synthesis of the bicyclic toxin norphalloin (38) developed[78] some years ago has been extended[79, 80] for the synthesis of the analogues (39)—(41). The method requires two cyclization steps at the positions A and B, both of which were achieved by the mixed anhydride procedure. The analogues (39)—(41), in contrast to (38), are non-toxic and exhibit different

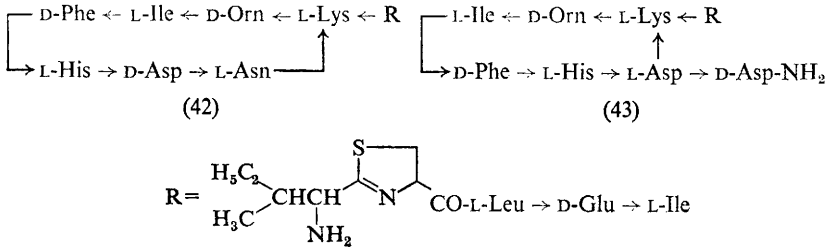


(38) R^1 = Me, R^2 = H, R^3 = OH
(39) R^1 = Me, R^2 = OH, R^3 = H
(40) R^1 = Me, R^2 = H, R^3 = H
(41) R^1 = H, R^2 = H, R^3 = OH

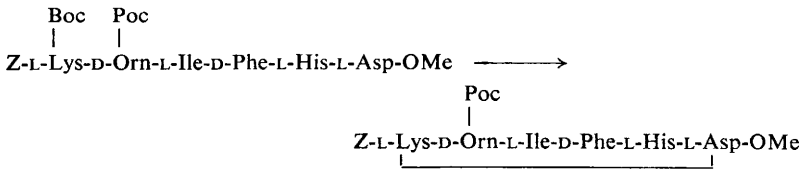
[74] J. Meienhofer, *Chem. Technol.*, 1973, **3**, 242.
[75] H. Sugano, H. Abe, M. Miyoshi, T. Kato, and N. Izumiya, *Experientia*, 1973, **29**, 1488.
[76] S. Matsuura, M. Waki, T. Kato, and N. Izumiya, *Bull. Chem. Soc. Japan*, 1973, **46**, 977.
[77] S. Sofuku, *Bull. Chem. Soc. Japan*, 1973, **46**, 968.
[78] Th. Wieland, H. Faulstich, and F. Fahrenholz, *Annalen*, 1971, **743**, 77, 83.
[79] E. Nebelin, H. Faulstich, and Th. Wieland, *Annalen*, 1973, 45.
[80] H. Faulstich, E. Nebelin, and Th. Wieland, *Annalen*, 1973, 50.

o.r.d. spectra which in turn differ from the spectrum of (38). Further studies aimed at defining the structural requirements for the antitoxic activity and metal-complexing ability of antamanide have resulted in the synthesis of [Phe[1],Ala[9]]-antamanide and [Val[6],Ala[9]]-antamanide by the active ester method.[81] Both analogues, along with their perhydro-derivatives, form more stable sodium and potassium complexes than antamanide itself but exhibit weaker biological activity.

The structure of the antibiotic bacitracin still remains something of an enigma and further work directed at solving this problem has been described. The structures (42) and (43) have been taken as the basis of an extensive synthetic study.[82] However, neither of these structures is firmly



(42) (43)



(44) Poc = cyclopentyloxycarbonyl (45)

established and there is evidence that the terminal amino-group of the leucine residue is not free but involved in a cyclol structures. Nevertheless, the linear derivative (44) has been prepared and cyclized by the azide method to the protected cyclic hexapeptide (45). As part of this investigation the cyclopentyloxycarbonyl group has been developed as a protecting group for the δ-amino of the ornithine residue. The intermediate (44) should prove of value both for the total synthesis of (42) and (43) and general structure–activity correlations.

As part of an investigation related to elucidating the mechanism of enzyme action, the cyclic decapeptide (46) has been prepared as an esterase model.[83] The small esterase-like activity observed for (46) in the hydrolysis

[81] A. I. Miroshnikov, K. Kh. Khalilulina, N. N. Uvarova, V. T. Ivanov, and Yu. A. Ovchinnikov, *Khim. prirod. Soedinenii*, 1973, 214.

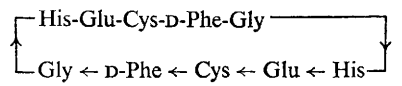
[82] E. Munekata, Y. Masui, T. Shiba, and T. Kaneko, *Bull. Chem. Soc. Japan*, 1973, **46**, 3187.

[83] K. Nakajima and K. Okawa, *Bull. Chem. Soc. Japan*, 1973, **46**, 1811.

Table *Syntheses of cyclic peptides achieved in* 1973 (*derived from linear analogues unless otherwise stated*)

Peptide	*Bond formed in cyclization step*	*Method for cyclization*	*Ref.*
cyclo-(L-Ala-D-Phe-L-Pro)$_2$	Pro-Ala	*a*	53
cyclo-(L-Orn-D-Phe-L-Pro-)$_2$	Pro-Orn	*a*	53
cyclo-(L-His-D-Phe-L-Pro-)$_2$	Pro-His	*a*	53
[MeLeu^3]-gramicidin S	Pro-Val	*b*	75
[MeLeu^3,MeLeu$^{3'}$]-gramicidin S	Pro-Val	*b*	75
cyclo-(L-Val-L-Orn-L-Leu-D-Phe-L-Pro-)$_3$ (sesquigramicidin S)	Pro-Val	*b*	76
cyclo-(L-Val-L-Orn-L-Leu-D-Phe-L-Pro-)$_4$ (digramicidin S)	Pro-Val	*b*	76
[δ-Ava$^{4\text{-}5}$]-gramicidin S [e]	Pro-Val	*c*	77
[Hyp^4]-norphalloin (39)	Nva-Ala	*d*	80
[Pro^4]-norphalloin (40)	Nva-Ala	*d*	80
[Gly^5]-norphalloin (41)	Nva-Ala	*d*	80
[Phe^1,Ala^9]-antamanide	Pro-Ala	*c*	81
[Val^6,Ala^9]-antamanide	Pro-Ala	*c*	81
cyclo-[L-Lys(N^α-Z)-D-Orn(N^δ-Poc)-L-Ile-D-Phe-L-His-L-Asp(α-OMe)-] (45)	Asp-Lys	*a*	82
cyclo-(L-His-L-Glu-L-Cys-D-Phe-Gly-) (46)	Gly-His	*a*	83

[a] Azide method; [b] DCCI–*N*-hydroxysuccinimide–pyridine; [c] *p*-nitrophenyl ester; [d] mixed anhydride; [e] δ-Ava = δ-aminovaleric acid.



(46)

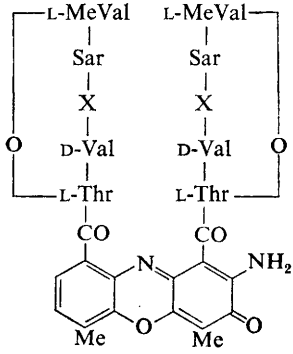
of *p*-nitrophenyl acetate was three times greater than that of the linear decapeptide precursor.

The Table summarizes the syntheses of cyclic peptides reported during 1973.

3 Depsipeptides and Other Heterodetic Cyclic Peptides

Actinomycin.—A number of new actinomycins which differ from actinomycin D (47) at the imino site of the molecule have been isolated from various micro-organisms.[84, 85] Three new actinomycins (48)—(50) are produced by *Streptomyces antibioticus*, when it is grown in the presence of DL-pipecolic acid, representing a further example of directed biosynthesis.[85] 3-Hydroxy-4-oxo-5-methyl-L-proline has been identified [86] as the imino-acid in actinomycin Z_1, whereas actinomycin Z_5 contains 5-methyl-L-proline.[87] The amino-acid sequence in these antibiotics has not yet been determined but it is clearly different from that in (47).

An attempt to distinguish by chemical means which of the lactone rings of actinomycin D is open in the monolactone isolated from *Streptomyces*



(47) X = Pro
(48) X = Pro and Pip
(49) X = Pip
(50) X = Pip and 4-oxo-Pip

[84] T. I. Orlova, N. V. Sorokina, V. D. Kuznetsov, and A. B. Silaev, *Antibiotiki*, 1973, **18**, 111.
[85] J. V. Formica and E. Katz, *J. Biol. Chem.*, 1973, **248**, 2066.
[86] H. Brockmann and E. A. Stahler, *Tetrahedron Letters*, 1973, 3685.
[87] E. Katz, K. T. Mason, and A. B. Mauger, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 819.

antibioticus has been unsuccessful.[88] Additional evidence concerning the role of 3-hydroxy-4-methylkynurenine as an intermediate in the biosynthesis of the actinomycins has been described.[89]

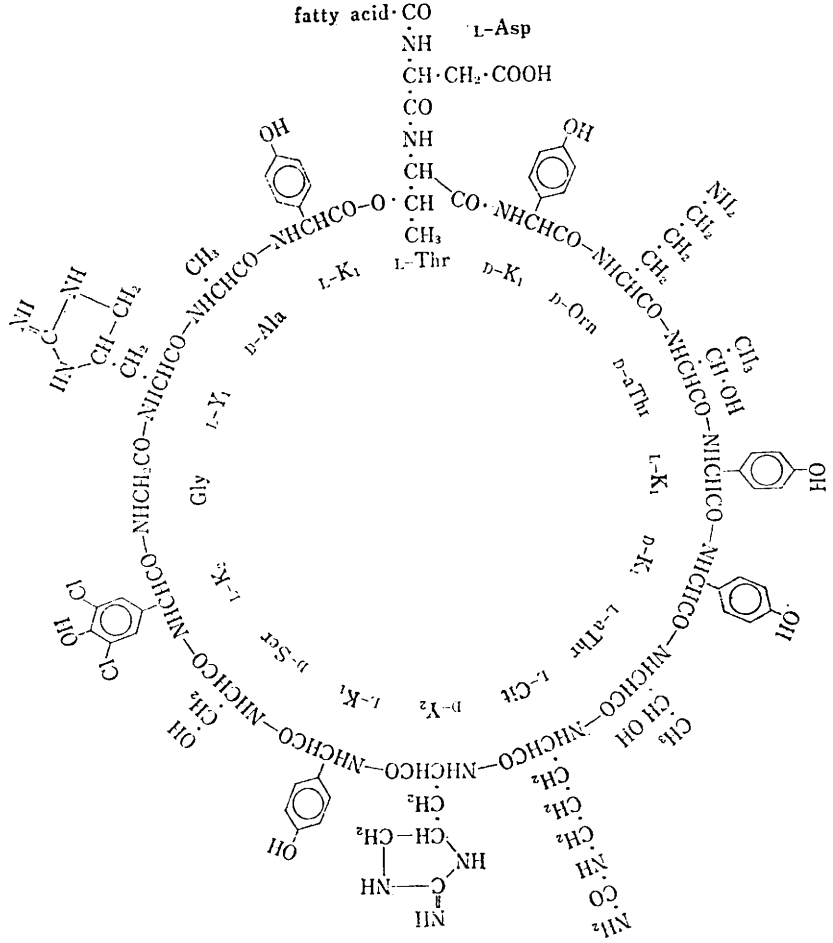
The ^{1}H n.m.r. spectra of several cyclic and acyclic peptides representing portions of the actinomycin structure have been determined and the lanthanide shift reagent Eu(fod)$_3$ was used to differentiate between *N*-methyl groups in various conformations.[90] A theoretical treatment of the molecular structure of actinomycin has been refined using energy minimization calculations.[91] The conformation of dimethyl actinocynil bis-(L-threonate), a synthetic model compound, has been investigated by a variety of physical methods including *X*-ray crystallographic analysis.[92] The differences observed in the physical data from those obtained from the naturally occurring antibiotic were accounted for in terms of structural differences around the phenoxazone chromophore. The interaction between actinomycin D and individual mononucleotides[93] and deoxydinucleotides[94] has been the subject of a ^{1}H n.m.r. investigation. The results indicate that the complexes consist of guanine bases stacked each side of actinomycin molecules. For dinucleotide complexes the geometry was inferred from the induced shifts of the actinomycin protons on complex formation.

Actinomycin D lactam, in which the threonine residues have been replaced with L-*threo*-$\alpha\beta$-diaminobutyric acid, has been synthesized as a potential antitumour chemotherapeutic agent.[95]

Valinomycin.—The interest in valinomycin continues to be centred mainly on conformational aspects. Evidence from Raman spectral data on the uncomplexed molecule accords with the *X*-ray crystallographic analysis concerning the conformation in the crystal lattice.[96, 97] A combination of conformational energy calculations and information derived from ^{1}H n.m.r. spectra has been used in a search for low-energy structures.[98] The most favoured conformer derived accords with that proposed earlier and recently substantiated.[99] The large conformation changes in the depsipeptide on the formation of the potassium complex have been followed by ^{13}C

[88] R. W. Rickards, K. L. Perlman, and D. Perlman, *J. Antibiotics*, 1973, **26**, 177.
[89] D. Perlman, S. Otani, K. L. Perlman, and J. E. Walker, *J. Antibiotics*, 1973, **26**, 289.
[90] A. B. Mauger, N. J. Rzeszotarski, and R. A. Ford, *Org. Magn. Resonance*, 1973, **5**, 231.
[91] P. K. Ponnuswanmy, R. F. McGuire, and H. A. Scheraga, *Internat. J. Peptide Protein Res.*, 1973, **5**, 73.
[92] B. Anastasi, F. Ascoli, P. Costantino, P. DeSantis, R. Rizzo, and M. Savino, *Biochemistry*, 1973, **12**, 1834.
[93] T. R. Krugh and J. W. Neely, *Biochemistry*, 1973, **12**, 1775.
[94] T. R. Krugh and J. W. Neely, *Biochemistry*, 1973, **12**, 4418.
[95] E. Atherton, R. P. Patel, Y. Sano, and J. Meienhofer, *J. Medicin. Chem.*, 1973, **16**, 355.
[96] W. L. Daux, H. Hauptman, C. M. Weeks, and D. A. Norton, *Science*, 1972, **176**, 911.
[97] K. J. Rothschild, I. M. Asher, E. Anastassakis, and H. E. Stanley, *Science*, 1973, **182**, 384.
[98] D. J. Patel and A. E. Tonelli, *Biochemistry*, 1973, **12**, 486.
[99] V. T. Ivanov, I. A. Laine, Yu. A. Ovchinnikov, I. I. Chervin, and G. I. Yakovlev, *Khim. prirod. Soedinenii*, 1973, 248.

n.m.r.[100, 101] The carbonyl resonances involved in complexation move downfield significantly on co-ordination with diamagnetic univalent metal ions. The kinetics and mechanism of the selective binding of alkali-metal ions with valinomycin and the related antibiotic enniatin B have been discussed in detail.[102] MO calculations on the conformations of valinomycin[103] and enniatin B[104] conform reasonably well with the *X*-ray data.

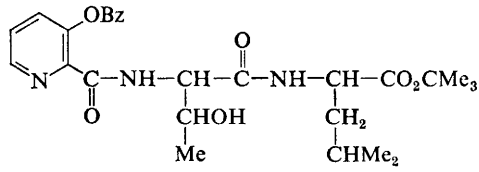


(51)

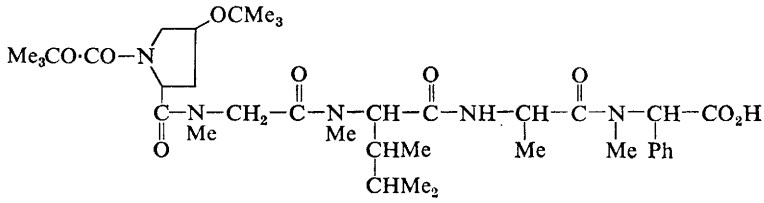
[100] D. J. Patel, *Biochemistry*, 1973, **12**, 496.
[101] E. Grell, T. Funck, and H. Sauter, *European J. Biochem.*, 1973, **34**, 415.
[102] T. Funck, F. Eggers, and E. Grell, *Chimia*, 1972, **26**, 637.
[103] N. L. Max, *Biopolymers*, 1973, **12**, 1565.
[104] B. Maigret and B. Pullman, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 908.

Other Cyclic Depsipeptides.—The microbial peptides enduracidin A and B differ from one another only with respect to the fatty acid component. A primary structure (51) for the common peptide entity has been proposed [105, 106] on the basis of an extensive chemical and physicochemical investigation. The amino-acid sequence was established by the analysis of peptide fragments from partial hydrolysis and the chirality of the amino-acids determined enzymatically. The position of the lactone linkage followed from the reduction of enduracidin with lithium borohydride and the subsequent identification of the resulting amino-alcohol. If the structure (51) is correct then the enduracidins represent some of the largest cyclic peptides yet isolated.

The antibiotic etamycin has been synthesized by joining the two protected peptide fragments (52) and (53). The formation of the ester bond between



(52)



(53)

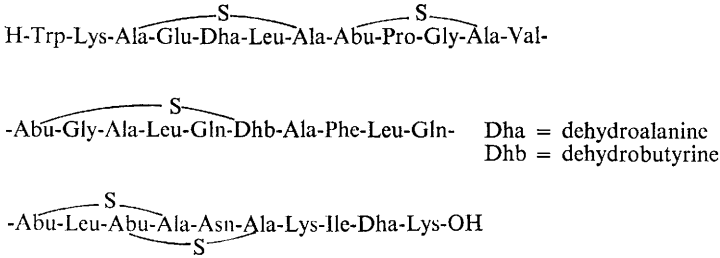
the phenylsarcosine and the threonine residues was mediated by carbonyl di-imidazole. The blocking groups of the coupled peptide were cleaved with trifluoroacetic acid and the free peptide cyclized in very dilute solution with an excess of carbodi-imide. Hydrogenation gave etamycin which was identical with the natural material.[107]

The solution conformation of the lactone antibiotic telomycin has been studied by high-resolution ^{1}H n.m.r. techniques.[108] The data obtained indicate that the threonine, hydroxyleucine, and β-methyltryptophan peptide protons are all solvent-shielded. A preferred conformer in which

[105] M. Hori, H. Iwasaki, S. Horii, I. Yoshida, and T. Hongo, *Chem. and Pharm. Bull. Japan*, 1973, **21**, 1175.
[106] H. Iwasaki, S. Horii, M. Asai, K. Mizuno, and J. Ueyanagi, *Chem. and Pharm. Bull. Japan*, 1973, **21**, 1184.
[107] J. C. Sheehan and S. L. Ledis, *J. Amer. Chem. Soc.*, 1973, **95**, 875.
[108] N. G. Kumar and D. W. Urry, *Biochemistry*, 1973, **12**, 3811, 4372.

there are three hydrogen-bonded rings of atoms, two of which are typical β-turns, is proposed. Further conformational studies reported concern cyclic octadepsipeptides with alternating α-hydroxy-acids and L-α-methyl amino-acids,[109] medium-size depsipeptide rings,[110] and the octadepsipeptide *cyclo*(Sar_3-O·CH_2·CO-Sar_3-O·CH_2·CO-).[55]

Subtilin and Nisin.—The important heterodetic polycyclic peptide subtilin from *Bacillus subtilis* has been assigned the structure (54) on the basis of



(54)

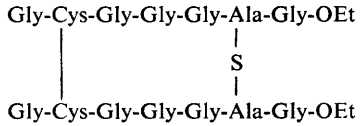
an extensive investigation.[111] The molecule is composed of 27 amino-acid residues, eight of which, *i.e.*, four 3-methyl-lanthionine, two dehydroalanine, and one each of dehydro-α-aminobutyric acid and lanthionine, are not commonly found in nature. The dehydroamino-acid residues could not be detected directly since they are degraded under acidic hydrolysis conditions to ammonia and the corresponding α-keto-acid. However, it was possible to detect them indirectly as *S*-benzylcysteine and *S*-benzyl-methylcysteine in the acid hydrolysate of the product resulting from the addition of benzylmercaptan to subtilin.[112] Tryptic and thermolytic cleavage of the molecule afforded a number of peptide fragments which lent themselves to structure elucidation by standard procedures.[113–115] The alanine moiety of lanthionine and all four aminobutyric acid moieties of the 3-methyl-lanthionine residues have been shown to possess the D-configuration. The complete stereochemistry of 3-methyl-lanthionine from both subtilin and the related antibiotic nisin has been established.[116, 117]

The biosynthesis of subtilin and nisin is of considerable interest since there is evidence[118] to suggest that nisin at least is formed initially under

[109] V. Z. Pletnev and E. P. Popov, *Khim. prirod. Soedinenii*, 1973, 220.
[110] N. D. Adullaev, L. I. Andreeva, V. K. Antonov, V. F. Bystrov, E. S. Efremov, and M. M. Shemyakin, *Zhur. org. Khim.*, 1972, **8**, 1859.
[111] E. Gross, H. H. Kiltz, and E. Nebelin, *Z. physiol. Chem.*, 1973, **354**, 810.
[112] E. Gross, H. H. Kiltz, and L. Craig, *Z. physiol. Chem.*, 1973, **354**, 799.
[113] H. H. Kiltz and E. Gross, *Z. physiol. Chem.*, 1973, **354**, 802.
[114] H. H. Kiltz and E. Gross, *Z. physiol. Chem.*, 1973, **354**, 805.
[115] E. Nebelin and E. Gross, *Z. physiol. Chem.*, 1973, **354**, 807.
[116] J. L. Morell and E. Gross, *J. Amer. Chem. Soc.*, 1973, **95**, 6480.
[117] J. R. Knox and P. C. Keck, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 567.
[118] L. Ingram, *Biochim. Biophys. Acta*, 1970, **224**, 263.

ribosomal control, whereas it is now generally accepted that small microbial peptides containing structural units not typical of protein are produced by enzymatic processes independent of protein biosynthesis. The thio-ether bridges in both molecules probably result from the intramolecular addition of cysteine thiol groups to dehydroalanine or dehydrobutyrine units which in turn are derived by the dehydration of serine and threonine, respectively.

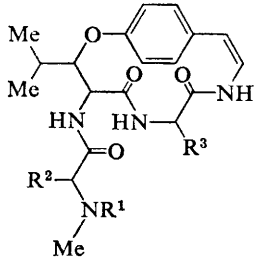
Lanthionine has been synthesized by a novel rearrangement of cysteinyl-serine,[119] and the homodetic-heterodetic cyclopeptide (55) containing a lanthionine residue has been prepared.[120]



(55)

4 Peptide Alkaloids

New peptide alkaloids continue to be isolated from the plants of the Rhamnaceae family. Texensine (56) has been isolated from *Colbrina texensine* in the course of a general search for tumour inhibitory substances.[121] The structure was deduced mainly from an analysis of its mass spectrum which showed the characteristic fragmentation pattern for this class of compounds. The known alkaloid frangulanine (57) is the major component of the alkaloids from *Hovenia dulcis* and *Hovenia tommentella*.[122] Two new compounds designated hovenine-A and -B were also isolated and hovenine-A shown to be de-*N*-methylfrangulanine (58). The structure of hovenine-B was not reported. The relative stereochemistry of the β-hydroxyleucine from the closely related alkaloid lasiodine B has



(56) R^1 = Me, R^2 = Bu^i, R^3 = β-indolylmethyl
(57) R^1 = Me, R^2 = Bu^s, R^3 = Bu^i
(58) R^1 = H, R^2 = Bu^s, R^3 = Bu^i

[119] L. Zervas and N. Ferderigos, *Experientia*, 1973, **29**, 262.
[120] A. Schöberl, M. Rimpler, and E. Graf, *Annalen*, 1973, 1379.
[121] M. C. Wani, H. L. Taylor, and M. E. Wall, *Tetrahedron Letters*, 1973, 4675.
[122] M. Tokai, Y. Ogihara, and S. Shibata, *Phytochemistry*, 1973, **12**, 2985.

been established by the stereospecific synthesis of both iosmers and corresponds to that assigned to the other members of the series.[123]

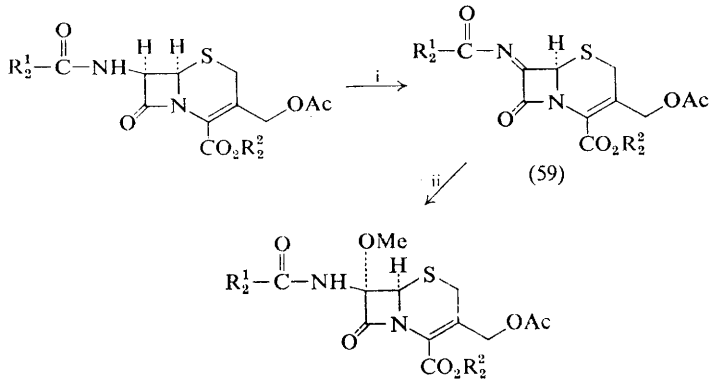
5 Penicillins and Cephalosporins

The considerable interest in this field has continued unabated and no attempt is made here to give a fully comprehensive survey of the vast amount of literature published throughout 1973. The reviewer has instead endeavoured to outline the most significant developments in the major areas most directly related to chemistry.

An excellent book covering the literature on the chemistry and biology of the cephalosporins and penicillins up to 1972 has been published.[124] In addition, two authoritative reviews on the recent advances concerning the chemical modification and medicinal aspects of penicillins have also appeared.[125, 126]

The recent isolation and structure determination of the 7-methoxycephalosporin antibiotics from *Streptomyces* described in last year's Report have aroused considerable attention. The possibility that they are derived biosynthetically from cephalosporins through the intermediacy of an acylimine has general implications in relation to the biosynthesis of microbial peptides. A number of synthetic approaches to 6-methoxypenicillins and 7-methoxycephalosporins employing an acylimine intermediate have been described.

An elegant and simple one-step synthesis involves generating the acylimine (59) by N-chlorination with t-butyl hypochlorite at low tempera-



Reagents: i, Bu^{t}OCl–LiOMe; ii, HOAc–MeOH

Scheme 3

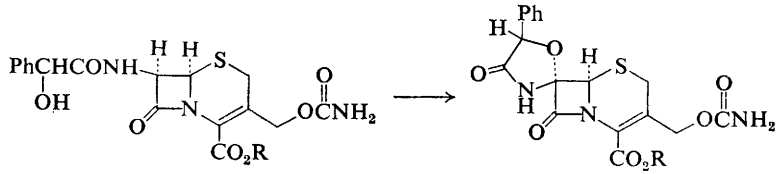
[123] J. Marchand, F. Rocchiccioli, M. Pais, and F. X. Jarreau, *Bull. Soc. chim. France*, 1972, 4699.

[124] 'Cephalosporins and Penicillins, Chemistry and Biology,' ed. E. H. Flynn, Academic Press, New York and London, 1972.

[125] J. H. C. Naylor, in 'Advances in Drug Research', ed. N. J. Harper and A. B. Simmonds, Academic Press, New York and London, 1973, Vol. 7, p. 1.

[126] R. J. Stoodley, *Progr. Org. Chem.*, 1973, **8**, 102.

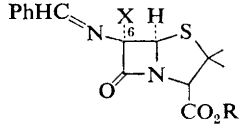
tures. Addition of methanol leads to the 7-α-methoxy-derivative in high yield.[127] It is surprising that the sensitive functional groups in the cephalosporin molecule survive these severe reaction conditions. The proposed acylimine intermediate can be trapped with an internal nucleophile. Oxidation of (60) under the same conditions afforded the spiro-compound (61).[128] In a similar independent approach to 6-α-methoxypenicillins, in



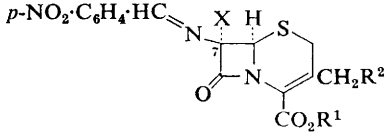
(60) (61)

which t-butyl hypochlorite was employed under different reaction conditions, it proved necessary to protect the sulphur atom as either the sulphoxide or the sulphone.[129]

An alternative route for the same transformation which is applicable to both penicillins and cephalosporins has been developed. The C-6 anion, generated at low temperature from the aldimine (62), reacts with *N*-bromosuccinimide to give the 6-bromo-derivative (63) of unspecified stereochemistry. Addition of silver oxide and methanol afforded only the 6-α-methoxypenicillin (64).[130] The anions derived from the Schiff bases from



(62) X = H
(63) X = Br
(64) X = OMe
(65) X = SMe



(66) X = SMe

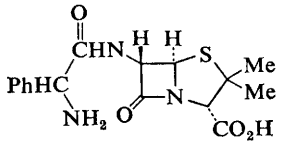
penicillinates and cephalosporanates also react with methoxycarbonylmethyl disulphide to yield the *S*-methyl compounds (65) and (66). The conversion of these Schiff bases into the free amines can be achieved with Girard's reagent and the amine acylated by standard procedures. The *S*-methyl derivatives may be converted into the corresponding α-methoxy-compounds on treatment with chlorine followed by triethylamine in methanol, again presumably through the intermediacy of an acylimine.[131]

[127] G. A. Koppel and R. E. Koehler, *J. Amer. Chem. Soc.*, 1973, **95**, 2403.
[128] G. A. Koppel and R. E. Koehler, *Tetrahedron Letters*, 1973, 1943.
[129] J. E. Baldwin, F. J. Urban, R. D. G. Cooper, and F. L. Jose, *J. Amer. Chem. Soc.*, 1973, **95**, 2401.
[130] L. D. Cama and B. G. Christensen, *Tetrahedron Letters*, 1973, 3505.
[131] W. A. Spitzer and T. Goodson, *Tetrahedron Letters*, 1973, 273.

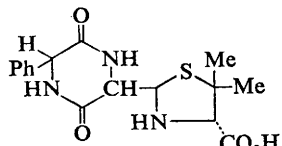
The C-6 (C-7) anions generated from these Schiff bases can be induced to undergo alkylation or aldol condensation. This has led to a variety of 6(7)-substituted derivatives all of which are assumed to be the α-isomer.[132]

As well as the considerable effort expended on producing C-6 substituted penicillins, interest has also been focused on C-6 epimerization. The direct base-catalysed epimerization of penicillins bearing a secondary amide side-chain is hampered by the preferential formation of the amide anion. This problem has been avoided by protecting the amide as the silyl imino-ether and subsequently epimerizing with a tertiary amine.[133] Alternatively, penicillin V methyl ester is epimerized at C-6 on treatment with two equivalents of di-isopropylamine followed by the quenching of the vicinal dianion with formic acid.[134]

6-Epi-ampicillin (67) in neutral aqueous conditions cyclizes to the dioxopiperazine (68) by intramolecular nucleophilic attack of the side-chain amino-group on the β-lactam carbonyl; under the same conditions ampicillin is stable.[135]



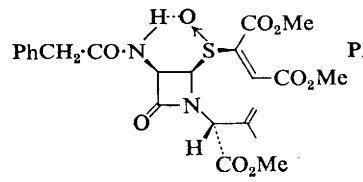
(67)



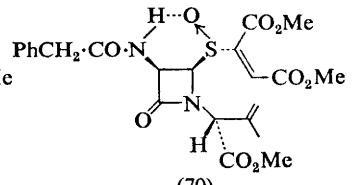
(68)

The rearrangements of penicillin sulphoxides still continue to be of importance. On heating, they form equilibrium mixtures with the corresponding sulphenic acids. These can be trapped intermolecularly with a variety of reagents: diketen, norbornadiene, and dimethyl acetylenedicarboxylate react by addition, whereas vinyl ethers react by displacement of the hydroxy-group.[136] A number of novel compounds have been obtained from the reaction of thermally produced sulphenic acids and acetylenic esters.[137] Penicillin G *S*-oxide methyl ester afforded, with dimethyl acetylenedicarboxylate, the isomeric sulphoxides (69) and (70) in high yield. On treatment with base only the isomer (69) cyclized to (71). Thiols react with sulphenic acids to give disulphides (72). The disulphide obtained by heating penicillin sulphoxides with 2-mercaptobenzothiazole has been used as a key intermediate in the conversion of penicillins into

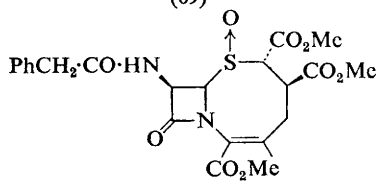
[132] G. H. Rasmusson, G. F. Reynolds, and G. E. Arth, *Tetrahedron Letters*, 1973, 145.
[133] P. Claes, A. Vlietinck, E. Roets, H. Vanderhaeghe, and S. Toppet, *J.C.S. Perkin I*, 1973, 932, 937.
[134] G. A. Koppel, *Tetrahedron Letters*, 1973, 4233.
[135] E. Roets, A. Vlietinck, G. A. Janssen, and H. Vanderhaeghe, *J.C.S. Chem. Comm.*, 1973, 484.
[136] I. Ager, D. H. R. Barton, D. G. T. Greig, G. Lucente, P. G. Sammes, M. V. Taylor, G. H. Hewitt, B. E. Looker, A. Mowatt, C. A. Robson, and W. G. E. Underwood, *J.C.S. Perkin I*, 1973, 1187.
[137] D. H. R. Barton, I. H. Coates, P. G. Sammes, and C. M. Cooper, *J.C.S. Chem. Comm.*, 1973, 303.



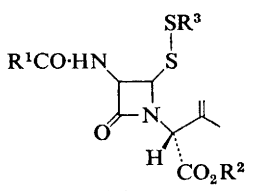
(69)



(70)



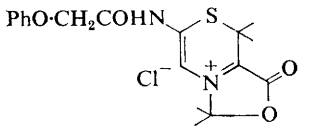
(71)



(72)

cephalosporins.[138] The compounds with the general structure (72) react readily with trialkylphosphites to produce the corresponding alkyl sulphides; the alkyl residue is derived from the phosphite reagent.[139]

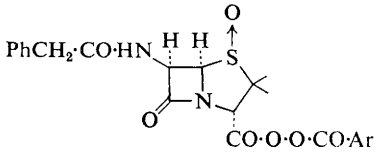
A further example of the well-established multiple rearrangement of phthalimidopenicillin 1-oxide with acetic anhydride has been described.[140] Penicillin V β-sulphoxide reacts with phenylacetyl chloride in acetone to give the unusual rearrangement product (73), the structure of which has been established by an *X*-ray crystallographic analysis.[141]



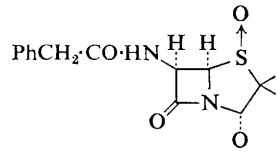
(73)

The advantages of using penicillin as a starting material for the synthesis of the basic azetidinone system continue to be explored. Rearrangement of the aroyl peroxide (74) affords the ester (75) in good yield. Hydrolysis of (75) to the corresponding alcohol (76) is achieved best with the *o*-nitrobenzoate ester using zinc dust and ammonium chloride. The possible value of this ester as a protecting group for alcohols and phenols has been noted.[142] This oxidative decarboxylation offers a convenient way of degrading the thiazolidine ring since the alcohol (76) is in equilibrium with the corresponding aldehyde. The aldehyde (77), prepared by an

[138] T. Kamiya, T. Teraji, Y. Saito, M. Hashimoto, O. Nakaguchi, and T. Oka, *Tetrahedron Letters*, 1973, 3001.
[139] R. D. Allan, D. H. R. Barton, M. Girijavallabhan, P. G. Sammes, and M. V. Taylor, *J.C.S. Perkin I*, 1973, 1182.
[140] D. O. Spry, *J.C.S. Chem. Comm.*, 1973, 259.
[141] R. Thomas and D. J. Williams, *J.C.S. Chem. Comm.*, 1973, 226.
[142] D. H. R. Barton, I. H. Coates, and P. G. Sammes, *J.C.S. Chem. Comm.*, 1973, 599.

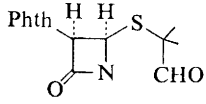


(74)

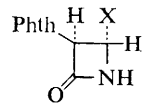


(75) R = COAr
(76) R = H

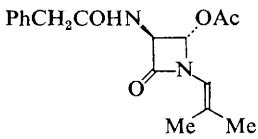
alternative method, is selectively degraded to the basic azetidinone (78) on treating with chlorine.[143] Various other degradative procedures for the modification of the thiazolidine ring of penicillins have been developed [144, 145] with the general aim of producing semi-synthetic antibiotics (see also ref. 2). Oxidation of penicillin with mercury(II) acetate leads directly to the monocyclic azetidinone (79).[146] It appears probable that the reaction is initiated by the cleavage of the 1,5-bond, since oxidation under the same conditions at room temperature afforded (80).[147]



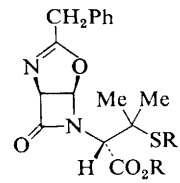
(77)



(78) X = Cl



(79)

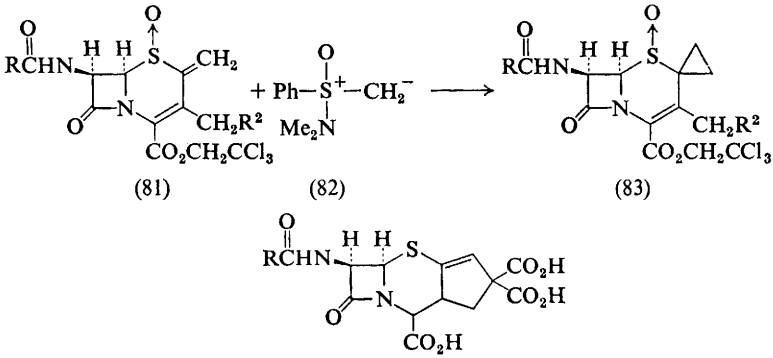


(80)

Structural modification of the cephalosporin molecule, other than substitution at C-7 which has already been described, has centred on the construction of tricyclic systems. The C-2 *exo*-methylene derivative (81), derived by a Mannich reaction on the corresponding sulphoxide, reacts with the sulphonium ylide (82) to give the C-2 spirocyclopropyl derivative (83) in high yield.[148] The tricyclic cephalosporin (84) has been synthesized by a standard Michael addition of malonate to (81) followed by cyclization.[149]

The cycloaddition of diazomethane to cephalosporins is now a well-documented process; further work related to cephalosporinates has been

[143] J. C. Sheehan, D. Ben-Ishai, and J. U. Piper, *J. Amer. Chem. Soc.*, 1973, **95**, 3064.
[144] S. Wolfe and P. Goeldner, *Tetrahedron Letters*, 1973, 5131.
[145] J. C. Sheehan and C. A. Panetta, *J. Org. Chem.*, 1973, **38**, 940.
[146] R. J. Stoodley and N. R. Whitehouse, *J.C.S. Perkin I*, 1973, 32.
[147] R. J. Stoodley and N. R. Whitehouse, *J.C.S. Chem. Comm.*, 1973, 477.
[148] D. O. Spry, *Tetrahedron Letters*, 1973, 2413.
[149] D. O. Spry, *J.C.S. Chem. Comm.*, 1973, 671.



(81) (82) (83)

(84)

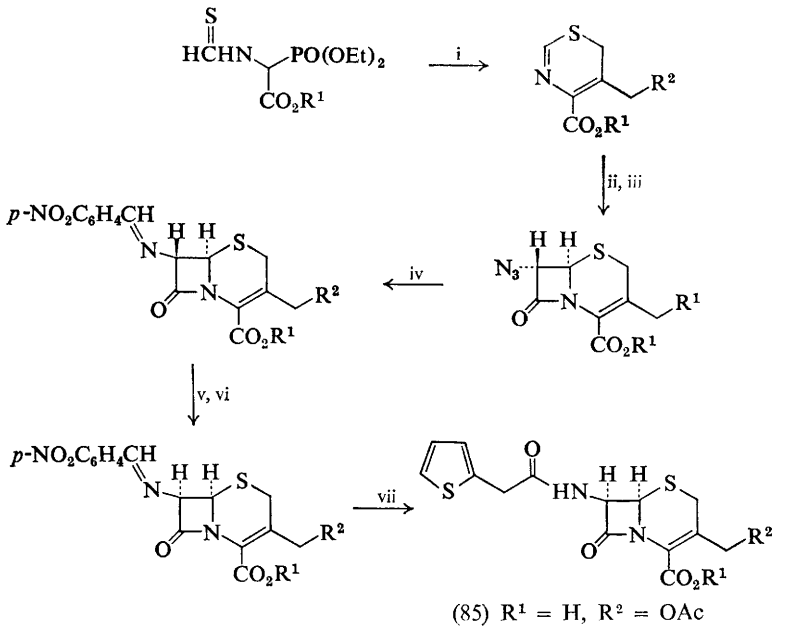
presented.[150] The elucidation of the stereochemistry at C-4 in the compounds resulting from the catalytic hydrogenation of a series of cepham derivatives has been reported.[151]

The total synthesis of β-lactam systems, as opposed to chemical modification of the naturally occurring compounds, is still an area of active endeavour. Undoubtedly one of the most outstanding synthetic achievements of the year has been the total synthesis of racemic cephalothin (85),[152, 153] details of which are summarized in Scheme 4. An extension[154] to the total synthesis of the 7-methoxycephalosporins uses a method already described.[130]

A new route for the synthesis of β-lactams employing a photolytically induced ring-contraction of 3-diazopyrrolidine-2,4-dione has been developed[155, 156] and is outlined in Scheme 5. The yields are acceptable and the method has been applied to simple bicyclic systems. Further nuclear analogues of penicillin and cephalosporin have been synthesized by the original diazo procedure,[157, 158] including the homopenicillin (86). The structure and stereochemistry of the intermediate (87) in this synthesis have been established by an *X*-ray analysis.[159] The pyruvyl derivatives (88) photocyclize to the 6-hydroxypenams (89) in moderate yield, affording another avenue to substituted azetidinones.[160]

An interesting structure–activity study which claims to correlate the biological activity of cephalosporin derivatives with the electron densities

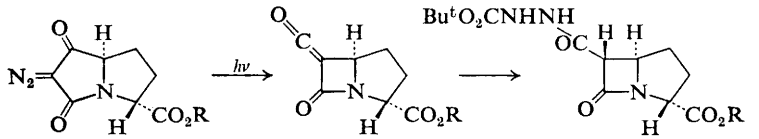
[150] E. R. Farkas, E. T. Gunda, and J. Cs. Jaszberenyi, *Tetrahedron Letters*, 1973, 5127.
[151] D. O. Spry, *Tetrahedron Letters*, 1973, 165.
[152] R. W. Ratcliffe and B. G. Christensen, *Tetrahedron Letters*, 1973, 4645.
[153] R. W. Ratcliffe and B. G. Christensen, *Tetrahedron Letters*, 1973, 4649.
[154] R. W. Ratcliffe and B. G. Christensen, *Tetrahedron Letters*, 1973, 4653.
[155] G. Lowe and D. D. Ridley, *J.C.S. Chem. Comm.*, 1973, 328.
[156] G. Lowe and H. W. Yeung, *J.C.S. Perkin I*, 1973, 2907.
[157] D. M. Brunwin and G. Lowe, *J.C.S. Perkin I*, 1973, 1321.
[158] G. Lowe and M. V. J. Ramsay, *J.C.S. Perkin I*, 1973, 479.
[159] K. Vijayan, B. F. Anderson, and D. C. Hodgkin, *J.C.S. Perkin I*, 1973, 484.
[160] K. R. Henrey-Logan and C. G. Chen, *Tetrahedron Letters*, 1973, 1103.



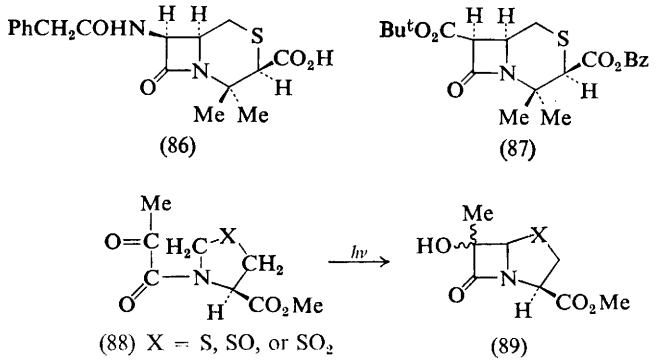
(85) R^1 = H, R^2 = OAc

Reagents: i, $R^2CH_2COCH_2Cl$; ii, N_3CH_2COCl–Et_3N; iii, Pt–H_2; iv, *p*-$NO_2C_6H_4CHO$; v, PhLi–HOAc; vi, 2,4-DNPH–TsOH; vii, 2-thienylacetyl chloride–pyridine

Scheme 4



Scheme 5

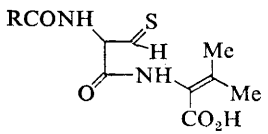


(86) (87)

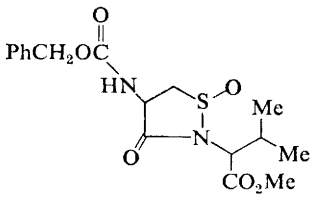
(88) X = S, SO, or SO_2 (89)

on the carbonyl carbon of the β-lactam and the bond strength indices of the carbonyl nitrogen bond has been reported.[161] The parameters were determined using two different MO methods which treat all valency electrons.

A number of attempts to achieve biomimetic syntheses of the penicillin skeleton from the supposed acyl bis-dehydrocysteinylvaline (90) have so



(90)



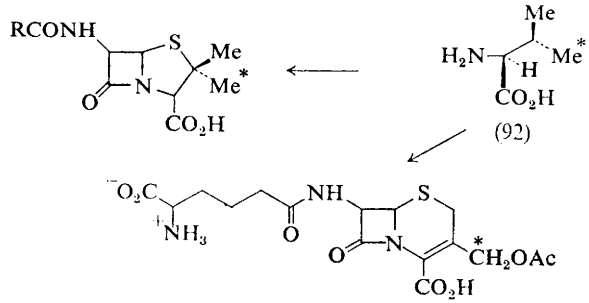
(91)

far been unsuccessful.[162–164] The isothiazolidone oxide (91) which represents a blocked form of (89) has been prepared and its chemistry extensively investigated.[162, 163] A further speculative mechanism of the formation of both penicillin and cephalosporin from (90) has been proposed.[164]

On the biosynthetic front a number of groups have independently reported on the synthesis of valine specifically labelled at the prochiral β-centre and its subsequent incorporation into penicillin and cephalosporin. Two groups have described [165, 166] the synthesis of the appropriately ^{13}C-labelled L-valine (92) and demonstrated that (2*S*,3*S*)-[4-^{13}C]valine is readily incorporated into both molecules. ^{13}C N.m.r. investigations clearly show that the label is located in the exocyclic methylene (C-17) in the cephalosporin molecule [166, 167] and in the α-methyl carbon of the thiazolidine ring in penicillin N [166] and penicillin V (Scheme 6).[167] Syntheses of valine prochirally labelled at the β-centre with deuterium have also been described [168, 169] in relation to the biosynthesis of the β-lactam antibiotics, but the feeding studies have not yet been reported.

Although the above-described investigations provide further important information concerning the biosynthesis, the major drawback to more esoteric studies with more complex precursors is hindered by the impermeability of the mycelium. In an attempt to overcome this problem,

[161] R. B. Hermann, *J. Antibiotics*, 1973, **26**, 223.
[162] R. B. Morin, E. M. Gordon, T. McGrath, and R. Shuman, *Tetrahedron Letters*, 1973, 2159.
[163] R. B. Morin, E. M. Gordon, and J. R. Lake, *Tetrahedron Letters*, 1973, 5213.
[164] J. E. Baldwin, S. B. Haber, and J. Hitchin, *J.C.S. Chem. Comm.*, 1973, 790.
[165] J. E. Baldwin, J. Loliger, W. Rastetter, N. Neuss, L. L. Huckstep, and N. De La Higuera, *J. Amer. Chem. Soc.*, 1973, **95**, 3796.
[166] H. Kluender, C. H. Bradley, C. J. Sih, P. Fawcett, and E. P. Abraham, *J. Amer. Chem. Soc.*, 1973, **95**, 6149.
[167] N. Neuss, C. H. Nash, J. E. Baldwin, P. A. Lemke, and J. B. Grutzner, *J. Amer. Chem. Soc.*, 1973, **95**, 3797.
[168] D. J. Aberhart and L. J. Lin, *J. Amer. Chem. Soc.*, 1973, **95**, 7859.
[169] R. K. Hill and S. Yan, *J. Amer. Chem. Soc.*, 1973, **95**, 7857.



Scheme 6

protoplasts have been prepared by the action of lytic enzymes on the mycelium of *Penicillium chrysogenum* and *Cephalosporium acremonium*.[170] Initial results with these systems are encouraging, but whether or not this approach will provide more definitive evidence must await further investigation.

6 Peptides Linked to Carbohydrates

This field covers a wide area of biological chemistry and the emphasis, as in previous years, is concentrated on the peptide constituents and in particular on those structures which contain unusual amino-acids or unusual peptide linkages. Other areas are covered by quoting appropriate review articles; a number of valuable general reviews on glycoproteins and glycopeptides have appeared during the year.[171–174]

Glycopeptides from Bacterial Cell Walls.—A comprehensive review on the biosynthesis of the bacterial cell wall covering the literature up to 1972 has been published.[175]

Considerable interest continues to be shown in the final steps in the biosynthesis of cell wall peptidoglycan. It is now well established that the terminal D-alanine residue present in the precursor is lost (Scheme 7).[176] The enzymes involved are membrane-bound in most organisms, but many *Streptomyces* excrete DD-carboxypeptidases. It was recently claimed that both these reactions are the properties of a single enzyme. This hypothesis has received further strong support from the observation that purified DD-carboxypeptidases from various strains of *Streptomyces* are also

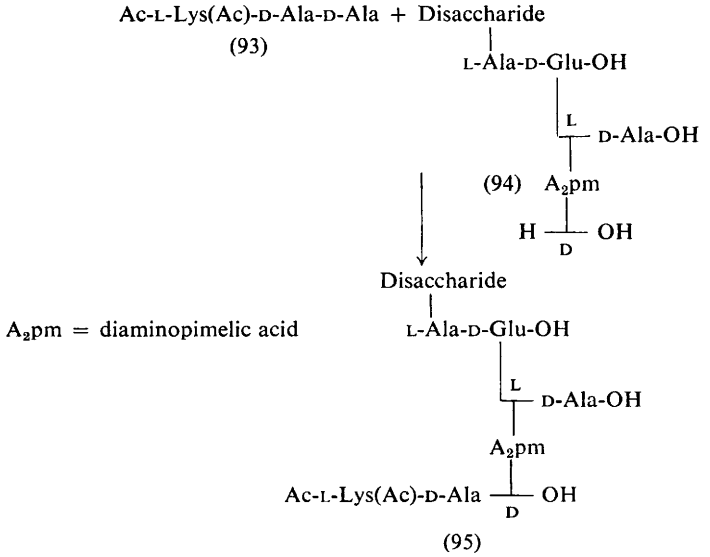
[170] P. A. Fawcett, P. B. Loder, M. J. Duncan, T. J. Beesley, and E. P. Abraham, *J. Gen. Microbiol.*, 1973, **79**, 293.
[171] 'Glycoproteins', Parts A and B, 2nd edition, ed. A. Gottschak, Elsevier, New York, 1972.
[172] R. Montgomery, MTP International Review of Science, Organic Chemistry Series One, ed. G. O. Aspinall, Butterworths, London, 1973.
[173] K. Heide and H. G. Schwick, *Angew. Chem.*, 1973, **85**, 803.
[174] P. J. Somers, *Carbohydrate Chem.*, 1970, **3**, 193.
[175] J. Baddiley, *Polm. Biol. Syst.*, CIBA Foundation Symposium, 1972, 87.
[176] H. R. Perkins and M. Nieto, *Pure Appl. Chem.*, 1973, **35**, 371.

$$R^1\text{-D-Ala-D-Ala} \xrightarrow{\text{DD-carboxypeptidase}} R^1\text{-D-Ala-OH} + \text{D-alanine}$$

$$\underset{\text{(donor)}}{R^1\text{-D-Ala-D-Ala-OH}} + \underset{\text{(acceptor)}}{NH_2R^2} \xrightarrow{\text{transpeptidase}} R^1\text{-D-Ala-D-Ala-NHR}^2 + \text{D-alanine}$$

Scheme 7

capable of performing transpeptidation reactions. An extensive *in vitro* study,[177–179] employing the synthetic peptide (93) as the donor, has shown that the substrate requirements of the enzyme for the peptide acceptor are similar to or identical with the natural peptide, *e.g.* (94) to (95).



The c.d. and fluorescence of one of these enzymes have been studied and a new theory concerning molecular mechanism of the interaction of the enzyme with penicillin was proposed.[180] The enzymes, which release D-alanyl-D-alanine from the substrate UDP–MurNAc–pentapeptide for *B. subtilis*[181] and *E. coli*,[182] have been purified and shown to be different from the enzyme which adds this dipeptide unit.

[177] J. M. Ghuysen, M. Leyh-Bouille, J. N. Campbell, R. Moreno, J. M. Frère, C. Duez, M. Nieto, and H. R. Perkins, *Biochemistry*, 1973, **13**, 1243.
[178] H. R. Perkins, M. Nieto, J. M. Frère, M. Leyh-Bouille, and J. M. Ghuysen, *Biochem. J.*, 1973, **131**, 707.
[179] J. M. Frère, J. M. Ghuysen, H. R. Perkins, and M. Nieto, *Biochem. J.*, 1973, **135**, 483.
[180] M. Nieto, H. R. Perkins, J. M. Frère, and J. M. Ghuysen, *Biochem. J.*, 1973, **135**, 493.
[181] A. Egans, P. Lawrence, and J. L. Strominger, *J. Biol Chem.*, 1973, **248**, 3122.
[182] B. Gondre, B. Flouret, and J. Van Heijenoort, *Biochimie*, 1973, **53**, 685.

A number of peptides which make up the partial structure of the peptide subunit (96) have been isolated from the autolysates of cell walls of *B. psychrophilus*. The cross-linkage between adjacent peptides is supplied by a second D-glutamic acid which is bound to the ε-amino-group of lysine

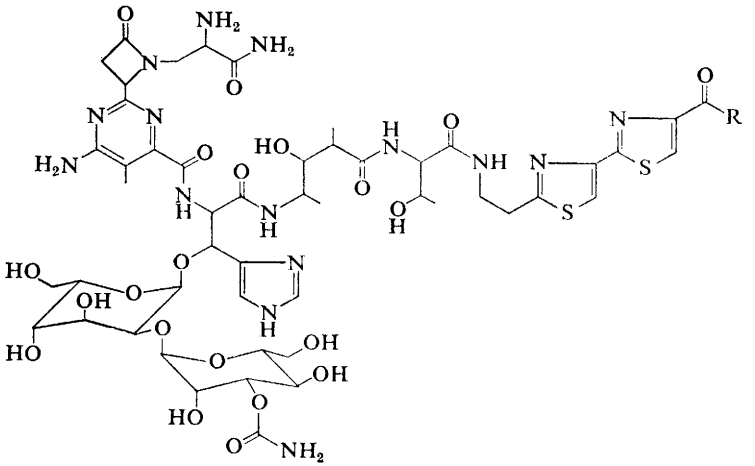
Muramyl-L-Ala-γ-D-Glu-L-Lys-D-Ala-OH
(96)

and the carboxyl of the D-alanine.[183] The primary structures of the peptidoglycan from various strains of *Peptococcus* have been investigated and four different types observed.[184] It is proposed that the occurrence of these different peptidoglycan types can be employed as a valuable criterion for classification. The peptidoglycans from a number of Gram-negative bacteria have all been shown to contain *meso*-diaminopimelic acid and belong to the same chemotype I, and it is suggested that all Gram-negative bacteria may be the same.[185] An investigation on the primary structure of the peptidoglycan from *Spirochaeta stenostrepta* has provided evidence for cross-linking of L-ornithine. The isolation of D-alanyl-N^{δ}-L-ornithine indicates direct cross-linking between the δ-amino-group of L-ornithine and the carboxy-group of the D-alanine of an adjacent subunit.[186]

The pentapeptide L-Ala-D-γ-Glu-L-Lys-D-Ala-D-Ala has been synthesized by classical methods and covalently linked to the random polypeptide $(Glu_{60}Ala_{40})_n$. The conjugate evoked an antibody response in rabbits and the antibody was absorbed specifically by a natural peptidoglycan from group C *Streptococcus*.[187]

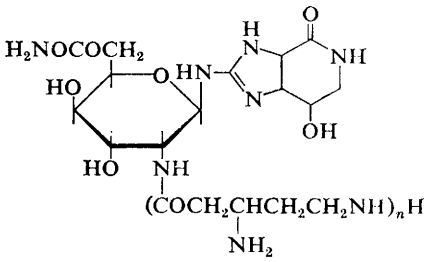
Glycopeptide Antibiotics.—The structures of the bleomycins, a group of closely related glycopeptide antitumour antibiotics, have been determined[188] after an extensive chemical and physicochemical investigation lasting several years. The general structure (97) illustrates that the major components differ only in the terminal amine entity. New bleomycins have been produced by providing the organism with specific amines which are incorporated into the molecule. Recently it has been shown that the amine group can be removed enzymatically to give the parent compound bleomycinic acid, from which semi-synthetic antibiotics have been prepared.[189] Perhaps the most notable feature of the molecule is the presence of the β-lactam ring. The evidence for this structural unit is largely circumstantial but if correct it represents a further example of this biologically important system.

[183] G. K. Best and S. J. Mattingly, *J. Bacteriol.*, 1973, **115**, 221.
[184] K. H. Scheiler and E. Nimmermann, *Arch. Mickrobiol.*, 1973, **93**, 245.
[185] J. P. Martin, J. Fleck, M. Mock, and J. M. Ghuysen, *European J. Biochem.*, 1973, **38**, 301.
[186] K. H. Schleifer and R. Joseph, *F.E.B.S. Letters*, 1973, **36**, 83.
[187] A. R. Zeiger and P. H. Maurer, *Biochemistry*, 1973, **13**, 3387.
[188] T. Takita, Y. Muraoka, T. Yoshioko, A. Fujii, K. Maeda, and H. Umezawa, *J. Antibiotics*, 1972, **25**, 755.
[189] T. Takita, A. Fujii, T. Fukuoka, and H. Umezawa, *J. Antibiotics*, 1973, **26**, 252.



(97) Bleomycin A_2 : R = $-NH-(CH_2)_3-S^+(CH_3)_2X^-$
Bleomycin B_2 : R = $-NH-(CH_2)_4-NH-C(=NH)-NH_2$

The streptothricin group (98) continues to be of interest and further investigations have confirmed the general structure for the whole group.[190, 191]



(98)

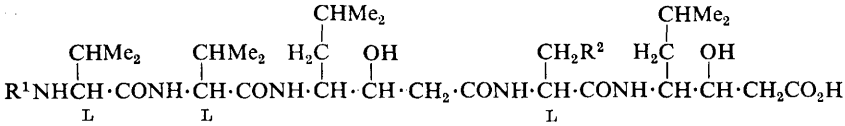
7 Other Peptides containing Unusual Structural Features

This section, as in previous years, is inevitably a miscellany, covering various naturally occurring peptides which in general contain anomalous amino-acid residues; in addition, a number of synthetic and conformational studies not conveniently covered under any other heading are included.

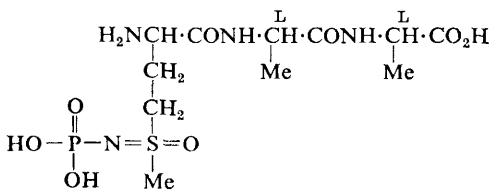
A number of new pepstatins which specifically inhibit acid proteases have been isolated from *Streptomyces parvisporogenes*. These compounds

[190] A. S. Khoklov and M. I. Shutova, *J. Antibiotics*, 1972, **25**, 501.
[191] K. Arima, T. Kuwamura, and T. Beppa, *J. Antibiotics*, 1972, **25**, 471.

(99)—(101) differ from those previously isolated in the nature of the acyl residue.[192] An additional pepstatin (102), isolated together with known pepstatins, contains L-serine in place of the alanine residue.[193] The novel phosphorus-containing peptide (103) possesses limited antibacterial



(99) R^1 = Pr, R^2 = H
(100) R^1 = Et, R^2 = H
(101) R^1 = Me, R^2 = H
(102) R^1 = Bu^i, R^2 = OH

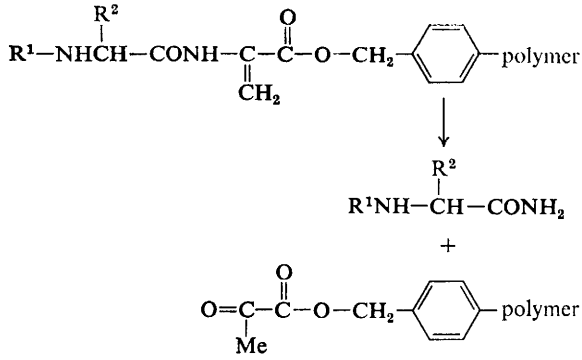


(103)

activity. Analysis of the physicochemical data and of the chemical and enzymatic degradation products led to the formulation (103).[194] The simple phenylalanine-containing peptides L-Arg-D-*allo*-Thr-L-Phe[195] and *N*-feruloyl-Gly-L-Phe[196] have been isolated from a fungus and barley, respectively. Full details of the *X*-ray crystallographic analysis on the chromophore of the pigment produced by *Azobacter vinelandii* have now been reported.[197]

The growing interest in dehydroamino-acid systems has now moved into synthesis and their potential synthetic utility has been successfully demonstrated. A considerable number of peptide hormones possess carboxyl-terminal amide groups and these can be prepared directly in a simple solid-phase procedure (Scheme 8). The peptide is built up on the resin in the normal manner and subsequently cleaved using dilute hydrochloric acid in acetic acid to give the amide and pyruvyl resin. As yet the method has only been used for the synthesis of the thyrotropin releasing factor.[198]

[192] T. Aoyagi, Y. Yagisawa, M. Kumagai, M. Hamada, H. Morishima, T. Takeuchi, and H. Umezawa, *J. Antibiotics*, 1973, **26**, 539.
[193] H. Umezawa, T. Miyano, T. Murakami, T. Takita, T. Aoyagi, T. Takeuchi, H. Naganawa, and H. Morishima, *J. Antibiotics*, 1973, **26**, 615.
[194] D. L. Pruess, J. P. Scanell, H. A. Ax, M. Kellett, F. Weiss, T. C. Demny, and A. Stempel, *J. Antibiotics*, 1973, **26**, 261.
[195] W. A. Koenig, W. Loeffler, W. H. Meyer, and R. Uhmann, *Chem. Ber.*, 1973, **106**, 816.
[196] C. F. van Sumere, H. De Pooter, H. Ali, and M. D. van Bussel, *Phytochemistry*, 1973, **12**, 407.
[197] K. Sasaki and Y. Hirata, *J.C.S. Perkin II*, 1973, 485.
[198] E. Gross, K. Noda, and B. Nisula, *Angew. Chem.*, 1973, **85**, 672.



Scheme 8

A number of other synthetic studies related to dehydroamino-acids have been described. Tripeptides containing a dehydroalanine residue have been cyclized to give cyclol type structures [199] and their spectroscopic properties investigated.[200] Peptides containing dehydrotryptophan have been synthesized, no doubt because of the naturally occurring peptides which possess this unit.[201] In addition, several dehydrophenylalanine-containing dipeptides have been made and extensive c.d. and o.r.d. studies undertaken.[202] The results so far obtained are not particularly enlightening concerning the conformation of these compounds. However, a very extensive and detailed investigation of simple dipeptides has appeared.[203]

Other areas which may be of general interest are the syntheses of peptides containing unusual amino-acids,[204] *N*-methylamino-acids,[205, 206] and α-benzylphenylalanine.[207]

[199] G. Lucente, A. Romeo, and G. Zanotti, *Gazzetta*, 1972, **102**, 941.
[200] F. Conti, G. Lucente, A. Romeo, and G. Zanotti, *Internat. J. Peptide Protein Res.*, 1973, **5**, 353.
[201] M. Bakhra, G. S. Katrakha, and A. B. Silaev, *Khim. prirod. Soedinenii*, 1973, 280.
[202] O. Pieroni, A. Fissi, and G. Montagnoli, *Biopolymers*, 1973, **12**, 1445.
[203] A. W. Burgess and H. A. Scheraga, *Biopolymers*, 1973, **12**, 2177.
[204] L. N. Veselova and E. S. Chaman, *Zhur. obshchei Khim.*, 1973, **43**, 1637.
[205] J. R. McDermott and N. L. Benoiton, *Canad. J. Chem.*, 1973, **51**, 2562.
[206] H. Sugano, K. Higaki, and M. Miyoshi, *Bull. Chem. Soc. Japan*, 1973, **46**, 231.
[207] H. L. Maia, B. Ridge, and H. N. Rydon, *J.C.S. Perkin I*, 1973, 98.

5

Chemical Structure and Biological Activity

BY R. CAMBLE AND A. S. DUTTA

1 Introduction

This chapter is organized in a similar manner to last year's,[1] and once again Current Contents Life Sciences (Institute for Scientific Information) has been used as the most convenient means of monitoring literature, scattered over many publications, which is directed towards correlation of the chemical structure of peptides with their biological activity. Several areas, such as relationships between the structural determinants for immunological and biological activities, the structural requirements for substrates and inhibitors (both natural and synthetic) of many proteolytic enzymes, and the structural demands for peptide transport through cell membranes as distinct from their interactions with receptors at the membrane surface, have been reluctantly excluded.

The complexity of the biological response elicited by a peptide hormone and the problems associated with its measurement and interpretation were discussed in the introduction to the last Report.[1] Oxytocin has been used as an example to illustrate convincingly the way in which synthetic analogues can be exploited to analyse the different stages of hormone action.[2] Even chemical structure becomes a complex dynamic quantity once a peptide is introduced into a biological environment. In the past few years we have grown accustomed to the idea of long-chain, relatively inactive, biosynthetic precursor 'prohormones', but now there are indications that several familiar hormone sequences themselves may be prohormones or may need to be enzymatically degraded before they interact most effectively with their target cell receptors. Could it be that a major milestone in many structure–activity investigations, *i.e.* the discovery of an active fragment, may have been reached by mimicking the biological situation rather than by outwitting it? The importance of the balance between activation and deactivation of peptide molecules in the control of hormonal activity has been considered in recent reviews.[3, 4a]

[1] H. D. Law in 'Amino-acids, Peptides, and Proteins', ed. R. C. Sheppard (Specialist Periodical Reports), The Chemical Society, London, Vol. 5, 1974, p. 384.
[2] J. Rudinger, V. Pliska, and I. Krejci, *Recent Progr. Hormone Res.*, 1972, **28**, 131.
[3] E. B. Knights, S. B. Baylin, and G. V. Foster, *Lancet*, 1973, 719.
[4] Proceedings 12th European Peptide Symposium, Reinhardsbrunn Castle, German Democratic Republic, 1972, ed. H. Hanson and H.-D. Jakubke, North Holland Publishing Co., Amsterdam, 1973: (*a*) R. Walter, p. 363; (*b*) R. Schwyzer, p. 424; (*c*) O. A. Kaurov, V. F. Martynov, Yu. D. Mihaylov, O. A. Popernaczky, and M.-P. Smirnova, p. 450; (*d*) R. Camble, R. Cotton, A. S. Dutta, J. J. Gormley, C. F. Hayward, J. S. Morley, and M. J. Smithers, p. 200; (*e*) J. T. Potts, jun., G. W. Tregear, J. Van Rietschoten, H. D. Niall, and H. T. Keutmann, p. 191; (*f*) L. Moroder, G. Borin, F. Marchiori, and E. Scoffone, p. 223.

2 Hypothalamic Regulatory Factors

Several reviews have appeared on the hypothalamic regulatory hormones.[5–10] No new releasing factors have been reported but the growth hormone release-inhibiting peptide, somatostatin, reported for the first time at the end of 1972, has now been evaluated in a number of laboratories. The biology of luteinizing hormone-releasing factor (LRF) and of thyrotropin-releasing factor (TRF) has been investigated in detail and much synthetic effort has been concentrated on LRF. Involvement of catecholamines in the secretion of hypothalamic factors has also been studied.[11–14]

Thyroid Stimulating Hormone-releasing Factor (TRF).—Many studies have been devoted to the effects of TRF in man. The peptide elicits a rise in plasma TSH and prolactin levels which is always greater in men than in women; the levels of LH and FSH are unaffected.[15] Serum growth hormone levels increase in women but not in men.[15] In pregnant women TRF has no effect on the levels of placental lactogen and chorionic thyrotropin, indicating that the placental hormones are controlled by a different mechanism.[16] When TRF (50—400 μg) is given to healthy volunteers by rapid intravenous injection there is a dose-related TSH response.[17]

The rise in TSH levels can be detected within 10 minutes and reaches a peak concentration after 15—30 minutes depending upon the dose. Peak levels of TSH after subcutaneous injection are obtained only after 30—60 minutes.[18] A more prolonged rise in serum TSH levels is observed after oral administration of TRF than after injection, but the results are variable.[17] Patients with primary hyperthyroidism or thyrotoxicosis show no increase in serum TSH levels but patients with primary hypothyroidism give an exaggerated and prolonged increase.[15] Most patients with tumorous or non-tumorous pituitary disorders show a positive response.[18, 19] In patients with mild hypothalamic hypothyroidism or an intractable depressive syndrome, TRF increases serum cholesterol levels but the effect is not dose-dependent.[20] TRF also causes a rapid release of TSH from human

[5] K. Folkers, N.-G. Johansson, F. Hooper, B. Currie, H. Sievertsson, J.-K. Chang, and C. Y. Bowers, *Angew. Chem. Internat. Edn.*, 1973, **12**, 255.
[6] A. V. Schally, A. Arimura, and A. J. Kastin, *Science*, 1973, **179**, 341.
[7] R. Blackwell and R. Guillemin, *Ann. Rev. Physiol.*, 1973, **35**, 357.
[8] V. L. Gay, *Fertility and Sterility*, 1972, **23**, 50.
[9] A. V. Schally, A. J. Kastin, and A. Arimura, *Amer. J. Obstet. Gynecol.*, 1972, **114**, 423.
[10] G. W. Harris, *J. Endocrinol.*, 1972, **53**, ii.
[11] Y. Grimm and S. Reichlin, *Endocrinology*, 1973, **93**, 626.
[12] S. P. Kalra and S. M. McCann, *Endocrinology*, 1973, **93**, 356.
[13] K.-H. Lu and J. Meites, *Endocrinology*, 1972, **91**, 868.
[14] A. P. Labhsetwar, *J. Reprod. Fert.*, 1973, **33**, 545.
[15] P. A. Torjesen, E. Haug, and T. Sand, *Acta Endocrinologica*, 1973, **73**, 455.
[16] J. M. Hershman, A. Kojima, and H. G. Friesen, *J. Clin. Endocrinol. Metab.*, 1973, **36**, 497.
[17] *Brit. Med. J.*, 1973, **3**, 465.
[18] M. Otsuki, H. Mori, S. Baba, and N. Hiroshige, *Acta Endocrinologica*, 1973, **73**, 233.
[19] G. Faglia, P. Beck-peccoz, C. Ferrari, B. Ambrosi, A. Spada, P. Travaglini, and S. Paracchi, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 595.
[20] M. J. E. Van Der Vis-Melsen and J. D. Wiener, *Brit. Med. J.*, 1973, **4**, 419.

anterior pituitary tissue *in vitro* and this release is completely inhibited by thyroxine or tri-iodothyronine.[21]

A single intravenous injection of TRF significantly increases serum prolactin levels in pro-oestrous female rats and in normal and oestrogen-primed male rats.[22] Prolactin release is also markedly stimulated by TRF in *in vitro* experiments with rat pituitary cells.[23] TRF increases the rate of synthesis of prolactin by cultured GH3 cells, a clonal strain of prolactin- and growth-hormone-secreting cells from rat pituitary.[24, 25] Binding of TRF to the GH3 cells is time-dependent, increasing linearly for 15 minutes and reaching a plateau in 60 minutes;[26] this is similar to the pattern of prolactin release.

In normal men and women, although the minimum effective dose of TRF is essentially the same for both TSH and prolactin release,[27] the prolactin response is more sensitive (6.25 μg of TRF gave 300% higher prolactin levels and only 50% higher TSH levels).[28] In addition, maximum prolactin release occurs at a lower dose of TRF (100 μg) than maximum TSH release (400 μg). These findings indicate that TRF is a prolactin-releasing hormone of physiological significance, but the existence of a separate prolactin-releasing factor has not been excluded.

When given intravenously, TRF exerts an immediate anti-depressant action. In normal women TRF first produces a sense of relaxation and later a mild euphoria, accompanied by an increase in energy and a sense of mental clarity; in schizophrenia patients, the peptide has a rapid overall beneficial effect.[29] Treatment of depression with electroconvulsive therapy is not thought to involve the release of TRF.[30] The anti-depressant activity of TRF was also studied in conscious dogs but in the test system employed (potentiation of the behavioural and autonomic effects of yohimbine) the peptide had no effect.[31]

Very few new analogues of TRF have been reported this year (Table 1). The stereoisomers (1)—(4) were synthesized in an attempt to increase the resistance of TRF to enzymic degradation and so give analogues with prolonged duration of action.[32] Unfortunately, (1)—(3) possessed very little activity and (4) was inactive; (4) also failed to inhibit the binding of tritiated TRF to anterior pituitary homogenates; (3) inhibited binding but

[21] P. B. May and R. K. Donabedian, *J. Clin. Endocrinol. Metab.*, 1973, **36**, 605.
[22] G. P. Mueller, H. J. Chen, and J. Meites, *Proc. Soc. Exp. Biol. Med.*, 1973, **144**, 613.
[23] P. S. Dannies and A. H. Tashjian, jun., *J. Biol. Chem.*, 1973, **248**, 6174.
[24] P. M. Hinkle and A. H. Tashjian, jun., *J. Biol. Chem.*, 1973, **248**, 6180.
[25] D. Gourdji, A. Tixier-Vidal, A. Morin, P. Pradelles, J. L. Morgat, P. Fromageot, and B. Kerdelhue, *Exp. Cell. Res.*, 1973, **82**, 39.
[26] W. Vale, R. Blackwell, G. Grant, and R. Guillemin, *Endocrinology*, 1973, **93**, 26.
[27] C. Y. Bowers, H. G. Friesen, and K. Folkers, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 512.
[28] L. S. Jacobs, P. J. Snyder, R. D. Utiger, and W. H. Daughaday, *J. Clin. Endocrinol. Metab.*, 1973, **36**, 1069.
[29] I. C. Wilson, P. P. Lara, and A. J. Prange, jun., *Lancet*, 1973, 43.
[30] J. I. Thorell and G. Adielsson, *Lancet*, 1973, 43.
[31] B. Hine, I. Sanghvi, and S. Gershon, *Life Sci.*, 1973, **13**, 1789.
[32] G. Flouret, R. Morgan, and R. Gendrich, *J. Medicin. Chem.*, 1973, **16**, 1137.

Table 1 *Synthetic analogues of TRF*

Compound number	*Structure*	TSH-*releasing activity*/% [a]	*Ref.*
(1)	Glp-D-His-Pro-NH_2 [b]	2—3	32
(2)	D-Glp-His-Pro-NH_2	0.1	32
(3)	Glp-His-D-Pro-NH_2	0.1	32
(4)	D-Glp-D-His-D-Pro-NH_2	inactive	32
(5)	Glp-Gly-Pro-OCH_3	inactive	34
(6)	Glp-Gly-Pro-NH_2	inactive	34
(7)	Glp-Leu-Pro-OCH_3	inactive	34
(8)	Glp-Leu-Pro-NH_2	0.2	34
(9)	Glp-Arg-Pro-OCH_3	inactive	34
(10)	Glp-Arg(NO_2)-Pro-OCH_3	inactive	34
(11)	Glp-Thi-Pro-OCH_3 [c]	inactive	34
(12)	Glp-Thi-Pro-NH_2	0.2	34
(13)	Glp-D-Thi-Pro-OCH_3	inactive	34
(14)	Glp-D-Thi-Pro-NH_2	0.1	34
(15)	Glp-His-pyrrolidine	inactive	35
(16)	Cpc-His-Pro-NH_2 [d]	inactive	35
(17)	Pro-His-Pro-NH_2	inactive	35
(18)	Cpc-His-pyrrolidine	inactive	35

[a] Expressed relative to TRF as 100%; [b] Glp = pyroglutamic acid; [c] Thi = β-2-thienyl-L-alanine; [d] Cpc = cyclopentylcarbonyl.

with only about 2% of the potency of TRF.[33] The histidine residue in TRF can be replaced by leucine (8) without eliminating all activity, but its replacement with glycine gives an inactive analogue (6). This indicates that neither the basic nor the aromatic properties of the imidazole group are essential for activity, but the side-chain cannot be eliminated completely.[34] Compounds (8), (12) and (14) also increased the serum levels of prolactin in monkeys. None of the analogues (5)—(14) inhibited the *in vivo* activity of TRF in mice or monkeys;[34] in another study,[35] compound (18) proved to be an effective *in vivo* antagonist of the effects of TRF.

The conformation of TRF has been studied in detail by several groups but the results are still not in agreement. Blagdon *et al.* have suggested a 'hairpin turn' conformation on the basis of semi-empirical energy calculations and n.m.r. studies.[36] This conformation is stabilized by two specific hydrogen bonds. The presence of one, that between the α-NH of the histidine residue and the π nitrogen of the imidazole ring, is supported by potentiometric titration data[37] but not by spin–lattice relaxation-time measurements.[38] The other proposed hydrogen bond, between the back-

[33] J. F. Wilber and M. J. Seibel, *Endocrinology*, 1973, **92**, 888.
[34] H. Sievertsson, S. Cartensson, C. Y. Bowers, H. G. Friesen, and K. Folkers, *Acta Pharm. Suecica*, 1973, **10**, 297.
[35] H. Lybeck, J. Leppaluoto, P. Virkkunen, D. Schafer, L. Carlsson, and J. Mulder, *Neuroendocrinology*, 1973, **12**, 366.
[36] D. E. Blagdon, J. Rivier, and M. Goodman, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1166.
[37] G. Grant, N. Ling, J. Rivier, and W. Vale, *Biochemistry*, 1972, **11**, 3070.
[38] R. Deslauriers, C. Garrigou-Lagrange, A. Bellocq, and I. C. P. Smith, *F.E.B.S. Letters*, 1973, **31**, 59.

bone carbonyl group of the pyroglutamic acid residue and the *C*-terminal amide, is similar to one suggested by Deslauriers *et al.*[39] in a β-turn model analogous to that proposed for Pro-Leu-Gly-NH_2. Other conformational energy calculations indicate that the central histidine residue of the TRF molecule is in an extended conformation and that the backbone retains considerable rigidity except for the pyrrolidone ring, which can rotate relatively freely.[40] Spin–lattice relaxation-time measurements also suggest some conformational flexibility in the ring of the pyroglutamic acid residue.[38] Similar energy calculations on some TRF analogues show that [2-phenylalanine]-TRF, [2-pyrazolyl-3′-alanine]-TRF, [2-N^{τ}-methylhistidine]-TRF and Glp-His-Pro-NHMe may have conformations almost identical with that of TRF. The low activity of [2-phenylalanine]-TRF (10%) and [2-pyrazolyl-3′-alanine]-TRF (5%) is therefore ascribed to poor affinity of the receptor site for phenyl or pyrazolyl rings [40] (note, however, that in a pituitary depletion assay [2-pyrazolyl-3′-alanine]-TRF has been found to have the same activity as TRF [41]). The most stable conformation of [2-N^{π}-methylhistidine]-TRF apparently has different side-chain orientations. TRF and [3-glycine]-TRF may have similar backbone conformations, but in the latter the imidazole ring is probably twisted into a different position by an intramolecular hydrogen bond between the N^{π}-imidazole and the NH of the glycine. Formation of this type of bond is also possible when the proline residue in position 3 is replaced by other amino-acids and probably accounts for the inactivity of most TRF analogues modified in position 3.[40]

Tritiated TRF competes with TRF for binding sites in pituitary, thyroid, and hypothalamic tissue homogenates.[42] The binding of tritiated TRF to an anterior pituitary extract is four times greater than to the corresponding fractions of liver, renal cortex, and myocardium, and is also much more specific.[33] This is indicated by the fact that binding of labelled TRF to anterior pituitary homogenates is inhibited significantly by a one-thousand-fold excess of unlabelled TRF whereas in other tissues no inhibiton is observed.[33] [1-Proline]-TRF, which induces a rapid accumulation of TSH in the pituitary but no release, has a high affinity for the receptors in the pituitary but not for those in the thyroid.[42] In bovine anterior pituitary extracts the TRF receptors are localized in the plasma membrane fraction.[43, 44] Tritiated TRF has a similar affinity (*ca.* 4×10^{-8} mol l^{-1}) for non-cloned TSH-secreting tumour cells and cloned GH3 tumour cells,

[39] R. Deslauriers, R. Walter, and I. C. P. Smith, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 244.
[40] A. W. Burgess, F. A. Momany, and H. A. Scheraga, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1456.
[41] H. Steiner, F. Piva, G. Gavazzi, R. O. Studer, D. Gillessen, and L. Martini, *Hormone Metab. Res.*, 1972, **4**, 484.
[42] H. Steiner, *Experientia*, 1973, **29**, 759.
[43] L. J. Eddy, J. M. Hershman, R. E. Taylor, jun., and S. B. Barker, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 140.
[44] N. Barden and F. Labrie, *J. Biol. Chem.*, 1973, **248**, 7601.

suggesting that thyrotropin- and prolactin-secreting cells may have very similar receptor sites.[26] This is supported by the finding that [1-proline]-TRF, [2-N^{π}-methylhistidine]-TRF, [2-N^{τ}-methylhistidine]-TRF, [1-*N*-methylpyroglutamic acid]-TRF and TRF free acid have the same relative binding affinities for TSH- and prolactin-secreting cells.[26] Pituitary thyrotrophs appear to have two receptor sites;[45] one with an affinity constant of 2×10^{-8} mol l^{-1}, and the other with an affinity constant of 5×10^{-7} mol l^{-1}. The high-affinity site is saturated at a TRF concentration ten-fold greater than that required to obtain maximal biological activity, possibly indicating the presence of a large number of spare TRF receptors. The assumption that these specific pituitary binding sites are related to the biologically important receptors is supported by the correlation which has been found[45] (Table 2) between the relative binding

Table 2 *Comparison of biological activity and affinity constants of TRF analogues (from ref.* 45)

Compound number	*Structure*	*Biological activity*/%	*Relative affinity constant* [a]/mol l^{-1}
(19)	Glp-His-Pro-NH_2(TRF)	100	2.0×10^{-8}
(20)	Glp-N^{τ}-MeHis-Pro-NH_2	800	3.0×10^{-9}
(21)	Glp-His-Pro-Gly-NH_2	35	6.0×10^{-8}
(22)	Glp-His-Pro-$NHCH_2CH_2OH$	16	2.0×10^{-7}
(23)	Glp-His-Pro-$NHCH_2CH_3$	14	6.0×10^{-7}
(24)	Glp-His-Pro-OCH_3	10	8.0×10^{-7}
(25)	MeGlp-His-Pro-NH_2 [b]	1.7	3.0×10^{-6}
(26)	Glp-His-Prolinol	1.2	3.2×10^{-6}
(27)	Glp-Met-Pro-NH_2	1.0	5.0×10^{-6}
(15)	Glp-His-pyrrolidine	0.8	1.2×10^{-6}
(28)	Glp-His-Pro-OC_2H_5	0.4	1.0×10^{-6}
(29)	Glp-His-hexamethylenimine	0.1	4.0×10^{-5}
(30)	Glp-His-morpholine	0.3	1.2×10^{-5}

[a] Concentration of analogue required to compete to the extent of 50% for saturating amounts of bound tritiated TRF; [b] MeGlp = *N*-methylpyroglutamic acid.

affinities of several TRF analogues and their relative biological potencies.[46, 47]

Luteinizing Hormone-releasing Factor (LRF).—The responses to LRF in diverse species under a variety of conditions have been the subject of intense study and ever more detailed measurement. Only a brief summary of these results, with particular emphasis on human studies, can be included here. The ability of LRF to stimulate the release of both LH and FSH has been demonstrated in animals,[48] in organ cultures of human foetal

[45] G. Grant, W. Vale, and R. Guillemin, *Endocrinology*, 1973, **92**, 1629.
[46] R. Guillemin, R. Burgus, and W. Vale, *Vitamins and Hormones*, 1971, **29**, 1.
[47] W. Vale, G. Grant, and R. Guillemin, *Frontiers in Neuroendocrinology*, 1973, 1.
[48] R. Blackwell, M. Amoss, jun., W. Vale, R. Burgus, J. Rivier, M. Monahan, N. Ling, and R. Guillemin, *Amer. J. Physiol.*, 1973, **224**, 170.

pituitaries,[49] in normal men and women,[50] and in patients with various forms of mental and testicular disorders.[51] The peptide has no effect on the secretion of prolactin;[52] decreases protein synthesis in the anterior hypothalamus;[53] increases protein synthesis in the anterior pituitary;[53] and induces mating behaviour in the female rat.[54] In women, LRF is least effective in increasing plasma LH levels during the early and mid-follicular phase and is most effective around the spontaneous LH surge;[55] the amount of LH released is always higher than the amount of FSH.

There is some disagreement about the minimum dose of LRF required to release LH or FSH in rats and in humans. Blackwell *et al.*[48] have shown in rats that although LH and FSH are released simultaneously, the amount of FSH is very small at lower doses. Ondo *et al.*[56] report that only LH is released at low concentrations (0.015 μmol l^{-1}) and a concentration of 0.15—15 μmol l^{-1} is required to release both LH and FSH. In humans, Rebar *et al.*[57] have observed the release of LH with a dose of 10 μg of LRF and the release of FSH only with 150 μg, whereas Hang and Torjesen [58] have detected release of FSH with 12.5 μg of LRF and of LH with 25 μg. It has been suggested that the threshold for the initiation of FSH release is lower than that of LH, although greater concentrations of LH are eventually achieved.[59] It has not been established if LRF is the only hormone controlling the release of both LH and FSH. Biosynthetic studies with [^{14}C]glutamic acid and [^{14}C]glutamine in a hypothalamic system have indicated the existence of a separate FSH-releasing factor.[60] This has been partially purified from porcine hypothalami and has been freed from LRF decapeptide. The purified fractions release both LH and FSH, but the amount of FSH released is 40—128 μg ml^{-1}.[61, 62] LRF under similar conditions releases only about 18 and never more than 35 μg ml^{-1} of FSH.

[49] G. V. Groom and A. R. Boyns, *J. Endocrinol.*, 1973, **59**, 511.
[50] K. Thomas, J. Donnez, and J. Ferin, *Contraception*, 1972, **6**, 55.
[51] K. Isurugi, K. Wakabayashi, K. Fukutani, H. Takayasu, B. Tamaoki, and M. Okada, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 533.
[52] R. Blackwell, W. Vale, M. Amoss, R. Burgus, M. Monahan, J. Rivier, N. Ling, and R. Guillemin, *Amer. J. Physiol.*, 1973, **224**, 176.
[53] J. A. Moguilevsky and J. Christot, *Proc. Soc. Exp. Biol. Med.*, 1973, **143**, 260.
[54] R. L. Moss and S. M. McCann, *Science*, 1973, **181**, 177.
[55] K. Thomas, M. Cardon, J. Donnez, and J. Ferin, *Contraception*, 1973, **7**, 289.
[56] J. G. Ondo, R. L. Eskay, R. S. Mical, and J. C. Porter, *Endocrinology*, 1973, **93**, 205.
[57] R. Rebar, S. S. C. Yen, G. Vandenberg, F. Naftolin, Y. Ehara, S. Engblom, K. J. Ryan, J. Rivier, M. Amoss, and R. Guillemin, *J. Clin. Endocrinol. Metab.*, 1973, **36**, 10.
[58] E. Haug and P. Torjesen, *Acta Endocrinologica*, 1973, **73**, 465.
[59] C. H. Mortimer, G. M. Besser, D. J. Goldie, J. Hook, and A. S. McNeilly, *Nature New Biol.*, 1973, **246**, 22.
[60] K. N. G. Johansson, B. L. Currie, K. Folkers, and C. Y. Bowers, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 8.
[61] B. L. Currie, K. N. G. Johansson, K. Folkers, and C. Y. Bowers, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 14.
[62] C. Y. Bowers, B. L. Currie, K. N. G. Johansson, and K. Folkers, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 20.

The route of administration of LRF in humans affects the timecourse of the response. After rapid intravenous injection maximum LH levels are reached in 16—32 min.[63] High levels are generally maintained for 60—90 min and normal levels are restored after 3—6 h. Intranasal administration of LRF also produces maximum LH levels after about 30 min but the dose required is much greater than by intravenous injection.[64] Subcutaneous and intramuscular injections produce maximum LH levels in *ca.* 64 min.[63, 65] Intramuscular injection is the most effective route with regard to increment and duration of LH secretion;[66] sublingual administration does not release sufficient LH to be of practical use.[66] The half-life of LRF is about 7 min in normal female rats,[67] and 2—4 min in man.[68, 69] In work with tritiated LRF most of the radioactivity was found in urine within a few hours. The radioactivity was associated with pyroglutamic acid and the dipeptide pyroglutamyl-histidine, indicating that one mechanism for inactivation of LRF is cleavage of the *N*-terminal pyroglutamyl-histidine moiety.[68]

N.m.r. studies with LRF suggest there is no appreciable stacking of the aromatic rings, no electrostatic intramolecular interactions involving the charged imidazole ring, and no special structure-forming intramolecular hydrogen bonds.[70] ^{13}C N.m.r. spectroscopy indicates that the arginyl–proline peptide bond is in the *trans* conformation.[71] C.d. studies also suggest a random conformation in acidic medium but there is a shift towards ordered structures when neutral aqueous solutions are heated or when trifluoroethanol is added.[72]

The analogues of LRF (31) reported this year are listed in Table 3. Elimination of the *C*-terminal glycine residue gives an analogue (32) with about 10% biological activity.[73, 74] The dose–response curves for this analogue are parallel to those of LRF, with similar maximum responses,

[63] D. Gonzalez-Barcena, A. J. Kastin, D. S. Schalch, J. A. Bermudez, D. Lee, A. Arimura, J. Ruelas, I. Zepeda, and A. V. Schally, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 481.
[64] D. R. London, W. R. Butt, S. S. Lynch, J. C. Marshall, S. Owusu, W. R. Robinson, and J. M. Stephenson, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 829.
[65] A. Arimura, M. Saito, Y. Yaoi, T. Kumasaka, H. Sato, T. Koyama, N. Nishi, A. J. Kastin, and A. V. Schally, *J. Clin. Endocrinol. Metab.*, 1973, **36**, 385.
[66] P. J. Keller and C. Gerber, *Hormone Metab. Res.*, 1973, **5**, 64.
[67] T. W. Redding and A. V. Schally, *Life Sci.*, part I, 1973, **12**, 23.
[68] T. W. Redding, A. J. Kastin, D. Gonzalez-Barcena, D. H. Coy, E. J. Coy, D. S. Schalch, and A. V. Schally, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 626.
[69] Y. Miyachi, R. S. Mecklenburg, J. W. Hansen, and M. B. Lipsett, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 63.
[70] P. L. Wessels, J. Feeney, H. Gregory, and J. J. Gormley, *J.C.S. Perkin II*, 1973, 1691.
[71] R. Deslauriers, R. Walter, and I. C. P. Smith, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 244.
[72] P. Marche, J. Morgat, and P. Fromageot, *European J. Biochem.*, 1973, **40**, 513.
[73] M. Fujino, S. Shinagawa, M. Obayashi, S. Kobayashi, T. Fukuda, I. Yamazaki, R. Nakayama, W. F. White, and R. H. Rippel, *J. Medicin. Chem.*, 1973, **16**, 1144; R. H. Rippel, E. S. Johnson, W. F. White, M. Fujino, I. Yamazaki, and R. Nakayama, *Endocrinology*, 1973, **93**, 1449.
[74] J. Rivier, W. Vale, R. Burgus, N. Ling, M. Amoss, R. Blackwell, and R. Guillemin, *J. Medicin. Chem.*, 1973, **16**, 545.

Table 3 *Synthetic analogues of LRF*

Compound number	*Structure*	*Biological activity*		*Ref.*
		LRF *type*	FSH-RF *type*	
(31)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	100	100	
(32)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-NH_2	11	significant	73, 74
(33)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-NH_2	< 0.01	not significant	74
(34)	Glp-His-Trp-Ser-Tyr-Gly-Leu-NH_2	< 0.01	not significant	74
(35)	Glp-His-Trp-Ser-Tyr-Gly-NH_2	< 0.01	not significant	74
(36)	Glp-His-Trp-Ser-Tyr-NH_2	< 0.01	not significant	74
(37)	Glp-His-Trp-Ser-NH_2	< 0.01	not significant	74
(38)	Glp-His-Trp-NH_2	< 0.01	not significant	74
		considerably less than 1.0	not detectable	75
		1.0	not detectable	76
		0.4		77
		inactive	inactive	78—80
(39)	Glp-His-Trp-OH	marginally active	inactive	75
		inactive	inactive	79
(40)	Glp-His-NH_2	< 0.01	inactive	74
(41)	Glp-NH_2	0.01	inactive	74
(42)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-$NHCH_3$	47	56	73
		80—100	—	81
(43)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-$N(CH_3)_2$	15	9	73
(44)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-$NHCH_2CH_3$	300	280	73
		500	—	81
(45)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-$NHCH_2CH_2CH_3$	190	210	73
		200—300	—	81
(46)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-$NHCH_2CH_2OH$	210	220	73
		100—150	—	81

Table 3 (*cont.*)

Compound number	*Structure*	*Biological activity*		*Ref.*
		LRF *type*	FSH-RF *type*	
(47)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-$NHCH(CH_3)_2$	150	100	73
(48)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-pyrrolidine	110	92	73
		70—80	—	81
(49)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-morpholine	17	16	73
		20—30	—	81
(50)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-piperidine	1.5	2	73
(51)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-cyclohexylamide	0.7	0.6	73
(52)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-$NHCH_2CH_2CH_2CH_3$	7.2	4.3	73
(53)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-$NHCH_2CH(CH_3)_2$	3.0	2.8	73
(54)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-OCH_3	6.2	2.9	73
(55)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-OCH_2CH_3	11.5	6.4	73
(56)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-$N(CH_3)_2$	14	—	82
(57)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Ala-NH_2	6	3.5	73
		0.85	—	82
		10	—	83
(58)	Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Pro-NH_2	0.18	0.14	73
(59)	Glp-His-Trp-Ser-Tyr-Ile-Leu-Arg-Pro-Gly-NH_2	0.034	—	82
(60)	Glp-His-Trp-Ser-Tyr-Ala-Leu-Arg-Pro-Gly-NH_2	4.0	—	84
(61)	Glp-His-Trp-Ser-Tyr-D-Ala-Leu-Arg-Pro-Gly-NH_2	357—474	—	84
(62)	Glp-His-Trp-Ser-Tyr-Val-Leu-Arg-Pro-Gly-NH_2	< 0.1	—	84
(63)	Glp-His-Trp-Ser-Tyr-D-Val-Leu-Arg-Pro-Gly-NH_2	32	—	84
(64)	Glp-His-Trp-Ser-Tyr-Pro-Leu-Arg-Pro-Gly-NH_2	1.5—3.0	—	84
(65)	Glp-His-Trp-Ser-Tyr-D-Pro-Leu-Arg-Pro-Gly-NH_2	3.3—10	—	84
(66)	Glp-His-Trp-Ser-Phe-Gly-Leu-Arg-Pro-Gly-NH_2	50	—	83
		64	98	85
		44.3	—	86
(67)	Glp-His-Trp-Ser-Tyr(Me)-Gly-Leu-Arg-Pro-Gly-NH_2	24	—	87

(68)	Glp-His-Trp-Ser-Phe(p-NH_2)-Gly-Leu-Arg-Pro-Gly-NH_2	37	—	87
(69)	Glp-His-Trp-Ser-Phe(p-NO_2)-Gly-Leu-Arg-Pro-Gly-NH_2	5	—	87
(70)	Glp-His-Trp-Ala-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	3—6	—	83
		8.6	—	86
(71)	Glp-His-Trp-Thr-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	19	—	82
		4	17	83
(72)	Glp-His-Trp-Ala-Phe-Gly-Leu-Arg-Pro-Gly-NH_2	1.01	—	86
(73)	Glp-His——Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.0026	—	86
(74)	Glp-His-Leu-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.0008	—	86
(75)	Glp-His-Phe-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.43	—	86
(76)	Glp-His-Tyr-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.13	—	82
(77)	Glp-Ser-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.032	—	88
(78)	Glp-Leu-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.18	—	88
(79)	Glp-Gln-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.01	—	88
(80)	Glp-Phe-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	1.4	—	88
(81)	Glp-Trp-His-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.0093	—	86
(82)	Glp——Trp-Ser-Phe-Gly-Leu-Arg-Pro-Gly-NH_2	0.007	—	86
(83)	——His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	very weakly active	—	90
(84)	————Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	very weakly active	—	90
(85)	Gly-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0	—	91
(86)	Pro-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.0088	—	83
		< 0.1	—	88
(87)	Leu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.003	—	82
(88)	Orotic acid-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.005	—	88
(89)	Glu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	5.7	—	88
(90)	Formyl-Gly-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	1.46	—	91
(91)	Acetyl-Gly-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.90	—	91
(92)	Propionyl-Gly-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0.22	—	91
(93)	Palmitoyl-Gly-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0	—	91
(94)	Acetyl-Ala-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2	0	—	91

indicating that the reduced activity results from lowered affinity for LRF receptor(s). Reducing the chain length even further by eliminating other amino-acids from the *C*-terminus (33)—(41) markedly decreases biological activity, showing that the full decapeptide sequence is required for high activity.[74–79] Disagreement continues regarding the biological activity of the *N*-terminal tripeptide amide sequence (38), possibly because of the difficulty in comparing the low potency of this compound with the highly active decapeptide.[74–80] Replacement of the *C*-terminal glycine amide with various linear, branched, or cyclic alkylamines provides analogues (42)—(53) with considerable biological activity.[73, 81] Some of these analogues (44)—(47) are in fact much more active than LRF; the reason for the enhanced activity is not clear – it may be that the analogues have increased affinity for the LRF receptors or it may be that they have a prolonged half-life by virtue of the protection afforded against a specific deamidase or other hydrolytic enzymes. It appears that compounds which have greatest conformational similarity to the N—C—C unit of the replaced glycine amide have the highest activity. The amides (42) and (44) are much more active than the comparable chain-length esters (54) and (55). Replacement of the glycine residue in position 10 of LRF with alanine (57) reduces the biological activity to 0.85—10%;[73, 82, 83] replacement with proline (58) results in almost complete loss of activity.[73]

Analogues (59), (60), (62) with L-isoleucine, L-alanine, or L-valine residues in place of the glycine at position 6 have low activity,[82, 84] but analogues (61) and (63) with D-alanine or D-valine in place of this glycine residue have high activity.[84] The low potency of [6-D-valine]-LRF (63) compared with [6-D-alanine]-LRF (61) can be understood on steric grounds. The high activity of the D-alanine analogue (61) and low activity of the

[75] R. S. Mecklenburg, K. Noda, Y. Miyachi, E. Gross, and M. B. Lipsett, *Endocrinology*, 1973, **93**, 993.

[76] W. Vale, M. Amoss, R. Blackwell, R. Burgus, G. Grant, N. Ling, M. W. Monahan, J. Rivier, and R. Guillemin, Abstracts, 4th International Congress of Endocrinology, Washington, D.C., 1972, *Excerpta Medica, International Congress Series* No. 256, p. 86 (Abstract 218).

[77] M. Amoss, J. Rivier, and R. Guillemin, *J. Clin. Endocrinol. Metab.*, 1972, **35**, 175.

[78] H. Matsuo, Y. Baba, R. M. G. Nair, A. Arimura, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1971, **43**, 1334.

[79] A. V. Schally, A. Arimura, W. H. Carter, T. W. Redding, R. Geiger, W. König, H. Wissman, G. Jaeger, J. Sandow, N. Yanaihara, C. Yanaihara, T. Hashimoto, and M. Sakagami, *Biochem. Biophys. Res. Comm.*, 1972, **48**, 366.

[80] L. Debeljuk, A. Arimura, and A. V. Schally, *Neuroendocrinology*, 1973, **11**, 130.

[81] M. Fujino, S. Kobayashi, M. Obayashi, S. Shinagawa, T. Fukuda, C. Kitada, R. Nakayama, I. Yamazaki, W. F. White, and R. H. Rippel, *Biochem. Biophys. Res. Comm.*, 1972, **49**, 863.

[82] D. H. Coy, E. J. Coy, and A. V. Schally, *J. Medicin. Chem.*, 1973, **16**, 1140.

[83] M. Fujino, S. Kobayashi, M. Obayashi, T. Fukuda, S. Shinagawa, I. Yamazaki, R. Nakayama, W. F. White, and R. H. Rippel, *Biochem. Biophys. Res. Comm.*, 1972, **49**, 698.

[84] (*a*) M. W. Monahan, W. Vale, J. Rivier, G. Grant, and R. Guillemin, Abstracts, 55th Annual Meeting of the Endocrine Society, Chicago, Illinois, U.S.A., June 1973; (*b*) M. W. Monahan, M. S. Amoss, H. A. Anderson, and W. Vale, *Biochemistry*, 1973, **12**, 4616.

L-alanine analogue (60) in comparison with LRF might be explained by conformational considerations – the presence of D- or L-alanine residues in position 6 could restrict the degrees of conformational freedom of LRF to give molecules with increased or decreased affinity for binding at the receptor site. The fact that LRF and [6-D-alanine]-LRF (61) exhibit parallel dose–response curves with similar response maxima is consistent with the idea that the increased specific activity of (61) is due to increased binding affinity for the pituitary receptors. It has been suggested that peptide (61) could exist more extensively than (60) in a conformation incorporating a β-II type bend.[84b]

The phenolic hydroxy-group of the tyrosine residue in position 5 is not essential for activity since it can be eliminated (66),[83, 85, 86] or replaced by other groups [(67)—(69)][87] with retention of activity. The serine hydroxy-group in position 4 is not essential for activity, but its elimination in [4-alanine]-LRF (70) causes a considerable loss of biological activity.[83, 86] Elimination of the tryptophan residue in position 3 (73) or its replacement by leucine (74) abolishes all activity whereas its replacement with phenylalanine (75) or tyrosine (76) results in very low but definite LRF activity.[82, 86] Analogues (77)—(79), with an aliphatic amino-acid replacing histidine in position 2, have very low biological activity, but when the histidine residue is replaced by phenylalanine (80) the activity is slightly higher.[88] These results, together with the almost complete loss of activity which occurs when LRF is photo-oxidized with or without added sensitizer,[89] indicate the importance of aromatic side-chains in positions 2 and 3. The histidine and tryptophan residues may contribute to the intrinsic activity as well as to the binding of LRF to its receptors.

Des-1-pyroglutamic acid-LRF (83) and des-1-pyroglutamic acid-des-2-histidine-LRF (84) are almost inactive.[90] Analogues of LRF in which the pyroglutamic acid residue is replaced by glycine (85), proline (86), leucine (87), or orotic acid (88) are almost inactive,[82, 83, 88, 91] but [1-glutamic acid]-LRF (89) retains 5.7% activity.[88] Some N^{α}-acyl-[1-glycine]-LRF analogues (90)—(92) show slight activity but (93) is completely inactive.[91] These results indicate that the *N*-terminal pyroglutamic acid residue is most important for biological activity and probably contributes to the intrinsic activity of LRF as well as to binding at the receptors.

[85] D. H. Coy, E. J. Coy, and A. V. Schally, *J. Medicin. Chem.*, 1973, **16**, 83.
[86] N. Yanaihara, T. Hashimoto, C. Yanaihara, K. Tsuji, Y. Kenmochi, F. Ashizawa, T. Kaneko, H. Oka, S. Saito, A. Arimura, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 64.
[87] D. H. Coy, E. J. Coy, and A. V. Schally, *J. Medicin. Chem.*, 1973, **16**, 827.
[88] N. Yanaihara, K. Tsuji, C. Yanaihara, T. Hashimoto, T. Kaneko, H. Oka, A. Arimura, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 165.
[89] J. F. Kennedy, C. J. Gray, S. Ramanvongse, L. Albrighton, and W. F. White, *Life Sci.*, part I, 1973, **12**, 533.
[90] N. Yanaihara, C. Yanaihara, M. Sakagami, K. Tsuji, T. Hashimoto, T. Kaneko, H. Oka, A. V. Schally, A. Arimura, and T. W. Redding, *J. Medicin. Chem.*, 1973, **16**, 373.
[91] Y. Okada, K. Kitamura, Y. Baba, A. Arimura, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 1180.

The residual activity retained by compounds (90)—(92) suggests that the —CONHCHCO— portion of the pyroglutamic acid residue may be the essential part for activity and that this is fixed into a desirable conformation by the ring structure.

The biological activity summarized in Table 3 generally refers to results in laboratory animals. Some analogues have also been tested in organ cultures of human foetal pituitaries.[92] In general, the results agree with those obtained in animals; analogues (32) and (33) showed some activity but peptide (84) was inactive. Compounds (95) and (96), which have been

Glp-Gln-Ala-NH_2 (95)

Glp-Tyr-Arg-Trp-NH_2 (96)

Glp-His-Trp-Ser-Tyr-Gly-OH (97)

Leu-Arg-Pro-Gly-NH_2 (98)

Glp-His-Pro-Gly-NH_2 (99)

claimed to possess low levels of LRF activity, were completely inactive. Peptides (97) and (98), representing the *N*-terminal and *C*-terminal regions of LRF respectively, were inactive individually, but a mixture of the two enhanced the release of both LH and FSH. This suggests that specific binding sites for the *N*- and *C*-terminal regions of LRF may exist in the receptors on pituitary gonadotrophic cells and that *N*- and *C*-terminal fragments at high concentration can occupy both these sites to produce a stimulation of LH and FSH release. This possibility was supported by the detection of activity with the tetrapeptide (99), which incorporates the *N*- and *C*-terminal dipeptide sequences of LRF.

The discovery of des-2-histidine-LRF and [2-glycine]-LRF as LRF antagonists[93] has stimulated further search for more potent antagonists since these could be useful in regulating fertility. Work has centred around the possibility that any modification in the LRF structure which results in an analogue with increased specific activity might be incorporated into known inhibitors of LRF to give increased antagonistic properties. In this way the enhanced agonist activities of [6-D-alanine]-LRF (61) and [9-proline ethylamide]-des-10-glycinamide-LRF (44) have prompted the synthesis of peptides (100) and (101)—(103), respectively. Compound (100) is about three times more potent as an inhibitor than des-2-histidine-LRF.[84] The octapeptide derivatives (101) and (102) are more potent inhibitors than the nonapeptide (103) but all three inhibit at very low

[92] G. V. Groom and A. R. Boyns, *F.E.B.S. Letters*, 1973, **33**, 57.
[93] R. Guillemin, M. Amoss, R. Blackwell, R. Burgus, G. Grant, N. Ling, M. W. Monahan, J. Rivier, and W. Vale, *Gynec. Invest.*, 1971/72, **2**, 2.

concentrations the release of LH stimulated by LRF.[94, 95] The inhibitory activity of these compounds may result from non-productive binding to pituitary receptors.

Glp——Trp-Ser-Tyr-D-Ala-Leu-Arg-Pro-Gly-NH_2
(100)

Glp——Trp-Ser-Tyr-Gly -Leu-Arg-Pro-NHC_2H_5
(101)

Glp——Leu-Ser-Tyr-Gly -Leu-Arg-Pro-NHC_2H_5
(102)

Glp-Gly-Leu-Ser-Tyr-Gly -Leu-Arg-Pro-NHC_2H_5
(103)

Tritiated LRF has been shown to bind to two distinct orders of binding sites on cultured normal rat anterior pituitary cells.[96] One is a high-affinity (2×10^{-9} mol l^{-1}), low-capacity site with a specificity for LRF agonist and antagonist analogues which reflects the known biological effects of these compounds; this site could be related to the physiological LRF receptor. The second is a low-affinity (2×10^{-8} mol l^{-1}), high-capacity site and lacks the ability to distinguish between active and inactive LRF analogues.[96]

The ability of anti-LRF serum to bind various analogues and fragments of LRF has also been studied;[97] cross-reactivity was shown by LRF free acid (0.2%), [4-threonine]-LRF (12%), [5-phenylalanine]-LRF (7%), [6-isoleucine]-LRF (6.4%), [8-lysine]-LRF (1.2%), des-1-pyroglutamic acid-LRF (20%), and des-1-pyroglutamic acid-des-2-histidine-LRF (2.4%). The antibody and the receptor appear to recognize different aspects of the molecule since weakly active [6-isoleucine]-LRF (59) and highly active [5-phenylalanine]-LRF (66) have similar low cross-reactivities and the weakly active des-1-pyroglutamic acid-LRF (83) has a comparatively high cross-reactivity. The anti-serum suppressed gonadotrophin secretion, and blocked ovulation in rats.[98, 99]

[94] D. H. Coy, J. A. Vilchez-Martinez, E. J. Coy, A. Arimura, and A. V. Schally, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 331.
[95] A. V. Schally, A. Arimura, D. H. Coy, E. J. Coy, R. Geiger, A. J. Kastin, W. König, and H. Wissman, Presented at XXIVth IUPAC Congress, Hamburg (Federal Republic of Germany), September 1973, p. 706.
[96] G. Grant, W. Vale, and J. Rivier, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 771.
[97] Y. Koch, M. Wilchek, M. Fridkin, P. Chobsieng, U. Zor, and H. R. Lindner, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 616.
[98] Y. Koch, P. Chobsieng, U. Zor, M. Fridkin, and H. R. Lindner, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 623.
[99] T. Makino, M. Takahashi, K. Yoshinaga, and R. O. Greep, *Contraception*, 1973, 133.

The involvement of an adenyl cyclase–phosphodiesterase system in the mechanism of action of LRF is supported by observations that the concentration of cyclic AMP in rat anterior pituitary cells is increased within 1 min of exposure to LRF, but is unaffected by a number of inactive analogues.[100]

Growth Hormone-releasing Factor (GH-RF).—It now seems unlikely that the decapeptide (104) isolated from porcine hypothalami is the same as the physiological GH-RF.[6] Though it stimulates the release of bioassayable growth hormone in certain tests it does not increase the plasma levels of immunoreactive growth hormone in rats, sheep, pigs, or monkeys. Similarly, studies in man with the substances isolated from porcine and human hypothalami and with the synthetic decapeptide revealed no convincing increase in plasma growth hormone levels.[101] In the first reported synthesis of the decapeptide,[102] a structural similarity with the proposed amino-terminal sequence (105) of the β-chain of porcine haemoglobin was noted. This suggested that the decapeptide may arise by the action of a proteolytic enzyme specific for cleavage of the β-chain of haemoglobin between residues 10 and 11, just as angiotensin I is generated by the renin cleavage of the α_2-globulin angiotensinogen between residues 10 and 11. The analogy has stimulated an attempt to prepare an inhibitor of the hypothetical 'GH-RF liberating enzyme' by synthesis of the hexapeptide amide (106) containing residues 10—15 of the β-chain of human haemoglobin [103] (biological results with this amide have not yet appeared). It is unlikely that the physiological GH-RF will arise from a haemoglobin precursor if, as has been suggested,[104] it is biosynthesized enzymatically by a non-ribosomal mechanism.

Val-His-Leu-Ser-Ala-Glu-Glu-Lys-Glu-Ala
(104)

Val-His-Leu-Ser-Ala-Glu-Glu-Lys-(Ala, Glx, Val) ...
(105)

Ala-Val-Thr-Ala-Leu-Trp-NH_2
(106)

Glp-Ser-Gly-NH_2
(107)

A tripeptide derivative (107) has been isolated from bovine hypothalami, synthesized, and claimed to stimulate the release of growth hormone from

[100] K. Wakabayashi, Y. Date, and B. Tamaoki, *Endocrinology*, 1973, **92**, 698; T. Kaneko, S. Saito, H. Oka, T. Oda, and N. Yanaihara, *Metabolism*, 1973, **22**, 77; Y. A. Fontaine, C. Salmon, E. Fontaine-Bertrand, and N. Delerue-Le Belle, *Hormone Metab. Res.*, 1973, **5**, 376.

[101] A. J. Kastin, A. V. Schally, C. Gual, S. Glick, and A. Arimura, *J. Clin. Endocrinol. Metab.*, 1972, **35**, 326.

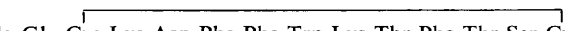
[102] D. F. Veber, C. D. Bennett, J. D. Milkowski, G. Gal, R. G. Denkewalter, and R. Hirschmann, *Biochem. Biophys. Res. Comm.*, 1971, **45**, 235.

[103] S. Bajusz, I. Fauszt, and J. Barvendeg, *Acta Chim. Acad. Sci. Hung.*, 1973, **76**, 431.

[104] S. Reichlin and M. Mitnick, *Proc. Soc. Exp. Biol. Med.*, 1973, **142**, 497.

rat pituitaries *in vitro*.[105] The same peptide synthesized in another laboratory was inactive in a number of *in vivo* and *in vitro* assay systems for growth hormone-releasing activity.[106]

Growth Hormone-release Inhibiting Factor (GH-RIF).—Further details have appeared of the primary sequence determination of GH-RIF (somatostatin) (108),[107] and four syntheses have been reported.[108–111] The secretion of radioimmuno-assayable growth hormone in a number of *in vitro* systems [112] was inhibited with equal potency by the natural, the synthetic linear (dithiol), and the synthetic cyclic peptides.[108] A synthetic product believed to be mainly dimerized peptide also possessed some inhibitory activity.[111] That a covalent cyclic structure is not essential to the activity of somatostatin was convincingly demonstrated by synthesis of the analogue (109) in which the cysteine residues at positions 3 and 14 have been replaced by alanine.[113] This showed about 0.01% of the potency of somatostatin. It was concluded that the disulphide bridge of somatostatin is not involved in the reaction of somatostatin with its receptor but that the cyclic structure may serve to increase the population of conformations which are recognized by the receptor and trigger the biological response.



Ala-Gly-Cys-Lys-Asn-Phe-Phe-Trp-Lys-Thr-Phe-Thr-Ser-Cys

(108)

Ala-Gly-Ala-Lys-Asn-Phe-Phe-Trp-Lys-Thr-Phe-Thr-Ser-Ala

(109)

The availability of synthetic somatostatin and the possible importance of a compound with growth hormone-release inhibiting activity in the management of diseases associated with excess growth hormone secretion (*e.g.* acromegaly, gigantism, and diabetes mellitus) have led to very prompt investigation of its effects in humans. Reduced somatostatin abolished the increase in serum growth hormone levels produced in normal subjects

[105] N. A. Yudaev, Z. F. Outecheva, T. E. Novikova, Y. P. Chvatchkin, and A. P. Smirnova, *Doklady Akad. Nauk S.S.S.R.*, 1973, **210**, 731.

[106] A. V. Schally, T. W. Redding, J. Takahara, D. H. Coy, and A. Arimura, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 556.

[107] N. Ling, R. Burgus, J. Rivier, W. Vale, and P. Brazeau, *Biochem. Biophys. Res. Comm.*, 1973, **50**, 127; R. Burgus, N. Ling, M. Butcher, and R. Guillemin, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 684.

[108] J. Rivier, P. Brazeau, W. Vale, N. Ling, R. Burgus, C. Gilon, J. Yardley, and R. Guillemin, *Compt. rend.*, 1973, **276**, *D*, 2737.

[109] D. Sarantakis and W. A. McKinley, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 234.

[110] D. Yamashiro and C. H. Li, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 882.

[111] D. H. Coy, E. J. Coy, A. Arimura, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 1267.

[112] W. Vale, P. Brazeau, G. Grant, A. Nussey, R. Burgus, J. Rivier, N. Ling, and R. Guillemin, *Compt. rend.*, 1972, **275**, *D*, 2913; P. Brazeau, W. Vale, R. Burgus, N. Ling, M. Butcher, J. Rivier, and R. Guillemin, *Science*, 1973, **179**, 77.

[113] D. Sarantakis, W. A. McKinley, and N. H. Grant, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 538.

by administration of arginine and L-dopa;[114] somatostatin was a potent inhibitor of (*a*) elevated plasma growth hormone levels stimulated by insulin-induced hypoglycaemia in normal males,[115] or by exercise in normal subjects and in diabetics,[116] and (*b*) high, unstimulated circulating growth hormone levels associated with acromegalic patients.[115] No effects were observed on the basal levels of thyrotrophin, prolactin, luteinizing hormone, follicle-stimulating hormone,[114, 115] adrenocorticotropin or corticosteroids,[115] in these subjects. The increases in serum thyrotropin and follicle-stimulating hormone concentrations in response to TRF were smaller than controls during somatostatin infusion, but the increases in serum prolactin after TRF administration, and in luteinizing hormone and follicle-stimulating hormone after LRF injection, were uninfluenced by the presence of somatostatin.[115] (These results support the possibility that different mechanisms are involved in the release of thyrotropin and prolactin by TRF and indicate that more than one type of pituitary cell receptor may be involved with follicle-stimulating hormone release.) However, exploitation of somatostatin in the treatment of acromegaly or in inhibiting the development of diabetic angiopathy may be complicated by the observation that it lowered basal plasma insulin levels in some normal subjects and markedly suppressed intravenous glucose-induced insulin release, with subsequent lowering of the glucose-disappearance rate.[117] *In vitro* perfusion experiments on isolated canine pancreas indicated that somatostatin acted directly on the beta cells, causing suppression of the initial and late sustained phases of insulin release.[117]

Another peptide, with the ability to inhibit growth hormone release *in vitro* but not *in vivo*, has been isolated from hypothalamic extracts, and shown to be identical with α-melanocyte stimulating hormone (α-MSH).[118]

Melanocyte Stimulating Hormone-release Inhibiting Factor (MSH-RIF).—There is still disagreement between various investigators as to the identity of MSH-RIF. Peptides (110)—(112), all fragments of oxytocin, were devoid of MSH-RIF activity when tested in rat pituitary cell cultures and the tripeptide amide (110) was also without effect on whole and hemi-pituitary preparations.[119] However, peptide (110), but not tocinoic acid (111), was observed to inhibit MSH release in rats with median eminence lesions (apparently by a direct action on the pituitary).[120]

[114] T. M. Siler, G. VandenBerg, S. S. C. Yen, P. Brazeau, W. Vale, and R. Guillemin, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 632.
[115] R. Hall, G. M. Besser, A. V. Schally, D. H. Coy, D. Evered, D. J. Goldie, A. J. Kastin, A. S. McNeilly, C. H. Mortimer, C. Phenekos, W. M. G. Tunbridge, and D. Weightman, *Lancet*, 1973, 581.
[116] Aa. Prange-Hansen, H. Ørskov, K. Seyer-Hansen, and K. Lundbaek, *Brit. Med. J.*, 1973, **3**, 523.
[117] K. G. M. M. Alberti, N. J. Christensen, S. Engkjaer-Christensen, Aa. Prange-Hansen, J. Iversen, K. Lundbaek. K. Seyer-Hansen, and H. Ørskov, *Lancet*, 1973, 1299.
[118] A. V. Schally, A. Arimura, A. J. Kastin, T. Uehara, D. H. Coy, E. J. Coy, and J. Takahara, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 1314.
[119] N. H. Grant, D. E. Clark, and E. I. Rosanoff, *Biochem. Biophys. Res. Comm.*, 1973, **51**, 100.
[120] M. E. Celis, R. Macagno, and S. Taleisnik, *Endocrinology*, 1973, **93**, 1229.

Pro-Leu-Gly-NH_2	Cys-Tyr-Ile-Gln-Asn-Cys	Cys-Tyr-Ile-Gln-Asn
(110)	(111)	(112)

The tripeptide amide (110) is believed to have an anti-depressant action which is independent of any MSH-RIF activity,[121] and it also decreases spontaneous locomotor activity and increases stereotyped behaviour in cats.[122] N.m.r.[123] and *X*-ray studies[124] suggest that this peptide has a ten-membered β-turn structure. A membrane-bound exopeptidase which degrades [9-glycinamide-1-^{14}C]-oxytocin to give labelled peptide (110) has been identified in rabbit and rat hypothalamus.[125] Studies with Pro-[^{14}C]Leu-Gly-NH_2 and [^{3}H]Pro-Leu-Gly-NH_2 in human plasma and in rats indicate that the major metabolic pathway for degradation of the tripeptide amide is cleavage between the proline and leucine residues, since proline and Leu-Gly-NH_2 are generated as the main products (75%) of inactivation.[126] Cleavage of the *C*-terminal amide, a major pathway in the inactivation of TRF, is of minor importance in the degradation of peptide (110).

3 Pituitary Hormones

Adrenocorticotropin (ACTH).—An immunoreactive ACTH component present in human plasma and pituitary extracts, and considerably larger and more acidic than the well-known 1—39 ACTH peptide, has been more fully characterized.[127] This 'Big ACTH' is not readily converted into ACTH either spontaneously or in 8M-urea but yields a component very similar to authentic ACTH on trypsin digestion. It is suggested that 'Big ACTH' could be a 'pro-ACTH' consisting of ACTH covalently linked at its amino-terminus to a lysine or arginine residue at the *C*-terminus of a larger, more acidic peptide. (Evidence for a form having higher molecular weight of biologically active ACTH has also been obtained in rats.)[128] On the other hand, isolation of a peptide resembling the *C*-terminal portion of ACTH (residues 18—39) from rat and pig *pars intermedia* has prompted a hypothesis that, in the intermediate lobe of the pituitary, ACTH may itself act as a prohormone for α-MSH.[129] It is proposed that an intracellular cleavage of 1—39 ACTH occurs in the 13—18 region of the molecule.

[121] N. P. Plotnikoff, A. J. Kastin, M. S. Anderson, and A. V. Schally, *Neuroendocrinology*, 1973, **11**, 67.
[122] R. B. North, S. I. Harik, and S. H. Snyder, *Brain Res.*, 1973, **63**, 435.
[123] R. Deslauriers, R. Walter, and I. C. P. Smith, *F.E.B.S. Letters*, 1973, **37**, 27.
[124] L. L. Reed and P. L. Johnson, *J. Amer. Chem. Soc.*, 1973, **95**, 7523.
[125] R. Walter, E. C. Griffiths, and K. C. Hooper, *Brain Res.*, 1973, **60**, 449.
[126] R. M. G. Nair, T. W. Redding, A. J. Kastin, and A. V. Schally, *Biochem. Pharmacol.*, 1973, **22**, 1915.
[127] R. S. Yalow and S. A. Berson, *J. Clin. Endocrinol. Metab.*, 1973, **36**, 415.
[128] R. E. Lang, H. L. Fehm, K. H. Voigt, and E. F. Pfeiffer, *F.E.B.S. Letters*, 1973, **37**, 197.
[129] A. P. Scott, J. G. Ratcliffe, L. H. Rees, J. Landon, H. P. J. Bennett, P. J. Lowry, and C. McMartin, *Nature New Biol.*, 1973, **244**, 65.

One cleavage product could then be transformed into α-MSH following acetylation, amidation, and, if necessary, carboxypeptidase digestion. The other cleavage product might provide the newly identified 18—39 peptide by sequential aminopeptidase digestion up to the enzyme-resistant Arg^{18}–Pro^{19} peptide bond.[129]

Peptides related to ACTH [(113), Table 4] have been used to explore further the contribution to biological activity of the basic amino-acid residues in positions 11, 17, and 18. Peptides (114)—(121) were compared in an *in vivo* ascorbic acid depletion test in dexamethasone-blocked rats.[130, 131] Replacement of one lysine residue in position 17 or 18 in peptide (114) by leucine [peptides (115) and (116)] gave about 50% reduction in biological activity. Compound (121) is only half as active as peptide (116), indicating that complete elimination of any side-chain at position 18 leads to a further reduction in activity; the presence of a *C*-terminal amide group at position 18 is not important since (117) is 50% more active again than (114). The lysine residue in position 11 of peptide (117) can be replaced by arginine (118) without effect on activity, but elimination of this basic side-chain (120) or its replacement by an aliphatic chain (119) leads to a drastic decrease in activity. This indicates an important binding role for a basic amino-acid residue at this point.

These conclusions from *in vivo* experiments are similar to those drawn from *in vitro* binding studies with a cortical membrane preparation[132] (see ref. 1). The binding role of the basic residues in positions 15—18 of ACTH has also been confirmed by measurement of steroidogenic activity with isolated intact adrenal cells.[133] The tetradecapeptide (122) possessed about 1/3000 of the potency of ACTH-(1—18)-octadecapeptide (123) on a molar basis but had the same intrinsic activity. In the same study, replacement of the L-phenylalanine residue at position 7 of peptide (124) by the D-isomer (125) reduced steroidogenic activity to about 3%. This drop was again ascribed to a decrease in affinity of peptide (125) for the adrenal receptor since both peptides gave almost identical maximum levels of corticosterone production and (125) did not act as a competitive inhibitor of (124).[133]

Replacement of the arginine residue in position 8 by homoarginine[134] gave peptides (126) and (127), which retained 10—50% of their steroidogenic and lipolytic activity both *in vivo* and *in vitro*.[135] This is in marked contrast to peptides substituted with lysine or ornithine in position 8

[130] R. Geiger and H.-G. Schroeder, 'Progress in Peptide Research', Proceedings of 2nd American Peptide Symposium, Cleveland, Ohio, August, 1970, ed. S. Lande, Gordon and Breach, 1972, p. 273.

[131] R. Geiger and H.-G. Schroeder, *Z. physiol. Chem.*, 1973, **354**, 156.

[132] F. M. Finn, C. C. Widnell, and K. Hofmann, *J. Biol. Chem.*, 1972, **247**, 5695.

[133] M. Nakamura, *J. Biochem.*, 1972, **71**, 1029.

[134] G. I. Tesser, A. W. J. Pleumekers, W. Bassie, and I. C. Balvert-Geers, *Rec. Trav. chim.*, 1973, **92**, 1210.

[135] G. I. Tesser, R. Maier, L. Schenkel-Hulliger, P. L. Barthe, B. Kamber, and W. Rittel, *Acta Endocrinol.*, 1973, **74**, 56.

Table 4 *Analogues of ACTH*

Compound number	*Structure* [a]
	1 5 10 15 20 24 39
(113)	Ser-Tyr-Ser-Met-Glu-His-Phe-Arg-Trp-Gly-Lys-Pro-Val-Gly-Lys-Lys-Arg-Arg-Pro-Val-Lys-Val-Tyr-Pro-...-Phe
(114)	β-Ala—Lys-Lys-NH_2
(115)	β-Ala—Leu-Lys-NH_2
(116)	β-Ala—Lys-Leu-NH_2
(117)	β-Ala—Lys-$NH(CH_2)_4NH_2$
(118)	β-Ala—Arg—Lys-$NH(CH_2)_4NH_2$
(119)	β-Ala—Nle—Lys-$NH(CH_2)_4NH_2$
(120)	β-Ala—Gly—Lys-$NH(CH_2)_4NH_2$
(121)	β-Ala—$NH\cdot(CH_2)_6\cdot NH_2$
(122)	Gly—OH
(123)	—OH
(124)	β-Ala—Orn—NH_2
(125)	β-Ala—D-Phe—Orn—NH_2
(126)	—Har—Lys-Lys-NH_2 [b]
(127)	—Har—OH
(128)	—OH
(129)	—Arg—OH
(130)	—NH_2
(131)	—OH
(132)	—OH
(133)	Nps —.....Phe [c]
(134)	Dnps —OH [d]
(135)	—Phe—Nva*—OH [e]
(136)	—MePhe—OH [f]

[a] Amino-acid residues are the same as in (113) unless indicated; [b] Har = homoarginine; [c] Nps = *o*-nitrophenylsulphenyl; [d] Dnps = 2,4-dinitrophenylsulphenyl; [e] Nva* = 4,5-ditritio-norvaline = NHCHCO— $\vert$ —CH_2—CHT—CH_2T [f] MePhe = *N*-methylphenylalanine.

which showed only residual biological activity. These results demonstrate that the presence of a guanidino-group in the functionally important position 8 of ACTH peptides is essential for the retention of high biological activity; the length of the aliphatic chain carrying the guanidino-group is of less importance.[135] When arginine was substituted for glutamic acid in the heptapeptide fragment (128), the product (129) gave the same maximal stimulation of lipolysis in isolated rabbit fat cells, but was about four times as potent as peptide (128).[136] Promised studies with different assay systems and other analogues incorporating the 5-arginine substitution will be needed to demonstrate the generality of this somewhat unexpected observation.

The octapeptide (130) and the tetradecapeptide (131) contain the important binding site -Lys-Lys-Arg-Arg- but lack the sequence -His-Phe-Arg-Trp-Gly- necessary for function. The tetradecapeptide (131) has been shown to act as a competitive antagonist of ACTH (113) and its biologically active analogues in isolated adrenal cells for both corticosteroid and cyclic-AMP production.[137] Contrary to expectation, other workers found the octapeptide (130) to potentiate the submaximal steroidogenesis induced by low concentrations of [1-glycine]-ACTH-(1—18)-octadecapeptide amide both in isolated adrenal cells and *in vivo*.[133] The results were interpreted as an indication of two types of binding for ACTH on the adrenal cell. It was suggested that under the conditions used, peptide (130) competes with corticotropic peptides for non-specific adsorption sites in preference to the specific, biologically productive receptors. Addition of peptide (130) is envisaged as liberating ACTH bound to non-specific sites, making it available for interaction with true receptors and so producing the observed increase in corticosterone production.[133] It will be interesting to see these results extended to include measurements of cyclic-AMP production.

Details have appeared of earlier reports (see ref. 1) on peptides able to elicit maximum corticosteroid production, but which do not induce maximum cyclic-AMP production [compound (132)][137] or which will inhibit ACTH-induced cyclic-AMP synthesis [compound (133)].[138] It was estimated that cyclic-AMP production need not exceed 20% of the maximum possible in order to induce maximum corticosterone production.[137] The relatively low affinity of peptide (131), and the relatively high intrinsic steroidogenic activity of peptide (133), restrict their value as inhibitors. A new derivative (134) has been prepared which has low intrinsic activity and a relatively high affinity for isolated adrenal cortex cells. It inhibits both cyclic-AMP and corticosterone production induced by 1—24 ACTH.[139]

[136] M. W. Draper, R. B. Merrifield, and M. A. Rizack, *J. Medicin. Chem.*, 1973, **16**, 1326.
[137] S. Seelig and G. Sayers, *Arch. Biochem. Biophys.*, 1973, **154**, 230.
[138] W. R. Moyle, Y. C. Kong, and J. Ramachandran, *J. Biol. Chem.*, 1973, **248**, 2409.
[139] S. Kumar, S. Seelig, and G. Sayers, *Biochem. Biophys. Res. Comm.*, 1972, **49**, 1316.

A strongly radio-labelled ACTH derivative (135) has been synthesized [140] and shown to have essentially the same intrinsic biological activity, but only about one-tenth of the potency, of 1—24 ACTH in isolated rat lipocytes and adrenal cortex cells.[141] The peptide has already been used to study the availability of receptors on isolated fat cells.[4b] More recently, a carefully designed synthesis has yielded [3,5-ditritio-Tyr^{23}]-ACTH(1—24).[142] This peptide, like compound (135), incorporates in a precise location the high specific activity required for biological studies, but does so with retention of the natural tetracosapeptide sequence.

In another new analogue (136), N-methylation of the *N*-terminal peptide bond did not prolong or enhance the *in vivo* steroidogenic response, suggesting that this modification did not increase resistance to digestion by circulatory exopeptidases.[143] Finally, a new total synthesis of 1—39 ACTH by the solid-phase procedure appeared in 1973.[144] The isolation by routine chromatographic means, without recourse to techniques such as affinity chromatography or selective enzymic degradation, of a product comparable with natural ACTH (by various chemical and biological criteria) is a significant achievement for solid-phase synthesis and promises easier access, hopefully, to synthetic analogues of ACTH.

As well as its primary role in stimulating the synthesis and secretion of adrenal cortical steroids, ACTH has a number of extra-adrenal activities. One of these, with a growing body of literature, relates to the effect of ACTH in restoring to rats the ability, lost after hypophysectomy, to acquire a conditioned avoidance response.[145] Synthetic fragments of ACTH which contain a minimum sequence comprising residues 4—10 of the parent hormone also show facilitation of avoidance conditioning in hypophysectomized rats, and treatment of intact rats with the same peptides delays the extinction of a conditioned avoidance response.[146, 147] [7-D-phenylalanine]-ACTH-(4—10)-heptapeptide showed an opposite effect; it facilitated extinction of avoidance behaviour in intact rats and failed to facilitate acquisition in hypophysectomized rats.[146] A review of these behavioural effects of ACTH peptides has appeared recently.[148] Since ACTH-(4—10) and [7-D-phenylalanine]-ACTH-(4—10) have different effects on conditioned avoidance they have been used in studies aimed at correlation of the behavioural phenomenon with changes in brain amine metabolism.[149, 150] The results are inconsistent, but indicate that there is

[140] R. Schwyzer and G. Karlaganis, *Annalen*, 1973, 1298.
[141] U. Lang, G. Karlaganis, S. Seelig, G. Sayers, and R. Schwyzer, *Helv. Chim. Acta*, 1973, **56**, 1069.
[142] D. E. Brundish and R. Wade, *J.C.S. Perkin I*, 1973, 2875.
[143] J. Blake and C. H. Li, *Internat. J. Peptide Protein Res.*, 1972, **4**, 343.
[144] D. Yamashiro and C. H. Li, *J. Amer. Chem. Soc.*, 1973, **95**, 1310.
[145] D. E. de Wied, *Amer. J. Physiol.*, 1964, **207**, 255.
[146] D. E. de Wied, in 'Frontiers in Neuroendocrinology', ed. W. F. Ganong and L. Martin, Oxford University Press, 1969, p. 97.
[147] B. Bohus, W. H. Gispen, and D. E. de Wied, *Neuroendocrinology*, 1973, **11**, 137.
[148] D. E. de Wied, *La Recherche*, 1973, **4**, 939.
[149] B. E. Leonard, *Brit. J. Pharmacol.*, 1973, **46**, 560P.
[150] D. H. G. Versteeg, *Brain Res.*, 1973, **49**, 483.

not a simple relationship between brain amine turnover and the behavioural effects.

Growth Hormone.—The regulation,[151] secretion and action,[152] relationship with diabetes mellitus,[153] and mechanism of the diabetogenic effects [154] of growth hormone, and the chemistry of growth hormone and the lactogenic hormones [155] have all been reviewed. Reports from a symposium on growth and growth hormones [156] and a monograph on human pituitary growth hormone [157] have been published.

The primary structures of bovine [158] and ovine [159] growth hormones and of human placental lactogen [160] have all been determined or revised and, like human growth hormone,[155, 160, 161] are now believed to contain 191 amino-acid residues. These structures relate to growth hormones isolated from pituitary tissue. However, there are indications that some kind of high molecular weight precursor (a possible pro-growth hormone) may exist,[162] and suggestions that, once released into the circulation, the growth hormone molecule may require further modification or cleavage before it can be fully biologically effective.[155, 163–165] An extension of this possibility is speculation that the familiar growth hormone molecule may itself be a circulating pro-hormone, that is a source of smaller peptides each of which could be the effective trigger for different aspects of the wide range of metabolic effects now attributed to the parent molecule.[154] It is also possible that many of the anabolic effects of growth hormone are not produced directly, but are mediated by one or more growth-hormone-dependent plasma factors, known as sulphation factor or somatomedin.[166]

[151] S. Reichlin, *Metabolism*, 1973, **22**, 987.
[152] T. J. Merimee and D. Rabin, *Metabolism*, 1973, **22**, 1235.
[153] J. D. Baird, W. M. Hunter, and A. W. M. Smith, *Post Grad. Med. J.*, 1973, **49** (Suppl. 1), 132.
[154] J. Bornstein, H. P. Taft, J. McD. Armstrong, F. M. Ng, and M. K. Gould, *Post Grad. Med. J.*, 1973, **49** (Suppl. 2), 219.
[155] H. D. Niall, M. L. Hogan, G. W. Tregear, G. V. Segre, P. Hwang, and H. Friesen, *Recent Progr. Hormone Res.*, 1973, **29**, 387.
[156] 'Growth and Growth Hormones', ed. A. Pecile and E. E. Muller, *Excerpta Medica Internat. Congress*, 1972, No. 244.
[157] A. W. Root, 'Human Pituitary Growth Hormone', Charles C. Thomas, Springfield, 1972.
[158] J. A. Santome, J. M. Dellacha, A. C. Paladini, C. Pena, M. J. Biscoglio, S. T. Daurat, E. Poskus, and C. E. M. Wolfenstein, *European J. Biochem.*, 1973, **37**, 164; M. Wallis, *F.E.B.S. Letters*, 1973, **35**, 11.
[159] C. H. Li, D. Gordon, and J. Knorr, *Arch. Biochem. Biophys.*, 1973, **156**, 493.
[160] C. H. Li, J. S. Dixon, and D. Chung, *Arch. Biochem. Biophys.*, 1973, **155**, 95.
[161] T. A. Bewley, J. S. Dixon, and C. H. Li, *Internat. J. Peptide Protein Res.*, 1972, **4**, 281.
[162] P. Gordon, C. M. Hendricks, and J. Roth, *J. Clin. Endocrinol. Metab.*, 1973, **36**, 178; P. Gordon, M. A. Lesniak, C. M. Hendricks, and J. Roth, *Science* 1973, **182**, 829; M. E. Stachura and L. A. Frohman, *Endocrinology*, 1973, **92**, 1708.
[163] R. A. Yadley and A. Chrambach, *Endocrinology*, 1973, **93**, 858; R. A. Yadley, D. Rodbard, and A. Chrambach, *ibid.*, p. 866.
[164] J. Schwartz, D. F. Nutting, H. M. Goodman, J. L. Kostyo, and R. E. Fellows, *Endocrinology*, 1973, **92**, 439.
[165] A. E. Wilhelmi, in discussion following ref. 155.
[166] W. H. Daughaday, K. Hall, M. S. Raben, W. D. Salmon, J. L. Van den Brande, and J. J. Van Wyk, *Nature*, 1972, **235**, 107; J. J. Van Wyk, L. E. Underwood, R. C. Lister, and R. N. Marshall, *Amer. J. Diseases of Children*, 1973, **126**, 705.

The recent isolation of highly purified somatomedin(s)[167] should permit a more exact definition of its contribution to the biological actions of growth hormone.

Whatever the state of endogenous growth hormone when it finally interacts with its various target cells, the search continues for chemically derived, biologically active modifications and fragments of the known molecule. Reduced, tetra-*S*-carbamidomethylated human growth hormone, previously shown to retain activity in the rat tibial cartilage assay and the pigeon crop-sac assay, has been tested in man.[168] It gave effects qualitatively and quantitatively similar to those of native human growth hormone (HGH) on nitrogen, calcium, phosphorus, and hydroxyproline excretion and on carbohydrate and lipid metabolism.[168] These results support the belief that the integrity of the two disulphide bridges is not important for many of the biological responses to growth hormone.

A tryptic cleavage peptide from bovine growth hormone which retains significant ability to increase both weight and tibial width has been isolated and its sequence determined.[169] It contains 37 amino-acid residues and is homologous to residues (95—134) of HGH. The demonstration of HGH-like metabolic effects after administration of this bovine growth hormone fragment to human hypopituitary dwarfs[170] has prompted several synthetic investigations of the corresponding region of HGH. Chillemi and co-workers[171] have prepared two peptides corresponding to residues (88—124) and (125—156)* and both showed a low, but significant, activity in the rat tibia growth assay (previously they had found similar activity in analogous peptides from an incorrect HGH sequence[172]). Blake and Li[173] have found some growth-promoting activity in the tibia test with the N^{α}-acetyl derivative of synthetic HGH-(95—136). Niall *et al.*, on the other hand, were unable to provoke significant biological activity with synthetic HGH-(95—133), but stressed the preliminary nature of the data and the problems of the assay system.[155] No significant activity was detected[174] in the rat tibia test with synthetic HGH-(95—124)-amide or its N^{α}-benzyloxycarbonyl derivative (this peptide coincides with the overlap region of the active products obtained by other workers).

[167] K. Uthne, *Acta Endocrinol.*, 1973, **73**, Suppl. 175, 3.
[168] M. H. Connors, S. L. Kaplan, C. H. Li, and M. M. Grambach, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 499.
[169] N. Yamasaki, K. Kangawa, S. Kobayashi, M. Kikutani, and M. Sonenberg, *J. Biol. Chem.*, 1972, **247**, 3874.
[170] L. S. Levine, M. Sonenberg, and M. I. New, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 607.
[171] F. Chillemi, A. Aiello, and A. Pecile, *Nature New Biol.*, 1972, **238**, 243; F. Chillemi, *Gazzetta*, 1973, **103**, 657.
[172] F. Chillemi and A. Pecile, *Experientia*, 1971, **27**, 385.
[173] J. Blake and C. H. Li, *Internat. J. Peptide Protein Res.*, 1973, **5**, 123.
[174] R. Camble, R. Cotton, and S. E. Jaggers, unpublished data.

* The numbering has been corrected to conform with the present 191 amino-acid residue sequence for HGH.

The properties, identification, and synthesis of amino- and carboxy-terminal peptides of HGH possessing, respectively, insulin-like (cataglykin) and insulin-antagonistic (somantin) properties (both regions are devoid of growth-promoting activity) have been comprehensively reviewed.[154] A pituitary peptidase that is believed to hydrolyse ovine growth hormone, specifically, to release a peptide with the activity of somantin has recently been identified.[175] If synthetic studies in other laboratories of the amino-terminal decapeptide[176] and the carboxy-terminal cyclic dodecapeptide[177] sequences of HGH are extended, independent evaluation of these activities should be forthcoming.

Oxytocin and Vasopressin.—Reviews have appeared on the biological activities of synthetic analogues of oxytocin and vasopressin,[178] radio-immuno-assay of posterior pituitary peptides,[179] and the action of vaso-pressin on the gastro-intestinal tract.[180]

A conformation for oxytocin involving two β-turns – one in the ring and one in the side-chain – was proposed by Urry and Walter (see ref. 1). ^{13}C N.m.r. studies with oxytocin and 8-lysine-vasopressin in dimethyl-sulphoxide–trifluoroethanol support the existence of a β-turn conforma-tion in the ring portion of the molecules, stabilized by a hydrogen bond between the peptide NH of the 5-asparagine residue and the carbonyl oxygen of the 2-tyrosine residue.[181] There is also an indication of an additional hydrogen bond between the peptide NH of the 2-tyrosine residue and the carbonyl oxygen of the 5-asparagine residue.[181] However, Brewster *et al.* have suggested that oxytocin is extremely flexible in aqueous solution,[182] and could have three favoured conformations of equivalent energy in DMSO:[183] (*a*) a 1—4 β-turn structure in which the 5-asparagine peptide NH is hydrogen-bonded to the 2-tyrosine carbonyl group (as proposed above), (*b*) a structure in which the 5-asparagine peptide NH is hydrogen-bonded to the 4-glutamine carboxamide carbonyl group (in [4-glycine]-oxytocin this hydrogen bond cannot exist and a conformation is suggested with hydrogen bonds between the 5-asparagine peptide NH and carbonyl groups and the 2-tyrosine carbonyl and peptide NH groups, respectively), and (*c*) a 1—5 turn structure in which the 5-asparagine peptide NH is hydrogen-bonded to the 1-half-cystine carbonyl group.

[175] C. J. Driver, J. McD. Armstrong, J. Bornstein, and F. M. Ng, *J. Endocrinol.*, 1973, **59**, 261.
[176] E. W. B. de Leer and H. C. Beyerman, *Rec. Trav. chim.*, 1973, **92**, 174; H. C. Beyerman, E. W. B. de Leer, and J. Floor, *ibid.*, p. 481.
[177] D. Yamashiro, R. L. Noble, and C. H. Li, *J. Org. Chem.*, 1973, **38**, 3561.
[178] W. H. Sawyer and M. Manning, *Ann. Rev. Pharmacol.*, 1973, **13**, 5; I. L. Schwartz, in 'The Chemistry of Polypeptides', ed. P. G. Katsoyannis, Plenum Publishing Corporation, New York, 1973.
[179] A. G. Robinson and A. G. Frantz, *Metabolism*, 1973, **22**, 1047.
[180] H. Schapiro and L. G. Britt, *Digestive Diseases*, 1972, **17**, 649.
[181] R. Walter and J. D. Glickson, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 1199.
[182] A. I. R. Brewster and V. J. Hruby, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 3806.
[183] A. I. R. Brewster, V. J. Hruby, J. A. Glasel, and A. E. Tonelli, *Biochemistry*, 1973, **12**, 5294.

No evidence was found for a hydrogen bond involving the *C*-terminal tripeptide amide but it was suggested that this sequence might have a *trans-cis* junction with the ring and be folded towards it (in [7-D-proline]-oxytocin the cysteinyl–proline peptide bond is *trans*,[184] but the *C*-terminal tripeptide amide is extended away from the ring).[183] [2-Valine]-oxytocin, which lacks the aromatic side-chain of the tyrosine residue, appears to assume a different ring conformation, with a buried 3-isoleucine peptide NH which may be weakly hydrogen-bonded to a sulphur atom.[183] In D_2O, the ^{13}C n.m.r. spectra of 8-arginine-vasopressin and 8-lysine-vasopressin are essentially identical but different from that of oxytocin; 8-arginine-vasotocin has an intermediate spectrum.[185] Details have appeared of earlier reports (see ref. 1, p. 414, Table 7) on the activities of ring structures related to oxytocin (137) and vasopressin, and some additional pharmacological effects have also been reported (Table 5).[186, 187] The uterotonic activity of analogues (139) and (141) is increased on deamination [compounds (140) and (142)], and the activity of all four analogues is blocked by *N*-carbamoyl-2-*O*-methyltyrosine-oxytocin, indicating that the effect is not non-specific.[186] The antidiuretic activity of the vasopressin ring peptide (141) is much lower than that of 8-lysine-vasopressin (138).[186] The antidiuretic activity of tocinamide is in some doubt – in one report it is similar [186] and in the other [187] much lower than that of oxytocin.

Oxytocin ring peptides (139), (140) have higher natriuretic (urinary sodium excretion) activity than vasopressin ring analogues (141)—(144), but even the most active analogue, deaminotocinamide (140), is less potent than oxytocin.[187] The oxytocin ring analogues show very high hydro-osmotic (water-flow along an osmotic gradient) and natriferic (active sodium ion transport) activities, and the maximum response induced by tocinamide (139) is comparable to that of oxytocin.[187] The vasopressin ring analogues have very low hydro-osmotic and natriferic activities and the deamino analogues of tocinamide and pressinamide have lower activities than the parent analogues.[187]

The natriuretic activity of oxytocin in mammals is increased in analogues where the glutamine residue in position 4 is replaced by leucine (145); several new analogues of this type have been prepared (Table 6). [4-Leucine]-arginine-vasotocin (146) [188] gave a three-fold increase in sodium excretion at a dose of 100 $\mu g\,kg^{-1}$ (s.c.) in saline-loaded rats, and a signi-

[184] A. I. R. Brewster, V. J. Hruby, A. F. Spatola, and F. A. Bovey, *Biochemistry*, 1973, **12**, 1643.
[185] R. Walter, K. U. M. Prasad, R. Deslauriers, and I. C. P. Smith, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2086.
[186] I. Krejčí, T. Barth, B. Kupková, L. Fruhaufová, M. Flegel, and M. Zaoral, *European J. Pharmacol.*, 1973, **24**, 179.
[187] W. Y. Chan, P. J. S. Chiu, and W. H. Sawyer, *Proc. Soc. Exp. Biol. Med.*, 1973, **143**, 954.
[188] D. Gillessen, R. O. Studer, and J. Rudinger, *Experientia*, 1973, **29**, 170.

Table 5 *Biological activity of ring structures related to oxytocin and vasopressin (from refs.* 186 *and* 187)

Compound number		*Structure*	*Biological activity* [a]			
			Isolated rat uterus	*Anti-diuretic*	*Toad urinary bladder Hydro-osmotic*	*Natriferic*
(137)	Oxytocin	Cys -Tyr-Ile -Gln-Asn-Cys-Pro-Leu-Gly-NH_2	450	3	450	450
(138)	8-Lysine-vasopressin	Cys -Tyr-Phe-Gln-Asn-Cys-Pro-Lys-Gly-NH_2	5	260	—	—
(139)	Tocinamide	Cys -Tyr-Ile -Gln-Asn-Cys-NH_2	2.5	1.3	232	320
(140)	1-Deamino-tocinamide	βMP-Tyr-Ile -Gln-Asn-Cys-NH_2 [b]	16.4	0.21	130	141
(141)	Pressinamide	Cys -Tyr-Phe-Gln-Asn-Cys-NH_2	0.25	0.49	3.84	6.63
(142)	1-Deamino-pressinamide	βMP-Tyr-Phe-Gln-Asn-Cys-NH_2	1.03	0.50	1.50	3.38
(143)	Pressinoic acid	Cys -Tyr-Phe-Gln-Asn-Cys-OH	—	—	0.70	1.17
(144)	1-Deamino-pressinoic acid	βMP-Tyr-Phe-Gln-Asn-Cys-OH	—	—	0.10	0.07

[a] Activities are expressed in terms of International Units per milligram; [b] βMP = β-mercaptopropionic acid.

Table 6 *4-Leucine analogues of vasotocin, oxytocin, and vasopressin*

Compound number	*Structure*	*Ref.*
(145)	Cys -Tyr-Ile -Leu-Asn-Cys-Pro-Leu-Gly-NH_2	
(146)	Cys -Tyr-Ile -Leu-Asn-Cys-Pro-Arg-Gly-NH_2	188, 189, 193
(147)	Cys -Tyr-Ile -Leu-Asn-Cys-Pro-Lys-Gly-NH_2	190
(148)	Cys -Tyr-Leu-Leu-Asn-Cys-Pro-Arg-Gly-NH_2	190
(149)	βMP-Tyr-Ile -Leu-Asn-Cys-Pro-Arg-Gly-NH_2	190
(150)	βMP-Tyr-Ile -Leu-Asn-Cys-Pro-Lys-Gly-NH_2	190
(151)	βMP-Tyr-Leu-Leu-Asn-Cys-Pro-Arg-Gly-NH_2	190
(152)	Cys -Phe-Ile -Leu-Asn-Cys-Pro-Arg-Gly-NH_2	190
(153)	Cys -Tyr-Phe-Leu-Asn-Cys-Pro-Arg-Gly-NH_2	190
(154)	Cys -Tyr-Ile -Phe-Asn-Cys-Pro-Arg-Gly-NH_2	190
(155)	Cys -Ile -Ile -Leu-Asn-Cys-Pro-Leu-Gly-NH_2	191
(156)	Cys -Phe-Ile -Leu-Asn-Cys-Pro-Leu-Gly-NH_2	191
(157)	Cys -Tyr-Phe-Leu-Asn-Cys-Pro-Lys-Gly-NH_2	192
(158)	βMP-Tyr-Phe-Leu-Asn-Cys-Pro-Lys-Gly-NH_2	192
(159)	Gly-Cys -Tyr-Ile -Leu-Asn-Cys-Pro-Arg-Gly-NH_2	190

ficant increase in cats and dogs.[189] Other vasotocin analogues (147)—(151) approximately doubled total sodium excretion at a dose of 30 $\mu g\,kg^{-1}$ (i.v.).[190] Peptides (152)—(158) had very little natriuretic activity,[190–192] and the N^{α}-glycyl-derivative (159), synthesized as a 'hormonogen', had a low but considerably protracted activity.[190] Generally, these 4-leucine substituted peptides showed decreased oxytocic and vasopressin effects. The activities of peptide (146) were: pressor (6.1 ± 0.4 I.U. mg^{-1}), antidiuretic (3.2 ± 0.5 I.U. mg^{-1}), and natriferic (5.85 ± 0.35 I.U. mg^{-1}); on isolated rat uterus in Mg^{2+}-free medium, it had no contractile action, but inhibited the contractile response to oxytocin, and this inhibition persisted for some time after the analogue had been washed out.[193] Peptides (155) and (156) had a weak pressor activity (0.5 I.U. mg^{-1}) and no antidiuretic or oxytocic activity.[191] Compounds (157) and (158) had low

[189] J. H. Cort, K. M. Strub, G. Häusler, and J. Rudinger, *Experientia*, 1973, **29**, 173.
[190] K. M. Strub, J. H. Cort, A. Daum, D. Gillessen, A. Trzeciak, R. O. Studer, and J. Rudinger, *Experientia*, 1973, **29**, 767.
[191] V. J. Hruby, F. Muscio, C. M. Groginsky, P. M. Gitu, D. Saba, and W. Y. Chan, *J. Medicin. Chem.*, 1973, **16**, 624.
[192] D. F. Dyckes, M. F. Ferger, V. du Vigneaud, and W. Y. Chan, *J. Medicin. Chem.*, 1973, **16**, 843.
[193] V. Pliška, J. Vašák, M. Rufer, and J. Rudinger, *Experientia*, 1973, **29**, 171.

Table 7 *Biological activities of vasopressins, vasotocins, 4-Thr-vasopressins, 4-Thr-vasotocins, and their deamino-analogues (from ref. 195)*

Compound number	*Structure*	*Rat uterus*					*Ratio antidiuretic : vasopressor*
		$-Mg^{2+}$	0.5 mmol l^{-1} Mg^{2+}	*Fowl depressor*	*Antidiuretic*	*Vasopressor*	
(160)	Cys -Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly-NH_2	13.9 ± 0.5	25.5 ± 0.6	105 ± 5	323 ± 16	369 ± 6	0.9
(161)	βMP-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly-NH_2	47 ± 3	66 ± 3	150 ± 4	1390 ± 136	370 ± 20	3.8
(162)	Cys -Tyr-Phe-Thr-Asn-Cys-Pro-Arg-Gly-NH_2	18.6 ± 0.4	65 ± 3	321 ± 15	231 ± 29	104 ± 2	2.2
(163)	βMP-Tyr-Phe-Thr-Asn-Cys-Pro-Arg-Gly-NH_2	10.2 ± 0.5	19.7 ± 1.8	195 ± 9	758 ± 50	30 ± 1	25
(138)	Cys -Tyr-Phe-Gln-Asn-Cys-Pro-Lys-Gly-NH_2	10.1 ± 0.3	22.1 ± 0.9	52 ± 6	284 ± 39	270 ± 15	1.1
(164)	βMP-Tyr-Phe-Gln-Asn-Cys-Pro-Lys-Gly-NH_2	12 ± 0.5	—	61 ± 2	301 ± 11	126 ± 2	2.4
(165)	Cys -Tyr-Phe-Thr-Asn-Cys-Pro-Lys-Gly-NH_2	12.9 ± 0.4	25.1 ± 0.8	267 ± 5	155 ± 17	49 ± 2	3.2
(166)	βMP-Tyr-Phe-Thr-Asn-Cys-Pro-Lys-Gly-NH_2	2.7 ± 0.2	7.5 ± 0.3	195 ± 9	544 ± 62	10 ± 0.4	54
(167)	Cys -Tyr-Ile -Gln-Asn-Cys-Pro-Arg-Gly-NH_2	127 ± 9	194 ± 15	493 ± 17	231 ± 30	160 ± 4	1.4
(168)	βMP-Tyr-Ile -Gln-Asn-Cys-Pro-Arg-Gly-NH_2	251 ± 12	206 ± 10	1174 ± 24	890 ± 100	256 ± 6	3.5
(169)	Cys -Tyr-Ile -Thr-Asn-Cys-Pro-Arg-Gly-NH_2	201 ± 10	260 ± 21	831 ± 27	279 ± 25	106 ± 25	2.6
(170)	βMP-Tyr-Ile -Thr-Asn-Cys-Pro-Arg-Gly-NH_2	63 ± 2	82 ± 3	786 ± 40	348 ± 33	33 ± 1	11

pressor and no oxytocic activity, and were one-tenth as strong inhibitors of oxytocin as [1-deaminopenicillamine]-oxytocin in an isolated rat uterus test.[192] The low antidiuretic potencies of compounds (157) and (158) (1—2 I.U. mg^{-1} and 5—6 I.U. mg^{-1}, respectively) were in marked contrast to the extremely high potencies (707 and 729 I.U. mg^{-1}, respectively) of [4-aminobutyric acid]-lysine-vasopressin and its 1-deamino-analogue (compounds which differ from the 4-leucine analogues only in the absence of two methyl groups in the 4 position).[192]

Substitution of threonine for glutamine in position 4 of oxytocin and mesotocin (8-isoleucine-oxytocin) increased the oxytocin-like activities and decreased the vasopressin-like activities of these hormones; in contrast to oxytocin, where the deamino-analogue shows enhanced biological activity, the deamino-analogues of [4-threonine]-oxytocin and [4-threonine]-mesotocin showed reduced biological activity (see ref. 1). Now, the corresponding 4-threonine analogues of the vasopressins and vasotocin have been described (Table 7), and again, in general, oxytocin-like properties are increased and vasopressin-like activities are diminished.[194, 195] As in the oxytocin series, the deamino-analogues (161), (164), and (168) of 8-arginine-vasopressin (160), 8-lysine-vasopressin (138), and 8-arginine-vasotocin (167) show increased biological activity. The deamino-analogues (163), (166), and (170) of the corresponding 4-threonine-substituted hormones (162), (165), and (169) show diminished potencies in the rat uterus, fowl vasodepressor, and rat vasodepressor assay systems but substantially increased antidiuretic activity.

The hydro-osmotic activity of 8-arginine-vasotocin (167), deamino-8-arginine-vasotocin (168), and oxytocin (137) is reduced by substitution of threonine in position 4 (Table 8). From these results it has been suggested that receptors which mediate the antidiuretic response are of a more lipophilic nature than those which mediate either the pressor or the bullfrog bladder hydro-osmotic response.

In addition to the 4-threonine analogues mentioned above, several other analogues of 8-arginine-vasopressin (160) with increased antidiuretic to pressor selectivity have been reported. Structural alterations which result in enhancement of the antidiuretic to pressor ratio (A/P) include (*a*) removal of amino-groups from position 1 [1-deamino-8-arginine-vasopressin (161), A/P 3.8]; (*b*) replacement of the 2-tyrosine residue with phenylalanine ([2-phenylalanine]-8-arginine-vasopressin, A/P 2.9); (*c*) substitution of the 4-glutamine residue by a more lipophilic residue {[4-threonine]-8-arginine-vasopressin (162), A/P 2.2, and [4-α-aminobutyric acid]-8-arginine-vasopressin, A/P 20}; and (*d*) substitution of D-arginine for L-arginine in position 8 ([8-D-arginine]-vasopressin, A/P 28). 1-Deamino-[8-D-arginine]-vasopressin, combining structural features (*a*) and

[194] M. Manning, E. J. Coy, W. H. Sawyer, and M. Acosta, *J. Medicin. Chem.*, 1973, **16**, 463.

[195] M. Manning, E. J. Coy, M. Acosta, and W. H. Sawyer, *J. Medicin. Chem.*, 1973, **16**, 836.

Table 8 *Antidiuretic and hydro-osmotic activities of* 8-*Arg-vasotocin and oxytocin, their* 4-*threonine-, deamino-, and deamino-*4-*threonine-analogues* (*from ref.* 195)

		Activities/I.U. mg^{-1}		
Compound number	*Analogue*	*Rat antidiuretic*	*Bullfrog bladder hydro-osmotic*	*Ratio hydro-osmotic: antidiuretic*
(167)	8-Arg-vasotocin	231 ± 30	59 000 ± 7900	255
(168)	[1-Deamino]-8-Arg-vasotocin	890 ± 100	35 000 ± 3100	39
(169)	[4-Thr]-8-Arg-vasotocin	279 ± 25	3 240 ± 280	12
(170)	[1-Deamino, 4-Thr]-8-Arg-vasotocin	348 ± 33	415 ± 34	1.2
(137)	Oxytocin	4 ± 0.8	520	130
(171)	Deamino-oxytocin	19	72	4
(172)	[4-Thr]-oxytocin	1.8 ± 0.3	13	7

(*d*), has A/P 2000,[196] and after a single intranasal dose produces antidiuresis for up to 20 h in patients with cranial diabetes insipidus (the same dose of 8-lysine-vasopressin produces a weaker antidiuretic effect lasting only 3—4 h).[197] Recently, 1-deamino-[4-valine, 8-D-arginine]-vasopressin has been synthesized and found to be a highly selective antidiuretic agent, A/P 125 000.[196]

Other vasopressin analogues reported this year include 1-deamino-[8-N^{ε}-toluene-*p*-sulphonyl-lysine]-vasopressin (negligible rat pressor activity),[198] 1-deamino-[9-ethylenediamine]-8-lysine-vasopressin (0.002 I.U. mg^{-1} rat pressor activity),[199] [8-homonorleucine]-vasopressin (8% rat pressor and 4% antidiuretic potency of 8-lysine-vasopressin),[200] and glycylglycylglycyl-8-lysine-vasopressin (a 'hormonogen' converted *in vivo* into 8-lysine-vasopressin and giving an antidiuretic effect which rose to a maximum in 30—50 min and was still detectable after 2 h).[201] The specifically radio-labelled derivatives [9-^{14}C-glycinamide]-8-lysine-vasopressin,[202] [2-^{3}H-tyrosine]-8-lysine-vasopressin,[203] and [3-^{3}H-phenylalanine]-8-arginine-vasopressin [204] have been synthesized. In oxytocin, replacement of the carbonyl group in the *C*-terminal amide of 1-deamino-oxytocin with a thiocarbonyl group gave an analogue, 1-deamino-[9-thioglycine]-oxytocin, with only 6% of the oxytocic and 1.5% of the avian vasodepressor activity of 1-deamino-oxytocin;[205] [2-pentafluorophenylalanine]-oxytocin,[4c, 206] [2-phenylglycine]-oxytocin,[4c] [2-*p*-ethoxy-D-phenylalanine]-oxytocin,[4c] [2-tryptophan]-oxytocin,[4c] and [2-(3-nitrotyrosine)]-oxytocin [207] all showed very low uterotonic activity (0.01, 0.01, 3.3, 0.24, and 1.1 I.U. mg^{-1}, respectively); [2-D-pentafluorophenylalanine]-oxytocin,[4c, 206] [2-D-phenylalanine]-oxytocin,[4c, 206] and [2-D-tryptophan]-oxytocin [4c] inhibited the action of oxytocin on the rat uterus.

The inhibitory properties of a number of oxytocin derivatives have been compared.[208] *N*-Acetyl[2-*O*-acetyltyrosine]-oxytocin, *N*-carbamoyl[2-*O*-acetyltyrosine]-oxytocin, and *N*-carbamoyl[2-*O*-carbamoyltyrosine]-oxytocin were significant oxytocin antagonists, but the phenol–ester linkage

[196] M. Manning, L. Balaspiri, M. Acosta, and W. H. Sawyer, *J. Medicin. Chem.*, 1973, **16**, 975.
[197] C. R. W. Edwards, M. J. Kitau, T. Chard, and G. M. Besser, *Brit. Med. J.*, 1973, **3**, 375.
[198] R. T. Havran, *Experientia*, 1973, **29**, 520.
[199] J. D. Glass and V. du Vigneaud, *J. Medicin. Chem.*, 1973, **16**, 160.
[200] M. L. Fink and M. Bodanszky, *J. Medicin. Chem.*, 1973, **16**, 1324.
[201] V. Pliška, T. Chard, J. Rudinger, and M. L. Forsling, *Experientia*, 1973, **29**, 756.
[202] E. J. Neer, *J. Biol. Chem.*, 1973, **248**, 1897.
[203] J. Vaněčková and T. Barth, *Coll. Czech. Chem. Comm.*, 1973, **38**, 2008.
[204] H. H. Holton, L. A. Branda, and B. M. Ferrier, *Canad. J. Chem.*, 1973, **51**, 1910.
[205] W. C. Jones, jun., J. J. Nestor, jun., and V. du Vigneaud, *J. Amer. Chem. Soc.*, 1973, **95**, 5677.
[206] O. A. Kaurov, V. F. Martynov, and M. P. Smirnova, *Zhur. obshchei Khim.*, 1973, **43**, 217.
[207] D. B. Hope and M. Wälti, *Biochem. J.*, 1973, **135**, 241.
[208] I. Krejčí, B. Kupková, T. Barth, and K. Jost, *Physiologia Bohemoslovaca*, 1973, **22**, 315.

was unstable under physiological conditions (the di-carbamoyl derivative slowly lost its inhibitory effect and developed a slight uterotonic activity). *N*-Carbamoyl[2-*O*-methyltyrosine]-oxytocin and *N*-acetyl[2-*O*-methyltyrosine]-oxytocin were stable inhibitors of the uterine effects of oxytocin, and the latter compound was also shown to inhibit the action of oxytocin on the mammary gland, and the rat pressor action of 8-lysine-vasopressin.[208]

Contrary to conclusions drawn from rat blood-pressure measurements, studies *in vitro* with isolated rat aorta and *in vivo* with rat mesenteric arterioles suggest that the length as well as the basicity of the amino-acid side-chain in position 8 may be an important determinant of the vasoconstrictor potency of vasopressin analogues (Table 9).[209] The microscopic

Table 9 *Effect of the structure of the side-chain in 8-substituted vasopressin analogues on contractile activity in rat blood vessels* [a]

Amino-acid residue in position 8	*Arteriole*		*Aorta*	
	Affinity [b]	*% maximal response*	*Affinity* [b]	*% maximal response*
Arginine	0.7 ± 0.05	85 ± 6.0	3.5 ± 0.2	100 ± 3.0
Ornithine	6.3 ± 0.4	100 ± 0.0	1.6 ± 0.2	99 ± 4.0
Lysine	9.0 ± 0.5	85 ± 2.0	2.9 ± 0.3	80 ± 5.0
Homolysine	5.6 ± 0.9	84 ± 2.0	3.0 ± 0.4	81 ± 4.0
Formyl-lysine	280 ± 32.0	80 ± 8.0	54 ± 2.3	98 ± 5.0

[a] Ref. 209; [b] Peptide concentration for half-maximum response (nmol l^{-1}).

arterioles do show greatest affinity for 8-arginine-vasopressin (which has the most basic side-chain), but give a higher maximal response with 8-ornithine-vasopressin, and the isolated aorta shows greatest affinity for 8-ornithine-vasopressin. Increasing the length of the side-chain in position 8 without significantly changing the basicity (compare 8-ornithine-vasopressin with 8-lysine- and 8-homolysine-vasopressin) decreases the maximal response of, and, with one exception, the affinity for, both types of blood vessel. [8-Diaminobutyric acid]-vasopressin, which has a shorter side-chain than the ornithine analogue, has an affinity similar to 8-lysine-vasopressin but gives a lower maximal contractile response. The very high maximal response which can be achieved with [8-N^{ε}-formyl-lysine]-vasopressin demonstrates that a basic side-chain in position 8 is not an absolute requirement for the vasoconstrictor action of vasopressin.[209]

Binding studies with tritiated oxytocin, [4-proline]-oxytocin, [4-threonine]-oxytocin, 8-lysine-vasopressin, and [8-valine]-vasopressin indicate the presence of specific receptors on uterus,[210] mammary gland,[211, 212]

[209] B. M. Altura, *Proc. Soc. Exp. Biol. Med.*, 1973, **142**, 1104.
[210] M. S. Soloff, T. L. Swartz, M. Morrison, and M. Saffran, *Endocrinology*, 1973, **92**, 104.
[211] M. S. Soloff, T. L. Swartz, and M. Saffran, *Endocrinology*, 1972, **91**, 213.
[212] M. S. Soloff and T. L. Swartz, *J. Biol. Chem.*, 1973, **248**, 6471.

and frog bladder epithelial cells.[213] The affinities of analogues were proportional to their biological activities. Vasopressin is known to release ACTH from the anterior pituitary, but there is much evidence that it is chemically distinct from the physiological corticotropin-releasing factor (CRF). The release of ACTH by vasopressin and several analogues has been studied on isolated pituitary cells (Table 10).[214] The CRF-like

Table 10 *ACTH-releasing activity of oxytocin and vasopressin analogues (from ref.* 214)

Compound	*Dose*/μg	ACTH (% *control*)
8-Arginine-vasopressin	0.25	154
	1.0	160
	10.0	158
8-Lysine-vasopressin	0.25	143
	1.0	148
[8-Homolysine]-vasopressin	0.05	161
	0.10	155
	1.0	158
1-Deamino-[8-homolysine]-vasopressin	0.5	162
	1.0	171
	10.0	142
Oxytocin	0.20	132
	1.0	146
	10.0	147
8-Arginine-vasotocin	1.0	160
	2.5	176
	10.0	150
Glycylglycylglycyl-8-lysine-vasopressin	—	Inactive
Pro-Leu-Gly-NH_2	—	Inactive
Pro-Arg-Gly-NH_2	—	Inactive
ACTH (4—10)-heptapeptide	—	Inactive
N-Carbamoyl-[2-*O*-methyl-tyrosine]-oxytocin	—	Inactive

activity of these peptides does not correspond to their relative pressor or antidiuretic activities, and the maximum rate of release of ACTH (< 200%) was significantly less than that induced by hypothalamic median eminence extracts (359 ± 29%).

Neurophysins, the low-molecular-weight proteins which act as physiological 'carriers' of neurohypophyseal hormones within the secretory neurons of the hypothalamo-neurohypophyseal system, have been extracted from posterior pituitary glands of mammals and other vertebrates. Although the specificity of the hormone-binding sites of neurophysins for oxytocin and vasopressin differs from that of the hormone receptors, their study provides useful information about peptide–protein binding interactions. Neurophysins I and II bind oxytocin and vasopressin with

[213] C. Roy, J. Bockaert, R. Rajerison, and S. Jard, *F.E.B.S. Letters*, 1973, **30**, 329.
[214] R. Portanova and G. Sayers, *Proc. Soc. Exp. Biol. Med.*, 1973, **143**, 661.

similar affinities; at pH 5.7 they seem to have one binding site per molecule for oxytocin, and perhaps two binding sites for 8-lysine-vasopressin (if a second binding site exists it may have a much lower binding affinity).[215, 216] Belief that the principal binding sites on the hormones are located in the *N*-terminal tripeptide region is supported by the small effect that changes in positions 4, 5, 8, and 9 have on the binding of oxytocin and vasopressin,[217] and by the great freedom of the *C*-terminal tripeptide amide portion of oxytocin in the hormone–neurophysin complex as suggested by n.m.r. studies.[218] C.d., proton titration, and n.m.r. data on the neurophysin complexes with oxytocin, vasopressin, and several di- and tri-peptides structurally related to their *N*-terminal regions indicate that residues 2 and 3 are bound to a hydrophobic region of the protein and that almost two-thirds of the binding free energy of the hormones is contributed by the first three residues.[216, 219, 220] Analogues of 8-lysine-vasopressin in which the peptide chain is extended at the amino end by one leucine, one phenylalanine, or one, two, or three glycine residues do not complex with neurophysin, suggesting a possible important binding role for the *N*-terminal α-amino-group.[221]

4 Pancreatic Hormones

Insulin.—Two new total syntheses of insulin have been reported,[222, 223] and the syntheses of human pro-insulin C-peptide [224, 225] and of its (9-Glu, 11-Gln) analogue [224] have been completed. Several derivatives of native insulin, in which the α-amino-group of the 1-glycine residue of the A-chain is intramolecularly cross-linked with the ε-amino-group of the 29-lysine residue of the B-chain, have been prepared using dicarboxylic acid reagents. After conversion into the corresponding hexa-*S*-sulphonates, the disulphide bridges of derivatives cross-linked with adipic,[226] suberic,[227] or

[215] M. Camier, R. Alazard, P. Cohen, P. Pradelles, J. Morgat, and P. Fromageot, *European J. Biochem.*, 1973, **32**, 207.
[216] E. Breslow, J. Weis, and C. J. Menendez-Botet, *Biochemistry*, 1973, **12**, 4644.
[217] E. Breslow and R. Walter, *Mol. Pharmacol.*, 1972, **8**, 75.
[218] J. A. Glasel, V. J. Hruby, J. F. McKelvy, and A. F. Spatola, *J. Mol. Biol.*, 1973, **79**, 555.
[219] J. H. Griffin, R. Alazard, and P. Cohen, *J. Biol. Chem.*, 1973, **248**, 7975.
[220] P. Balaram, A. A. Bothner-By, and E. Breslow, *Biochemistry*, 1973, **12**, 4695.
[221] I. Kluh, E. Sedláková, T. Barth, and J. H. Cort, *Mol. Pharmacol.*, 1973, **9**, 414.
[222] N. A. Yudaev, Yu. P. Shvachkin, M. G. Poznyak, V. P. Fedotov, R. G. Vdovina, E. N. Voluiskaya, M. N. Ryabtsev, V. F. Krivtsov, A. K. Gracheva, S. P. Krasnoshchekova, V. A. Novoselov, V. S. Gruzdev, A. M. Oleinik, Z. B. Kalinkina, and A. I. Ivanova, *Biokhimiya*, 1973, **38**, 221.
[223] Yu. P. Shvachkin, R. G. Vdovina, M. G. Poznyak, E. N. Voluiskaya, M. N. Ryabtsev, V. F. Krivtsov, A. K. Gracheva, S. P. Krasnoshchekova, V. A. Novoselov, V. S. Gruzdev, A. M. Oleinik, Z. B. Kalinkina, V. P. Fedotov, A. I. Ivanova, and N. A. Yudaev, *Zhur. obshchei Khim.*, 1973, **43**, 216.
[224] R. Geiger, G. Jäger, and W. König, *Chem. Ber.*, 1973, **106**, 2347.
[225] V. K. Naithani, *Z. physiol. Chem.*, 1973, **354**, 659.
[226] D. Brandenburg and A. Wollmer, *Z. physiol. Chem.*, 1973, **354**, 613.
[227] S. M. L. Robinson, I. Beetz, O. Loge, D. G. Lindsay, and K. Lubke, *Tetrahedron Letters*, 1973, 985.

succinic[227] acid can be re-formed in good yield, giving products with almost unchanged chemical, spectral, and biological properties. Thus these simple, non-peptide, inter-chain cross-links seem able to mimic the role of the connecting peptide in pro-insulin in presenting correct alignment of the half-cystine residues. The possible application of such derivatives for improving yields in the final chain-combination steps of chemical syntheses of insulin and its analogues requires a cross-link which can be removed selectively from the insulin molecule. Towards this end, α,α'-diaminosuberoyl-insulin was prepared,[228, 229] and successfully reduced and reoxidized and the cross-link (together with the 1-phenylalanine residue of the B-chain) removed by Edman degradation to give crystallizable des-Phe^{B1}-insulin.[228] Separated natural A-chain tetra-*S*-sulphonate, and $N^{\alpha B1}$-trifluoroacetylated natural B-chain di-*S*-sulphonate were also successfully linked together.[229] The correct disulphide bridges were formed in high yield (60—75%) and biologically fully active $N^{\alpha B1}$-trifluoroacetyl-insulin was obtained after Edman degradation of the α,α'-diaminosuberoyl cross-link.[229]

Analogues of sheep insulin in which the intra-chain disulphide bridge between the cysteine residues in positions 6 and 11 of the A-chain has either been eliminated ([A-6,11-alanine]-insulin),[230] or enlarged by two methylene groups ([A-6,11-homocystine]-insulin),[231] have been synthesized and purified. The products possessed 10% and 30%, respectively, of the biological activity of the natural hormone in the mouse convulsion assay. An uncharacterized [A-6,11-alanine]-insulin preparation had previously been reported to have only 0.04—0.2% activity.[232] These results suggest that the intra-chain ring contributes to the establishment of a conformation essential for full activity, but is not functionally involved in the hormone-receptor interaction,[230, 231] a conclusion supported by *X*-ray analysis findings that in the three-dimensional structure of insulin the intra-chain disulphide bridge is completely buried.

When the *C*-terminal tripeptide sequence is removed from the B-chain of natural bovine or porcine insulin by treatment with sodium–liquid ammonia, the resulting des-B(28—30)-insulins (with *C*-terminal alcohol or aldehyde groups) are fully active biologically.[233] Synthetic des-B(27—30)-human insulin possesses about 50% of the biological activity of the native hormone.[234] The *C*-terminal tetrapeptide sequence of the B-chain does not, therefore, seem to be critical to the biological action of insulin

[228] D. Brandenburg, W. Schermutzki, and H. Zahn, *Z. physiol. Chem.*, 1973, **354**, 1521.
[229] R. Geiger and R. Obermeier, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 60.
[230] P. G. Katsoyannis, Y. Okada, and C. Zalut, *Biochemistry*, 1973, **12**, 2516.
[231] A. Cosmatos and P. G. Katsoyannis, *J. Biol. Chem.*, 1973, **248**, 7304.
[232] U. Weber, S. Hörnle, P. Köhler, G. Nagelschneider, K. Eisele, and G. Weitzel, *Z. physiol. Chem.*, 1968, **349**, 512.
[233] P. G. Katsoyannis, C. Zalut, A. Harris, and R. J. Meyer, *Biochemistry*, 1971, **10**, 3884.
[234] P. G. Katsoyannis, J. Ginos, A. Cosmatos, and G. Schwartz, *J. Amer. Chem. Soc.*, 1973, **95**, 6427.

but the absence of a carboxy-group in the proximity of residue 26, or the presence of the side-chain of residue 27, is necessary for full activity. At the other end of the B-chain, des-B(1—3)-[B-4-pyroglutamyl]-insulin, prepared by repeated selective Edman degradation, still possesses high blood-sugar-lowering activity in the rabbit.[235]

Further studies have been made with peptides related to the sequence Arg-Gly-Phe-Phe (residues 22—25) of the insulin B-chain, which was observed to elicit an insulin-like effect *in vitro*. A series of analogues (Table 11) has been assessed *in vivo* by measuring the incorporation of

Table 11 *Peptides related to residues* 20—26 *of the insulin B-chain* (*ref.* 236)

Compound number	*Structure*	
(173)	Arg——Phe-NH_2	
(174)	Arg-Gly-Phe-NH_2	(or -OH, -OMe)
(175)	Arg-Gly-Phe-Phe-NH_2	(or -OH, -OMe)
(176)	Ac-Arg-Gly-Phe-Phe-NH_2	
(177)	Arg-Gly-Phe-Phe-Tyr-Ala-NH_2	(or -OH)
(178)	Glu-Arg-Gly-Phe-Phe-Tyr-Ala-NH_2	(or -OH)
(179)	Glu-Glu-Arg-Gly-Phe-Phe-Tyr-Ala-NH_2	(or -OH)

radioactivity into diaphragm and other tissues after injection of peptide together with [*U*-^{14}C]glucose.[236] The dipeptide amide (173) was inactive but the tri- and tetra-peptide derivatives (174) and (175) stimulated ^{14}C uptake into diaphragm and epididymal fat tissue; the activity decreased in the order amides > esters > free carboxy-compounds. The tetra-peptides were more potent than the tripeptides, and Arg-Gly-Phe-Phe-NH_2 and its *N*-acetyl derivative (176) could increase ^{14}C uptake into diaphragm to a level comparable with maximal stimulation by insulin. The hexapeptide (177) was about as active as the tetrapeptide (175) in the diaphragm and twice as active in epididymal fat tissue. The hepta- and octa-peptide derivatives (178) and (179) were too insoluble to be assayed. The effect of compounds (174) and (175) on rat diaphragm seems to be independent of the insulin receptor since, when pre-incubated with trypsin *in vitro*, diaphragm did not respond to insulin for 20—30 min, but the response to compounds (174) and (175) was unaffected.[236] Thus, any relevance of these interesting observations to the structure–activity properties of insulin itself must remain obscure.

Experiments with fat cell membranes,[237] liver membranes,[238] and isolated hepatocytes [238] from obese hyperglycaemic mice suggest that the increase in insulin resistance in these animals may be related to an impaired

[235] R. Geiger and D. Langner, *Z. physiol. Chem.*, 1973, **354**, 1285.
[236] G. Weitzel, K. Eisele, V. Schulz, and W. Stock, *Z. physiol. Chem.*, 1973, **354**, 321.
[237] P. Freychat, M. H. Laudat, P. Laudat, G. Rosselin, C. R. Kahn, P. Gordon, and J. Roth, *F.E.B.S. Letters*, 1972, **25**, 339.
[238] C. R. Kahn, D. M. Neville, and J. Roth, *J. Biol. Chem.*, 1973, **248**, 244.

insulin-receptor interaction due to a decrease in the number of receptor sites.[238] Insulin bound covalently to agarose has been used to bind, selectively, intact fat cells and their ghosts, presumably through the insulin-receptor interaction, and this could provide a valuable tool in the purification of receptor-containing plasma membranes.[239] The known properties of insulin receptors have been summarized [240, 241] and the receptors on human lymphocytes [242] and rat fat-cell membranes [243] have been further characterized. In each case kinetic data were interpreted as indicating two major classes of binding sites, one of high affinity and low capacity, and the other of low affinity and high capacity. However, more recent kinetic experiments suggest that insulin-filled receptor sites may lower the affinity of other sites for this hormone.[244] Thus, what had been identified as two distinct sites may actually be similar sites participating in negative co-operative interactions. Such site–site interactions seem to be hormone-specific and it is possible that the region of the insulin molecule involved is distinct from the regions involved in binding to the receptor and in producing a biological response.[244] The heterogeneous populations of binding sites postulated for several other hormone-receptor systems may require re-evaluation.

Glucagon.—Studies on the biosynthesis, secretion, and biological effects of glucagon presented at British Diabetic Association medical and scientific section meeting have been published [245] and other reviews on glucagon have also appeared.[246, 247] Evidence has been presented for the existence of a proglucagon in anglerfish [248] and pigeon.[249] Purification of crude bovine and porcine glucagon has given a peptide of high molecular weight (37 residues) containing the sequence, Lys-Arg-Asn-Asn-Lys-Asn-Ile-Ala attached to the *C*-terminus of glucagon; this is thought to be a fragment of proglucagon.[250]

Several derivatives of glucagon have been synthesized and some of them have been found to be highly active. In a hyperglycaemic assay [13-nitrotyrosine]-glucagon and [13-dinitrotyrosine]-glucagon are more active

[239] D. D. Sodermann, J. Germershausen, and H. M. Katzen, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 792.
[240] B. Desbuquois and P. Cuatrecasas, *Ann. Rev. Medicine*, 1973, **24**, 233.
[241] P. Cuatrecasas, *Fed. Proc.*, 1973, **32**, 1838.
[242] J. R. Gavin, P. Gordon, J. Roth, J. A. Archer, and D. N. Buell, *J. Biol. Chem.*, 1973, **248**, 2202.
[243] J. M. Hammond, L. Jarett, I. K. Mariz, and W. H. Daughaday, *Biochem. Biophys. Res. Comm.*, 1972, **49**, 1122.
[244] P. de Meyts, J. Roth, D. M. Neville, J. R. Gavin, and M. A. Lesniak, *Biochem. Biophys. Res. Comm.*, 1973, **55**, 154.
[245] *Postgrad. Med. J.*, 1973, **49**, Supplement 6, 601–623.
[246] P. P. Foa, *Amer. Zoologist*, 1973, **13**, 613.
[247] 'Glucagon: Molecular Physiology, Clinical and Therapeutic Implications', ed. P. J. Lefebvre and R. H. Unger, Pergamon Press, Oxford, 1972.
[248] B. D. Noe and G. E. Bauer, *Proc. Soc. Exp. Biol. Med.*, 1973, **142**, 210.
[249] A. K. Tung, *Hormone Metab. Res.*, 1973, **5**, 416.
[250] H. S. Tager and D. F. Steiner, *Proc. Nat. Aced. Sci. U.S.A.*, 1973, **70**, 2321.

(231% and 145%, respectively) than glucagon, but both analogues are less potent in stimulating adenyl cyclase (77% and 22%, respectively). The corresponding [13-aminotyrosine]-glucagon and [13-diaminotyrosine]-glucagon are less active in the hyperglycaemic (56% and 28%, respectively) and adenyl cyclase (40% and 14%, respectively) assays.[251] Iodination of glucagon enhances both activation of rat liver adenyl cyclase and biological activity as measured by hyperglycaemia in rabbits (adenyl cyclase activity increases about five-fold after mono-iodination and about ten-fold after more extensive iodination).[252] Guanidinated, N^{α}-carbamoyl-, and N^{α}-acetyl-[12-N^{ε}-acetyl-lysine]-glucagon largely retain the ability to stimulate formation of cyclic AMP in rat liver homogenates, but N^{α}-acetyl-[1-N^{im}-ethoxycarbonylhistidine, 12-N^{ε}-acetyl-lysine]-glucagon is inactive.[253] This suggests that the imidazole group of the amino-terminal histidine residue, but not the α- or ε-amino-groups of glucagon, is essential for activity (the critical role for activity of the histidine imidazole group has been reported earlier, see ref. 1). Some fragments of glucagon obtained by cleavage of amino-acids from the carboxy-terminal portion have also been tested for their ability to stimulate adenyl cyclase activity.[254] Fragment 1—27 (with homoserine in place of methionine at position 27), and fragment 1—23 were active although the concentrations required to produce enhanced activity were very high. In another study synthetic fragment 1—23 was inactive in hyperglycaemic and liver adenyl cyclase tests.[255] C.d. and ultracentrifuge studies suggest that fragments 1—27, 1—23, and 1—21 have structures similar to that of the native molecule although somewhat less ordered.[254]

5 Gastrointestinal Hormones

Gastrin.—Derivatives (180)—(185) related to the *C*-terminal tetrapeptide amide (186) of gastrin have been found to stimulate gastric acid secretion in conscious dogs in a dose range of 100—1500 $\mu g\,kg^{-1}\,h^{-1}$.[256] The tetrapeptide ester (180) and the tripeptide amide (182) were, respectively, about 1/25—30 and 1/3000 as potent as compound (186). The methyl

Trp-Met-Asp-Phe-OMe
(180)

Trp-Met(O_2)-Asp-Phe-NH_2
(181)

Met-Asp-Phe-NH_2
(182)

Met-Asp-Phe-OMe
(183)

[251] J. M. Patterson and W. W. Bromer, *J. Biol. Chem.*, 1973, **248**, 8337.
[252] W. W. Bromer, M. E. Boucher, and J. M. Patterson, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 134.
[253] R. M. Epand, R. F. Epand, and V. Grey, *Arch. Biochem. Biophys.*, 1973, **154**, 132.
[254] R. M. Epand and V. Grey, *Canad. J. Physiol. Pharmacol.*, 1973, **51**, 243.
[255] B. Guy-Grand and R. Assan, *Hormone Metab. Res.*, 1973, **5**, 60.
[256] T.-M. Lin and L. Southard, *Fed. Proc.*, 1973, **32** (3, part 1), Abstr. No. 1065, p. 410.

Met-Asp-Phe-OH (184)

Met-Asp(OBzl)-Phe-OMe (185)

Trp-Met-Asp-Phe-NH_2 (186)

Asp-Phe-NH_2 (187)

Asp-Phe-OMe (188)

esters were less potent than the corresponding amides, but this difference was greater in the tetrapeptides (180) and (186) than in the tripeptides (183) and (182). The tripeptides (182)—(184) and the dipeptide derivatives (187) and (188) all stimulated motility of the pylorus, duodenum, and colon in the dog at doses much lower than those required for stimulation of acid secretion. Similar peptide derivatives were reported previously to be inactive,[257] but lower doses (50—100 μg per dog) may have been used.

Replacement of the β-alanine residue in pentagastrin* (189) with L- or D-serine gave analogues (190) and (191) possessing 75% and 62%, respectively, of the potency of pentagastrin (stimulation of acid secretion).[258]

Boc-β-Ala-Trp-Met-Asp-Phe-NH_2 (189)

Boc-L-Ser-Trp-Met-Asp-Phe-NH_2 (190)

Boc-D-Ser-Trp-Met-Asp-Phe-NH_2 (191)

Glp-Gly-Pro-Try-Met-$(Glu)_2$-Ala-Tyr-Gly-Trp-Met-Asp-Phe-NH_2 (192)

Glp-Gly-Pro-Try-Met-$(Glu)_5$-Ala-Tyr-Gly-Trp-Met-Asp-Phe-NH_2 (193)

It was reported some years ago[257] that the shortened gastrin (192) is biologically indistinguishable from porcine gastrin (193) even though it lacks three of the five glutamyl residues. Peptides (194)—(197) (Table 12) have now been synthesized[259] to study further the contribution to biological activity of the carboxy-groups of the glutamic acid residues in positions 9 and 10 of the gastrin sequence. Introduction into the weakly active heptapeptide (194) of a single γ-carboxy-group in the form of a 4-carboxybutyryl residue (195) causes a considerable increase in activity; a second carboxy-group (197) only slightly increases this biological activity. Comparison of compounds (194)—(196) would indicate that much of the increased

[257] H. J. Tracy and R. A. Gregory, *Nature*, 1964, **204**, 935.
[258] M. Portelli and G. Renzi, *Il. Pharmaco*, 1973, **28**, 316.
[259] H. Wissmann, R. Schleyerbach, B. Schoelkens, and R. Geiger, *Z. physiol. Chem.*, 1973, **354**, 1591.

* Pentagastrin, Peptavlon, ICI Ltd. (B.P. 1 042 487).

Table 12 *Relative activities of gastrin analogues*

Compound number	*Structure*	*Relative activity*	
		Weight basis	*Molar basis*
(189)	Boc-β-Ala-Trp-Met-Asp-Phe-NH_2	1.0	1.0
(194)	H-Ala-Tyr-Gly-Trp-Leu-Asp-Phe-NH_2	0.6	0.6
(195)	R-CO-Ala-Tyr-Gly-Trp-Leu-Asp-Phe-NH_2 [a]	2.3	2.6
(196)	H-Glu-Ala-Tyr-Gly-Trp-Leu-Asp-Phe-NH_2	0.7	0.85
(197)	R-CO-Glu-Ala-Tyr-Gly-Trp-Leu-Asp-Phe-NH_2 [a]	2.5	3.3
(198)	[15-Leucine]-human gastrin 1	3.0	7.6

[a] R = $-(CH_2)_3CO_2H$.

potency of the deamino-glutamyl derivative (195) results from its resistance to aminopeptidase digestion. However, comparison of compounds (195) and (197) with pentagastrin, which is also resistant to aminopeptidase, might suggest that a γ-carboxy-group in position 10 of the gastrin molecule does make a significant contribution to biological activity, possibly by increasing the affinity of the molecule for the receptor (in natural bovine and ovine gastrin the glutamic acid residue in position 10 is replaced by alanine [260]).

The value of radio-immunoassay for the study of factors involved in the secretion of gastrin and in the detection and characterization of different molecular forms of gastrin has been summarized.[261, 262] More recently, radio-immunoassay has revealed raised plasma gastrin levels in patients with rheumatoid arthritis [263] and in rats after induction of adjuvant arthritis,[264] suggesting that gastrin may play a role in the inflammatory process.

Intestinal Hormones and Peptides.—The chemistry and pharmacology of secretin and cholecystokinin-pancreozymin (CCK) (and of gastrin) have been extensively reviewed in the past year,[265] as have secretion of gastrointestinal hormones and peptides [266] and the effects of these polypeptides on intestinal secretion.[267] Yet another new peptide, motilin, has been isolated from the upper intestine and shown to contain 22 amino-acid residues in a sequence (199) quite distinct from those of the other intestinal polypeptides.[268] Its most distinguishing action is stimulation of motor

[260] K. L. Agarwal, J. Beacham, P. H. Bentley, R. A. Gregory, G. W. Kenner, R. C. Sheppard, and H. J. Tracy, *Nature*, 1968, **219**, 614.
[261] M. I. Grossman, *Metabolism*, 1973, **22**, 1033.
[262] J. F. Rehfeld, *Scand. J. Gastroenterol.*, 1973, **8**, 577.
[263] P. J. Rooney, J. Vince, A. S. Kennedy, J. Webb, P. Lee, W. C. Dick, K. D. Buchanan, J. R. Hayes, J. Ardill, and F. O'Connor, *Brit. Med. J.*, 1973, **2**, 752.
[264] P. J. Rooney, W. C. Dick, R. C. Imrie, and K. D. Buchanan, *Nature*, 1973, **246**, 497.
[265] *Handbook of Experimental Pharmacology*, 1973, **34**, ed. J. E. Jorpes and V. Mutt, Springer-Verlag, Berlin, 1973; J. Dupre, *Modern Trends in Endocrinology*, 1972, **4**, 278; J. Katz, *Medical Clinics of North America*, 1973, **57**, 893.
[266] S. Andersson, *Ann. Rev. Physiol.*, 1973, **35**, 431.
[267] G. O. Barbezat, *Scand. J. Gastroenterol.*, 1973, **8** (Suppl. 22), 3.
[268] J. C. Brown, M. A. Cook, and J. R. Dryburgh, *Canad. J. Biochem.*, 1973, **51**, 533; H. Schubert and J. C. Brown, *ibid.*, 1974, **52**, 7.

1 10
Phe-Val-Pro-Ile-Phe-Thr-Tyr-Gly-Glu-Leu-Gln-Arg-Met-Gln-Glu-Lys-Glu-
20
Arg-Asn-Lys-Gly-Gln

(199)

activity in pouches of both antral and fundic gland areas of dog stomach A total synthesis of [13-norleucine, 14-glutamic acid]-motilin has been reported briefly;[269] the product exhibited over 90% of the activity of the natural peptide.

The *C*-terminal octapeptide fragment of CCK was estimated to be five times as potent as the structurally similar caerulein decapeptide, ten times as potent as CCK itself and 1800 times as potent as pentagastrin in its ability to contract guinea-pig gall-bladder muscle.[270] All four peptides appeared to act directly on the muscle and not by acetylcholine mediation. CCK and its octapeptide fragment suppress feeding in intact and gastric fistula rats; this property was not shared by the octapeptide with an unsulphated tyrosine residue.[271] It remains to be established whether this satiety effect is a truly physiological function of CCK.

Secretin has been found to stimulate guanyl cyclase activity in homogenates of rat liver; glucagon, which shows extensive sequence homology with secretin (Figure 1), was without effect on this activity either in the

	1	2	3	4	5	6	7	8	9	10	11	12	14	13	15
Secretin	His	Ser	Asp	Gly	Thr	Phe	Thr	Ser	Glu	Leu	Ser	Arg	Leu	Arg	Asp
Glucagon	His	Ser	Gln	Gly	Thr	Phe	Thr	Ser	Asp	Tyr	Thr	Lys	Tyr	Leu	Asp
VIP	His	Ser	Asp	Ala	Val	Phe	Thr	Asp	Asn	Tyr	Thr	Arg	Leu	Arg	Lys
GIP	Tyr	Ala	Glu	Gly	Thr	Phe	Ile	Ser	Asp	Tyr	Ser	Ile	Ala	Met	Asp

	16	17	18	19	20	21	22	23	24	25	26	27	28	29
Secretin	Ser	Ala	Arg	Leu	Gln	Arg	Leu	Leu	Gln	Gly	Leu	Val	$-NH_2$	
Glucagon	Ser	Arg	Arg	Ala	Gln	Asp	Phe	Val	Gln	Trp	Leu	Met	Asp	Thr
VIP	Gln	Met	Ala	Val	Lys	Lys	Tyr	Leu	Asn	Ser	Ile	Leu	Asn	$-NH_2$
GIP	Lys	Ile	Arg	Gln	Gln	Asp	Phe	Val	Asn	Trp	Leu	Leu	Ala	Gln ...[a]

[a] Only 29 residues of GIP are shown.

Figure 1 *The amino-acid sequences of homologous porcine intestinal peptides and pancreatic glucagon*

presence or absence of secretin.[272] Unlike the situation reported previously with a plasma membrane preparation of rat liver,[1] secretin does stimulate adenyl cyclase activity in homogenates of rat liver and to about 50% of the level produced by glucagon.[272] The secretin and glucagon stimulations

[269] E. Wunsch, J. C. Brown, K.-H. Deimer, F. Drees, E. Jaeger, J. Musiol, R. Scharf, H. Stocker, P. Thamm, and G. Wendelberger, *Z. Naturforsch.*, 1973, **28c**, 235.
[270] Y. M. Yau, G. M. Makhlouf, L. E. Edwards, and J. T. Farrar, *Gastroenterology*, 1973, **65**, 451.
[271] J. Gibbs, R. C. Young, and G. P. Smith, *Nature*, 1973, **245**, 323.
[272] W. J. Thompson, R. H. Williams, and S. A. Little, *Biochim. Biophys. Acta*, 1973, **302**, 329.

are additive and it was suggested that secretin activates adenyl cyclase through a fragile receptor distinct from that of glucagon. This ability of secretin to stimulate liver adenyl cyclase activity has now been confirmed with a rat cell membrane preparation.[273] Vasoactive intestinal polypeptide (VIP), another peptide with a structure closely related to those of secretin and glucagon, also stimulates the activity of adenyl cyclase in liver and fat cell membranes.[273] The effects of combinations of VIP with glucagon and secretin at concentrations that maximally activate adenyl cyclase suggest that in adipose tissue all three hormones act on the same enzyme; in liver, VIP activates the same enzyme as secretin and this enzyme is distinct from that responding to glucagon. Studies of competitive inhibition by unlabelled VIP, glucagon, and secretin of the binding of radio-iodinated derivatives of VIP and glucagon to both membrane preparations indicate that VIP and glucagon interact with different receptors.[273] Secretin and VIP, which give essentially the same maximal stimulation of adenyl cyclase activity, appear to compete for a common receptor. However, the apparent affinity of secretin for adenyl cyclase and for the receptor is about two orders of magnitude lower than that of VIP. In liver membranes the maximal enzyme activity achievable with VIP is about one-fifth of that attainable with glucagon, but in fat cell membranes VIP produces about twice the maximal activity generated by glucagon.[273] These relative effects on adenyl cyclase activity correlate with biological activity, VIP being less potent than glucagon in promoting glycogenolysis and hyperglycaemia (60% and 30%, respectively)[274] but more effective than glucagon in stimulating lipolysis.[275]

Synthesis of the complete octacosapeptide amide sequence of VIP has been achieved.[276, 277] Biological activities of the synthetic VIP and of several shorter chain synthetic fragments have been compared with natural VIP in both vasodilator and smooth muscle relaxation assay systems (Table 13).[276] The synthetic material exhibited activity similar to that of natural VIP. Fragments containing the *C*-terminal eleven residues showed weak activity and this increased with increasing chain length; it will be interesting to examine these fragments in tests more characteristic of secretin- and glucagon-like activity (in which fragments of secretin and glucagon have not shown significant activity).

Gastric inhibitory polypeptide (GIP), a forty-three amino-acid residue peptide, which also shows considerable sequence homology with secretin and glucagon in the *N*-terminal region (Figure 1), is the subject of syn-

[273] B. Desbuquois, M. H. Laudat, and P. Laudat, *Biochem. Biophys. Res. Comm.*, 1973, **53**, 1187.
[274] C. Kerins and S. I. Said, *Proc. Soc. Exp. Biol. Med.*, 1973, **142**, 1014.
[275] M. H. Laudat and P. Laudat, quoted in ref. 273.
[276] M. Bodanszky, Y. S. Klausner, and S. I. Said, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 382.
[277] Y. S. Klausner, C. Y. Lin, V. Mutt, and M. Bodanszky, *Bio-organic Chem.*, 1973, **2**, 345; Y. S. Klausner and M. Bodanszky, *ibid.*, 1973, **2**, 354.

Table 13 *Biological activities of synthetic peptides and natural VIP (from ref. 276)*

Effect	*Relative potency*						
	VIP_{1-6}	VIP_{18-28}	VIP_{15-28}	VIP_{14-28}	VIP_{7-28}	VIP_{1-28}	$VIP_{(natural)}$
Vasodilatation Systemic hypotension	0.2	0.05	0.1	1—2	2—10	80—100	100
Tracheal relaxation	0.01	0.05—0.1	0.2	0.1	5—10	200—250 [a]	100
Stomach relaxation	0.04	0	0.1	[b]	10—15	100	100

[a] No explanation has been found for this high activity; [b] This peptide fragment usually induced contraction, followed by relaxation.

thetic studies.[4d, 278] Originally characterized as an inhibitor of gastric acid secretion in dogs, GIP has now been shown to be a potent stimulator of insulin secretion and of glucose clearance in man when given in physiological doses together with glucose by i.v. infusion.[279] Only a small, statistically insignificant, increment in immuno-reactive insulin in serum was observed when the peptide was given to fasted subjects, and there was no detectable change in plasma glucose concentrations. From these results it is possible that GIP is related to the partially purified 'Intestinal Insulin-releasing Polypeptide (IRP)' whose properties correspond to those of the hypothetical intestinal hormone 'Incretin' or 'Glucose-dependent Insulin-releasing Hormone (GIRH)' (invoked to account for the greater insulin secretory response to oral as compared to intravenous glucose).[280]

It has been suggested that the hexapeptide (200), liberated in the duodenum during conversion of bovine pancreatic trypsinogen into the

Val-Asp-Asp-Asp-Asp-Lys
(200)

active enzyme trypsin, may be a physiological regulator of gastric acid secretion.[281] When given by intravenous or intraduodenal infusion this peptide was found significantly to inhibit gastrin- or pentagastrin-stimulated gastric acid secretion. The presence of four adjacent aspartic acid residues in the structure (200) and the presence of four or five contiguous glutamic acid residues in the gastrins led to speculation that a cluster of acidic amino-acid side-chains is important for binding to the gastrin receptor and

[278] K. Kovacs, J. Kovacs-Petres, G. Wendelberger, and E. Wunsch, *Z. physiol. Chem.*, 1973, **354**, 890; J. Kovacs-Petres, K. Kovacs, K.-H. Deimer, and E. Wunsch, *ibid.*, 1973, **354**, 894.

[279] J. Dupre, S. A. Ross, D. Watson, and J. C. Brown, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 826.

[280] D. S. Turner, A. Shabaan, L. Etheridge, and V. Marks, *Endocrinology*, 1973, **93**, 1323.

[281] J. P. Abita, A. Moulin, M. Ladzunski, G. Hage, G. Palasciano, A. Brasca, and O. Tiscornia, *F.E.B.S. Letters*, 1973, **34**, 251.

that the peptide (200) therefore acts as a gastrin antagonist.[281] The idea is attractive since it would link the duodenal activation of pancreatic zymogens, secreted in response to the presence of acid in the duodenum, with the gastric secretion responsible for that acidification.

6 Thyroid and Parathyroid Hormones

Calcitonin.—The hormone calcitonin produces hypocalcaemia and hypophosphataemia and has been reported to influence bone metabolism by interfering with the bone resorptive process. The hypophosphataemic effects observed after calcitonin treatment may be mediated by inhibition of bone resorption, and by effects on other tissues (*e.g.* the gut) involved in the overall homeostasis of phosphate and calcium.[282] The hypocalcaemic response to calcitonin is markedly retarded with advancing age.[283]

A solid-phase synthesis of calcitonin has been reported,[284] and the synthetic peptide had a specific biological activity of approximately 25 MRC units mg^{-1} against pure human MRC standard calcitonin (assigned potency, 100 MRC units mg^{-1}). Mono-iodinated salmon calcitonin retaining full biological activity has also been prepared [285] and shown to bind specifically to renal and skeletal receptors.[286]

Parathyroid Hormone.—In the search for a minimum active fragment of bovine parathyroid hormone (an 84 amino-acid single-chain polypeptide), the amino-terminal peptide fragment (1—34) has been found to be highly active (80%, on a molar basis, of the activity of the parent hormone *in vitro*, and 130% *in vivo*).[4e, 287] Further shortening of the peptide chain resulted in considerable loss of activity; fragments (1—28) and (1—27) showed weak activity (5% and 2%, respectively, of the parent hormone in *in vitro* tests), but smaller fragments (1—26, 1—12, 13—34, and an equimolar mixture of the latter two) were inactive. Deletion of the amino-terminal alanine residue (fragment 2—34) caused a marked decrease in *in vitro* activity, but the fragment showed 65% activity in the intravenous chick assay. Fragment 3—34 was inactive in all tests.

The active fragment (1—34) of human parathyroid hormone has been synthesized by two groups.[288, 289] Andreatta *et al.* synthesized, by a frag-

[282] F. S. Tanzer and J. M. Navia, *Nature New Biol.*, 1973, **242**, 221.
[283] H. Orimo and P. F. Hirsch, *Endocrinology* 1973, **93**, 1206.
[284] D. A. Ontjes, J. C. Roberts, J. F. Hennessey, H. J. Burford, and C. W. Cooper, *Endocrinology*, 1973, **92**, 1780.
[285] R. Ardaillou, J. Drouet, N. Loreau, and P. Corvol, *J. Clin. Endocrinol. Metab.*, 1973, **37**, 776.
[286] S. J. Marx, C. Woodward, G. D. Aurbach, H. Glossmann, and H. T. Keutmann, *J. Biol. Chem.*, 1973, **248**, 4797.
[287] B. Reit, B. Rafferty, J. A. Parsons, G. W. Tregear, J. Van Rietschoten, H. T. Keutmann, and J. T. Potts, jun., *J. Endocrinology*, 1973, **59**, xlii; G. W. Tregear, J. Van Rietschoten, E. Greene, H. T. Keutmann, H. D. Niall, B. Reit, J. A. Parsons, and J. T. Potts, jun., *Endocrinology*, 1973, **93**, 1349.
[288] R. H. Andreatta, A. Hartmann, A. Jöhl, B. Kamber, R. Maier, B. Riniker, W. Rittel, and P. Sieber, *Helv. Chim. Acta*, 1973, **56**, 470.
[289] G. Milhaud, P. Rivaille, and J. F. Staub, *Lancet*, 1973, 440.

ment-condensation procedure, the sequence reported by Brewer *et al.*,[290] and reported weak hypercalcaemic activity in thyroparathyroidectomized rats (100 units mg^{-1}). Milhaud *et al.* synthesized, by a solid-phase method, the amide of the (1—34) sequence determined by Niall *et al.*[291] and reported a much higher biological activity (1200 units mg^{-1}). This sequence differs from the one reported by Brewer *et al.* in having a glutamic acid in place of glutamine in position 22, leucine in place of lysine in position 28, and aspartic acid in place of leucine in position 30.

Several papers have appeared on the existence of a precursor of parathyroid hormone (pro-parathyroid hormone or Calcemic Fraction A).[292–296] A hexapeptide sequence, Lys-Ser-Val-Lys-Lys-Arg, linked to the amino end of parathyroid hormone has been identified but the exact nature of the prohormone remains to be established.

7 Vasoactive Plasma Peptides

Bradykinin.—Bradykinin was earlier predicted to have an extended conformation (see ref. 1). Circular dichroism studies in aqueous solution have now suggested that the secondary structure is a time average of two interconverting structures – one disordered and the other partially ordered by formation of a hydrogen bond between the 6-serine carbonyl group and the 8-phenylalanine peptide NH (the ordered structure predominates at higher temperatures).[297] In 90% dioxone–water a more ordered structure, stabilized by three intramolecular hydrogen bonds, is indicated.

Bradykininyl-Val-Ala-Pro-Ala-Ser (isolated from the skin of *Rana nigromaculata*)[298] and bradykininyl-Gly-Lys-Phe-His (isolated from the skin of *Bombina orientalis*)[299] have been synthesized, and shown to exert some smooth-muscle-contracting activity. The sequence of bradykinin potentiator A, one of five bradykinin-potentiating peptides isolated from

[290] H. B. Brewer, T. Fairwell, R. Ronan, G. W. Sizemore, and C. D. Arnaud, *Proc. Nat. Acad. Sci. U.S.A.*, 1972, **69**, 3585.
[291] H. D. Niall, R. T. Sauer, J. W. Jacobs, H. T. Keutmann, G. V. Segre, J. L. H. O'Riordan, G. D. Aurbach, and J. T. Potts, jun., *Proc. Nat. Acad. Sci. U.S.A.*, 1974, **71**, 384.
[292] J. F. Habener, B. Kemper, J. T. Potts, jun., and A. Rich, *Endocrinology*, 1973, **92**, 219.
[293] B. Kemper, J. F. Habener, J. T. Potts, jun., and A. Rich, *Proc. Nat. Acad. Sci. U.S.A.*, 1972, **69**, 643.
[294] D. V. Cohn, R. R. MacGregor, L. L. H. Chu, J. R. Kimmel, and J. W. Hamilton, *Proc. Nat. Acad. Sci. U.S.A.*, 1972, **69**, 1521.
[295] L. L. H. Chu, R. R. MacGregor, P. I. Liu, J. W. Hamilton, and D. V. Cohn, *J. Clin. Invest.*, 1973, **52**, 3089.
[296] J. W. Hamilton, H. D. Niall, H. T. Keutmann, J. T. Potts, jun., and D. V. Cohn, *Fed. Proc.*, 1973, **32**, 269 (abstract).
[297] J. R. Cann, J. M. Stewart, and G. R. Matsueda, *Biochemistry*, 1973, **12**, 3780.
[298] N. Yanaihara, C. Yanaihara, M. Sakagami, T. Nakajima, T. Nakayama, and K. Matsumoto, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 616.
[299] T. Yasuhara, M. Hira, T. Nakajima, N. Yanaihara, C. Yanaihara, T. Hashimoto, N. Sakura, S. Tachibana, N. Araki, M. Bessho, and T. Yamanaka, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 1388.

Table 14 *Biological activity of bradykinin analogues,* X-Arg-Pro-Y-Gly-Phe-Ser-Pro-Phe-Arg-OH (*from refs.* 301 *and* 302)

Compound number	X	Y	*Rat blood pressure* [a]	*Rat uterus* [b]	*Guinea-pig ileum* [b]	*% Inhibition of prolyl hydroxylase activity* [c]
Bradykinin	H	L-Pro	1.00	1.00	1.00	—
(203)	H	3,4-dehydro-L-Pro	0.71	0.97	0.82	32
(204)	H	3,4-dehydro-DL-Pro	0.26	0.54	0.70	—
(205)	H	*trans*-4-hydroxy-L-Pro	0.63	1.14	0.61	51
(206)	H	*cis,trans*-4-methyl-DL-Pro	0.55	0.32	0.45	37
(207)	H	L-thiazolidine-4-carboxyl	0.15	0.41	0.51	65
(208)	H	L-azetidine-2-carboxyl	0.26	0.46	0.68	7
(209)	L-Tyr	L-Pro	0.74	0.40	0.33	—
(210)	L-Glu	L-Pro	0.32	0.41	0.08	—
(211)	L-Glu	3,4-dehydro-L-Pro	0.32	0.39	0.44	100
(212)	L-Glu	*trans*-4-hydroxy-L-Pro	—	—	—	67
(213)	L-Lys	L-Pro	0.01	0.004	0.001	—

[a] Bradykinin dose/analogue dose to produce a 10 mmHg decrease in mean arterial blood pressure in the rat; [b] Bradykinin dose/analogue dose to produce a 1g force of contraction 15 s after drug administration; [c] At a concentration of 4×10^{-4} mol l^{-1} in a standard hydroxylating system.

the venom of *Agkistrodon halys blomhoffii*, has been determined.[300] This peptide (201), which unlike potentiators B, C, and E lacks a prolyl-proline sequence at the *C*-terminus, has a weak potentiating activity. A synthetic analogue (202) with an additional proline residue at the *C*-terminus shows strong bradykinin-potentiating activity on the guinea-pig ileum [it is about 200 times as potent as (201)].

Glp-Gly-Arg-Pro-Pro-Gly-Pro-Pro-Ile-Pro
(201)

Glp-Gly-Arg-Pro-Pro-Gly-Pro-Pro-Ile-Pro-Pro
(202)

New analogues of bradykinin and their biological activities are shown in Table 14. Replacement of the proline residue in position 3 by 3,4-dehydro-L-proline gave a peptide (203) with almost the same activity as bradykinin, but the other analogues listed were all less active than bradykinin.[301] Compounds (203), (205)—(207), (211), and (212) showed significant ability to inhibit prolyl hydroxylase activity; the most effective being compounds (207), (211), and (212).[302] Bradykinin, and analogues obtained by extending bradykinin at the *N*-terminus (Table 15), have been found to

Table 15 *Relative potencies of bradykinin analogues (from ref.* 303)

	Relative potencies [a]		
Structure	*Guinea-pig ileum*	*Vascular permeability*	*Antinociceptive effect*
Bradykinin	1.00	1.00	1.00
Lysyl-bradykinin	0.30	1.00	0.28
Methionyl-lysyl-bradykinin	0.08—0.10	10.00	0.060
Glycyl-arginyl-methionyl-lysyl-bradykinin	0.06—0.07	11.0—20.0	0.047

[a] All data are corrected for molar concentrations.

produce an increase in the threshold voltage required to stimulate tooth pulp (antinociceptive effect) when given by the intraventricular route in rabbits.[303] Activities on the guinea-pig ileum and on the threshold of electrical stimulation decreased as the chain length of the peptides increased, but the effect on vascular permeability increased with increasing chain length.

[300] H. Kato, T. Suzuki, K. Okada, T. Kimura, and S. Sakakibara, *Experientia*, 1973, **29**, 574.
[301] A. M. Felix, M. H. Jimenez, R. Vergona, and M. R. Cohen, *Internat. J. Peptide Protein Res.*, 1973, **5**, 201.
[302] J. O'D. McGee, M. H. Jimenez, A. M. Felix, G. J. Cardinale, and S. Udenfriend, *Arch. Biochem. Biophys.*, 1973, **154**, 483.
[303] S. A. Ribeiro and M. Rocha e Silva, *Brit. J. Pharmacol.*, 1973, **47**, 517.

Bradykinin and some of its fragments have been tested for oedema-producing and pain-producing activities.[304] When injected into the dog femoral or superior mesenteric artery, bradykinin increased 'vocalization' in a dose-dependent manner. Three synthetic fragments, Phe-Ser-Pro-Phe-Arg, Phe-Ser-Pro-Phe, and Arg-Pro, showed oedema-producing activity but failed to induce pain.

Angiotensin.—The effects of angiotensin on renal pharmacology,[305] the role of renal factors in the pathogenesis of experimental hypertension,[306] and the effects of angiotensin on the central nervous system [307] have been reviewed. The renin–angiotensin system does not seem to be involved in the pathogenesis of pregnancy hypertension since the concentrations of renin, renin substrate, angiotensin II, and aldosterone in the plasma of patients are all lower than in normotensive pregnant women.[308] Evidence has been presented that the effects of angiotensin II on heart muscle result from two independent actions, a direct stimulation of the resting tension and force of contraction, and an indirect stimulation (by release of endogenous catecholamines) of the force and velocity of contraction.[309] The many actions of angiotensin II on blood vessels are not restricted to the control of blood pressure. The contracting effect on the endothelial cells of rat aorta and coronary arteries and of rabbit dermal capillaries leads to a widening of the interendothelial space, and this increase in the gap between cells is associated with increased vascular permeability to plasma macromolecules.[310] The effect is inhibited by simultaneous injection of angiotensin II analogues known to be antagonists of its myotropic and vasopressor activities.[310]

In the ^1H n.m.r. spectrum of [1-asparagine, 5-valine]-angiotensin II, all the CH proton resonances observed in D_2O, and the NH resonances of the *cis* and *trans* primary amide hydrogens of the asparagine residue and the four equivalent guanidino hydrogens of the arginine residue observed in water, have been unambiguously assigned.[311] Tentative assignments, on the basis of titration data and the rates of hydrogen–deuterium exchange,[311, 312] of the peptide amide NH resonances to specific amino-acid residues have been confirmed by decoupling experiments of the α-protons in H_2O at pH 3 using a specially adapted spectrometer.[313] From the appro-

[304] N. Taira and K. Hashimoto, *Tohoku J. Exp. Med.*, 1973, **110**, 191.
[305] F. Gross and J. Mohring, *Ann. Rev. Pharmacol.*, 1973, **13**, 57.
[306] S. Koletsky, *Internat. Rev. Exp. Pathol.*, 1973, **12**, 203.
[307] W. B. Severs and A. E. Daniels-Severs, *Pharmacol. Rev.*, 1973, **25**, 415.
[308] R. J. Weir, J. J. Brown, R. Frazer, A. Kraszewski, A. F. Lever, G. M. McIlwaine, J. J. Morton, J. I. S. Robertson, and M. Tree, *Lancet*, 1973, 291.
[309] J. Drimal and D. Boska, *European J. Pharmacol.*, 1973, **21**, 130.
[310] A. L. Robertson and P. A. Khairallah, *Circulation Res.*, 1973, **31**, 923.
[311] J. D. Glickson, W. D. Cunningham, and G. R. Marshall, *Biochemistry*, 1973, **12**, 3684.
[312] H. E. Bleich, R. E. Galardy, and M. P. Printz, *J. Amer. Chem. Soc.*, 1973, **95**, 2041; H. E. Bleich, R. E. Galardy, M. P. Printz, and L. C. Craig, *Biochemistry*, 1973, **12**, 4950.
[313] J. D. Glickson, J. Dadok, and G. R. Marshall, *Biochemistry*, 1974, **13**, 11.

priate coupling constants it was possible to calculate values of dihedral angles ϕ for the NH—CH bond of each of the residues in angiotensin II.[314] After allowing for the errors and assumptions involved in these calculations, it was felt that the data exclude all models proposed to date (see ref. 1 for a summary) from being major contributors to the conformational equilibrium except that of an equilibrium between various ordered and disordered forms.[315] Whatever the conformation of angiotensin II at this acid pH, the molecule undergoes a conformational transition in aqueous solution with a pK_a 6.6 ± 0.2 [311] (also detectable by ^{19}F n.m.r. with fluorine-labelled analogues of angiotensin [316]), and so the conformation in physiological fluid or at the receptor site will probably be different.[313, 314]

A study of the interactions between angiotensin II and its receptors on rat stomach and rabbit aortic strips has confirmed the applicability of drug receptor theory to these systems.[317] It was first established that the myotropic action of angiotensin II on these tissues is a direct one resulting from stimulation of specific receptors and is not mediated through release of acetylcholine on the stomach or of noradrenaline on the aorta. In addition, the contractile action of angiotensin II on either tissue does not appear to involve increased biosynthesis or release of prostaglandins (since it is unaffected by the presence of indomethacin, a potent inhibitor of prostaglandin biosynthesis). Second, dose–response curves for angiotensin II on both tissues were constructed and found to be close to the theoretical curves predicted by the law of mass-action. Third, a linear relationship between stimulus (as a percentage of maximum stimulus) and effect (as a percentage of maximum effect) was found, suggesting that the extent of contraction is proportional to the number of receptors occupied. There was no indication of a threshold effect or of a receptor reserve capacity in either tissue. It was concluded that the concepts of affinity and intrinsic activity can be validly applied to quantitative measurement of relations between chemical structure and biological activity in the interaction of angiotensin with its receptors in these tissues.

A microsomal membrane extract from rabbit aorta which contains specific angiotensin receptors has been separated into five fractions on a density gradient.[318] Binding of tritiated angiotensin II was greatest in the lightest band of membranes (density less than 1.01) and could be inhibited by unlabelled angiotensin II but not by L-noradrenaline or 8-lysine-vasopressin. Very low levels of tritiated angiotensin II bound to the other

[314] G. R. Marshall, H. E. Bosshard, W. H. Vine, and J. D. Glickson, *Nature New Biol.*, 1973, **245**, 125.
[315] S. Fermandjian, P. Fromageot, A.-M. Tistchenko, J. P. Leickman, and M. Lutz, *European J. Biochem.*, 1972, **28**, 174.
[316] W. H. Vine, D. A. Brueckner, P. Needleman, and G. R. Marshall, *Biochemistry*, 1973, **12**, 1630.
[317] F. Rioux, W. K. Park, and D. Regoli, *Canad. J. Physiol. Pharmacol.*, 1973, **51**, 665.
[318] M.-A. Devynck, M.-G. Pernollet, P. Meyer, S. Fermandjian, and P. Fromageot, *Nature New Biol.*, 1973, **245**, 55.

membrane fractions, and these were not significantly inhibited by unlabelled peptide. Though 5′-nucleotidase and adenyl cyclase activity could be detected in the lightest membrane fraction, the highest content of these enzymes was found in the adjacent band of the gradient (density 1.01—1.04); calcium binding was also highest in this second membrane band. It is suggested that disruption of rabbit aorta cells might cause the plasma membrane to separate into different types of vesicles each carrying a different complement of the functions of the intact surface. The surface of the cells is imagined as made up of a mosaic of large and functionally discrete macromolecular assemblies.

The ability of various analogues and fragments of angiotensin II to displace tritiated peptide from the lightest membrane fraction increased with their tendency to assume a cross-beta type conformation[318] (as proposed for angiotensin II in concentrated aqueous solution or in polar organic solvents[315]). In general the binding characteristics of these various peptides correlated with their biological effects on intact rabbit aorta,[318] an important point if specific angiotensin II binding sites, present only on the lightest membrane fraction, are to be related to the receptor sites involved in the biological response. However, the (3—8)-hexapeptide of angiotensin II, which has very little intrinsic activity, and [4-phenylalanine]-angiotensin II, which has a greatly reduced agonist activity, were both very effective inhibitors of the binding of tritiated angiotensin II. The *C*-terminal penta- and tetra-peptide fragments of angiotensin II, thought to have no activity on intact aorta, showed some ability to inhibit binding. This suggests that the four *C*-terminal amino-acids of the hormone provide the minimal stereocomplementarity necessary for binding.[318] Interestingly an analogous tetrapeptide sequence (see below) is a weak antagonist of the myotropic effect of angiotensin II.[319]

Angiotensin II has no effect on the activity of two cyclic-AMP phosphodiesterases detected in rat uterine homogenates.[320] Since it also has no effect on basal uterine adenyl cyclase, it is concluded that the uterine-contracting response is not mediated by either of the enzymes controlling the cellular level of cyclic-AMP.[320] In contrast, the steroidogenic action of angiotensin II on intact isolated bovine adrenal cells is associated with an increase in cyclic-AMP levels, even though adenyl cyclase is not stimulated in broken cell preparations.[321]

The last Report[1] described how substitution of the *C*-terminal phenylalanine residue in angiotensin II [Table 16, compound (214)] with an aliphatic amino-acid residue produced competitive antagonists to the parent molecule [*e.g.* compounds (215) and (216)]. This antagonistic

[319] M. C. Khosla, M. M. Hall, R. R. Smeby, and F. M. Bumpus, *J. Medicin. Chem.*, 1973, **16**, 1184.
[320] G. A. d'Auriac and P. Meyer, *Life Sci.*, 1973, **12**, 233.
[321] A. Peytermann, W. E. Nicholson, R. D. Brown, G. W. Liddle, and J. G. Hardman, *J. Clin. Invest.*, 1973, **52**, 835.

Table 16 *Antagonist analogues of angiotensin*

Compound number	*Structure*	*Ref.*
(214)	Asp-Arg-Val-Tyr-Ile -His-Pro-Phe	—
(215)	Asp-Arg-Val-Tyr-Val-His-Pro-Ala	—
(216)	Asp-Arg-Val-Tyr-Ile -His-Pro-Ile	—
(217)	Sar -Arg-Val-Tyr-Val-His-Pro-Ala	—
(218)	Sar -Arg-Val-Tyr-Ile -His-Pro-Ile	—
(219)	Suc-Arg-Val-Tyr-Ile -His-Pro-Ala [a]	332
(220)	Asp-Arg-Val-Ala-Ile -His-Pro-Ile	332
(221)	Asp-Arg-Val-Tyr-Ile -His-Pro-D-*a*Ile [b]	332
(222)	Asp-Arg-Pro-Tyr-Ile -His-Pro-Ile	319
(223)	Sar -Arg-Pro-Tyr-Ile -His-Pro-Ile	319
(224)	Asp-Arg-Val-Phe-Ile -His-Pro-Ile	319
(225)	Ile -His-Pro-Ile	319
(226)	Asp-Arg-Val-Tyr-Ile -His-Pro-Gly	333
(227)	Asp-Arg-Val-Tyr-Ile -His-Pro-Abu [c]	333
(228)	Asp-Arg-Val-Tyr-Ile -His-Pro-Val	333
(229)	Asp-Arg-Val-Tyr-Ile -His-Pro-Leu	333
(230)	Sar -Arg-Val-Tyr-Ile -His-Pro-Leu	333
(231)	Sar -Arg-Val-Tyr-Ile -His-Pro-Phe	333*b*, 334
(232)	Asn-Arg-Val-Tyr-Val-His-Pro-Cys	335
(233)	Asp-Arg-Val-Tyr-Ile -His-Pro-Val-His-Leu	337
(234)	Asp-Arg-Val-Tyr-Ile -His-Pro-Ile -His-Leu	337
(235)	Chl-Arg-Val-Tyr-Ile -His-Pro-Val-His-Leu [d]	337
(236)	Chl-Arg-Val-Tyr-Ile -His-Pro-Phe	338

[a] Suc = succinyl; [b] *a*Ile = *allo*isoleucine; [c] Abu = aminobutyryl; [d] Chl = chlorambucil.

activity was potentiated, and its duration of action was increased, when the *N*-terminal residue in the 8-substituted analogues was replaced by sarcosine. Two of the compounds, (217) and (218), were received with particular enthusiasm and used in numerous investigations; the collective results indicate a significant role for the renin–angiotensin system in the maintenance of certain forms of hypertension, and in the maintenance of normal blood pressure in sodium-depleted rats [322–325] and dogs.[324, 326–328] In man, compound (217) has been shown to reduce the elevated blood pressure markedly in hypertensive patients with high plasma renin levels.[329] Both peptides seem to antagonize the central and peripheral nervous

[322] J. Bing and K. Nielsen, *Acta Path. Microbiol. Scand.*, 1973, **81A**, 254.
[323] D. T. Pals and R. W. Fulton, *Arch. Internat. Pharmacodyn.*, 1973, **204**, 20.
[324] F. M. Bumpus, S. Sen, R. R. Smeby, C. S. Sweet, C. M. Ferrario, and M. C. Khosla, *Circulation Res.*, 1973, **32** (suppl. 1), 150.
[325] H. Gavras, H. R. Brunner, E. D. Vaughan, and J. H. Laragh, *Science*, 1973, **180**, 1369.
[326] D. T. Pals and F. D. Masucci, *European J. Pharmacol.*, 1973, **23**, 115.
[327] J. A. Johnson, J. O. Davis, W. S. Spielman, and R. H. Freeman, *Physiologist*, 1973, **16**, 354.
[328] J. A. Johnson and J. O. Davis, *Circulation Res.*, 1973, **32** (suppl. 1), 159; *Science*, 1973, **179**, 906.
[329] H. R. Brunner, H. Gavras, J. H. Laragh, and R. Keenan, *Lancet*, 1973, 1045.

interactions of angiotensin II as well as the direct vasoconstrictor effects.[323, 330, 331]

New analogues (219)—(224) have been prepared[319, 332] in order to study the contribution of residues in the remainder of the peptide chain to inhibitory activity. Each exhibited lower antagonistic activity against angiotensin II than the parent molecules (215) and (216). The level of antagonistic activity decreased still further when the L-histidine residue in position 6 of peptides (219), (220), and (221) was replaced by D-histidine.[332] These decreases in potency may result from changes in conformation, or binding to the receptor, or from a decreased resistance to enzymic degradation. As mentioned above, a tetrapeptide (225) was also found to be an inhibitor, albeit a very weak one.[319]

Detailed studies using rat isolated stomach strips have confirmed that compounds (216) and (226)—(229) are specific and competitive antagonists of the effects of angiotensin I and angiotensin II, but not of 5-hydroxytryptamine and bradykinin.[333] Peptides (216) and (229) are about 12 times more potent than compound (226), and the similarity of the pA_2* values to the pD_2* values for angiotensin II indicates that one molecule of agonist is blocked by one molecule of antagonist. The action of the more potent 1-sarcosine-substituted derivative (230) is specific, and competitive in the sense that the angiotensin II receptor site is involved. However, probably because of an increased resistance to enzymic inactivation, the derivative does not satisfy the criteria of a 'competitive' inhibitor and is better described as a 'long-acting' antagonist.[333] These conclusions regarding 1-sarcosine-substituted antagonists are supported by studies with the potent agonist analogue (231).[333b, 334]

The competitive antagonist (232) has been shown to block the release of prostaglandin from isolated perfused rabbit spleen when angiotensin I or II, but not adrenaline, is used as stimulant.[335] Similarly, this and other antagonists will inhibit, specifically, the release of prostaglandin from isolated rabbit kidney by angiotensin II.[336] In both organs the release by angiotensin I is independent of its conversion into angiotensin II, but angiotensin I has only 2—10% of the potency of angiotensin II.

[330] C. S. Sweet, C. M. Ferrario, M. C. Khosla, and F. M. Bumpus, *J. Pharmacol. Exp. Therap.*, 1973, **185**, 35.
[331] B. G. Zimmerman, *J. Pharmacol. Exp. Therap.*, 1973, **185**, 486.
[332] M. C. Khosla, M. M. Hall, R. R. Smeby, and F. M. Bumpus, *J. Medicin. Chem.*, 1973, **16**, 829.
[333] (*a*) F. Rioux, W. K. Park, and D. Regoli, *Canad. J. Physiol. Pharmacol.*, 1973, **51**, 108; (*b*) D. Regoli, W. K. Park, and F. Rioux, *ibid.*, p. 114.
[334] M. M. Hall, M. C. Khosla, P. A. Khairallah, and F. M. Bumpus, *I.C.R.S. J. Internat. Res. Comm.*, 1973, **1**, 10.
[335] J. R. Douglas, E. M. Johnson, G. R. Marshall, B. M. Jaffe, and P. Needleman, *Prostaglandins*, 1973, **3**, 67.
[336] P. Needleman, A. H. Kauffman, J. R. Douglas, E. M. Johnson, and G. R. Marshall, *Amer. J. Physiol.*, 1973, **224**, 1415.

* See ref. 1, p. 442, for definition of this term.

The decapeptide analogues (233)—(235) are specific competitive antagonists of the effects of angiotensin I and II on rat blood pressure, isolated guinea-pig ileum, rat uterus, and rabbit aortic strip.[337] They are comparable in potency to compounds (215) and (216) as inhibitors of the myotropic activity of angiotensin II, and more potent [but less potent than compound (217)] as inhibitors of the pressor effect of angiotensin II. The reversible and competitive inhibition shown by the decapeptide (235) contrasts with the irreversible and non-competitive inhibition exhibited by the octapeptide (236).[338] Although the former has high affinity for the receptors it seems that the interaction does not permit the *N*-terminal alkylating group to approach a suitable nucleophilic site on the receptor.[337]

Few reports in 1973 have been concerned with further clarification of the role of individual amino-acid residues in the agonist activity of angiotensin. Details of results (reviewed last year [1]) reflecting the importance of stereochemical factors at position 5,[339] and of aromatic and nucleophilic character in the side-chain at position 6 [340] have appeared. [6-N^{τ}-Methylhistidine]- and [6-N^{π}-methylhistidine]-analogues (237) and (238) of angiotensin II have been prepared and their biological effects evaluated.[341] The N^{τ}-methylhistidine analogue was about 5% as active as angiotensin II in contracting isolated smooth-muscle preparations, as a vasopressor agent, and in releasing prostaglandin from isolated perfused rabbit kidneys. The N^{π}-methylhistidine analogue possessed negligible biological activity, and neither analogue showed any antagonistic activity. The sensitivity of uterine strips to angiotensin is increased at alkaline pH, and a similar increase in sensitivity is observed with [6-pyrazolyl-3′-alanine]-angiotensin II,[342] with the methylhistidine analogues (237) and (238), and with *N*-acetylangio-

Asp-Arg-Val-Tyr-Ile-His(τMe)-Pro-Phe
(237)

Asp-Arg-Val-Tyr-Ile-His(πMe)-Pro-Phe
(238)

tensin II.[341] This suggests that the pH-dependent sensitivity is not a consequence of change in the state of ionization of the histidine imidazole group or of the *N*-terminal α-amino-group, and it was speculated that the increased sensitivity may result from some change in the receptor itself, or from changes in some subsequent step in the stimulus–response coupling.[341]

[337] A. C. M. Paiva, V. L. A. Nouailhetas, M. E. Miyamoto, G. B. Mendes, and T. B. Paiva, *J. Medicin. Chem.*, 1973, **16**, 6.
[338] T. B. Paiva, A. C. M. Paiva, R. J. Freer, and J. M. Stewart, *J. Medicin. Chem.*, 1972, **15**, 6.
[339] E. C. Jorgensen and G. C. Windridge, *J. Medicin. Chem.*, 1973, **16**, 467.
[340] R. J. Freer and J. M. Stewart, *J. Medicin. Chem.*, 1973, **16**, 733.
[341] P. Needleman, G. R. Marshall, and J. Rivier, *J. Medicin. Chem.*, 1973, **16**, 968.
[342] P. Needleman, R. J. Freer, and G. R. Marshall, *Arch. Internat. Pharmacodyn.*, 1972, **200**, 118.

Several recent studies have been directed towards the synthesis of inhibitors of the renin–angiotensin pathway (Figure 2) at stages preceding the interaction of angiotensin II with its target-cell receptors. Renin is released from the juxtaglomerular cells of the kidney in response to a variety of stimuli; it cleaves a leucyl-leucine bond in a specific, circulating

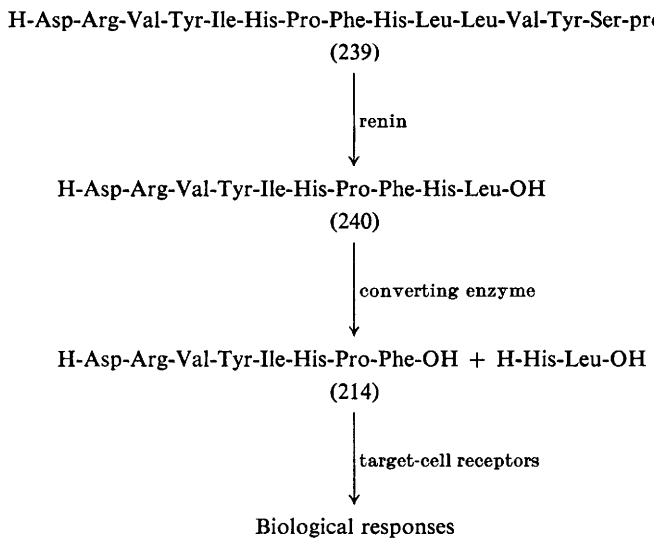


Figure 2 *Outline of the renin–angiotensin system*

α_2-globulin substrate angiotensinogen (239), to give the decapeptide angiotensin I (240). Further digestion by converting enzyme, a dipeptidyl-carboxypeptidase, releases the *C*-terminal dipeptide from (240), generating the highly active octapeptide angiotensin II.

The *N*-terminal tetradecapeptide fragment (241) (Table 17) of angiotensinogen (239) is as effective as the natural substrate for renin, and kinetic studies with shorter synthetic peptide congeners have shown the octapeptide (242) to be the smallest fragment still able to serve as a useful substrate.[343] Using this knowledge, peptides (243)—(248) were prepared, and tested as competitive inhibitors of the reaction between renin and tetradecapeptide substrate.[344] The octapeptides (242)—(248) all acted as true competitive inhibitors but only the D-leucine analogues (245) and (246) did not also act as substrates for the enzyme. Peptide (246) also inhibited the reaction between human renin and the protein substrate (239) in human plasma.[344] Analogous replacement of L-leucine by D-leucine

[343] L. T. Skeggs, K. E. Lentz, J. R. Kahn, and H. Hochstrasser, *J. Exp. Med.*, 1968, **128**, 13.

[344] K. Poulsen, J. Burton, and E. Haber, *Biochemistry*, 1973, **12**, 3877.

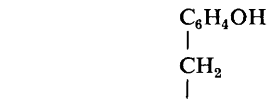
Table 17 *Competitive inhibitors of renin*

Compound number	*Structure*	K_i/μmol l^{-1}
(241)	H-Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu-Leu-Val-Tyr-Ser-OH	—
(242)	H———OH	39
(243)	H-D-His———OH	6
(244)	H———Phe-OH	3
(245)	H———D-Leu———OH	25
(246)	H———D-Leu———OH	3
(247)	H———Ile———OH	7
(248)	H———Ile———OH	> 25
(249)	———D-Leu———OH	7 [a]
(250)	———D-Leu———OH	15 [a]
(251)	———D-Leu-D-Leu———OH	15 [a]
(252)	H———OH	
(253)	H———Phe-OH	

[a] A different assay system was used for this peptide.

residues in the tridecapeptide corresponding to des-Asp[1]-tetradecapeptide substrate also yielded potent competitive inhibitors (249)—(251), with compound (249) being slightly more effective than the others.[345] Different assay systems were used in the two investigations so it is not possible to assess the relative potency of the octapeptide and tridecapeptide series. However, replacement with D-leucine on the amino side of the critical amide bond in the octapeptide gave an inhibitor (246) with greater affinity for the enzyme than when the same replacement was made on the carboxyl side of the amide bond (245), whereas the reverse relationship seems to exist in the tridecapeptide series [*cf.* compounds (249) and (250)]. It is hoped that these inhibitors may prove valuable in unravelling the physiological role of renin and the mechanism of its action in various forms of essential hypertension.

C-Terminal ester, amide, and carbinol derivatives of the tetrapeptides (252) and (253) also competitively inhibit the renin–plasma angiotensinogen reaction,[346, 347] but they are effective only at very high concentrations.[348] The carbinol (254) has an inhibitory constant three orders of magnitude greater than that of the octapeptide (246).[344] Structural requirements for



(254)

[345] I. Parikh and P. Cuatrecasas, *Biochem. Biophys. Res. Comm.*, 1973, **54**, 1356.
[346] K. Shigezane and T. Mizoguchi, *Chem. and Pharm. Bull (Japan)*, 1973, **21**, 972.
[347] T. Kokubu, K. Hiwada, T. Ito, E. Ueda, Y. Yamamura, T. Mizoguchi, and K. Shigezane, *Biochem. Pharmacol.*, 1973, **22**, 3217.
[348] D. A. Ontjes, J. Majstoravich, and G. Frye, *Fed. Proc.*, 1973, **32**, Abstr. 3077, p. 760.

anti-renin activity in these tetrapeptides are two adjacent *N*-terminal leucine residues, the first having an L-configuration, and an aromatic amino-acid residue at the *C*-terminus.[347] Replacement of the valine residue by leucine or addition of a histidine residue at the *N*-terminus did not affect the activity, but *N*-acylation of the tetrapeptide derivatives markedly reduced the activity.

Angiotensin-converting enzyme activity has been detected in blood and several other tissues, for example the lung. The action of the 'converting enzyme' is not limited to conversion of angiotensin I into angiotensin II; it will cleave *C*-terminal dipeptides from many peptides, including bradykinin. The most potent and specific inhibitors of converting enzyme presently known are structurally similar peptides, initially recognized for their bradykinin-potentiating activity, and isolated from the venoms of the South American pit viper *Bothrops jararaca* and the Japanese pit viper *Agkistrodon halys blomhoffii*. Many analogues (Table 18) of three of the peptides

Table 18 *Inhibition of angiotensin-converting enzyme by synthetic analogues of peptides from* Bothrops jararaca *venom (from ref.* 49)

Compound number	*Structure*	I_{50}/μg ml^{-1}
(255)	Glp-Lys-Trp-Ala-Pro	0.05
(256)	Glp-Lys-Phe-Ala-Pro	0.05
(257)	Cpc-Lys-Phe-Ala-Pro	0.06
(258)	Boc-Lys-Phe-Ala-Pro	5.2
(259)	Glp-Nle-Phe-Ala-Pro	0.2
(260)	Glp-Gln-Phe-Ala-Pro	0.4
(261)	Glp-Glu-Phe-Ala-Pro	3.0
(262)	Glp-Thr-Phe-Ala-Pro	5.5
(263)	Glp-Lys-Pro-Ala-Pro	1.1
(264)	Glp-Lys-Ile -Ala-Pro	1.6
(265)	Glp-Lys-Ser -Ala-Pro	2.4
(266)	Glp-Lys-*trp* -Ala-Pro	72
(267)	Glp-Lys-Phe-Lac-Pro	0.06
(268)	Glp-Lys-Phe-Gly-Pro	0.1
(269)	Glp-Lys-Phe-Pro-Pro	3.3
(270)	Glp-Lys-Phe-Ala-Ala	0.06
(271)	Glp-Lys-Phe-Ala-Glu	2.0
(272)	Glp-Lys-Phe-Ala-Pyn	> 200
(273)	Lys-Trp-Ala-Pro	1.2
(274)	Phe-Ala-Pro	1.4
(275)	Ala-Pro	50
(276)	Glp-Phe-Ala-Pro	2.7
(277)	Glp-Lys-Phe-Ala-Pro-Pro	4.8
(278)	Glp-Lys-Phe-Ala	7.5
(279)	Boc-Phe-Ala-Pro	17
(280)	Glp-Lys-Phe	50
(281)	Glp-Lys-Phe-Pro	> 200
(282)	Glp-Ile -Pro-Pro-Lys-Phe-Ala-Pro	0.3
(283)	Glp-Lys-Phe-Ala-Pro-Gln-Ile -Pro-Pro	3.3
(284)	Glp-Trp-Pro-Arg-Pro-Lys-Phe-Ala-Pro	0.05
(285)	Glp-Trp-Pro-Arg-Pro-Gln-Ile -Pro-Pro	1.1

Table 18 (*cont.*)

Compound number	*Structure*	I_{50}/μg ml^{-1}
(286)	Cpc-Trp-Pro-Arg-Pro-Gln-Ile -Pro-Pro	0.2
(287)	Glp-Tyr-Pro-Arg-Pro-Gln-Ile -Pro-Pro	2.6
(288)	Glp-Phe-Pro-Arg-Pro-Gln-Ile -Pro-Pro	2.8
(289)	Glp-Leu-Pro-Arg-Pro-Gln-Ile -Pro-Pro	6.5
(290)	Glp-Gly-Pro-Arg-Pro-Gln-Ile -Pro-Pro	6.5
(291)	Glp-*trp* -Pro-Arg-Pro-Gln-Ile -Pro-Pro	7.0
(292)	Glp-Trp-Pro-Lys-Pro-Gln-Ile -Pro-Pro	1.5
(293)	Glp-Trp-Pro-His-Pro-Gln-Ile -Pro-Pro	1.9
(294)	Glp-Trp-Pro-Gly-Pro-Gln-Ile -Pro-Pro	1.9
(295)	Glp-Trp-Pro-Orn-Pro-Gln-Ile -Pro-Pro	2.5
(296)	Glp-Trp-Pro-*arg* -Pro-Gln-Ile -Pro-Pro	9.0
(297)	Glp-Trp-Pro-Arg-Pro-Gln-Phe-Pro-Pro	0.3
(298)	Glp-Trp-Pro-Arg-Pro-Gln-Ile -Ala-Pro	0.4
(299)	Glp-Trp-Pro-Arg-Pro-Gln-Ile -Pro-Pyn	> 200
(300)	Glp-Asn-Trp-Pro-Arg-Pro-Gln-Ile -Pro-Pro	3.7
(301)	Glp-Glu-Trp-Pro-Arg-Pro-Gln-Ile -Pro-Pro	4.5
(302)	Glp-Lys-Trp-Pro-Arg-Pro-Gln-Ile -Pro-Pro	5.0
(303)	Pro-Arg-Pro-Gln-Ile -Pro-Pro	9.6
(304)	Arg-Pro-Gln-Ile -Pro-Pro	32
(305)	Z-Pro-Gln-Ile -Pro-Pro	78
(306)	Glp-Ile -Pro-Pro	> 200
(307)	Glp-Trp-Pro-Arg-Pro	14
(308)	Pro-Arg-Pro-Gln-Ile	77
(309)	Glp-Arg-Pro-Gln-Ile -Pro	> 200
(310)	Glp-Trp-Pro	> 200
(311)	Glp-Trp	> 200

D-Amino-acids are written uncapitalized and in italics; Cpc = cyclopentylcarbonyl; Lac = Lactyl; Pyn = pyrrolidinyl.

from *Bothrops jararaca*, (255), (285), and (300), have been synthesized [349, 350] and the structural requirements for inhibitory activity defined using crude or homogeneous preparations of angiotensin-converting enzyme from rabbit lung.[351, 352] The nonapeptide (285) is only about one-tenth as potent on a molar basis as the pentapeptide (255) but the inhibitory activity persists for a much longer period. This arises because the nonapeptide, unlike the pentapeptide, does not act as an alternative substrate for the enzyme. A brief summary follows of the conclusions drawn from this work (the original paper [351] should be consulted for a more detailed analysis of the structure–activity data).

As with substrates, the presence in an inhibitor of a *C*-terminal carboxy-group is an absolute requirement for activity [compare compounds (256)

[349] M. A. Ondetti, J. Pluscec, E. R. Weaver, N. Williams, E. F. Sabo, and O. Kocy; 'Proceedings 3rd American Peptide Symposium, Boston, Mass., June 1972', ed. J. Meienhofer, Ann Arbor Science Publishers Inc., December 1972, p. 525.

[350] J. Pluscec, E. R. Weaver, N. Williams, E. F. Sabo, O. Kocy, and M. A. Ondetti, ref. 4, p. 403.

[351] D. W. Cushman, J. Pluscec, N. J. Williams, E. R. Weaver, E. F. Sabo, O. Kocy, H. S. Cheung, and M. A. Ondetti, *Experientia*, 1973, **29**, 1032.

with (272) and (285) with (299); the methyl ester or amide derivatives of (256) retain only 5% and 2% respectively of its activity]. Similarly, the presence of a penultimate proline residue [compounds (269), (277), and most analogues of compounds (285) and (300)], or of a *C*-terminal glutamic acid residue [compound (271)], or the absence of an aromatic residue in the third position from the *C*-terminus [compounds (263)—(265) and most analogues of compounds (285) and (300)], greatly decreases the inhibitory potency. These structure–activity correlations, together with kinetic data,[352] indicate that the three *C*-terminal amino-acid residues of all the venom peptides are bound to angiotensin-converting enzyme in a manner analogous to the binding of corresponding residues of substrates for the enzyme. However, the affinity for the enzyme of this *C*-terminal tripeptide 'competitive portion' is increased, at least 50-fold, by the presence of the remaining *N*-terminal regions of the inhibitors [compare compound (256) with compounds (273) and (274); the *C*-terminal tripeptide fragment of compound (285) produces no significant inhibition]. It is a combination of binding interactions from both parts of the molecules which gives these venom peptides their potency and specificity as inhibitors. Many substitutions in the *N*-terminal Glp-Lys sequence of peptide (256) lead to a decrease in potency [compounds (258), (261), and (262)] but some give highly active analogues [compounds (257), (259), and (260)]. Comparison of compounds (264) and (269) with compounds (298) and (297), respectively, suggests that the *N*-terminal portion of the nonapeptide (285) enhances affinity for the enzyme more than the *N*-terminal part of the pentapeptide (256). The tryptophan residue in compound (285) seems to be of particular importance.

The nonapeptide (285), frequently referred to only by the code name 'SQ 20881', has been shown to reduce the pressor effects of angiotensin I, but not those of angiotensin II, in various species,[353–355] including man.[356] It has already been used in a number of investigations on the role of the renin–angiotensin system and the regulation of hypertension.[357–361]

One of the responses to centrally administered angiotensin which has received attention in 1973 is its dipsogenic effect. CNS injection of nanogram quantities of angiotensin into many species produces a rapid increase in water intake; given intravenously, the minimum effective doses

[352] H. S. Cheung and D. W. Cushman, *Biochem. Biophys. Acta*, 1973, **293**, 451.
[353] A. Bianchi, D. B. Evans, M. Cobb, M. T. Peschka, T. R. Schaeffer, and R. J. Laffan, *European J. Pharmacol.*, 1973, **23**, 90.
[354] J. Di Salvo, S. Britton, P. Galvas, and T. W. Sanders, *Circulation Res.*, 1973, **32**, 85.
[355] T. R. Schaeffer, S. L. Engel, and B. Rubin, *Pharmacologist*, 1973, **15**, Abstr. 475, p. 240.
[356] J. G. Collier, B. F. Robinson, and J. R. Vane, *Lancet*, 1973, 72.
[357] A. I. Samuels, E. D. Miller, C. S. Fray, E. Haber, and A. C. Barger, *Fed. Proc.*, 1973, **32**, part 1, Abstr. 910, p. 380.
[358] E. E. Muirhead, B. Brooks, and K. K. Arora, *Circulation*, 1973, **48**, suppl. 4, Abstr. 284, p. 72.
[359] S. L. Engel, T. R. Schaeffer, M. H. Waugh, and B. Rubin, *Proc. Soc. Exp. Biol. Med.*, 1973, **143**, 483.
[360] J. A. Barbour and M. D. Bailie, *Proc. Soc. Exp. Biol. Med.*, 1973, **143**, 400.
[361] J. Bing, *Acta Path. Microbiol. Scand.*, 1973, **81A**, 376.

are at least 1000 times higher (see ref. 307 for a review of earlier work). The effect has even been seen with domestic cats, which normally do not readily drink water.[362] In rats, the speed and sensitivity of the drinking response when angiotensin II is applied directly to the subfornical organ, and the observation that lesions in the body of the subfornical organ block the response induced by angiotensin II applied to the preoptic area, suggest that this organ may contain the central dipsogenic receptors for circulating angiotensin II.[363] Such a hypothesis is attractive since the organ is a part of the central nervous system lying outside the blood–brain barrier (through which angiotensin II would probably penetrate only slowly and with difficulty). Renin, synthetic tetradecapeptide renin-substrate, and angiotensin I are all effective dipsogens, and it is not yet clear if angiotensin II is the common mediator for these substances. Intracranial injection of specific anti-angiotensin II serum blocks the dipsogenic effect of both intracranial angiotensin II and intracranial tetradecapeptide renin substrate (241).[364] The converting enzyme inhibitor (285) is reported to block the drinking response to peripheral renin in nephrectomized rats,[365] and when given intraventricularly 15 min before the stimulus, it blocked the drinking response to an intraventricular injection of angiotensin I but not the response to angiotensin II.[366] These results support the idea that angiotensin II is a common mediator, but other reports [367, 368] indicate that angiotensin I can act as a dipsogen independently of its conversion into angiotensin II. For example, the inhibitor (285) given intracerebrally 1 min before the stimulus did not block the response to intracerebral injection of angiotensin I.[367]

Intracerebral administration of three peripherally antagonistic, 8-substituted angiotensin II analogues {compounds (215) and (216), and [4-phenylalanine, 8-tyrosine]-angiotensin II} 1 min before intracerebral injection of angiotensin II failed to inhibit the dipsogenic response.[367] Indeed, under the conditions used, these compounds were themselves potent agonists. However, the long-acting antagonist (217) did act as a potent inhibitor of the drinking effect when given centrally to cats 30 min before angiotensin II [369] or when infused intravenously into rats during intraperitoneal injection of a crude renal extract.[370] Any similarity between the central dipsogenic and peripheral receptors for angiotensin II thus remains in question. In the same way, evidence has been presented for [369]

[362] M. J. Cooling and M. D. Day, *Brit. J. Pharmacol.*, 1973, **49**, 150P.
[363] J. B. Simpson and A. Routtenberg, *Science*, 1973, **181**, 1172.
[364] A. N. Epstein, J. T. Fitzsimons, and A. K. Johnson, *J. Physiol.*, 1973, **230**, 42P.
[365] D. Lehr, H. W. Goldman, and P. Casner, *Science*, 1973, **182**, 1031.
[366] W. B. Severs, J. Summy-Long, and A. Daniels-Severs, *Proc. Soc. Exp. Biol. Med.*, 1973, **142**, 203.
[367] L. W. Swanson, G. R. Marshall, P. Needleman, and L. G. Sharpe, *Brain Res.*, 1973, **49**, 441.
[368] R. W. Bryant and J. L. Falk, *Pharmacol. Biochem. Behaviour*, 1973, **1**, 469.
[369] M. J. Cooling and M. D. Day, *J. Pharm. Pharmacol.*, 1973, **25**, 1005.
[370] E. D. Vaughan, H. Gavras, J. H. Laragh, and M. N. Koss, *Nature*, 1973, **242**, 334.

and against[367] the involvement of catecholaminergic pathways in the mediation of angiotensin-induced drinking.

8 Tachykinins

Most of the numerous biologically active peptides which have been identified in the skin of amphibia over the past fifteen years can be classified in four major groups headed by caerulein, bradykinin, physalaemin, and bombesin. The pharmacology of these peptides, their similarity to active peptides from mammalian tissues, and the conclusions drawn from structure–activity studies with synthetic analogues have been conveniently restated in a recent lecture.[371]

Members of the group related to physalaemin (312) bear a close structural similarity to eledoisin (313) (isolated from the posterior salivary gland of the Mediterranean octopus) and substance P (314) (isolated from the

Glp-Ala-Asp-Pro-Asn-Lys-Phe-Tyr-Gly-Leu-Met-NH_2

(312)

Glp-Pro-Ser-Lys-Asp-Ala-Phe-Ile-Gly-Leu-Met-NH_2

(313)

Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH_2

(314)

hypothalamus and present in the intestinal wall of mammals). Not surprisingly, they share many of the same biological properties – powerful effects on vascular smooth muscle (producing vasodilatation and hypotension) and extravascular smooth muscle, and potent stimulation of lachrymal and salivary glands. Structure–activity relationships in physalaemin and eledoisin are similar, though some changes have a more marked effect on potency in one than in the other. The *C*-terminal pentapeptides are the smallest fragments with appreciable activity (1—3%); the *C*-terminal hexapeptides are considerably more active (20—50%) and the *C*-terminal octapeptides are as active as the natural peptides.[371] Because of the prompt nature of their contractile action on extravascular smooth muscle these peptides have been named 'tachykinins' to distinguish them from the group of slow-acting kinins, the true bradykinins.[372]

Several analogues of the *C*-terminal octapeptide of eledoisin with backbone modifications (hydrazino-, substituted hydrazino-, and aza-amino-acids in positions 5, 7, and 9) were reported last year.[373] Their enzymic

[371] V. Erspamer and P. Melchiorri, *Pure Appl. Chem.*, 1973, **35**, 495.

[372] V. Erspamer, *Ann. Rev. Pharmacol.*, 1971, **11**, 327.

[373] P. Oehme, J. Bergmann, M. Falck, J. G. Reich, W. E. Vogt, H. Niedrich, J. Pirrwitz, C. H. Berseck, and F. Jung, *Acta Biol. Med. Germ.*, 1972, **28**, 109; P. Oehme, J. Bergmann, H. G. Müller, R. Grupe, H. Niedrich, W. E. Vogt, and F. Jung, *ibid.*, p. 121; H. Niedrich and J. Pirrwitz, *J. prakt. Chem.*, 1972, **314**, 735; H. Niedrich and P. Oehme, *ibid.*, p. 759.

stability has also been investigated.[374, 375] Analogues of the *C*-terminal hexapeptide reported this year are listed in Table 19.[375–378] Peptide (330) lowered the blood pressure in anaesthetized dogs at a dose of 0.01 μg kg^{-1} (i.v.); in order to produce a similar effect the related *N*-dimethylamino-

Table 19 *Analogues of eledoisin*

Compound number	*Structure*	*Relative potency* [a]	*Ref.*
(315)	Lys-Phe-Ile-Gly-Leu-Met-NH_2	100	375
(316)	Lys-Phlac-Ile-Gly-Leu-Met-NH_2 [b]	90	375
(317)	Lys-Phe-HyMeV-Gly-Leu-Met-NH_2 [c]	10	375
(318)	Lys-Phe-HyIv-Gly-Leu-Met-NH_2 [d]	6	377
(319)	Lys-Phe-Ile-Glyc-Leu-Met-NH_2 [e]	30	375
(320)	Lys-Phe-Ile-Gly-HyIc-Met-NH_2 [f]	120	375
(321)	Lys-Phlac-Ile-Gly-HyIc-Met-NH_2	30	375
(322)	Lys-Phe-Ile-Gly-Leu-HyMtb-NH_2 [g]	0.1	377
(323)	Lys-MePhe-Ile-Gly-Leu-Met-NH_2	10	376
(324)	Lys-Phe-MeIle-Gly-Leu-Met-NH_2	0.1	376
(325)	Lys-Phe-Ile-Sar-Leu-Met-NH_2	200	376
(326)	Lys-Phe-Ile-Gly-MeLeu-Met-NH_2	120	376
(327)	Lys-MePhe-Ile-Gly-MeLeu-Met-NH_2	10	376
(328)	Lys-Phe-Ile-Gly-Leu-MeMet-NH_2	0.1	377
(329)	Lys-Phe-Ile-Gly-MeIle-Met-NH_2	< 0.2	376
(330)	Ala-Phe-Val-Gly-Leu-Met-NH_2		378
(331)	Boc-Ala-Phe-[NCHCO]-Gly-Leu-Met-NH_2 (N(CH$_3$)$_2$; HC(CH$_3$)$_2$)		378
(332)	Ala-Phe-Val-Gly-[NCHCO]-DL-Met-NH_2 (N(CH$_3$)$_2$; CH$_2$CH(CH$_3$)$_2$)		378

[a] Expressed as a percentage of peptide (315) on the basis of the weight required to cause a 20 mmHg fall in rabbit blood pressure; [b] Phlac = β-phenyl-lactic acid; [c] HyMeV = α-hydroxy-β-methylvaleric acid; [d] HyIv = α-hydroxyisovaleric acid; [e] Glyc = glycolic acid; [f] HyIc = α-hydroxyisocaproic acid; [g] HyMtb = α-hydroxy-γ-methylthiobutyric acid.

compounds (331) and (332) required 2—10 μg kg^{-1} and 50—100 μg kg^{-1}, respectively.

Substance P isolated from horse intestine is identical with substance P (314) isolated from bovine hypothalami.[379] A conventional synthesis of the

[374] P. Oehme, W. E. Vogt, H. Niedrich, S. Katzwinkel, J. Bergmann, and F. Jung, *Acta Biol. Med. Germ.*, 1973, **30**, 415; P. Oehme, S. Katzwinkel, W. E. Vogt, and H. Niedrich, *Experientia*, 1973, **29**, 215.
[375] H. Sugano, K. Higaki, and M. Miyoshi, *Bull. Chem. Soc. Japan*, 1973, **46**, 226.
[376] H. Sugano, K. Higaki, and M. Miyoshi, *Bull. Chem. Soc. Japan*, 1973, **46**, 231.
[377] H. Sugano, *Bull. Chem. Soc. Japan*, 1973, **46**, 2168.
[378] H. Immer, V. Nelson, W. Robinson, and M. Götz, *Annalen*, 1973, 1789.
[379] R. O. Studer, A. Trzeciak, and W. Lergier, *Helv. Chim. Acta*, 1973, **56**, 860.

undecapeptide has been reported [380, 381] and the product, and the *C*-terminal tetra-, penta-, hexa-, hepta-, and octa-peptide amides, have been assayed for hypotensive and smooth-muscle contracting effects.[381] In comparison with substance P, the shorter peptides were less active in reducing blood pressure and heart rate in the anaesthetized rat, and the heptapeptide amide was more active in contracting the isolated guinea-pig ileum.

Extracts of natural substance P have long been associated with an action on the central nervous system. The synthetic peptide (314) also affects the central nervous system; when administered intramuscularly it apparently penetrates the blood–brain barrier,[382] it reduces aggressive behaviour in mice,[382] abolishes abstinence symptoms in morphinized mice,[382] potentiates tremor caused by lysergic acid diethylamide,[383] and abolishes the excitement induced by amphetamine.[383] Substance P, eledoisin, and physalaemin (Table 20) exert a powerful depolarizing action on the motor neurons of

Table 20 *Depolarizing activity of tachykinin peptides (from ref.* 384)

	Structure	*Relative potency* [a]
(Physalaemin)	Glp-Ala-Asp-Pro-Asn-Lys-Phe-Tyr-Gly-Leu-Met-NH_2	1500
(Eledoisin)	Glp-Pro-Ser-Lys-Asp-Ala-Phe-Ile-Gly-Leu-Met-NH_2	2000
(Substance P)	Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH_2	200
	Lys-Phe-Ile-Gly-Leu-Met-NH_2	770
	Lys-Phe-Tyr-Gly-Leu-Met-NH_2	150
	$NH_2(CH_2)_5CO$-Phe-Tyr-Gly-Leu-Met-NH_2	130
	Lys-Phe-Ile-Gly-MeLeu-Met-NH_2	50
	Lys-Phe-Ile-Gly-HyIc-Met-NH_2	30
	Lys-Phe-Gly-Gly-Leu-Met-NH_2	2
	Lys-Phe-Lac-Gly-Leu-Met-NH_2	2
	Lys-MePhe-Ile-Gly-Leu-Met-NH_2	1
	Lys-MePhe-Ile-Gly-MeLeu-Met-NH_2	1

[a] Ratio of the molar concentration of L-glutamate to that of the test peptide which caused the same amplitude of depolarization.

isolated spinal cord of the frog.[384] The depolarizing action is not affected when synaptic transmission is blocked by tetrodotoxin or by lowered Ca^{2+} levels; this indicates that the action is direct on the motor neurons and supports suggestions that substance P is, or is related to, the excitatory transmitter of primary sensory neurons (bradykinin and angiotensins I and II also depolarize motor neurons, but indirectly, by liberating the excitatory transmitter from nerve terminals synapsing with the motor-neurons). Several analogues are also active (Table 20); it seems that the

[380] H. Yajima and K. Kitagawa, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 682.
[381] H. Yajima, K. Kitagawa, and T. Segawa, *Chem. and Pharm. Bull. (Japan)*, 1973, **21**, 2500.
[382] P. Stern and J. Hadzovic, *Arch. Internat. Pharmacodyn.*, 1973, **202**, 259.
[383] P. Stern, *J. Pharm. Pharmacol.*, 1973, **25**, 259.
[384] S. Konishi and M. Otsuka, *Brain Res.*, 1974, **65**, 397.

amino-acid residues at positions 7 and 8 are particularly important for activity.

9 Selected Proteins

Each year many publications are concerned with the effects of structural modifications on the biological activity of natural proteins. This review is limited to work with smaller proteins (or large polypeptides) in which chemical synthesis has made a contribution to the unravelling of structure–function relationships.

The enzyme ribonuclease A remains the pre-eminent example; the story of the total synthesis of this enzyme[385] and the exploitation of the S-peptide–S-protein system as a model for the interaction of hormones with their receptors[386] have been the subjects of new reviews. Syntheses of S-peptide-(1—15) analogues in which the phenylalanine residue in position 8 was substituted with L-*p*-fluorophenylalanine or uniformily enriched L-[^{13}C]phenylalanine were recently described.[387] When mixed with natural S-protein (residues 21—124 of ribonuclease A) the products were about as active as normal synthetic S-peptide-(1—15) or native S-peptide-(1—20) in regenerating enzymic activity. The specifically labelled peptides and complexes were then employed in ^{13}C and ^{19}F n.m.r. studies, where the precisely located signals helped overcome the problems of resonance resolution and assignment normally associated with n.m.r. studies of proteins. Differences observed between the spectra of the S-peptide-(1—15) analogues and of their complexes with S-protein were interpreted as reflecting the formation, in the complexes, of an 8-phenylalanine-containing amino-terminal α-helix, and the introduction of this phenylalanine residue into a non-polar, sterically rigid environment. Addition of 2′-cytidine monophosphate to the complexes produced further spectral changes suggestive of localized conformational changes around the phenylalanine side-chain.[387]

Enzymic removal of four to six of the *C*-terminal residues from ribonuclease A gave inactive molecules which could be reactivated by admixture with the overlapping synthetic nonapeptide (residues 116—124) (see ref. 1 for summary). The product with six *C*-terminal residues removed, *i.e.* ribonuclease-(1—118), has been further digested to give (1—115)- and (1—114)-ribonuclease fragments.[388] These were largely unfolded at 25 °C whereas ribonuclease-(1—118) had a transition temperature of 33 °C. Both bound with the synthetic nonapeptide (116—124) to give non-overlapping complexes in which the peptide is bound about 10 times as

[385] R. B. Merrifield, *The Harvey Lectures*, (1971–72), Academic Press Inc. p. 143, published 1973.
[386] F. M. Finn and K. Hofmann, *Accounts Chem. Res.*, 1973, **6**, 169.
[387] I. M. Chaiken, M. H. Freedman, J. R. Lyerla, and J. S. Cohen, *J. Biol. Chem.*, 1973, **248**, 884.
[388] R. Hayashi, S. Moore, and R. B. Merrifield, *J. Biol. Chem.*, 1973, **248**, 3889.

strongly as in the overlapping complex with ribonuclease-(1—118). The maximum regenerated enzymic activity (about 50%) was the same for all three complexes. However, the synthetic tetradecapeptide (residues 111—124) regenerated 98% activity with ribonuclease-(1—118), but only 50% activity with the shorter sequences (1—114) and (1—115). The removal of residues 116—118 seems therefore to unfold the chain to an extent not completely reversible by complex formation with the tetradecapeptide; consequently the precise orientation of the 119-histidine residue, necessary for full catalytic activity, is not achieved. The tyrosine residue in position 115 is apparently not essential for complex formation or enzymic activity.[388]

The contribution made by this semi-synthetic approach to the study of structure–activity relationships in staphylococcal nuclease, and the forces involved in folding of the protein chain, is clearly illustrated in the 1972 Nobel Chemistry Prize lecture.[389] The work has been extended to a series of analogues of nuclease-T-(6—48) with progressively longer deletions at the carboxy-terminus.[390] The binding characteristics of these analogues with native nuclease-T-(49—149) to produce active complexes indicate that the residues 43—48 are not necessary for complex formation. However, the presence of a residue at position 44 is critical for regeneration of normal enzymic activity; some activity was retained when an alanine residue was substituted for the normal threonine at this position, indicating involvement of the peptide linkage with the catalytically important 43-glutamic acid residue rather than a major contribution of the threonine side-chain.[390] An analogue of nuclease-T-(6—48) with alanine replacing aspartic acid in position 19 appeared to form a complex with nuclease-T-(49—149) but no enzymic activity was generated.[390]

Fourteen analogues and fragments of the 1—74 sequence of *E. coli* apo-acyl carrier protein (77 amino-acids) have been prepared by total synthesis, and a 4′-phosphopantetheine prosthetic group has been introduced enzymically into each.[391] The ability of the products to act as substrates for the enzymes malonyl-coenzyme A–acyl carrier-protein transacylase and β-ketoacyl–acyl carrier-protein synthetase was examined. The single methionine residue in position 44, and the lysine residues in positions 8, 9, and 18, could be replaced by norleucine residues with retention of activity. Replacement of the single arginine residue at position 6 with N^{ε}-nitroarginine did not abolish activity but replacement with norleucine gave an analogue inactive towards malonyl-coenzyme A–acyl carrier-protein transacylase. The fragment (1—61) retained significant activity, indicating that much of the *C*-terminal region is not essential for activity; omission of more than three residues from the *N*-terminus resulted in a significant fall in activity.[391]

[389] C. B. Anfinsen, *Science*, 1973, **181**, 223.
[390] G. R. Sanchez, I. M. Chaiken, and C. B. Anfinsen, *J. Biol. Chem.*, 1973, **248**, 3653.
[391] W. S. Hancock, G. R. Marshall, and P. R. Vagelos, *J. Biol. Chem.*, 1973, **248**, 2424.

Apolipoprotein-alanine, a 79 amino-acid residue polypeptide isolated from the very low-density lipoproteins of human plasma, has been used as a simple system to study phospholipid–protein interactions. Fragments corresponding to the sequences 61—79, 55—79, 48—79, and 41—79 of the protein have been synthesized and their ability to bind phosphatidylcholine has been assessed by c.d. ultracentrifugation, and inhibition of β-hydroxybutyrate apodehydrogenase reactivation measurements.[392] The sequences 61—79 and 55—79 did not bind significantly but the sequence 48—79 had significant lipid-binding properties, and these were markedly increased in the longest fragment, 41—79. The *N*-terminal region of this apolipoprotein would thus seem inessential for the binding with phosphatidyl choline.

Attempts to synthesize the enzyme lysozyme have given disappointing results owing to the synthetic difficulties involved.[393] The enzymic activities of the final products were too low to provide a meaningful basis for the synthesis and assay of analogues for structure–activity correlations.[393] Progress has been reported towards the total synthesis of ribonuclease T_1 [394] and bakers' yeast iso-1-cytochrome *c*.[4f, 395]

Toxic proteins of low molecular weight isolated from snake venoms have been extensively investigated in the past few years and figure prominently in a current series of multidisciplinary articles on toxinology [396] (an updated review of their chemistry has also appeared [397]). They contain between 57 and 74 amino-acid residues in a single polypeptide chain cross-linked by four or five disulphide bridges. The complete sequences of some 30 members of the groups are known, and they exhibit considerable structural homology. Biologically, they fall into two distinct categories, cardiotoxins and neutrotoxins. Knowledge of structure–activity relationships has been acquired only by chemical modification of specific functional groups and by partial enzymic digestion.[397] This, together with physicochemical data, has been used to construct a model of a possible three-dimensional structure of the neurotoxins.[398] Only one synthetic paper, briefly reporting the synthesis of a product with about 20% of the neurotoxic activity of cobratoxin, has appeared.[399]

[392] J. T. Sparrow, A. M. Gotto, and J. D. Morrisett, *Proc. Nat. Acad. Sci. U.S.A.*, 1973, **70**, 2124.
[393] J. J. Sharp, A. B. Robinson, and M. D. Kamen, *J. Amer. Chem. Soc.*, 1973, **95**, 6097.
[394] K. Kawasaki, R. Camble, G. Dupuis, H. Romovacek, H. T. Storey, C. Yanaihara, and K. Hofmann, *J. Amer. Chem. Soc.*, 1973, **95**, 6815.
[395] L. Moroder, G. Borin, F. Marchiori, and E. Scoffone, *Biopolymers*, 1973, **12**, 477; L. Moroder, F. Marchiori, G. Borin, and E. Scoffone, *ibid.*, p. 729.
[396] Various authors, *Experientia*, 1973, **29**, 1317–1334, 1453–1471; *ibid.*, 1974, **30**, 2–12, 121–129.
[397] A. T. Yu, *Ann. Rev. Biochem.*, 1973, **42**, 235.
[398] L. Ryden, D. Gabel, and D. Eaker, *Internat. J. Peptide Protein Res.*, 1973, **5**, 261.
[399] H. Aoyagi, H. Yonezawa, N. Takahashi, T. Kato, N. Izumiya, and C.-C. Yang, *Biochim. Biophys. Acta*, 1972, **263**, 823.

Author Index