Amino-acids, Peptides, and Proteins—Volume 3

Specialist Periodical Reports

The Chemical Society

SBN: 85186 024 9

A Specialist Periodical Report

Amino-acids, Peptides, and Proteins

Volume 3

A Review of the Literature Published during 1970

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

SBN: 85186 024 9

The Chemical Society
Burlington House, London, W1V 0BN

Reporters

C. C. F. Blake, *University of Oxford*
J. R. Brocklehurst, *University of Oxford*
B. W. Bycroft, *University of Nottingham*
D. G. Dalgleish, *University of Oxford*
J. S. Davies, *University College of Swansea, University of Wales*
R. D. Gillard, *University of Kent*
R. Henson, *University of Oxford*
J. H. Jones, *University of Oxford*
S. H. Laurie, *Leicester Polytechnic*
P. H. Lloyd, *University of Oxford*
A. R. Peacocke, *University of Oxford*
R. Perham, *Cambridge University*
N. C. Price, *University of Oxford*
R. M. Stephens, *Portsmouth Polytechnic*
J. O. Thomas, *Cambridge University*

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 third Report continues in the pattern of the earlier volumes, reviewing papers relevant to the chemistry of amino-acids, peptides, and proteins appearing in the main journals in 1970; in all, some 2300 references are quoted. The unfortunate cessation of publication of *Current Chemical Papers* has made systematic coverage more difficult, and we shall be grateful if our attention is drawn to any important omissions in the bibliography; some December issues of journals were delayed by the postal strike, and references from these journals will be included in the next volume. It will be recalled that the extent of new work on metal derivatives was considered insufficient for annual review and this field was not covered in Volume 2; Chapter 5 therefore surveys this literature for 1969 and 1970. Work on the synthesis of cyclic peptides, formerly appearing in Chapter 3 (Peptide Synthesis) is now included in Chapter 4 (Peptides with Structural Features Not Typical of Proteins), in which structural work on such peptides is discussed. We would draw attention to the collected references now a feature of several chapters, *e.g.* the list of amino-acids whose syntheses were first reported during the year (Chapter 1), the list of peptide syntheses, and the (selective) list of new intermediates useful in synthesis (Chapter 3). We would remind readers that extracts relevant to this field from the Tentative Rules of the I.U.P.A.C.–I.U.B. Commission on Biochemical Nomenclature were reprinted in Volume 2.

As in earlier volumes, there is an author index but instead of a subject index (the preparation of which would delay publication unduly) there is an extended list of contents; an inspection of this will, we hope, enable the reader to see readily which sections are likely to contain the material he seeks.

Once more I record my gratitude to the contributors who have taken time from their research to produce this survey for their colleagues in the field.

G. T. YOUNG

Contents

Abbreviations

Abbreviations for amino-acids and their use in the formulation of derivatives follow the relevant Tentative Rules of the I.U.P.A.C.–I.U.B. Commission on Biochemical Nomenclature, extracts of which were reprinted in Chapter 5 of Volume 2 of this Series.

Other abbreviations which have been used without definition 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-biphenyl)-isopropoxycarbonyl
BSA	bovine serum albumin
c.d.	circular dichroism
Cha	cyclohexylamine
Cm	carboxymethyl
Cmc	*S*-carboxymethylcysteine
Dce	2,2-diethoxycarbonyl
Dcha	dicyclohexylamine
DMF	*NN*-dimethylformamide
DNA	deoxyribonucleic acid
Dnp	2,4-dinitrophenyl
Dns	1-dimethylaminonaphthalene-5-sulphonyl (dansyl)
Dopa	3,4-dihydroxyphenylalanine
Ec	ethylcarbamoyl
EDTA	ethylenediamine tetra-acetate
e.s.r.	electron spin resonance
Gal	galactose
g.l.c.	gas–liquid chromatography
Glc	glucose
Glp or Pca	pyrrolid-2-one-5-carboxylic acid
Glx	glutamic acid or glutamine (not yet determined)
GTP	guanosine 5′-triphosphate
i.r.	infrared
Man	mannose
NAD	nicotinamide-adenine dinucleotide (NAD^+, oxidised, NADH, reduced)
n.m.r.	nuclear magnetic resonance
ONSu	succinimido-oxy
OPcp	pentachlorophenoxy
OPic	4-picolyloxy

o.r.d.	optical rotatory dispersion
OTcp	2,4,5-trichlorophenoxy
Pipoc	piperidino-oxycarbonyl
Pth-Gly	the phenylthiohydantion 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 B. W. BYCROFT

Amino-acids continue to attract attention from all branches of chemistry and related disciplines. Unfortunately, this wide diversity of interest poses considerable problems in reviewing the year's literature since only in a limited number of cases can any particular development be considered in depth. However, there can be no doubt that this year some of the most interesting advances have occurred within the field of asymmetric synthesis, and these have been considered worthy of more detailed coverage. The pattern already established for this chapter is maintained with a few minor exceptions, and the emphasis remains on α-amino-acids. Regrettably, it has only been possible to cover biochemical aspects when they relate directly to the chemistry.

1 Naturally Occurring Amino-acids

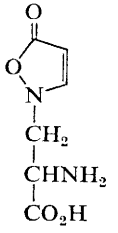
A. Introduction.—Amino-acids with novel structures continue to be isolated from natural sources both in the free state and from peptide and protein hydrolysates. Spectroscopic methods have played an important rôle in structure determination, in particular n.m.r. spectroscopy and mass spectrometry. However, it is interesting to note the increasing application of *X*-ray crystallographic analysis in cases where other spectroscopic methods have proved ambiguous (see Section 3). Those amino-acids whose structures have been confirmed by synthesis are presented in the list of newly synthesised amino-acids in Section 2.

The presence of amino-acids in marine sediments is attracting increasing attention. These acids are undoubtedly of natural origin since they are optically active, and the degree of racemisation observed with increasing age appears to offer an alternative method of dating.[1, 2]

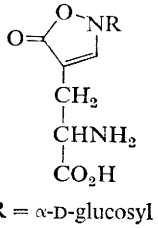
B. New Natural Free Amino-acids.—A number of new plant amino-acids have been described. Further work on the amino-acids formed in the early stages of germination of pea seedlings has resulted in the isolation of the isoxazolinones (1) and (2).[3, 4] The structures are based on extensive

[1] K. A. Kvenvolden, E. Peterson, and F. S. Brown, *Science*, 1970, **169**, 1079.
[2] J. L. Bada, B. P. Luyendyk, and J. B. Maynard, *Science*, 1970, **170**, 730.
[3] F. Lambein, N. Schamp, L. Vandendriessche, and R. van Parijs, *Biochem. Biophys. Res. Comm.*, 1969, **37**, 375.
[4] F. Lambein and R. van Parijs, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 557.

spectroscopic data and degradative evidence. Treatment of (1) with mild base followed by acid hydrolysis affords αβ-diaminopropionic acid, whereas similar treatment of (2) gives D-glucose and glutamic acid. Both (1) and (2) are uncommonly sensitive to u.v. radiation and it has been suggested that they may play a rôle in some photobiological mechanism. It is also reported that they are present in the seedlings of several other leguminous plants.[4] γ-Cyano-L-α-aminobutyric acid has been identified

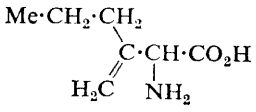


(1)



R = α-D-glucosyl

(2)



(3)

for the first time as a biological substance.[5] It accumulates when inorganic cyanide is administered to young cultures of *Chromobacterium violaceum*. β-Methylene-L-norleucine (3) has been isolated from the carpophores of *Amanita vaginata*[6] and is a further example of a dehydro-amino-acid produced by either a fungus or a micro-organism. *N*-Jasmonoyl- and *N*-dihydrojasmonoyl-isoleucine (stereochemistry at the amino-acid centre not defined) are produced by the fungus *Gibberella fujikuroi*[7] and 2-*N*,6-*N*-di-(2,3-dihydroxybenzoyl)-L-lysine has been isolated from an iron-deficient culture of *Azobacter vinelandii*.[8]

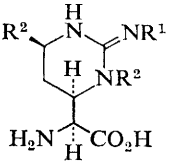
A number of new amino-acids have been detected in human urine.[9] Both *guanidino-NN*-dimethylarginine and *-NN'*-dimethylarginine were isolated in crystalline form; their structures were established by detailed spectroscopic analysis and chemical degradation to ornithine, and finally confirmed by synthesis. In addition, $N^{\varepsilon}N^{\varepsilon}$-dimethyl-lysine and $N^{\varepsilon}N^{\varepsilon}N^{\varepsilon}$-trimethyl-lysine were observed; although these had been previously obtained from certain protein hydrolysates, they had not been previously encountered in the free state.

β-Putreamine, a β-amino-acid, occurs in relatively large amounts in bovine brain tissue[10] and the quaternary β-amino-acids anodendrine and *allo*-anodendrine are present in the plant *Anodendron affine*.[11]

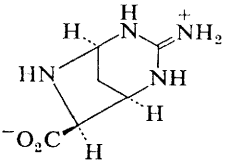
C. New Amino-acids from Peptide Hydrolysates.—Two new guanidino-amino-acids have been reported. Hydrolysis of the tuberculostatic anti-

[5] M. M. Brysk and C. Ressler, *J. Biol. Chem.*, 1970, **245**, 1156.
[6] R. Vervier and J. Casimir, *Phytochemistry*, 1970, **9**, 2059.
[7] B. E. Cross and G. R. B. Webster, *J. Chem. Soc.* (*C*), 1970, 1839.
[8] J. L. Corbin and W. A. Bulen, *Biochemistry*, 1969, **8**, 757.
[9] Y. Kakimoto and S. Akazawa, *J. Biol. Chem.*, 1970, **245**, 5751.
[10] T. Shiba, I. Kubota, and T. Kaneko, *Tetrahedron*, 1970, **26**, 4307.
[11] K. Sasaki and Y. Hirata, *Tetrahedron*, 1970, **9**, 2119.

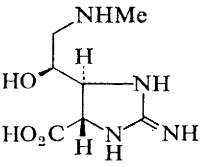
biotic tuberactinomycin affords tuberactidine (4) as well as viomycidine (6),[12] whereas hydrolysis of a new streptothricin-type antibiotic yields *N*-methylstreptolidine (7).[13] The relative and absolute configurations of stendomycidine have been established and are as shown (5).[14] It is noteworthy that the above-mentioned amino-acids, as well as enduracididine



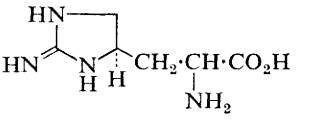
(4) $R^1 = H$, $R^2 = OH$
(5) $R^1 = Me$, $R^2 = H$



(6)



(7)



(8)

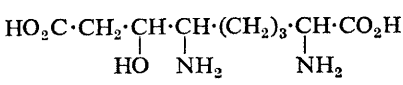
(8), reported last year, are all derived from microbial peptides and are related both structurally and stereochemically to L-arginine. But, as yet, there is no well-authenticated report of the isolation of arginine itself from microbial peptide hydrolysate.

The antibiotics edeine A and B, produced by a strain of *Bacillus brevis*, give on hydrolysis 2,6-diamino-7-hydroxyazelaic acid (9).[15] The stereochemistry has not been fully defined but it is suggested that the two amino-groups have the same relative chirality as in *meso*-pimelic acid. Acid hydrolysis of cycloheptamycin yields L-β-hydroxynorvaline [16] and *N*-methyl-*allo*-isoleucine, and L-β-hydroxyglutamic acid has been isolated from the hydrolysate of a peptide antibiotic complex.[17]

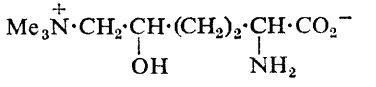
Further work on the hydrolysates of diatom cell walls has resulted in the isolation of $N^{\epsilon}N^{\epsilon}N^{\epsilon}$-trimethyl-δ-hydroxy-L-lysine (10). The structure has been determined by a detailed analysis of the spectral data, including 220 MHz ^{1}H n.m.r. spectra, and subsequently confirmed by synthesis.[18] Possible

[12] T. Nakamiya, T. Shiba, T. Kaneko, H. Sakakibara, T. Take, and J. Abe, *Tetrahedron Letters*, 1970, 3497.
[13] D. B. Borders, K. J. Sax, J. E. Lancaster, W. K. Hausmann, L. A. Mitscher, E. R. Wetzel, and E. L. Patterson, *Tetrahedron*, 1970, **26**, 3123.
[14] G. G. Marconi and M. Bodanszky, *J. Antibiotics*, 1970, **23**, 120.
[15] T. P. Hettinger and L. C. Craig, *Biochemistry*, 1970, **9**, 1224.
[16] W. O. Godtfredsen, S. Vangedal, and D. W. Thomas, *Tetrahedron*, 1970, **26**, 4931.
[17] J. Shoji and R. Sakazahi, *J. Antibiotics*, 1970, **23**, 418.
[18] T. Nakajima and B. E. Volcani, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 28.

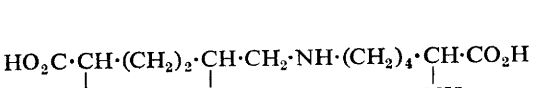
similarities between the structure and function of the cell-wall protein and collagen were noted. Sodium borohydride reduction of the intermolecular cross-links in collagen fibrils, followed by acid hydrolysis, affords N^{ε}-(5-amino-5-carboxypentyl)-δ-hydroxylysine (11) and its corresponding six-membered lactone,[19] as well as δε-dihydroxynorleucine.[20] It is debatable



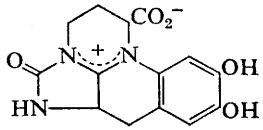
(9)



(10)



(11)



(12)

whether these can be classified as true natural products, but this method has proved valuable in locating the sites of cross-linking in collagen and it was felt that the structures are of sufficient interest to warrant inclusion. The same argument applies also to the isolation, and characterisation by an *X*-ray analysis, of (12) from the hydrolysate of a fluorescent peptide produced by iron-deficient cultures of *Azobacter vinelandii*.[21]

A number of β-amino-acids have been isolated for the first time from peptide hydrolysates: δ-hydroxy-β-lysine from tuberactinomycin;[12] iso-serine and β-tyrosine from edeine A and B;[15] and β-amino-β-phenyl-propionic acid from a cyclic tetrapeptide produced by the lichen *Roccella canariensis*.[22]

D. Occurrence of Known Amino-acids.—It has been decided to include in this section only those amino-acids which are rarely encountered or which exhibit interesting biological activity. The new basic amino-acids isolated from human urine have been described (see above); in fact the three possible N^{ε}-methyl derivatives of lysine were obtained, as well as glucosyl-galactosyl- and galactosyl-5-hydroxylysines.[9] The concentrations of the *N*-methyl derivatives of arginine and lysine were unchanged either by oral loading of these amino-acids or by a protein-free diet, and it was tentatively suggested that these compounds are derived from tissue protein. N^{ε}-Trimethyl-lysine has also been obtained from the hydrolysate of the cytochrome *c* derived from *Saccharomyces cerevisiae*.[23] It has previously been

[19] M. L. Tanzer and G. Mechanic, *Biochem. Biophys. Acta*, 1970, **207**, 548.
[20] G. Mechanic and M. L. Tanzer, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 1597.
[21] J. L. Corbin, I. L. Karle, and J. Karle, *Chem. Comm.*, 1970, 186.
[22] G. Bohman, *Tetrahedron Letters*, 1970, 3065.
[23] R. J. Delange, A. N. Glazer, and E. L. Smith, *J. Biol. Chem.*, 1970, **245**, 3325.

observed in plant cytochromes but is absent in the cytochrome *c* produced by animal tissue. N^{ω}-Methylarginine has been detected in histones from rat-liver cell nuclei.[24]

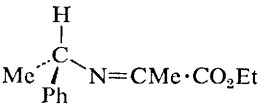
The West African legume *Griffonia simplicifola*, reputed to possess marked physiological activity, has been shown to contain relatively large amounts of 5-hydroxy-L-tryptophan in the free amino-acid pool.[25] N^{δ}-Acetylornithine has been isolated from the seeds of the bush bean, *Phaseolus vulgaris*.[26] Acid hydrolysis of a peptide produced by the plant *Canthium euryoides* gives *NN*-dimethyl-L-phenylalanine and L-*threo*-β-phenylserine,[27] whereas the antibiotic alamethicin, on hydrolysis, affords 2-amino-isobutyric acid.[28]

2 Chemical Synthesis and Resolution of Amino-acids

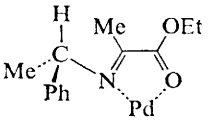
A. Introduction.—The problem of asymmetric synthesis of α-amino-acids has received considerable attention, and important advances have been made both as regards optical efficiency and yield of material. The majority of the work has centred on a continuation and extension of asymmetric syntheses of α-amino-acids from their corresponding α-keto-acids, for which several methods have already been employed. These include: (*a*) hydrogenation of C=N double bonds using an optically active catalyst; (*b*) reduction of Schiff bases derived from an α-keto-acid derivative and an optically active amine, and (*c*) reduction of Schiff bases obtained from an amine and an optically active α-keto-acid derivative. The earlier literature on hydrogenation using asymmetrically-modified catalysts has been reviewed [29, 30] and further work using Raney-nickel modified with histidine has been described.[31] A detailed investigation on the course of the reaction is reported, but in general the optical efficiency in the process is relatively poor. An extensive study on the sodium borohydride reduction of the Schiff bases of α-keto-esters with optically active α-ethyl- and α-methylbenzylamine has revealed that the optical purities of the resulting amino-acids are lower than those obtained by catalytic hydrogenation.[32] The effect of both temperature [33] and solvent [34] on the catalytic hydrogenation of the above-mentioned Schiff bases has been studied. Low-temperature hydrogenation of (13) and hydrolysis affords (*S*)(*i.e.* L)-alanine (optical purity 60%). The optical activity decreases sharply with a rise in the reaction

[24] W. K. Paik and S. Kim, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 224.
[25] L. E. Fellows and E. A. Bell, *Phytochemistry*, 1970, **9**, 2389.
[26] R. M. Zacharius, *Phytochemistry*, 1970, **9**, 2047.
[27] G. Boulvin, R. Ottinger, M. Pais, and G. Chiurdoglu, *Bull. Soc. chim. belges*, 1970, **78**, 583.
[28] J. Payne, R. Jakes, and B. S. Hartley, *Biochem. J.*, 1970, **117**, 757.
[29] Y. Izumi, *Tampakushitsu Kakusan Koso*, 196, **12**, 301.
[30] E. I. Klabunovskii and E. S. Levivina, *Uspekhi Khim.*, 1970, **39**, 2154.
[31] Y. Izumi, H. Takizawa, K. Nakagawa, R. Imamura, M. Imaida, T. Ninomiya, and S. Yajima, *Bull. Chem. Soc. Japan*, 1970, **43**, 1792.
[32] K. Harada and J. Okhashi, *Bull. Chem. Soc. Japan*, 1970, **43**, 960.
[33] K. Harada and T. Yoshida, *Chem. Comm.*, 1970, 1071.
[34] K. Harada and T. Yoshida, *Bull. Chem. Soc. Japan*, 1970, **43**, 921.

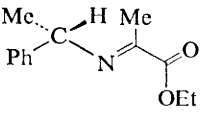
temperature, becoming zero at about 17 °C. Then the configuration is inverted and the optical activity of the resulting D-alanine increases steadily until it reaches a maximum (optical purity 43%) at about 50 °C, finally decreasing at higher temperatures. It is suggested that at lower temperatures the preferred conformation of the substrate on the catalyst



(13)



(14)

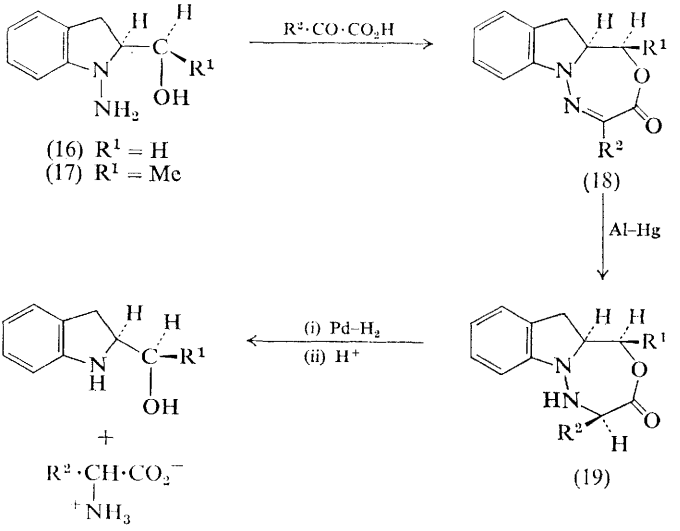


(15)

surface is as shown in (14) and that hydrogenation occurs from the least hindered side, generating an L-alanine derivative; at higher temperatures the conformer (15) predominates, which on reduction yields a D-alanine ester, and further increase in temperature results in complete conformational mobility and the consequent fall in optical activity. Similar changes in conformer population are invoked to account for the increase in optical efficiency in solvents with a low dielectric constant. It is claimed that with polar solvents there is an increase in the concentration of the conformer (15).

The inherent problem of conformational mobility of the substrate in all the previously described asymmetric syntheses has been elegantly solved in a new important synthesis which is essentially a combination of methods (*b*) and (*c*), and is outlined in Scheme 1.[35, 36] The chiral reagents (16) and (17) have been synthesised and resolved, so that both enantiomers can be employed in the synthesis. Condensation of either (16) or (17) with an α-keto-ester affords the corresponding cyclic hydrazono-lactone (18), a chiral compound with limited conformational mobility. Reduction of (18) cannot be achieved under catalytic conditions but, with aluminium amalgam, (19) is formed in high yield. As expected, the addition of the hydrogen, to the α-carbon atom occurs from the least hindered side, *i.e.* from a direction which is *cis* to the hydrogen at C-2 of the indoline. The stereochemical efficiency of the synthesis of (19) is 80—90% when (16) is employed but rises to 96—99% for (17). Reduction and hydrolysis of (19) in the manner indicated affords the optically pure amino-acid and regenerates the

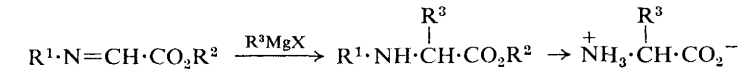
[35] E. J. Corey, R. J. McCaully, and H. S. Sachdev, *J. Amer. Chem. Soc.*, 1970, **92**, 2476.
[36] E. J. Corey, H. S. Sachdev, J. Z. Gougoutas, and W. Saenger, *J. Amer. Chem. Soc.*, 1970, **92**, 2488.



Scheme 1

chiral reagent. From a practical viewpoint this synthesis makes possible the facile and economical synthesis of 100% optically pure amino-acids.

A previously reported synthesis of optically active alanine,[37] employing a Strecker sequence from hydrogen cyanide and Schiff base between acetaldehyde and a chiral amine, has been extended to other amino-acids, and the optical efficiency and material yield have been substantially improved.[38] A new synthesis, involving the addition of a Grignard reagent to a Schiff base derived from an optically active amine and a glyoxylate ester, has been reported [39] (Scheme 2). The optical efficiency is not high but it offers a new general route to amino-acids.

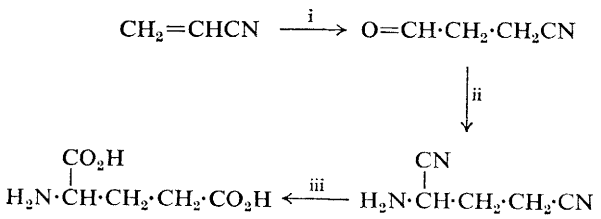


Scheme 2

Interest continues on the origin of amino-acids. Further work describing the synthesis of amino-acids in simulated primitive environments [40] and a speculative article on the origin of chiral compounds, with particular reference to amino-acids,[41] have been published.

[37] K. Harada and S. W. Fox, *Naturwiss.*, 1964, **51**, 106.
[38] M. S. Patel and M. Worsley, *Canad. J. Chem.*, 1970, **48**, 1881.
[39] J. C. Fiaud and H. B. Kagan, *Tetrahedron Letters*, 1970, 1813.
[40] A. Bar-Nun, N. Bar-Nun, S. H. Bauer, and C. Sagan, *Science*, 1970, **168**, 470.
[41] K. Harada, *Naturwiss.*, 1970, **57**, 114.

B. Protein and Other Naturally Occurring Amino-acids.—The methods described in the general introduction have, in many cases, been applied to the synthesis of protein amino-acids; in addition, further syntheses for D-alanine,[42] DL-aspartic acid,[43] and DL-glutamic acid [43] have been described. An industrial preparation of L-glutamic acid, although performed by a well-established route (Scheme 3),[44] is of considerable interest since it must



Reagents: i, CC–H_2; ii, NH_3,KCN; iii, H_3O^+

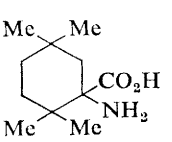
Scheme 3

presumably compete commercially with the naturally obtained material. The use of synthetic amino-acids in fortifying food materials is becoming increasingly important and has been reviewed.[45]

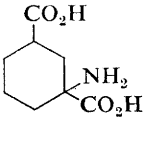
Interest in N^{ε}-methyl derivatives of lysine continues; a further synthesis of the trimethyl derivative has been reported,[46] as well as a new synthesis of the mono- and di-methyl derivatives.[9] The work on hydroxylation of phenylalanine to tyrosines under physiological conditions has been extended, and effective means of separation now appear to be available.[47] The unusual amino-acid α-amino-β-phenylbutyric acid, obtained from the hydrolysate of a microbial peptide, has been synthesised and the relative configuration determined.[48] Photocatalytic oxidation of glucose in the presence of a nitrate has been shown to give a variety of α-amino-acids.[49]

C. *C*-Alkyl- and Substituted *C*-Alkyl-α-amino-acids.—The synthesis of the novel cyclohexane-amino-acids (20) and (21) has been achieved by the Strecker procedure [50, 51] from the corresponding ketones and, in the case of (21), the *cis*- and *trans*-isomers have been separated.[51] The *cis*-isomer readily forms an anhydride, thus allowing a definitive assignment of relative stereochemistry. A novel general synthesis of perfluoroalkyl-α-amino-acids

[42] H. Matsuo, H. Kobayashi, and T. Tatsuno, *Chem. and Pharm. Bull. (Japan)*, 1970, **18**, 1693.
[43] S. Zen and E. Kaji, *Bull. Chem. Soc. Japan*, 1970, **43**, 2277.
[44] T. Yoshida, *Chem.-Ing.-Tech.*, 1970, **42**, 641.
[45] H. H. Ottenheym and P. J. Jenneskens, *J. Agric. Food Chem.*, 1970, **18**, 1010.
[46] J. Puskas and E. Tyihak, *Periodica Polytech.*, 1969, **13**, 261.
[47] M. Viscontini and G. Mattern, *Helv. Chim. Acta.*, 1970, **53**, 372.
[48] H. Arold, M. Eule, and S. Reissmann, *Z. Chem.*, 1969, **9**, 447.
[49] N. R. Dhar and S. K. Arora, *Proc. Nat. Acad. Sci., India*, 1969, **39**, 451.
[50] J. W. Cremlyn, R. M. Ellam, and T. K. Mitra, *J. Indian Chem. Soc.*, 1970, **8**, 218.
[51] J. D. Gass and A. Meister, *Biochemistry*, 1970, **9**, 842.

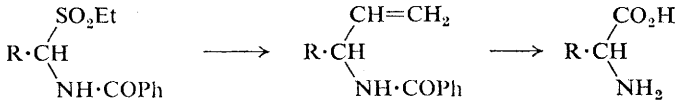


(20)



(21)

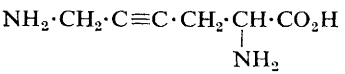
employs a perfluorocarboxylic acid anhydride as the starting material.[52] The anhydride is converted into the sulphone (22) *via* an oxazolone intermediate and treated with a Grignard reagent to give (23), which on oxidation and hydrolysis yields the required amino-acid (24). L-ω-Fluoro-*allo*-isoleucine has been prepared[53] and an improved synthesis of fluorinated



(22) (23) (24)

valine and norvaline derivatives reported.[54] Photochlorination of alanine affords a mixture of isomers, from which β-chloroalanine can be isolated.[55]

The D and L forms of the acetylenic acid (25) are conveniently prepared from 1,4-dichlorobutyne and acetamidomalonate.[56] It is suggested that (25) may be of value for the synthesis of lysine derivatives labelled in the

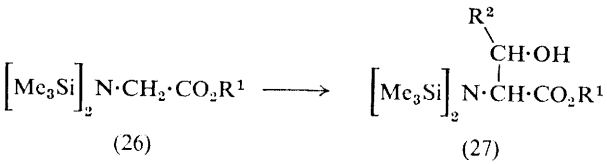


(25)

4- and 5-positions. Considerably enhanced yields of ornithino-alanine are obtained by condensing *N*-benzoyldehydroalanine with *N*-benzoylornithine and subsequent hydrolysis.[57]

D. α-Amino-acids with Aliphatic Hydroxy-groups in the Side-chain.—The synthesis of β-hydroxy-amino-acids by the reaction of a suitably protected glycine derivative with an aldehyde is a well-established method, but two interesting modifications have been reported. Treatment of *NN*-bis-(trimethylsilyl)glycine ester (26) with base, followed by reaction of the resultant carbanion with an aldehyde, affords (27).[58] Good yields are obtained if the aldehyde lacks α-hydrogen atoms, but enolisable aldehydes

[52] F. Weygand, S. Wolfgang, and W. Oettmeier, *Chem. Ber.*, 1970, **103**, 818.
[53] M. Hudlicky, V. Jelinek, K. Eisler, and J. Rudinger, *Coll. Czech. Chem. Comm.*, 1970, **35**, 498.
[54] R. M. Babb and F. W. Bollinger, *J. Org. Chem.*, 1970, **35**, 1438.
[55] T. Zaima, K. Mitsuhashi, I. Sasaji, and T. Asahara, *J. Chem. Soc. Japan, Ind. Chem. Sect.*, 1970, **73**, 319.
[56] A. C. A. Jansen, K. E. T. Kerling, and E. Havinga, *Rec. Trav. chim.*, 1970, **89**, 861.
[57] M. A. Febrer and P. Miro, *Invest. Inform. Text.*, 1969, **12**, 293.
[58] K. Rühlmann, K. D. Kaufmann, and K. Ickert, *Z. Chem.*, 1970, **10**, 393.



undergo aldol condensation under the basic conditions of the reaction. Similar observations had previously been made in the reaction between an aldehyde and copper bis glycinate. However, it is now claimed that, using the copper complex derived from the Schiff base of glycine and pyruvic acid, reaction occurs readily with mild bases and is applicable to a wide variety of aldehydes.[59]

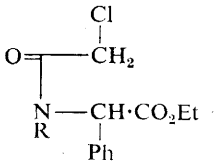
A number of *O*-glycosides of β-hydroxy-amino-acids have been prepared,[60–62] and also several phosphoglycerides of threonine.[63] Interest continues in L-homoserine and its derivatives [64] and a facile enzymic synthesis of *O*-alkyl-homoserines from *O*-acetyl-homoserine in the presence of an alcohol has been described.[65] An improved method for the synthesis of *threo*- and *erythro*-β-hydroxy-DL-aspartic acids from *cis*- and *trans*-epoxysuccinic acids has been claimed.[66]

E. Aromatic and Heterocyclic α-Amino-acids.—Interest continues in substituted phenylalanines because of their potential biological activity. Detailed accounts of the synthesis of various L-cyclodopa (5,6-dihydroxyindolin-2-carboxylic acid) derivatives by oxidative cyclisation of the corresponding L-3,4-dihydroxyphenylalanine have been published.[67, 68] The indane isostere of L-cyclodopa has been prepared,[69] as well as L-6-hydroxydopa by hydrobromic demethylation of the corresponding trimethoxyphenylalanine,[70] and L-*N*-bis-(2-chloroethyl)dopa.[71] Several new halogenated phenylalanines have been reported,[72] and in this context it is of interest to note that an earlier claim that *p*-chlorophenylalanine methyl ester is an aphrodisiac has been discounted.[73]

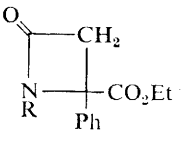
[59] T. Ichikawa, S. Maeda, Y. Araki, and Y. Ishido, *J. Amer. Chem. Soc.*, 1970, **92**, 5514.
[60] K. Kum, *Carbohydrate Res.*, 1970, **11**, 269.
[61] M. G. Vafina, E. M. Klimov, and T. G. Alieva, *Khim. Biokhim. Uglevodov, Mater. Vses. Konf., 4th* (1967) published 1969, p. 171.
[62] N. K. Kochetkov, V. A. Derevitskaya, L. M. Kikhosherstov, V. M. Kalinevich, and O. S. Novikova, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1969, 2509.
[63] J. W. Moore and M. Szelke, *Tetrahedron Letters*, 1970, 4423.
[64] A. Kase, K. Nakayama, and S. Kinoshita, *Agric. and Biol. Chem.* (*Japan*), 1970, **34**, 274, 282.
[65] Y. Murooka, K. Seto, and T. Harada, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 407.
[66] C. W. Jones, D. E. Leyden, and C. H. Stammer, *Canad. J. Chem.*, 1969, **47**, 4363.
[67] V. Wölke, A. Kaiser, W. Koch, and M. Scheir, *Helv. Chim. Acta*, 1970, **53**, 1704, 1708.
[68] N. Fischer and A. S. Dreiding, *Helv. Chim. Acta*, 1970, **53**, 1937.
[69] J. B. Taylor, J. W. Lewis, and M. Jacklin, *J. Medicin. Chem.*, 1970, **13**, 1226.
[70] B. A. Berkowitz, S. Spector, A. Brossi, A. Facella, and S. Teitel, *Experientia*, 1970, **26**, 982.
[71] M. N. Vasileva, V. S. Martynov, and A. Y. Berlin, *Zhur. org. Khim.*, 1970, **6**, 1677.
[72] R. E. Counsell, P. Desai, T. D. Smith, P. S. Chan, P. A. Weinhold, V. B. Rethy, and D. Burke, *J. Medicin. Chem.*, 1970, **13**, 1040.
[73] R. E. Whalen and W. G. Lutege, *Science*, 1970, **169**, 1000.

A number of *N*-uridyl-phenylalanine derivatives have been synthesised as analogues of naturally occurring nucleotides for an investigation into the mechanism of action of various enzyme systems.[74] Thyronine derivatives with an isopropyl group in the positions usually occupied by iodine atoms possess increased biological activity and it is claimed that this is related to the size of the isopropyl group. Several new derivatives have been prepared by standard methods and are listed.[75–77]

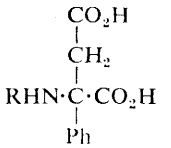
Intramolecular cyclisation of *N*-chloroacetyl-2-phenylglycine ester (28) with base produces the corresponding azetidinone (29) which undergoes



(28)



(29)



(30)

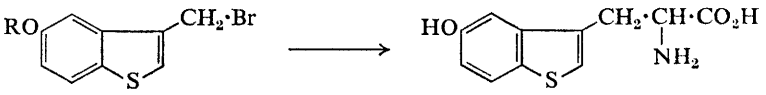
facile ring cleavage to yield a series of novel 2-phenylaspartic acid derivatives (30).[78] Reasonably high yields were obtained and it appears likely that this synthesis could be extended. An interesting modification of the *N*-formylaminomalonate route, which appears to offer a new general synthesis of substituted tryptophans, has been applied to the synthesis of a number of monofluorotryptophans. Alkylation of *N*-formylaminomalonate with morpholine and formaldehyde gives the expected Mannich base, which reacts with monofluoro-indoles to give, after hydrolysis, the required tryptophan in good yield.[79] A considerable number of new aromatic and heterocyclic amino-acids have been reported which were prepared by well-established synthetic routes, and those that have been synthesised for the first time are included in the list in Section H below.[80–84]

F. *N*-Alkyl-α-amino-acids.—Methylation of the amide nitrogen of *N*-benzyloxycarbonyl- and *N*-t-butoxycarbonyl-amino-acids with methyl iodide

[74] N. G. Shinskii, N. N. Preobrazhenskaya, Z. A. Shabarova, and M. A. Prokof'ev, *Zhur. obshchei Khim.*, 1970, **40**, 1114.
[75] T. Matsuura, T. Nagamachi, and A. Nishinaga, *Chem. and Pharm. Bull. (Japan)*, 1969, **17**, 2176.
[76] E. C. Jorgensen and J. Wright, *J. Medicin. Chem.*, 1970, **13**, 745.
[77] E. C. Jorgensen and J. Wright, *J. Medicin. Chem.*, 1970, **13**, 367.
[78] T. A. Martin, W. T. Cower, C. M. Combs, and J. Q. Carrigan, *J. Org. Chem.*, 1970, **35**, 3814.
[79] M. Beutov and C. Roffman, *Israel J. Chem.*, 1969, **7**, 835.
[80] I. I. Grandberg, L. F. Morozova, V. A. Moskalenko, and A. N. Kost, *Khim. geterotsikl. Soedinenii*, 1969, 1049.
[81] S. J. Norton and P. T. Sullivan, *J. Heterocyclic Chem.*, 1970, **7**, 699.
[82] J. D. Milkowski, F. M. Miller, E. M. Johnson, and N. Zenker, *J. Medicin. Chem.*, 1970, **13**, 741.
[83] M. Y. Lidak, Y. Y. Shluke, S. Y. Poritere, and Y. P. Shvachkin, *Khim. geterotsikl. Soedinenii*, 1970, 529.
[84] V. V. Kiselev and G. P. Menshikov, *Zhur. obshchei Khim.*, 1970, **40**, 914.

and silver oxide in dimethylformamide gives the corresponding *N*-methyl-amino-acid derivatives in excellent yield.[85] The methylation occurs without racemisation of the α-centre and the reaction could no doubt be extended to other alkyl halides. Reference to the novel synthesis of *N*-alkyl-2-phenyl-aspartic acid derivatives has been made above, and several *N*-pyridyl derivatives of the same acid have been obtained by addition of an appropriate amine to a maleate ester.[86] Several new substituted proline derivatives have been reported.[87, 88]

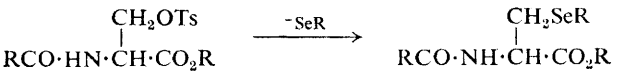
G. α-Amino-acids containing Sulphur or Selenium.—Alkylation of *N*-formyl-aminomalonate with the bromo-derivative (31) affords, after hydrogenolysis and hydrolysis, the novel benzothiophen isostere (32) of 5-hydroxytryptophan.[89] An extensive investigation into the synthesis of thialysine and its



(31) R = Bz (32)

sulphoxide and sulphone has been reported.[90] The antileukaemic activity of *S*-trityl-L-cysteine has led to the synthesis of a wide variety of these compounds by the condensation of the corresponding carbinol and cysteine in the presence of boron trifluoride;[91] most of these compounds are new but only a selection have been included in the list of newly synthesised amino-acids.

Optically active selenium-containing amino-acids can be prepared in good yield by nucleophilic displacement of a tosyl group by either a benzyl selenoate or selenide anion from a suitably protected *O*-tosylserine derivative (33).[93, 94] Compound (34) can be further modified to give a range of



(33) (34)

selenium-containing amino-acids. In this way L-selenocystine, L-seleno-lanthionine, L-selenomethionine, and L-selenoethionine have been prepared; it is claimed that the optical purities are greater than those previously obtained by alternative methods.

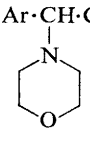
[85] R. K. Olsen, *J. Org. Chem.*, 1970, **35**, 1912.
[86] C. J. Abshire and R. Pineau, *Experientia*, 1970, **26**, 752.
[87] G. Gallina, F. Pentrini, and A. Romeo, *J. Org. Chem.*, 1970, **35**, 2425.
[88] A. J. Verbiscar and B. Witkop, *J. Org. Chem.*, 1970, **35**, 1924.
[89] E. Campaigne and A. Dinner, *J. Medicin. Chem.*, 1970 **13**, 1205.
[90] P. Hermann, K. Stalla, J. Schwimmer, I. Willhardt, and I. Kutschera, *J. prakt. Chem.*, 1970, **311**, 1018.
[91] K. Z. Cheng and C. C. Cheng, *J. Medicin. Chem.*, 1970, **13**, 414.
[92] C. Lee and G. S. Serif, *Biochemistry*, 1970, **9**, 2068.
[93] J. Roy, N. Gordon, I. L. Schwartz, and R. Walter, *J. Org. Chem.*, 1970, **35**, 510.
[94] C. S. Pande, J. Rudick, and R. Walter, *J. Org. Chem.*, 1970, **35**, 1440.

H. A List of α-Amino-acids which have been Synthesised for the First Time.—

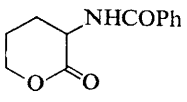
Compound	*Ref.*
N^GN^G-dimethyl-L-arginine	9
$N^GN'^G$-dimethyl-L-arginine	9
$N^\varepsilon N^\varepsilon N^\varepsilon$-trimethyl-δ-hydroxy-L-lysine	18
N^ε-(5-amino-5-carboxypentyl)-5-hydroxy-L-lysine	19
δε-dihydroxy-L-norleucine	20
DL-1-amino-3,3,5,5-tetramethylcyclohexane-1-carboxylic acid	50
DL-1-aminocyclodecane-1-carboxylic acid	50
DL-1-aminocyclo-octane-1-carboxylic acid	50
DL-*cis*-1-amino-1,3-dicarboxycyclohexane (cycloglutamic acid)	51
DL-*trans*-1-amino-1,3-dicarboxycyclohexane	51
DL-3,3,4,4,4,-pentafluoro-2-aminobutyric acid	52
DL-3,3,4,4,5,5,5-heptafluoro-2-aminovaleric acid	52
ω-fluoro-DL- and L-*allo*-isoleucine	53
D- and L-2,6-diamino-4-hexynoic acid	54
O-(α-D-glucopyranosyl)-L-serine	60
O-(1-oleyl-glycero-3-phosphoryl)-L-threonine	63
O-(1,2-dioleyl-glycero-3-phosphoryl)-L-threonine	63
DL-2-amino-5,6-dihydroxyindan-2-carboxylic acid	69
N-bis-(2-chloroethyl)-3,4-dihydroxy-L-phenylalanine	71
DL-3-(*m*-fluorophenyl)-2-methylalanine	72
DL-3-(*m*-bromophenyl)-2-methylalanine	72
DL-3-(*m*-iodophenyl)-2-methylalanine	72
DL-2-[(*m*-iodophenyl)methyl]glycine	72
DL-4-(*m*-iodophenyl)-2-methyl-2-aminobutyric acid	72
3,5,3′-tri-isopropyl-DL-thyronine	75
3,5-dimethyl-3′-isopropyl-DL-thyronine	76
3,5-di-isopropyl-DL-thyronine	77
3,5-di-isopropyl-4′-amino-DL-thyronine	77
3,5-di-isopropyl-3′-bromo-DL-thyronine	77
3,5-di-isopropyl-3′-methyl-DL-thyronine	77
3,5-di-s-butyl-DL-thyronine	77
3,5-di-s-butyl-4′-amino-DL-thyronine	77
3,5-di-s-butyl-3′-bromo-DL-thyronine	77
3,5-di-s-butyl-3′-iodo-DL-thyronine	77
N-phenyl-2-phenyl-DL-aspartic acid	78
N-methyl-2-phenyl-DL-aspartic acid	78
N-benzyl-2-phenyl-DL-aspartic acid	78
4-fluoro-DL-tryptophan	79
5-fluoro-DL-tryptophan	79
6-fluoro-DL-tryptophan	79
β-(5-hydroxy-6-iodo-2-pyridyl-1-oxide)-DL-alanine	81
β-(5-hydroxy-6-iodo-2-pyridyl)-DL-alanine	81
β-(benzimidazol-5-yl)-DL-alanine	82
β-(2-amino-6-hydroxypurin-9-yl)-DL-alanine	83
β-(2-amino-6-mercaptopurin-9-yl)-DL-alanine	83
N-dicolchicidyl-L-lysine	84
N-(2-pyridyl)-DL-aspartic acid	86
N-(3-methyl-2-pyridyl)-DL-aspartic acid	86
N-(4-methyl-2-pyridyl)-DL-aspartic acid	86
N-(6-methyl-2-pyridyl)-DL-aspartic acid	86
cis-p-methoxybenzylmercapto-L-proline	88
trans-p-methoxybenzylmercapto-L-proline	88

Compound	*Ref.*
β-(5-hydroxy-3-benzo[*b*]thienyl)-DL-alanine	89
N^{ε}-acetyl-L-thialysine	90
L-thialysine sulphoxide	90
L-thialysine sulphone	90
S-[(diphenyl-α-naphthyl)methyl]-L-cysteine	91
S-[(diphenyl-β-naphthyl)methyl]-L-cysteine	91
S-(9-methyl-9-fluorenyl)-L-cysteine	91
DL-2-amino-6-(methylthio)caproic acid	92

I. Labelled Amino-acids.—An interesting new method for labelling aryl aldehydes with deuterium or tritium in the formyl group, which also provides a novel route to labelled amino-acids, has been described.[95] Aryl aldehydes are converted into the crystalline morpholinoacetonitrile derivatives (35). The benzylic hydrogen is readily exchanged with deuterium oxide or tritiated water, and hydrolysis with aqueous acid affords the formyl-labelled aldehydes in high yield without loss of the label. The utility of these aldehydes for the synthesis of labelled amino-acids has



(35)



(36)

been demonstrated with synthesis of (±)-3,4-dihydroxy-[β-$^{2}H_{2}$]phenylalanine, but the wider synthetic potential of these aldehydes for specific labelling at prochiral centres is apparent. Tritioammonia has been employed to label 5-hydroxytryptophan, asparagine, and glutamic acid.[96] Reaction of the lactone (36) with radioactive cyanide in dimethylformamide, followed by reduction and hydrolysis, affords [6-^{14}C]-DL-lysine.[97] [2-^{14}C]- and [^{15}N]-DL-2-amino-6-(methylthio)caproic acid has been prepared from the correspondingly labelled aminomalonate,[92] and several DL-[^{35}S]thialysine derivatives have been synthesised by standard procedures.[90] An improved biochemically ^{14}C- and ^{35}S-labelled methionine has been reported[98] and incubation of *O*-acetylhomoserine in the presence of a radioactive alcohol gives the corresponding labelled *O*-alkylhomoserine.[65] It is claimed that the reaction of tyrosine with hydrochloric acid solutions of radioactive potassium iodide and potassium iodate gives good yields of [^{135}I]iodotyrosine,[99] but previous experience with this type of compound suggests that they are very unstable.

[95] D. J. Bennett, G. W. Kirby, and V. A. Moss, *J. Chem. Soc.* (*C*), 1970, 2049.
[96] K. Bloss, *J. Labelled Compounds*, 1969, **5**, 555.
[97] J. Pichat, J. Tostain, and C. Baret, *Bull. Soc. chim. France*, 1970, 1837.
[98] K. Samochocka and J. Kowalczyk, *Radiochem. Radioanalyt. Letters*, 1970, **4**, 131.
[99] M. El-Garhy, Y. M. Megahed, A. A. Kassem, and E. Abdullah, *Internat. J. Appl. Radiation Isotopes*, 1970, **21**, 240.

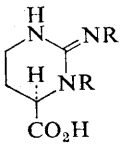
J. Resolution of α-Amino-acids.—It is well established that certain amino-acids can be resolved by paper adsorption chromatography, and further applications of this particular technique have been reported.[100] The possibility of employing alternative optically active adsorbents for resolution has recently attracted considerable attention. It is claimed that the ion-exchange resin, prepared by co-polymerisation of ethyl-*N*-acryloyl-L-polyglutamate and divinylbenzene, will resolve basic amino-acids such as lysine and ornithine with up to 90% efficiency.[101] Di- and tri-peptide derivatives of L-valine are also very effective stationary phases for the gas–liquid partition chromatographic separation of the enantiomers of racemic *N*-trifluoroacetyl-α-amino-esters.[102, 103] This technique is to some extent limited to neutral amino-acids and, as yet, is intended more for analytical application than for preparative use. (See also Section 5.)

Chemical and enzymic methods of resolution of synthetic racemates are still the most commonly employed, and detailed accounts of the use of ephedrine [104] and phenylethylamine [105] for the resolution of *N*-benzyloxycarbonyl-DL-amino-acids have been published.

3 Physical and Stereochemical Studies of Amino-acids

A. Determination of Absolute Configuration.—Lengthy chemical correlations for the establishment of absolute configuration are fortunately becoming less essential. The extent of application of o.r.d. and c.d. spectra and *X*-ray analytical techniques is increasing (see below), and chemical methods are only employed where direct correlation is relatively facile. Enzymic resolution of synthetic racemic amino-acids is, in effect, a method of determining the absolute configuration, and is constantly employed.

The o.r.d. spectra of L-amino-acids exhibit a positive Cotton effect at about 225 nm.[106] This has now been shown to hold for amino-acids containing a second chiral centre and has been used to assign the L-configuration at the α-centre of stendomycidine (5). Oxidation of (5) with *N*-bromosuccinimide gives (37; R = Me), the o.r.d. curve of which is

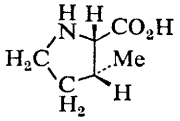


(37)

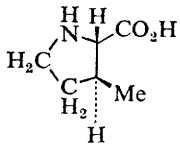
[100] R. Weichert, *Arkiv. Kemi*, 1970, **31**, 517.
[101] H. Suda, Y. Hosono, Y. Hosokawa, and T. Seto, *J. Chem. Soc. Japan, Ind. Chem. Sect.*, 1970, **73**, 1250.
[102] S. Nakaparkan, P. Birrell, E. Gil-Av, and J. Oro, *J. Chromatog. Sci.*, 1970, **8**, 177.
[103] B. Feibush and E. Gil-Av, *Tetrahedron*, 1970, **26**, 1361.
[104] K. Oki, K. Suzuki, S. Tuchida, T. Saito, and H. Kotake, *Bull. Chem. Soc. Japan*, 1970, **43**, 2554.
[105] E. Felder, D. Pitre, and S. Boveri, *Z. physiol. Chem.*, 1970, **351**, 943.
[106] J. P. Jennings, W. Klyne, and P. Scopes, *J. Chem. Soc.*, 1965, 294.

identical with that obtained from the hexahydropyrimidine derivative (37; R = H) derived from L-glutamic acid. The absolute chirality of stendomycidine is therefore as shown in (5)[14] and corresponds to that of capreomycidine and viomycidine at the α- and β-centres. The β-methyl-aspartic acid derived from the antibiotic amphomycin has been assigned the L-*threo*-configuration[107] and *N*-malonylmethionine, present in tobacco plants, is tentatively assigned the D configuration on the basis of the selective incorporation of D-methionine.[108]

Enzymic deacylation of the *N*-acetyl derivatives of ω-fluoro-DL-isoleucine and ω-fluoro-DL-*allo*-isoleucine affords the respective L-isomers which can be readily transformed into the corresponding *cis*-3-methyl-L-proline (38)



(38)

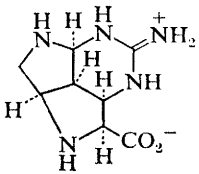


(39)

and *trans*-3-methyl-L-proline (39).[53] These are readily distinguished by their n.m.r. spectra, thus allowing a complete stereochemical assignment. A similar application of n.m.r. spectroscopy and enzymic resolution has allowed configurational assignments to be made to all the isomers of β-methyl-leucine and β-methylnorleucine.[109]

B. Crystal Structures of Amino-acids.—(See also Chapter 2, Part II, Section 2.) The crystal structures of L-isoleucine,[110] L-valine,[111] L-arginine hydrochloride,[112] DL-histidine hydrochloride,[113] 3,4-dihydroxy-L-phenyl-alanine[114] (L-dopa), and 5-hydroxy-DL-tryptophan,[115] as well as the derivatives *N*-chloroacetyl-DL-alanine[116] and *O*-phosphoryl-DL- and -L-serine,[117] have been published. A further *X*-ray analysis of the basic amino-acid viomycidine[12] and the details of the analysis of viocidic acid (40),[118] the other basic component isolated from the hydrolysate of the antibiotic viomycin, have been reported. The proposed structure and

[107] M. Bodanszky and G. G. Marconi, *J. Antibiotics*, 1970, **23**, 238.
[108] B. Ladešić, M. Pokorny, and D. Keglević, *Phytochemistry*, 1970, **9**, 2105.
[109] K. Okubi and Y. Izumi, *Bull. Chem. Soc. Japan*, 1970, **43**, 1541.
[110] B. Khawas, *Acta Cryst.*, 1970, **26B**, 1385.
[111] K. Torii and Y. Iitaka, *Acta Cryst.*, 1970, **26B**, 1317.
[112] J. Dow, L. H. Jensen, S. K. Mazumdar, R. Srinivasan, and G. N. Ramachandran, *Acta Cryst.*, 1970, **26B**, 1662.
[113] I. Bennett, A. G. H. Davidson, M. M. Harding, and I. Morelle, *Acta Cryst.*, 1970, **26B**, 1722.
[114] A. Mostad, T. Ottersen, and C. Romming, *Acta Chem. Scand.*, 1970, **24**, 1864.
[115] A. Wakahara, M. Kido, T. Fujiwara, and K. Tomita, *Tetrahedron Letters*, 1970, 3003.
[116] F. E. Cole, *Acta Cryst.*, 1970, **26B**, 622.
[117] M. Sundaralingam and E. Putkey, *Acta Cryst.*, 1970, **26B**, 782.
[118] P. Coggon, *J. Chem. Soc.* (*B*), 1970, 838.



(40)

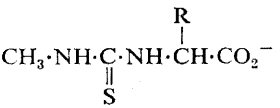
stereochemistry of the proline derivative obtained from diatom cell walls have been confirmed by *X*-ray crystallographic analysis.[119]

C. Optical Rotatory Dispersion (o.r.d.) and Circular Dichroism (c.d.).— (See also Chapter 2, Part III, Section 3B.) It has been confirmed that α-amino- and α-hydroxy-acids of the L-configuration exhibit, in addition to the strong positive c.d. maximum at 210—215 nm which is used for configurational assignments, a weak, negative, long-wavelength band at 235—240 nm.[120] The former is due to the n–π^* transition of the carboxy-group and the latter is attributed to the coupling of the non-bonding heteroatom with the chromophoric transition of the carbonyl,[121, 122] and not to intramolecular or intermolecular hydrogen-bonding as originally supposed. The fact that the hydrochlorides of L-α-amino-esters lack the weak, negative, long-wavelength band provides support for this proposal.[123]

The c.d. spectra of a number of α-methylamino-acids have been measured in acidic, alkaline, and neutral media and it appears probable that for the simple neutral amino-acids the shorter wavelength band can be used for configurational assignments.[124] Further work on the pH dependence of a number of aromatic amino-acids has led to the suggestion that the 220 nm band is due to the interaction of one of the transitions associated with the benzene ring and the π–π^* transition of the carbonyl, and not merely to a summation of the contributions of each chromophore.[125] *N-o*-Nitrobenzoyl derivatives of α-amino-acids exhibit a negative Cotton effect centred at 350 nm and it is claimed that they can be used for configurational assignments.[126]

In weakly alkaline solution α-amino-acids react with methyl isothiocyanate to give the derivatives (41). The c.d. data of the *N*-methylthiocarbamyl derivatives of all the common amino-acids with the L configuration exhibit positive Cotton effects centred around 260 nm.[127] These are

[119] I. L. Karle, *Acta Cryst.*, 1970, **26B**, 765.
[120] C. Toniolo, *J. Phys. Chem.*, 1970, **74**, 1390.
[121] J. C. Craig and W. E. Pereira, *Tetrahedron*, 1970, **26**, 3457.
[122] G. Barth, W. Voelter, E. Bunnenberg, and C. Djerassi, *Chem. Comm.*, 1969, 355.
[123] J. C. Craig and W. E. Pereira, *Tetrahedron Letters*, 1970, 1563.
[124] S. Yamada, K. Achiwa, S. Terashima, H. Mizuno, N. Takamara, and M. Legrand, *Chem. and Pharm. Bull. (Japan)*, 1969, **17**, 2608.
[125] N. Sakota, K. Okita, and Y. Matsui, *Bull. Chem. Soc. Japan*, 1970, **43**, 1138.
[126] U. Nagai and M. Kurumi, *Chem. and Pharm. Bull. (Japan)*, 1970, **18**, 831.
[127] C. Toniolo, *Tetrahedron*, 1970, 5479.



(41)

unaffected by the presence of a second chiral centre or a further dichroic centre. This method would appear to offer a very facile means of determining the configuration of amino-acids derived from microbial peptides. A similar, but more limited, approach employing isothiocyanate derivatives has also been reported.[128]

D. Nuclear Magnetic Resonance (n.m.r.) Spectra.—(See also Chapter 2, Part III, Section 4B.) The literature on the n.m.r. spectra of amino-acids has been critically reviewed[129] in an extremely valuable article which covers analysis and configurational determination, as well as the n.m.r. data of various derivatives including metal complexes. A detailed investigation of chemical shifts and spin–spin coupling constants of amino-acids, in relation to studies on peptide conformation, has been described[130] and rotational isomerism, as determined by variation of the vicinal couplings, has been shown to be dependent on solute–solvent and solute–solute interactions.[131]

As part of a more general programme related ultimately to the study of polypeptides, the ^{13}C n.m.r. spectra of ^{13}C-enriched amino-acids have been determined. The ^{13}C nuclei were noise decoupled from protons and the spread in the chemical shifts is sufficiently large for unambiguous assignment.[132] Interest in the application of ^{15}N n.m.r. data of amino-acids continues[133] and the vicinal coupling constants have been used to estimate conformer populations.[134] In addition to these more esoteric applications of n.m.r. spectroscopy, the technique is now generally employed for characterisation of new amino-acids and is particularly important in differentiating between diastereoisomers (see, *e.g.*, references 12, 14, 53, 66).

E. Mass Spectrometry.—The value of mass spectrometry as a structural tool appears to be increasing and it has been applied to at least two new naturally occurring amino-acids (see reference 10). Protection of amino-acids as either the *N*-trifluoroacetyl ester[135] or *N*-trimethylsilyl ester[136] continues to be the preferred method, but perhaps the most interesting

[128] B. Halpern, W. Patton, and P. Crabbé, *J. Chem. Soc.* (*B*), 1969, 1143.
[129] J. Rowe, M. Julian, J. Hinton, and K. L. Rowe, *Chem. Rev.*, 1970, **70**, 1.
[130] M. Nagai, A. Nishioka, and J. Yoshimura, *Bull. Chem. Soc. Japan*, 1970, **43**, 1323.
[131] J. R. Cavanaugh, *J. Amer. Chem. Soc.*, 1970, **92**, 1488.
[132] W. J. Horsley, H. Sternlicht, and J. Cohen, *J. Amer. Chem. Soc.*, 1970, **92**, 680.
[133] R. L. Lichter and J. D. Roberts, *Spectrochim. Acta*, 1970, **26**, 1813.
[134] R. L. Lichter and J. D. Roberts, *J. Org. Chem.*, 1970, **35**, 2806.
[135] M. S. Manhas, R. S. Hsieh, and A. K. Bose, *J. Chem. Soc.* (*C*), 1970, 116.
[136] K. Bergstrom, J. Gurtler, and R. Bloomstrand, *Analyt. Biochem.*, 1970, **34**, 74.

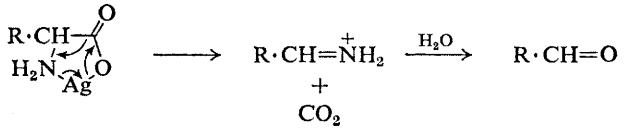
development in this area is the application of chemical ionisation mass spectrometry.[137] In this method ion formation is effected by protonation rather than loss of an electron. Consequently the resulting even-electron species is relatively stable and the quasimolecular ion at *m/e* ($M + 1$) is the most intense in the spectrum, except in the cases of glutamic acid or ornithine where cyclisation occurs, with loss of water and ammonia respectively. The technique could possibly be valuable for molecular weight determination since it avoids the inconvenient protection of amino-acids. A novel differentiation of *meso*- and racemic diaminopimelic acid, which involves an initial preferential thermal dehydration of the racemic form, has been reported[138] and it is suggested that the method can be extended to other diamino-acids. The mass spectra of trimethylsilyl derivatives of deuteriated[139] and ^{13}C-enriched[140] amino-acids have been determined, in relation to their possible application in biosynthetic studies.

F. Other Physical and Stereochemical Studies.—The internal rotation in crystalline glycine has been estimated from heat capacity data[141] and the i.r. spectrum of matrix-isolated glycine supports the suggestion that the molecules are not in the zwitterion form in this state.[142] The conformations of a number of amino-acids have been calculated theoretically using extended MO theory.[143] These studies come within the framework of a broader approach to the conformation of polypeptides. The dissociation constants in deuterium oxide of several amino-acids have been determined[144] and a detailed study on the racemisation of α-amino-acids and their derivatives in acetic acid has been reported.[145–147]

4 Chemical Studies of Amino-acids

A. Oxidation and Reduction.—Oxidation of α-amino-acids by silver(II) picolinate gives almost quantitative yields of the nor-aldehyde, while the action of silver(II) oxide affords, in most cases, the nor-acid.[148] The α-keto-acid is not an intermediate in these reactions and the mechanism outlined in Scheme 4 has been proposed. Oxidation of α-amino-esters, under the

[137] G. W. A. Milne, T. Axenrod, and H. M. Fales, *J. Amer. Chem. Soc.*, 1970, **92**, 5170.
[138] H. Falter, M. Madaiah, and R. A. Day, *Tetrahedron Letters*, 1970, 4463.
[139] W. J. A. van den Heuvel, J. L. Smith, I. Patter, and J. S. Cohen, *J. Chromatog.*, 1970, **50**, 405.
[140] W. J. A. van den Heuvel, J. L. Smith, and J. S. Cohen, *Biochim. Biophys. Acta*, 1970, **208**, 251.
[141] R. C. J. Li and N. S. Berman, *J. Phys. Chem.*, 1970, **74**, 1643.
[142] Y. Grence, J. C. Lasseques, and C. Carrigou-Laqrange, *J. Chem. Phys.*, 1970, **53**, 2980.
[143] J. M. George and L. B. Kier, *Experientia*, 1970, **26**, 952.
[144] I. N. Gordon and B. M. Lowe, *Chem. Comm.*, 1970, 803.
[145] M. Satō, T. Tatsuno, and H. Matsuo, *Chem. and Pharm. Bull. (Japan)*, 1970, **18**, 1794.
[146] H. Matsuo, Y. Kawazoe, M. Satō, M. Ohnishi, and T. Tatsuno, *Chem. and Pharm. Bull. (Japan)*, 1970, **18**, 1788.
[147] M. Satō, T. Tatsuno, and H. Matsuo, *J. Pharm. Soc. Japan*, 1970, **90**, 1160.
[148] T. G. Clarke, N. A. Hampson, J. B. Lee, J. R. Morley, and B. Scanlon, *J. Chem. Soc. (C)*, 1970, 815.



Scheme 4

same conditions, yields the corresponding α-keto-esters. Alternatively, electrochemical oxidation of α-amino-acids at a silver electrode in aqueous electrolytes gives the corresponding nor-nitriles and small amounts of the nor-aldehyde.[149] In this case it is argued that the intermediate imine is formed and further oxidation to the nitrile occurs before it is released from the electrode surface.

Kinetic studies on oxidation using cobalt(III) in aqueous perchloric acid solution have not provided any evidence for the formation of chelate complexes in the course of reaction, and a radical mechanism is proposed for the formation of the nor-aldehyde.[150] Kinetic data have also been reported on the uncatalysed oxidation of glycine by potassium persulphate.[151] *N*-Bromosuccinimide oxidises aspartic acid to the corresponding nor-aldehyde, which then undergoes α-bromination and decarboxylation to give bromoform in good yield.[152] *N*-Acyl-β-hydroxyamino-acids (42) are cleaved on oxidation with lead tetra-acetate to *N*-acyl-hydroxyglycine

(42) (43)

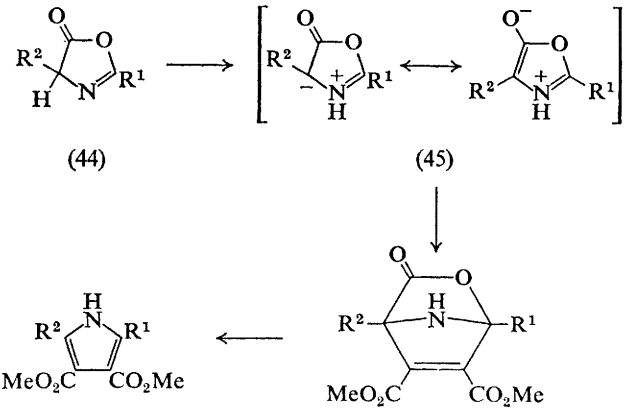
derivatives (43). The yields are high and this reaction offers a simple route to this interesting class of compounds.[153] A new general route has also been devised for the synthesis of *N*-phenylacetyl-β-alkoxyglycine derivatives by the reaction of 2-benzylidene pseudo-oxazolone with various alcohols.[154]

Reduction of α-amino-amides and esters with lithium aluminium hydride under a variety of conditions affords both primary alcohols and aldehydes in varying ratios,[155] whereas reduction of tertiary amides of amino-acids with sodium borohydride in pyridine gives the corresponding diamines in moderate yield.[156]

[149] N. A. Hampson, J. B. Lee, K. I. MacDonald, and M. J. Shaw, *J. Chem. Soc.* (*B*), 1970, 1766.
[150] R. A. Sheikh and W. A. Waters, *J. Chem. Soc.* (*B*), 1970, 988.
[151] K. Kumar and L. K. Saxena, *J. Indian. Chem. Soc.*, 1970, **47**, 435.
[152] W. L. Parker, C. Aklonis, and J. A. Last, *Experientia*, 1970, **26**, 242.
[153] W. Oettmeier, *Chem. Ber.*, 1970, **103**, 2314.
[154] G. Lucente, G. M. Lucente, F. Pantonella, and A. Romeo, *Ann. Chim.* (*Italy*), 1970, **60**, 259.
[155] M. P. Duhamel, L. Duhamel, and P. Siret, *Compt. rend.*, 1970, **270**, *C*, 1750.
[156] I. Saitō, Y. Kikugawa, and S. Yamada, *Chem. and Pharm. Bull.* (*Japan*), 1970, **18**, 1731.

B. General Reactions.—The literature on the protonation of amino-acids in strong acid solutions has been reviewed.[157] In superacids (*e.g.* fluorosulphonic acid–antimony pentafluoride) protonation of both the amino- and carboxy-functions occurs as well as at other basic sites in the molecule.[158] The protonated species have been studied by n.m.r. spectroscopy and are generally quite stable, unlike aliphatic carboxylic acids, which dehydrate to give the corresponding oxo-carbonium ions. The *N*-nitroso-derivatives of a number of common secondary amino-acids have been prepared in high yield under conditions approximating to those found in the mammalian stomach.[159] Evidence concerning the populations of the *syn*- and *anti*-conformers of these nitroso-amino-acids is derived from n.m.r. data. A previous report that the reaction of glucose with amino-acids affords nitrosamines has been refuted by a very extensive analytical investigation of the reaction products.[160]

α-Amino-acids continue to act as convenient starting materials for heterocyclic syntheses. The azlactones (44), prepared *in situ* from the corresponding *N*-acylamino-acid, react readily with dimethyl acetylenedicarboxylate to give substituted pyrrole-3,4-dicarboxylates.[161] The reaction proceeds as outlined in Scheme 5 *via* a 1,3-dipolar cycloaddition of the dimethyl acetylenedicarboxylate to the tautomeric oxazolium-5-oxide (45). An extensive investigation of this general type of reaction, employing a variety of dipolarophiles and azlactones, has been reported.[161–164] In

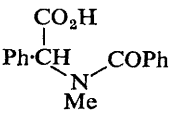


(44) (45)

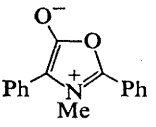
Scheme 5

[157] G. A. Olah, A. M. White, and D. H. O'Brien, *Chem. Rev.*, 1970, **70**, 561.
[158] G. A. Olah, D. L. Brydon, and R. H. Schlosberg, *J. Org. Chem.*, 1970, **35**, 317.
[159] W. Lijinsky, L. Keefer, and J. Loo, *Tetrahedron*, 1970, **26**, 5137.
[160] K. Heyns and H. Koch, *Tetrahedron Letters*, 1970, 741.
[161] H. O. Bayer, H. Gotthardt, and R. Huisgen, *Chem. Ber.*, 1970, **103**, 2356.
[162] R. Huisgen, H. Gotthardt, and H. O. Bayer, *Chem. Ber.*, 1970, **103**, 2368.
[163] H. O. Bayer, R. Huisgen, R. Knorr, and F. C. Schaefer, *Chem. Ber.*, 1970, **103**, 2581.
[164] R. Huisgen, H. Gotthardt, H. O. Bayer, and F. C. Schaefer, *Chem. Ber.*, 1970, **103**, 2611.

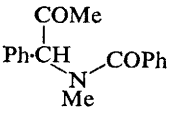
relation to these reactions, the mechanism of the Dakin–West reaction has also come under scrutiny.[165, 166] Reaction of (46) with acetic anhydride containing appreciable amounts of acetic acid yields (47) which then



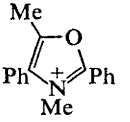
(46)



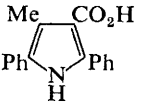
(47)



(48)



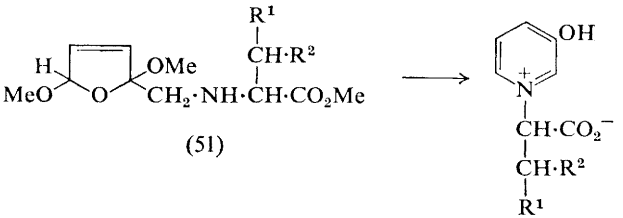
(49)



(50)

affords the normal Dakin–West product (48). However, if the reaction is conducted in acetic anhydride with a very low concentration of acetic acid a number of other products are formed, including the oxazolium salt (49) and the pyrrolecarboxylic acid (50). A reaction sequence involving acetylation of (47) followed by nucleophilic ring-opening is proposed to account for these products.[166]

N-Furfuryl-amino-acids have been synthesised by reductive alkylation of amino-acids using furfural. These compounds, on electrolytic oxidation in alcohol, yield compounds of the type (51), which readily rearrange in

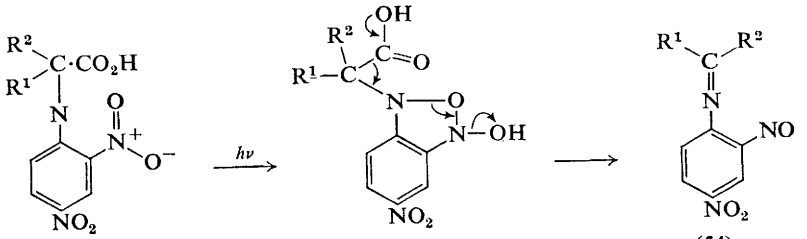


(51) (52)

acid solution to give the novel pyridinium derivative (52) of the starting amino-acid.[167] Treatment of *N*-benzoylserine methyl ester with phosgene gives the predicted oxazolidine in good yield.[168]

[165] N. I. Aronova, N. N. Makhova, and S. I. Zavialov, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1970, 1835.
[166] R. Knorr and R. Huisgen, *Chem. Ber.*, 1970, **103**, 2598.
[167] K. Unaheim and M. Gacek, *Acta Chem. Scand.*, 1969, 2488, 2475.
[168] T. Invi, S. Tanaka, and M. Takino, *Bull. Chem. Soc. Japan*, 1970, **43**, 1582.

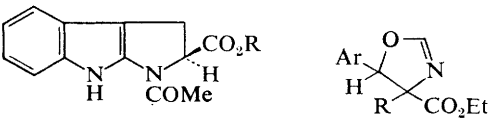
The use of orthoesters is still attracting considerable attention both for the acylation and esterification of amino-acids.[169] It is now possible to prepare the diazoketones from *N*-benzyloxycarbonyl or *N*-t-butoxycarbonyl amino-acids by the reaction of the mixed anhydride or carbodiimide adduct with diazomethane.[170] Previous attempts to prepare these compounds *via* the acid chloride had been unsuccessful. The suggested key step in the photolytic breakdown of *N*-2,4-dinitrophenyl derivatives of amino-acids is the formation of the intermediate (53) which then undergoes



(53) (54)

ring-opening and decarboxylation to (54). This intermediate can yield either 2-nitroso-4-nitroaniline and an aldehyde, or the benzimidazole-*N*-oxide, which are known products of photolysis, depending on the conditions of the reaction.[171]

C. Specific Reactions.—Oxidation of *N*-acetylated derivatives of tryptophan with t-butyl hypochlorite gives the acid-labile indole compounds (55), the cyclisation presumably occurring *via* the β-halogeno-indolenine.[172] Hydrogenolysis of aryl-2-oxazolines of the type (56) affords *N*-formylphenylalanine ethyl esters in good yield. Only a preliminary account of this work



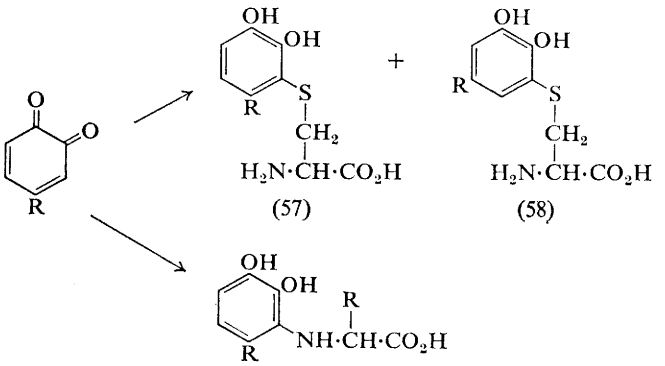
(55) (56) R = H
R = Me

has been published but it appears to offer an interesting new route to phenylalanines.[173] Interest in the reaction of L-cysteine with carbonyl compounds continues,[174] and several new thiazolidines have been prepared by the condensation of cysteine with monosaccharides.[175] Reduction of

[169] S. V. Rogozhin, Y. A. Davidovich, and V. V. Korshak, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1970, 727, 956, 2858.
[170] B. Penke, J. Czombos, L. Baláspiri, J. Petres, and K. Kovács, *Helv. Chim. Acta*, 1970, **53**, 1057.
[171] O. Meth-Cohn, *Tetrahedron Letters*, 1970, 1235.
[172] M. Ohno, T. F. Spande, and B. Witkop, *J. Amer. Chem. Soc.*, 1970, **92**, 343.
[173] U. Schöllkopf and D. Hoppe, *Angew. Chem.*, 1970, **82**, 459.
[174] N. Hellström, S. Almqvist, and M. Aamisepp, *J. Chem. Soc.* (*B*), 1969, 1103.
[175] R. Bognar, L. Somogyi, and Z. Gyorgydeak, *Annalen*, 1970, **738**, 68.

α-nitro-acrylates [176] and condensation of an amide with an α-keto-acid [177] are still the main sources of dehydro-amino-acids.

D. Non-enzymic Models of Biochemical Processes Involving Amino-acids.— The reactions of *o*-quinones with amino-acids continue to attract attention in relation to the biosynthesis of products resulting from quinone–amino-acid and quinone–protein interactions. It is now fairly well established that phaeomelanins are formed *in vivo* by the 1,6-addition of cysteine to dopaquinone produced by enzymic oxidation of tyrosine.[178, 179] The first step is stated to involve the formation of 2-*S*-cysteinyldopa (57) and 5-*S*-cysteinyldopa (58) in the ratio 95 : 5 (Scheme 6) which then undergo oxidative cyclisation. These proposals are supported by the observation that the model compound (59) is oxidised by oxygen in buffered solution to give a high yield of the dihydrobenzothiazine (60).[180] It is claimed that the *o*-quinone produced by oxidation of caffeic acid reacts with the α-amino-group of amino-acids [181] (Scheme 6). The *in vitro* oxidation of L-dopa with

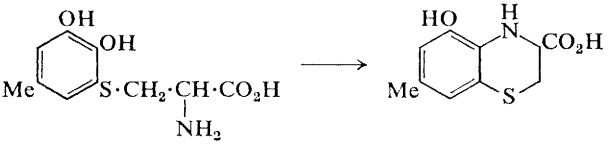


Scheme 6

p-benzoquinone under physiological conditions leads to the same products as are obtained from inorganic oxidants and phenol oxidase.[182]

The interaction of amino-acids with pyridoxal can be conveniently followed by n.m.r. spectroscopy and appears to be a valuable probe for examining structure and equilibria, as well as the reactivity of the azomethine bond towards various functional groups in polyfunctional amino-acids.[183] The formation of a thiazolidine derivative from cysteine is readily observed, as is the formation of a bis-Schiff-base by homocystine.

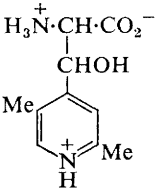
[176] C. Shin, M. Masaki, and M. Ohta, *Bull. Chem. Soc. Japan*, 1970, **43**, 3219.
[177] A. Kaneda and R. Sudo, *Bull. Chem. Soc. Japan*, 1970, **43**, 2159.
[178] G. Mizuraca, R. A. Nicolaus, G. Prota, and G. Ghiara, *Experientia*, 1969, **25**, 920.
[179] L. Minale, E. Fattorusso, S. De Stefano, and R. A. Nicolaus, *Gazzetta*, 1970, **100**, 461.
[180] G. Prota, S. Crescenzi, G. Misuraca, and G. A. Nicolaus, *Experientia*, 1970, **26**, 1058.
[181] C. H. Brieskorn and A. Mosande, *Tetrahedron Letters*, 1970, 109.
[182] A. Hikosaka and J. Kumanotani, *Bull. Chem. Soc. Japan*, 1970, **43**, 2620.
[183] E. H. Abbott and A. E. Martell, *J. Amer. Chem. Soc.*, 1970, **92**, 1745.



(59) (60)

Cystine is known to degrade when it reacts with pyridoxal phosphate, giving thiocysteine, ammonia, and pyruvic acid, but homocystine is stable under these conditions, giving, as observed by n.m.r. spectroscopy, a bis-Schiff-base.[184]

Transamination reactions conducted with 1-methyl-3-hydroxy-4-formylpyridinium chloride indicate that quaternisation of the ring nitrogen in pyridoxal models markedly increases the amount of transamination.[185] The decarboxylation of aminomalonic acid has been studied as a function of pH and in the presence of 5-deoxypyridoxal.[186] In acid solution, (61)

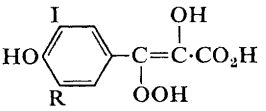


(61)

is formed by the interaction of aminomalonic acid with two molecules of 5-deoxypyridoxal.

The non-enzymic addition of ammonia to fumaric acid has been shown to be non-stereospecific, unlike the enzymic reaction.[187]

The hydroperoxide (62) is still regarded as the precursor of thyroxine in non-enzymic model systems, involving the oxidative coupling of di-iodotyrosine. Further work supports this claim and indicates that (63) is the precursor of the other thyroid hormone, 3,5,3′-tri-iodothyronine.[188]

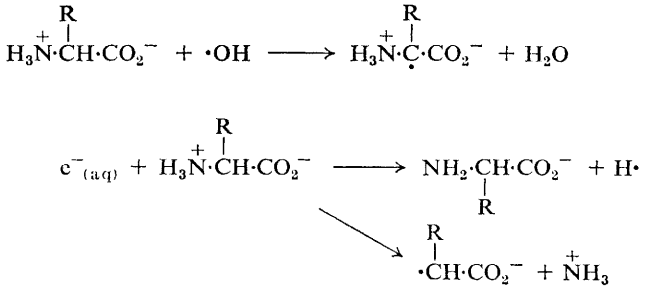


(62) R = H
(63) R = I

[184] A. Rinaldi and C. De Marco, *Arch. Biochem. Biophys.*, 1970, **138**, 697.
[185] J. R. Maley and T. C. Bruice, *Arch. Biochem. Biophys.*, 1970, **136**, 187.
[186] J. W. Thanassi, *Biochemistry*, 1970, **9**, 525.
[187] J. L. Bada and S. L. Miller, *J. Amer. Chem. Soc.*, 1970, **92**, 2774.
[188] H. J. Cahanmann and K. Funakoshi, *Biochemistry*, 1970, **9**, 90.

A direct conversion of di-iodotyrosine to thyroxine can be achieved by aerial oxidation in the presence of glyoxylic acid and copper acetate.[189] It is suggested that 4-hydroxy-3,5-di-iodophenylpyruvic acid is formed *in situ* by a transamination reaction with the glyoxylic acid and that this is oxidised to the intermediate (62). A kinetic investigation of the iodination of tyrosine suggests that a non-polar environment favours di-iodination relative to mono-iodination and it is concluded that this may be significant in relation to the internal environment of thyroglobin.[190]

E. Effects of Electromagnetic Radiation on Amino-acids.—There continues to be much activity in research on the effect of ionising radiation on amino-acids, and the observed radiolytic products and postulated reaction mechanisms have recently been reviewed.[191] Although a number of studies in the solid state have been reported,[192–195] the emphasis again remains on reactions in solution. For radiolysis of neutral amino-acids in aqueous solution, the products formed by their reaction with (·OH) radicals and hydrated electrons and the presumed mechanistic pathways are shown in Scheme 7. An e.s.r. study of the deamination step by hydrated electrons

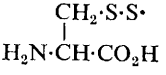


Scheme 7

has been reported.[196] Irradiation with high-energy electrons was carried out directly in the e.s.r. cavity and either ethyl formate or methanol was used as an (·OH) radical scavenger in order to eliminate its reacting with the amino-acids. General confirmation of the pathways outlined in Scheme 7 has been obtained from the transient optical absorption spectra of the intermediates in the dehydrogenation and deamination steps.[197]

[189] T. Shiba, M. Kajiwara, Y. Kato, K. Inoue, and T. Kaneko, *Arch. Biochem. Biophys.*, 1970, **140**, 90.
[190] W. E. Mayberry and T. J. Hockert, *J. Biol. Chem.*, 1970, **245**, 697.
[191] W. M. Garrison in 'Current Topics in Radiation Research', ed. M. Ebert and A. Howard, North Holland, Amsterdam, 1968, p. 43.
[192] V. G. Pasoyan, M. K. Pulatova, and L. P. Kavush, *Biofizika*, 1970, **15**, 12.
[193] V. G. Krivenko and M. K. Pulatova, *Biofizika*, 1969, **14**, 986.
[194] T. M. Chen, *Photochem. and Photobiol.*, 1970, **12**, 81.
[195] M. D. Sevilla, *J. Phys. Chem.*, 1970, **74**, 2096.
[196] P. Neta and R. W. Fessenden, *J. Phys. Chem.*, 1970, 2263.
[197] P. Neta, M. Simic, and E. Hayon, *J. Phys. Chem.*, 1970, **74**, 1214.

The radiolysis of aqueous solutions of cysteine and of cystine, and of phenylalanine and tyrosine, involves qualitatively different reactions from those outlined, due to the comparatively higher reactivity of the thiol group and the benzene ring towards hydroxyl radicals, and of sulphide and thiol group towards hydrated electrons. Several further studies on cysteine and cystine have been reported.[198–200] The radiolysis products of these amino-acids have been attributed to the breakdown or reaction of the sulphur radical of cysteine, but evidence for the participation of the radical (64) in the radiolysis of cystine has now been given.[201]

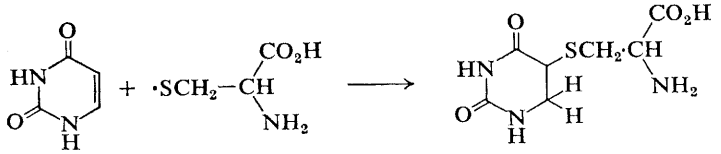


(64)

Chemically produced hydroxyl radicals continue to be used to simulate the effects of radiation. The generation of hydroxyl radicals (by the reaction of titanous chloride with hydrogen peroxide) in the presence of amino-acids leads to the same free radicals as are formed by radiolysis, and these can be conveniently studied by e.s.r. spectroscopy.[202, 203] The free radicals produced in the reaction between amino-acids and peptides with ninhydrin are very much dependent on the water and oxygen concentrations in the reaction mixture. Analysis of the e.s.r. spectra reveals hyperfine structure which, it is suggested, permits their possible use in identifying certain amino-acids and peptides.[204]

The mechanism for the photochemical addition of L-cysteine to uracil is believed to proceed through the triplet excited state of uracil, which can abstract an hydrogen atom from cysteine to form (65) (see Scheme 8).[205] The photo-oxidation of methionine to methionine oxide has been investigated using acetone[206] and proflavine[207] as sensitizers. In the case of proflavine it is suggested that the triplet state of the sensitizer is an intermediate and a mechanism is proposed in which methionine reacts with the first singlet state of oxygen, produced by energy transfer from the triplet sensitizer. A detailed scheme for the sensitized photo-oxidation of histidine

[198] D. Giles and D. W. Grant, *Chem. and Ind.*, 1970, 1437.
[199] L. I. Grossweiner and Y. Usui, *Photochem. and Photobiol.*, 1970, **11**, 53.
[200] H. Nishimura, S. Kawakishi, and M. Namiki, *Agric. and Biol. Chem. (Japan)*, 1970, **34**, 609.
[201] C. J. Dixon and D. W. Grant, *J. Phys. Chem.*, 1970, **74**, 941.
[202] V. B. Ilyasova, E. P. Busel, E. A. Burshtein, and O. A. Azizova, *Biofizika*, 1970, **15**, 265.
[203] P. Smith, W. M. Fox, D. J. McGinty, and R. D. Stevens, *Canad. J. Chem.*, 1970, **48**, 480.
[204] V. P. Yuferov, W. Froncisz, I. G. Kharitonenkov, and A. E. Kalmanson, *Biochem. Biophys. Acta*, 1970, **200**, 160.
[205] T. Jellinek and R. B. Johns, *Photochem. and Photobiol.*, 1970, **11**, 349.
[206] G. Gennari and G. Jori, *F.E.B.S. Letters*, 1970, **10**, 129.
[207] G. Jori and G. Cauzzo, *Photochem. and Photobiol.*, 1970, **12**, 231.



(65)

Scheme 8

and *N*-benzoylhistidine to aspartic acid and urea has been proposed and substantiated to a considerable degree by the isolation of a number of the proposed intermediates.[208]

5 Analytical Methods

The number of papers on the subject of amino-acids which are devoted to analytical methods continues to run at approximately a quarter of the total published. As pointed out in previous Reports, these mainly cover modifications of established techniques or the determination of specific amino-acids under certain conditions (*e.g.* amino-acids in biological fluids). The improvements and developments in amino-acid analysis in relation to structural studies of proteins and peptides are also covered elsewhere in this Report (Chap. 2, Part I, Section 2A). Therefore the majority of references will be presented under the appropriate heading without discussion, and only a few advances of general interest will be dealt with more fully.

A. Gas–Liquid Chromatography.—The application of g.l.c. for amino-acids protected either as the *N*-trifluoroacetyl n-butyl esters or the trimethylsilyl derivatives is now a well-established technique and is being used more often for routine analysis.[209–211] Per(methylsilyl) derivatives have been employed extensively [212–214] and the various methods of silylation have been reviewed.[215] The difficulties imposed by injection of the derivatives onto the column while in solution have to some extent been circumvented either by direct solid injection [216] or by preparation of the derivative on a solid support followed by direct introduction on to the column.[217] The separation of enantiomers of racemic *N*-trifluoroacetyl-α-amino-esters by g.l.c. has been described earlier, but several more detailed accounts have

[208] M. Momita, M. Irie, and T. Ukita, *Biochemistry*, 1969, **8**, 5149.
[209] D. J. Casagrande, *J. Chromatog.*, 1970, **49**, 537.
[210] R. W. Zumwalt, D. Roach, and C. W. Gehrke, *J. Chromatog.*, 1970, **53**, 195.
[211] W. Parr, C. Yang, J. Pleterski, and E. Bayer, *J. Chromatog.*, 1970, **50**, 510.
[212] C. W. Gehrke, H. Nakamoto, and R. Zumwalt, *J. Chromatog.*, 1969, **45**, 24.
[213] C. W. Gehrke and K. Leimer, *J. Chromatog.*, 1970, **53**, 195.
[214] C. W. Gehrke and K. Leimer, *J. Chromatog.*, 1970, **53**, 201.
[215] J. F. Klebe, *Accounts Chem. Res.*, 1970, **3**, 299.
[216] A. Darbre and A. Islam, *J. Chromatog.*, 1970, **49**, 293.
[217] B. Teuwissen and A. Darbre, *J. Chromatog.*, 1970, **49**, 298.

been reported.[218–220] A further study on g.l.c. separation of methylthiohydantoins has been presented.[221]

B. Ion-exchange Chromatography.—Hydrolysis of proteins with a mixture of oxalic and hydrochloric acids in a sealed tube is claimed to be more efficient than the classical procedure, and analysis can be conducted without initial neutralisation.[222] A large number of improvements in instrumentation and techniques,[223–226] as well as alternative buffer systems,[227–229] for automatic amino-acid analysers have been described. The ninhydrin colour-constant for *N*-methylamino-acids can be increased if the eluting buffer flow is slowed down, and the optical purity can be established by analysing the diastereoisomeric dipeptides obtained by coupling L-alanine *N*-carboxy-anhydride with the *N*-methylamino-acid.[229] The increasing application of computers for handling automatic analysis data is widely apparent and the whole area has been critically reviewed.[230] Alternative internal standards for amino-acid analysis have been reported.[231]

C. Thin-layer Chromatography.—An English translation of a well-known handbook on the t.l.c. of amino-acids is now available.[232] A new method of detection of amino-acids and amines on thin-layer chromatograms uses the fact that primary amines readily condense with 2,5-dimethoxytetrahydrofuran to yield *N*-substituted pyrroles which with *p*-dimethylaminobenzaldehyde in acid solution give an intense violet-red colour.[233]

A further modification of the Dragendorff reaction [234] for the visualisation of amino-acids and detailed investigations on techniques for rapid separation with various absorbents and detection with various reagents have been reported.[235, 236] Other papers on t.l.c. have described an improved technique for determining dansyl derivatives,[237] iodo-amino-acids,[238] and

[218] S. Nakaparkan, P. Birrell, E. Gil-Av, and J. Oro, *J. Chromatog. Sci.*, 1970, **8**, 177.
[219] W. A. Koenig, W. Parr, H. A. Lichenstein, E. Bayer, and J. Oro, *J. Chromatog. Sci.*, 1970, **8**, 193.
[220] W. Parr, C. Yang, J. Pleterski, and E. Bayer, *J. Chromatog.*, 1970, **50**, 510.
[221] K. Tuzimura, *Agric. and Biol. Chem. (Japan)*, 1969, **33**, 1566.
[222] N. Maravalhas, *J. Chromatog.*, 1970, **50**, 413.
[223] J. G. Heathcote, R. J. Washington, C. Haworth, and S. Bell, *J. Chromatog.*, 1970, **51**, 267.
[224] S. Jacobs, *Analyst*, 1970, **95**, 370.
[225] P. D. Meyer, L. D. Stegnik, and H. W. Shipton, *J. Chromatog.*, 1970, **48**, 538.
[226] A. Mondino, *J. Chromatog.*, 1970, **50**, 260.
[227] R. G. Redman, *J. Chromatog.*, 1970, **46**, 107.
[228] A. Vega and P. B. Nann, *Analyt. Biochem.*, 1969, **32**, 446.
[229] J. R. Coggins and N. L. Benoiton, *J. Chromatog.*, 1970, **52**, 251.
[230] B. Sheldrick, *Quart. Rev.*, 1970, **24**, 454.
[231] J. F. Cavins and M. Friedman, *Analyt. Biochem.*, 1970, **35**, 489.
[232] G. Pataki, 'Techniques of Thin-layer Chromatography in Amino-acid and Peptide Chemistry', Ann Arbor–Humphrey Sci. Publishers, Ann Arbor, Michigan, 1969.
[233] P. Haefelfinger, *J. Chromatog.*, 1970, **48**, 184.
[234] E. Tyihak and D. Vagujfalvi, *J. Chromatog.*, 1970, **49**, 343.
[235] J. G. Heathcote, R. J. Washington, C. Haworth, and S. Bell, *J. Chromatog.*, 1970, **51**, 267.
[236] D. T. N. Pillay and R. Mehdri, *J. Chromatog.*, 1970, **47**, 119.
[237] P. Nedkov and K. Gaucev, *Pharmazie*, 1970, **25**, 159.
[238] C. Wu and R. C. Ling, *Analyt. Biochem.*, 1970, **37**, 313.

peptide hydrazides.[239] A preliminary account of the separation of ^{14}C-labelled amino-acids and their detection by autoradiography on indium oxide plates indicates that this could be a valuable technique for the separation of very small amounts.[240]

D. Other Methods.—High-voltage electrophoresis continues to be a valuable analytical technique.[241] A new electrolytic procedure for the detection of L-amino-acids, employing an electrode covered with L-amino-acid oxidase, has been developed.[242] A preliminary account of the possible application of isotachophoresis of amino-acids in the presence of formaldehyde suggests that the technique may have some qualitative application.[243] Many other topics have been discussed, including fluorimetric determinations,[244] Sephadex chromatography,[245] and ion-exchange paper electrophoresis.[246]

E. Determination of Specific Amino-acids.—Papers on the determination of the following amino-acids have appeared: L-leucine,[247] L-phenylalanine,[248] L-glutamic acid,[249] L-tryptophan,[250] arginine,[251, 252] cystine,[253] hydroxylysine,[254] and lysine.[255]

[239] H. J. Goren and M. Fridkin, *J. Chromatog.*, 1970, **47**, 519.
[240] E. Cremer and E. Seidl, *Monatsh.*, 1970, **101**, 1614.
[241] P. J. Peterson and H. Fowden, *J. Chromatog.*, 1970, **48**, 575.
[242] G. G. Guilbault and E. Hrabankova, *Analyt. Chem.*, 1970, **42**, 1779.
[243] F. M. Everaerts and A. J. M. van der Pat, *J. Chromatog.*, 1970, **52**, 415.
[244] J. Borko and R. W. Frei, *Mikrochim. Acta*, 1969, 1144.
[245] A. D. Williams, D. E. Freeman, and W. H. Florsheim, *J. Chromatog.*, 1969, **45**, 391.
[246] E. Selegny, J. C. Fenyo, G. Broun, F. Matray, and C. De Bernady, *J. Chromatog.*, 1970, **47**, 552.
[247] S. Pande and M. L. Verma, *Microchem. J.*, 1970, **15**, 6.
[248] E. Jellum, V. A. Close, W. Patton, W. Pereira, and B. Halpern, *Analyt. Biochem.*, 1969, **31**, 227.
[249] A. Zamorani and G. Roda, *J. Chromatog.*, 1970, **47**, 261.
[250] F. Lingens and B. Sprössler, *Analytische Chemie*, 1970, **252**, 232.
[251] C. Saxena, *Microchem. J.*, 1970, **15**, 391.
[252] K. Bahadur and M. H. Verma, *Microchem. J.*, 1969, **14**, 547.
[253] P. Grisonni, *Biochim. appl.*, 1970, **15**, 227.
[254] E. Moczar and M. Moczar, *J. Chromatog.*, 1970, **54**, 277.
[255] O. A. M. Lewis and B. M. G. Shanley, *J. Agric. Food. Chem.*, 1970, **18**, 1178.

2
Structural Investigation of Peptides and Proteins

BY R. N. PERHAM, J. O. THOMAS, C. C. F. BLAKE, AND A. R. PEACOCKE

Part I: Primary Structures and Chemical Modification *by R. N. Perham and J. O. Thomas*

1 Introduction

One of the pleasures in compiling this Report of advances in the study of primary structure and chemical modification of proteins during the year 1970 is the occasional rediscovery of a delightful piece of work that, somehow, escaped one's proper attention at the time of its publication. The horror comes in having to attempt to condense the mass of relevant material logically and fairly into an article of reasonable (or even unreasonable) length. So much so that one is reduced to the state of Mrs. Craster's centipede [1] by the mere prospect. Whether the editors of the 'Atlas of Protein Sequence and Structure' suffer the same sorry affliction we do not know but they continue to publish their admirable annual compendium [2] and deserve every praise for their efforts.

For reasons of his own, none of them connected with protein chemistry, Bernard Levin has termed the Sixties 'The Pendulum Years.'[3] Not surprisingly, perhaps, protein chemistry has followed much the same pattern divined by Levin for society at large. Thus, in a decade which started in some excitement with proteins of 100 or so residues having their sequences determined, there are reports this year of complete sequences in excess of 500 residues and promises of others approaching 1000. All this has been achieved with gathering speed and has been nicely set in its historical context.[4]

[1] The Centipede was happy quite,
Until the Toad in fun
Said 'Pray which leg goes after which?'
And worked her mind to such a pitch
She lay distracted in the ditch
Considering how to run.
Attributed to Mrs. Edmund Craster.

[2] 'Atlas of Protein Sequence and Structure,' ed. M. O. Dayhoff, National Biomedical Research Foundation, Maryland, 1970, Vol. 5.

[3] 'The Pendulum Years' by Bernard Levin, Jonathan Cape, London, 1970.

[4] B. S. Hartley, 'British Biochemistry Past and Present', Biochemical Society Symposium No. 30, ed. T. W. Goodwin, Academic Press, London, 1970, p. 29.

2 Methods

The dansyl–Edman degradation, paper electrophoretic diagonal techniques, and reversible reactions of protein amino-groups have served the protein chemist so well in the analysis of primary structure that a review by Hartley of their introduction and current application is most timely.[5] A common use of amino-acid sequence information is the search for homology between proteins. Computer methods for accelerating the search are much in vogue.[6]

A. Amino-acid Analysis (see also Chap. 1 Section 5).—It has been reported [7] that acid hydrolysis of ribonuclease gives the same results if carried out at 145 °C for 4 h or 110 °C for 26 h. Although acid hydrolysis of peptides and proteins is not the rate-limiting step in sequence determination, there may well be little point in using the longer, conventional procedure. Alternatively, it has been suggested [8] that hydrolysis with crystallised oxalic acid containing some added hydrochloric acid is superior, since the hydrolysate may be applied direct to ion-exchange columns without the need to remove the acid required after hydrolysis with hydrochloric acid. A new procedure for alkaline hydrolysis of proteins in plastic tubes and subsequent analysis for tryptophan using potato-starch columns on a conventional analyser has been described.[9]

It has also been reported [10] that tryptophan analysis can be carried out on the intact protein by treatment with ninhydrin in a mixture of formic and hydrochloric acids for 10 min at 100 °C, followed by estimation of the absorbance at 390 nm. The use of sulphenyl halides for tryptophan and cysteine analysis in proteins has been given in further detail.[11] Direct spectral analysis for tryptophan is easy for proteins with no cysteine. However, since the reaction with cysteine residues can be reversed in 0.1N-NaOH, this provides a basis for the separate estimation of cysteine and tryptophan in proteins containing both sorts of residue. The reaction of a new water-soluble azomercurial, 4-(*p*-sulphophenylazo)-2-mercuriphenol, with the thiol groups of proteins has been described.[12] The reagent can be used for spectrophotometric titration of the thiol groups and also as a reporter group, once bound.

[5] B. S. Hartley, *Biochem. J.*, 1970, **119**, 805.
[6a] A. J. Gibbs and G. A. McIntyre, *European J. Biochem.*, 1970, **16**, 1.
[6b] S. B. Needleman and C. D. Wunsch, *J. Mol. Biol.*, 1970, **48**, 443.
[6c] W. M. Fitch, *J. Mol. Biol.*, 1970, **49**, 15; *ibid.*, 1970, **49**, 15.
[6d] J. E. Haber and D. E. Koshland, jun., *J. Mol. Biol.*, 1970, **50**, 617.
[7] D. Roach and C. W. Gehrke, *J. Chromatog.*, 1970, **52**, 393.
[8] N. Maravalhas, *J. Chromatog.*, 1970, **50**, 413.
[9] F. J. Oelshlegel, jun., J. R. Schroeder, and M. A. Stahmann, *Analyt. Biochem.*, 1970, **34**, 331.
[10] M. K. Gaitonde and T. Dovey, *Biochem. J.*, 1970, **117**, 907.
[11] E. Boccù, F. M. Veronese, A. Fontana, and C. A. Benassi, *European J. Biochem.*, 1970, **13**, 188.
[12] M. Jonian, J. Bloemmen, and R. Lontie, *Biochim. Biophys. Acta*, 1970, **214**, 468.

An unusual approach to the estimation of amino-acids has been described [13] involving an isotope dilution effect on the enzymatic synthesis of aminoacyl–tRNA. The method is capable of estimating some 5 nmol of amino-acid, comparable with the amount readily estimated by current high-sensitivity ion-exchange chromatography. Special methods for the determination of methionine,[14] proline,[15] and glycine [16] in crude protein hydrolysates have been reported and a refinement of the Sakaguchi reaction for the microdetermination of arginine has been described.[17]

Ion-exchange Chromatography. The use of lithium citrate buffers and the inclusion of n-propanol in the eluting buffers have received systematic study and increased resolution of asparagine, glutamine, and other rare ninhydrin-positive substances was observed.[18] A detailed analysis of the optimum resin column dimensions has been published [19] in which the use of the cheaper crushed (rather than spherical) resin is advocated and full analyses in 1.5 h reported. A warning has been sounded [20] that ninhydrin-positive peaks eluting between cysteic acid and aspartic acid may be artefacts derived from 'caramelisation' of carbohydrate during hydrolysis.

It has been suggested [21] that cysteine and cystine can be accurately estimated as *S*-β-(4-pyridylethyl)cysteine, which elutes just before arginine on the analyser. The derivative is produced by acid hydrolysis of the protein following reduction of the disulphide bridges and alkylation of the cysteine residues with 4-vinylpyridine. An alternative method [22] involves conversion of the cystine residues to *S*-sulphocysteine *after* the hydrolysis of the protein, with estimation of the *S*-sulphocysteine on the analyser. However, it is difficult to see either of these methods displacing the present analytical techniques for routine use.

Improved conditions have been reported [23] for the chromatography of the N^{ε}-carboxymethyl-lysines, increasing their separation from the common amino-acids. The use of the amino-acid analyser in the determination of *N*-methylamino-acids and the estimation of their optical purity has been described [24] with the warning that longer than normal reaction times with ninhydrin may be required for these modified amino-acids to generate their full colour. Further descriptions have been given of special conditions suitable for the analysis of collagen hydrolysates [25] and for the resolution

[13] M. T. F. de Fernández, S. M. de Sesé, and O. A. Scornik, *Biochemistry*, 1970, **9**, 2280.
[14] M. S. Vasantha, M. N. Moorjani, and K. S. Sreenivasan, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 568.
[15] I. Bergman and R. Loxley, *Analyt. Chem.*, 1970, **42**, 702.
[16] S. Suzuki, Y. Hachimori, and U. Yaoeda, *Analyt. Chem.*, 1970, **42**, 101.
[17] L. P. Alexeenko and V. N. Orekhovich, *Internat. J. Protein. Res.*, 1970, **2**, 241.
[18] G. E. Atkin and W. Ferdinand, *Analyt. Biochem.*, 1970, **38**, 309.
[19] A. Mondino, *J. Chromatog.*, 1970, **50**, 260.
[20] I. E. P. Taylor, *J. Chromatog.*, 1970, **50**, 331.
[21] M. Friedman, L. H. Krull, and J. F. Cavins, *J. Biol. Chem.*, 1970, **245**, 3868.
[22] A. S. Inglis and T.-Y. Liu, *J. Biol. Chem.*, 1970, **245**, 112.
[23] G. Ronca, R. Chiti, and A. Lucacchini, *J. Chromatog.*, 1970, **47**, 114.
[24] J. R. Coggins and N. L. Benoiton, *J. Chromatog.*, 1970, **52**, 251.
[25] R. C. Page, D. Jones, and R. Hansen, *Analyt. Biochem.*, 1970, **37**, 293.

of desmosine, isodesmosine, and lysinonorleucine in the analysis of elastin.[26] An improved buffer system that allows separation of aspartic acid and methionine sulphone on analysers of the Technicon type has also been reported.[27] The estimation of lysine by simple non-automatic ion-exchange chromatography particularly suitable for nutritional studies has been described.[28]

To limit the tedium in analyser calculations, the application of a desk-top computer to amino-acid analysis has been investigated[29a] and such uses of computers have been reviewed.[29b]

High-voltage Electrophoresis and Thin-layer Chromatography. The separation of nucleotides by high-voltage electrophoresis on ion-exchange paper is now widely practised,[30] and a new report[31] described the rapid separation of amino-acids by the same technique. Good resolution of the free amino-acids in purified plant extracts by t.l.c. in a single solvent system on cellulose-coated plates[32] is a typical example of the power of a basically simple method.

It has been suggested[33] that substances with primary amino-groups may be detected on thin-layer plates by reaction with 2,5-dimethoxytetrahydrofuran and *p*-dimethylaminobenzaldehyde. The reaction is akin to the Ehrlich estimation of tryptophan and the sensitivity is comparable with that of ninhydrin. More importantly, perhaps, the use of the Dragendorff reagent for the detection of *N*-methyl-lysines and *N*-methyl-histidines has been advocated.[34] Since the sensitivity falls off in the order N^{ε}-trimethyl-lysine > N^{ε}-dimethyl-lysine > N^{ε}-methyl-lysine > N^{α}-methyl-lysine > lysine, *i.e.* the reverse order to ninhydrin, this system has much potential value. The reaction works on paper and, with greater sensitivity, on thin-layer plates.

Another t.l.c. method for the microdetermination of amino-acids following their conversion to Dnp-derivatives has been described[35] but, since the sensitivity is similar to modern ion-exchange chromatography, its major merit would seem to be cheapness. A conceptually similar method is to separate the dansylated amino-acids and estimate them by their fluorescence. Details of such a technique have now been given[36] in

[26] S. Y. Yu, *Analyt. Biochem.*, 1970, **37**, 212.
[27] D. G. Redman, *J. Chromatog.*, 1970, **46**, 107.
[28] J. A. Bell and V. C. Mason, *J. Chromatog.*, 1970, **46**, 317.
[29a] A. P. Damoglou, *J. Chromatog.*, 1970, **47**, 257.
[29b] B. Sheldrick, *Quart. Rev.*, 1970, **24**, 454.
[30] F. Sanger and G. G. Brownlee, 'British Biochemistry Past and Present,' Biochemical Society Symposium No. 30, ed. T. W. Goodwin, Academic Press, London, 1970, p. 183.
[31] E. Selegny, J. C. Fenyo, G. Brown, F. Matray, and C. de Bernardy, *J. Chromatog.*, 1970, **47**, 552.
[32] D. T. N. Pillay and R. Melidi, *J. Chromatog.*, 1970, **47**, 119.
[33] P. Haefelfinger, *J. Chromatog.*, 1970, **48**, 184.
[34] E. Tyihák and D. Vagujfalvi, *J. Chromatog.*, 1970, **49**, 343.
[35] E. Heymann, *Z. physiol. Chem.*, 1970, **351**, 927.
[36] J. P. Zanetta, G. Vincendon, P. Mandel and G. Gombos, *J. Chromatog.*, 1970, **51**, 441.

which the Dns-amino-acids are separated by t.l.c. on silica gel and then eluted for assay. The necessity of rapid elution from the gels is stressed. Comparable t.l.c. can also be used to separate dansylated tryptic peptides on a micro scale.[36]

Gas–Liquid Chromatography. The separation of amino-acids by g.l.c. continues to receive attention. A further study of analysis of the *N*-trifluoroacetyl derivatives of their n-butyl esters, using simplified apparatus, has been described.[37] Analysis of amino-acids as their trifluoroacetylated phenylthiohydantoins has also been suggested,[38] the trifluoroacetylation increasing the volatility of the Pth-amino-acid.

Pertrimethylsilylated amino-acids are also popular derivatives for analysis by g.l.c. Some of the difficulties observed in analyses for glycine and lysine have been elucidated[39] and it has been shown that the trimethylsilyl derivatives of deuterium-containing amino-acids have smaller retention times on g.l.c. columns than do their ^{1}H-analogues.[40] G.l.c. analysis of ^{13}C-labelled amino-acids after pertrimethylsilylation has also been reported.[41]

Analysis of some 35 rare and common amino-acids (and two amino-sugars) as the corresponding *N*-trifluoroacetyl derivatives of the n-butyl esters has been recorded[42] and the same derivative of the methyl esters has been used to separate and estimate glutamic acid, 2-pyrrolidone-5-carboxylic acid, and 2-pyrrolidone.[43]

B. End-group Analysis and Sequential Degradation.—The enzyme pyrrolidonyl peptidase, which specifically removes the blocked *N*-terminal residue pyrrolidonecarboxylic acid from proteins, has been partially purified from a number of animal and plant tissues.[44]

The *N*-terminal residues of soluble proteins from prokaryotic and eukaryotic cells have been determined by the FDNB method, using tritiated reagent for increased sensitivity.[45] In some cases, Edman degradation and a second treatment with ^{3}H-FDNB allowed *N*-terminal sequences to be established. In agreement with well-known results of earlier workers, methionine was found to be the major *N*-terminal residue in prokaryotic proteins, with alanine and serine also favoured. With eukaryotes the results were not so clear, although alanine and serine were again found to be common at the *N*-terminus. Mass spectrometry has been used[46] to

[37] N. Lachovitzki and B. Björklund, *Analyt. Biochem.*, 1970, **38**, 446.
[38] G. Roda and A. Zamorani, *J. Chromatog.*, 1970, **46**, 315.
[39] K. Bergström, J. Gurtler, and R. Blomstrand, *Analyt. Biochem.*, 1970, **34**, 74.
[40] W. J. A. VandenHeuvel, J. L. Smith, I. Putter, and J. S. Cohen, *J. Chromatog.*, 1970, **50**, 405.
[41] W. J. A. VandenHeuvel and J. S. Cohen, *Biochim. Biophys. Acta*, 1970, **208**, 251.
[42] D. J. Casagrande, *J. Chromatog.*, 1970, **49**, 537.
[43] A. Zamorani and G. Roda, *J. Chromatog.*, 1970, **47**, 261.
[44] A. Szewczuk and J. Kwiatkowska, *European J. Biochem.*, 1970, **15**, 92.
[45] J. L. Brown, *Biochim. Biophys. Acta*, 1970, **221**, 480.
[46] M. H. Studier, L. P. Moore, R. Hayatsu, and S. Matsuoka, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 894.

identify Dnp-amino-acids separated by paper chromatography, even when the resolution on the paper was not quite perfect. Components in a mixture of Dnp-amino-acids were also readily identified. Paper chromatography and t.l.c. have been used [47] to identify the Dnp-derivatives of amino-acids from the peptidoglycans of bacterial cell walls. A mixture of cellulose and silica gel was reported to be preferable to silica gel alone for t.l.c. since it gave a less fluorescent background when detecting Dnp-amino-acids under u.v. light. A new, volatile solvent system was also recommended for paper chromatography of Dnp-amino-acids.

Amino-acid analysis by quantitation of the separated Dns-amino-acids has already been mentioned.[36] Another report of quantitative analysis of Dns-amino-acids has also been given,[48] which may enable the dansyl technique to be used as a reliable quantitative *N*-terminal method. In this connexion, it is worth noting that care is required in handling dansyl amino-acids and peptides since they are subject to photolysis in u.v. and 'visible' light.[49]

Sequential degradation in sequence analysis has recently been reviewed.[50] Without doubt, the *N*-terminal stepwise degradation due to Edman is by far the most commonly employed, in its original form and under many guises. Some *S*-alkyl derivatives of cysteine suitable for the Edman degradation have been described:[51] the *S*-methyl, *S*-ethyl, and *S*-vinyl derivatives are all very suitable for identification of the corresponding Pth-amino-acids by t.l.c. Recent reports suggest that the original phenyl isothiocyanate can be replaced with advantage. Thus, *p*-bromophenyl isothiocyanate has been used in sequence analysis of pig pancreatic trypsin inhibitor because the resulting thiohydantoins give good mass spectra.[52] In particular, the spectra are more easily interpreted since the natural bromine isotopes impart a characteristic doublet nature to the peaks. In other investigations [53, 54] methyl isothiocyanate has been reported to be the reagent of choice. It is more volatile than its phenyl counterpart and all the methyl thiohydantoins, with the exception of arginine, can be resolved by g.l.c.[53] Separation of the methyl thiohydantoins by g.l.c. after trimethylsilylation has also been described [54] and it is noted that most methyl thiohydantoins are sufficiently volatile for such analysis even without trimethylsilylation.

However, it is in the automated Edman degradation, carried out in the 'sequenator', that most change is evident this year; not a change in quality,

[47] W. D. Grant and A. J. Wicken, *J. Chromatog.*, 1970, **47**, 126.
[48] V. A. Spivak, V. M. Orlov, V. V. Shcherbukhin, and J. M. Varshavsky, *Analyt. Biochem.*, 1970, **35**, 227.
[49] L. D'Souza, K. Bhatt, M. Madaiah, and R. A. Day, *Arch. Biochem. Biophys.*, 1970, **141**, 690.
[50] G. R. Stark, *Adv. Protein Chem.*, 1970, **24**, 261.
[51] C. Rochat, H. Rochat, and P. Edman, *Analyt. Biochem.*, 1970, **37**, 259.
[52] H. Tschesche and E. Wachter, *European J. Biochem.*, 1970, **16**, 187.
[53] M. Waterfield and E. Haber, *Biochemistry*, 1970, **9**, 832.
[54] D. E. Vance and D. S. Feingold, *Analyt. Biochem.*, 1970, **36**, 30.

though that is evident too, but rather one in quantity. Many of the sequences described elsewhere in this Report have been established with the help of the sequenator and such machines, in good hands, would seem to be generally capable of granting the first 40 or so residues from the *N*-terminus of most proteins. This implies an average of approximately 98% efficiency at each step of the degradation and the usual starting requirement is for about 0.1—0.5 μmol of protein. In fact, the sequenator may truly be said to have come of age since its use has now led to the substantial revision of the amino-acid sequence of growth hormone established earlier by conventional means (see Section 4). Most of this work has been carried out on commercial machines, the cost of which puts them beyond the means of many individual laboratories. To repeat the suggestion in last year's Report, laboratories would do well to consider sharing an instrument and its associated costs, the same suggestion applying, for the same reasons, to mass spectrometry in amino-acid sequence determination (Section 2C). To bring down the cost there have been several attempts to build sequenators with little more than the average laboratory workshop facilities and one such machine has been described in detail.[55] The possible advantages in terms of volatility, *etc.* that might result from using methyl isothiocyanate in the sequenator have also been indicated.[53]

At the other end of the cost scale, the idea of the Edman degradation on a paper support has been resurrected.[56] The method described is capable of very high sensitivity, 0.1 nmol of Pth-amino-acid being sufficient for detection and identification on t.l.c. plates. However, the dansyl–Edman degradation continues to hold pride of place as the method of choice for determining the amino-acid sequence of peptides and proteins at high sensitivity. It, too, is cheap. The analysis can be speeded up by carrying out parallel degradations on a number of samples of the same peptide, stopping the degradation after 1 step, 2 steps, 3 steps, *etc.* on the various samples. *N*-Terminal analysis of the resultant progressively shortened peptides then gives the amino-acid sequence of the starting peptide.[57] Extraction of excess of phenyl isothiocyanate and of the side-products of the degradation is not carried out until the requisite number of steps has been performed on the given sample of peptide and the method is particularly suitable for peptides of about 5—6 residues.

Following on the astonishing identification of Dns-amino-acids in the picomole range by t.l.c. on polyamide layer sheets,[58] amino-acid sequence determination has been pushed down to the nanomole range for peptides.[5, 59] The *N*-terminal sequence of proteins can also be determined by a neat trick.[59] The protein is reacted with ^{14}C-maleic anhydride, digested with a proteolytic

[55] M. D. Waterfield, C. Corbett, and E. Haber, *Analyt. Biochem.*, 1970, **38**, 475.
[56] J. M. Boigne, N. Boigne, and J. Rosa, *J. Chromatog.*, 1970, **47**, 238.
[57] W. R. Gray and J. F. Smith, *Analyt. Biochem.*, 1970, **33**, 36.
[58] V. Neuhoff, F. von der Haar, E. Schlimme, and M. Weise, *Z. physiol. Chem.*, 1969, **350**, 121.
[59] C. J. Bruton and B. S. Hartley, *J. Mol. Biol.*, 1970, **52**, 165.

enzyme, and the digest applied to a column of Zeokarb-225 in the H^+ form. The *N*-terminal peptide now carries no free α-amino-group and, uniquely, is not retained by the column. The maleyl group is next removed at mildly acidic pH and the sequence of the *N*-terminal peptide determined by the dansyl–Edman procedure. Very little protein is required for this analysis. Using this method, the *N*-terminal sequence of the met-tRNA synthetase of *E. coli* was shown to be Ala-Gly-Gly-Thr-.[59]

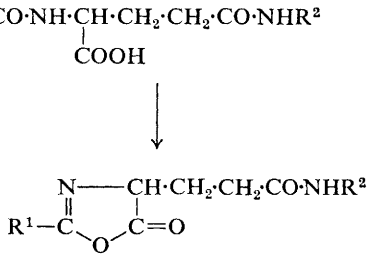
Other high-sensitivity variations of the Edman procedure have been suggested. Fluorescent thiohydantoins can be generated direct (using 1-naphthyl isothiocyanate as the degrading reagent),[60] and these can then be identified and detected with a sensitivity comparable with that for the Dns-amino-acid. Further reports of experience with this reagent will be of considerable interest (*cf.* fluorescein isothiocyanate[61]). An 'additive' method of carrying out the Edman degradation on small peptides has also been reported.[62] After each round of the degradation with methyl isothiocyanate, a sample of the Mth-amino-acids present is analysed by g.l.c. After an additional step of the degradation, the process is repeated and the additional Mth-amino-acid is detected in the g.l.c. analysis. A sequencing technique relying entirely on a radioactive label in conjunction with the Edman degradation has been applied to the determination of the *N*-terminal sequences of bacteriophage (Qβ, R17, f2) proteins synthesised in a cell-free system.[63] In parallel experiments the proteins are allowed to incorporate either methionine labelled with ^{35}S or other amino-acids labelled with ^{14}C. Since the bacterial proteins all have methionine as their *N*-terminal residue when newly synthesised, the ^{35}S-label can be used to locate the *N*-terminal peptide in proteolytic digests, which in turn locates the *N*-terminal peptide in the digests of the ^{14}C-labelled protein. The composition of the ^{14}C-labelled peptide is determined by adding ^{3}H-standards after hydrolysis, separating the amino-acids on an analyser column, and counting each fraction in a dual channel scintillation counter. The sequence of the peptide is determined by conventional subtractive Edman degradation, except that the composition of the peptide remaining at each step is determined by the counting technique.

Cyanomethyl dithiobenzoate has been suggested as a new reagent for the sequential degradation of peptides,[64] yielding a derivative that will cyclise under less drastic conditions than those required for the phenyl isothiocyanate method (*cf.* the work of Barrett[65]). A subtractive analysis of a pentapeptide was successfully achieved.

For *C*-terminal analysis by the tritiation method (reviewed in the 1969 Report), a warning has been given[66] that labelling of non-*C*-terminal

[60] Z. Deyl, *J. Chromatog.*, 1970, **48**, 231.
[61] H. Maeda and H. Kawauchi, *Biochem. Biophys. Res. Comm.*, 1968, **31**, 188.
[62] D. E. Vance and D. S. Feingold, *Nature*, 1971, **229**, 121.
[63] M. Osborn, K. Weber, and H. F. Lodish, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 748.
[64] A. Previero and J.-F. Pechere, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 549.
[65] G. C. Barrett, *Chem. Comm.*, 1967, 487.
[66] G. Ramponi, G. Cappugi, and P. Nassi, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 642.

glutamic acid can occur if the γ-glutamyl peptide bond is present in the protein, presumably by the formation of the oxazolone as shown in Scheme 1. It is suggested that the same considerations might apply to β-aspartyl



Scheme 1

linkages, should these be present naturally or generated by mistreatment of the protein.

C. Mass Spectrometry.—This has been a period of consolidation in what might almost be called the 'conventional' use of mass spectrometry for sequence analysis of peptides and proteins. Further efforts have been made to circumvent the problems posed by particular amino-acid residues when peptides are derivatised by *N*-acetylation and esterification, and made volatile by permethylation (for which the methyl iodide–sodium hydride–dimethyl sulphoxide method seems to have become the accepted procedure). Certain complications arising from permethylation have been reported: *N*-methylation of side-chain N atoms can occur, as can *C*-methylation of glycine, tryptophan, and histidine, with predictable complication of the spectra.[67] Permethylation cannot be used with the *N*-acetylacetonyl group (and the usefulness of this group is thereby limited) since the eneamine-ketone can undergo *O*- and *N*-alkylation and form a non-volatile quaternary ammonium salt.[68] To date, cysteine peptides have not been persuaded to give interpretable mass spectra if permethylation is employed; and formation of involatile sulphonium salts has meant that the sequence of methionine peptides cannot be obtained beyond the methionine residue. Desulphurisation is one way around the problem, and Lederer and his co-workers have recently improved on their original method with the development of much milder conditions for treatment with the Raney nickel catalyst.[69] This is more rapid than the still milder two-day treatment with this catalyst which has been recommended.[70] Cysteine peptides were found to desulphurise more readily than those containing methionine.[69]

[67] M. M. Shemyakin, Yu. A. Ovchinnikov, E. I. Vinogradova, A. A. Kiryushkin, M. Yu. Feigina, N. A. Aldanova, Yu. B. Alakhov, V. M. Lipkin, A. I. Miroshnikov, B. V. Rosinov, and S. A. Kazaryan, *F.E.B.S. Letters*, 1970, **7**, 8.

[68] V. Bacon, E. Jellum, W. Patton, W. Pereira, and B. Halpern, *Biochem. Biophys. Res. Comm.*, 1969, **37**, 878.

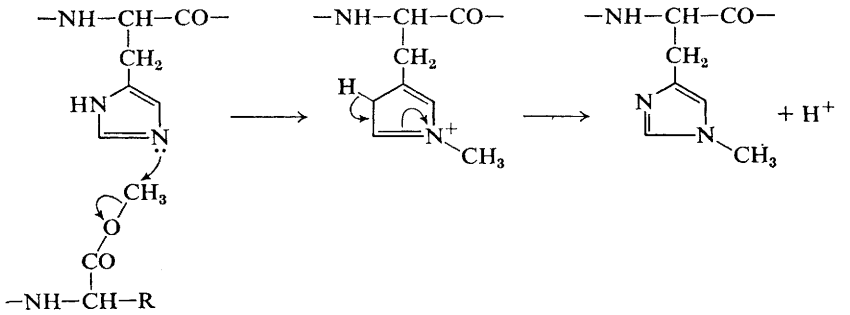
[69] R. Toubiana, J. E. G. Barnett, E. Sach, B. C. Das, and E. Lederer, *F.E.B.S. Letters*, 1970, **8**, 207.

An alternative catalyst (Ni_2B) which is easily prepared has also been suggested.[71] An attempt to take advantage of the formation of *S*-methylcysteinylsulphonium iodide during permethylation of cysteine peptides, by using an excess of methyl iodide in the hope of producing the corresponding dehydroalanine peptide and an interpretable spectrum, was unsuccessful. However, with a strictly limited amount of methyl iodide for the permethylation step, interpretable spectra of the *S*-methylcysteine peptides were obtained.[72] There was evidence that methionine peptides may also succumb to such treatment. An alternative to desulphurisation for a peptide containing methionine is its temporary conversion into methionine sulphoxide while permethylation is carried out, with subsequent reduction of the *N*-methylmethionine sulphoxide to the *N*-methylmethionine derivative for mass spectrometry.[73] This gives good spectra, with the usual advantages of lower temperatures and simpler fragmentations that accompany permethylation. A procedure for converting arginine-containing peptides into ornithine derivatives before *N*-acetylation, *etc.* has been described briefly: 1,1,2,2-tetramethoxypropane in methanolic hydrogen chloride gives the *N*-pyrimidinylornithine methyl ester derivative.[67]

Peaks at 14 mass units above the molecular ion ($M + 14$) in the spectra of *N*-acyl peptide esters containing histidine or tryptophan, which are sometimes stronger than the parent ion, have been attributed to intermolecular alkylation reactions in the mass spectrometer followed by proton loss.[74] Similar satellite peaks occur throughout the spectrum. Peptides containing two histidine residues had peaks at $M + 14$ and $M + 28$. Such intermolecular reactions are temperature dependent and are rarely observed below 200 °C. In methyl ester derivatives the reaction may proceed as follows with abstraction by histidine of a methylcarbonium ion from the methyl ester grouping (Scheme 2).[74] Compounds devoid of methyl groups *per se* sometimes exhibit the anomalous ($M + 14$) peaks, suggesting that methyl carbonium ions can arise from several points in a molecule. This is observed in the mass spectra of alkaloids.[75] Tryptophan can undergo similar reaction, as can *N*-pyrimidyl- or *N*-imidazolinedione-ornithine peptides,[74] but the phenomenon is easily recognised and need not complicate interpretation. It is comforting that a peak at ($M - 4$) [74] can be explained away in terms of dehydration, intermolecular methylation, and proton loss – provided, of course, that the peptide under investigation does contain a hydroxyamino acid and tryptophan or histidine.

[70] A. A. Kiryushkin, V. A. Gorlenko, B. V. Rosinov, Yu. A. Ovchinnikov, and M. M. Shemyakin, *Experientia*, 1969, **25**, 913.
[71] M. A. Paz, A. Bernath, E. Henson, O. O. Blumenfeld, and P. M. Gallop, *Analyt. Biochem.*, 1970, **36**, 527.
[72] M. L. Polan, W. J. McMurray, S. R. Lipsky, and S. Lande, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 1127.
[73] P. Roepstorff, K. Norris, S. Severinson, and K. Brunfeldt, *F.E.B.S. Letters*, 1970, **9**, 235.
[74] G. W. A. Milne, A. A. Kiryushkin, Yu. A. Alakhov, V. M. Lipkin, and Yu. A. Ovchinnikov, *Tetrahedron*, 1970, **26**, 299.
[75] D. W. Thomas and K. Biemann, *J. Amer. Chem. Soc.*, 1964, **87**, 5447.

Computer analysis of high resolution mass spectra in sequence analysis has not swept the board in the way one might have expected when the potential of the method was first investigated a few years ago. It is becoming clear that with improved methods of volatilisation (such as permethylation) and of derivatisation of 'awkward' amino-acid residues, low resolution



Scheme 2

spectra may be perfectly adequate to reveal the sequence of peptides simply by inspection. The hurdle now is the size of peptide that can be made sufficiently volatile to give a spectrum at acceptable source temperatures. With cleavage restricted very largely to the CO—NMe bond by permethylation, the use of a suitable acyl group gives a series of strong sequence ions based on the *N*-acyl derivative of the *N*-terminal amino-acid. Considerable information about the sequence of small peptides can be derived from as little as 10—20 nmol.[76] Impurities generated during the derivatisation of very small amounts of peptide can be a nuisance and it might be wise to do a preliminary purification on t.l.c. Alternatively, it might sometimes be sufficient to raise the temperature very gradually in the hope of volatilising the impurities before the sample.[76] If 20 nmol of material is required to determine the sequence of a small peptide then the method falls short of the dansyl–Edman method in terms of sensitivity, especially if the usually desirable preliminary step of determining the amino-acid composition is included, but the time and labour saved are enormous. It has been suggested that chemical or enzymic shortening of a peptide chain is useful in extending the range of mass spectrometry, *e.g.* Edman degradation is useful for removing 2 or 3 amino-acids from the *N*-terminus, particularly if these reduce the volatility, or need protection or modification before satisfactory spectra can be obtained:[67] this would, of course, presuppose some knowledge of the sequence. Similarly, carboxypeptidase B would be useful for removing *C*-terminal lysine and arginine from tryptic peptides, thereby decreasing the number of subsequent manipulations required to produce a satisfactory spectrum.

[76] J. Lenard and P. M. Gallop, *Analyt. Biochem.*, 1970, **34**, 286.

A rather elegant trick for making the *N*-terminal sequence of a protein directly available for mass spectrometry has been described.[77] *N*-Acetylation of the protein with acetic anhydride and its deuteriated counterpart is followed by digestion with a suitable protease (*e.g.* subtilisin) and the whole digest then permethylated. The *N*-terminal peptide will now be very hydrophobic and can be extracted into chloroform, while the others, as their quaternary ammonium salts, will be retained in the aqueous phase. The *N*-terminal sequence can thus be obtained very simply on as little as 1 mg of protein, which is not easily accomplished by standard methods (but see Section 2B). However, the method is unlikely to define more than a few residues from the *N*-terminus and is in no way a rival to the sequenator, which will handle almost equally small amounts of whole proteins and yield many more residues from the *N*-terminus. The sequenator and the mass spectrometer have not yet become part of the standard equipment of a protein chemistry laboratory, but as methods improve and faith in them increases there are signs that the attractions of these techniques will be thought to justify the cost of the machines – even if there are secret fears of redundancy in the minds of the more committed sequencers of proteins!

The possibility of being able to obtain sequence information for peptides and proteins present in mixtures is obviously attractive, and progress is being made in this direction. McLafferty and his co-workers[78] have been able to derive at least partial sequence information for mixtures of oligopeptides, acylated and esterified (although not permethylated), by computer analysis of the high-resolution spectrum. The method is not free from ambiguities, *e.g.* Gly-Leu and Ala-Val are identical in terms of elemental composition, and there are also the usual complicating non-amide cleavages. However, metastable ions can aid interpretation and partial vaporisation of the sample is useful in revealing those sets of peaks for which the ratio of abundances is constant and which therefore derive from a common peptide.[78] It should be borne in mind that partial vaporisation is just as useful for low-resolution spectra and, in fact, one would hope that this, in conjunction with a set of suitable derivatised peptides, would be sufficient to permit extensive amino-acid sequence analysis in peptide mixtures without recourse to computers. In any event a series of derivatives can often also clear up ambiguities in the computer analysis; a knowledge of the amino-acid composition is invaluable in minimising these. Full details have now appeared[79] of a method for sequence analysis of complex protein mixtures in which mass spectrometry is used in conjunction with isotope dilution to identify volatile thiohydantoins of *N*-terminal amino-acids generated using methyl isothiocyanate. A standard mixture of ^{15}N-enriched amino-acids or their thiohydantoins is added to the protein

[77] W. R. Gray and U. E. del Valle, *Biochemistry*, 1970, **9**, 2134.
[78] F. W. McLafferty, R. Venkataraghavan, and P. Irving, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 274.
[79] T. Fairwell, W. T. Barnes, F. F. Richards, and R. E. Lovins, *Biochemistry*, 1970, **9**, 2260.

mixture before the Edman degradation and the *N*-terminal amino-acid identified by virtue of its effect on the $^{15}N : {}^{14}N$ ratio estimated in the mass spectrometer. In this way 10 or so residues from the *N*-terminus can be identified for mixtures of peptides by quantitating the residues at each locus. There are certain limitations with very minor components, and where two components are present in roughly equal amounts. The method could prove to be extremely valuable for analysis of the heterogeneity of antibody sequences, and for investigating the homogeneity of subunits in multimeric proteins. We already have the first example of its use in this context in establishing that the dimeric enzyme enolase has the unique *N*-terminal sequence:[79]

Ala-Gly-Lys-Val-Gly-Asp-Thr-Gln-

This work was done on a single-focusing mass spectrometer and the point was made that the cost is roughly equivalent to that of an automatic amino-acid analyser. The chief limitation of the method seems to derive from the inevitable impurities generated during the stepwise degradation; there is every hope that this will cease to be a problem if the chemistry is delegated to a sequenator.[78]

While chemical ionisation mass spectrometry does not seem to offer any advantages over conventional electron impact mass spectrometry for the identification of amino-acids,[80] it does look promising for sequence analysis of peptides.[81] Acetylated, permethylated samples are used in the usual way. Ionisation through bombardment with carbonium ions derived from the reagent gas (*e.g.* methane) requires a much lower temperature than in the electron impact method, so that fragmentation is less extensive and is virtually limited to the peptide bonds. A major advantage lies in the more or less uniform intensity of the sequence-determining peaks, so that the sensitivity of the method is increased and interpretation greatly simplified. For instance, in the electron impact method 200 nmol of penta-alanine were apparently required to give discernible intensities of the highest sequence ions, since there is a 900-fold fall-off in peak intensity; in the chemical ionisation method 2 nmol were sufficient.[81] Thus, low-resolution spectra are perfectly adequate and easily interpreted. The method is very much in its infancy and only very simple peptides have so far been investigated; the lower temperatures needed to produce a spectrum, and the small amounts of material that will suffice, are clearly advantageous, particularly for larger, hence less volatile, peptides.

D. Cleavage of Protein Chains.—For a detailed account of procedures for selective cleavage of protein and peptide chains, particularly those effected chemically, the reader is directed to an excellent recent review.[82]

[80] G. W. A. Milne, T. Axenrod, and H. M. Fales, *J. Amer. Chem. Soc.*, 1970, **92**, 5170.
[81] W. R. Gray, L. H. Wojcik, and J. H. Futrell, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 1111.
[82] T. F. Spande, B. Witkop, Y. Degani, and A. Patchornik, *Adv. Protein Chem.*, 1970, **24**, 97.

Enzymic Cleavage. It is apparent that thermolysin is being increasingly used in overlapping initial points of cleavage by, usually, trypsin or chymotrypsin. For example, it was particularly useful in the determination of the complete primary structure of phospholipase A from pig pancreas [83] (see Section 5D). Attention is being directed to study of the cleavage reaction using synthetic oligopeptides [84] and dipeptides;[85] in the latter case use of acryloyl substrates enabled the cleavage to be monitored spectrophotometrically. Various other proteolytic enzymes are being investigated: peptidase β from brewer's yeast [86] resembled chymotrypsin in cleaving on the carboxyl side of tyrosine, phenylalanine, and leucine when tested on the perennial B-chain of oxidised insulin. The major protease of the venom of *Trimeresorus flavoviridis* has been purified (mol. wt. 24 000). It cleaves on the *N*-terminal side of leucyl residues in certain peptides.[87] An interesting enzyme isolated from the culture filtrates of *Streptomyces albus G* can hydrolyse the *C*-terminal D-Ala—D-Ala linkage of cell wall peptidoglycans,[88] and is analogous to the recently recognised enzyme of *E. coli.* It is proposed that the carboxypeptidase acts as a transpeptidase when integrated into the cell membrane. In penicillin-insensitive strains it is unaffected by penicillin thereby, perhaps, explaining the molecular basis of penicillin action.[89] Another instance of cleavage of D-Ala residues is rather more surprising: carboxypeptidase A has been reported to remove a *C*-terminal D-Ala from a tightly bound substrate [90] and it is concluded that the enzyme shows a higher degree of specificity towards the penultimate amino-acid residue than towards the *C*-terminal one. On a more mundane note, it is worth bearing in mind that the ability of chymotrypsin to cleave tyrosyl peptide bonds may be impaired if the ring [91] or the phenolic function [92] is substituted. Evidence is limited to ring iodine atoms and *O*-Dnp groups, respectively. In the case of model compounds containing iodotyrosine, at least, substitution enhanced the efficiency of binding [91] and lack of hydrolysis was apparently due to impairment of catalytic function.

Restriction of Enzymic Cleavage. Chemical modification of lysine residues in such a way as to restrict tryptic cleavage to arginyl peptide bonds has been discussed in some detail in the previous Reports. Modifications such as amidination [83] and succinylation [93] which are not reversible continue

[83] G. H. De Haas, A. J. Slotboom, P. P. M. Bonsen, L. L. M. VanDeenen, S. Maroux, A. Puigserver, and P. Desnuelle, *Biochem. Biophys. Acta*, 1970, **221**, 31.
[84] K. Morihara and H. Tsuzuki, *European J. Biochem.*, 1970, **15**, 374.
[85] J. Feder and J. M. Schuck, *Biochemistry*, 1970, **9**, 2784.
[86] F. Felix and D. Brissieux, *Biochim. Biophys. Acta*, 1970, **214**, 148.
[87] T. Takanashi and A. Ohsaka, *Biochim. Biophys. Acta*, 1970, **198**, 293.
[88] J. M. Ghuysen, M. Leyh-Bouille, R. Bonaly, M. Nieto, H. R. Perkins, K. H. Schleifer, and O. Kandler, *Biochemistry*, 1970, **9**, 2955.
[89] J. M. Ghuysen, M. Leyh-Bouille, R. Bonaly, M. Nieto, H. R. Perkins, K. H. Schleifer, and O. Kandler, *Biochemistry*, 1970, **9**, 2961, 2971.
[90] I. Schechter, *European J. Biochem.*, 1970, **14**, 516.
[91] C. J. Garratt and D. M. Harrison, *F.E.B.S. Letters*, 1970, **11**, 17.
[92] N. Kundu and S. Roy, *Nature*, 1970, **226**, 1171.
[93] K. D. Hapner and E. P. Wilcox, *Biochemistry*, 1970, **9**, 4470.

to be used occasionally for this purpose. The most recent of the reversible anhydrides is *exo-cis*-3,6-endoxo-Δ^4-tetrahydrophthalic anhydride (ETPA)[94] which is designed not to alkylate cysteine residues, a side reaction which accompanies the use of maleic and citraconic anhydrides.[95] While this is not a problem from the point of view of restricting tryptic cleavage in, say, *S*-carboxymethylated proteins, it becomes of prime importance if the reagents are used in studies of reversible dissociation of multimeric proteins (*e.g.* aldolase[95]). Further details of a method for non-reversible blocking of arginine residues have been reported.[96] The adducts of the biacetyl trimer with the arginine side-chain are stable at neutral pH. However, modification of lysine side-chains may also occur (*e.g.* 38% of the lysine residues in bovine serum albumin), although it is possible to cut this to 3% and still ensure 50—60% reaction at the arginine residues. It seems doubtful that this method will fill the gap that still exists for a suitable specific blocking group for arginine residues in proteins.

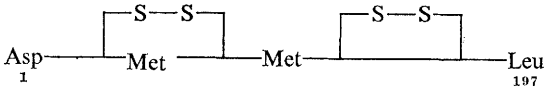
Chemical Cleavage. Spande, Witkop, Degani, and Patchornik[82] discuss many methods of specific cleavage of model peptides, but show that in the great majority of cases they are, sadly, not applicable to proteins because of numerous and various side reactions with different functional groups. Cleavage at methionyl peptide bonds with cyanogen bromide remains the only really satisfactory chemical method of cleaving protein chains. It is known that the lower homologue of methionine, *S*-methylcysteine, does not react with cyanogen bromide in the same way that methionine does, since the necessary β-lactone does not form (see review by Spande *et al.*[82]). However, alternative modes of reaction can result in cleavage on the amino side of the original cysteine. It is possible that this might become more widely used for cleavage of protein chains with the report by Heinriksen[97] of methylation of cysteine residues in proteins using methyl *p*-nitrobenzenesulphonate at pH 8.6. The interesting suggestion was also made that a 'thiol enzyme' might be converted to a 'hydroxyl enzyme' under suitable conditions (which are known for model compounds) by the sequence:

$$\text{E—SH} \longrightarrow \text{E—S—CH}_3 \xrightarrow{\text{CNBr}} \text{E—OH}$$

This would, of course, require a thiol enzyme that did not contain susceptible methionine residues. It is worth recalling that Nakagawa and Bender[98] used the same reagent to react with His-57 in *native* α-chymotrypsin (see p. 82); in reaction with denatured proteins, however, it is reported to be specific for cysteine residues.[97] Cyanogen bromide cleavage of peptide chains in the 'conventional' manner continues to be used so widely

[94] M. Riley and R. N. Perham, *Biochem. J.*, 1970, **118**, 733.
[95] I. Gibbons and R. N. Perham, *Biochem. J.*, 1970, **116**, 843.
[96] J. A. Yankeelov, jun., *Biochemistry*, 1970, **9**, 2433.
[97] R. L. Heinrikson, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 967.
[98] Y. Nakagawa and M. L. Bender, *Biochemistry*, 1970, **9**, 259.

that it would be an almost impossible task to list every instance. Of particular interest is the production of very large fragments (200—300 residues) from collagen, which can be ordered in the sequence by electron microscopy [99] (see Section 3). Cyanogen bromide fragments, often large, are sometimes insoluble but can usually be rendered soluble by citraconylation or maleylation. Such treatment can also facilitate fractionation of cyanogen bromide fragments by gel filtration, presumably by decreasing aggregation.[100] The presence in soybean trypsin inhibitor of two conveniently placed methionine residues (1) permits cleavage of the molecule into two fragments,



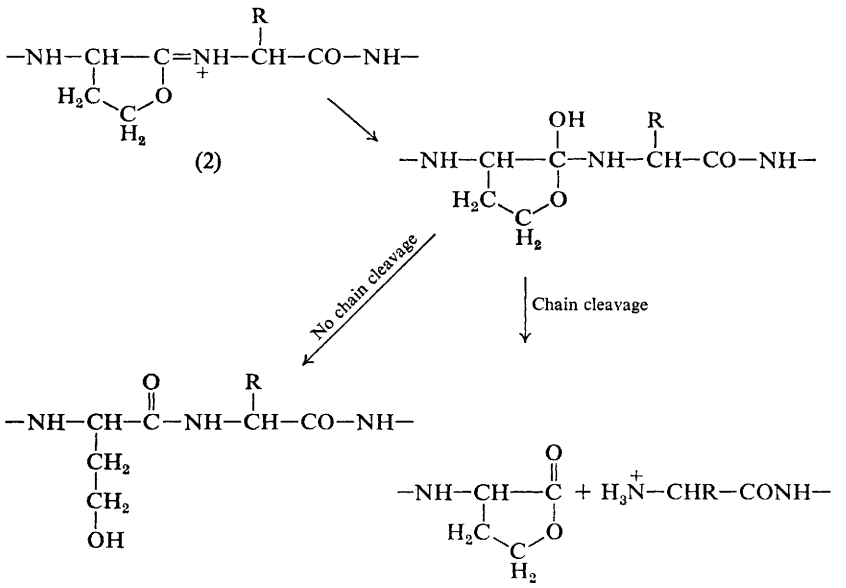
(1)

both devoid of anti-tryptic activity. These could be reconstituted to give a non-covalently bonded species with 80% of the anti-tryptic activity of the native inhibitor, despite rupture of one of the disulphide loops.[101]

Several cases have come to light recently where cyanogen bromide does not cleave protein chains completely at methionine residues. DeLange [102a] reported another case of limited cleavage where methionine is followed in the chain by either serine or threonine; in such cases this is attributed to iminolactone formation [102b] involving neighbouring group participation by the hydroxy side-chain. This mechanism cannot, however, explain incomplete cleavage at both Met-65 and Met-80 in cytochrome *c*,[103] since in neither case does a hydroxy-amino-acid follow in the sequence. An alternative mechanism analogous to that suggested for the hydrolysis of *N*-phenyliminolactone [104] is proposed [103] (see Scheme 3). The cyclic imidate, of the form (2), has been isolated and identified by Inglis and Edman [105] who followed the reaction by g.l.c. analysis of the methyl thiocyanate released. They observed that hydrolysis of the imidate almost always, but not inevitably, occurs, which would explain the replacement of methionone by homoserine without chain cleavage in certain instances.

A noteworthy example of successful cleavage of a protein using *N*-bromosuccinimide (NBS) has been reported by Rall and Cole for the lysine-rich histone from rabbit thymus.[106] This protein must have been designed with the NBS reaction in mind – it has only one tyrosine and no tryptophan, histidine, or sulphur-containing amino-acids – and it is heartening to find

[99] S. Igarashi, A. H. Kang, and J. Gross, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 697.
[100] I. Gibbons, P. J. Anderson, and R. N. Perham, *F.E.B.S. Letters*, 1970, **10**, 49.
[101] I. Kato and N. Tominaga, *F.E.B.S. Letters*, 1970, **10**, 313.
[102] (*a*) R. J. DeLange, *J. Biol. Chem.*, 1970, **245**, 907; (*b*) W. A. Schroeder, J. B. Shelton, and J. R. Shelton, *Arch. Biochem. Biophys.*, 1969, **130**, 551.
[103] G. Corradin and H. A. Harbury, *Biochim. Biophys. Acta*, 1970, **221**, 489.
[104] G. L. Schmir and B. A. Cunningham, *J. Amer. Chem. Soc.*, 1965, **87**, 5692; B. A. Cunningham and G. L. Schmir, *ibid.*, 1966, **88**, 551.
[105] A. S. Inglis and P. Edman, *Analyt. Biochem.*, 1970, **37**, 73.
[106] S. C. Rall and R. D. Cole, *J. Amer. Chem. Soc.*, 1970, **92**, 1800.



Scheme 3

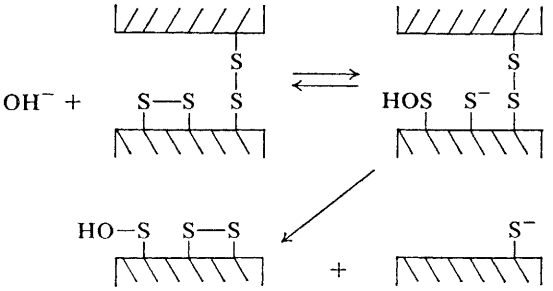
that cleavage at tyrosine was nearly quantitative. These workers tender the warning that applying samples to Sephadex columns in sucrose can result in poor resolution (although in fairness one should point out that histone fragments are hardly typical peptides).

Cleavage of peptide bonds where none was intended has also been reported. The lability of Asp-Pro bonds at low pH, where other aspartyl bonds are stable, was confirmed [107] during work on the primary structure of glutamate dehydrogenase. A mechanism involving intramolecular catalysis by a carboxylate ion on an *N*-protonated peptide bond was proposed, and the increased basicity of a proline nitrogen relative to that of a primary amine would facilitate this as well as providing a better leaving group. The further suggestion was made that the acidic conditions caused an $\alpha \rightarrow \beta$ aspartyl shift, and that the attacking species was the α-carboxylate. This shift presumably proceeds through the cyclic imine, which in the case of neighbouring proline would have a quaternary nitrogen, thereby facilitating isomerisation; the point was made that this effect would be important only if the rate of cleavage of the β-Asp-Pro bond was greater than that of the α-linkage, so that isomerisation was rate-limiting, and this is not yet clear. The cautionary tale [108] of hydrolysis of proteins during dialysis and ultrafiltration can be placed on a rather less rational footing. The obvious

[107] D. Piszkiewicz, M. Landon, and E. L. Smith, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 1173.
[108] G. Di Prisco and H. J. Strecker, *F.E.B.S. Letters*, 1970, **6**, 89.

controls seem to have been done but the explanation of the weird effects is not immediately apparent!

Though something of a digression from chain cleavage in proteins, it is worth reporting studies of disulphide bond cleavage in alkaline media. Spectrophotometric detection at 335 nm of persulphide (R—SSH) groups in alkaline insulin solutions [109] was taken as indication that, in this protein at least, β-elimination of cystinyl sulphur, rather than direct hydrolysis of the S—S bond, was mainly responsible for degradation of cystinyl sulphur. Further studies on the cleavage of dimers of bovine serum albumin at high pH suggest [110] the involvement of disulphide exchange reactions according to Scheme 4:



Scheme 4

This explains the inhibition by, for example, *p*-mercuribenzoate which would facilitate the simple hydrolytic reaction (Scheme 5).

$$RSSR + OH^- \rightleftharpoons RS^- + RSOH$$

$$2RSOH \longrightarrow RSO_2H + RSH$$

Scheme 5

Insolubilised Enzymes.[111a] The influence of the matrix on the properties of the matrix-bound enzyme has recently been discussed.[111b] Enzymes chemically attached to water-insoluble polymer supports are recognised as often being superior to the free enzyme in terms of stability, re-usability, and ease of manipulation (*e.g.* in a column). A wide range, particularly of insolubilised proteolytic enzymes, is now available commercially. The reader is referred to a commercial publication for selected references to earlier literature on the subject.[112] The variety of supports, and of procedures for coupling the enzymes to them, is apparent in reports which

[109] D. Cavallini, G. Federici, E. Barboni, and M. Marcucci, *F.E.B.S. Letters*, 1970, **10**, 125.

[110] L.-O. Andersson, *Biochim. Biophys. Acta*, 1970, **200**, 363.

[111] (*a*) I. H. Silman and E. Katchalski, *Ann. Rev. Biochem.*, 1966, **35**, 873; K. Mosbach, *Sci. Amer.*, 1971, **224**, 26. (*b*) E. M. Crook, *F.E.B.S. Symposia*, Vol. 19, ed. A. Sols and S. Grisolia, Academic Press, London and New York, 1970, p. 297.

[112] 'Insolubilized Enzymes,' Miles-Seravac, Miles-Yeda, December 1970, No. 3.

have appeared in the past year; again a representative selection only is included. Sepharose has been used for attachment of a protease from *Arthrobacter*[113] and for rennin.[114] An interesting case is reported of attachment of active monomers of aldolase to Sepharose[115] by reacting with the native enzyme in its tetrameric form, washing with urea, and then dialysing extensively. The preparation of insolubilised trypsin (G-200 Sephadex) with activity in 8M urea[116] could prove useful in sequence analysis.

The coupling of α-chymotrypsin to a variety of supports (cellulose, agarose, and dextran) using 2-amino-4,6-dichloro-*s*-triazine has been described[117] in a continued study using this method of coupling. Diazotisation of the support has frequently been used, *e.g.* for coupling of papain, trypsin, and subtilopeptidase to starch modified by introduction of amino-groups;[118] for attachment of glucose oxidase to polystyrene (so providing an automated method of analysis for glucose);[119] for coupling of deoxyribonuclease I to porous glass;[120] and for the linkage of trypsin to nylon tubes.[121] Catalase[122] and urease[123] have been coupled to polymers with glutaraldehyde; urease has also been immobilised on a polymer with alkylating properties.[124] The alternative approach for insolubilising enzymes, that of immobilising them in the interstices of a cross-linked gel,[125] has not been neglected: for example cholinesterase insolubilised in polyacrylamide has been prepared.[126] Increasing industrial application of insolubilised enzymes is virtually guaranteed; it is likely that the same will be true of laboratory use.

E. Fractionation Methods.—This section will be confined to fractionation of proteins and peptides; procedures for amino-acids have been dealt with in Section 2A.

Chromatography. A critical and comprehensive account of the principles underlying the use of gel chromatography as an analytical technique has recently appeared[127] and this aspect will not be dealt with further. The behaviour of molecules in both gel electrophoresis and gel filtration has also been the subject of theoretical analysis.[128]

[113] D. Gabel and B. v. Hofsten, *European J. Biochem.*, 1970, **15**, 410.
[114] T. Seki, T. A. Jenssen, Y. Levin, and E. G. Erdös, *Nature*, 1970, **225**, 864.
[115] W. W.-C. Chan, *Biochem. Biophys. Res. Comm.*, 1970, **4**, 1198.
[116] D. Gabel, P. Vretblad, R. Axén, and J. Porath, *Biochim. Biophys. Acta*, 1970, **214**, 561.
[117] G. Kay and M. D. Lilly, *Biochim. Biophys. Acta*, 1970, **198**, 276.
[118] L. Goldstein, M. Pecht, S. Blumberg, D. Atlas, and Y. Levin, *Biochemistry*, 1970, **9**, 2323.
[119] W. E. Hornby, H. Filippusson, and A. McDonald, *F.E.B.S. Letters*, 1970, **9**, 8.
[120] A. R. Neurath and H. H. Weetall, *F.E.B.S. Letters*, 1970, **8**, 253.
[121] W. E. Hornby and H. Filippusson, *Biochim. Biophys. Acta*, 1970, **220**, 343.
[122] A. Schejter and A. Bar-Eli, *Arch. Biochem. Biophys.*, 1970, **136**, 325.
[123] P. V. Sundaran and W. E. Hornby, *F.E.B.S. Letters*, 1970, **10**, 325.
[124] E. Brown, A. Racois, and H. Gueniffey, *Tetrahedron Letters*, 1970, 2139.
[125] P. Bernfeld and J. Wan, *Science*, 1963, **142**, 678.
[126] Y. Degani and T. Miron, *Biochim. Biophys. Acta*, 1970, **212**, 362.
[127] G. K. Ackers, *Adv. Protein Chem.*, 1970, **24**, 343.
[128] D. Rodbard and A. Chrambach, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **65**, 970.

It comes as no surprise to find that basic proteins such as myelin and the histones behave anomalously on gel filtration. The molecular weights of certain myelin proteins have been redetermined by gel filtration at acid pH,[129] using as markers proteins which have similar conformations under the conditions of study. The relationship between elution volume and molecular weight is dependent on the charge density of the molecule at low ionic strength so that suitable standards for determination of the molecular weight of myelin were lysine-rich histone and other proteins in this category, *e.g.* myoglobin, reduced ribonuclease, and cytochrome *c*. Anomalous behaviour of histones on gel filtration and gel electrophoresis at low pH was correlated [130] with the proline content and its effect on conformation. A system has been described [131] for determination of the molecular weights of peptides and proteins by chromatography on polyacrylamide gel (BioGel P-100) in phenol–acetic acid–water (1 : 1 : 1 v/v). The empirical relationship between log(mol. wt.) and elution volume holds for about 30 single-chain proteins studied, with the exception of cytochrome *c* which eluted early (see remarks above) and insulin which was retarded. The method has so far not been tested on dissociated proteins, but there seems to be no reason why the scope of the method should not be extended by inclusion of 2-mercaptoethanol. The high solvent power of this system [131] makes it an attractive one for certain problems, *e.g.* the study of membranes, *etc.* Reports that proteins interact with the polyacrylamide or dextran matrix during gel filtration occur quite frequently,[132, 133] and illustrate the importance of hydrophobic character and net charge in determining elution behaviour. A further report has appeared from Tanford's laboratory on gel chromatography in denaturing solvents.[134] In aqueous solutions of 6M guanidine hydrochloride or sodium dodecyl sulphate (SDS), reduced proteins assume conformations such that the hydrodynamic shape and molecular weight are related in a predictable way. Gel chromatography can therefore be used to measure both the effective hydrodynamic radius and the molecular weight of reduced polypeptide chains in these solvents. The most recent innovation from Porath's laboratory is the development of dipolar ion adsorbents for chromatography of proteins;[135] these are formed by coupling glycine, β-alanine, or ε-aminocaproic acid to Sephadex G-75. Molecular sieving was reported to be more prominent with glycine–Sephadex, while the weak cation effect dominated with the other two derivatives. Satisfactory results were obtained in separations of cobra venom toxins. Relative insensitivity to pH variations but a definite dependency on ionic strength means that good separations

[129] G. E. Deibler, R. E. Martenson, and M. W. Kies, *Biochim. Biophys. Acta*, 1970, **200**, 342.
[130] D. M. P. Phillips and M. Clark, *J. Chromatog.*, 1970, **46**, 320.
[131] A. Pusztai and W. B. Watt, *Biochim. Biophys. Acta*, 1970, **214**, 463.
[132] C. A. Bonilla, *J. Chromatog.*, 1970, **47**, 449.
[133] R. I. Gregerman, T. Weaver, jun., and M. A. Kowatch, *J. Chromatog.*, 1970, 499.
[134] W. W. Fish, J. A. Reynolds, and C. Tanford, *J. Biol. Chem.*, 1970, **245**, 5166.

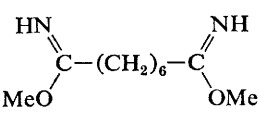
can be achieved at constant pH with shallow buffer concentrations. Dipolar ion adsorbents have several other advantages [135] over conventional ion exchangers, and it will be interesting to watch developments in this field. A highly unusual micro-method for the estimation of proteins by chromatography on PVC membranes (3—4 mm by 10—20 mm) has been described.[136] The method is reported to be applicable to a very wide range of molecular weights (12 000—400 000) and to operate at least in the pH range 3.7—9.06. The procedure takes 1—3 min, is very reproducible (± 3%), and requires 1—10 μg. Apart from the membranes (Sartorius, Göttingen, Germany) no special apparatus is necessary. Whether the method will find general acceptance remains to be seen.

Electrophoresis. Polyacrylamide gels are being used with success in a number of enterprising ways. Pinder and Gratzer have described an elegant separation of proteins by two-dimensional sedimentation and gel electrophoresis.[137] Sedimentation is carried out in a sucrose gradient (poured in the dark) containing acrylamide, bis-acrylamide, initiator, and riboflavin catalyst. After ultracentrifugation, the contents of the tube are photopolymerised, and longitudinal sections of the gel implanted into the sample groove of a flat-bed polyacrylamide gel. Electrophoresis is then carried out perpendicular to the direction of sedimentation. Serum proteins run in this system showed a marked sharpening of bands which was attributed to the role of the sedimentation gel as a spacer gel for electrophoresis. With the inclusion of markers of known molecular weight, the gel 'fingerprint' can be used to calculate the sedimentation coefficient of a protein, and could be especially useful for small amounts of protein present in complex mixtures. Pinder and Gratzer suggest that the scope of the technique could be further extended by performing the sedimentation on a micro scale, and by including SDS in the electrophoresis gel – thereby giving both the sedimentation coefficient of a protein and the molecular weight of its subunits! Polyacrylamide gel electrophoresis in two dimensions has been found to give good separation of ribosomal proteins,[138] and this method too should be of general value. Proteins rich in tyrosine or tryptophan, and glycoproteins, can be selectively stained in ordinary analytical polyacrylamide gels using modifications of standard procedures.[139] Microgels [140] (1 mm diameter) are recommended when staining for tyrosine, to avoid problems arising from gas evolution. A procedure for gel staining and autoradiography has also been documented.[141] There is an improved method for the separation of globin chains by starch gel electrophoresis in

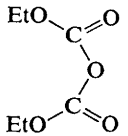
[135] J. Porath and L. Fryklund, *Nature*, 1970, **226**, 1169.
[136] T. I. Přistoupil, *J. Chromatog.*, 1970, **49**, 550; T. I. Přistoupil and M. Kramlová, *Experientia*, 1970, **26**, 1045.
[137] J. C. Pinder and W. B. Gratzer, *Biochim. Biophys. Acta*, 1970, **200**, 429.
[138] E. Kaltschmidt and H. G. Wittmann, *Analyt. Biochem.*, 1970, **36**, 401.
[139] K. Felgenhauer, A. Weis, and G. G. Glenner, *J. Chromatog.*, 1970, **46**, 116.
[140] K. Felgenhauer, *Biochim. Biophys. Acta*, 1967, **133**, 165.
[141] M. J. Daniels and D. G. Wild, *Analyt. Biochem.*, 1970, **35**, 544.

6M urea,[142] but urea is avoided altogether in an alternative procedure in which intact haemoglobin is made to move through a zone of *p*-chloromercuribenzenesulphonate by electrophoresis on a starch gel;[143] it is thereby dissociated into its constituent chains which then migrate separately. A potentially useful method has been developed for preparative electrophoresis in isoelectric buffers using Ampholine carrier electrolytes (LKB-Producter A.B., Stockholm, Sweden) in a sucrose density gradient.[144] The isoelectric buffer systems have much lower conductances than 'conventional' buffer systems (without loss of buffering capacity) and hence higher voltage gradients and shorter times of separation may be used.

The usefulness and reliability of gel electrophoresis in the presence of SDS as a method of determining subunit molecular weight is now beyond question. Instances in which it has been used to resolve conflicting values, and hence to establish quaternary structure, are legion; many of these are included in Table 2 (Section 5G). An expanded account of the technique has now appeared.[145] Reynolds and Tanford[146] have recently shown how the method works: a wide variety of proteins bind identical amounts of SDS on a gram/gram basis at concentrations of SDS monomer (the micelle does not bind) greater than 0.5 mmol l^{-1}. Binding is independent of ionic strength and is principally hydrophobic in nature. It is not surprising, therefore, that electrophoretic mobility in SDS gels is dependent only on molecular size. These studies of the binding of SDS to proteins could also have some relevance to membrane models.[146] Davies and Stark[147] have reported the use of the technique to study the effect of the bifunctional imidoester, dimethyl suberimidate (3), on multimeric proteins. Intra-oligomeric cross-linking leads to a predominant set of bands corresponding



(3)



(4)

to multiples of the subunit molecular weight, and if the subunits are identical the number of bands is equal to the number of subunits in the original molecule. Along the same lines, diethyl pyrocarbonate (4) reacts with proteins to give a series of bands on SDS-gels at integral values of the monomer molecular weight.[148] In addition to providing a useful set of

[142] M. D. Garrick, R. H. Balzer, jun., and J. P. Charlton, *Analyt. Biochem.*, 1970, **34**, 312.
[143] A. R. Schwantes and M. L. B. Schwantes, *Experientia*, 1970, **26**, 928.
[144] J. R. Fullerton and A. J. Kenny, *Biochem. J.*, 1970, **116**, 147.
[145] J. V. Maizel, jun., 'Fundamental Techniques in Virology,' ed. C. Hobel and N. P. Salzman, Academic Press, New York and London, 1969, p. 334.
[146] J. A. Reynolds and C. Tanford, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 1002.
[147] G. E. Davies and G. R. Stark, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 651.
[148] B. Wolf, P. M. Lausarot, J. A. Lesnaw, and M. E. Reichmann, *Biochim. Biophys. Acta*, 1970, **200**, 180.

'polymeric' markers for use as standards in measurements of molecular weight, it is suggested that an unknown molecular weight can be determined with increased accuracy by inspection of the multiple bands generated if the protein is first treated with diethyl pyrocarbonate. A range of gel concentrations was studied and some general deviations from linearity in the relation between log (mol. wt.) and mobility were reported.[148] In an interesting study, SDS-gel electrophoresis has been used to investigate the size distribution of polypeptide chains in eukaryotic and prokaryotic cells.[149] It was found that only 5% of the proteins in HeLa cells, and only 3% of those in *E. coli* cells, had molecular weights greater than 80 000. The average size for *E. coli* proteins was calculated as 24 000; for HeLa cells 31 700. This is a remarkable illustration of the amount of information that such a simple technique can yield, and it is difficult to see how else it could easily be obtained.

Substitution of *NN'*-diallyltartardiamide (DATD) for bisacrylamide as cross-linker gives polyacrylamide gels that can be solubilised in 2% periodic acid.[150] This does not quench and the method is useful for scintillation counting of gels on which radioactive proteins have been run and which are then sliced laterally. It is also possible to recover the proteins from the gel although one would need to be convinced that there were no deleterious effects of periodate. Inclusion of DATD does not interfere with SDS, and the linear relationship between log(mol. wt.) and mobility still holds, at least on 7.5% gels.[151] It has been suggested that protein molecular weight can be determined by investigating the mobility in gels of different porosity,[152, 153] but it seems unlikely that this will rival the far more convenient method using SDS-gels.

Peptide Detection and Identification. A number of reports have been concerned with peptide detection and identification. Peptide mapping can be carried out on the column of an automatic amino-acid analyser;[154] this was accomplished for oxidised ribonuclease on 0.15 mg of material, with preparative separation of pure peptides for sequence studies. It is possible to predict the composition of nucleotides from their mobility in two-dimensional electrophoresis.[30] Similarly, a comprehensive account has now been given of the prediction of the composition of dipeptides from their mobility on paper and column chromatography.[155] Since dipeptidases (see ref. 155) which work sequentially from the *N*-terminus of a protein are known, it is not inconceivable that such separations could form the basis of a new method of protein sequencing.

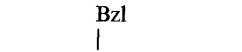
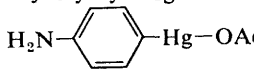
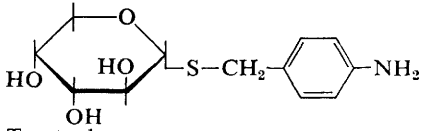
[149] E. D. Kiehn and J. J. Holland, *Nature*, 1970, **226**, 544.
[150] H. S. Anker, *F.E.B.S. Letters*, 1970, **7**, 293.
[151] A. R. Hunter and J. O. Thomas, unpublished work.
[152] C. R. Parish and J. J. Marchalonis, *Analyt. Biochem.*, 1970, **34**, 436.
[153] D. P. Blatter and F. J. Reithel, *J. Chromatog.*, 1970, **46**, 286.
[154] D. J. Bennett and E. H. Creaser, *Analyt. Biochem.*, 1970, **37**, 191.
[155] P. X. Callahan, J. A. Shepard, T. J. Reilly, J. K. McDonald, and S. Ellis, *Analyt. Biochem.*, 1970, **38**, 330.

Trinitrobenzensulphonic acid (TNBS) has been suggested as a reagent for detection of peptides on paper chromatograms [156] and the peptides can then be eluted, hydrolysed, and analysed. 2,2′-Dithiobis-(5-nitrobenzoic acid) (DTNB) has been used as a spray reagent for the detection of thiol compounds on chromatograms [157] [*cf.* 2,2′-dithiobis-(5-nitropyridine)] [158] and can also be used to detect disulphides (after reduction with borohydride) and thioesters (after hydrolysis on the paper). It is particularly useful for locating the disulphide-bridged peptides in enzymic digests.[159] DTNB has also been incorporated into a procedure for automatic detection of cystine peptides in column effluents, in the presence of pyridine–acetate buffers.[160] The method is based on the reduction of disulphides with dithioerythritol and detection of the resulting thiol groups with DTNB in the presence of arsenite. The most recent addition to the battery of 'diagonal' techniques (see Vol. 1, p. 61) provides another method for isolation of the *C*-terminal fragment of peptides and proteins.[161] The polypeptide is converted to the methylamide and digested enzymically. The *C*-terminal peptide, present as its methylamide, does not alter in mobility during the two steps of two-dimensional electrophoresis, first at pH 6.5 and then at pH 1.8, and will thus lie on the diagonal. This separates it from all the other peptides which will lie off the diagonal in the direction of the negative electrode. The method cannot be used for histidine-containing *C*-terminal peptides since these will also have different mobilities at the different pH values. Side-chain carboxy-groups presumably are converted into methylamides at the first stage and do not interfere.

Affinity Chromatography. Purification of a protein by affinity chromatography is based on specific and reversible interaction of the protein with a ligand. For an enzyme the ligand can be an inhibitor (an active-site-directed small molecule, or a protein inhibitor) or an activator; for an antibody it would be a hapten or an antigen; and for an antigen the antibody would serve as ligand. The ligand is covalently attached to an insoluble matrix, usually agarose (Sepharose, Pharmacia), and the protein is purified by selective adsorption on to a column of the insolubilised ligand. This powerful and elegant method is becoming so frequently used that it is not inappropriate to include here some of the uses to which it has been put since the brief coverage given in last year's Report. A remarkable array of adsorbents for affinity chromatography is already available commercially; one catalogue contains a useful bibliography relating to those described.[112] A list is given here of most of the uses to which affinity

[156] C. Wilson, R. F. Kibler, and R. Shapira, *Analyt. Biochem.*, 1970, **35**, 371.
[157] C. B. Glaser, H. Maeda, and J. Meienhofer, *J. Chromatog.*, 1970, **50**, 151.
[158] D. R. Grassetti and J. F. Murray, jun., *J. Chromatog.*, 1969, **41**, 121.
[159] H. Maeda, C. B. Glaser, and J. Meienhofer, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 1211.
[160] K. A. Walsh, R. M. McDonald, and R. A. Bradshaw, *Analyt. Biochem.*, 1970, **35**, 193.
[161] A. Furka, F. Sebestyén, and T. Karácsonyi, *F.E.B.S. Letters*, 1970, **6**, 34.

Table 1 *Some recent applications of affinity chromatography of proteins*

Substance purified	*Ligand*	*Elution*	*Ref.*
Papain		Water	162
Papain		$HgCl_2$ or 2-mercaptoethanol	163
Acetylcholinesterase		1M-NaCl	164
β-D-Xylosidase		2M-D-xylose	165
Chorismate mutase	Tryptophan	1mM-L-tryptophan	166
Lactose synthetase A protein	α-Lactalbumin	Omission of glucose from buffer	167
Lactose synthetase A protein	α-Lactalbumin	Omission of NAG[a] from buffer	168
Trypsin	Chicken ovomucoid	0.2M-KCl, pH2	169
Chymotrypsin and trypsin	Turkey ovomucoid	0.2M-KCl, pH2	170
Trypsin	Soybean trypsin inhibitor[b]	pH 2.3	171

[a] NAG = *N*-acetylglucosamine. [b] Attached to *p*-diazobenzoylcellulose

[162] S. Blumberg, I. Schechter, and A. Berger, *European J. Biochem.*, 1970, **15**, 97.
[163] L. A. A. Sluyterman and J. Wijdenes, *Biochim. Biophys. Acta*, 1970, **200**, 593.
[164] N. Kalderon, I. Silman, S. Blumberg, and Y. Dudai, *Biochim. Biophys. Acta*, 1970, **207**, 560.
[165] M. Claeyssens, H. Kersters-Hilderson, J.-P. Van Wauwe, and C. K. De Bruyne, *F.E.B.S. Letters*, 1970, **11**, 336.
[166] B. Sprossler and F. Lingens, *F.E.B.S. Letters*, 1970, **6**, 233.
[167] I. P. Trayer, P. Mattock, and R. K. Hill, *Fed. Proc.*, 1970, **29**, 597.
[168] P. Andrews, *F.E.B.S. Letters*, 1970, **9**, 297.
[169] G. Feinstein, *F.E.B.S. Letters*, 1970, **7**, 353.
[170] G. Feinstein, *Biochim. Biophys. Acta*, 1970, **214**, 224.
[171] V. V. Mosolov and E. V. Lushnikova, *Biokhimyia* (*Eng.*), 1970, **35**, 399.

Table 1 (*cont.*)

Substance purified	*Ligand*	*Elution*	*Ref.*
Ribonuclease inhibitor	Ribonuclease[c]	0.9M-NaCl	172
β-Galactosidase (*Aer. cloacae*)	Anti-*E. coli* β-galactosidase antibody	0.1M-NaCl, 10mM-$MgCl_2$	173
β-Galactosidase (*Aer. cloacae*)	Anti-*Aer. cloacae* β-galactosidase antibody	1.0M-NaCl	173
β-Galactosidase mutants	Anti-wild type β-galactosidase antibody	Urea	174
Mouse myeloma IgA (MOPC 315)	N^{ε}-Dnp—lysine	Dnp—glycine	175
^{32}P-tRNA	Anti-tRNA antibodies	tRNA (unlabelled)	176
Proteinase c (contaminant of phaseolin)	Anti-proteinase *c*	None (contaminant)	177
Angiotensin antibody	Angiotensin	0.1M acetic acid, pH 2.2	178
Human chorionic somatomammotropin (HCS)	Anti-HCS antibody	6M guanidine hydrochloride	179
Insulin	Anti-insulin antibody	1M acetic acid	180
Cells producing anti-phenyl-β-lactoside antibody	Phenyl β-lactoside[d]	Phenyl β-lactoside	181
Highly branched polysaccharides and glycoproteins	Concanavalin A	Methyl α-D-mannopyranoside	182

[c] Attached to carboxymethylcellulose. [d] Attached to polyacrylamide beads.

[172] A. A. M. Gribnau, J. G. G. Schoenmakers, M. van Kraaikamp, and H. Bloemendal, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 1064.
[173] R. P. Erickson and E. Steers, *Arch. Biochem. Biophys.*, 1970, **137**, 399.
[174] F. Melchers and W. Messer, *European J. Biochem.*, 1970, **17**, 267.
[175] E. J. Goetzl and H. Metzger, *Biochemistry*, 1970, **9**, 1267.
[176] B. Bonavida, S. Fuchs, and M. Sela, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 1335.
[177] W. F. Carey and J. R. E. Wells, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 574.
[178] E. Hurwitz, F. M. Dietrich, and M. Sela, *European J. Biochem.*, 1970, **17**, 273.
[179] B. D. Weintraub, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 83.
[180] Y. Akanuma, T. Kuzuya, M. Hayashi, T. Ide, and N. Kuzuya, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 947.
[181] P. Truffa-Bachi and L. Wofsy, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 685.
[182] K. O. Lloyd, *Arch. Biochem. Biophys.*, 1970, **137**, 460.

chromatography has been put during the past twelve months. Some special applications are dealt with later. Unless otherwise indicated the support was Sepharose (for the particular grade the reader is referred to the original paper), activated by the now standard cyanogen bromide method for covalent attachment of the ligand through its amino-groups.

One or two of the listed examples merit further comment. In general, the small molecules used as ligands in specific adsorption of enzymes are potent inhibitors of the enzymes, and presumably bind at the active site. Tryptophan, however, is a very strong activator of the allosteric enzyme chorismate mutase, which has binding sites for the activator and for the inhibitors L-Phe and L-Tyr as well as for the substrate.[166] The purification of the lactose synthetase A protein on a column of α-lactalbumin is interesting.[167, 168] Adsorption onto the column was in the presence of *N*-acetylglucosamine [168] or glucose,[167] and when these were omitted from the buffer the A protein was released from the column. This illustrates nicely that the interaction between the two components of the lactose synthetase system is enhanced in the presence of a substrate. It is apparently stronger in the presence of NAG, since retardation of A protein is greater in buffers containing NAG than those containing glucose.[168] Turkey ovomucoid-agarose retained both trypsin and chymotrypsin [170] but the former less efficiently. It was suggested that this was so because linkage of ligands to Sepharose occurs through amino-groups, and the trypsin inhibitory site on ovomucoid is thought to contain a lysine residue. It was possible to confirm the presence of anti-tRNA antibodies in goat and rabbit immunised with covalently bound tRNA conjugates by coupling γ-globulin to Sepharose and then observing selective adsorption of ^{32}P-tRNA.[176] The selective adsorption of whole cells [181] is remarkable.

An elegant method for isolating active-site peptides based on affinity chromatography has been demonstrated by Givol *et al.*[183] A protein is affinity-labelled and digested, and the digest applied to a Sepharose column to which the native protein itself is bound as ligand. Selective adsorption of the active-site peptide follows, by virtue of the interaction between the affinity label (bound to the peptide) and the active site of the protein. The method was successfully applied to peptides from the active site of an antibody (goat anti-DNP) and an enzyme (bovine pancreatic ribonuclease). This approach has also been used by Wilchek [184] to isolate two different affinity-labelled peptides from staphylococcal nuclease. A third labelled peptide resulting from a much less effective affinity label was not retained by the Sepharose-nuclease column, demonstrating the important point that only peptide-bound ligands having a high affinity for the column will be adsorbed. Biotin-containing peptides from methylmalonyl–oxaloacetic

[183] D. Givol, Y. Weinstein, M. Gorechi, and M. Wilchek, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 825.
[184] M. Wilchek, *F.E.B.S. Letters*, 1970, **7**, 161.

transcarboxylase have been bound to an avidin-Sepharose column,[185] using the biotin on the peptides as a built-in affinity label.

Cuatrecasas, one of the innovators of affinity chromatography, has recently given a comprehensive account[186] of the preparation of a large number of derivatives of agarose and polyacrylamide beads that will extend the scope of the method enormously. To minimise interactions of the ligand with the matrix an ω-aminoalkyl arm is attached by treatment of the activated support with an $\alpha\omega$-diamine. The ω-amino-group can then be used either for direct attachment of ligands (*e.g.* through their carboxy-groups) or it can be modified in a variety of ingenious ways so that ligands can be attached through amino-, carboxy-, phenolic, or imidazole groups, and there is no longer the restriction of mandatory coupling through amino-groups. The versatility of this approach is exemplified by the elution of a ligand–protein complex from a column where the ligand had been attached through its carboxy-groups to thiol groups in the matrix. This could provide an alternative to the drastic conditions sometimes needed (see list) to elute very tightly bound proteins from affinity columns. The presence of a flexible arm appears to be more important when the affinity of the ligand for the protein is low.

A recent note of an enterprising variation on the theme of 'biospecific adsorption' described the simultaneous isolation of trypsin inhibitor and anti-A phytohaemagglutinin from a suspension of ground seeds (*Vicia cracca*).[187] These were specifically adsorbed on to insolubilised trypsin and blood group A substance, respectively, contained in nylon mesh bags sitting in the stirred suspension. It was noted that 'bags' of insoluble proteolytic enzymes would be useful for controlled degradation. Not an example of affinity chromatography, but embodying the same principle of specific interactions, is the report of the purification of glucose-6-phosphate dehydrogenase, pseudocholinesterase, lactate dehydrogenase, and alanine dehydrogenase by substrate elution from ion exchange columns.[188]

3 Structural Proteins

The term 'structural proteins' is taken here to include the fibrous proteins, proteins connected with motility, and certain other globular proteins that have no identified enzymic activity. Immunoglobulins are dealt with separately in Section 6.

A. The Proteins of Motility.—The molecular weight of G-actin, previously put at *ca.* 45 000 by other workers, has been confirmed by light scattering and SDS-gel electrophoresis.[189] Studies on the primary structure of rabbit

[185] A. Bodanszky and M. Bodanszky, *Experientia*, 1970, **26**, 327.
[186] P. Cuatrecasas, *J. Biol. Chem.*, 1970, **245**, 3059.
[187] L. Sundberg, J. Porath, and K. Aspberg, *Biochim. Biophys. Acta*, 1970, **221**, 394.
[188] A. Yoshida, *Analyt. Biochem.*, 1970, **37**, 357.
[189] I. Sakakibara and K. Yagi, *Biochim. Biophys. Acta*, 1970, **207**, 178.

muscle actin are compatible with this figure.[190, 191] Thus, 17 peptides that together account for the whole molecule have been isolated after cyanogen bromide cleavage, and 3-methylhistidine was shown to occur in a single peptide.[190] The *N*- and *C*-terminal peptides were also identified. Other studies [192] have also shown that the 3-methylhistidine of actin is located in a single tryptic peptide and suggest that modified histidine residues are not essential to the biological activities of actin or myosin. Similarly, it has been found [193] that 3-methylhistidine is absent from red fibre, cardiac, and foetal myosins of the cat but is present (one residue per heavy chain) in white fibre myosin. It is not clear yet whether the absence in red fibre myosin is due to a failure to effect methylation or the absence of the susceptible histidine residue(s). There is adduced, however, a direct correlation between histidine methylation in white fibre myosin and the increasing maturity of the animal.

Cyanogen bromide has been used to effect limited cleavage of the myosin rod *ca.* 950 Å from the *C*-terminus, as seen in the electron microscope.[194] Chemical evidence has been produced [195] for heterogeneity in the two chains (mol. wt. 35 000) that comprise the molecule of rabbit tropomyosin. Either the two chains are similar, but not identical, or there is more than one form of tropomyosin.

An elegant method has been described [196] for the selective precipitation of microtubule protein from high-speed supernatants of homogenates of cultured mouse neuroblastoma cells by the addition of vinblastine. This is comparable with the effect of vinblastine in causing intracellular crystallisation of cytoplasmic microtubules with bound vinblastine. Fractionation of outer fibre doublet microtubules of cilia and flagella into A- and B-subfibre components has also been effected.[197] The A-subfibre is a complete cylinder to which is attached the ATPase, dyenin, in the form of arms, whereas the smooth B-subfibre shares part of the wall of the A-subfibre.[198] The subunit proteins A- and B-tubulin of the respective subfibres are nearly identical in terms of amino-acid analysis, peptide maps, *etc.* but show characteristic differences enough to suggest that they are the products of non-allelic genes.[197] Despite earlier suggestions to the contrary, it now seems likely that there is little or no sequence homology between the actin of striated muscle (subunit mol. wt. 45 000) and the A-tubulin (subunit mol. wt. 59 000), as shown by a comparison of the two isolated from the same organism, *Pecten irradians*.[199]

[190] R. S. Adelstein and W. M. Kuehl, *Biochemistry*, 1970, **9**, 1355.
[191] M. Elzinga, *Biochemistry*, 1970, **9**, 1365.
[192] P. Johnson and S. V. Perry, *Biochem. J.*, 1970, **119**, 293.
[193] W. M. Kuehl and R. S. Adelstein, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 956.
[194] M. V. King and M. Young, *J. Mol. Biol.*, 1970, **50**, 491.
[195] R. S. Hodges and L. B. Smillie, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 987.
[196] J. B. Olmsted, K. Carlson, R. Klebe, F. Ruddle, and J. Rosenbaum, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **65**, 129.
[197] R. E. Stephens, *J. Mol. Biol.*, 1970, **47**, 353.
[198] O. Behnke and A. Forer, *J. Cell Sci.*, 1967, **2**, 169.
[199] R. E. Stephens, *Science*, 1970, **168**, 845.

Some properties of the flagellin from *B. megaterium* KM have been reported.[200] The monomeric mol. wt. is 34 000 and the protein contains no tryptophan, cysteine, or cystine, and no *N*-methylated lysine. It has also been found [201] that after treatment with carboxypeptidase B, flagellin does not polymerise, suggesting that the *C*-terminal region of the molecule is part of a binding site essential to the polymerisation process. This is supported by the fact that the *C*-terminal residues become inaccessible to carboxypeptidase when the flagellin has polymerised to flagella, although the alternative explanation, a conformational change accompanying the polymerisation, cannot be excluded.

B. Collagen.—*Primary Structure.* Though most of the studies of the primary structure of collagen have been confined to simple mammalian sources, *e.g.* calf, rat, and human, the occurrence of Gly-Hyp-Ser sequences has been reported for earthworm cuticle collagen,[202] and the operculum of the whelk, *Buccinum undatum*, has been shown to consist of a protein whose amino-acid composition suggests it belongs to the collagen group.[203] A further study [204] of the specificity of bacterial collagenase A confirms that this enzyme splits only between X and Gly in the sequence -Pro-X-Gly-Pro-Y- where X can be any residue and Y is usually hydroxyproline or alanine. The same study further confirms the broader specificity of collagenase B which, in addition, cleaves the Y-Gly bond in the sequences -Gly-X-Y-Gly-Pro-Z- or -Gly-X-Y-Gly-Z-Hyp-. *O*-Acetylation of the hydroxy-groups in collagen has been claimed to render the molecule more susceptible to digestion with pronase,[205] perhaps by disturbance of the hydrogen-bonding at the susceptible sites.

However, in view of the large size of the collagen chain, most work on the primary structure starts with the peptides produced by cyanogen bromide cleavage. Thus, the sequence of the major hexose-containing peptide (37 residues) of the α1-chain of rat skin collagen [206] and the *N*-terminal sequence (55 residues) of the α1-chain of chick skin collagen [207] have been established (see last year's Report, p. 40) and the isolation and characterisation of five fragments that together account for the whole α2-chain of calf skin collagen has been reported.[208] Considerable homology with the corresponding peptides from rat skin, chick bone, and chick skin collagen is evident.

[200] R. Mirsky, *Arch. Biochem. Biophys.*, 1970, **139**, 97.
[201] W. Bode and H. Olossman, *Z. physiol. Chem.*, 1970, **351**, 1285.
[202] A. Goldstein and E. Adams, *J. Biol. Chem.*, 1970, **245**, 5478.
[203] S. Hunt, *Biochim. Biophys. Acta*, 1970, **207**, 347.
[204] E. Harper and A. H. Kang, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 482.
[205] D. Fujimoto, *Biochim. Biophys. Acta*, 1970, **200**, 172.
[206] W. T. Butler, *Biochemistry*, 1970, **9**, 44.
[207] A. H. Kang and J. Gross, *Biochemistry*, 1970, **9**, 796.
[208] P. P. Fietzek, M. Münch, D. Breitkreutz, and K. Kühn, *F.E.B.S. Letters*, 1970, **9**, 229.

Marked similarity is also shown between the *N*-terminal sequences of the α1-chains of human skin and rat tendon collagen:[209]

Human	**Glp-Leu-Ser-Tyr-Gly-Tyr-Asp-Glu-Lys-Ser-Thr-Gly-**	
Rat	**Glp-Met-Ser-Tyr-Gly-Tyr-Asp-Glu-Lys-Ser-Ala-Gly-**	
	1	**10**
Human	**Gly-Ser-Val-Ile-Pro-Gly-Pro-Met-**	
Rat	**Val-Ser-Val——Pro-Gly-Pro-Met-**	
	15	

this is good evidence of common evolutionary ancestry.

Some success has been achieved in ordering such cyanogen bromide fragments by attempting to renature the peptides and examining the products in the electron microscope. This has been reported for the α1- and α2-chain of chick skin collagen [99] and for the α2-chain of calf skin collagen.[208] The order of the cyanogen bromide peptides from the α2-chain of rat collagen has been partially established by chemical means and confirmed and completed by examining the pulse labelling of the collagen in culture.[210a] The order is determined by examining the specific radioactivity of the peptides in the manner originally described in the classic experiments of Dintzis on the biosynthesis of haemoglobin.[210b] Direct evidence has been produced [211] for a correlation between the distribution of polar residues in the amino-acid sequence and the cross-striation pattern of collagen observed in the electron microscope, by examining the banding pattern for a large peptide (112 residues) of known amino-acid sequence derived from the *C*-terminus of α1-chains of calf skin collagen.

Further work has been reported on the peculiar hydroxylamine-sensitive bonds in collagen. Thus, with peptide CB8 (mol. wt. 24 000) from near the middle of the α1-chain of rat collagen, treatment with 1M hydroxylamine at pH 10.5 and 35 °C for 90 min caused substantial cleavage of an Asx—Gly bond in the sequence:[212]

Gly-Ala-Asx-Gly-(Hyp$_2$,Gly$_2$,Ala$_3$,Ile)-(Hyp,Gly$_2$,Ala,Phe)-Arg

Cleavage of the α1- and α2-chains of calf skin collagen with hydroxylamine has also been reported:[213] glycine was the predominant new *N*-terminus in the products of the α1-chain. An examination of the treatment of bovine ribonuclease A with 2M hydroxylamine at pH 9.0 showed [214] that partial cleavage had occurred at the Asn(67)—Gly(68) bond and, to a lesser extent, at the Asn(34)—Leu(35) bond. It was suggested that cyclisation to the αβ-imide, particularly in Asn—Gly sequences, might be responsible (see Chapter 2, Part I, Section 2C of last year's Report).

[209] E. M. Click and P. Bornstein, *Biochemistry*, 1970, **9**, 4699.
[210] (*a*) J. Vuust, J. M. Lane, P. P. Fietzek, E. J. Miller, and K. A. Piez, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 703. (*b*) H. M. Dintzis, *Proc. Nat. Acad. Sci. U.S.A.*, 1961, **48**, 247.
[211] K. Mark, P. Wendt, F. Rexrodt, and K. Kühn, *F.E.B.S. Letters*, 1970, **11**, 105.
[212] P. Bornstein, *Biochemistry*, 1970, **9**, 2408.
[213] E. Heidemann and W. Heinrich, *European J. Biochem.*, 1970, **14**, 61.
[214] P. Bornstein and G. Balian, *J. Biol. Chem.*, 1970, **245**, 4854.

Proline hydroxylation takes place mainly in completed collagen chains.[215] The enzyme responsible, collagen proline hydroxylase, has been purified from chick embryo [216a] and from newborn rat skin.[216b]

Cross-links. The complete structural basis of collagen cross-linking is not yet clear. A detailed proof has been given [217] of the existence of the cross-link (syndesine) derived by aldol condensation of the δ-semialdehydes of lysine and hydroxylysine and of the cross-link 6,7-dehydrohydroxylysino–norleucine, both of which were described in detail in last year's Report. Lysino–norleucine, a cross-link previously found only in elastin, has also now been reported for chick skin [218] and calf skin [219] collagen. Another new cross-link, δ-hydroxylysino–δ′-hydroxynorleucine, has been identified in borohydride-reduced collagen of bovine Achilles tendon and the reduced form of the precursor aldehyde, δε-dihydroxynorleucine has been isolated.[220] The same workers have also recognised hydroxylysino–norleucine as a cross-link in borohydride-reduced collagen.[221]

These cross-links are all based on δ-semialdehydes derived from lysine, and the enzyme that oxidises lysine to its δ-semialdehyde (allysine) has been partially purified.[222, 223] It is irreversibly inhibited by the lathyrogen β-aminopropiononitrile *in vivo* and *in vitro*. Other studies of the enzyme [224] suggest that it is dependent on pyridoxal phosphate and copper, in accord with the fact that copper and pyridoxine deficiency in the diet impair cross-linking in collagen and elastin.

It has been suggested [225] that the progressive insolubility of collagen gels *in vitro* is due to the development of intermolecular cross-links that are independent of the intramolecular aldol cross-links based on lysine semi-aldehydes near the *N*-terminus of the molecule. Other studies on the formation of intermolecular cross-links *in vitro* indicate [226] that the intramolecular aldol cross-link does not remain intact during fibril formation and, further,[227] that the intramolecular cross-links are not separate entities but only intermediates of an intermolecular link. Their existence in solubilised collagen is seen as a result of the extraction procedure. An electron

[215] B. N. Bachra and A. J. van der Eb, *Biochemistry*, 1970, **9**, 3001.
[216] (*a*) M. Pänkäläinen, H. Aro, K. Simons, and K. I. Kivirikko, *Biochim. Biophys. Acta*, 1970, **221**, 559. (*b*) R. E. Rhoads and S. Udenfriend, *Arch. Biochem. Biophys.*, 1970, **139**, 329.
[217] A. J. Bailey, C. M. Peach, and L. J. Fowler, *Biochem. J.*, 1970, **117**, 819.
[218] A. H. Kang, B. Faris, and C. Franzblau, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 175.
[219] M. L. Tanzer and G. Mechanic, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 183.
[220] G. Mechanic and M. L. Tanzer, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 1597.
[221] M. L. Tanzer, G. Mechanic, and P. M. Gallop, *Biochim. Biophys. Acta*, 1970, **207**, 548.
[222] R. C. Siegel and G. R. Martin, *J. Biol. Chem.*, 1970, **245**, 1653.
[223] R. C. Siegel, S. R. Pinnell, and G. R. Martin, *Biochemistry*, 1970, **9**, 4486.
[224] L. J. Fowler, C. M. Peach, and A. J. Bailey, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 251.
[225] A. Shuttleworth and M. J. Glimcher, *Biochim. Biophys. Acta*, 1970, **200**, 332.
[226] C. Franzblau, A. H. Kang, and B. Faris, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 437.
[227] A. H. Kang and J. Gross, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 1307.

microscopic examination of cross-links in calf collagen [228] showed that bonds linking molecules side to side in the quarter-staggered array were easily cleaved by pepsin, whereas those linking molecules head to tail were not. Earlier conclusions that there might be α-amino-aldehydes have now been disproved [229] by the observation that the α-amino-alcohols obtained after sodium borohydride reduction of collagen are derived from peptide chain cleavages during the reduction. Thus, the analysis of collagen cross-links is at a particularly interesting stage at the moment and, in view of the fearsome size of the polypeptide chains involved, it is impressive that so much has been learned in such a short time.

C. Fibrinogen.—Human fibrinogen has been studied by SDS-gel electrophoresis.[230] Heterogeneity was observed in the α-chain, the major component having mol. wt. 73 000 and the minor component mol. wt. 70 000. The mol. wts. of the β- and γ-chains were 58 000 and 47 000, respectively. It is likely that plasmin digestion is responsible for the heterogeneity in the α-chain.[230] The action of plasmin on fibrinogen and fibrin appears to result in the formation of almost identical high molecular weight end-products.[231] The enzyme arvin from snake venom digests fibrinogen in a somewhat similar fashion to thrombin, except that only fibrinopeptide A is produced.[232] Since thrombin usually promotes the fibrin clot by activating the fibrin-stabilising factor, whereas arvin does not, arvin may find some use as an anti-coagulant.

In human fibrinogen, the *N*-terminal sequence of the α-chain is known and the *N*-terminal residue of the β-chain is pyrrolidone carboxylic acid. Edman degradation of the intact molecule (using ^{35}S-labelled phenyl isothiocyanate) therefore allows the *N*-terminal sequence of the γ-chain to be deduced. Using this technique, the sequence has been shown [233] to be Tyr-Val-Ala-Thr-Arg-Asp-Asn-, in confirmation of earlier results. Other work [234, 235] has revealed that the *C*-terminal residue of the β- and γ-chains is valine and that of the α-chain is proline.

Comparative studies [236] of A-chain fragments from ox, sheep, rabbit, horse, and human fibrinogens suggest that there are two highly conserved regions of primary structure, Phe(8)-Arg(19) and Cys(45)-Met(51). Since the first region includes the site of thrombin attack and the second contains elements of disulphide bridges, this may be of functional significance in

[228] B. K. Zimmermann, J. Pikkarainen, P. P. Fietzek, and K. Kühn, *European J. Biochem.*, 1970, **16**, 217.
[229] M. A. Paz, E. Henson, R. Rombauer, L. Abrash, O. A. Blumenfeld, and P. M. Gallop, *Biochemistry*, 1970, **9**, 2123.
[230] D. Mills and S. Karpatkin, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 206.
[231] G. A. Dudek, M. Klockewiak, A. Z. Budzykski, Z. S. Latallo, and M. Kopec, *Biochim. Biophys. Acta*, 1970, **214**, 44.
[232] W. H. Holleman and L. J. Coen, *Biochim. Biophys. Acta*, 1970, **200**, 587.
[233] E. Irion and B. Blombäck, *F.E.B.S. Letters*, 1970, **7**, 143.
[234] M. Okude and S. Iwanaga, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 450.
[235] K. Kasai, S. Yokoyama, and S. Ishii, *F.E.B.S. Letters*, 1970, **11**, 298.
[236] T. Söderquist and B. Blombäck, *F.E.B.S. Letters*, 1970, **7**, 321.

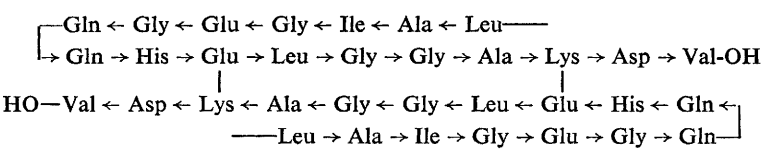
terms of biological activity and chain conformation. The fibrinopeptides of the gibbon (an Asian ape) have been compared with those of man and chimpanzee:[237]

Gibbon A
Ala-Asp-Thr-**Gly-Glu**-(**Gly**, Glu)-**Phe**-(**Leu**, **Ala**, **Glu**, **Gly**, **Gly**, **Gly**, **Val**)-**Arg**
Human A
Ala-Asp-Ser-**Gly-Glu**- **Gly**-Asp -**Phe** -**Leu**-**Ala**-**Glu**-**Gly**-**Gly**- **Gly**-**Val** -**Arg**
Chimpanzee A
Ala-Asp-Ser-**Gly-Glu**-(**Gly**, Asp)-**Phe**-(**Leu**, **Ala**, **Glu**, **Gly**, **Gly**, **Gly**, **Val**)-**Arg**

Gibbon B
Glp-(**Gly**, **Val**, Asx, Asx, Asx, Glx)——(Gly, Leu)-**Phe**-(Ser, Gly) **Ala**)-**Arg**
Human B
Glp- **Gly**-**Val** -Asn-Asp- Asn-Glu———**Gly**-Phe -**Phe**- Ser -**Ala** -**Arg**
Chimpanzee B
Glp-(**Gly**, **Val**, Asn, Asp, Asn, Glu, Glu, **Gly**, Phe)-**Phe**-(Ser, **Ala**)-**Arg**

These results are in line with the view that man is more closely related to the African apes (chimpanzee) than to the Asian apes. It is also interesting to note that there is an allelic variation (Ser, Gly) near the *C*-terminus of the gibbon B peptide, particularly so in view of the fact that the human/chimpanzee class has serine and the Old World monkeys have glycine in that position.

The cross-linking induced by fibrin-stabilizing-factor has been studied by SDS-gel electrophoresis.[238] Rapid formation of γ-γ links was observed together with a slower formation of higher polymers of the α-chain. The β-chains appear not to be involved. The amino-acid sequence around the γ-γ links of bovine fibrin has now been worked out.[239] The isopeptide links are found in overlapping *C*-terminal segments of neighbouring γ-chains:



The relationship of such covalent links to the folding of the polypeptide chains in molecular models needs careful consideration.

D. Chromosomal Proteins.—Among the remarkable properties of the histones are their limited heterogeneity and the high conservation of the amino-acid sequences of corresponding histones from various organisms

[237] G. A. Mross, R. F. Doolittle, and B. F. Roberts, *Science*, 1970, **170**, 468.
[238] P. A. McKee, P. Mattock, and R. L. Hill, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 738.
[239] R. Chen and R. F. Doolittle, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 472.

during evolution (see last year's Report). Thus, the histones F2a1 from normal and neoplastic tissues (foetal calf thymus, calf thymus; bovine lymphosarcoma, Novikoff hepatoma, and human leukaemic lymphoblasts) have very similar structures.[240] Similarly, the single cysteine-containing histone (histone III) of the five major histone fractions has a constant electrophoretic mobility when isolated from numerous sources.[241] On the other hand, characteristic differences are observed in the very lysine-rich histones F1 isolated from calf thymus, echinoderms, and molluscs.[242] The same major histone components are present in calf thymus, pea seedling, and undifferentiated HeLa cells,[243] but differences in the ratios of N^{ϵ}-dimethyl-lysine to N^{ϵ}-monomethyl-lysine in histones might be significant in gene regulation.[240]

Further studies [244] on the *N*- and *C*-terminal sequences of the glycine-rich, arginine-rich histone (histone IV) of calf thymus are in agreement with the whole sequence reported last year. The complete sequence of the slightly lysine-rich histone (F2b) of calf thymus has been determined [245] and is shown in Figure 1. The asymmetric clustering of the basic residues in this

Pro-Gln-Pro-Ala-Lys-Ser-Ala-Pro-Ala-Pro-Lys-Lys-Gly-Ser-Lys-Ala-Val-Thr-
1 **10**
Lys-Lys-Ala-Gln-Lys-Lys-Asp-Gly-Lys-Lys-Arg-Lys-Arg-Ser-Arg-Lys-Glu-Ser-
20 **30**
Tyr-Ser-Val-Tyr-Val-Tyr-Lys-Val-Leu-Lys-Gln-Val-His-Pro-Asp-Thr-Gly-Ile-
40 **50**
Ser-Ser-Lys-Ala-Met-Gly-Ile-Met-Asn-Ser-Phe-Val-Asn-Asp-Ile-Phe-Glu-Arg-
60 **70**
Ile-Ala-Gly-Glu-Ala-Ser-Arg-Leu-Ala-His-Tyr-Asn-Lys-Arg-Ser-Thr-Ile-Thr-
80 **90**
Ser-Arg-Glu-Ile-Gln-Thr-Ala-Val-Arg-Leu-Leu-Leu-Pro-Gly-Glu-Leu-Ala-Lys-
100
His-Ala-Val-Ser-Glu-Gly-Thr-Lys-Ala-Val-Thr-Lys-Tyr-Thr-Ser-Ser-Lys
110 **120**

Figure 1 *Amino-acid sequence of slightly lysine-rich histone* (F2b) *from calf thymus*

molecule, particularly in the *N*-terminal region, has been pointed out.[245, 246] A comparable asymmetry has been observed for the very lysine-rich histone of rabbit thymus,[247] although the biological significance remains obscure. At its most obvious, it might relate to the interaction with DNA.

[240] L. S. Desai and G. E. Foley, *Biochem. J.*, 1970, **119**, 165.
[241] S. Panyim, R. Chalkley, S. Spiker, and D. Oliver, *Biochim. Biophys. Acta*, 1970, **214**, 216.
[242] J. A. Subirana, J. Palau, C. Cozcolluela, and A. Ruiz-Carrillo, *Nature*, 1970, **228**, 992.
[243] A. Sadgopal and J. Bonner, *Biochim. Biophys. Acta*, 1970, **207**, 206.
[244] P. Sautierre, Y. Moschetto, M. Dautrevaux, and G. Bisterte, *European J. Biochem.*, 1970, **12**, 222.
[245] K. Iwai, K. Ishikawa, and H. Hayashi, *Nature*, 1970, **226**, 1056.
[246] L. S. Hnilica, H. A. Kappler, and J. J. Jordan, *Experientia*, 1970, **26**, 353.
[247] M. Bustin and R. D. Cole, *J. Biol. Chem.*, 1970, **245**, 1458.

Partial sequences of calf thymus histone III have been published.[248] The *N*-terminal sequence is Arg-Ala- and that at the *C*-terminus is:

-Arg-Val-Thr-Ile-Met-Pro-Lys-Asp-Ile-Gln-Leu-Ala-Arg-Arg-Ile-Arg-Gly
-Glu-Arg-Ala

Little similarity is yet apparent with the *C*-terminal sequences of the histones IV and F2b.

Modified amino-acids continue to be discovered in the histones. *N*-Acetylated and *N*-methylated lysine residues were found in calf thymus histone III, although the modification of the susceptible residues was observed to be incomplete.[248] Hydroxyproline has been detected in the lysine-rich histone (F1) of tobacco callus tissue [249] and ω-*N*-methylarginine was found in histones from rat liver nuclei.[250] The lysine-rich histones of mouse ascites tumour cells are heterogeneous because of phosphorylation [251] and, at a late stage of spermatogenesis in rainbow trout testis, the histones are extensively phosphorylated and acetylated and replaced by newly synthesised protamine.[252] It is possible [251, 252] that these modifications affect the interaction between histone and DNA and a detailed molecular model has been suggested.[252] Two enzymes that catalyse specific acetylation of histones have been partially purified from rat liver nuclei [253] and a protease tightly associated with thymus nucleohistone has been observed.[254] Tissues having a high rate of cell turnover show an increased rate of proteolysis of nucleohistone and the enzyme possibly is connected with autolysis and cell destruction.[254]

Limited heterogeneity of the major non-histone chromosomal proteins has also been demonstrated by gel electrophoresis of material from many different sources, suggesting that some at least of these proteins may have common roles, enzymic or structural, in the chromosome.[255]

E. Miscellaneous.—*Serum and Egg Proteins.* The *C*-terminal sequence of bovine serum albumin has been shown,[256] by investigating a number of model peptides, to be -Thr-Ala-Leu-Ala. The amino-acid sequence around the single tryptophan residue of human serum albumin has been found to be -Lys-Ala-Trp-Ala-Val-Ala-Arg.[257] The tryptophan residue is probably part of the strong binding site of this protein for steroids and long-chain aliphatic anions.

[248] R. J. DeLange, E. L. Smith, and J. Bonner, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 989.
[249] B. I. Sahai Srivastava, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 1357.
[250] W. K. Paik and S. Kim, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 224; W. K. Paik and S. Kim, *J. Biol. Chem.*, 1970, **245**, 88.
[251] D. Sherod, G. Johnson, and R. Chalkley, *Biochemistry*, 1970, **9**, 4611.
[252] M. T. Sung and G. H. Dixon, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 1616.
[253] D. Gallwitz, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 236.
[254] J. Bartley and R. Chalkley, *J. Biol. Chem.*, 1970, **245**, 4286.
[255] S. C. R. Elgin and J. Bonner, *Biochemistry*, 1970, **9**, 4440.
[256] C. Varsel, E. S. Nuwayser, and A. J. Richard, *Biochim. Biophys. Acta*, 1970, **200**, 417.
[257] J. B. Swaney and I. M. Klotz, *Biochemistry*, 1970, **9**, 2570.

Different amino-acid compositions have been recorded for the ovalbumins of chicken, duck, and turkey.[258] Of particular interest are the differences in the molar contents of sulphur-containing amino-acids:

	Cysteine	*Cystine*	*Methionine*
Chicken	4	1	15
Duck	2	1	23
Turkey	3	3	14

Ovalbumin is one of those rare proteins in which cysteine and cystine residues co-exist. A similar analysis for the hen ovalbumin has also been reported by other workers [259] and the *C*-terminal sequence demonstrated to be -Cys-Val-Ser-Pro, containing one end of the disulphide bridge.

Two recent reports [260] agree that transferrin can exist with 0, 1, or 2 iron atoms bound, contrary to earlier belief. Modification of the carboxy-groups of transferrin with glycine methyl ester and a water-soluble carbodiimide suggests [261] that carboxy-groups do not participate directly in the iron binding and that the two iron-binding sites are not identical. On the other hand, in a demonstration [262] that human transferrin comprises a single polypeptide chain of mol. wt. 77 000, there is evidence that some portions of the chain may be duplicated to provide the two binding sites.

SDS-Gel electrophoresis of the other iron-binding protein, apoferritin (from horse spleen), yields a subunit mol. wt. of *ca.* 18 000,[263] which suggests in turn that there may be 24 rather than 20 subunits in the intact molecule. Some preliminary studies on the peptides derived by cyanogen bromide cleavage have also been reported.[264]

Milk Proteins. The almost complete sequence of bovine α_{S1}-casein has been published (Figure 2).[265] The residue at position 191 is glutamic acid in the B variant and glycine in the C variant. The A variant lacks the residues from positions 14—26 of the B and C variants.[266] Preliminary work on bovine β-casein (mol. wt. 24 000) has also been reported.[267]

A new and better method for the isolation of κ-casein has been reported.[268] The rennin-sensitive bond of bovine κ-casein has now been shown not to be

[258] M. B. Smith and J. F. Back, *Austral. J. Biol. Sci.*, 1970, **23**, 1221.
[259] L. A. Fothergill and J. E. Fothergill, *Biochem. J.*, 1970, **116**, 555.
[260] (*a*) J. Williams, C. Phelps and J. M. Lowe, *Nature*, 1970, **226**, 858. (*b*) P. Aisen, S. H. Koenig, W. E. Schillinger, I. H. Scheinberg, K. G. Mann, and W. Fish, *Nature*, 1970, **226**, 859.
[261] A. Bezkorovainy and D. Grohlich, *Biochim. Biophys. Acta*, 1970, **214**, 37.
[262] K. G. Mann, W. W. Fish, A. C. Cox, and C. Tanford, *Biochemistry*, 1970, **9**, 1348.
[263] R. R. Crichton and C. F. A. Bryce, *F.E.B.S. Letters*, 1970, **6**, 121.
[264] R. R. Crichton and V. Barbiroli, *F.E.B.S. Letters*, 1970, **6**, 134.
[265] J.-C. Mercier, F. Grosclaude, and B. Ribadeau-Dumas, *European J. Biochem.*, 1970, **16**, 453.
[266] F. Grosclaude, M.-F. Mahé, J.-C. Mercier, and B. Ribadeau-Dumas, *F.E.B.S. Letters*, 1970, **11**, 109.
[267] B. Ribadeau-Dumas, F. Grosclaude, and J.-C.Mercier, *European J. Biochem.*, 1970, **14**, 451.
[268] K. K. Tripathi and C. W. Gehrke, *J. Chromatog.*, 1970, **46**, 280.

an ester but the Phe-Met bond linking the *C*-terminus of para-κ-casein to the *N*-terminus of the macropeptide.[269] The action of rennin on κ-casein produces at least seven genetically defined macropeptides and recent work [270] demonstrates that in the B variant an alanine and isoleucine residue replace an aspartic acid and threonine residue, respectively, of the A variant. The changes are not yet located in the primary structure but are definitely on the *C*-terminal side of the Lys(11)-Thr(12) bond.

Arg-Pro-Lys-His-Pro-Ile-Lys-His-Gln-Gly-Leu-Pro-Gln-Glu-Val-Leu-Asn-Glu-
1 **10**
Asn-Leu-Leu-Arg-Phe-Phe-Val-Ala-Pro-Phe-Pro-Gln-Val-Phe-Gly-Lys-Glu-Lys-
20 **30**
Val-Asn-Glu-Leu-Ser-Lys-Asp-Ile-Gly-(Asx, Thr, Ser_2, Glx_3, P_3)-Ala-Met-Glu-
40 **50**
Asp-Ile-Lys-Glu-Met-Glu-Ala-(Asx, Ser_4, Glx_5, Pro, Val_2, Ile_2, P_5)-Lys-His-Ile-Gln-
60 **70** **80**
Lys-Glu-Asp-Val-Pro-Ser-Glu-Arg-Tyr-Leu-Gly-Tyr-Leu-Glu-Gln-Leu-Leu-Arg-
90
Leu-Lys-Lys-Tyr-Lys-Val-Pro-Gln-Leu-Glu-Ile-Val-Pro-Asp-Ser-Ala-Glu-Glu-
100 **110** |
P
Arg-Leu-His-Ser-Met-Lys-Glu-Gly-Ile-His-Ala-Gln-Gln-Lys-Glu-Pro-Met-Ile-
120 **130**
Gly-Val-Asn-Gln-Glu-Leu-Ala-Tyr-Phe-Tyr-Pro-Glu-Leu-Phe-Arg-Gln-Phe-Tyr-
140 **150**
Gln-Leu-Asp-Ala-Tyr-Pro-Ser-Gly-Ala-Trp-Tyr-Tyr-Val-Pro-Leu-Gly-Thr-Gln-
160 **170**
Tyr-Thr-Asp-Ala-Pro-Ser-Phe-Ser-Asp-Ile-Pro-Asn-Pro-Ile-Gly-Ser-Glu-Asn-
180
Ser-Glu-Lys-Thr-Thr-Met-Pro-Leu-Trp
Gly (variant C)
190 **198**

Figure 2 *Partial amino-acid sequence of bovine* α_{S1}*-casein*

The isolation and characterisation of a β-lactoglobulin from pig milk has been described.[271] The mol. wt. is 18 500 and the *N*- and *C*-terminal residues are both valine. The column chromatography of bovine whey proteins has been reported [272] and β-lactoglobulin Droughtmaster, a unique carbohydrate-containing variant from Droughtmaster cattle, has been characterised.[273] A partial amino-acid sequence for bovine β-lactoglobulin has been reported.[274]

Virus Proteins. The molecular weights of ^{14}C-labelled capsid proteins of SV40 have been estimated on SDS-gels to be *ca.* 48 000, 33 000, and

[269] C. A. Macdonald and M. A. W. Thomas, *Biochim. Biophys. Acta*, 1970, **207**, 139.
[270] R. J. Hill, M. A. Naughton, and R. G. Wake, *Biochim. Biophys. Acta*, 1970, **200**, 267.
[271] E. Kessler and K. Brew, *Biochim. Biophys. Acta*, 1970, **200**, 449.
[272] J. M. Armstrong, K. E. Hopper, H. A. McKenzie, and W. H. Murphy, *Biochim. Biophys. Acta*, 1970, **214**, 419.
[273] K. Bell, H. A. McKenzie, W. H. Murphy, and D. C. Shaw, *Biochim. Biophys. Acta*, 1970, **214**, 427.
[274] G. von Frank and G. Braunitzer, *Biochim. Biophys. Acta*, 1970, **198**, 222.

24 000.[275] Other minor bands were also observed. Similarly, three major proteins of Kilham rat virus have molecular weights of 72 000, 62 000, and 55 000, respectively, that of mol. wt. 62 000 probably being the capsid protein.[276]

SDS-Gels have also been used to estimate the mol. wt. of protein subunits in viruses of the potato virus X group.[277] The values for white clover mosaic virus, clover yellow mosaic virus, papaya mosaic virus, and two strains of cactus virus X were all approximately 20 000. A value of 29 800 was obtained for potato virus X when it was rapidly isolated from leaf homogenates, but this was observed to fall to 24 000 if the virus was allowed to remain in the homogenate overnight at room temperature. Treatment of the pure virus with trypsin produced the same fall in subunit molecular weight and it therefore seems likely that proteolysis in the unfractionated homogenate is responsible. Some preliminary work on the primary structure of the coat protein of turnip yellow mosaic virus has also been reported.[278]

The amino-acid sequence of the coat protein of Qβ bacteriophage has been established and compared with the corresponding proteins from the bacteriophages f2 and R17.[279] As shown in Figure 3, enough homology exists to indicate that the proteins are related and to suggest that they probably fold similarly.[279] Peptide-mapping of the R17 coat protein synthesised *in vitro* by cell-free extracts of *E. coli* demonstrates the high fidelity of translation in this system.[280] In another bacteriophage, T4, there is evidence from SDS-gel electrophoresis that precursor proteins are converted into proteins of lower mol. wt. during capsid formation. Mutations that prevent normal capsid formation appear to inhibit this conversion.[281] The effects of treatment with carboxypeptidase B suggest that the T4D tail plate contains one or more proteins with a *C*-terminal arginine residue which is necessary for specific and essential interaction with another phage component during tail assembly.[282]

Other Proteins. A low-sulphur protein has been purified from a limited tryptic digest of reduced wool [283] and N^{ε}-(γ-L-glutamyl)-L-lysine cross-links have been observed in native wool keratins.[284] It has been suggested [285]

[275] M. Girard, L. Marty, and F. Suarez, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 97.
[276] L. A. Salzman and W. L. White, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 1551.
[277] R. Koenig, H. Stegemann, H. Francksen, and H. L. Paul, *Biochim. Biophys. Acta*, 1970, **207**, 184.
[278] J. Reinbolt, R. Peter, D. Stehelin, D. Collot, and H. Duranton, *Biochim. Biophys. Acta*, 1970, **207**, 532.
[279] W. Konigsberg, T. Maita, J. Katze, and K. Weber, *Nature*, 1970, **227**, 271.
[280] D. J. W. Burns and P. L. Bergquist, *Biochemistry*, 1970, **9**, 1310.
[281] J. Hosoda and R. Cone, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 1275.
[282] D. Shapiro and L. M. Kozloff, *J. Mol. Biol.*, 1970, **51**, 185.
[283] M. E. Hilburn, P. T. Speakman, and R. E. Yarwood, *Biochim. Biophys. Acta*, 1970, **214**, 245.
[284] R. S. Asquith, M. S. Otterburn, J. H. Buchanan, M. Cole, J. C. Fletcher, and K. L. Gardner, *Biochim. Biophys. Acta*, 1970, **207**, 342.
[285] Y. Tashiro and E. Otsuki, *Biochim. Biophys. Acta*, 1970, **214**, 265.

that the native fibroin molecule is composed of two subunits of approximately equal molecular weight linked by a disulphide bond. In other experiments, the gland fibroin was extracted from the silkworm, *Bombyx mori*, and the amino-acid sequences around the cystine residues studied

```
Qβ  Ala-Lys-Leu-Glu-Thr-Val-Thr-Leu-Gly-Asn-Ile-Gly-Lys-Asp-Gly-
f2      Ala-Ser-Asn-Phe-Thr-Gln-Phe-Val-Leu-Val-Asn-Asp-Gly-
    1                                   10

Qβ      Lys-Gln-Thr-Leu-Val-Leu-Asp-Pro-Arg-Gly-Val-Asn-Pro-Thr-
f2      Gly-Thr-Gly-Asn-Val-Thr-Val-Ala-Pro———Ser-Asn-Phe-Ala-
                        20

Qβ      Asn-Gly-Val-Ala-Ser-Leu-Ser-Gln-Ala-Gly-Ala-Val-Pro-Ala-
f2      Asn-Gly-Val-Ala-Glu-Trp-Ile-Ser-Ser-Asn-Ser———Arg-Ser-
        30                                      40

Qβ      Leu-Glu-Lys-Arg-Val-Thr-Val-Ser-Val-Ser-Gln-Pro-Arg———
f2      Gln-Ala-Tyr-Lys-Val-Thr-Cys-Ser-Val-Arg-Gln-Ser-Ser-Ala-
                                50

Qβ      ——Asn-Arg-Lys———Asn-Tyr-Lys-Val-Gln-Val———Lys-Ile-
f2      Gln-Asn-Arg-Lys-Tyr-Thr-Ile-Lys-Val-Glu-Val-Pro-Lys-Val-
                            60

Qβ      Gln-Asn-Pro-Thr-Ala-Cys-Thr-Ala-Asn-Gly-Ser-Cys-Asp-Pro-
f2      Ala-Thr-Gln-Thr-Val-Gly———Gly-Val-Glu-Leu————Pro-
            70                  75                  80

Qβ      Ser-Val-Thr-Arg-Gln-Ala-Tyr-Ala-Asp-Val-Thr-Phe-Ser-Phe-
f2      Val-Ala-Ala-Trp-Arg-Ser-Tyr-Leu-Asn-Leu-Glu-Leu-Thr-Ile-
                                    90

Qβ      Thr-Gln-Tyr-Ser-Thr-Asp-Glu-Glu-Arg-Ala-Phe———Val-Arg-
f2      Pro-Ile-Phe-Ala-Thr-Asn-Ser-Asp-Cys-Glu-Leu-Ile-Val-Lys-
                        100

Qβ      Thr-Glu-Leu-Ala-Ala-Leu-Leu-Ala-Ser-Pro-Leu-Leu-Ile-Asp-
f2      Ala-Met-Gln-Gly-Leu-Leu-Lys-Asp-Gly-Asn-Pro-Ile-Pro-Ser-
            110                                 120

Qβ      Ala-Ile-Asp-Gln-Leu-Asn-Pro-Ala-Tyr
f2      Ala-Ile-Ala-Ala-Asn-Ser-Gly-Ile-Tyr
                            130
```

Figure 3 *Amino-acid sequences of the coat proteins of bacteriophages* Qβ *and* f2

after their oxidation to cysteic acid.[285] This was made easier by labelling *in vivo* with ^{35}S. It is proposed that the only half-cystine containing sequence in gland fibroin is:

-Gly-Ala-Gly-Ala-Gly-Cys-Asx-Ser-Ala-Val-Cys-(Pro, Leu)-

and that the two half-cystine residues form an intra-chain disulphide bridge. This sequence may be at or near the *C*-terminus and new calcula-

tions based on the cysteic acid content of performic-oxidised fibroin give a revised mol. wt. of 103 000 for the peptide chain.[286]

There has been a substantial revision of the mol. wt. of the peptide chains of α-crystallin, from 20 000 to 11 500.[287] Some interesting studies have been reported [288] of the freezing-point-depressing glycoproteins of the antarctic fish, *Trematomus borchgrevinki*. These remarkable proteins consist largely of a backbone of alanine and threonine with carbohydrate in *O*-glycosidic linkage with the β-hydroxy-group of the threonine residues.

4 Peptides and Hormones

A. Pancreatic Hormones.—The fascinating proinsulin story has now been reviewed.[289] Proinsulin (like insulin) interacts with zinc to form a hexameric complex (mol. wt. 55 000) with a minimum of two zinc ions per hexamer.[290a] There is speculation that the zinc may stabilise the proinsulin so that the proteolysis leading to insulin is confined to the appropriate bonds. It now appears [290b] that residue 35 of pig proinsulin (in the C-peptide) is glutamic acid and not glutamine as originally assigned. Des-Gly^{A1}-des-Phe^{B1}-insulin has been prepared by limited Edman degradation of the molecule without effect on the ε-amino-group.[291] The derivative retains some hormonal activity.

Details of the amino-acid sequence analysis of pig secretin have now been published.[292] This hormone, produced by the duodenal mucosa, shows considerable homology with the pancreatic hormone glucagon, despite the fact that their biological activities are totally different. This evidence of a common genetic ancestry is pointed up by the similar embryonic origin of pancreas and duodenum.

B. Pituitary Hormones.—The amino-acid sequence of melanocyte-stimulating hormone (MSH) from dogfish *Squalus acanthias* has been shown to be:[293]

Dogfish	**Ser-Met-Glu-His-Phe-Arg-Trp-Gly-Lys-Pro-Met**
Human α-MSH	Ac-Ser-Tyr-**Ser-Met-Glu-His-Phe-Arg-Trp-Gly-Lys-Pro-Val-NH$_2$**

Half the molecules of dogfish MSH have a free α-carboxy-group and half are amidated. None has an *N*-acetyl group, though 20% carry an extra tyrosine at the *N*-terminus. The homology with human α-MSH is very strong, stronger in fact than with β-MSH, suggesting that α-MSH may be

[286] A. Robson, J. M. Woodhouse, and Z. H. Zaidi, *Internat. J. Protein Res.*, 1970, **2**, 181.
[287] K. De Groot, H. J. Hoenders, J. J. T. Gerding, and H. Bloemendal, *Biochim. Biophys. Acta*, 1970, **207**, 202.
[288] S. K. Komatsu, A. L. De Vries, and R. E. Feeney, *J. Biol. Chem.*, 1970, **245**, 2909.
[289] P. T. Grant and T. L. Coombs, *Essays in Biochemistry*, 1970, **6**, 69.
[290a] B. H. Frank and A. J. Veros, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 284.
[290b] R. E. Chance, cited by G. Jäger and R. Geiger, *Chem. Ber.*, 1970, **103**, 1727.
[291] B. Africa and F. H. Carpenter, *Biochemistry*, 1970, **9**, 1962.
[292] V. Mutt, J. E. Jorpes, and S. Magnusson, *European J. Biochem.*, 1970, **15**, 513.
[293] P. J. Lowry and A. Chadwick, *Nature*, 1970, **226**, 219.

more primitive than β-MSH.[294] Chemical modification of the single tryptophan residue in the structurally related corticotropin (ACTH) with sulphenyl halides results in loss of the lipolytic activity of the hormone [295] without, however, affecting the melanocyte-stimulating activity.[296]

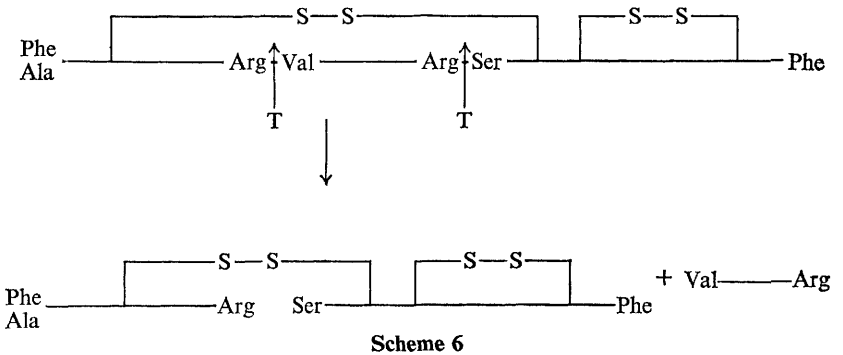
Three neurophysins (proteins that bind the hormones oxytocin and vasopressin) have been purified from pig pituitaries.[297] All have *N*-terminal alanine, although they range in mol. wt. from 9400 to 14 000. Lipolytic peptides that resemble the neurophysins have also been isolated from pig pituitary glands.[298]

The preparation of luteinising hormone (LH) from human pituitary glands has been reported,[299] yielding a protein of mol. wt. 27 000, and separation of the two subunits of sheep [300] and ox [301] LH by chromatography and counter-current distribution, respectively, has been described. The ox and sheep hormones resemble one another closely [301] and, after the subunits of the ox hormone have been separated, they can be recombined with regeneration of *ca.* 50% of the hormonal activity.[301] The two subunits of LH, though similar in mol. wt., are dissimilar in sequence. The same is true of thyroid-stimulating hormone (TSH).[302, 303] In fact, there is some evidence [303] that TSH and LH may each have one subunit derived from a common ancestor, associated appropriately with one or other of two quite different chains. In support of this idea, perhaps, it is possible [302] to form biologically active molecules by hybridisation of subunits from LH and TSH. The thyrotropin-releasing hormone of pig has been shown to have the amino-acid sequence Glp-His-Pro-NH_2.[304] If the synthetic peptide containing *N*-terminal glutamine is prepared, it is found [305] that the *N*-terminal residue converts to pyrrolidonecarboxylic acid under very mild conditions, but there is no evidence that *N*-terminal glutamine occurs in the natural peptide. The same structure is found for the hormone of several other species; many synthetic analogues have markedly reduced biological activity.[306]

There has been considerable activity in the study of growth hormones and prolactins. Partial sequences have been reported for the cyanogen

[294] P. J. Lowry and A. Chadwick, *Biochem. J.*, 1970, **118**, 713.
[295] J. Ramachandran and V. Lee, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 507.
[296] J. Ramachandran, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 353.
[297] L. O. Uttenthal and D. B. Hope, *Biochem. J.*, 1970, **116**, 899.
[298] D. Rudman, A. E. Del Rio, L. A. Garcia, J. Barnett, C. H. Howard, W. Walker, and G. Moore, *Biochemistry*, 1970, **9**, 99.
[299] P. Rathnam and B. B. Saxena, *J. Biol. Chem.*, 1970, **245**, 3725.
[300] W. M. Lamkin, M. Fujino, J. D. Mayfield, G. N. Holcomb, and D. N. Ward, *Biochim. Biophys. Acta*, 1970, **214**, 290.
[301] H. Papkoff and J. Gan, *Arch. Biochem. Biophys.*, 1970, **136**, 522.
[302] G. Hennen, G. Maghuin-Rogrster, and G. Mamoir, *F.E.B.S. Letters*, 1970, **9**, 20.
[303] T.-H. Liao and J. G. Pierce, *J. Biol. Chem.*, 1970, **245**, 3275.
[304] R. M. G. Nair, J. F. Barrett, C. Y. Bowers, and A. V. Schally, *Biochemistry*, 1970, **9**, 1103.
[305] K. Folkers, J.-K. Chang, B. L. Currie, C. Y. Bowers, A. Weil, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 110.
[306] C. Y. Bowers, A. Weil, J.-K. Chang, H. Sievertsson, F. Enzmann, and K. Folkers, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 683.

bromide fragments of pig growth hormone,[307] and the sheep and ox hormones have been shown to be highly homologous.[308] The *C*-terminal sequence of the sheep protein [308] -(Phe, Gly, Glu)-Ala-Ser-Cys-Ala-Phe is identical with that of ox. Limited tryptic digestion of ox growth hormone (Scheme 6) releases a peptide of mol. wt. 5000 which contains no disulphide



Scheme 6

bridges but retains some biological activity.[309] Similarly, a large peptide starting near the *N*-terminus of the molecule, produced by cyanogen bromide cleavage of the ox hormone after reduction and aminoethylation, also has slight biological activity.[310] Since these two active fragments produced by different methods have a relatively short stretch of amino-acid sequence in common, a comparison of further work with these systems will be of much interest.

The complete sequence of sheep pituitary prolactin (SP) has now been reported (Figure 4).[311] Considerable homology has been demonstrated with the sequence of human growth hormone (HGH),[312] suggesting that these molecules have elements of tertiary structure in common. A study of the conformations of ox growth hormone and sheep prolactin by o.r.d. and c.d. would support this.[313] Since HGH is active both as a growth hormone and as a lactogenic hormone, it seems likely that the areas of homology between human growth hormone and sheep prolactin will be connected with lactogenic activity.[313] In fact, doubts have been voiced in the past as to the existence of a human pituitary prolactin but recent

[307] J. B. Mills, S. C. Howard, S. Scapa, and A. E. Wilhelmi, *J. Biol. Chem.*, 1970, **245**, 3407.
[308] J. M. Dellacha, M. A. Enero, J. A. Santomé, and A. C. Paladini, *European J. Biochem.*, 1970, **12**, 289.
[309] N. Yamasaki, M. Kikutani, and M. Sonenberg, *Biochemistry*, 1970, **9**, 1107.
[310] D. F. Nutting, J. L. Kostyo, J. B. Mills, and A. E. Wilhelmi, *Biochim. Biophys. Acta*, 1970, **200**, 601.
[311] C. H. Li, J. S. Dixon, T.-B. Lo, K. D. Schmidt, and Y. A. Pankov, *Arch. Biochem. Biophys.*, 1970, **141**, 705.
[312] T. A. Bewley and C. H. Li, *Science*, 1970, **168**, 1361.
[313] S. M. Aloj and H. Edelhoch, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 830.

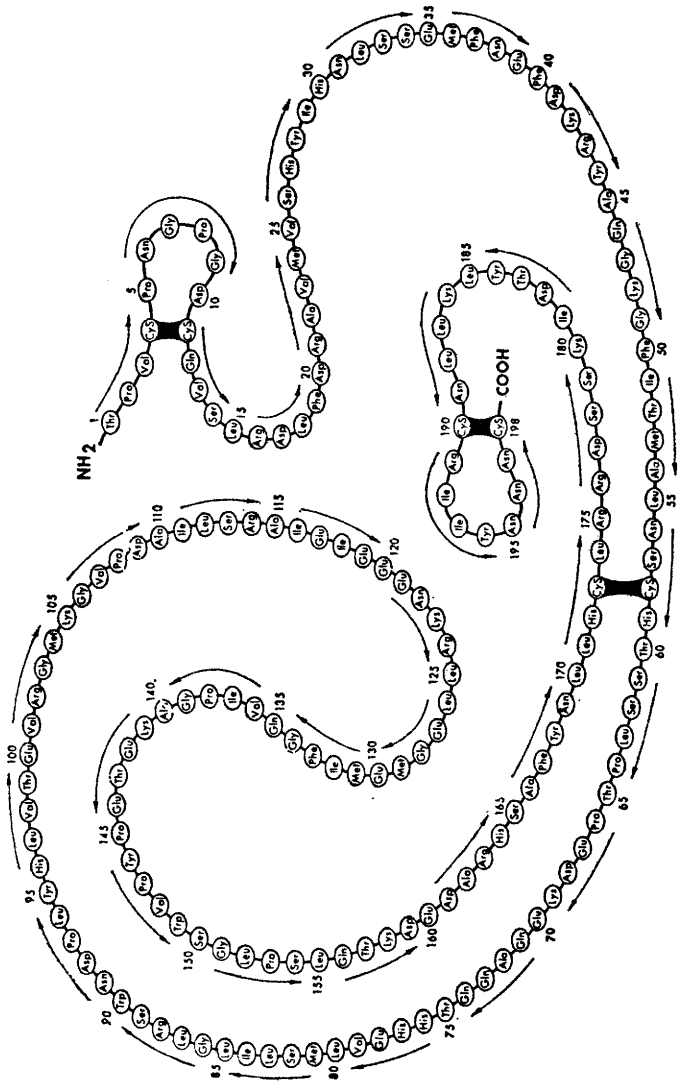


Figure 4 *Amino-acid sequence of the sheep prolactin molecule*
(Reproduced by permission from *Arch. Biochem. Biophys.*, 1970, **141**, 705)

experiments do favour such a possibility.[314] In other experiments[315] it has been shown that all seven tyrosine residues in sheep prolactin can be modified with tetranitromethane without loss of biological activity, despite detectable changes in hormone conformation. However, a major disturbance has been caused in this apparently orderly corner of the hormone field by the determination of the amino-acid sequence of the related hormone, human placental lactogen (HPL).[316] A comparison of the amino-acid sequences of HPL, HGH, and SP reveals[316] that the expected homology between HPL and HGH exists but that the published primary structure of HGH[317] is wrong, certainly at the *N*-terminus and probably elsewhere. A re-examination of the *N*-terminal sequence of HGH suggests[316] that the mistake almost certainly lies in a transposition of two major segments of the original sequence. The corrected *N*-terminal sequence for HGH is shown in Figure 5, together with that for HPL and SP. The close homology

HGH	Phe[1]	-Pro	-Thr	-Ile	-Pro	-Leu	-Ser	-Arg	-Leu	-Phe[10]	-Asp	-Asn	-Ala	-Met-					
HPL	Val[1]	-Gln	-Thr	-Val	-Pro	-Leu	-Ser	-Arg	-Leu	-Phe[10]	-Asp	-His	-Ala	-Met-					
SP	Asp[10]	-Cys	-Gln	-Val	-Ser	-Leu	-Arg	-Asp	-Leu	-Phe[19]	-Asp	-Arg	-Ala	-Val -					
HGH	Leu	-Arg	-Ala	-His	-Arg	-Leu[20]	-His	-Gln	-Leu	-Ala	-Phe	-Asp	-Thr	-Tyr -					
HPL	Leu	-Gln	-Ala	-His	-Arg	-Ala[20]	-His	-Gln	-Leu	-Ala	-Ile	-Asp	-Thr	-Tyr -					
SP	Met	-Val	-Ser	-His	-Tyr	-Ile[29]	-His	-Asn	-Leu	-Ser	-Ser	-Glu	-Met	-Phe -					
HGH	Glu	-Glu[30]	-Phe	-Glu	-Glu	-Ala	-Tyr[35]-												
HPL	Gln	-Glu[30]	-Phe	-Glu	-Glu	-Thr	-Tyr[35]-												
SP	Asn	-Glu[39]	-Phe	-Asp	-Lys	-Arg	-Tyr[44]-												

Figure 5 *Comparative* N*-terminal amino-acid sequences of human growth hormone (HGH), human placental lactogen (HPL), and sheep prolactin (SP)*

is obvious. But – hills peep o'er hills, and Alps on Alps arise[318] – the most significant consequence is, perhaps, yet to come, for Li and Yamashiro[319] have now completed the synthesis of human growth hormone utilising the erroneous sequence and demonstrated that it has some biological activity. What the future holds in store can best be described as uncertain.

C. Other Hormones.—Gastrin and its activities have been reviewed,[320] and the decapeptide Leu-Ala-Ala-Gly-Lys-Val-Glu-Asp-Ser-Asp has been isolated from normal human gastric juice.[321] The physiological significance of this peptide is obscure, though it is possible that it is involved in binding vitamin B_{12}.

[314] A. G. Frantz and D. L. Kleinberg, *Science*, 1970, **170**, 745.
[315] L. Ma, J. Brovetto-Cruz, and C. H. Li, *Biochemistry*, 1970, **9**, 2302.
[316] H. D. Niall, *Nature New Biology*, 1971, **230**, 90.
[317] C. H. Li, J. S. Dixon, and W.-K. Liu, *Arch. Biochem. Biophys.*, 1969, **133**, 70.
[318] Alexander Pope, *Essay on Criticism.*
[319] C. H. Li and D. Yamashiro, *J. Amer. Chem. Soc.*, 1970, **92**, 7608.
[320] M. I. Grossman, *Nature*, 1970, **228**, 1147.
[321] J. G. Heathcote and R. J. Washington, *Internat. J. Protein Res.*, 1970, **2**, 117.

The complete amino-acid sequence of ox parathyroid hormone has been established in two independent investigations [322, 323] and happily these agree (Figure 6). Much use was made of the sequenator in both cases, a

Ala-Val-Ser-Glu-Ile-Gln-Phe-Met-His-Asn-Leu-Gly-Lys-His-Leu-Ser-Ser-Met-
1 **10**
Glu-Arg-Val-Glu-Trp-Leu-Arg-Lys-Lys-Leu-Gln-Asp-Val-His-Asn-Phe-Val-Ala-
20 **30**
Leu-Gly-Ala-Ser-Ile-Ala-Tyr-Arg-Asp-Gly-Ser-Ser-Gln-Arg-Pro-Arg-Lys-Lys-
40 **50**
Glu-Asp-Asn-Val-Leu-Val-Glu-Ser-His-Gln-Lys-Ser-Leu-Gly-Glu-Ala-Asp-Lys-
60 **70**
Ala-Asp-Val-Asp-Val-Leu-Ile-Lys-Ala-Lys-Pro-Gln
80

Figure 6 *Amino-acid sequence of ox parathyroid hormone*

stepwise degradation for 66 residues actually being achieved by Brewer and Ronan. There is a suggestion [323] (based partly on a study of synthetic peptides) that the *N*-terminal 30 residues contain the structural requirements for biological activity. There is no obvious sequence similarity with the physiological antagonist, thyrocalcitonin,[322] detailed evidence for the already-published structure of which has been given.[324] Thyroglobulin, which provides a matrix for the synthesis of thyroid hormones, is a glycoprotein of mol. wt. 660 000. It has been reported [325] that thyroxine is to be found in the sequence -Ala-Ser-Thyroxine-Glx-Asx-.

The evolution of the neurohypophysial hormones has been discussed [326] in terms of the active peptides of a primitive bony fish, *Polypterus bichir*. Isotocin (Ser^4-Ile^8-oxytocin) and vasotocin (Arg^8-oxytocin) were identified and, since these molecules are replaced in mammals by oxytocin and vasopressin, are thought to be ancestor molecules. Similarly, another group of bony fishes has mesotocin (Ile^8-oxytocin), bridging the gap to the amphibians (mesotocin) and thence to the oxytocin of mammals. Various chemical modifications of oxytocin suggest [327] that the phenolic hydroxy group of Tyr-2 is important in binding the hormone to its receptor, and that the α-amino group of the hormone plays a role in the biological activity that follows binding. Vasotocin has been identified [328] as a gonadotropin-inhibitor from the pineal gland of the ox.

[322] H. B. Brewer, jun., and R. Ronan, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 1862.
[323] H. D. Niall, H. Keutmann, R. Sauer, M. Hogan, B. Dawson, G. Aurbach, and J. Potts, jun., *Z. physiol. Chem.*, 1970, **351**, 1506.
[324] (*a*) P. H. Bell, D. F. Colucci, C. Dziobkowski, E. H. Snedeker, W. F. Barg, jun., and R. Paul, *Biochemistry*, 1970, **9**, 1665. (*b*) W. F. Barg, jun., M. E. Englert, M. C. Davies, D. F. Colucci, E. H. Snedeker, C. Dziobkowski, and P. H. Bell, *Biochemistry*, 1970, **9**, 1671.
[325] J. T. Dunn, *J. Biol. Chem.*, 1970, **245**, 5954.
[326] R. Acher, J. Chauvet, and M.-T. Chauvet, *F.E.B.S. Letters*, 1970, **11**, 332.
[327] D. G. Smyth, *Biochim. Biophys. Acta*, 1970, **200**, 395.
[328] D. W. Cheesman, *Biochim. Biophys. Acta*, 1970, **207**, 247.

D. Kinins, Encephalitogenic Protein, and Toxins.—Cyanogen bromide treatment of ox kininogen-I gives two kinin-containing peptides.[328] Since treatment of the kininogen with carboxypeptidase B destroys half the potential kinin activity, it is concluded [329] that ox kininogen-I contains two -Met-Lys-bradykinin- sequences, one at the *C*-terminus and the other in the interior of the molecule. On the basis of amino-acid composition, it is probable [330] that the sequence of a new plasma kinin from the turtle (*Pseudemys scripta elegans*) is Thr^6-bradykinin, *i.e.* Arg-Pro-Pro-Gly-Phe-Thr-Pro-Phe-Arg, this being the first example of a modified bradykinin occurring naturally. Various bradykinin-potentiating peptides have been described. One from the venom of *Bothrops jararaca* has the sequence Glp-Lys-Trp-Ala-Pro.[331] Others from the venom of *Agkistrodon halys blanhoffii* have the sequences Glp-Gly-Leu-Pro-Pro-Arg-Pro-Lys-Ile-Pro-Pro [332] and Glp-Lys-Trp-Asp-Pro-Pro-Pro-Val-Ser-Pro-Pro.[333] Dipeptidyl carboxypeptidase can inactivate bradykinin and also convert angiotensin I into the active angiotensin II.[334]

A simplified purification of the basic encephalitogenic protein of myelin has been given.[335] Some studies of the cyanogen bromide cleavage products have been reported [336] and it has been demonstrated [337] that the A1 proteins from human and ox brain are very similar but distinct. This is confirmed by a comparison of the complete sequences of the two proteins.[338] Ox encephalitogenic protein contains 170 residues, that of human containing 172 [because of an insertion of the dipeptide -His-Gly- between Arg(10)-Ser(11) of the ox protein]. Apart from this, the differences are confined to some 10 residues. It is perhaps worth remarking that this protein comprises some 30% of the myelin membrane.

The complete amino-acid sequence of cobrotoxin has been reported (Figure 7).[339] The disulphide bridges were established by the diagonal procedure and for this purpose the protein had to be cleaved with acid protease A since pepsin was ineffective. It is of interest that the loop between residues 55 and 60 is the same size as that of oxytocin and vasopressin, although whether this of any significance is not yet clear.[339] *N*-Terminal sequence analysis of the scorpion neurotoxins shows them to be a family of homologous proteins [340] and modification with 2-hydroxy-5-nitrobenzyl

[329] M. Yano, S. Nagasawa, and T. Suzuki, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 914.
[330] R. S. Dunn and A. M. Perks, *Experientia*, 1970, **26**, 1220.
[331] S. H. Ferreira, D. C. Bartelt, and L. J. Greene, *Biochemistry*, 1970, **9**, 2583.
[332] H. Kato and T. Suzuki, *Proc. Japan Acad.*, 1970, **46**, 176.
[333] H. Kato and T. Suzuki, *Experientia*, 1970, **26**, 1205.
[334] H. Y. T. Yang, E. G. Erdös, and Y. Levin, *Biochim. Biophys. Acta*, 1970, **214**, 374.
[335] H. Hirshfeld, D. Teitelbaum, R. Arnon, and M. Sela, *F.E.B.S. Letters*, 1970, **7**, 317.
[336] L.-P. Chao and E. R. Einstein, *J. Biol. Chem.*, 1970, **245**, 6397.
[337] Y. Oshiro and E. H. Eylar, *Arch. Biochem. Biophys.*, 1970, **138**, 606.
[338] E. H. Eylar, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 1425.
[339] C. C. Yang, H. J. Yang, and R. H. C. Chiu, *Biochim. Biophys. Acta*, 1970, **214**, 355.
[340] H. Rochat, C. Rochat, C. Kupeyan, F. Miranda, S. Lissitzky, and P. Edman, *F.E.B.S. Letters*, 1970, **10**, 349.

bromide of the tryptophan residue of sea snake erabutoxin *a* causes loss of activity.[341]

The purification of the type A toxin (mol. wt. 150 000) from *Clostridium botulinum* has been described[342] and the complete amino-acid sequence (239 residues) of staphylococcal enterotoxin B has been established.[343] The single disulphide bridge in this molecule can be reduced and methylated without loss of biological activity or of conformation.

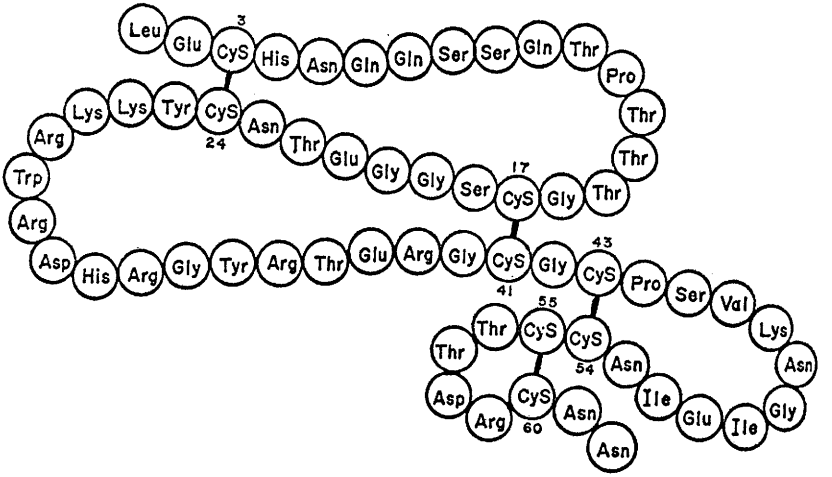


Figure 7 *Amino-acid sequence of cobrotoxin*
(Reproduced by permission from *Biochim. Biophys. Acta*, 1970, **214**, 355)

Mellitin (a toxic peptide from bee venom) has been subjected to a variety of separate chemical modifications.[344] The tryptophan has been reacted with 2-hydroxy-5-nitrobenzyl bromide, the amino-groups have been acetylated or succinylated, and guanidination of the ε-amino-groups has been achieved with *S*-methylisothiourea. When tested for haemolytic activity, the derivatives decreased in activity in the order guanidino > hydroxynitrobenzyl > native > acetyl > succinyl.

E. Peptide Antibiotics.—The *N*-terminal sequence of the anti-tumour antibiotic neocarzinostatin has been shown to be Ala-Ala-Pro-Thr-, while the *C*-terminal sequence is -Ser-Val-Ala-Ile-Phe-Asn.[345] The polypeptide has been reported to contain two tryptophan residues per molecule and to possess disulphide bridges that are resistant to reduction even in 8M urea.[346]

[341] A. Seto, S. Sato, and N. Tamiya, *Biochim. Biophys. Acta*, 1970, **214**, 483.
[342] B. R. Dasgupta, L. J. Berry, and D. A. Boroff, *Biochim. Biophys. Acta*, 1970, **214**, 343.
[343] I.-Y. Huang and M. S. Bergdoll, *J. Biol. Chem.*, 1970, **245**, 3518.
[344] E. Habermann and H. Kowallek, *Z. physiol. Chem.*, 1970, **351**, 884.
[345] H. Maeda and J. Meienhofer, *F.E.B.S. Letters*, 1970, **9**, 301.
[346] H. Maeda and J. Meienhofer, *Internat. J. Protein Res.*, 1970, **2**, 135.

The amino-acid sequence of the antibiotic alamethicin from *Trichoderma viride* has been reported [see Chapter 4, Section 2D, Structure (12),].[347] It is interesting in that the ring is closed by a peptide link between the γ-carboxy group in Glu-17 and the imino-group of Pro-1. One possible structure derived from model-building stacks to form a tunnel with lipophilic exterior and hydrophilic interior and flexible internal arms. Since the antibiotic transports cations and induces action potentials in synthetic membranes, this structure, though only one of several possible, might help to explain its biological activity.[347]

5 Enzymes

This section is largely concerned with the results of studies of the primary structure of enzymes and certain other proteins. Some discussion of chemical modification is included here: more is to be found in Section 8. Two Symposium volumes relating to enzymes have appeared: one, relating to structure and function, after some delay;[348] the other,[349] which would perhaps more appropriately preface Section 8, rather more promptly.

A. Proteolytic Enzymes.—Last year the complete primary structure of bovine carboxypeptidase A emerged from Neurath's laboratory.[350] The same workers have now produced chemical evidence for a single disulphide bridge between Cys-138 and Cys-161 in the enzyme,[351] in accord with the *X*-ray crystallographic evidence.[352] Stoicheiometric reaction of the water-soluble carbodi-imide, 1-cyclohexyl-3-(2-morpholinoethyl)carbodi-imide methotoluene-*p*-sulphonate, with carboxypeptidase A at pH 6.0 abolished both the esterase and the peptidase activities.[353] The results were consistent with specific reaction of a single carboxy group, thought to be probably Glu-270, a component of the active site. Similar results were obtained[354] with *N*-ethyl-5-phenylisoxazolium-3′-sulphonate.

Mention was made last year of the elucidation of the sequence of porcine elastase,[355] and of the extensive homology with the primary structures of chymotrypsin and trypsin that confirmed its membership of the serine protease club. The sequence of another proteolytic enzyme, α-lytic pro-

[347] J. W. Payne, R. Jakes, and B. S. Hartley, *Biochem. J.*, 1970, **117**, 757.
[348] F.E.B.S. Symposia, ed. D. Shugar, Academic Press, London and New York, Vol. 18, 1970.
[349] Biochemical Society Symposia, Number 31, ed. R. M. S. Smellie, Academic Press, London and New York, 1970, p. 49.
[350] R. A. Bradshaw, L. H. Ericsson, K. A. Walsh, and H. Neurath, *Proc. Nat. Acad. Sci. U.S.A.*, 1969, **63**, 1389.
[351] K. A. Walsh, L. H. Ericsson, R. A. Bradshaw, and H. Neurath, *Biochemistry*, 1970, **9**, 219.
[352] W. N. Lipscomb, J. A. Hartsuck, G. N. Roeke, F, A. Quiocho, P. Bethge, M. L. Ludwig, T. A. Steitz, H. Muirhead, and J. C. Coppola, *Brookhaven Symp. Biol.*, 1968, **21**, 24.
[353] J. F. Riordan and H. Hayashida, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 122.
[354] P. H. Pétra and H. Neurath, *Fed. Proc.*, 1970, **29**, 666 Abs.
[355] D. M. Shotton and B. S. Hartley, *Nature*, 1970, **225**, 802.

tease from Myxobacter 495 (*Sporangium* sp.),[356] is now complete but the homology is not nearly as high. However, it is clearly apparent in certain localised regions, particularly those containing the catalytically important residues. In order to maximise the homology between the sequence of α-lytic protease and the mammalian proteases it is necessary to recognise deletions or insertions of blocks of amino-acid residues.[356] A comparison[357] of α-lytic protease and elastase (which has similar specificity) is shown in Figure 8. Thus, if α-lytic protease does belong to the family it is by far the most distant relation.

The connexion between three-dimensional structure and amino-acid sequence for the members of a family of enzymes is by now well established, and the natural next step has already been taken for α-lytic protease. A tentative space-filling model of α-lytic protease has been built,[357] following the main-chain conformation of elastase and chymotrypsin. The case in favour of α-lytic protease being related to the others is very much strengthened by the finding that an almost identical structural 'core' can be built without any difficulty; insertions or deletions of whole loops occur at the surface. Only 3 of the 23 ionisable groups are found to be buried in the proposed model, and these are none other than Asp-102, Asp-194, and the *N*-terminal residue, which are of functional importance. 30% of the internal positions but only 10% of the surface residues are identical with the corresponding residues in the mammalian enzymes. We now await the results of *X*-ray work on α-lytic protease,[358] to see whether the usefulness of extrapolative model-building exercises [357] will again be proven.

The suggestion that α-lytic protease and the mammalian proteases derive from a common ancestral gene is not the only explanation, even if the three-dimensional structure of the bacterial enzyme is as predicted, but this idea is borne out by the discovery that another trypsin-like protease from *Streptomyces griseus* has sequences about its disulphide bridges that are homologous with those of trypsin.[356, 359] This suggests a class of bacterial serine proteases which share with the mammalian proteases a common ancestor that underwent gene-duplication and separate evolution very early on.

A trypsin-like protease of the crayfish *Astacus leptodactylus* [360] also has an -Asp-Ser-Gly- sequence, containing the serine easily labelled by di-isopropyl fluorophosphate (DFP). A protease that resembles trypsin in specificity has also been isolated from the starfish *Evastericas trochelii* [361] (although this differs from trypsin in not needing calcium ions for stability, and in having no zymogen) and from another species of starfish, *Der-*

[356] M. O. J. Olson, N. Nagabhushan, M. Dzwiniee, L. B. Smillie, and D. R. Whitaker, *Nature*, 1970, **228**, 438.
[357] A. D. McLachlan and D. M. Shotton, *Nature New Biology*, 1971, **229**, 202.
[358] M. N. G. James and L. B. Smillie, *Nature*, 1969, **224**, 694.
[359] L. Jurašek, D. Fackre, and L. B. Smillie, *Biochem. Biophys. Res. Comm.*, 1969, **37**, 99.
[360] V. Tomášek, F. Šorm, R. Zwilling, and G. Pfleiderer, *F.E.B.S. Letters*, 1970, **6**, 229.
[361] W. P. Winter and H. Neurath, *Biochemistry*, 1970, **9**, 4673.

α ALA-(Asn,Ile, **Val, Gly, Gly**————————————————————
E VAL-**Val**-**Gly**-**Gly**-Thr-Glu-Ala-Gln-Arg-Asn-Ser-Trp-Pro-
14

α Ile)-Glu-Tyr-**Ser**-Ile-Asn-(Asn————————————**Ala**)-Ser-
E Ser-Gln-Ile-**Ser**-Leu-Gln-Tyr-Arg-Ser-Gly-Ser-Ser-Trp-**Ala**-His-

α Leu-**Cys**-Ser-Val-Gly-Phe-Ser-Val-Thr-Arg-Gly-Ala-Thr-Lys-Gly-
E Thr-**Cys**-Gly-Gly-Thr-Leu-Ile-Arg————————Gly-Asn-Trp-

α Phe-Val-**Thr**-**Ala**-Gly-**HIS**-**Cys**-Gly-Thr-Val-Asn-Ala-**Thr**-(Ala———
E Val-Met-**Thr**-**Ala**-Ala-**HIS**-**Cys**-Val-Asp-Arg-Glu-Leu-**Thr**-Phe-Arg-
57

α ————————————————————————————————————
E Val-Val-Val-Gly-Glu-His-Asn-Leu-Asn-Gln-Asn-Asn-Gly-Thr-Glu-

α ————————Arg)-Ile-Gly-Gly-Ala-**Val**-**Val**-Gly———Thr-Phe-Ala-
E Gln-Tyr-Val-Gly-Val-Gln-Lys-Ile-**Val**-**Val**-His-Pro-Tyr-Trp-Asn-

α Ala-Arg-Val-Phe-Pro———**Gly**-Asn-**ASP**-Arg-**Ala**-Trp-Val-Ser-**Leu**-
E Thr-Asp-Asp-Val-Ala-Ala-**Gly**-Tyr-**ASP**-Ile-**Ala**-Leu-Leu-Arg-**Leu**-
102

α (Thr,Ser, Ala, Gln————————————————————Thr)-Leu-**Leu**-
E Ala-Gln-Ser-Val-Thr-Leu-Asn-Ser-Tyr-Val-Gln-Leu-Gly-Val-**Leu**-

α **Pro**-**Arg**-Val-Ala-Asn-Gly-(Ser———Ser————Phe)-Val-Thr-Val-
E **Pro**-**Arg**-Ala-Gly-Thr-Ile-Leu-Ala-Asn-Asn-Ser-Pro-Cys-Tyr-Ile-

α Arg-(**Gly**————————————————————————————
E Thr-**Gly**-Trp-Gly-Leu-Thr-Arg———Thr-Asn-Gly-Gln-Leu-Ala-Gln-

α ——Ser)-Thr-Glu-**Ala**-Ala-Val-Gly-Ala-(Ala————————Val)-Cys-
E Thr-Leu-Gln-Gln-**Ala**-Tyr-Leu-Pro-Thr-Val-Asp-Tyr-Ala-Ile-Cys-

α (Arg————————————**Ser**)-Gly-Arg-Thr-Thr-Gly-Tyr-Gln-
E Ser-Ser-Ser-Ser-Tyr-Trp-Gly-**Ser**-Thr-Val-Lys-Asn-Ser-Met-Val-

α **Cys**-Gly-Thr-Ile-Thr-Ala-(Lys,Asn,Val, Thr, Ala, Asn, Tyr, Ala,Glu,
E **Cys**-Ala-Gly-Gly-Asn————————————————————————

α Gly, Ala, Val, Arg, Gly, Leu, Thr, Gln, Gly)———Asn-Ala-**Cys**-Met-**Gly**-
E ————————————Gly-Val-Arg-Ser-Gly-**Cys**-Gln-**Gly**-

α (Arg-Gly)-**ASP**-**SER**-Ser-**Gly**-Ser-Trp-Ile-Thr-Ser-Ala-Gly-Gln-Ala-
E ————**ASP**-**SER**-Gly-**Gly**-Pro-Leu-His-Cys-Leu-Val-Asn-Gly-Gln-

α Gln-Gly-**Val**-Met-Ser-Gly-Gly-Asn-Val-Gln-**Ser**-Asn-Gly-Asn-Asn-
E Tyr-Ala-**Val**-His-Gly-Val-Thr-Ser-Phe-Val-**Ser**-Arg———Leu-Gly-

α **Cys**-Gly-Ile-Pro-Ala-Ser-Gln-Arg-Ser-Ser-Leu-**Phe**-Glu-**Arg**-Leu-
E **Cys**-Asn-Val————Thr-Arg-Lys-Pro-Thr-Val-**Phe**-Thr-**Arg**-Val-

α Gln-Pro-Ile-Leu-Ser-Gln-Tyr-Gly-Leu-Ser-Leu-Val-Thr-Gly
E Ser-Ala-Tyr-Ile———Ser-Trp-Ile-Asn-Asn-Val-Ile-Ala-Ser-Asn
245

Figure 8 *Comparison (ref. 357) of the amino-acid sequences of α-lytic protease (α) and elastase (E). The numbering scheme is that for chymotrypsinogen A. Catalytically important residues are in capitals*

masterias imbricata.[362] In the latter case two such proteases (mol. wt. 25 000—26 000) have been purified; they are inhibited by DFP but, strangely, not by soybean trypsin inhibitor, are activated spontaneously at 20 °C, and again are not dependent on calcium for stability. A study of the effect of calcium ions on trypsin revealed a temperature-dependent activation of the enzyme by calcium.[363] The calcium–trypsin complex is more active and has greater thermal stability than the enzyme alone. Inactive material present is degraded to small components.[363] It is of evolutionary interest that the enzymes of mammalian pancreas (proteinases or their zymogens, amylases, and nucleases) have their counterparts in the African lungfish *Protopterus aethiopicus*.[364]

The two chains of plasmin have mol. wts. 35 000 and 24 000; the *N*-terminal sequence of the light chain has been established [365] as Ile-Val-Gly-Gly-, and is homologous with the *N*-terminal sequence of trypsin, chymotrypsin, *etc.* Carefully purified plasminogen, which still contains components of different electrophoretic mobility, gave as its *N*-terminal sequence [366] Glu-Pro-Leu-Asp-Asp-Tyr-. The *N*-terminus of bovine prothrombin has been reported as threonine [367] in contradiction to earlier reports of *N*-terminal alanine. A report has appeared of the isolation of pig proelastase.[368] A comparison of the partial amino-acid sequences of dogfish and bovine trypsinogen reveals considerable homology,[369] as one might, by now, expect. A recent revision of an amide assignment in the sequence of bovine trypsinogen has been confirmed:[370] residue 177 now becomes Asp.

As chemical studies of the active site of trypsin and chymotrypsin continue, evidence in accord with the picture presented by *X*-ray crystallography accumulates. Confirmation that Asp-177 was at the ionic binding site that determines trypsin specificity was possible [371] by protecting it with the competitive inhibitor benzamidine while other accessible carboxy-groups were modified and then coupling with radioactive glycineamide, using a water-soluble carbodi-imide. Carboxy-groups at the binding site of trypsin were also modified [372] with triethyloxonium fluoroborate ($Et_3O^+BF_4^-$). Bender and his colleagues have reported some interesting chemical modifications of trypsin and chymotrypsin. The active-site reagent methyl *p*-nitrobenzenesulphonate, an analogue of non-specific

[362] Z. Camacho, J. R. Brown, and G. B. Kitto, *J. Biol.Chem.*, 1970, **245**, 3964.
[363] T. Sipos and J. R. Merkel, *Biochemistry*, 1970, **9**, 2766.
[364] G. R. Reeck, W. P. Winter, and H. Neurath, *Biochemistry*, 1970, **9**, 1398.
[365] S. Nagasawa and T. Suzuki, *Biochem. Biophys. Res. Comm.*, 1970, **3**, 562.
[366] P. Wallén and B. Wiman, *Biochim. Biophys. Acta*, 1970, **221**, 20.
[367] K. E. B. Platzer and H. A. Scheraga, *Biochim. Biophys. Acta*, 1970, **207**, 262.
[368] A. Gertler and Y. Birk, *European J. Biochem.*, 1970, **12**, 170.
[369] R. A. Bradshaw, H. Neurath, R. W. Tye, K. A. Walsh, and W. P. Winter, *Nature*, 1970, **226**, 237.
[370] V. Tomášek, J. Strmeň, and F. Šorm, *F.E.B.S. Letters*, 1970, **8**, 176.
[371] A. Eyl and T. Inagami, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 149.
[372] H. Nakayama, K. Tanizawa, and Y. Kanaoka, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 537.

substrates (*cf.* tosylphenylalanylchloromethyl ketone: analogous to specific substrates), resulted in specific methylation of His-57 of α-chymotrypsin; isolation of 3-methylhistidine confirmed that the nitrogen in position 3 of the imidazole ring is involved in catalysis.[98] Inhibition of trypsin with the irreversible inhibitor methyl *m*-guanidiniumbenzenesulphonate fluoroborate [373] was dependent on the basic form of an ionisable group with $pK_a \simeq 7$: it is likely, though not yet confirmed, that this is histidine at the active site. The reaction is entirely analogous with the inhibition of chymotrypsin mentioned above. Reagents of this type have the advantage [in contrast with tosyl-L-phenylalanyl chloromethyl ketone (TPCK) for chymotrypsin, and tosyl-L-lysine chloromethyl ketone (TLCK) for trypsin] that the bulky group is released after the inhibition reaction, thereby permitting loss of activity to be correlated with modification of an essential functional group rather than with steric hindrance of the active site. It is interesting that methyl *p*-guanidiniumbenzenesulphonate fluoroborate is a good *competitive* inhibitor of trypsin, thus emphasising that specificity of orientation is obviously necessary in order that irreversible inhibition may occur.[373] It has been reported [374] that the essentially irreversible inhibition of trypsin, and to successively lesser extents plasmin and thrombin, by benzyl 4-guanidinobenzoate and 4′-nitrobenzyl 4-guanidinobenzoate is due to acylation of the enzyme, followed by very slow deacylation. Urokinase is also irreversibly inhibited by the latter reagent, and by DFP, but not by TLCK.[375] It has been suggested that TLCK and nitrophenyl *p*-guanidinobenzoate may find some use as antifertility agents.[376]

Modification of Ile-16 in δ-chymotrypsin confirms the importance of the salt-bridge between Ile-16 and Asp-194 in maintaining the conformation of the active site: acetylation abolishes activity towards specific ester and amide substrates but still permits binding and hydrolysis of *p*-nitrophenyl acetate.[377] Spectral analysis of indoleacryloylchymotrypsin suggests that the catalytic mechanisms in solution and in the wet crystal are identical.[378] However, the accessibility of surface groups can apparently differ in the two states: iodination of α-chymotrypsin occurred preferentially at Tyr-171 in the crystal [379] whereas Tyr-146, known to be iodinated most readily in solution, became inaccessible. Its phenolic group was actually embedded in the dyad-related neighbour.[379] Evidence has been presented in support of an oxazoline intermediate in chymotrypsin hydrolysis of

[373] M. B. Jackson and M. L. Bender, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 1157.
[374] F. Markwardt, M. Richter, P. Walsmann, and H. Landmann, *F.E.B.S. Letters*, 1970, **8**, 170.
[375] H. Landmann and F. Markwardt, *Experientia*, 1970, **26**, 147.
[376] L. J. D. Zaneveld, R. T. Robertson, and W. L. Williams, *F.E.B.S. Letters*, 1970, **11**, 345.
[377] C. Ghélis, J.-R. Garel, and J. Labouesse, *Biochemistry*, 1970, **9**, 3902.
[378] G. L. Rossi and S. A. Bernhard, *J. Mol. Biol.*, 1970, **49**, 85.
[379] P. B. Sigler, *Biochemistry*, 1970, **9**, 3609.

specific substrates.[380] It lies on the path between the Michaelis complex and the acylenzyme and derives from intramolecular reaction of the substrate. The lesson is that functional groups responsible for enzyme catalysis may already be present in the substrate, in its specific orientation within the enzyme–substrate complex.

Immunochemical cross-reaction is widely used to investigate similarities between homologous proteins (see also Section 5B). The structural and chemical homology between chymotrypsin and trypsin was confirmed by cross-reaction of a fragment of chymotrypsin containing two disulphide bridges with anti-trypsin antiserum.[381] A difference in the accessibility of the disulphide bridges of chymotrypsin and chymotrypsinogen to reduction with 2-mercaptoethanol has been reported.[382]

Further progress has been made in Šorm's laboratory with the sequence of pig pepsin; they now report the *C*-terminal sequence.[383] Sequences have also appeared for the *C*-terminal regions of human pepsin and gastricsin,[384] thereby permitting comparison of these enzymes with pig pepsin:

Pig pepsin
-Tyr-Gly-Thr-Gly-Ser-Met-Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu-

Pig pepsin
-Leu-Trp-Ile **-Leu-Gly-Asp-Val-Phe-Ile -Arg-Gln**-Tyr-Tyr-Thr-
Human pepsin
-Ile **-Leu-Gly-Asp-Val-Phe-Ile -Arg-Gln**-Phe-Tyr-Thr-
Human gastricsin
-Gln-Phe-Tyr-Thr-

Pig pepsin
-Val-Phe-Asp-Arg-Ala-Asn-Asn-Lys-Val-**Gly-Leu-Ala-Pro-Val-Ala**
Human pepsin
-Val-Phe-Asp-Arg-Ala-Asn-Asn-Gln-Val-**Gly-Leu-Ala-Pro-Val-Ala**
Human gastricsin
-Val-Phe-Asp-Arg-Ala-Asn-Asn-Lys-Glu-**Gly-Leu-Ala-Pro-Val-Ala**

With the exception of three positions these sequences are identical in the 19 residues that can be compared, thereby providing further support for the existence of another 'family' of enzymes. Tang and Hartley[385] have used their diagonal techniques to advantage in establishing the sequences around the three disulphide bridges of pig pepsin, and around the four methionine residues; the latter are invaluable in ordering the fragments obtained from cleavage with cyanogen bromide. The homology with rennin was confirmed and extended. Homology between these two enzymes at the *N*-terminus was not immediately apparent when only the first five

[380] M.-A. Coletti-Previero, C. Axelrud-Cavadore, and A. Previero, *F.E.B.S. Letters*, 1970, **11**, 213, 218.
[381] M. M. Sanders, K. A. Walsh, and R. Arnon, *Biochemistry*, 1970, **9**, 2356.
[382] B. Mesrob and B. Keil, *F.E.B.S. Letters*, 1970, **6**, 17.
[383] V. Kostka, L. Moravek, and F. Šorm, *European J. Biochem.*, 1970, **13**, 447.
[384] W.-Y. Huang and J. Tang, *J. Biol. Chem.*, 1970, **245**, 2189.
[385] J. Tang and B. S. Hartley, *Biochem. J.*, 1970, **118**, 611.

residues of pepsin were known, but the sequence in this region has been extended[386] and the homology with rennin in this region too is now unambiguous:

Rennin	↓Gly-**Glu**-Val-Ala-Ser -Val-**Pro**-**Leu**-Thr-**Asn**-Tyr-	
	1	**10**
Pepsin	-Ala-**Glu**-Ile -Gly-Asp-Glu-**Pro**-**Leu**-Glu-**Asn**-Tyr-	
	↑**1**	**9**

The arrows indicate the points at which the zymogens are cleaved. Preliminary results on the sequence of the *N*-terminal region of pig pepsin have also come from another laboratory.[387]

Further details have been reported[388] of the reaction of pepsin with the diazo inactivator 1-diazo-4-phenyl-2-butanone. The inhibitor was found to be attached in ester linkage to the aspartyl side-chain in the sequence:

-Ile-Val-Asp-Thr-(Gly, Thr)-Ser-Leu-

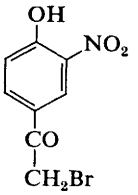
This is in accord with the results obtained by other workers[389] after labelling with *N*-diazoacetyl-L-phenylalanine methyl ester. The isolation of an enzymically active fragment from the autolysis of pepsin has been reported.[390]

Considerable heterogeneity of the peptic zymogens of human gastric mucosa has been reported,[391] and the warning sounded that previous work in this field, carried out on material assumed to be homogeneous, may need re-evaluation. Three pepsinogens (mol. wt. *ca.* 42 000) and the corresponding pepsins, isolated from chicken gastric mucosa, have been compared.[392] The A and D forms are very similar both as the zymogens and as the free enzymes. The enzymes have mol. wts. in the region of 42 000 and lose about 15 residues on activation. They contain over two-thirds of the basic amino-acids in the zymogens, a feature to which their stability at neutral pH has been attributed. Pepsin C, on the other hand, has a mol. wt. of 38 500 and the loss of peptide material in the activation step seems to resemble the generation of the pig enzyme from its zymogen.

Corrections to the sequence of papain have now been published,[393] bringing it into line with *X*-ray crystallographic results.[394] A reporter group may be attached to the active thiol group using the alkylating agent (5).[395] Like other thiol proteases, papain suffers from lack of a

[386] J. Tang, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 697.
[387] V. M. Stepanov, M. M. Amirkhanyan, B. G. Belen'kii, R. A. Valyulis, E. A. Vakhitova, I. B. Pugacheva, and L. G. Senyutenkova, *Biokhimyia* (*Eng.*), 1970, **35**, 230.
[388] K. T. Fry, O.-K. Kim, J. Spona, and G. A. Hamilton, *Biochemistry*, 1970, **9**, 4624.
[389] R. S. Bayliss, J. R. Knowles, and G. B. Wybrandt, *Biochem. J.*, 1969, **113**, 377.
[390] H. Determann and R. Kotitschke, *Z. physiol. Chem.*, 1970, **351**, 1169.
[391] M. D. Turner, J. C. Mangla, I. M. Samloff, L. L. Miller, and H. L. Segal, *Biochem. J.*, 1970, **116**, 397.
[392] S. T. Donta and H. Van Vunakis, *Biochemistry*, 1970, **9**, 2791, 2798.
[393] R. E. J. Mitchel, I. M. Chaiken, and E. L. Smith, *J. Biol. Chem.*, 1970, **245**, 3485.
[394] J. Drenth, J. N. Jansonius, R. Koekoek, L. A. A. Sluyterman, and B. G. Wolthers, *Phil. Trans.*, 1970, **B257**, 231.
[395] R. W. Furlanetto and E. T. Kaiser, *J. Amer. Chem. Soc.*, 1970, **92**, 6980.

satisfactory assay procedure. Those concerned with estimating the normality of the active site of this enzyme for accurate kinetic studies now have a choice of two new methods: Williams and Lucas [396] recommend the *p*-nitrophenyl ester of *N*-benzyloxycarbonyl-D-norleucine, which they say can



(5)

be used to estimate active site normalities accurately to less then 0.2 μmol l^{-1}. It fulfils the requirements for a good titrant in that deacylation of the inactivated enzyme occurs only very slowly, and permits titration of papain over a pH range including its pH optimum. Brocklehurst and Little [397] use an irreversible inhibitor of papain which they consider preferable. The reagent is 2,2′-dipyridyl disulphide which inhibits papain by disulphide exchange with the active site thiol group. Reaction of the disulphide reagent with papain is optimal at pH 3.75, and in acidic media is very much faster than with the denatured enzyme, L-cysteine, or 2-mercaptoethanol. This means that the titration can be carried out in the presence of other thiols (including denatured enzyme). A mechanism is suggested for the unusually low pH optimum of the reaction of the disulphide with papain: a conformational change resulting from hydrogen-bonding of one of the pyridine rings of the reagent with Asp-158 permits the formation of a strong hydrogen bond between the active site thiol and His-159, thereby generating nucleophilic character in this thiol group even at low pH. The 2-mercaptopyridine anion is then displaced. A speculative article [398] on the evolution of papain has looked for internal regions of homology in the amino-acid sequence of the two 'halves' of the molecule shown in the *X*-ray model.

Husain and Lowe have now extended their studies with dibromoacetone and have isolated the cross-linked peptides containing the active site cysteine and histidine residues of both ficin [399] and stem bromelain.[400] There is extensive homology [400] with the sequence of the similarly isolated regions of papain, and these three enzymes very clearly constitute yet another family:

[396] A. Williams and E. C. Lucas, *Analyt. Chem.*, 1970, **42**, 1491.
[397] K. Brocklehurst and G. Little, *F.E.B.S. Letters*, 1970, **9**, 113.
[398] B. Weinstein, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 441.
[399] S. S. Husain and G. Lowe, *Biochem. J.*, 1970, **117**, 333.
[400] S. S. Husain and G. Lowe, *Biochem. J.*, 1970, **117**, 341.

Papain
-**Pro**-Val-Lys-Asn-**Gln**-Gly-Ser -**Cys**-**Gly**-Ser -**CYS**-**Trp**-Val-**Gly**-**Pro** -**Cys**-
Ficin
-**Pro**-Ile -Arg-Gln-**Gln**-Gly-Gln-**Cys**-**Gly**-Ser -**CYS**-**Trp**-Thr-**Gly**-**Pro** -**Cys**-
Stem bromelain
-Asn-**Gln**-Asp-Pro -**Cys**-**Gly**-Ala-**CYS**-**Trp**-

Papain
Gly-Asn-Lys-Val -**Asp**-**HIS** -**Ala**-**Val**-Ala-Ala-Val-**Gly**-**Tyr**-
Ficin
Gly-Thr-Ser -Leu-**Asp**-**HIS** -**Ala** -**Val**-Ala-Leu-
Stem bromelain
-**HIS** -**Ala** -**Val**-Thr-Ala-Ile-**Gly**-**Tyr**-

Both ficin and papain have three disulphide bridges. However, the only free cysteine residue of papain is at its active site, while ficin has a second (buried) thiol. The amino-acid sequence adjoining this residue has been shown [399] to be homologous with the region around Ser-206 in papain:

Ficin	-Cys-Leu-**Tyr**-**Pro**-**Val**-**Lys**-
Papain	-Ser -Phe-**Tyr**-**Pro**-**Val**-**Lys**-
	206

Inspection of the *X*-ray model confirms that Ser-206 is not a 'surface' residue. Five ficins from *Ficius glabrata* latex have been purified.[401] They all have molecular weights of about 25 000—26 000 and *N*-terminal leucine. Preliminary peptide mapping indicates that they are homologous but different. The structure of the carbohydrate moiety of stem bromelain is known.[402]

An interesting method of zymogen activation is shown by the streptococcal proteinase under study at the Rockefeller Institute:[403] an active enzyme is generated by reduction of a mixed disulphide bond with a small molecule in the zymogen, without the need for proteolytic cleavage. One tryptophan residue in the enzyme (mol. wt. 44 000) reacts with 2-hydroxy-5-nitrobenzyl bromide, with loss of enzymic activity.[404] The amino-acid sequence around this tryptophan residue [404] shows some slight but interesting homology with a tryptophan-containing sequence in papain (mol. wt. 23 000):

Streptococcal proteinase
-His-Val -Asn-**Trp**-**Gly**-Phe-**Gly**-**TRP**-**Gly**-(Val, Ser) -Asn-Gly-Phe-**Arg**-

Papain
-Lys-Asn-Ser -**Trp**-**Gly**-Thr-**Gly**-**TRP**-**Gly**- Glu-Asn -Gly-Tyr-Ile -**Arg**-

[401] I. K. Jones and A. N. Glazer *J. Biol. Chem.*, 1970, **245**, 2765.
[402] Y. Yasuda, N. Takahashi, and T. Murachi, *Biochemistry*, 1970, **9**, 25.
[403] M. Bustin, M. C. Lin, W. H. Stein, and S. Moore, *J. Biol. Chem.*, 1970, **245**, 846.
[404] G. W. Robinson, *J. Biol. Chem.*, 1970, **245**, 4832.

There seem to have been as many publications this year relating to inhibitors of proteolytic enzymes as to the proteases themselves. A method of estimating the dissociation of the trypsin-inhibitor complex using an active site titrant for trypsin, *p*-nitrophenyl *p'*-guanidinobenzoate, has been described.[405] Almost 20 varieties of trypsin inhibitor have been studied by chemical modification with lysine- and arginine-blocking reagents (maleic anhydride and 2,3-butanedione respectively).[406] They fell clearly into two classes: those with anti-tryptic activity abolished by maleylation (and which, therefore, presumably had a lysine residue at the active site) and those that were inactivated by the diketone reagent (hence arginine at the active site). This was in accord with the observation [407] that all trypsin inhibitors have either an Arg-X bond or a Lys-X bond at the active site, inhibition occurring with cleavage of these bonds. Modification of the lysine residues of lima bean inhibitor abolished its anti-tryptic activity but not its anti-chymotryptic activity;[406] it has now been shown [408] that inhibition of chymotrypsin is associated with cleavage of a Leu-Ser bond in the sequence -Thr-Leu-Ser-Ile-Pro-, 29 residues from the *C*-terminus of the inhibitor. Chemical modification in the presence and absence of trypsin permitted identification of Lys-18 as the active-site lysine in the 67-residue inhibitor from cow colostrum.[409a] In this inhibitor too, the active sites for anti-tryptic and anti-chymotryptic activity were shown to be different.[409b] However, cocoonase (an exotic trypsin-like protease from the mouth parts of the silk moth) combines with soybean trypsin inhibitor at the anti-tryptic site, with cleavage of the same Arg(64)-Ile(65) bond.[410] It is of interest that at neutral pH both trypsin and cocoonase catalyse substantial formation of the Arg-Ile bond in an already cleaved inhibitor: evidently the equilibrium for this particular reaction is not enormously in favour of cleavage.[410] This is an unusual situation for the hydrolysis of a peptide bond. Some attention is being directed towards the identification of other groups in the combining sites of trypsin inhibitors.[411, 412] *N*-Acetylimidazole has been useful in showing that several amino-groups and tyrosyl phenolic groups from both the enzyme and the inhibitor are shielded in the trypsin–soybean inhibitor complex.[411]

Cleavage of soybean trypsin inhibitor with cyanogen bromide to yield two inactive fragments has already been mentioned [101] (Section 2D). More than 80% of the native anti-tryptic activity could be regenerated without formation of covalent bonds. A study of the regeneration of native

[405] J. C. Zahnley and J. G. Davis, *Biochemistry*, 1970, **9**, 1428.
[406] H. Fritz, E. Fink, M. Gebhardt, K. Hochstrasser, and E. Werle, *Z. physiol. Chem.*, 1969, **350**, 933.
[407] K. Ozawa and M. Laskowski, jun., *J. Biol. Chem.*, 1966, **241**, 3955.
[408] J. Krahn and F. C. Stevens, *Biochemistry*, 1970, **9**, 2646.
[409] (*a*) D. Cechova and G. Moszynska, *F.E.B.S. Letters*, 1970, **8**, 84. (*b*) G. Muszynska and D. Cechova, *F.E.B.S. Letters*, 1970, **8**, 274.
[410] H. F. Hixson, jun. and M. Laskowski, jun., *Biochemistry*, 1970, **9**, 166.
[411] S. E. Papaioannou and I. E. Liener, *J. Biol. Chem.*, 1970, **245**, 4931.
[412] M. J. Gorbunoff, *Biochim. Biophys. Acta*, 1970, **221**, 314.

structure and activity by air oxidation of reduced pancreatic trypsin inhibitor and des-Thr-Ser-Pro-Gln-Arg-inhibitor (*i.e.* inhibitor from which the first five residues had been removed) showed that the inhibitor fragment was, in fact, regenerated more rapidly than the intact inhibitor.[413] Trypsin inhibitors are known to be remarkably resistant to general proteolytic digestion; however, it has been possible [414] to achieve this using thermolysin at 60—80 °C. It is interesting that a trypsin inhibitor from bovine liver is identical with the inhibitor from pancreas.[415] Its presence in bovine lungs and parotid glands as well as liver led to the speculation that these organs might also produce trypsin, or trypsin-like enzymes.[415] Two trypsin inhibitors have been isolated from the body walls of *Ascaris lumbricoides* var. *suis* [416] and two kallikrein inhibitors from potatoes.[417] Chymotrypsin inhibitor from potatoes (known to have a mol. wt. of 39 000 and to bind four molecules of chymotrypsin per molecule of inhibitor) has been shown to have subunit structure despite its low mol. wt.[418] It is a tetramer but the identity, or otherwise, of its subunits has not yet been established. Dissociation of this inhibitor into monomers is accomplished by treatment with guanidine hydrochloride, and does not require the presence of chymotrypsin since covalent bond cleavage is not involved.

Primary structure has not been neglected. The sequences of trypsin inhibitors I and II from pig pancreas (56 and 52 residues, respectively) have been elucidated.[52, 419] Sequences have also been established for trypsin inhibitors from maize seeds [420] (65 residues) and from seeds of *Arachis hypogaea* [421] (48 residues).

B. Lysozyme and α-Lactalbumin.—Several points of interest relating to lysozyme have been reviewed.[422] The primary structure of duck egg-white lysozyme II, now complete,[423] shows 19 differences when compared with hen egg-white lysozyme. The duck enzyme contains no histidine, but all the tryptophan and cystine residues are conserved in identical positions in the sequence, as are the catalytically important residues Glu-35 and Asp-52. All the changes that do occur can be accommodated on the surface of the crystallographic model of hen egg lysozyme. The pK of Asp-52 has now been determined [424] and is reported as 4.5. Peptide mapping of lysozyme from turkey egg-white [425] suggests that it differs from the chicken enzyme

[413] H. Tschesche and H. Haenisch, *F.E.B.S. Letters*, 1970, **11**, 209.
[414] T.-W. Wang and B. Kassell, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 1039.
[415] J. Chauvet and R. Acher, *Internat. J. Protein Res.*, 1970, **2**, 165.
[416] U. Kucich and R. J. Peanasky, *Biochim. Biophys. Acta*, 1970, **200**, 47.
[417] H. Moriya, Y. Hojima, C. Moriwaki, and T. Tajima, *Experientia*, 1970, **26**, 720.
[418] J. C. Melville and C. A. Ryan, *Arch. Biochem. Biophys.*, 1970, **138**, 700.
[419] H. Tschesche and E. Wachter, *Z. physiol. Chem.*, 1970, **351**, 1449.
[420] K. Hochstrasser, K. Illchmann, and E. Werle, *Z. physiol. Chem.*, 1970, **351**, 721.
[421] K. Hochstrasser, K. Illchmann, E. Werle, R. Hössl, and S. Schwarz, *Z. physiol. Chem.*, 1970, **351**, 1503.
[422] P. Jollès, *Angew. Chem. Internat. Edn.*, 1964, **3**, 28.
[423] J. Hermann and J. Jollès, *Biochim. Biophys. Acta*, 1970, **200**, 178.
[424] S. M. Parsons and M. A. Raftery, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 45.
[425] J. N. LaRue and J. C. Speck, jun., *J. Biol. Chem.*, 1970, **245**, 1985.

in at least seven positions. The most interesting is the apparent replacement at position 101 of glycine for aspartic acid, the carboxy-group of which was implicated (by *X*-ray analysis) in substrate binding. The possibility of multiple genes for lysozyme in birds has been reported,[426] together with a plea for care in drawing inferences about the rate of protein evolution from changes in amino-acid sequence of the protein from various sources: multiple gene loci might mean that different loci were being expressed in different species. What prompted this warning was the finding of two lysozymes with different antigenic properties in the egg white of the black swan (*Cygnus atratus*), but availability of starting material was expected to hinder chemical characterisation of the two enzymes! In contrast, the multiple forms of duck lysozyme are almost certainly due to multiple alleles of a single locus.[427]

Further details have now appeared of the determination of the primary structure of T4 bacteriophage lysozyme,[428] which shows no apparent homology with the enzyme from egg-white. Amber mutation of each of the three tryptophan codons in the lysozyme gene of phage T4, followed by reversion to replace the amber mutation with a codon for tyrosine (*i.e.* UGG → UAG → UAU or UAC), results in an interesting lysozyme molecule in which the three tryptophans have been replaced by tyrosine.[429] The enzyme is still 50% active and is immunologically cross-reactive with native phage lysozyme. Other mutant lysozymes show that serine or glutamine (the other amino-acids that result from suppression of the amber mutation) can replace tryptophan at positions 126 and 158, but not at position 138: aromaticity is apparently essential at this position.[429] Replacement of tryptophan by tyrosine at position 138 is by itself sufficient to cause 50% loss in enzymic activity.

The complete amino-acid sequence of α-lactalbumin is now available.[430] This confirms the homology with hen egg-white lysozyme that was already suspected, and which had led to the fitting of the available α-lactalbumin sequence to the model of lysozyme obtained by *X*-ray crystallography.[431] Of the 123 residues in α-lactalbumin, 49 are identical with the corresponding positions in lysozyme and there are 23 conservative replacements. The four disulphide bridges are arranged identically.[432] Until the actual *X*-ray structure becomes available, comparisons of chemical reactivity are useful in probing the similarities and differences in the native structures of lysozyme and α-lactalbumin. Carboxymethylation of ox α-lactalbumin[433] shows

[426] N. Arnheim and R. Steller, *Arch. Biochem. Biophys.*, 1970, **141**, 656.
[427] E. Prager and A. C. Wilson, *J. Biol. Chem.*, 1971, **246**, 523.
[428] M. Inouyé, M. Imada, and A. Tsugita, *J. Biol. Chem.*, 1970, **245**, 3479.
[429] M. Inouyé, E. Akaboshi, M. Kuroda, and A. Tsugita, *J. Mol. Biol.*, 1970, **50**, 71.
[430] K. Brew, F. J. Castellino, T. C. Vanaman, and R. L. Hill, *J. Biol. Chem.*, 1970, **245**, 4570.
[431] W. J. Browne, A. T. C. North, D. C. Phillips, K. Brew, T. C. Vanaman, and R. L. Hill, *J. Mol. Biol.*, 1969, **42**, 65.
[432] T. C. Vanaman, K. Brew, and R. L. Hill, *J. Biol. Chem.*, 1970, **245**, 4583.
[433] F. J. Castellino and R. L. Hill, *J. Biol. Chem.*, 1970, **245**, 417.

that the order of reactivity of the residues alkylated is Met-90 > His-68 > His 32 ≫ His 107, which is in general accord with the predicted conformation.[431] The six tryptophan residues of lysozyme react with 2-hydroxy-5-nitrobenzyl bromide in the manner expected from the known three-dimensional structure.[434] However, although three out of the four tryptophan residues of α-lactalbumin (residues 63, 108, and 123) behave in the manner dictated by the proposed structure, Trp-28 reacts with the reagent when it would appear that it should be buried.[434] Certain differences were found in the reactivities of the carboxy-groups in lysozyme and α-lactalbumin when they were coupled with glycineamide using a soluble carbodi-imide.[435] None of the carboxy-groups in α-lactalbumin had the reactivity of the catalytically important Glu-35 of lysozyme; in general, all carboxy-groups in α-lactalbumin appear to be more accessible than those of lysozyme. The proposal [436] that lysozyme and avidin may be genetically related in the same way as are lysozyme and α-lactalbumin is by no means ruled out yet but it is not apparent from the regions of the primary structure so far defined;[102a] *i.e.* the *N*-terminal 26 and the *C*-terminal 33 residues. The completion of the avidin sequence is awaited with interest. The carbohydrate moiety is attached to Asn-17 in the sequence Asn-Met-Thr, -Asn-X-Thr/Ser- being a common point of carbohydrate attachment (see Section 6); this is the same Met-Thr bond that is not wholly cleaved by cyanogen bromide (see Section 2D).

Immunological cross-reaction as a test for common features in protein structure can sometimes be misleading. It has been suggested [437] that the lack of cross-reaction between lysozyme and α-lactalbumin in precipitin analysis is due to marked conformational changes which are reflected in differences in susceptibility of the disulphide bonds to reduction. However, these results would seem to be equally explicable in terms of small local changes of conformation. For example, the same group of workers [438] demonstrate antigenic differences for a series of myoglobins, where it seems most unlikely that there will be major differences in main-chain conformation. In another study,[439] no cross-reaction was found between ox α-lactalbumin and any of the six lysozymes tested, suggesting either that the three-dimensional structures are different or that the amino-acid sequence changes form part of and, therefore, alter the antigenic determinants. Strosberg and his colleagues, however, were able to demonstrate [440] a low but reproducible inhibition by α-lactalbumin of the precipitation of lysozyme with anti-lysozyme antibodies. But they, too, point out

[434] T. E. Barman, *J. Mol. Biol.*, 1970, **52**, 391.
[435] T.-Y. Lin, *Biochemistry*, 1970, **9**, 984.
[436] N. M. Green, *Nature*, 1968, **217**, 254.
[437] M. Z. Atassi, A. F. S. A. Habeeb, and L. Rydstedt, *Biochim. Biophys. Acta*, 1970, **200**, 184.
[438] M. Z. Atassi, D. P. Tarlowski, and J. H. Paull, *Biochim. Biophys. Acta*, 1970, **221**, 623.
[439] A. Faure and P. Jollès, *F.E.B.S. Letters*, 1970, **10**, 237.
[440] A. D. Strosberg, C. Nihoul-Deconinck, and L. Kanarek, *Nature*, 1970, **227**, 1241.

that immunological results can disguise many homologies between proteins. In particular, the evolutionarily constant part of the molecule will, by virtue of its constancy, not evoke the production of antibodies in a foreign species. Conversely, antigenic determinants, since they are generally composed of surface residues, occur in regions of a protein that can usually tolerate extensive change in amino-acid sequence without major change in protein conformation. Perhaps studies of antibodies directed against particular peptides would circumvent this.[439]

An elegant immunological comparison of bird and human lysozymes and their 'loop' peptides takes up this point.[441] Lysozyme (or loop peptide) was chemically linked to bacteriophage T4 [442] and inactivation of the modified phage by antiprotein antibodies was tested, substantially increasing the sensitivity of the system. It was thus possible to decide whether the loop peptide was expressed immunologically.

The persistent heterogeneity of highly purified α-lactalbumin from bovine milk has now been explained as being partly due to the presence of glyco-α-lactalbumin,[443] which differs from the normal protein in having 11—12 sugar residues per molecule. Two α-lactalbumins have been isolated from Droughtmaster cattle:[444] they differ in the replacement of a glutamine in the A form by an arginine in the B form.

C. Ribonucleases.—A simple but highly efficient purification of ribonuclease T_1 has been published.[445] The amino-acid sequence of pig pancreatic ribonuclease has been shown [446] to be sufficiently similar to that of the ox to expect that the folding is very similar too. This enzyme, unlike that of ox, is a glycoprotein and two glycopeptides have been obtained [447] from proteolytic digests of the enzyme. These contain carbohydrate bound to the amide group of asparagine in the sequences -Arg-Asn-(Met, Thr, Glu, Gly)-Arg and -Ser-Asn-Ser-Thr-. Whale pancreatic ribonuclease W_2 also contains carbohydrate, this time attached to an asparagine residue in the sequence -Asn-(Thr, Ser)-.[448] Some preliminary results on the inactivation of whale pancreatic ribonuclease W_1 with iodoacetate have been given,[449] and a detailed analysis of the reaction with histidine, methionine, and lysine

[441] E. Maron, R. Arnon, M. Sela, J.-P. Perin, and P. Jollès, *Biochim. Biophys. Acta*, 1970, **214**, 222.
[442] J. Haimovich, E. Hurwitz, N. Novik, and M. Sela, *Biochim. Biophys. Acta*, 1970, **207**, 115.
[443] T. E. Barman, *Biochim. Biophys. Acta*, 1970, **214**, 242.
[444] K. Bell, K. E. Hopper, H. A. McKenzie, W. H. Murphy, and D. C. Shaw, *Biochim. Biophys. Acta*, 1970, **214**, 437.
[445] R. Fields, H. B. F. Dixon, G. R. Law, and C. Yui, *Biochem. J.*, 1971, **121**, 591.
[446] R. L. Jackson and C. H. W. Hirs, *J. Biol. Chem.*, 1970, **245**, 637; J. J. Phelan and C. H. W. Hirs, *J. Biol. Chem.*, 1970, **245**, 654.
[447] T. Tsuruo, S. Yamashita, T. Terao, and T. Ukita, *Biochim. Biophys. Acta*, 1970, **200**, 544.
[448] T. Tsuruo, K. Sudo, T. Terao, and T. Ukita, *Biochim. Biophys. Acta*, 1970, **200**, 560.
[449] T. Terao, S. Yamashita, and T. Ukita, *Biochim. Biophys. Acta*, 1970, **198**, 45.

residues in ox ribonuclease A has been reported.[450] Much evidence already exists that Lys-41 is at the active site of the latter enzyme. A study of the reactivity of the lysine residues towards 1-guanidino-3,5-dimethylpyrazole nitrate (GDMP) confirms this [451] and shows that the protein in which only Lys-41 is left unmodified is enzymically active. Moreover, the modified enzyme reacted with bromoacetate in exactly the same way as the native enzyme. So, too, did the fully guanidinated but inactive enzyme, thereby separating the special reactivity of the histidine residues from the catalytic activity. The ribonuclease from *Bacillus amyloliquefaciens* has the *C*-terminal sequence -Thr-Thr-Asp-His-Tyr-Gln-Thr-Phe-Thr-Lys-Ile-Arg;[452] removal of up to three residues by carboxypeptidase treatment leaves substantial enzymic activity, but after removal of five residues the enzyme is inactive.

The ribonucleases are proving admirable proteins for studies on protein folding and enzyme activity. The interactions of various segments of the protein molecule in ribonuclease A [453] and staphylococcal nuclease [454, 455] can be studied by complementing one fragment (synthetic or natural) with another, to produce enzymically active aggregates. The kinetics of folding of staphylococcal nuclease have been investigated by following the fluorescence of the single tryptophan residue per molecule.[456] This can be approximated by two sequential first-order rate processes, with half-times of about 55 ms and 350 ms, respectively. These are several orders of magnitude slower than the conformational isomerisation of enzymes during catalysis and the helix-coil transition. Anfinsen showed some years ago that limited peptic digestion of ox pancreatic ribonuclease A removes the *C*-terminal four residues (-Asp-Ala-Ser-Val). A recent study shows [457] that the modified enzyme is, in fact, 0.5% active towards 2′,3′-cytidylic acid and that the inhibitor, 2′-cytidylic acid, is still strongly bound. Moreover, the special reactivity of the active-site histidine residues towards iodoacetate is preserved. However, removal of the new *C*-terminal amino-acid Phe-120 with carboxypeptidase destroys all these properties and it seems reasonable to conclude that this residue, which normally fits into a hydrophobic pocket in the molecule, plays an important part in the maintenance of the native enzyme conformation. Other experiments [458a] suggest that the same *C*-terminal four residues contain information critical to the correct folding of the protein. Thus, if the disulphide bridges of the pepsin-modified enzyme are reduced and then re-oxidised, they re-form in

[450] H. J. Goren and E. A. Barnard, *Biochemistry*, 1970, **9**, 959, 974.
[451] D. M. Glick and E. A. Barnard, *Biochim. Biophys. Acta*, 1970, **214**, 326.
[452] R. W. Hartley, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 263.
[453] M. C. Lin, B. Gutte, S. Moore, and R. B. Merrifield, *J. Biol. Chem.*, 1970, **245**, 5169.
[454] I. M. Chaiken and C. B. Anfinsen, *J. Biol. Chem.*, 1970, **245**, 2337.
[455] I. M. Chaiken and C. B. Anfinsen, *J. Biol. Chem.*, 1970, **245**, 4718.
[456] A. N. Schechter, R. F. Chen, and C. B. Anfinsen, *Science*, 1970, **167**, 886.
[457] M. C. Lin, *J. Biol. Chem.*, 1970, **245**, 6726.
[458a] H. Taniuchi, *J. Biol. Chem.*, 1970, **245**, 5459.

a random manner, which is not true for the same treatment of the intact enzyme. A similar effect is observed for staphylococcal nuclease (149 residues and no disulphide bridges), where residues 1—126 cannot fold correctly from denaturing solvents.[458b] Similarly, the *N*-terminal region of ox ribonuclease A also carries important folding information:[459] the S-protein will not re-fold properly except in the presence of the S-peptide (residues 1—20) to which, of course, it is not covalently linked. Moreover, the microsomal sulphydryl–disulphide interchange enzyme ('re-arrangease') will scramble the disulphide bridges of native S-protein in the absence of S-peptide and re-form them correctly in its presence. Studies of the synthetic S-protein and its derivatives[460] show that the *N*-terminal four residues of S-protein are not essential for these interactions. Comparable experiments with other proteins are in accord with these sorts of observations. The zymogen of streptococcal proteinase, a single chain with no cross-links, can be reversibly unfolded in guanidine hydrochloride, but proteolytic removal of 100 residues from the *N*-terminus leads, not surprisingly perhaps, to irreversible denaturation under these conditions.[461] It is now well known that while insulin itself cannot be reversibly reduced and re-oxidised in high yield, its biosynthetic precursor proinsulin undergoes this perfectly easily.[289] Recent experiments[462] show that 're-arrangease' correctly re-forms the disulphide bridges of scrambled insulin only in the presence of the connecting peptide (position 5 is Glu not Gln; see p. 71) that originally linked the A and B chains in proinsulin. These experiments on a variety of proteins would lend support, therefore, to the idea that, for these proteins at least, a majority of the primary structure is required to determine folding. Proteins (not to mention protein chemists) being what they are, exceptions are bound to follow. Promises and pie-crust are made to be broken.[463]

D. Dehydrogenases.—The proceedings of the Konstanz symposium on dehydrogenases have now been published.[464] The amino-acid sequence of the ethanol-active (E) form of horse-liver alcohol dehydrogenases has been completed[465] and has confirmed a dimer of identical chains, each containing 374 residues. Slight homology with glyceraldehyde 3-phosphate dehydrogenase is detected in the *N*-terminal region. The sequence changes in the steroid-active (S) isoenzyme have also been established[466] and found to occur mostly in the *N*-terminal half of the molecule:

[458b] H. Taniuchi and C. B. Anfinsen, *J. Biol. Chem.*, 1969, **244**, 3864.
[459] I. Kato and C. B. Anfinsen, *J. Biol. Chem.*, 1969, **244**, 1004.
[460] B. Gutte and R. B. Merrifield, *Fed. Proc.*, 1970, **29**, 727.
[461] M. C. Lin and M. Bustin, *J. Biol. Chem.*, 1970, **245**, 3384.
[462] P. T. Varandani and M. A. Nafz, *Arch. Biochem. Biophys.*, 1970, **141**, 533.
[463] Jonathan Swift, *Polite Conversation.*
[464] 'Pyridine Nucleotide Dependent Dehydrogenases,' ed. H. Sund, Springer–Verlag, Berlin, Heidelberg, and New York, 1970.
[465] H. Jörvnall, *European J. Biochem.*, 1970, **16**, 25.
[466] H. Jörnvall, *European J. Biochem.*, 1970, **16**, 41.

Position	*E*	*S*
17	Glu	Gln
94	Thr	Ile
101	Arg	Ser
110	Phe	Leu
115	Asp	Ser (tentative)
366	Glu	Lys

this should be of considerable interest when the crystal structure of alcohol dehydrogenase is available.

The longest dehydrogenase sequence to be published to date (506 residues) is that of ox-liver glutamate dehydrogenase (GDH) (Figure 9).[467] Homology with alcohol dehydrogenase is not obvious and that with glyceraldehyde 3-phosphate dehydrogenase (GPDH) appears to be limited to a 12-residue sequence:

GDH	-Gly-Gly-Ala-Lys-Ala-Gly-Val-Lys-Ile -Asn-Pro-Lys-	
	94	**105**
GPDH	-Gly-Ala-Ala-Lys-Ala-Val-Gly-Lys-Val-Ile -Pro-Glu-	
	209	**220**

This region includes Lys-97 of GDH which is the specific site of attachment of the inhibitor pyridoxal phosphate[468] and is probably the same residue that reacts with another inhibitor, *N*-(*N'*-acetyl-4-sulphamoylphenyl)-maleimide,[469] although the sequences reported differ slightly. Nitration of Tyr-412 leads to loss of allosteric inhibition by GTP.[467] The identity of the subunits of ox liver GDH has also been suggested by *N*- and *C*-terminal group analysis.[470] Whether or not the different dehydrogenases have elements of structure in common must, for the moment, remain an open question, particularly in the absence of three-dimensional structures from *X*-ray crystallography. Only the crystal structure of dogfish lactate dehydrogenase is available at high resolution[471] and, regrettably, little amino-acid sequence analysis of that protein has been published. The *N*-terminal sequence of the enzyme, however, has been shown to be *N*-acetyl-Thr-Ala-Leu-[472] and that, in more ways than one, is a start. The *N*-terminal sequence of the corresponding enzyme from rat liver is *N*-acetyl-Ala-Ala-.[473]

It has been proposed[474] that the malate dehydrogenase of rat liver mitochondria comprises two non-identical subunits of mol. wt. 32 000 and

[467] E. L. Smith, M. Landon, D. Piszkiewkz, W. J. Brattin, T. J. Langley, and M. D. Melamed, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 724.
[468] D. Piszkiewicz, M. Landon, and E. L. Smith, *J. Biol. Chem.*, 1970, **245**, 2622.
[469] J. J. Holbrook and R. Jeckel, *Biochem. J.*, 1969, **111**, 689.
[470] M. Pagé and C. Godin, *Biochim. Biophys. Acta*, 1970, **200**, 598.
[471] M. J. Adams, G. C. Ford, R. Koekoek, P. J. Lentz, jun., R. W. Schevitz, and A. J. Wonacott, *Nature*, 1970, **227**, 1098.
[472] W. S. Allison, J. Admiraal, and N. O. Kaplan, *J. Biol. Chem.*, 1969, **244**, 4743.
[473] B. M. Sanborn, M. C. Brummel, L. D. Stegink, and C. S. Vestling, *Biochim. Biophys. Acta*, 1970, **221**, 125.
[474] K. G. Mann and C. S. Vestling, *Biochemistry*, 1970, **9**, 3020.

Ala-Asp-Arg-Glu-Asp-Asp-Pro-Asn-Phe-Phe-Lys-Met-Val-Glu-Gly-Phe-Phe-
1 **10**

Asp-Arg-Gly-Ala-Ser-Ile-Val-Glu-Asp-Lys-Leu-Val-Glu-Asp-Leu-Lys-Thr-Arg-
20 **30**

Gln-Thr-Gln-Glu-Gln-Lys-Arg-Asn-Arg-Val-Arg-Gly-Ile-Leu-Arg-Ala-Gln-(His,
40 **50**

Ser)-His-Gln-Arg-Thr-Pro-Cys-Lys-Gly-Gly-Ile-Arg-Tyr-Ser-Thr-Asp-Val-Ser-
60 **70**

Val-Asp-Glu-Val-Lys-Ala-Leu-Ala-Ser-Leu-Met-Thr-Tyr-Lys-Cys-Ala-Val-Val-
80

Asp-Val-Pro-Phe-Gly-Gly-Ala-Lys-Ala-Gly-Val-Lys-Ile-Asn-Pro-Lys-Asn-Tyr-
90 * **100**

Thr-Asp-Glu-Asp-Leu-Glu-Lys-Ile-Thr-Arg-Thr-Arg-Phe-Met-Glu-(Leu, Thr,
110 **120**

Thr, Ala)-Met-Glu-Leu-Ala-Lys-Lys-Gly-Phe-Ile-Gly-Pro-Gly-Val-Asp-Val-Pro-
130 **140**

Ala-Pro-Asn-Met-Ser-Thr-Gly-Glu-Arg-Glu-Met-Ser-Trp-Ile-Ala-Asp-Thr-Tyr-
150 **160**

Ala-Ser-Thr-Ile-Gly-His-Tyr-Asp-Ile-Asn-Ala-His-Ala-Cys-Val-Thr-Lys-Pro-
170

Gly-Ile-Ser-Gln-Gly-Gly-Ile-His-Gly-Arg-Ile-Ser-Ala-Thr-Gly-Arg-Gly-Val-Phe-
180 **190**

Gly-His-Ile-(Glu, Asn)-Phe-Ile-Glu-Asn-Ala-Ser-Tyr-Met-Ser-Ile-Leu-Gly-Met-
200 **210**

Thr-Pro-Gly-Phe-Gly-Asp-Lys-Thr-Phe-Ala-Val-Gln-Gly-Phe-Gly-Asn-Val-Gly-
220 **230**

Leu-His-Ser-Met-Arg-Tyr-Leu-His-Arg-Phe-Gly-Ala-Lys-Cys-Val-Ala-Val-Gly-
240 **250**

Glu-Ser-Asp-Gly-Ser-Ile-Trp-Asn-Pro-Asp-Gly-Ile-Asp-Pro-Lys-Glu-Leu-Glu-
260

Asp-Phe-Lys-Leu-Gln-(His, Gly)-Thr-Ile-Leu-Gly-Phe-Pro-Lys-Ala-Lys-Ile-Tyr-
270 **280**

Glu-Gly-Ser-Ile-Leu-Glu-Val-Asp-Cys-Asp-Ile-Leu-Ile-Pro-Ala-Ala-Ser-Glu-
290 **300**

Lys-Gln-Leu-Thr-Lys-Ser-Asn-Ala-Pro-Arg-Val-Lys-Ala-Lys-Ile-Ile-Ala-Glu-
310 **320**

Gly-Ala-Asn-Gly-Pro-Thr-Thr-Pro-Glx-Ala-Asp-Lys-Ile-Phe-Leu-Glu-Arg-Ile-
330 340

Ile-Lys-Pro-Cys-Asn-His-Val-Leu-Ser-Leu-Ser-Phe-Pro-Ile-Arg-Arg-Asp-Asp-
350

Gly-Ser-Trp-Glu-Val-Ile-Glu-Gly-Tyr-Arg-(Ile, Glx)-Met-Val-Ile-Pro-Asp-Leu-
360 370

Tyr-Leu-Asn-Ala-Gly-Gly-Val-Thr-Val-Ser-Tyr-Phe-Glx-Leu-Lys-Asn-Leu-Asn-
380 390

His-Val-Ser-Tyr-Gly-Arg-Leu-Thr-Phe-Lys-Tyr-Glu-Arg-Asp-Ser-Asn-Tyr-His-
400 410 *

Leu-Leu-Met-Ser-Val-Gln-Glu-Ser-Leu-Glu-Arg-Lys-Phe-Gly-Lys-His-Gly-Gly-
420 430

Thr-Ile-Pro-Ile-Val-Pro-Thr-Ala-Glu-Phe-Gln-Asp-Arg-Ile-Ser-Gly-Ala-Ser-Glu-
440 450

Lys-Asp-Ile-Val-His-Ser-Gly-Leu-Ala-Tyr-Thr-Met-Glu-Arg-Ser-Ala-Arg-Gln-
460

Ile-Met-Arg-Thr-Ala-Met-Lys-Tyr-Asn-Leu-Gly-Leu-Asp-Leu-Arg-Thr-Ala-Ala-
470 480

Tyr-Val-Asn-Ala-Ile-Glu-Lys-Val-Phe-Arg-Val-Tyr-Asn-Glu-Ala-Gly-Val-Thr-
490 500

Phe-Thr
506

Figure 9 *Amino-acid sequences of ox liver glutamate dehydrogenase*

that hybrids of these (XX, XY, YY), not conformational isomers,[475] are the explanation of isoenzymes. The malate dehydrogenase from pig heart mitochondria is inhibited by iodoacetamide,[476] through alkylation of a single histidine residue per subunit at the N-3 position of the imidazole ring. The enzyme is not inactivated by iodoacetic acid[477] while NADH as well as abortive ternary complex formation with oxaloacetate and NAD protect against inhibition by iodoacetamide. Modification of histidine residues by photo-oxidation or with diazo-1-*H*-tetrazole has suggested their presence also in the active site of α-glycerophosphate dehydrogenase.[478]

The *N*-terminal residues of glucose 6-phosphate dehydrogenase from *Candida utilis* have been shown to be glycine and alanine:[479] the *C*-terminal

[475] G. B. Kitto, P. M. Wassarman, and N. O. Kaplan, *Proc. Nat. Acad. Sci. U.S.A.*, 1966, **56**, 578.
[476] B. H. Anderton, *European J. Biochem.*, 1970, **15**, 562.
[477] B. H. Anderton and B. R. Rabin, *European J. Biochem.*, 1970, **15**, 568.
[478] R. Apitz-Castro and Z. Suarez, *Biochim. Biophys. Acta*, 1970, **198**, 176.
[479] G. F. Domagk, W. Domschke, J. Meyer, and M. Weise, *Z. physiol. Chem.*, 1970, **351**, 923.

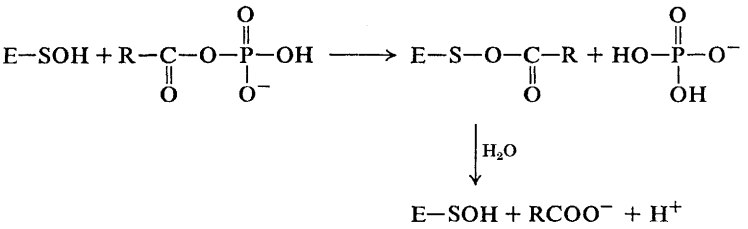
residue is alanine. SDS-Gel electrophoresis also suggests [479] the presence of different subunits of mol. wt. 10 000 and 14 000. It has been reported [480] that the FAD at the active site of succinate dehydrogenase is attached to histidine in the sequence -Ser-His-Thr-Val-Ala-, and the amino-acid sequence around the active centre disulphide bridge of thioredoxin reductase from *E. coli* has been shown [481] to be -Ala-Cys-Ala-Thr-Cys-Asp-Gly-Phe-. The thioredoxin induced in *E. coli* B by bacteriophage T4 differs from that of the uninfected cell, *e.g.* around the active site disulphide bridge:[482]

E. coli -Trp-Cys-Gly-Pro-Cys-Lys-Met-Ile-

E. coli/T4 -Cys-Tyr-Val-Cys-Asx-Asx-Ala-Lys-

It is interesting that the two thioredoxins and the reductase all contain a disulphide loop of identical size in their active sites, although there is no obvious detailed homology.

It has been known for some years that oxidation of rabbit muscle glyceraldehyde 3-phosphate dehydrogenase with iodine or iodosobenzoate inactivates the dehydrogenase activity and stimulates the acyl phosphatase activity. The oxidation was shown last year [483] to take place *via* the conversion of the catalytic cysteine residue, Cys-149, to the corresponding sulphenic acid (*cf.* creatine kinase [484]). It is now proposed [485] that the stabilised sulphenic acid residue participates directly in the hydrolysis of acyl phosphates (Scheme 7):



Scheme 7

E. Other Enzymes.—In view of the efforts to elucidate both the primary and three-dimensional structure of triosephosphate isomerase,[486] it is of some interest that the three electrophoretic forms of the enzyme from

[480] W. C. Kenney, W. H. Walker, E. B. Kearney, E. Zeszotek, and T. P. Singer, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 488.
[481] L. Thelander, *J. Biol. Chem.*, 1970, **245**, 6026.
[482] O. Berglund and A. Holmgren, *J. Biol. Chem.*, 1970, **245**, 6036.
[483] D. J. Parker and W. S. Allison, *J. Biol. Chem.*, 1969, **244**, 180.
[484] D. Trundle and L. W. Cunningham, *Biochemistry*, 1969, **8**, 1919.
[485] W. S. Allison and M. J. Connors, *Arch. Biochem. Biophys.*, 1970, **136**, 383.
[486] D. C. Phillips, *Abs. Eighth Internat. Congress of Biochemistry*, Switzerland, 1970, p. 116; R. E. Offord, A. J. Furth, and J. D. Priddle, *ibid.*, p. 116; J. C. MacGregor and S. G. Waley, *ibid.*, p. 116.

rabbit muscle (and perhaps from other sources) may be due to isoenzymes,[487] hence laying the conformer ghost. The enzyme from both rabbit muscle and rabbit liver has *N*-terminal alanine, while that from yeast shows two *N*-terminal groups, valine and alanine.[487] Hybridisation experiments and peptide mapping of the enzyme from rabbit muscle suggest[487] that two non-identical subunits (the products of two different structural genes) give rise to three isoenzymes of the form AA, BB, and AB.

The active centre of triose phosphate isomerase contains a glutamic acid residue which several workers have made a point of attack with suitably conceived reagents[488, 489, 490] (see also Section 8). The two sequences advanced for this region of the rabbit muscle enzyme are at slight variance:

-Ala-Tyr-Glu-Pro-Val-Trp-[488]
-Trp-Val-Leu-Ala-Tyr-Glu-Pro-Val-Ala-Trp-(Thr, Ile)-Gly-Gly-Lys-[489]

This minor discrepancy should soon be resolved: isolation of a peptide from the chicken muscle enzyme[490] with sequence identical with that in reference 488 would seem to favour that alternative.

The relative intensity of the two bands obtained for *S*-carboxymethylated rabbit muscle aldolase on gel electrophoresis in urea varies with the age of the rabbit;[491] a single band is observed for very young animals and two for mature rabbits. The change in mobility is consistent with deamidation of some asparaginyl or glutaminyl side-chains, and it is possible that this is the answer to the behaviour on gels. A difference in amide content of the two chains has, in fact, been reported[492] in the *C*-terminal hexapeptide:

α- -Ile-Ser-Asn-His-Ala-Tyr
β- -Ile-Ser-Asp-His-Ala-Tyr

No evidence is offered for the preferential deamidation in one chain, and the alternative explanation, the existence of two structural genes for muscle aldolase, cannot be ruled out. However, a study of the number, location, and reactivity of the thiol groups in the enzyme[349, 493, 684] shows that the two types of polypeptide chain must at least be very similar.

A recent redetermination[494] of the sequence in the region of the substrate-binding lysine residue resulted in the inversion of the previous[495] -Asn-Pro-

[487] W. K. G. Krietsch, P. G. Pentcher, H. Klingenbürg, T. Hofstatter, and T. Bücher, *European J. Biochem.*, 1970, **14**, 289; W. K. G. Krietsch, P. G. Pentchev, W. Machleidt, and H. Klingenbürg, *F.E.B.S. Letters*, 1970, **11**, 137.
[488] S. G. Waley, J. C. Miller, I. A. Rose, and E. L. O'Connell, *Nature*, 1970, **227**, 181.
[489] F. C. Hartman, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 384.
[490] A. F. W. Coulson, J. R. Knowles, J. D. Priddle, and R. E. Offord, *Nature*, 1970, **227**, 180.
[491] M. Koida, C. Y. Lai, and B. L. Horecker, *Arch. Biochem. Biophys.*, 1969, **134**, 623.
[492] C. Y. Lai, C. Chen, and B. L. Horecker, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 461.
[493] B. Szajani, M. Sajgó, E. Biszku, P. Friedrich, and G. Szabolcsi, *European J. Biochem.*, 1970, **15**, 171; H. M. Steinman, and F. M. Richards, *Biochemistry*, 1970, **9**, 4360.
[494] I. Gibbons, P. J. Anderson, and R. N. Perham, *F.E.B.S. Letters*, 1970, **10**, 49.
[495] C. Y. Lai, P. Hoffee, and B. L. Horecker, *Arch. Biochem. Biophys.*, 1965, **112**, 567; C. Y. Lai and C. Chen, *ibid.*, 1968, **128**, 212.

sequence on the *C*-terminal side of the lysine (Figure 10, centre line) and alteration of the amide assignments at four positions (4, 10, 32, and 33). Determination of the amino-acid sequence of the corresponding region in aldolase from sturgeon muscle [494] revealed a surprisingly high degree of homology, reminiscent of the homologies observed for another glycolytic

Rabbit liver	Ala	-Leu	-Asn	-Asn	-His	-His	-Val	-Tyr	-Leu	-Ser	-Gln	-Gly	-Thr-
Rabbit muscle	Ala	-Leu	-Ser	-Asp	-His	-His	-Ile	-Tyr	-Leu	———	Glu	-Gly	-Thr-
Sturgeon muscle	Ala	-Leu	-Ser	-Asp	-His	-His	-Val	-Tyr	-Leu	———	Glu	-Gly	-Thr-
	1	2	3	4	5	6	7	8	9	9a	10	11	12

Rabbit liver	Leu	-Leu	-LYS	-Asn	-Pro	-Met	-Val	-Thr	-Ala	-Gly	-His	-Ala	-Cmc-
Rabbit muscle	Leu	-Leu	-LYS	-Pro	-Asn	-Met	-Val	-Thr	-Pro	-Gly	-His	-Ala	-Cmc-
Sturgeon muscle	Leu	-Leu	-LYS	-Pro	-Asn	-Met	-Val	-Thr	-Ala	-Gly	-Gln	-Ala	-Cmc-
	13	14	15	16	17	18	19	20	21	22	23	24	25

Rabbit liver	Thr	-Lys									
Rabbit muscle	Thr	-Gln	-Lys	-Tyr	-Ser	-His	-Glu	-Glu	-Ile	-Ala	-Met-
Sturgeon muscle	Thr	-Lys	-Lys	-Tyr	-Thr	-Ser	-Gln	-Glu	-Ile	-Ala	-Met-
	26	27	28	29	30	31	32	33	34	35	36

Figure 10 *The amino-acid sequence around the substrate-binding lysine residue of aldolase from rabbit muscle, rabbit liver, and sturgeon muscle. The rabbit liver sequence is that proposed by Morse and Horecker (ref.* 496). *The substrate-binding residue is at position* 15 *and Cmc* = S-*carboxymethylcysteine*

enzyme (glyceraldehyde 3-phosphate dehydrogenase) in a wide variety of species.[348] Homologies between the two muscle enzymes and that of rabbit liver [496] are shown in Figure 10. Care should be taken during preparation of the liver enzyme to avoid loss of the *C*-terminal tyrosine residue by proteolysis.[497]

Despite the homology in sequence in the region of the active site lysine residue, the rabbit liver and muscle enzymes show differences in substrate specificity. This is also true for the rat enzymes, and it is interesting that the properties of aldolase from Novikoff hepatoma are virtually identical with those of the muscle enzyme and significantly different from those of the liver enzyme;[498] this supports the conclusion that liver aldolase is replaced by the muscle enzyme during development of the hepatoma. Two isoenzymes of transaldolase from *Candida utilis* have been characterised.[499] They are α_2 or β_2 dimers, have *C*-terminal phenylalanine and dimer mol. wts. of about 76 000 and 65 000, respectively. Despite the differences in size it is reported that they hybridise ($\alpha\beta$), and that a mono-β-glyceryl monomer will also hybridise with an unmodified monomer.

[496] D. E. Morse and B. L. Horecker, *Arch. Biochem. Biophys.*, 1968, **125**, 942.
[497] A. G. Lacko, L. W. Brox, R. W. Gracy, and B. L. Horecker, *J. Biol. Chem.*, 1970, **245**, 2140.
[498] R. W. Gracy, A. G. Lacko, L. W. Brox, R. C. Adelman, and B. L. Horecker, *Arch. Biochem. Biophys.*, 1970, **136**, 480.
[499] O. Tsolas and B. L. Horecker, *Arch. Biochem. Biophys.*, 1970, **136**, 303.

Interest has focussed on several regions of the muscle glycogen phosphorylase molecule. Earlier work had defined the sequences around the nine cysteine residues. In both phosphorylase *b*(dimer) and *a*(tetramer) the reactive thiol groups (as measured by reaction with chlorodinitrobenzene) [500] reside in the sequences:

-Ile-Cys-Gly-Gly-Trp-Gln-Met-Glu-Glu-Ala-Asp-Asp-Trp-Leu-Arg-

and

-Phe-Gly-Cys-Arg-Asp-Pro-Val-Arg-

In both forms of the enzyme, the activator AMP protects against reaction. A study of the reaction of the enzyme with iodoacetamide [501] showed that alkylation of two thiol groups in the sequences

-Ala-Cys-Ala-Phe- and -Asn-Ala-Cys-Asp-

resulted in inactivation and dissociation into monomers. These thiol groups were very sensitive to alterations in other parts of the molecule, such as the pyridoxal phosphate binding site, nicely demonstrating conformational changes in the protein. Alkylation of the cysteine in the second of these sequences was prevented by AMP. Two other thiol groups labelled rapidly without effect on activity or structure; these occur in the sequences:

-Asn-Glu-Lys-Ile-Cys-Gly-Gly- and -Gly-Cys-Arg-Asp-

and appear to be the same thiols that react with chlorodinitrobenzene. It is reported [502] that, when pyridoxal phosphate is removed from rabbit muscle phosphorylase, thiol groups that are otherwise buried become exposed. Treatment with [^{14}C]FDNB resulted in isolation of three labelled peptides assumed to derive from two cysteine-containing sequences:

-Asx-Glx-Lys-Cys-Gly-Gly- and -Asx-Ala-Cys-Asp-

The second of these is thought to be part of the pyridoxal phosphate binding site.[502] However, the allosteric properties of this enzyme necessitate great care in the interpretation of results such as these. The phosphorylase from rabbit liver has also been studied.[503] Both the active and the inactive forms of this enzyme are dimers (mol. wt. 185 000) and activation occurs through phosphorylation. The amino-acid sequence at the site of phosphorylation was determined and compared with that of the muscle enzyme:

Liver	-Arg-Gln-Ile-Ser(P)-Ile-Arg-
Muscle	-Lys-Gln-Ile-Ser(P)-Ile-Arg-

[500] A. M. Gold and D. Blackman, *Biochemistry*, 1970, **9**, 4408.
[501] O. Avramovic-Zikic, L. B. Smillie, and N. B. Madsen, *J. Biol. Chem.*, 1970, **245**, 1558.
[502] S. Shaltiel and Y. Zaidenzaig, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 1003.
[503] D. P. Wolf, E. H. Fisher, and E. G. Krebs, *Biochemistry*, 1970, **9**, 1923.

A comparison of the amino-acid compositions of the liver and muscle enzymes from several species (human, rabbit, and rat) indicated greater similarity between liver enzymes or muscle enzymes from different species than between the liver enzyme and muscle enzyme in a single species.[503] It is likely that further light will be shed on the structure and action of phosphorylase from studies with the fluorogenic probe introduced at thiol groups by reaction with 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole (NBD chloride).[504]

Determination of the amino-acid sequence (123 residues) of phospholipase A from pig pancreas is now complete,[505a] and the six disulphide bridges have been placed.[505b] The structure of the zymogen also now becomes available since the seven-residue activation peptide derived from tryptic attack at the *N*-terminus has already been defined;[506] the sequence of the zymogen is shown in Figure 11. Disulphide bridges connect the

Glu-Glu-Gly-Ile-Ser-Ser-Arg-Ala-Leu-Trp-Gln-Phe-Arg-Ser-Met-Ile-Lys-Cys-
1 **10**

Ala-Ile-Pro-Gly-Ser-His-Pro-Leu-Met-Asp-Phe-Asn-Asn-Tyr-Gly-Cys-Tyr-Cys-
20 **30**

Gly-Leu-Gly-Gly-Ser-Gly-Thr-Pro-Val-Asn-Glu-Leu-Asn-Arg-Cys-Glu-His-Thr-
40 **50**

Asp-Asn-Cys-Tyr-Arg-Asp-Ala-Lys-Asn-Leu-Asn-Asp-Ser-Cys-Lys-Phe-Leu-Val-
60 **70**

Asp-Asn-Pro-Tyr-Thr-Glu-Ser-Tyr-Ser-Tyr-Cys-Ser-Ser-Asn-Thr-Glu-Ile-Thr-
80 **90**

Cys-Asn-Ser-Lys-Asn-Asn-Ala-Cys-Glu-Ala-Phe-Ile-Cys-Asn-Asp-Arg-Asn-Ala-
100

Ala-Ile-Cys-Phe-Ser-Lys-Ala-Pro-Tyr-Asn-Lys-Glu-His-Lys-Asn-Leu-Asn-Thr-
110 **120**

Lys-Lys-Tyr-Cys
130

Figure 11 *The complete amino-acid sequence of reduced zymogen of phospholipase A*

following residues: 18–83, 34–130, 36–51, 57–111, 68–98, and 91–103. *X*-Ray studies are said to be in progress and it will be particularly interesting to see how the three-dimensional structure fits the enzyme for its lipolytic

[504] D. J. Birkett, N. C. Price, G. C. Radda, and A. G. Salmon, *F.E.B.S. Letters*, 1970, **6**, 346.

[505] (*a*) G. H. De Haas, A. J. Slotboom, P. P. M. Bonsen, L. L. M. Van Deenen, S. Maroux, A. Puigserver, and P. Desnuelle, *Biochim. Biophys. Acta*, 1970, **221**, 31; (*b*) G. H. De Haas, A. J. Slotboom, P. P. M. Bonsen, W. Nieuwenhuizen, L. L. M. Van Deenen, S. Maroux, V. Dlouha, and P. Desnuelle, *ibid.*, p. 54.

[506] G. H. De Haas, F. Franček, B. Keil, D. W. Thomas, and E. Lederer, *F.E.B.S. Letters*, 1969, **4**, 25.

function. Amino-acid sequences from the *N*- and *C*-terminal regions of snake venom phospholipase A have also been reported:[507]

Glp-Phe-Glu-Thr-Leu-Ile-Met-Ser-Leu-Met-Lys-Ile-Ala-Gly-Arg-Tyr-Tyr——
——Asn-Cys-Gln-Cys-Glu-(Pro, Ser, Glx)-Glu, Cys

Surprisingly, apart from the *C*-terminal cystine, there appears to be no homology with the pig pancreatic enzyme.[505]

Amino-acid decarboxylases commonly have as prosthetic group pyridoxal phosphate, in Schiff base combination with a lysine residue on the enzyme. The Schiff base can be stabilised by reduction and the binding site thereby located. For *E. coli* glutamate decarboxylase [508] it was shown to be:

-Ser-Ile-Ser-Ala-Ser-Gly-His-$\overset{*}{\text{L}}$ys-Phe-

The labelled peptide was purified by an enterprising column 'diagonal' method on anion exchange resin, using alkaline phosphatase treatment as the intermediate step to generate the charge change. The decarboxylase has six subunits of mol. wt. 50 000, has *N*-terminal methionine, and the *C*-terminal sequence -Lys-His-Thr.[509] An unusual situation obtains in histidine decarboxylase of *Lactobacillus* 30a; the enzyme (mol. wt. 190 000) contains five pyruvoylphenylalanyl residues at the *N*-terminus of five of the ten chains. It is now shown that the pyruvoyl residues derive from serine.[510] The two types of chain are characterised as follows:

Chain I: mol. wt. 9000; *N*-terminal Ser; *C*-terminal Ser; no pyruvate, cystine, His, or Phe
Chain II: mol. wt. 29 000; *N*-terminal pyruvate; *C*-terminal Tyr

It is proposed [511] that the role of the pyruvate residue in the enzymic reaction is analogous to that of pyridoxal phosphate in other amino-acid decarboxylases. It is worth nothing here that a study has been made of the transamination of *N*-terminal residues *in vitro*,[512] and a new micro method has recently been described for the determination of reactive carbonyl groups in proteins and peptides, using 2,4-dinitrophenylhydrazine.[513]

The pyridoxal binding site (one per chain of mol. wt. 55 000) of tryptophanase from *E. coli* [514] occurs in the sequence:

-Ser-Ala-Lys-$\overset{*}{\text{L}}$ys-Asp-Ala-Met-Val-Pro-Met-

[507] Y. Samejima, S. Iwanaga, T. Suzuki, and S. Kawauchi, *Biochim. Biophys. Acta*, 1970, **221**, 417.
[508] P. H. Strausbauch and E. H. Fischer, *Biochemistry*, 1970, **9**, 233.
[509] P. H. Strausbauch and E. H. Fischer, *Biochemistry*, 1970, **9**, 226.
[510] W. D. Riley and E. E. Snell, *Biochemistry*, 1970, **9**, 1485.
[511] P. A. Recsei and E. E. Snell, *Biochemistry*, 1970, **9**, 1492.
[512] H. B. F. Dixon, *Biochem. J.*, 1964, **92**, 661; *ibid.*, 1967, **103**, 38P.
[513] R. Fields and H. B. F. Dixon, *Biochem. J.*, 1971, **121**, 587.
[514] H. Kagamiyama, Y. Morino, and E. E. Snell, *J. Biol. Chem.*, 1970, **245**, 2819.

The authors examined the known sequences of the pyridoxal phosphate binding sites in several other enzymes [514] but (not too surprisingly, perhaps) found no recognisable pattern of amino-acid residues, as shown:

Aspartate aminotransferase (extra mitochondrial)	-Ser -Lys-Asn-Phe-
(mitochondrial)	-Ala -Lys-Asn-Met-
Glutamate decarboxylase	-Ser-Ile-Ser-Ala-Gly-His -Lys-Phe-
Phosphorylase	-Met-Lys-Phe-Met-
Pyridoxamine pyruvate transaminase	-Val-Thr-Gly-Pro-Asp-Lys-Cys-Leu-
Tryptophanase	-Ser-Ala-Lys -Lys-Asp-Ala-Met-Val-Pro-Met-

It has been suggested that protection against the generation of multiple forms of cytoplasmic aspartate aminotransferase is afforded if dithiothreitol is included at the beginning of a preparation.[515] The multiple forms are thought to be 'conformers' which have arisen through oxidation of sulphydryl groups: reduction at a later stage appears to regenerate activity but not the native conformation.[515] This might be a useful point to bear in mind when dealing with other enzymes; some may regenerate neither conformation nor activity once oxidised.

The sequence of 89 residues at the *N*-terminus of human carbonic anhydrase B has been determined (Figure 12).[516] It is stated [516] that this agrees

Acetyl-Ala-Ser-Pro-Asp-Trp-Gly-Tyr-Asp-Asp-Lys-Asn-Gly-Gln-Pro-Glu-Trp-
1 **10**

Ser-Lys-Leu-Tyr-Pro-Ile-Ala-Asp-Gly-Asn-Asn-Gln-Ser-Pro-Val-Asp-Ile-Lys-
20 **30**

Thr-Ser-Glu-Thr-Lys-His-Asn-Thr-Ser-Leu-Lys-Pro-Ile-Ser-Val-Ser-Tyr-Asn-Pro-
40 **50**

Ala-Thr-Ala-Lys-Glu-Ile-Ile-Asn-Val-Gly-His-Ser-Phe-His-Val-Phe- (Asx, Asx,
60 **70**

Glx, Asx)-Asx-Asx-Arg-Ser-Val-Leu-Lys-Gly-Gly-Pro-Phe-Ser-Asp-Ser-Tyr-Arg
80 **89**

Figure 12 N-*Terminal amino-acid sequence of human carbonic anhydrase* B

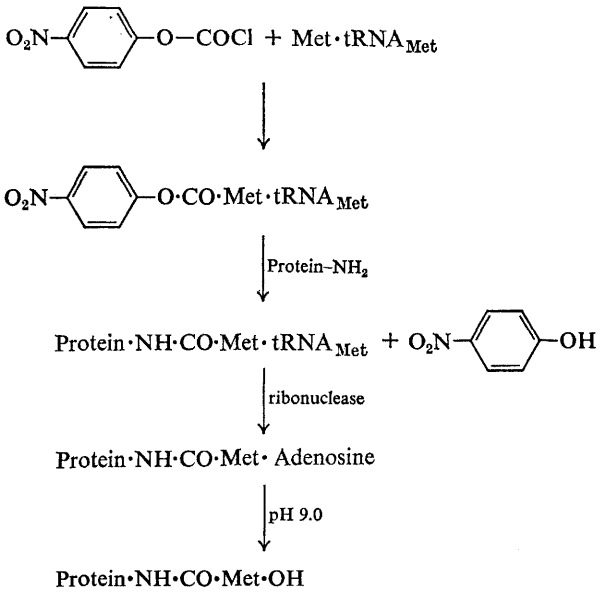
with that determined by another group of workers [517] with the exception of an inversion of residues 69 and 70. A comparison has been made of the primary structures of carbonic anhydrase B from man, great apes, and Old World monkeys.[518]

[515] M. Arrio-Dupont, I. Cournil, and P. Duie, *F.E.B.S. Letters*, 1970, **11**, 144.
[516] M. Sciaky, N. Limozin, N. Giraud, C. Marriq, M. Charrel, and G. Laurent, *Biochim. Biophys. Acta*, 1970, **221**, 665.
[517] B. Andersson, P. O. Nyman, and L. Strid, 'Atlas of Protein Sequence and Structure, ed. M. O. Dayhoff, 1970, vol. 5 (in preparation). (Cited in ref. 516.)
[518] R. E. Tashian and S. R. Stroup, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 1457.

Work is in progress on the primary structure of thiolase from pig heart. The tetrameric enzyme has one active thiol group per monomer to which substrate becomes bound in thiol ester linkage. The sequence around this cysteine residue was shown to be:[519]

-Gln-Ala-Val-Leu-Gly-Ala-Gly-Leu-Pro-(Cmc, Asn, Thr_3, Ser, Pro, Ile_2)-Lys-Val-Cys-Ala-Ser-Gly-Met-Lys-

Structural work is also under way on methionine tRNA synthetase from *E. coli*. The unique *N*-terminal sequence is Ala-Gly-Gly-Thr- and was elucidated[59] using the micro-method already described (Section 2B). A pleasing application of cross-linking (using *p*-nitrophenyl chloroformate) enabled the amino-group of methionine, attached to its tRNA, to be covalently linked with an amino-group on the activating enzyme as shown in Scheme 8. Any mono-substituted product hydrolyses to a carbamic acid



Scheme 8

derivative which immediately decarboxylates to regenerate the original amino-group. Using this procedure, a uniquely-reacting lysine residue in methionine tRNA synthetase was shown[59] to occur in the sequence:

-Phe-Thr-Tyr-Gln-Lys-Leu-His-Asn-

This lysine residue is presumably near the amino-acid recognition site on the enzyme.

[519] U. Gehring and J. I. Harris, *European J. Biochem.*, 1970, **16**, 492.

There is some heterogeneity in sequence around the serine residue that can be labelled with [^{32}P]DFP in microsomal carboxyl esterase:[520]

$$\text{-Gly-Glu-Ser(P)-Ala-Gly-}\begin{matrix}\text{Glu}\\\text{Asp}\end{matrix}\text{-Gly-Ser-}$$

The same mixture of peptides is obtained from the pig liver and pig kidney enzymes. Heterogeneity is also reported for *E. coli* alkaline phosphatase [521] but in this case it is attributed to aminopeptidase activity resulting in removal of the *N*-terminal histidine and possibly the next residue (valine). AMP is an inhibitor of the allosteric enzyme glutamine synthetase. The AMP becomes bound to a tyrosyl side-chain on each of the 12 subunits of the enzyme by phosphodiester linkages. The sequence around these tyrosine residues has now been defined:[522]

-Asp-Leu-Tyr-Asp-Leu-Pro-Pro-Glu-(Ala, Glu, Gly)-Lys-

The *N*-terminal sequence of asparaginases A and B from *E. coli* has been shown to be:[523]

Leu-Pro-Asn-Ile-Thr-Ile-Leu-Ala-Thr-Gly-

The enzyme has a mol. wt. of 120 000 and may be a hexamer of subunit mol. wt. 20 000.

Work is in progress on β-galactosidase (*E. coli*) which is a tetramer with subunit mol. wt. 135 000; the isolation and composition of tryptic peptides accounting for over half of the 1170 residues per chain has been reported.[524] The two non-identical subunits of ribulose 1,5-diphosphate carboxylase from spinach leaf have been characterised:[525] A, mol. wt. 60 000 with *C*-terminal valine; and B, mol. wt. 20 000 with *C*-terminal tyrosine. Enolase from salmon [526] is likewise a dimer, of mol. wt. 100 000; as judged by an unique *C*-terminal sequence, -Ile-Asn, it is possible that the subunits may be identical (mol. wt. about 50 000). There is evidence for a phosphohistidine intermediate in the phosphoglyceromutase reaction.[527] SDS-Gels showed that there is 1 g atom of ^{32}P bound per 27 000 g of protein; this is half the reported mol. wt. so it appears that the molecule is a dimer.[527] The purification, characterisation, and crystallisation of porcine myokinase have been described.[528] This is a single-chain enzyme of mol. wt. 21 000. The enzyme from pig is reported to give better crystals for *X*-ray purposes than the rabbit enzyme, and crystallographic studies are

[520] E. Heymann, K. Krisch, and E. Pahlich, *Z. physiol. Chem.*, 1970, **351**, 931.
[521] S. Natori and A. Garen, *J. Mol. Biol.*, 1970, **49**, 577.
[522] R. L. Heinrikson and H. S. Kingdon, *J. Biol. Chem.*, 1970, **245**, 138.
[523] A. Areus, E. Raienbusch, E. Irion, O. Wagner, K. Bauer, and W. Kaufmann, *Z. physiol. Chem.*, 1970, **351**, 197.
[524] A. V. Fowler and I. Zabin, *J. Biol. Chem.*, 1970, **245**, 5032.
[525] T. Sugiyama and T. Akazawa, *Biochemistry*, 1970, **9**, 4499.
[526] R. C. Ruth, D. M. Soja, and F. Wold, *Arch. Biochem. Biophys.*, 1970, **140**, 1.
[527] Z. B. Rose, *Arch. Biochem. Biophys.*, 1970, **140**, 508.
[528] I. Schirmer, R. H. Schirmer, G. E. Schulz, and E. Thuma, *F.E.B.S. Letters*, 1970, **10**, 333.

already under way.[528] Finally, the purification and properties of the initiation factor F_3 have been described;[529] this factor is required for formation of the initiation complex in protein biosynthesis when natural messenger RNA is used. All three initiation factors have thus now been obtained virtually pure. F_3 has a mol. wt. on SDS-gels of 21 000 and appears to have a single reactive thiol group.[529]

F. Haem Proteins.—*Haemoglobin and Myoglobin.* The primary structure (153 residues) of ox heart myoglobin has been deduced,[530] partly by homology with the closely related horse protein, and the differences between dolphin and sperm whale myoglobin have been discussed.[531] Of the 13 differences, only one (Gly/Ser) involves 'internal' residues. A comparison of the methoxycarbonylation of sperm whale myoglobin in the crystal[532] and in solution[533] indicates a close similarity for the two structures, most of the differences being ascribable to the presence of neighbouring protein molecules in the crystal lattice (*cf.* chymotrypsin, p. 83). The only significant difference unexplainable in these terms is that His-36 becomes unreactive when the protein passes from crystal to solution, implying a structural change affecting the accessibility of this residue.

Many aspects of the structure and function of haemoglobin have recently been reviewed.[534] The two haemoglobins from lamprey have been shown to differ simply by the formylation of the *N*-terminal proline residue of one of them,[535] and the *N*-terminal sequences of three of the multiple haemoglobins from the insect larvae *Chironomus thummi*[536] and of the α-chain of grey kangaroo,[537] have been reported:

```
Kangaroo  Val-Leu-Ser-Ala-Ala-Asp-Lys-Gly-His-Val-Lys-Ala-Ile-Trp-Gly-
CT-1      Gly-Pro-Ser-Gly-Asp-Gln-Ile-Ala-Ala-Ala-Lys-Ala-Ser-Trp-Asn-
CT-3          Leu-Ser-Ala-Asp-Gln-Ile-Ser-X-Val-Glx-Ala-Ser-Phe-Asp-
CT-4          Leu-Thr-Ala-Asp-Gln-Ile-Ser-Thr-Val-Gln-Ser-X-Phe-X-
Human α   Val-Leu-Ser-Pro-Ala-Asp-Lys-Thr-Asn-Val-Lys-Ala-Ala-Trp-Gly-
              1                                  10
Kangaroo  Lys-
CT-1      Thr-Val-Lys-Asn----------Asn-Gln-Val-Ser-Ile-Leu-Tyr-
CT-3      Lys-Val-Lys-Gly----------Asx-Pro-Val-Gly-Ile-Leu-Tyr-
CT-4      X-Val-Lys-Gly----------Asp-Ala-Val-Gly-Ile-Leu-Tyr-
Human α   Lys-Val-Gly-Ala-His-Ala-Gly-Glu-Tyr-Gly-Ala-Glu-Ala-Leu-Glu
                        20                                    30
```

[529] S. Sabol, M. A. G. Sillero, K. Iwasaki, and S. Ochoa, *Nature*, 1970, **228**, 1269.
[530] K.-K. Han, M. Dautrevaux, X. Chaila, and G. Biserte, *European J. Biochem.*, 1970, **16**, 465.
[531] M. Karadjova, P. Nedkov, A. Bakardjieva, and N. Genov, *Biochim. Biophys. Acta*, 1970, **221**, 136.
[532] T. E. Hugli and F. R. N. Gurd, *J. Biol. Chem.*, 1970, **245**, 1930.
[533] T. E. Hugli and F. R. N. Gurd, *J. Biol. Chem.*, 1970, **245**, 1939.
[534] E. Antonini and M. Brunori, *Arch. Biochem. Biophys.*, 1970, **39**, 977.
[535] H. Fujiki, G. Braunitzer, and V. Rudloff, *Z. physiol. Chem.*, 1970 **351**, 901.
[536] G. Braunitzer, H. Neuwirth, H. Mussnig, and B. Schrank, *Z. physiol. Chem.*, 1970, **351**, 1289.
[537] J. M. Beard and E. O. P. Thompson, *Austral. J. Biol. Sci.*, 1970, **23**, 185.

Details have now been given [538] of the determination of the sequences of the α- and β-chains of rhesus monkey haemoglobin and the sequence of the β-chain of frog (*Rana esculenta*) haemoglobin has been reported.[539] Corrections have been made to the amino-acid sequences around the histidine residues of leghaemoglobin [540] and it has been shown in a number of laboratories that methionine is the first amino-acid incorporated in the biosynthesis of haemoglobin and (probably) all other eukaryotic proteins.[541] This *N*-terminal residue is then cleaved off enzymically before the polypeptide chain is released from the ribosome.

The probable sequences for the α- and β-chains of the major form of the haemoglobin from the monkey *Macaca irus* have been deduced by analogy with other haemoglobins,[542] and there is evidence for the existence in the minor form of a different type of α-chain arising by gene duplication.[543] Duplication of the α-chain gene has also been observed in barbary sheep [544] (see last year's Report), and two and three non-allelic genes, respectively, have been postulated for the α-chain and β-chain of the haemoglobin of the golden hamster.[545] Multiple forms of haemoglobin occur in the Atlantic salmon but in this instance tetraploidy seems a more likely explanation than isolated gene duplication,[546] since multiple forms were also observed for various enzymes. The three forms of haemoglobin detected in mouse embryos can be explained [547] in terms of the different α-like and β-like chains that comprise them and a third embryonic form of human haemoglobin has been reported,[548] with the composition $\alpha_2\zeta_2$. The other two embryonic forms are ε_4 and $\alpha_2\varepsilon_2$.

Further abnormal human haemoglobins have been described: among them are G *Taichung* (α74 Asp → His),[549] G *Georgia* (α95 Pro → Leu),[550] J *Tongariki* (α115 Ala → Asp),[551] G *Makassar* (β6 Glu → Ala),[552] *Nagasaki*

[538] G. Matsuda, T. Maita, H. Ota, and H. Takei, *Internat. J. Protein Res.*, 1970, **2**, 99.
[539] J.-P. Chauvet and R. Acher, *F.E.B.S. Letters*, 1970, **10**, 136.
[540] S. J. Aggarwal and A. Riggs, *Acta Chem. Scand.*, 1970, **24**, 2234.
[541] A. E. Smith and K. A. Marcker, *Nature*, 1970, **226**, 607; J. C. Brown and A. E. Smith, *ibid.*, p. 610; D. B. Wilson and H. M. Dintzis, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 1282; J. P. Leis and E. B. Keller, *Biochim. Biophys. Res. Comm.*, 1970, **40**, 416; R. Jackson and T. Hunter, *Nature*, 1970, **227**, 672; D. T. Wigle and G. H. Dixon, *ibid.*, p. 676.
[542] P. T. Wade, N. A. Barnicot, and E. R. Huehns, *Biochim. Biophys. Acta*, 1970, **221**, 450.
[543] N. A. Barnicot, P. T. Wade, and P. Cohen, *Nature*, 1970, **228**, 379.
[544] J. B. Wilson, R. N. Wrightstone, and T. H. J. Huisman, *Nature*, 1970, **226**, 354.
[545] Y. Yasukochi, *Biochim. Biophys. Acta*, 1970, **221**, 1.
[546] N. P. Wilkins, *Biochim. Biophys. Acta*, 1970, **214**, 52.
[547] G. Vulpis and A. Bank, *Biochim. Biophys. Acta*, 1970, **207**, 390.
[548] G. L. Capp, D. A. Rigas, and R. T. Jones, *Nature*, 1970, **228**, 278.
[549] R. Q. Blackwell and C.-S. Liu, *Biochim. Biophys. Acta*, 1970, **200**, 70.
[550] T. H. J. Huisman, H. R. Adams, J. B. Wilson, G. D. Efremov, C. A. Reynolds, and R. N. Wrightstone, *Biochim. Biophys. Acta*, 1970, **200**, 578.
[551] R. K. Abramson, D. L. Rucknagel, D. C. Shreffler, and J. J. Saave, *Science*, 1970, **169**, 194.
[552] R. Q. Blackwell, S. Oemijati, W. Pribadi, M.-I. Weng, and C.-S. Liu, *Biochim. Biophys. Acta*, 1970, **214**, 396.

(β17 Lys → Glu),[553] and *Shepherd's Bush* (β74 Gly → Asp).[554] Haemoglobin *Shepherd's Bush* is an unstable haemoglobin associated with mild haemolytic anaemia: the other mutations give rise to no adverse clinical symptoms. This is of particular interest with haemoglobin *Makassar* since the change, β6 Glu → Val, results in the well-known sickle-cell haemoglobin. A preliminary report[555] of an inheritable abnormal β-chain in rabbit haemoglobin has also appeared.

When human β-chains in which the cysteine residue at position 112 has been methoxycarbonylated are combined with normal α-chains and when normal β-chains are combined with α-chains in which the cysteine residue at position 104 has been modified, dimers and tetramers result, all of which show increased oxygen affinity and decreased Bohr effect and a loss of co-operative interactions.[556, 557] Thus, modification at the $\alpha_1\beta_1$ interface can seriously affect the function of the haemoglobin molecule, presumably by distorting the interactions between subunits.[554] Removal of the *C*-terminal histidine residues from the β-chains of human haemoglobin by careful digestion with carboxypeptidase confirms that the modified haemoglobin exhibits haem–haem interaction but lacks half the alkaline Bohr effect.[558] The results of *X*-ray crystallographic and chemical analysis of haemoglobin have been brought together in a substantial review of the structure and activity of this molecule, leading to an hypothesis for its mechanism of action.[559]

Cytochromes. The primary structures of cytochrome *c* from tobacco horn worm moths[560] and carp (*Cyprinus carpio*)[561] have been established; that from carp shows heterogeneity (Asp or Glu) at position 4, presumably as a result of genetic polymorphism. The amino-acid sequences (111 residues) of cytochrome *c* from mung-bean (*Phaseolus aureus* L.)[562] and sunflower (*Helianthus annuus* L.)[563] have been described and compared with that from wheat germ. All three proteins contain ε-*N*-trimethyl-lysine at positions 72 and 86 and their sequences are generally very similar, suggesting a common ancestor for the cytochromes *c* of higher plants. ε-*N*-Trimethyl-lysine has also been reported to occupy position 72 of yeast iso-I-cytochrome *c*.[564]

[553] M. Maekawa, T. Maekawa, N. Fujiwara, K. Tabara, and G. Matsuda, *Internat. J. Protein Res.*, 1970, **2**, 147.
[554] J. M. White, M. C. Brain, P. A. Lorkin, H. Lehmann, and M. Smith, *Nature*, 1970, **225**, 939.
[555] B. Luppis and F. Conconi, *Experientia*, 1970, **26**, 488.
[556] E. J. Neer, *J. Biol. Chem.*, 1970, **245**, 564.
[557] E. J. Neer and G. Guidotti, *J. Biol. Chem.*, 1970, **245**, 570.
[558] J. V. Kilmartin and J. F. Wootton, *Nature*, 1970, **228**, 766.
[559] M. F. Perutz, *Nature*, 1970, **228**, 726.
[560] S. K. Chan, *Biochim. Biophys. Acta*, 1970, **221**, 497.
[561] L. Gürtler and H. J. Horstmann, *European J. Biochem.*, 1970, **12**, 48.
[562] E. W. Thompson, M. V. Laycock, J. A. M. Ramshaw, and D. Boulter, *Biochem. J.*, 1970, **117**, 183.
[563] J. A. M. Ramshaw, E. W. Thompson, and D. Boulter, *Biochem. J.*, 1970, **119**, 535.
[564] R. J. DeLange, A. N. Glazer, and E. L. Smith, *J. Biol. Chem.*, 1970, **245**, 3325.

The tyrosine residues at positions 46, 48, 67, and 74 in tuna heart cytochrome *c* can be iodinated whereas Tyr-97 is inaccessible.[565] However, of the four tyrosine residues in horse heart cytochrome *c* (positions 48, 67, 74, and 97), only Tyr-67 and Tyr-74 are significantly iodinated at pH 7.0 and pH 9.5.[566] Moreover, only Tyr-48 and Tyr-67 are nitrated by tetranitromethane at pH 8.[567]

A partial purification has been reported[568] for a soluble cytochrome from pig kidney with spectral properties similar to those of microsomal cytochrome b_5. The amino-acid sequence of calf liver microsomal cytochrome b_5 has been corrected[569] and that of the rabbit liver protein has been reported.[570] It is four residues longer at the *N*-terminus than the calf cytochrome but is otherwise very similar.[570] The primary structures of human, ox, chicken, and rabbit cytochrome b_5 have also been determined by other workers.[571] A few differences remain between the ox and calf sequences, even after the correction[569] to the calf sequence is allowed for.[571]

Bacterial cytochromes have been the subject of a recent review.[572] Cytochrome c_{550} has been purified from *Spirillum itersonii*[573] and two *c*-type cytochromes have been isolated from *Rhodospirillum molischianum*.[574] One is of the c_2-type present in most photosynthetic bacteria and contains 122 residues, whereas the other is only 93 residues long. Both have the *N*-terminal sequence Ala-Asp-Ala-Pro-Pro- and peptide maps suggest they are structurally related.[574] Cytochrome c_3 from *Desulphovibrio desulphuricans* has been characterised in terms of amino-acid composition and shown to contain three haem groups in each protein molecule.[575]

G. Non-haem Electron Transport Proteins.—Details have now been published of the determination of the amino-acid sequence of *Chromatium* ferredoxin.[576] The protein shows homology with the ferredoxins of non-photosynthetic bacteria. Similarly, the determination of the amino-acid sequence of taro ferredoxin (from *Colocasia esculenta*) allows a comparison

[565] E. Stellwagen and E. B. McGowan, *Biochemistry*, 1970, **9**, 3765.
[566] E. B. McGowan and E. Stellwagen, *Biochemistry*, 1970, **9**, 3047.
[567] K. Skov, T. Hofmann, and G. R. Williams, *Canad. J. Biochem.*, 1969, **47**, 750.
[568] J. H. Mangum, M. D. Klingler, and J. A. North, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 1520.
[569] J. Ozols and P. Strittmatter, *J. Biol. Chem.*, 1969, **244**, 6617.
[570] J. Ozols, *J. Biol. Chem.*, 1970, **245**, 4863.
[571] A. Tsugita, M. Kobayashi, S. Tani, S. Kyo, M. A. Rashid, Y. Yoshida, T. Kajihara, and B. Hagihara, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 442.
[572] M. D. Kamen and T. Horio, *Ann. Rev. Biochem.*, 1970, **39**, 673.
[573] G. D. Clark-Walker and J. Lascelles, *Arch. Biochem. Biophys.*, 1970, **136**, 153.
[574] K. Dus, T. Flatmark, H. deKlerk, and M. D. Kamen, *Biochemistry*, 1970, **9**, 1984.
[575] H. Drucker, E. B. Trousil, L. L. Campbell, G. H. Barlow, and E. Margoliash, *Biochemistry*, 1970, **9**, 1515.
[576] H. Matsubura, R. M. Sasaki, D. K. Tsuchiya, and M. C. W. Evans, *J. Biol. Chem.*, 1970, **245**, 2121.

to be made with other plant-type ferredoxins.[577] The sequences are closely related: in particular, the sequence

$$\text{-Leu-Pro-Tyr-Ser-Cys-Arg-Ala-Gly-}\underset{\text{Ala}}{\overset{\text{Ser}}{}}\text{-Cys-Ser-Ser-Cys-Ala-Gly-Lys-}$$

is a highly conserved region near the middle of the molecule (residues 35—50).[577]

The amino-acid sequence of ox adrenodoxin, another iron-binding protein in an electron transport chain concerned with steroid hydroxylation, has almost been completed.[578] The single chain of 118 residues contains five cysteines, most of which are involved in binding iron. No homology is observed with the bacterial and plant ferredoxins, nor with mammalian haem proteins. The five cysteine residues occur in the sequences:

-Cys-Glu-Gly-Thr-Leu-Ala-Cys-Ser-Thr-Cys- (46, 55) and -Cys-Glu-Ile-Cys- (95, 98)

The sequence -Cys-X-Y-Cys- is present in many of the non-haem iron proteins for which sequences are available, suggesting that it may commonly be involved in binding the iron.[578]

Two blue copper proteins, each containing one copper atom per molecule, have been described. One, a plastocyanin from *Phaseolus vulgaris*, has a mol. wt. of 10 700 and the *N*-terminal sequence Leu-Glx-Val-Leu.[579a] The other, umecyanin, from the horseradish root, has a mol. wt. of 14 600.[579b]

H. Studies on Quaternary Structure.—The quaternary structures of many proteins have recently been reviewed.[580] Some newer ones have been mentioned elsewhere in this Report: others are listed in Table 2.

Table 2 *Quaternary structure of some individual enzymes*

Enzyme	*Source*	*Molecular weight*	*No. of subunits*
Adenosine triphosphatase[a, b]	Ox heart mitochondria	280 000	6
	Strep. faecalis	385 000	12
Alcohol dehydrogenase[c]	*Drosophila*	60 000	8
Aldehyde dehydrogenase[d]	Yeast	200 000	4
α-Amylase[e]	*B. subtilis*	48 000	2
Asparaginase[f]	*E. coli*	133 000	4
Aspartate β-decarboxylase[g, h]	*Pseud. dacunhae*	800 000 or 675 000	16 12
	Alicaligenes faecalis	675 000	12

[577] K. K. Rao and H. Matsubura, *Biochem. Biophys. Res. Comm.*, 1970, **38**,500.
[578] M. Tanaka, M. Haniu, and K. T. Yasunobu, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 1182.
[579] (*a*) P. R. Milne and J. R. E. Wells, *J. Biol. Chem.*, 1970, **245**, 1566. (*b*) K.-G. Paul and T. Stigbrand, *Biochim. Biophys. Acta*, 1970, **221**, 255.
[580] I. M. Klotz, N. R. Langerman, and D. W. Darnell, *Ann. Rev. Biochem.*, 1970, **39**, 25.

Table 2 (*cont.*)

Enzyme	*Source*	*Molecular weight*	*No. of subnits*
Pyruvate decarboxylase[qq]	*E. coli*	180 000	2
Pyruvate kinase[rr]	Yeast	168 000	4
Rhodanese[ss]	Ox liver	37 000	2
D-Serine dehydratase[tt]	*E. coli*	45 500	1
Succinyl-CoA synthetase[uu]	*E. coli*	146 000	4
Tamm–Horsfall glycoprotein[vv]	Human urine	23×10^6	230
Thiolase[ww]	Pig heart	170 000	4
Tyrosine aminotransferase[xx]	Rat liver	120 000	4
tRNA synthetase			
Glycyl[yy]	*E. coli*	226 000	$4(\alpha_2\beta_2)$
Isoleucyl[zz aaa]	*E. coli*	114 000	1
Leucyl[bbb]	*E. coli*	105 000	1
Phenyl[ccc]	*E. coli*	180 000	4
Seryl[ddd]	*E. coli*	100 000	2
Valyl[zz]	*E. coli*	110 000	1

[a] G. Forrest and S. J. Edelstein, *J. Biol. Chem.*, 1970, **245**, 6468. [b] H. P. Schelsli, A. E. Vatter, and A. Abrams, *J. Biol. Chem.*, 1970, **245**, 1122. [c] K. B. Jacobson and P. Pfuderer, *J. Biol. Chem.*, 1970, **245**, 3938. [d] J. F. Clark and W. B. Jakoby, *J. Biol. Chem.*, 1970, **245**, 6072. [e] J. M. Connellan and D. C. Shaw, *J. Biol. Chem.*, 1970, **245**, 2845. [f] R. C. Jackson and R. E. Handschumacher, *Biochemistry*, 1970, **9**, 3585; B. H. Frank, A. H. Pekar, A. J. Veros, and P. P. K. Ho, *J. Biol. Chem.*, 1970, **245**, 3716; E. Irion and W. H. Voigt, *Z. physiol. Chem.*, 1970, **351**, 1154. [g] T. Kakimoto, J. Kato, T. Shibatani, N. Nishimura, and I. Chibata, *J. Biol. Chem.*, 1970, **245**, 3369. [h] S. S. Tate and A. Meister, *Biochemistry*, 1970, **9**, 2626; W. F. Bowers, V. B. Czubaroff, and R. H. Haschemeyer, *Biochemistry*, 1970, **9**, 2621. [i] C. Biswas, E. Gray, and H. Paulus, *J. Biol. Chem.*, 1970, **245**, 4900. [j] C. Oriol, M. F. Landon, and N. Van Thoai, *Biochim. Biophys. Acta*, 1970, **207**, 514. [k] M. C. Grant-Greene and F. Friedberg, *Internat. J. Protein. Res.*, 1970, **2**, 235. [l] N. M. Green and E. J. Toms, *Biochem. J.*, 1970, **118**, 67. [m] G. L. E. Koch, D. C. Shaw, and F. Gibson, *Biochim. Biophys. Acta*, 1970, **212**, 387. [n] J. C. Melville and C. A. Ryan, *Arch. Biochem. Biophys.*, 1970, **138**, 700. [o] M. Singh, G. C. Brooks, and P. A. Srere, *J. Biol. Chem.*, 1970, **245**, 4636; J.-Y. Wu and J. T. Yang, *J. Biol. Chem.*, 1970, **245**, 212. [p] D. Cavallini, C. Cannella, G. Federici, S. Dupré, A. Fiori, and E. Del Grosso, *European J. Biochem.*, 1970, **16**, 537. [q] Y. Nagata, T. Yamanaka, and K. Okunuki, *Biochim. Biophys. Acta*, 1970, **221**, 668. [r] R. M. Metrione, Y. Okuda, and G. F. Fairclough, jun., *Biochemistry*, 1970, **9**, 2427. [s] M. R. Rossi Fanelli, E. Chiancone, P. Vecchini, and E. Antonini, *Arch. Biochem. Biophys.*, 1970, **141**, 278. [t] K. G. Mann, F. J. Castellino, and P. A. Hargrave, *Biochemistry*, 1970, **9**, 4002; R. A. Harrison, Part II thesis, Department of Biochemistry, University of Cambridge, 1970. [u] R. C. Ruth, D. M. Soja, and F. Wold, *Arch. Biochem. Biophys.*, 1970, **140**, 1. [v] J. M. Brewer, L. Ljungdahl, T. E. Spencer, and S. H. Neece, *J. Biol. Chem.*, 1970, **245**, 4798. [w] N. Feliss and M. Martinez-Carrion, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 932. [z] P. H. Strausbauch and E. H. Fischer, *Biochemistry*, 1970, **9**, 226. [y] T. F. Deuel, A. Ginsburg, J. Yeh, E. Shelton, and E. R. Stadtman, *J. Biol. Chem.*, 1970, **245**, 5195. [z] R. E. Amelunxen, M. Noelken, and R. Singleton, *Arch. Biochem. Biophys.*, 1970, **141**, 447; F. Heinz and K. D. Kulbe, *Z. physiol. Chem.*, 1970, **351**, 249. [aa] V. L. Seery, E. H. Fischer, and D. C. Teller, *Biochemistry*, 1970, **9**, 3591. [bb] L D. Kohn, W. A. Warren, and W. R. Carroll, *J. Biol. Chem.*, 1970, **245**, 3821. [cc] R. E. Ferrell and G. B. Kitto, *Biochemistry*, 1970, **9**, 3053. [dd] H. D. Ellerton, D. E. Carpenter, and K. E. Van Holde, *Biochemistry*, 1970, **9**, 2225. [ee] J. R. Pringle, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 46. [ff] N. A. Giorgio, jun., A. T. Yip, J. Fleming, and G. W. E. Plaut, *J. Biol. Chem.*, 1970, **245**, 5469; R. F. Colman, R. C. Szeto, and P. Cohen, *Biochemistry*, 1970, **9**, 4945; J. Illingworth and K. F. Tipton, *Biochem. J.*, 1970, **118**, 253; R. L. Howard and R. R. Becker, *J. Biol. Chem.*, 1970, **245**, 3186. [gg] D. J. Prescott and P. R. Vagelos, *J.*

Table 2 (*cont.*)

Biol. Chem., 1970, **245**, 5484. [hh] E. J. Massaro, *Science*, 1970, **167**, 994; R. K. Selander and S. Y. Yang, *ibid.*, 1970, **169**, 179. [ii] C. Jacq and F. Lederer, *European J. Biochem.*, 1970, **12**, 154. [jj] F. J. Castellino, W. W. Fish, and K. G. Mann, *J. Biol. Chem.*, 1970, **245**, 4269. [kk] P. Melius, M. H. Moseley, and D. M. Brown, *Biochim. Biophys. Acta*, 1970, **221**, 62. [ll] F. C. Stevens, D. M. Brown, and E. L. Smith, *Arch. Biochem. Biophys.*, 1970, **136**, 413. [mm] V. Zappia and H. A. Barker, *Biochim. Biophys. Acta*, 1970, **207**, 505. [nn] K. G. Mann and C. S. Vestling, *Biochemistry*, 1970, **9**, 3020. [oo] J. Spina, jun., H. J. Bright, and J. Rosenbloom, *Biochemistry*, 1970, **9**, 3794. [pp] Z. B. Rose, *Arch. Biochem. Biophys.*, 1970, **140**, 508. [qq] G. Dennert and D. Eaker, *F.E.B.S. Letters*, 1970, **6**, 257. [rr] R. T. Kuzcenski and C. H. Suelter, *Biochemistry*, 1970, **9**, 2043. [ss] F. Detoma and J. Westley, *Biochim. Biophys. Acta*, 1970, **207**, 144. [tt] W. Dowhan, jun., and E. E. Snell, *J. Biol. Chem.*, 1970, **245**, 4618. [uu] C. Leitzmann, J.-Y. Wu, and P. D. Boyer, *Biochemistry*, 1970, **9**, 2338. [vv] F. K. Stevenson and P. W. Kent, *Biochem. J.*, 1970, **116**, 791. [ww] U. Gehring and J. I. Harris, *European J. Biochem.*, 1970, **16**, 487. [xx] F. Auricchio, F. Valeriote, G. Tomkins, and W. D. Riley, *Biochim. Biophys. Acta*, 1970, **221**, 307. [yy] D. L. Ostrem and P. Berg, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 1967. [zz] F. Berthelot and M. Yaniv, *European J. Biochem.*, 1970, **16**, 123. [aaa] D. J. Arndt and P. Berg, *J. Biol. Chem.*, 1970, **245**, 665. [bbb] P, Rouget and F. Chapeville, *European J. Biochem.*, 1970, **14**, 498; H. Hayashi, J. R. Knowles. J. R. Katze, J. Lapointe, and D. Söll, *J. Biol. Chem.*, 1970, **245**, 1401. [ccc] M. H. J. E. Kosakowski and A. Böck, *European J. Biochem.*, 1970, **12**, 67. [ddd] J. R. Katze and W. Konigsberg, *J. Biol. Chem.*, 1970, **245**, 923.

6 Immunoglobulins

Rational proposals for ending the confusion in immunoglobulin nomenclature have been made.[581] The whole subject is moving so rapidly that there is always room for good reviews [582] since different authors emphasise different aspects. Details have now been given [583] of the determination of the complete amino-acid sequence of the myeloma IgG protein, Eu, mentioned in last year's Report. A warning has been given that a single molecular species of IgG synthesised by a mouse plasma cell tumour rapidly becomes heterogeneous after synthesis [584] and a human pathological IgG (Po) was found [585] to fragment during purification into components that resembled the Fab and Fc pieces derived from IgG by papain digestion.

A. Light Chains.—The *N*-terminal sequences of the heavy and light chains of immunoglobulins from three lower vertebrates, *viz.* marine toad (*Bufo marinus*), the paddlefish (*Polyodon spathula*), and the gar (*Lepisosteus osseus*), have been examined and shown [586] to be homologous with those of other species. Similarly, the *C*-terminal sequence of ox light chains, -Thr-Val-Lys-Pro-Ser-Glu-Cys-Ser,[587] is closely comparable with that of λ-chains from other animals. The *N*-terminal sequences of pig λ-chains,

[581] (*a*) D. S. Rowe, *Nature*, 1970, **228**, 509. (*b*) *Bulletin W.H.O.*, 1969, **41**, 975.
[582] (*a*) C. Milstein and J. R. L. Pink, *Prog. Biophys. Mol. Biol.*, 1970, **21**, 211. (*b*) T. T. Wu and E. A. Kabat, *J. Exp. Med.*, 1970, **132**, 211; H. Metzger, *Ann. Rev. Biochem.*, 1970, **39**, 889.
[583] P. D. Gottlieb, B. A. Cunningham, U. Rutishauser, and G. M. Edelman, *Biochemistry*, 1970, **9**, 3155; W. E. Gall and G. M. Edelman, *ibid.*, p. 3188.
[584] Z. L. Awdeh, A. R. Williamson, and B. A. Askonas, *Biochem. J.*, 1970, **116**, 241.
[585] C. J. Brackenridge, *Internat. J. Protein. Res.*, 1970, **2**, 265,
[586] R. T. Acton, P. F. Wienheimer, M. Wolcott, E. E. Evans, and J. C. Bennett, *Nature*, 1970, **228**, 991.
[587] D. Beale and M. Squires, *Nature*, 1970, **226**, 1056.

although they carry species-specific residues, are clearly homologous with the corresponding region of human λ-chains:[588] the same is true of pig and human κ-chains.[589] The constant region of pig κ-chains, as one would expect, also shows homology with the constant region of human κ-chains.[590] There is evidence[591] that the amino-acid sequence differences between the b4 and b5 allotypes of rabbit light chains are more extensive than those near the *C*-terminus reported last year. It will, of course, be interesting to discover whether such differences can occur in the variable regions.

Several human Bence Jones proteins have been examined and assigned to variable region sub-groups.[592a] The sequence of a κ-type Bence Jones protein (Hau) belonging to sub-group 1 has been described[592b] and sequences have been reported for the human myeloma λ-chains Ha,[593] Bo,[594] and Sh.[595] Another example of genetic polymorphism, comparable with Oz, has been reported[596] for the constant region of human λ-chains, involving a Ser/Gly interchange at position 154. The *N*-terminal sequences of 20 κ-chains from BALB/c mouse myelomas have been examined in the sequenator and arranged in nine sub-groups (Figure 13).[597] These sub-groups are felt to be sufficiently distinct to merit the postulation of a separate germ line gene for each sub-group.[597]

The *N*-terminal sequence of the amyloid protein from the liver of a patient suffering from amyloidosis has been shown to be Asp-Ile-Gln-Met-Thr-Gln-Ser-, which is sufficient to identify it as being related to the *N*-terminal region of κ-chains.[598] Since the mol. wt. of the protein is only 7500 it must be restricted entirely to the variable region of the light chain, but whether all amyloid proteins derive from light chains remains to be proved.

It is also interesting that in an examination of the attachment of carbohydrate to the variable region of human myeloma light chains, it became clear[599] that the carbohydrate was always attached to an asparagine residue in the sequence -Asn-X-Ser- or -Asn-X-Thr-, suggesting that this particular tripeptide sequence is recognised by the transglycosylase catalysing the attachment. The same conclusion has been reached by the

[588] (*a*) F. Franěk and F. Šorm, *Biochim. Biophys. Acta*, 1970, **221**, 593; (*b*) F. Franěk, *F.E.B.S. Letters*, 1970, **8**, 269.
[589] J. Novotný, F. Franěk, and F. Šorm, *European J. Biochem.*, 1970, **14**, 309; *ibid.*, **16**, 278.
[590] J. Novotný and F. Franěk, *F.E.B.S. Letters*, 1970, **9**, 33.
[591] B. Frangione and M. E. Lamm, *F.E.B.S. Letters*, 1970, **11**, 339.
[592] (*a*) L. Alescio-Zonta and C. Baglioni, *European J. Biochem.*, 1970, **15**, 450. (*b*) S. Watanabe and N. Hilschmann, *Z. physiol. Chem.*, 1970, **351**, 1291.
[593] T. Shinoda, K. Titani, and F. W. Putnam, *J. Biol. Chem.*, 1970, **245**, 4475.
[594] M. Wikler and F. W. Putnam, *J. Biol. Chem.*, 1970, **245**, 4488.
[595] K. Titani, M. Wikler, T. Shinoda, and F. W. Putnam, *J. Biol. Chem.*, 1970, **245**, 2171.
[596] M. Hess and N. Hilschmann, *Z. physiol. Chem.*, 1970, **351**, 67.
[597] L. E. Hood, M. Potter, and D. J. McKean, *Science*, 1970, **170**, 1207.
[598] G. G. Glenner, J. Harbaugh, J. I. Ohms, M. Harada, and P. Cuatrecasas, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 1287.
[599] H. C. Sox, jun., and L. Hood, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 975.

Subgroup	*Protein*	*Class*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
$V_{\kappa I}$	41	BJ	Asp	Ile	Gln	Met	Thr	Gln	Ser	Pro	Ser	Ser	Leu	Ser	Ala	Ser	Leu	Gly	Glu	Arg	Val	Ser	Leu	Thr	Cys
	173	IgG	Asp	Ile	Glx	Met	Thr	Gln	Thr	Thr	Ser	Ser	Leu	Ser	Ala	Ser	Leu	Gly	Asp	( )	Val	Thr	Ile		
	31c	IgF	Asp	Ile	Gln	Met	Thr	Gln	Ser	Pro	Ala	Ser	Leu	Ser	Ala	Ser	Val	Gly	Glu	Arg	Val	Thr	Ile	Thr	Cys
	23	IgF	Asp	Ile	Gln	Met	Thr	Gln	Ser	Pro	Ala	Ser	Leu	Ser	Val	Ser	Val	Gly	Glx	( )	Val	Thr	Ile		
	149	BJ	Asp	Ile	Gln	Met	Thr	Gln	Ser	Pro	Asp	Tyr	Leu	Ser	Ala	Ser	Val	Gly	Glu	Thr	Val	Thr	Ile	Thr	Cys
$V_{\kappa II}$	70	IgF	Asp	Ile	Val	Leu	Thr	Gln	Ser	Pro	Ala	Ser	Leu	Ala	Val	Ser	Leu	Gly	Gln	Arg	Ala	Thr	Ile	Ser	Cys
	321	BJ	Asp	Ile	Val	Leu	Thr	Gln	Ser	Pro	Ala	Ser	Leu	Ala	Val	Ser	Leu	Gly	Gln	Arg	Ala	Thr	Ile	Ser	Cys
	63	BJ	Asp	Ile	Val	Leu	Thr	Gln	Ser	Pro	Ala	Ser	Leu	Ala	Val	Ser	Leu	Gly	Gln	Arg	Ala	Thr	Ile	Ser	Cys
$V_{\kappa III}$	603	IgA	Asp	Ile	Val	Met	Thr	Glx	Ser	Pro	Ser	Ser	Leu	Ser	Val	Ser	Ala	Gly	Glx	( )	Val	Thr	Met	Ser	Cys
	870	IgA	Asp	Ile	Val	Met	Thr	Gln	Ser	Pro	Ser	Ser	Leu	Ser	Val	Ser	Ala	Gly	Glu	Lys	Val	Thr	Met	Ser	Cys
	384	IgA	Asp	Ile	Val	Met	Thr	Gln	Ser	Pro	Ser	Ser	Leu	Ser	Val	Ser	Ala	Gly	Glu	Lys	Val	Thr	Met	Ser	Cys
$V_{\kappa IV}$	467	IgA	Asp	Val	Leu	Met	Thr	Gln	Thr	Pro	Leu	Ser	Leu	Pro	Val	(Ser)	Leu	Gly	Asp	Glu	Ala	( )	Ile	(Ser)	Cys
	47	IgF	Glu	Val	Val	Met	Thr	Gln	Thr	Pro	Leu	Ser	Leu	Ala	Val	(Ser)	Leu	Gly	( )	Glx	Ala	(Ser)			
	37	IgH	Asp	Val	Leu	Met	Thr	Gln	Thr	Pro	Leu	( )	Leu	Pro	Val	(Ser)	Leu	Gly	Asp	Glu	Ala	(Ser)			
$V_{\kappa V}$	843	BJ	Asp	Val	Val	Met	Thr	Gln	Thr	Pro	Leu	Thr	Leu	Ser	Val	Thr	Ile	Gly	Glu	Pro	Ala	( )	Ile	Ser	Cys
	674	BJ	Asp	Val	Val	Met	Thr	Gln	Thr	Pro	Leu	Thr	Leu	Ser	Val	Thr	Ile	Gly	Glu	Pro	Ala	Ser	Leu	Ser	Cys
$V_{\kappa VI}$	8	IgA	Asp	Ile	Val	Met	Thr	Gln	Ser	Pro	Thr	Phe	Leu	Ala	Val	Thr	Ala	Ser	Lys	Lys	Val	Thr	Ile	Ser	Cys
	15	IgA	Asp	Ile	Val	Met	Thr	Gln	Ser	Pro	Thr	Phe	Leu	Ala	Val	Thr	Ala	Ser	Lys	Lys	Val	Thr	Ile	Ser	Cys
$V_{\kappa VII}$	773	BJ	Glu	Thr	Thr	Val	Thr	Gln	Ser	Pro	Ala	Ser	Leu	Ser	Met	Ala	Ile	Gly	Glu	Lys	Val	Thr	Ile	Ser	Cys
	265	BJ	Glu	Thr	Thr	Val	Thr	Gln	Ser	Pro	Ala	Ser	Leu	Ser	Met	Ala	Ile	Gly	Glu	Lys	Val	Thr	Ile	Ser	Cys
$V_{\kappa VIII}$	LPC-1	IgG	Asp	Ile	Val	Met	Thr	Gln	Ser	Pro	Ser	Ser	Met	Gln	Ala	Ser	Ile	Gly	Glu	Lys	Val	Thr	Ile	Ser	Cys
$V_{\kappa IX}$	167	IgA	Asp	Ile	Val	Ile	Thr	Gln	Asx	Glu	Leu	Ser	Asp	Pro	Val	Thr	Ser	Gly	Glu	( )	Val				
											Ala														
												Ser													
			Asp		Val	Met				Pro	Ser														
Pool				Ile			Thr	Gln	Ser			Leu	Leu												
			Glu		Gln	Leu				Thr	Leu														
												Phe													
											Phe														

Figure 13 *Arrangement of mouse κ-chains into variable region sub-groups* (*ref.* 597)

Editors of the 'Atlas' [600] from general analysis of published glycoprotein structures. They also point out that this sequence would have to be absent in other proteins if carbohydrate attachment is to be avoided in the presence of the enzyme.

B. Heavy Chains.—A study of a mouse myeloma protein (IgG2a, MOPC 173) is in progress [601] and a partial sequence of the variable region of the heavy chain has been reported.[602] The amino-acid sequences of the Fd fragments of two human $\gamma 1$ heavy chains (Daw and Cor) have been established and compared (Figure 14).[603] These results show clearly that the

```
Daw  Glp-Val-Thr-Leu-Arg-Glu-Ser-Gly-Pro-Ala-Leu-Val-Arg-Pro-
Cor  Glp-Val-Thr-Leu-Arg-Glu-Ser-Gly-Pro-Ala-Leu-Val-Lys-Pro-
     1                                   10

Daw  Thr-Gln-Thr-Leu-Thr-Leu-Thr-Cys-Thr-Phe-Ser-Gly-Phe-Ser-
Cor  Thr-Gln-Thr-Leu-Thr-Leu-Thr-Cys-Thr-Phe-Ser-Gly-Phe-Ser-
                          20

Daw  Leu-Ser-Gly-Glu-Thr-Met-Cys-Val-Ala-Trp-Ile-Arg-Gln-Pro-
Cor  Leu-Ser-Ser-Thr-Gly-Met-Cys-Val-Gly-Trp-Ile-Arg-Gln-Pro-
     30                                        40

Daw  Pro-Gly-Glu-Ala-Leu-Glu-Trp-Leu-Ala-Trp-Asp-Ile-Leu-Asn-
Cor  Pro-Gly-Lys-Gly-Leu-Glu-Trp-Leu-Ala-Arg-Ile-Asp-Trp-Asp-
                               50

Daw  Asp-Asp-Lys-Tyr-Tyr-Gly-Ala-Ser-Leu-Glu-Thr-Arg-Leu-Ala-
Cor  Asp-Asp-Lys-Tyr-Tyr-Asx-Thr-Ser-Leu-Glu-Thr-Arg-Leu-Thr-
                          |
                     carbohydrate
               60                                        70

Daw  Val-Ser-Lys-Asp-Thr-Ser-Lys-Asn-Gln-Val-Val-Leu-Ser-Met-
Cor  Ile-Ser-Lys-Asp-Thr-Ser-Arg-Asn-Gln-Val-Val-Leu-Thr-Met-
                               80

Daw  Asn-Thr-Val-Gly-Pro-Gly-Asp-Thr-Ala-Thr-Tyr-Tyr-Cys-Ala-
Cor  Asp-Pro-Val----------Asp-Thr-Ala-Thr-Tyr-Tyr-Cys-Ala-
                     90

Daw  Arg-Ser-Cys-Gly-Ser-Gln----------Tyr-Phe-Asp-Tyr-
Cor  Arg-Ile-Thr-Val-Ile-Pro-Ala-Pro-Ala-Gly-Tyr-Met-Asp-Val-
     100

Daw  Trp-Gly-Gln-Gly-Ile-Leu-Val-Thr-Val-Ser-Ser-Ala-Ser-Thr-
Cor  Trp-Gly-Arg-Gly-Thr-Pro
     110                  115                  120
```

[600] L. T. Hunt and M. O. Dayhoff, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 757.
[601] A. Bourgois and M. Fougereau, *European J. Biochem.*, 1970, **12**, 558.
[602] A. Bourgois and M. Fougereau, *F.E.B.S. Letters*, 1970, **8**, 265.
[603] E. M. Press and N. M. Hogg, *Biochem. J.*, 1970, **117**, 641.

Daw Lys-Gly-Pro-Ser-Val-Phe-Pro-Leu-Ala-Pro-Ser-Ser-Lys-Ser-
Cor
130

Daw Thr-Ser-Gly-Gly-Thr-Ala-Ala-Leu-Gly-Cys-Leu-Val-Lys-Asp-
Cor
140 **150**

Daw Tyr-Phe-Pro-Glu-Pro-Val-Thr-Val-Ser-Trp-Asn-Ser-Gly-Ala-
Cor
160

Daw Leu-Thr-Ser-Gly-Val-His-Thr-Phe-Pro-Ala-Val-Leu-Gln-Ser-
Cor
170

Daw Ser-Gly-Leu-Tyr-Ser-Leu-Ser-Ser-Val-Val-Thr-Val-Pro-Ser-
Cor
180 **190**

Daw Ser-Leu-Gly-Thr-Gln-Thr-Tyr-Ile-Cys-Asn-Val-Asn-His-(Lys-
Cor
200

Daw Pro-Ser)-Asn-Thr-Lys-Val-Asp-Lys-Lys-Val-Glu-Pro-Lys-Ser-
Cor
210 **220**

Daw Cys-Asp-Lys-Thr-His
Cor
225

Figure 14 *Amino-acid sequences of the Fd regions of two human γ1 heavy chains (Daw and Cor) (ref. 603). The numbering used is that of the protein Daw*

variable region is similar in length to that of a light chain (*ca.* 115 residues) and the existence of heavy chain variable region sub-groups is deduced [603] by comparison of these sequences with partial sequences published for other γ-chains. Moreover, these sub-groups are found associated with the constant regions of α- and μ-chains, as well as those of γ-chains, of various sub-classes.[602–604] This is in striking contrast with the variable region sub-groups of light chains, where κ-chain variable regions are never found associated with λ-chain constant regions, nor *vice-versa*. In this connexion it may be relevant that the genes that code for κ- and λ-chains are not closely linked (at least in the rabbit) whereas those that code for the heavy chain sub-classes are. However, there appears so far to be no restriction on the combination of light and heavy chain variable region sub-groups in antibody molecules.[604a, 604b, 605]

[604] (*a*) H. Köhler, A. Shimizu, C. Paul, V. Moore, and F. W. Putnam, *Nature*, 1970, **227**, 1318. (*b*) A. C. Wang, J. R. L. Pink, H. H. Fudenberg, and J. Ohms, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 657. (*c*) H. Ponstingl, J. Schwarz, W. Reichel, and N. Hilschmann, *Z. physiol. Chem.*, 1970, **351**, 1591; S. E. Pitcher and W. Konigsberg, *J. Biol. Chem.*, 1970, **245**, 1267.

[605] H. Köhler, A. Shimazu, C. Paul, and F. W. Putnam, *Science*, 1970, **169**, 56.

The partial sequence of the constant region of a human γ4-chain has been reported and shown to be very similar to the γ1-chain and rabbit Fc region, indicating the recent evolutionary origin of the γ-chain sub-classes.[606] The α1- and α2-chains of IgA1 and IgA2 appear to differ in size[607] and the mol. wts. of the heavy and light chains of Waldenström macroglobulins are consistent[608] with the view that IgM contains five sub-units, each of which contains two heavy and two light chains. A similar pentameric structure is entertained for sting ray immunoglobulins[609] which supports the idea that only antibodies of the IgM class occur in primitive vertebrates. A study of the cleavage of human μ-chains with cyanogen bromide has been described[610] and the *N*-terminal sequences of some horse and human γ-chains have been reported.[611]

The site of carbohydrate attachment in the light and heavy chains of several human IgG1 myeloma proteins has been investigated.[612]

C. Disulphide Bridges.—A study has been made[613] of the selective reduction of disulphide bonds in human IgGM with 2-mercaptopyridine; the heavy and light chains of IgGA2 can be separated as stable disulphide-bridged dimers (H—H and L—L) without reduction since the molecules lack disulphide bridges between heavy and light chains.[607] An unusual component, J, which appears to function by linking the monomeric units that comprise the IgA dimer, has been described.[614] It is synthesised in the same plasma cell as the heavy and light chains, has a mol. wt. of approximately 23 000 and is linked by disulphide bridges to the α-chain.

It has now been proved[615] that the disulphide bridge at the *C*-terminus of the μ-chains of IgM is an intra sub-unit bridge and does not link sub-units together. That linkage is a disulphide bridge between cysteine residues in the heavy chain but in the sequence:[616]

-Ser-Ala-Val-Gly-Glu-Ala-Ser-Ile-Cys-Glu-Asp-Asn-Asn-Trp-

The inter-chain disulphide bridges of mouse IgG1 have been examined,[617] and the inter-chain bonds of human and mouse IgG have been discussed in detail.[618] Mouse and human IgG1 are so far unique in having the linkage

[606] J. R. L. Pink, S. H. Buttery, G. M. de Vries, and C. Milstein, *Biochem. J.*, 1970, **117**, 33.
[607] K. J. Dorrington and J. H. Rockey, *Biochim. Biophys. Acta*, 1970, **200**, 584.
[608] P. Johnson and J. N. Miller, *Biochim. Biophys. Acta*, 1970, **207**, 297.
[609] J. J. Marchalonis and S. A. Schonfield, *Biochim. Biophys. Acta*, 1970, **221**, 604.
[610] E. Razafimahaleo and R. Bourrillon, *Biochim. Biophys. Acta*, 1970, **200**, 89.
[611] P. C. Montgomery, A. C. Bello, and J. H. Rockey, *Biochim. Biophys. Acta*, 1970, **200**, 258.
[612] H. L. Spiegelberg, C. A. Abel, B. G. Fishkin, and H. M. Grey, *Biochemistry*, 1970, **9**, 4217.
[613] F. Dolder, *Biochim. Biophys. Acta*, 1970, **207**, 286.
[614] M. S. Halpern and M. E. Koshland, *Nature*, 1970, **228**, 1276.
[615] D. Beale and A. Feinstein, *F.E.B.S. Letters*, 1970, **7**, 175.
[616] S. I. Miekka and H. F. Deutsch, *J. Biol. Chem.*, 1970, **245**, 5534.
[617] J. Svasti and C. Milstein, *Nature*, 1970, **228**, 932.
[618] C. de Préval, J. R. L. Pink, and C. Milstein, *Nature*, 1970, **228**, 930.

between light and heavy chains only a few residues away from the hinge region: in all other sub-classes of human and mouse IgG the linkage is more than 90 residues nearer the *N*-terminus of the heavy chain. The rare human IgD has only one disulphide bridge between the two heavy (δ-) chains, which is unique for human immunoglobulins.[619, 620] The light chain is linked to the δ-chain through a disulphide bridge which shows marked homology with other immunoglobulin molecules,[620] as shown in the following sequences for this region of heavy chains:

Human γ2, γ3, and γ4	-Pro-Leu-Ala-Pro-Cys-Ser- Arg-
Human IgD	-Pro-Ile- Ile- Ser- Gly-Cys-Arg-
Mouse γ2b	-Pro-Leu-Ala-Pro-Gly-Cys-Gly-
Rabbit	-Pro-Leu-Ala-Pro-Cys-Cys-Gly-

The single inter-heavy chain bridge is in the sequence:[620]

-Thr-Pro-Glu-Cys-Pro-Ser-His-Thr-Gln-Pro-Leu-Gly-Val-
|
-Thr-Pro-Glu-Cys-Pro-Ser-His-Thr-Gln-Pro-Leu-Gly-Val-

The complete disulphide bridge arrangement in the rabbit IgG molecule has now been worked out (Figure 15).[621] As with human IgD, a single disulphide bridge links the heavy chains. The sequences around the inter-chain bridges are as follows:[621]

(Light chain)
-Asn-Arg-Gly-Asp-Cys
(Heavy chain) |
-Pro-Leu-Ala-Pro-Cys-Cys-Gly-Asp-Thr-Pro-Ser-Ser-Thr——
130 **140**

——Thr-Val-Ala-Pro-Ser-Thr-Cys-Ser-Lys-Pro-Met/Thr-Cys-Pro-Pro-Pro-Glu-Leu-
215 |
Heavy chain

The pattern of four intrachain disulphide loops of approximately 60 residues each appears to be common to most γ-chains. However, variations on this basic theme are permitted. For example, an additional intrachain disulphide bridge is found in the *N*-terminal half of rabbit γ-chains[621] (Figure 15) and in the variable region of the human γ1-chain Daw.[603] It is interesting that in the human γ1-chain Cor one-half of this disulphide bridge remains (Cys-35), whereas the other half (Cys-101) is replaced by threonine (Figure 14).[603]

D. Antibody Binding Sites.—Further evidence that antibody specificity resides in the variable regions of the light and heavy chains has been provided by showing[622] that there are differences in amino-acid composition

[619] H. L. Spiegelberg, J. W. Prahl, and H. M. Grey, *Biochemistry*, 1970, **9**, 2115.
[620] M. B. Perry and C. Milstein, *Nature*, 1970, **228**, 934.
[621] I. J. O'Donnell, B. Frangione, and R. R. Porter, *Biochem. J.*, 1970, **116**, 261.
[622] M. E. Koshland, P. Ochoa, and N. J. Fujita, *Biochemistry*, 1970, **9**, 1880.

between antibodies raised against the very similar haptens *p*-azophenyl arsonate and *p*-azophenyl phosphonate. Some preliminary sequence studies of horse antihapten antibodies have been reported [623] and several hopeful steps have been taken in the direction of chemically homogeneous antibodies. Thus, the *N*-terminal sequence of light chains from a rabbit antibody against *p*-azobenzenearsonate shows much less heterogeneity than is found with pooled normal light chains [624] and electrophoretically homogeneous anti-Dnp antibodies have been elicited in mice by using as

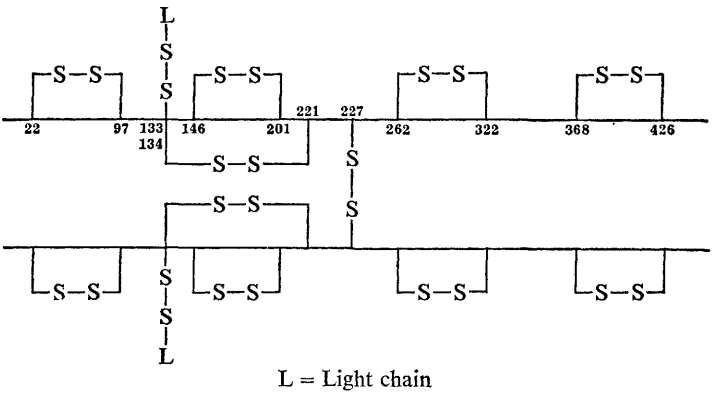


Figure 15 *The inter- and intra-chain disulphide bridges in the heavy chain of rabbit IgG (ref. 621). The numbering used is that of the human protein Daw (ref. 603) and ignores deletions postulated in the rabbit chain in order to maximise the homology (ref. 621)*

immunogen papain which has been reacted at Cys-25 with N^{α}-iodoacetyl-N^{ε}-2,4-dinitrophenyl-L-lysine.[625] A structurally homogeneous antibody from rabbit anti-pneumococcal antisera has also been reported.[626] The *N*-terminal sequence of the light chain was found to be Asp-Val-Val-Met-Thr-Glx-Thr-Pro-Ala-Thr-Val-.

Affinity labelling of a mouse IgA myeloma protein (MOPC 315) with anti-Dnp activity is yielding interesting results, again with the advantage that one is dealing with a chemically homogeneous protein. If the bromoacetyl derivative of the hapten Dnp-ethylene diamine is used to label the 'antibody', exclusive reaction with tyrosine residues in the light chains is observed.[627] If, however, the bromoacetyl derivative of N^{ε}-Dnp-lysine is used, the labelling of lysine residues exclusively in the heavy chain is noted, suggesting that the heavy and light chains both contribute to the antibody

[623] J. H. Rockey, P. C. Montgomery, and K. J. Dorrington, *Biochemistry*, 1970, **9**, 4310.
[624] K. J. Fraser and P. Edman, *F.E.B.S. Letters*, 1970, **7**, 99.
[625] G. N. Trump and S. J. Singer, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 411.
[626] J.-C. Jaton, M. D. Waterfield, M. N. Margolies, and E. Haber, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 959.
[627] J. Haimovich, D. Givol, and H. N. Eisen, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 1656.

binding site.[627] Moreover, affinity labelling of the same mouse myeloma protein with *m*-nitrobenzenediazonium fluoroborate gives unique reaction with a tyrosine in the light chain in the sequence:[628]

-Ser-Asn-Thr-Gly-Ala-Val-Thr-Thr-Ser-Asp-TYR-(Ala, Ser)-Trp-Ile-Glu-Glu-Pro-Asp-Lys-His-Leu-Phe-Thr-Gly-Leu-Ile-Gly-Gly-Thr-Ser-Asn-Arg-

Comparison with other mouse light chain sequences suggests[628] that this is a λ-chain (albeit a slightly unusual one) and that the reactive tyrosine residue is at position 34. There is also reason to hope that *X*-ray crystallographic analysis of Bence Jones proteins will be possible.[629]

E. Some Implications for Biosynthesis.—A detailed account of current theories of antibody biosynthesis is beyond the scope of this Report and the reader is referred to recent reviews.[582] However, a few separate points do, perhaps, merit abbreviated consideration here. For example, there seems to be general accord that the variable and constant regions of immunoglobulin polypeptide chains require separate genes to code for them ('two genes – one polypeptide chain') and that these genes are integrated by some form of somatic translocation. Studies of the heavy chain disease proteins Zu[630a] and Hi[630b] show that they contain internal deletions of different size that overlap the junction of constant and variable regions, making it almost certain, therefore, that fusion of the genetic information for V- and C-regions occurs at the level of DNA. Further, in the heavy chains of rabbit IgG, the *a* allotypic specificity (correlating with sequence changes in the *N*-terminal region) is inherited from the same parent as the A11 or A12 allotypic specificity (correlating with sequence changes in the constant region), suggesting that the genes for the variable and constant regions must be on the same chromosome of one parent.[631]

However, while there is also general agreement that there must be at least one gene for each variable region sub-group, there are differences of opinion as to exactly how many there are. Some[632] favour a germ-line theory, with a separate gene for each variable region, involving gene duplication on a massive scale.[633] Others[634] feel that a limited number of genes is more likely, with further variability introduced by some mechanism of somatic mutation. In support of this latter view, it seems likely that allotypic (allelic) variations occur in the variable region of rabbit heavy

[628] E. J. Goetzel and H. Metzger, *Biochemistry*, 1970, **9**, 3862.
[629] (*a*) W. H. Palm, *F.E.B.S. Letters*, 1970, **10**, 46. (*b*) A. Solomon, C. L. McLaughlin, C. H. Wei, and J. R. Einstein, *J. Biol. Chem.*, 1970, **245**, 5289.
[630] (*a*) B. Frangione and C. Milstein, *Nature*, 1969, **224**, 597. (*b*) W. D. Terry and J. Ohms, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 558.
[631] T. J. Kindt, W. J. Mandy, and C. W. Todd, *Biochemistry*, 1970, **9**, 2028.
[632] L. Hood and D. W. Talmage, *Science*, 1970, **168**, 325.
[633] L. Hood, K. Eichmann, H. Lackland, R. M. Krause, and J. J. Ohms, *Nature*, 1970, **228**, 1040.
[634] (*a*) J. A. Gally and G. M. Edelman, *Nature*, 1970, 227, 341. (*b*) M. G. Weigert, I. M. Cesari, S. J. Yonkovich, and M. Cohn, *Nature*, 1970, **228**, 1045. (*c*) G. M. Edelman, *Biochemistry*, 1970, **9**, 3197.

chains [635] and mouse light chains.[636] Since allotypic markers segregate in classical Mendelian fashion, this is difficult to reconcile with a germ-line theory that requires a large number of variable region genes. 'Who shall decide, when doctors disagree?' asked Pope. More experiments and lapse of time would seem to be the answer.

7 Membrane Proteins

'. . . brevity can never . . . do justice to all the facts of a complex situation. On such a theme one can be brief only by omission and simplification.'[637] Some recent advances in protein chemical technique, *e.g.* SDS-gel electrophoresis and gel-filtration in denaturing solvents, are now enabling membrane proteins to be more easily separated and characterised. It is this aspect only, the character of proteins in membranes, rather than any discussion of their various roles, that falls within the scope of this chapter.

Membranes of various sorts are under study: erythrocyte membranes (presumably because of ease of isolation and purification), membranes of mitochondria and chloroplasts, endoplasmic and sarcoplasmic reticulum, and bacterial membranes. In summary of previous work on the characterisation of membrane proteins, suffice it to say that in 1961 Green and his co-workers [638] postulated that mitochondrial membranes contained a simple homogeneous structural protein which accounted for 30% of the total mitochondrial protein, and that much of the literature in this field since then has been concerned with the search for, and denial of, 'structural proteins'. The case for structural proteins has been reviewed.[639, 640] The case against is expressed in most of the papers cited in the remainder of this section: evidence is accumulating that there is no membrane component with the simple character ascribed to 'structural protein'. The properties and function of proteins in excitable membranes have been reviewed.[641] Myelin A1 protein has been mentioned earlier (Section 4).

Broadly speaking, there are two approaches to the study of membrane proteins and there is value in both. In the first, the whole mixture of solubilised proteins is fractionated for further study of the various components: in the second, a specific component of the membrane is labelled with a specific 'active-site-directed' reagent which then provides a 'handle' for further purification of this particular protein. A comparable method allows proteins on the 'outside' of membranes to be selectively labelled and identified.

[635] J. M. Wilkinson, *Biochem. J.*, 1969, **112**, 173.
[636] G. M. Edelman and P. D. Gottlieb, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 1192.
[637] A. Huxley, foreword to 'Brave New World Revisited,' Harpers, New York, 1958.
[638] D. E. Green, H. D. Tisdale, R. S. Criddle, P. Y. Chen, and R. M. Bock, *Biochem. Biophys. Res. Comm.*, 1961, **5**, 109.
[639] R. S. Criddle, *Ann. Rev. Plant Physiol.*, 1969, **20**, 239.
[640] L. Rothfried and A. Finkelstein, *Ann. Rev. Biochem.*, 1968, **37**, 463.
[641] D. Nachmansohn, *Science*, 1970, **168**, 1059.

A. Red Blood Cell Membranes.—The heterogeneity of proteins in these membranes is now well established.[642, 643] Differential extraction under various conditions gave at least 12 different membrane proteins with mol. wts. in the range 10 000—150 000.[642b] A recent report [643] describes the analysis by SDS-gel electrophoresis of an unfractionated membrane preparation; this procedure, which does not involve prior removal of lipids, ensures that no membrane protein is lost before analysis, thereby escaping detection. Fourteen classes of protein, judged by mol. wt., were observed. Four major protein components, together accounting for 60—65% of the total membrane protein, had mol. wts. of 255 000, 240 000, 108 000, and 86 000; only that of 108 000 contained substantial amounts of carbohydrate.

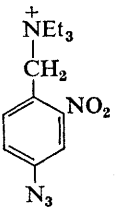
There have been several other reports that erythrocyte membranes contain large proteins. Gel-filtration under strongly reducing and denaturing conditions [644] (see Section 2E) gave the mol. wt. of one component of human erythrocytes as 200 000; for this size of protein it was necessary to use 4% agarose rather than the usual 6%. A protein (spectrin) with mol. wt. 140 000 in denaturing conditions, comprising 20% of the membrane protein, has been isolated [645] from the erythrocytes of several species; it is interesting that this was assigned the rôle of an 'inner structural protein'.

At the other end of the size scale come the controversial 'mini-proteins' (mol. wt. 5000) of Dreyer and his co-workers.[646] They report that as much as 50% of the protein is in this form in the membranes of red blood cells and mitochondria, and in the rhodopsin-containing membranes of the outer segments of retinal rods. No *N*-terminal amino-acids could be detected (implying that the peptides are blocked or cyclic) and the proteins were claimed to aggregate to give complexes that could not be dissociated by any known means. [Note added in proof: Dreyer has withdrawn his suggestion that 'mini-proteins' are the basic structural units of membranes; the mini-proteins turned out to contain much phosphorus and little amino-acid (*Nature New Biol.*, 1971, **231**, 227).]

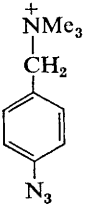
While the fractionation of membrane proteins continues, affinity-labelling of proteins in the intact membrane brings its rewards; the usual fractionation procedures are applied once the label has been specifically attached. Acetylcholinesterase in red blood cell membranes was labelled [647] with tritiated DFP, having been protected with the substrate analogue butyryl choline while other accessible sites were reacted with non-radio-

[642] (*a*) S. A. Rosenberg and G. Guidotti, *J. Biol. Chem.*, 1968, **243**, 1985. (*b*) S. A. Rosenberg and G. Guidotti, *J. Biol. Chem.*, 1969, **244**, 5118.
[643] J. Lenard, *Biochemistry*, 1970, **9**, 1129.
[644] J. T. Gwynne and C. Tanford, *J. Biol. Chem.*, 1970, **245**, 3269.
[645] T. W. Tillack, S. L. Marchesi, V. T. Marchesi, and E. Steers, jun., *Biochim. Biophys. Acta*, 1970, **200**, 125; S. L. Marchesi, E. Steers, V. T. Marchesi, and T. W. Tillack, *Biochemistry*, 1970, **9**, 50.
[646] M. T. Laico, E. I. Rvoslahti, D. S. Papermaster, and W. J. Dreyer, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 120.
[647] M. B. Bellhorn, O. O. Blumenfeld, and P. M. Gallop, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 267.

active reagent. SDS-Gels run under normal reducing conditions showed a radioactive peak corresponding to a mol. wt. *ca.* 90 000; in the absence of 2-mercaptoethanol, however, the mol. wt. was 180 000, implying the possible presence of disulphide-bridged dimers. The point is made [647] (and this is a general one where denatured proteins are the species examined) that there may be other components of acetylcholinesterase that do not become labelled. Acetylcholinesterase in excitable membranes has been reviewed.[641] The elegant technique of photo-affinity labelling with aryl azides, successfully used in affinity labelling of antibodies [648] (see Vol. 2 of these Reports, p. 65), has now been used to label membrane components.[649] Using the quaternary ammonium aryl azides (6) and (7), which are analogues



(6)



(7)

of acetylcholine, it was possible to label the acetylcholinesterase of intact red blood cell membranes and the acetylcholine receptor at the neuromuscular junction of frog sartorius muscle.[649] Specific labelling of the sodium–potassium ATPase of human red blood membranes has been accomplished [650] with tritiated ouabain (an inhibitor of active transport). When the membrane was solubilised the ouabain–membrane component and the sodium–potassium ATPase appeared to be identical; the radioactive label can thus be used to follow the purification of the ATPase.

Another popular approach is to study the distribution of proteins between the inside and outside of the erythrocyte membrane. The *p*-diazonium salt of [^{35}S]sulphanilic acid labelled predominantly a membrane protein of mol. wt. 140 000 as measured in SDS-gels;[651] this was therefore assigned to the outside of the membrane. Among the proteins that were not labelled were two with mol. wt. around 300 000 and 270 000. A danger in using small molecules to investigate selectively the outside of membranes is that the membrane may not be wholly impermeable to them, and that partial labelling of the inside will result. A very elegant method for iodinating outer proteins has been described:[652] the enzyme lactoperoxidase

[648] G. Fleet, J. R. Knowles, and R. R. Porter, *Nature*, 1969, **224**, 511.
[649] H. Kiefer, J. Lindstrom, E. S. Lennox, and S. J. Singer, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 1688.
[650] P. B. Dunham and J. F. Hoffman, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 936.
[651] H. C. Berg, *Biochim. Biophys. Acta*, 1969, **183**, 65.
[652] D. R. Phillips and M. Morrison, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 284.

(mol. wt. 78 000) is used to effect iodination with radioactive iodine, generated *in situ* by the enzyme from radioactive potassium iodide, under the influence of hydrogen peroxide. Iodination occurs within the enzyme–substrate (membrane) complex, so that there is no danger of penetration of the membrane. When human erythrocytes were labelled in this way, analysis of the membrane proteins on SDS-gels showed only one labelled band (mol. wt. 90 000).

Cross-linking of erythrocyte membranes with the bifunctional imido-ester dimethyl adipimidate appeared to have been successful,[653] although the products were not fractionated. While this approach is clearly capable of yielding considerable information about the arrangement of molecules within the membrane, it is probably unreasonable to expect any startling results until the protein components are more fully characterised.

B. Mitochondrial and Other Membranes.—The reader is reminded of putative structural proteins (mol. wt. in the region of 22 000) in mitochondrial membranes.[639] A study [654] of the proteins of the inner and outer mitochondrial membranes of rat liver, and of the rough and smooth microsomal membranes, showed them all to be highly heterogeneous. The major protein in each case had a mol. wt. in the range 50 000—68 000 and did not represent more than 15% of the total protein. The outer mitochondrial and smooth microsomal membranes appeared to have at least three proteins in common on SDS-gels, and it was speculated that they might perhaps have a common origin.[654] In another study [655] the major protein component of mitrochondrial membranes is reported to have a mol. wt. of 57 000. A protein of mol. wt. *ca.* 56 000 (25% of the total protein) has been extensively purified from the inner membrane of rat liver mitochondria and has already been subjected to peptide mapping.[656] Further structural work is awaited with interest.

A protein fraction comprising about 18% of the total microsomal fraction has been isolated from rat liver endoplasmic reticulum.[657] It appeared to have no *N*-terminal amino-acids. *C*-Terminal analysis and sedimentation studies suggested that it consisted of a group of closely related proteins. The ATPase of sarcoplasmic reticulum has been identified [658] as a single band in SDS-gels by selective labelling of the active thiol group of the protein with ^{14}C-labelled *N*-ethylmaleimide, after reacting other accessible thiol groups with unlabelled reagent in the presence of substrate (ATP) to protect the enzyme. It is reported [659] that more than 95% of the protein of sarcotubular membranes from rabbit skeletal muscle has mol. wt. 6500—10 000 and consists of identical or very similar molecules; *N*-terminal Gly

[653] W. G. Niehaus, jun., and F. Wold, *Biochim. Biophys. Acta*, 1970, **196**, 170.
[654] C. A. Schnaitman, *Proc. Nat. Acad. Sci. U.S.A.*, 1969, **63**, 412.
[655] P. J. Curtis, *Biochim. Biophys. Acta*, 1970, **211**, 575.
[656] J. H. J. Tsai, H. Tsai, and G. von Ehrenstein, *F.E.B.S. Letters*, 1970, **10**, 277.
[657] D. Kaplan, *Biochim. Biophys. Acta*, 1970, **211**, 396.
[658] R. Panet and Z. Seliger, *European J. Biochem.*, 1970, **14**, 440.
[659] B. P. Yu and E. J. Masoro, *Biochemistry*, 1970, **9**, 2909.

and *C*-terminal Ala were determined. It was suggested that it had a rôle in calcium transport.[659] These proteins are not quite as small as the 'mini-proteins' mentioned earlier but their existence again appears to be disputed.[660]

An interesting observation has been made on the composition of a bacterial membrane.[661] Ghosts from *Bacillus* PP, isolated from a culture of *B. megatherium* KM (and assumed to be a mutant of it), were found to contain a protein with mol. wt. of 30 000 (SDS-gels) representing 90% of the total membrane protein. The ghosts of the presumed parent *B. megatherium* KM, however, contained an electrophoretically identical protein as a minor component only. In this instance at least, there appears to be a case for 'structural' proteins.

8 Chemical Modification

Many examples of chemical modification have already arisen in other sections. In this section the aim will be to summarise some of the recent uses to which group-specific reagents have been put and to pay some attention to active-site-directed inhibitors and the preparation of modified proteins for *X*-ray crystallography. In a very useful review Stark[662] tabulates reagents used for modifications of particular types of functional groups in proteins, with comments on stability, reversibility, and side-reactions where these are known. No attempt will be made here to duplicate this material. Another review[82] in the same volume has already been mentioned. The burden is also lessened by the appearance of an excellent Symposium volume[340] dealing with the reactivity of functional groups in proteins. Modification of single residues in proteins, using active-site-directed inhibitors, has been ably reviewed[663] by Shaw. An interesting article[664] discusses the chemical probing of enzyme–substrate complexes as a means of detecting intermediates with particular chemical reactivities.

A. Amino-groups.—*Anhydrides.* The reversible blocking of amino-groups to restrict tryptic attack during enzymic digestion has already been discussed (Section 2D). A careful evaluation[665] of some reversible amino-blocking reagents (tetrafluorosuccinic, maleic, and citraconic anhydrides, and diketen) has now been reported with lysozyme (no complicating thiol groups) as the test system. Only citraconic anhydride was wholly satisfactory; tetrafluorosuccinic anhydride and diketen were particularly unsatisfactory in most respects. None of the reagents was absolutely specific for amino-groups, but the reversibility of the side-reactions (apparently with hydroxy side-chains) ensured a homogeneous deblocked

[660] A. Martonosi, *Biochem. Biophys. Res. Comm.*, 1969, **36**, 1039.
[661] P. H. Patterson and W. J. Lennarz, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 408.
[662] G. R. Stark, *Advances in Protein Chemistry*, 1970, **24**, 261.
[663] E. Shaw, *Physiol. Rev.*, 1970, **50**, 244.
[664] Ph. Christen, *Experientia*, 1970, **26**, 1.
[665] A. F. S. A. Habeeb and M. Z. Atassi, *Biochemistry*, 1970, **9**, 4939.

product from citraconylated and, to a lesser extent, maleylated material. The enzyme regenerated after citraconylation showed complete regain of activity and immunological properties. Another study[666a] demonstrated the expansion in protein structure that accompanies the introduction of negative charge through maleylation, citraconylation, or succinylation; it was again noted that reaction with citraconic anhydride was easily reversed. A study[666b] of the limited digestion of citraconylated bovine serum albumin with α-chymotrypsin has been reported, and the effects of succinylation and maleylation on ATP creatine phosphotransferase have been described.[667] Schachman and his co-workers report[668] some interesting studies of the effects of succinylation on multimeric proteins and describe the hybridisation of modified and native subunits. When glyceraldehyde 3-phosphate dehydrogenase[668a] was examined in the ultracentrifuge after succinylation, dimers (but not monomers) were detected, further support perhaps for a pairing of dimers in the native enzyme. Succinylation of aspartate transcarbamylase[668b] produced inactive catalytic subunits after about 50% modification of amino-groups. These combined normally with the regulatory subunits and there appeared to be little unfolding of the protein chain. Hybridisation of the native and succinylated catalytic subunits suggested that the catalytic subunit exists as an unprecedented trimer.

Other Reagents. In contrast with the large net gain of negative charge from reaction with the anhydrides, reaction of amino-groups with imidoesters preserves the *status quo*. These reagents continue to find occasional use for lysine-blocking in primary structural work.[505a] Modification of about 50 of the 60 amino-groups in horse liver alcohol dehydrogenase with methyl picolinimidate gives a 19-fold increase in enzyme activity,[669] whereas the activation is only two-fold if modification is carried out in the presence of substrate. It was suggested that activation might be due to increased affinity for the nicotinamide portion of the coenzyme NAD (presumably through interaction with the heterocyclic ring of the reagent). Some examples of the use of bifunctional reagents for cross-linking amino-groups have already been mentioned (pp. 52, 105; see also p. 126). The *N*-terminal valine residues of human carbonmonoxyhaemoglobin have been cross-linked with *pp'*-difluoro-*mm'*-dinitrophenylsulphone.[670] The modified haemoglobin has high oxygen affinity and no haem–haem interaction.

An interesting study[671] of the guanidination of lactic dehydrogenase suggests that the guanidinium groups interact with the carboxylate groups

[666] (*a*) A. Jonas and G. Weber, *Biochemistry*, 1970, **9**, 4729. (*b*) A. Jonas and G. Weber, *Biochemistry*, 1970, **9**, 5092.
[667] M. C. Grant-Greene and F. Friedberg, *Internat. J. Protein Res.*, 1970, **2**, 235.
[668] (*a*) E. A. Meighen and H. K. Schachman, *Biochemistry*, 1970, **9**, 1177. (*b*) E. A. Meighen, V. Pigiet, and H. K. Schachman, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **65**, 234.
[669] B. V. Plapp, *J. Biol. Chem.*, 1970, **245**, 1727.
[670] R. M. Macleod and R. J. Hill, *J. Biol. Chem.*, 1970, **245**, 4875.
[671] P. C. Yang and G. W. Schwert, *J. Biol. Chem.*, 1970, **245**, 4886.

of substrates and inhibitors and stabilise the enzyme–coenzyme–inhibitor (or substrate) ternary complex. Inactivation of the enzyme after guanidination is attributed to the reduced rate of dissociation of substrates. A slightly altered charge distribution on the surface of the protein when homoarginine replaces lysine could account for the stronger interactions.[671] The effect of pH, temperature, and urea on the kinetics of the reaction of amino groups with trinitrobenzenesulphonic acid (TNBS) (see last year's Report, p. 71) has now been studied for human serum albumin.[672] Lysine residues in lysozyme were modified with pyridoxal phosphate followed by reduction, to investigate the role of lysine residues in the enzyme.[673] The modification of single lysine residues in heavy meromyosin[674] and in lysozyme[675] with dansyl chloride has been described. The dansyl group is potentially useful as a reporter group.

Introduction of Thiol Groups by Substitution at Amino-groups. The desire to introduce (extra) thiol groups into proteins stems largely from a wish to assist the crystallographer in preparing isomorphous heavy atom derivatives of proteins for *X*-ray analysis, by providing points of attachment for mercurials. While chemical tampering may not be ideal for the task in hand, it is possible that it will prove useful in some cases, in particular where diffusion of heavy atoms into crystals by soaking fails.

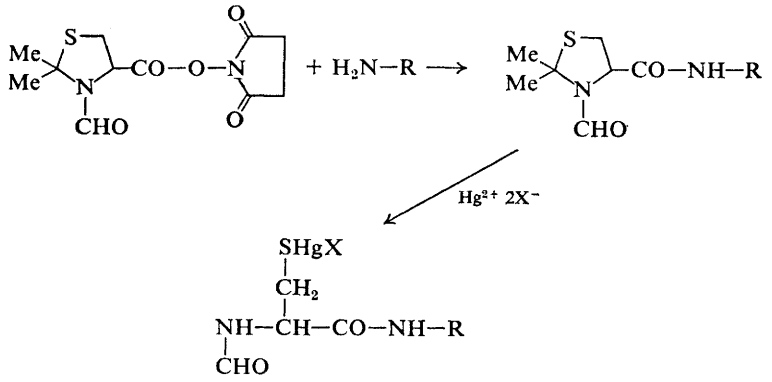
Stark has reviewed[662] two methods that can be used for the introduction of thiol-groups: reaction with thiolactones and cyclic anhydrides. In the former case the thiol group is generated directly by nucleophilic attack of the amino-group on the thiolactone ring; in the latter a thiolester substituent can be cleaved after acylation of the amino-group has occurred. The reaction of *N*-acetylhomocysteine thiolactone with ribonuclease was described in last year's Report (p. 72). It is of interest that active *N*-acetyl homocysteinyl ribonuclease was inhibited by silver ions[676] and there is evidence that inhibition of this 'thiol enzyme' is due to interaction of silver ions with a histidine residue in the protein.

A site for the attachment of heavy metals has been introduced into insulin[677] by acylation with the *N*-hydroxysuccinimide ester of 2,2′-dimethyl-3-formyl-L-thiazolidine-4-carboxylic acid. Ring-opening on treatment with mercuric ions generates the thiomercurial directly (Scheme 9). The three possible mono-substituted thiazolidine derivatives of insulin could be separated. Unfortunately, however, they were not the answer to the crystallographer's prayer: only one of the three crystallised and this disintegrated on exposure to mercury ions. Acylation with the hydroxysuccinimide ester of chloromercuriacetic acid[678] (8) might prove useful in

[672] A. R. Goldfarb, *Biochim. Biophys. Acta*, 1970, **200**, 1.
[673] J. E. Churchich and R. Irwin, *Biochim. Biophys. Acta*, 1970, **214**, 157.
[674] H. Rakashina, *Biochim. Biophys. Acta*, 1970, **200**, 319.
[675] R. F. Chen, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 1117.
[676] D. W. Hough and S. Shall, *F.E.B.S. Letters*, 1970, **8**, 243.
[677] D. G. Lindsay and S. Shall, *European J. Biochem.*, 1970, **15**, 547.
[678] G. Fölsch, *Acta Chem. Scand.*, 1970, **24**, 1115.

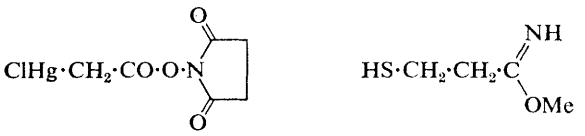
Table 2 (*cont.*)

Enzyme	*Source*	*Molecular weight*	*No. of subunits*
Aspartokinase[i]	*B. polymyxa*	116 000	(?) $4(\alpha_2\beta_2)$
ATP:arginine phosphotransferase[j]	Lobster, crab	40 000	1
ATP:creatine phosphotransferase[k]	Rabbit muscle	84 000	2
ATP:lombricine phosphotransferase[j]	Earthworm	80 000	2
ATP:taurocyamine phosphotransferase[j]	Arenicola	80 000	2
Avidin[l]	Hen egg-white	66 000	4
Chorismate mutase—prephenate dehydrogenase[m]	*Aer. aerogenes*	80 000	2
Chymotrypsin inhibitor[n]	Potatoes	39 000	4
Citrate synthase[o]	Pig heart	100 000	2
Cysteamine oxygenase[p]	Horse kidney	100 000	2
Cytochrome oxidase[q]	*Pseud. aerugenosa*	67 500	2
Dipeptidyl transferase[r]	Ox spleen	197 000	$8(\alpha_6\beta_2)$
Erythrocruorin[s]	Earthworm	*ca.* 3×10^6	144
Enolase[t, u]	Yeast	88 000	2
	Salmon	100 000	2
Formyl tetrahydrofolate synthetase[v]	*Clost. thermoaceticum*	244 000	4
Glutamate–aspartate transaminase[w]	Pig heart	92 000	2
Glutamate decarboxylase[x]	*E. coli*	310 000	6
Glutamine synthetase[y]	*B. subtilis*	600 000	12
Glyceraldehyde 3-phosphate dehydrogenase[z]	*B. stearothermophilus*	144 000	4
	Ox liver	142 000	4
Glycogen phosphorylase *a*[aa]	Rabbit muscle	370 000	4
Glyoxylic acid reductase[bb]	Spinach leaf	97 500	2
Haemerythrin[cc]	*Dendrostomum pyroides*	100 000	8
Haemocyanin[dd]	*Cancer magister*	940 000	(?) 36
Hexokinase[ee]	Yeast	51 000	1
Isocitrate dehydrogenase[ff]			
(NAD-linked)	Ox heart	330 000	8
(NADP-linked)	Pig heart	58 000	1
(NADP-linked)	Pig liver	75 000	2
(NADP-linked)	*B. stearothermophilus*	92 500	2
β-Ketoacyl carrier protein synthetase[gg]	*E. coli*	66 000	2
Lactate dehydrogenase[hh]	Horseshoe crab	140 000	4
		or 65 000	2
Lactate dehydrogenase (cytochrome b_2)[ii]	Yeast	210 000	4
Lactoferrin[jj]	Ox	77 000	1
Leucine aminopeptidase[kk]	Pig kidney	255 000	4
Lipoxygenase[ll]	Soybean	108 000	2
Lysine-2,3-aminomutase[mm]	*Clostridium* SB4	285 000	6
Malate dehydrogenase[nn]	Rat liver	66 000	2
Malic enzyme[oo]	*E. coli*	550 000	8
Phosphoglycerate mutase[pp]	Rabbit muscle	54 000	2



Scheme 9

the direct introduction of heavy atoms. A thiol-containing imidoester (9) has recently been synthesised [679] and found to react as expected with the B chain of oxidised insulin and with tobacco mosaic virus, and to provide points of attachment for mercurials. A simple procedure has been described [680] for identifying with certainty the half-cystine residues involved in bonds of the type R—S—Hg—S—R (see last year's Report, p. 70) that



(8) (9)

have been artificially generated in proteins; mercury is removed by treatment with EDTA and the liberated thiols are methoxycarbonylated *in situ* with [^{14}C]iodoacetic acid.

B. Carboxy-groups.—In several instances (*e.g.* trypsin,[371] transferrin,[261] α-lactalbumin,[435] and insulin [681]) water-soluble carbodi-imides have been used to couple nucleophiles (such as glycineamide or glycine esters) with the carboxy-groups of proteins. Modification of carboxy-groups by carbodi-imides in the absence of added nucleophiles has been summarised;[662] a single carboxy-group in carboxypeptidase was modified in this way.[353] Even with carbodi-imides, however, which are generally regarded as specific for carboxy-groups (although phenolic hydroxy-groups do react reversibly; see reference 662) complications can arise. It has been reported that reaction can occur with protein thiol groups and this is not reversed in

[679] R. N. Perham and J. O. Thomas, unpublished work.
[680] Y. Burstein and R. Sperling, *Biochim. Biophys. Acta*, 1970, **221**, 412.
[681] H. Ozawa, *Biochemistry*, 1970, **9**, 2158.

alkaline conditions.[682] Thus, care is called for in interpreting the effects of modification of proteins when carbodi-imides are used. Some compensation for the poor nucleophilicity of the carboxylate ion is afforded by the use of reagents with good leaving groups, *e.g.* triethyloxonium fluoroborate,[683] which has been used for trypsin,[372] or diazo-compounds. The latter are particularly effective when used as affinity labels, *e.g.* in pepsin (see Section 5A).

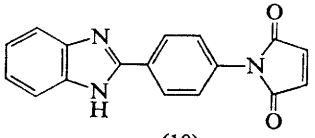
C. Thiol Groups.—Several articles in a single volume [349] illustrate the ways in which studies of the reactivity of thiol groups can provide information about the structure of proteins in which they occur. Different thiol groups in the same protein often react at different rates; it is important to remember that the reactivity of a particular group may well vary with the reagent used as probe and that the order of reactivity of groups may often be reversed.

Alkylation with halo-acetate or -acetamide continues to be a popular method, *e.g.* for aldolase,[684, 493a] lactic [685] and malic [686] dehydrogenases, and adenosine deaminase.[687] The cysteine residues of aldolase have also been investigated by reaction with disulphide monosulphoxides to give protein mixed disulphides:[493b] in this case it was found that two of the four reactive thiol groups could be protected by the competitive inhibitor hexitol diphosphate. DTNB [5,5′-dithiobis-(2-nitrobenzoate)] is also widely employed, *e.g.* for the thiol groups of aldolase,[684] jack bean urease,[688] met-tRNA synthetase,[59] pancreatic lipase,[689] and the NADP-linked isocitrate dehydrogenase from *Azotobacter vinelandii*.[690] A report [691] that the alkylation with *N*-ethylmaleimide of glutathione in red blood cells appears to be reversible should be borne in mind.

It has now been confirmed [692] by isolation of the labelled peptide that fluorodinitrobenzene reacts specifically with the catalytically active thiol group in glyceraldehyde 3-phosphate dehydrogenase (see also last year's Report, p. 69). A study [693] of pig heart lactate dehydrogenase in which the 'essential' thiol group has been converted into an *S*-sulpho group suggests that loss of activity is due to the inability of the binary enzyme–coenzyme complex to bind substrate.

[682] K. L. Carraway and R. B. Triplett, *Biochim. Biophys. Acta*, 1970, **200**, 564.
[683] S. M. Parsons, L. Jao, F. W. Dahlquist, C. L. Borders, jun., T. Groff, J. Racs, and M. A. Raftery, *Biochemistry*, 1969, **8**, 700.
[684] P. J. Anderson and R. N. Perham, *Biochem. J.*, 1970, **117**, 291.
[685] J. Südi and M. G. Khan, *F.E.B.S. Letters*, 1970, **6**, 253.
[686] E. M. Gregory and J. H. Harrison, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 995.
[687] J. Phelan, F. McEvoy, S. Rooney, and T. G. Brady, *Biochim. Biophys. Acta*, 1970, **200**, 370.
[688] A. T. DeB. Andrews and F. J. Reithel, *Arch. Biochem. Biophys.*, 1970, **141**, 538.
[689] R. Verger, L. Sarda, and P. Desnuelle, *Biochim. Biophys. Acta*, 1970, **207**, 377.
[690] J. E. Braginski, J.-S. Franzen, and A. E. Chung, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 644.
[691] E. Beutler, S. K. Srivastava, and C. West, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 341.
[692] S. Shaltiel and M. Tauber-Finkelstein, *F.E.B.S. Letters*, 1970, **8**, 345.
[693] J. J. Holbrook and R. A. Stinson, *Biochem. J.*, 1970, **120**, 289.

Two fluorogenic reagents for thiol groups have been described. The thiol groups of several enzymes are under investigation with NBD chloride[349, 504] (7-chloro-4-nitrobenzo-2-oxa-1,3-diazole) (see p. 102); work with BIPM [*N*-(*p*-2-benzimidazolylphenyl)-maleimide, (10)] is still at a preliminary stage.[694]



(10)

D. Tyrosine.—*Nitration.* Last year several instances of the successful use of tetranitromethane (TNM) to modify tyrosine residues were reported, but it was also necessary to warn against certain side-reactions. Much the same can be said this year, except that the side-reactions have now been more clearly defined. Riordan, Sokolowsky, and their co-workers continue their efforts to analyse the scope and limitations of their reagent. They conclude[695] that the acid-quenching procedure commonly used to terminate the reaction is largely responsible for the net loss of tyrosine sometimes noted. This is attributed to destruction of tyrosine by nitrous acid generated from nitrite by-product, and the original procedure of gel-filtration at pH 8 for terminating the reaction is recommended. Reports of cross-linking of tyrosine residues as a further side-reaction continue to appear, for example when TNM was used to nitrate insulin,[696] and several other proteins.[697] Nitration of a model peptide suggested[698] that this might be avoided under conditions where most of the reaction proceeded in acid solution. Sokolowsky, Riordan, and their co-workers also confirmed[698] reaction of tryptophan residues with TNM. It is worth bearing in mind that this reaction is very slow below pH 7; at pH 7 nitration of tyrosine residues might still be possible.

The novel idea of modification of intermediates in enzymic catalysis has already been mentioned.[664] Chemical probing of the enzyme–substrate complex with TNM was explored, and proved fruitful, for yeast and muscle aldolase[664] and for aspartate aminotransferase.[664, 699] For the aldolases very similar carbanion intermediates were detected, which is interesting in view of the differences between Class I and Class II aldolases. In aspartate aminotransferase, modification of an essential tyrosine residue occurred only in the presence of the substrate-pair glutamate/α-ketoglutarate and, since the coenzyme of the inactivated enzyme was found in the pyridoxamine

[694] Y. Kanaoka, M. Machida, K. Ando, and T. Sekine, *Biochim. Biophys. Acta*, 1970, **207**, 269.
[695] M. Sokolovsky and J. F. Riordan, *F.E.B.S. Letters*, 1970, **9**, 239.
[696] R. W. Boesel and F. H. Carpenter, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 678.
[697] J. P. Vincent, M. Lazdunski, and M. Delaage, *European J. Biochem.*, 1970, **12**, 250.
[698] M. Sokolovsky, M. Fuchs, and J. F. Riordan, *F.E.B.S. Letters*, 1970, **7**, 167.
[699] P. Christen and J. F. Riordan, *Biochemistry*, 1970, **9**, 3025.

form, it was concluded that inactivation occurred during or after the transition of the aldimine to the ketimine intermediate. These reactions contrast nicely with those in which substrates protect, and offer clear indication that dynamic processes take place within the enzyme–substrate complex.

Nitration of the functional tyrosine residue in arginine kinase was studied by first blocking thiol groups reversibly with tetrathionate.[700] Nitration of haemerythrin suggested that tyrosine played some rôle in liganding the iron, and that the environments of the two iron-binding sites per subunit were not identical.[701] TMN has also been used for reaction with tyrosine residues in the following: glucose-6-phosphate dehydrogenase;[702] human serum albumin, and bovine and human goitre thyroglobulins;[703] and insulin.[704] In α-amylase one tyrosine residue per subunit reacted with loss of activity.[705]

Iodination. Iodine monochloride or iodine–potassium iodide continue to find some use in studies of the degree of exposure of tyrosine residues in proteins, although histidine may also react. It has now been shown that the relative rates of iodination of tyrosine to mono-iodotyrosine and of mono- to di-iodotyrosine in proteins depend, not unexpectedly, on the particular tyrosine residue being iodinated;[706] the differences are attributed to microenvironment. The same workers studied the relative rates of iodination of the four tyrosine residues in the constant region of a Bence Jones protein, using their very useful paired-label technique.[707] They were also able to conclude that the microenvironment may affect differently the introduction of the first and second iodine atoms into tyrosine. It has been reported that hydrolysis of iodinated proteins with methanolic sodium hydroxide for 16 h at 110 °C does not result in the destruction of iodo-tyrosines, unlike the conditions normally used for acid hydrolysis of proteins, so that the mono- and the di-iodinated forms can be quantitatively estimated on the automatic amino-acid analyser.[708]

The reaction of tyrosine residues in glyceraldehyde 3-phosphate dehydrogenase with iodine in potassium iodide has been reported from two laboratories. In one case[709] the holoenzyme (from lobster) was studied and oxidative side-reactions at the active thiol group were precluded by monomethoxycarbonylation, and in the other[710] the apoenzyme (from pig)

[700] R. Kassab, A. Fattoum, and L. A. Pradel, *European J. Biochem.*, 1970, **12**, 264.
[701] R. L. Hill and I. M. Klotz, *Arch. Biochem. Biophys.*, 1970, **136**, 507.
[702] W. Domschke, C. Von Hinueber, and G. F. Domagk, *Biochim. Biophys. Acta*, 1970, **207**, 485.
[703] P. G. Malan and H. Edelhoch, *Biochemistry*, 1970, **9**, 3205.
[704] J. W. S. Morris, D. A. Mercola, and E. R. Arquilla, *Biochemistry*, 1970, **9**, 3930.
[705] J. M. Connellan and D. C. Shaw, *J. Biol. Chem.*, 1970, **245**, 2845.
[706] B.-K. Seon, O. A. Roholt, and D. Pressman, *Biochim. Biophys. Acta*, 1970, **221**, 114.
[707] B.-K. Seon, O. A. Roholt, and D. Pressman, *Biochim. Biophys. Acta*, 1970, **200**, 81.
[708] D. M. Harrison and C. J. Garratt, *F.E.B.S. Letters*, 1970, **11**, 14.
[709] J. O. Thomas and J. I. Harris, *Biochem. J.*, 1970, **119**, 307.
[710] S. Libor and P. Elödi, *European J. Biochem.*, 1970, **12**, 336, 345.

was studied. It is not too surprising, therefore, that the observed reactivities of the tyrosine residues differed. Oxidation of thiol groups complicated the reaction of lactic dehydrogenases with iodine[711] at pH 9. It has been known for some time that increased iodination of thyroglobulin confers on the molecule increased resistance to dissociation by sodium dodecyl sulphate: the cross-links now appear to be no more than disulphide bridges, and dissociation can be brought about under reducing and denaturing conditions.[712]

Acetylation and Other Methods. Acetylation with *N*-acetylimidazole (NAI) seems to be fairly free from side-reactions, and the modification of tyrosine residues can be reversed by treatment with hydroxylamine. Some irreversible side-reactions with amino-groups have, however, been reported for α-amylase from *B. subtilis*.[705] Octameric glutamine synthetase (from sheep brain) dissociated into two tetramers when acetylated with NAI; the native structure was regenerated when blocking groups were partially removed, and tyrosine residues were thereby implicated in stabilisation of the oligomeric structure.[713] NAI acetylates the active serine residue of trypsin with concomitant inactivation. The derivative can be isolated at low pH but at neutral pH it deacylates and reactivates.[714] At a later stage, a further acetyl group is introduced, perhaps at histidine, and this greatly retards the deacetylation of the serine. These acetyl groups can be distinguished from those on lysine and other serine residues by virtue of their removal with dilute imidazole.

Neither nitration nor acetylation of the single tyrosine residue of bovine neurophysin-II affected its hormone-binding power;[715] and both TNM and NAI were used with iodine and cyanuric fluoride in a study of a bacterial protease.[716] Cyanuric fluoride has been recommended for selective modification of tyrosine residues, and a comparison has been drawn with *N*-acetylimidazole.[717]

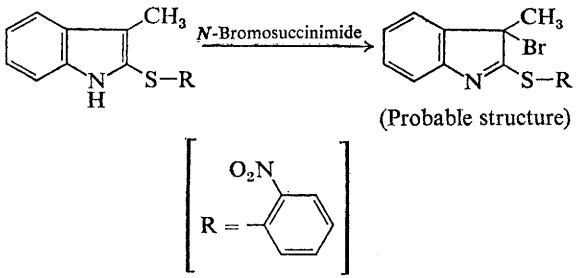
E. Tryptophan, Histidine, and Arginine.—Incorporation of 4-, 5-, and 6-fluorotryptophan into bacterial proteins *in vivo* will be useful in studies of the environment of tryptophan residues in proteins by means of ^{19}F n.m.r. spectroscopy.[718]

Further details of the reaction of tryptophan residues with 2-hydroxy-5-nitrobenzyl bromide have now been given[719] (see last year's Report, p. 67). Uses of this reagent to investigate the disposition of the tryptophan

[711] M. C. Shen and P. M. Wassarman, *Biochem. Biophys. Acta*, 1970, **221**, 407.
[712] M. Rolland and S. Lissitzky, *Biochim. Biophys. Acta*, 1970, 214, 282; P. A. Charlwood, R. Pitt-Rivers, and H. L. Schwartz, *Biochem. J.*, 1970, **116**, 769.
[713] S. Wilk, A. Meister, and R. H. Haschemeyer, *Biochemistry*, 1970, **9**, 2039.
[714] L. L. Houston and K. A. Walsh, *Biochemistry*, 1970, **9**, 156.
[715] A. J. Furth and D. B. Hope, *Biochem. J.*, 1970, **116**, 545.
[716] D. Tsuru, T. Yoshida, T. Hirose, T. Yoshimoto, and J. Fukumoto, *Internat. J. Protein Res.*, 1970, **2**, 257.
[717] M. J. Gorbunoff, *Arch. Biochem. Biophys.*, 1970, **138**, 684.
[718] D. T. Browne, G. L. Kenyon, and G. D. Hegeman, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 13.

residues of α-lactalbumin [434] and to label a single tryptophan in streptococcal proteinase [404] have already been mentioned. In addition to the known side-reactions with sulphydryl groups, it has now been reported [720] that modification of the α-amino-group of carboxypeptidase A by this reagent is rapid and apparently specific. However, this side-reaction should not interfere with the determination of total tryptophan carried out in acid solution.[720]

An interesting new reagent for tryptophan residues has been used to modify Trp-140 of staphylococcal nuclease;[721] the final step in its preparation is bromination with *N*-bromosuccinimide (Scheme 10). It is better than



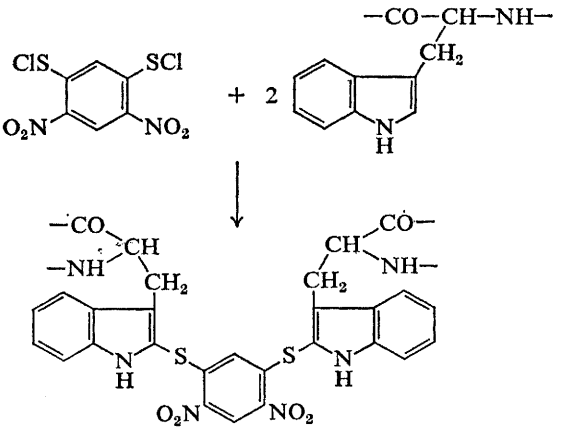
Scheme 10

N-bromosuccinimide itself for tryptophan modification in that it can be used without reaction of histidine and tyrosine, although at high molar excesses of reagent tyrosine and cysteine are modified. Tryptophan is converted to an oxindole derivative; prolonged exposure to the reagent results in cleavage of tryptophyl peptide bonds in low but acceptable yields (in the region of 15%), thus making it potentially useful for preparative chain cleavage. A bifunctional sulphenyl halide, 2,4-dinitro-1,5-phenyldisulphenyl chloride (DNPDS-Cl), is under scrutiny as a cross-linking reagent for tryptophan residues in proteins [722] (Scheme 11). It appears to work well on model compounds and to be sufficiently similar to *o*-nitrophenylsulphenyl chloride (NPS-Cl) to make it likely that it, too, can be used to modify proteins in aqueous solution. The reagent will also react with cysteine side-chains but this reaction is reversed by thiols.

Photo-oxidation of histidine residues in proteins is mentioned below. Histidine continues to show evidence of being difficult to modify specifically in proteins other than by affinity-labelling. Diazo-1*H*-tetrazole reacted preferentially with the thiol groups of myokinase, but even when these were protected tyrosine and lysine reacted before histidine.[723] It is known

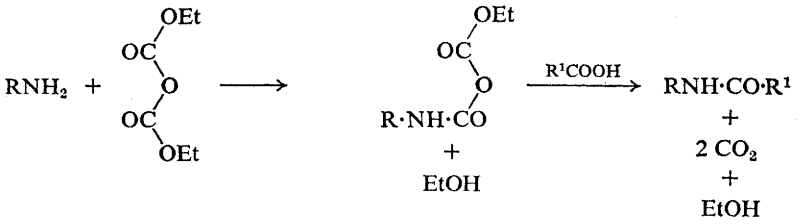
[719] G. M. London and D. E. Koshland, jun., *J. Biol. Chem.*, 1970, **245**, 2247.
[720] T. M. Radhakrishnan, R. A. Bradshaw, D. A. Deranleau, and H. Neurath, *F.E.B.S. Letters*, 1970, **7**, 72.
[721] G. S. Omenn, A. Fontana, and C. B. Anfinsen, *J. Biol. Chem.*, 1970, **245**, 1895.
[722] F. M. Veronese, E. Boccú, and A. Fontana, *Internat. J. Protein Res.*, 1970, **2**, 67.
[723] R. H. Schirmer, I. Schirmer, and L. Noda, *Biochim. Biophys. Acta*, 1970, **207**, 165.

that ethoxyformic anhydride (diethyl pyrocarbonate) (see last year's Report, p. 69) reacts with histidine in several proteins but preferentially with lysine in others (see reference 662). A recent report[724] shows that accessible imidazole and amino side-chains in proteins are rapidly acylated by this reagent, even at pH 4. However, the reagent can be used successfully in some cases, *e.g.* histidine residues in sheep heart phosphofructokinase were modified with loss of the allosteric properties of the enzyme; the reaction was reversed, and the allosteric properties regained, by treatment with hydroxylamine.[725] For ribonuclease, however, irreversible



Scheme 11

inactivation and polymer formation ensued,[726] and an explanation involving the formation of amide cross-links was proposed (Scheme 12). This is obviously a general point, particularly since treatment of proteins with diethyl pyrocarbonate has already been suggested (see Section 2E) as a method of preparing polymeric proteins for gel electrophoresis.[148] It has also been shown that the reagent can react with tryptophan side-chains in proteins and that reaction of buried residues may be favoured.[727]



Scheme 12

[724] W. B. Melchior, jun., and D. Fahrney, *Biochemistry*, 1970, **9**, 251.
[725] B. Setlow and T. E. Mansour, *J. Biol. Chem.*, 1970, **245**, 5524.
[726] B. Wolf, J. A. Lesnaw, and M. E. Reichmann, *European J. Biochem.*, 1970, **13**, 519.
[727] C.-G. Rosén, T. Gejvall, and L.-O. Andersson, *Biochim. Biophys. Acta*, 1970, **221**, 207.

The reaction of arginine residues with oligomers of 2,3-butanedione (biacetyl) is mentioned in Section 2D and the chemical modification of arginine in the active site of trypsin inhibitors in Section 5A. Treatment of glucose 6-phosphate dehydrogenase from *Candida utilis* with 1,2-cyclohexanedione resulted in complete inactivation when 15 out of the 42 arginine residues were modified;[728] substrate and coenzyme together protected against the inactivation. Unfortunately, it remains true that, in general, good selective modification of arginine residues in proteins is still beyond the protein chemist.

F. Photo-oxidation.—Photosensitised oxidation of amino-acids can be made specific with appropriate choice of conditions. This has recently been reviewed by Scoffone *et al.*[729] Photo-oxidation of glyceraldehyde 3-phosphate dehydrogenase from rabbit muscle in the presence of Rose Bengal resulted in reaction of a single histidine residue (His-38) with 50—60% loss of enzymic activity.[730] Photo-oxidisable histidine residues in collagenase A were protected from reaction mediated by Methylene Blue when calcium ions were present, leading to speculation on histidine–metal interaction in the enzyme.[731] When haematoporphyrin was used as sensitiser, photo-oxidation was confined to methionine residues in ribonuclease A;[732] measurement of photo-oxidation as a function of increasing strength of denaturing solvent enabled the accessibility of these residues to be estimated. Photo-oxidation in the presence of low concentrations of acetone is reported to be highly specific for methionine residues,[733] and in ribonuclease gave selective reaction at Met-29.

Photo-oxidation can be made extremely specific for particular residues in a protein if the photo-sensitiser is protein-bound: the life-time of the excited intermediate is such that only photo-oxidisable residues in the immediate vicinity will react.[729] If the photo-sensitiser is a substrate analogue, a coenzyme, or an inhibitor, photo-oxidation should be confined to residues at or near the active site. When Scoffone and his co-workers used a dinitrophenyl group attached to Lys-41 of ribonuclease A as photo-sensitiser, they found[734] specific modification of Met-30, His-12, and Tyr-97, in accord with the topography of the active site as revealed by *X*-ray crystallography. 'Active-site-directed' photo-oxidation was also achieved for α-chymotrypsin when the inhibitor *N*-acetyl-3-nitrotyrosine served as photo-sensitiser for the oxidation of Met-192 and His-57.[735]

[728] G. F. Domagk, W. Domschke, and C. von Hinüber, *Z. physiol. Chem.*, 1970, **351**, 718.
[729] E. Scoffone, G. Jori, and G. Galiazzo, ref. 349, p. 163.
[730] J. S. Bond, S. H. Francis, and J. H. Park, *J. Biol. Chem.*, 1970, **245**, 1041.
[731] S. Takahashi and S. Seifter, *Biochim. Biophys. Acta*, 1970, **214**, 557.
[732] G. Jori, G. Galiazzo, A. M. Tamburro, and E. Scoffone, *J. Biol. Chem.*, 1970, **245**, 3375.
[733] G. Gennari and G. Jori, *F.E.B.S. Letters*, 1970, **10**, 129.
[734] G. Jori, G. Galiazzo, F. Marchiori, and E. Scoffone, *Internat. J. Protein Res.*, 1970, **2**, 247.
[735] G. Gennari, G. Jori, G. Galiazzo, and E. Scoffone, *J. Amer. Chem. Soc.*, 1970, **92**, 4140.

Similarly, spinach leaf aldolase was photo-oxidised in the presence of the cofactor pyridoxal phosphate;[736] specific reaction of one histidine residue per subunit was reported although the particular residue is not yet identified. The modified enzyme was inactive but could still bind substrate. Two histidines were oxidised in glucose 6-phosphate dehydrogenase from *Candida utilis* when Rose Bengal, a potent competitive inhibitor of this enzyme with respect to NADP, mediated the reaction;[737] and photo-oxidation of horse heart cytochrome *c*, taking advantage of the porphyrin group as sensitiser,[738] resulted in reaction with His-18 and Met-80. The former residue is known to be one of the two ligands bonded to iron in ferricytochrome *c*, and it is suggested that Met-80 may be the other.[738]

Table 3 *Some recently reported active-site-directed inhibitors*

Enzyme	*Active-site-directed inhibitor*	*Residue modified*	*Ref.*
Adenosine deaminase	9-(*p*-Bromoacetamidobenzyl) adenine	Lys	739
Aldolase	Iodoacetol phosphate (1-hydroxy-3-iodo-2-propanone)	Cys	740
Alkaline phosphatase	Chloroacetyl phosphate and monochloro-β-glycerophosphate	His	741
Carnitine acetyltransferase	Bromoacetyl-L-carnitine	His	742
Glutamate dehydrogenase	4-Iodoacetamidosalicylic acid	Cys	743
Pepsin	1-Diazo-4-phenyl-2-butanone	Asp	388
Staphylococcal nuclease	Deoxythymidine 3′-*p*-aminophenyl phosphate 5′-phosphate	Tyr-115	744
Staphylococcal nuclease	Deoxythymidine 5′-*p*-aminophenyl phosphate	Tyr-85	744
Staphylococcal nuclease	Deoxythymidine 3′-*p*-aminophenyl phosphate	Trp-140 and His-46	744
Triosephosphate isomerase	Glycidol phosphate[a] (2,3-epoxypropanol phosphate)	Glu	488
Triosephosphate isomerase	Iodoacetol phosphate	Glu	745
Triosephosphate isomerase	Chloroacetol phosphate	Glu	489
Triosephosphate isomerase	Bromohydroxyacetone phosphate	Glu	490

[a] Glycidol phosphate was designed to resemble the transition state for the reaction (see reference 488).

[736] L. C. Davis, L. W. Brox, R. W. Gracy, G. Ribereau-Gayon, and B. L. Horecker, *Arch. Biochem. Biophys.*, 1970, **140**, 215.
[737] M. Rippa, C. Picco, and S. Pontremoli, *J. Biol. Chem.*, 1970, **245**, 4977.
[738] G. Jori, G. Gennari, G. Galiazzo, and E. Scoffone, *F.E.B.S. Letters*, 1970, **6**, 267.
[739] G. Ronca, M. F. Saettone, and A. Lucacchini, *Biochim. Biophys. Acta*, 1970, **206**, 414.
[740] F. C. Hartman, *Biochemistry*, 1970, **9**, 1783.
[741] H. Csopak and G. Fölsch, *Acta Chem. Scand.*, 1970, **24**, 1025.
[742] J. F. A. Chase and P. K. Tubbs, *Biochem. J.*, 1970, **116**, 713.
[743] A. D. B. Malcolm and G. K. Radda, *European J. Biochem.*, 1970, **15**, 555.
[744] P. Cuatrecasas, *J. Biol. Chem.*, 1970, **245**, 574.
[745] F. C. Hartman, *J. Amer. Chem. Soc.*, 1970, **92**, 2170.

A useful 'destruction–reduction' procedure is reported for the estimation of methionine sulphoxide in photo-oxidised myokinase:[723] the photo-oxidised protein is treated with cyanogen bromide to destroy unoxidised methionine, without effect on methionine sulphoxide, and the methionine sulphoxide is then estimated as methionine after carrying out acid hydrolysis of the protein in 6N hydrochloric acid in the presence of mercaptans.

G. Active-site-directed Inhibitors.—Many aspects of the specific modification of proteins by affinity labelling have recently been discussed in an excellent review[663] and will not be considered further here. Photo-oxidation of residues near the active site of enzymes has been described in the preceding section, and the specific labelling of certain residues in membrane proteins has been considered in Section 7.

Modification of residues at the active sites of trypsin and chymotrypsin has been outlined in Section 5A and some other instances of active-site labelling that have been reported in the past year are collected in Table 3.

PART II: *X*-Ray Studies *by C. C. F. Blake*

1 Introduction

1970 has seen three important events in protein crystallography that with uncommon neatness sum up the main achievements of the past and indicate the direction the subject will take in the future. The publication of the Royal Society Discussion on the structures and functions of proteolytic enzymes has given us the most detailed overall view of a group of enzymes so far available. It provides numerous fascinating details of the mechanisms of these enzymes, their specificities, and, possibly most interesting of all, their evolutionary relationships. This has been achieved by equal contributions from the crystallographic and the chemical–biochemical approaches and demonstrates the interdependence of the two methods and their power when they are used jointly on a single problem. The second summing-up paper is one in which Perutz appears to have found the structural explanation of the co-operative effects in haemoglobin. This has been one of the most outstanding problems of the past decade and its solution appears to vindicate the Monod–Wyman–Changeux view of allostery. The outstanding feature of Perutz's hypothesis is the beautiful simplicity of the change of state of the haemoglobin molecule: a small change in the radius of the haem-iron atom on oxygenation is amplified and transmitted to the other haem-iron atoms by the globin parts of the molecule. Finally, the announcement of the high-resolution structure of lactate dehydrogenase represents the first step along the new path of metabolic enzymes which more crystallographers will take in future. In another few years, perhaps, the Royal Society, or some other body, will be able to mount detailed discussion meetings on *the* dehydrogenases, *the* kinases, or *the* transaminases.

2 Amino-acids and Oligopeptides (See also Chapter 1, Section 3B)

A. Arginine.—Crystals of L-arginine hydrochloride [1] have been found to contain two molecules of the amino-acid in the asymmetric unit. The two molecules have very similar bond lengths for equivalent bonds, for example both the guanidino groups are planar with an average C–N bond length of 1.325 Å, but their conformations are different. In the notation of Edsall *et al.*,[2] the conformational angles are:

	ψ_1	ψ_2	χ_1	χ_2	χ_3	χ_4	χ_{51}	χ_{52}
Molecule I	154.9°	332.0°	306.1°	173.9°	185.2°	96.7°	346.0°	166.3°
Molecule II	174.1°	353.6°	301.1°	195.6°	179.3°	276.0°	10.6°	189.7°

B. Histidine.—The imidazole ring of the histidine molecule in crystals of DL-histidine hydrochloride dihydrate [3] is in its charged state. The molecule was found to be fully extended with the imidazole group *trans* to the carboxy-group across the C^{α}–C^{β} bond. The dihedral angles, $\psi_1 = 162.6°$, $\psi_2 = -16.7°$, $\chi_1 = -61.9°$, $\chi_{21} = -70.6°$, $\chi_{22} = 107.7°$, are rather different from those found in the study of L-histidine hydrochloride, but the bond lengths and bond angles are very similar.

C. *O*-Phosphoryl-serine.—The molecular structure of *O*-phosphoryl-serine is of considerable interest because of its existence in membrane proteins, metabolic enzymes such as phosphoglucomutase, and phospholipids such as phosphotidylserine. The structures of both phosphoryl-DL-serine monohydrate [4] and *O*-phosphoryl-L-serine [5] have been accurately determined. In both cases, the molecules are in the zwitterion form with the α-amino-group protonated while the phosphate group carries a negative charge. An unusual feature is that there are very short hydrogen bonds in both structures; in the DL-serine derivative intermolecular hydrogen bonds of about 2.50 Å link phosphate groups together, while the analogous L-serine structure contains an intermolecular carboxy–phosphate hydrogen bond of 2.49 Å. Similar very short hydrogen bonds involving phosphates have been observed elsewhere, while another feature characteristic of phosphate esters, that the phosphate ester oxygen is not involved in hydrogen bonding, is also maintained in these two structures.

D. Valine.—Like the crystals of L-arginine hydrochloride described above, crystals of L-valine [6] contain two molecules in different conformations in the asymmetric unit. One of the two molecules is in the so-called

[1] J. Dow, L. H. Jensen, S. K. Mazumdar, R. Srinivasan, and G. N. Ramachandran, *Acta Cryst.*, 1970, **B26**, 1662.
[2] J. T. Edsall, P. J. Flory, J. C. Kendrew, A. Liquori, G. Nemethy, G. N. Ramachandran, and H. A. Scherga, *J. Mol. Biol.*, 1930, **15**, 339.
[3] I. Bennett, A. G. H. Davidson, M. M. Harding, and I. Morelle, *Acta Cryst.*, 1970, **B26**, 1722.
[4] E. Putkey and M. Sundaralingam, *Acta Cryst.*, 1970, **B26**, 782.
[5] M. Sundaralingam and E. Putkey, *Acta Cryst.*, 1970, **B26**, 790.
[6] K. Torii and Y. Iitaka, *Acta Cryst.*, 1970, **B26**, 1317.

Gauche I conformation with dihedral angles $\psi_2 = -19.5°$, $\chi_{11} = 206.4°$, and $\chi_{12} = 81.8°$, while the other is in the *Trans* conformation with $\psi_2 = -43.7°$, $\chi_{11} = 300.7°$, and $\chi_{12} = 179.2°$.

E. Perdeuterio-α-glycylglycine.—This work [7] is the first three-dimensional structure determination of a peptide by neutron diffraction that accurately locates the hydrogen, or rather deuterium, atoms. The dimensions and conformation of the molecule are the same as those reported for the *X*-ray study of α-glycylglycine. The C(methylene)–D, N(amino)–D, and N(peptide)–D bond lengths are 1.085, 1.03, and 1.02 Å respectively. All four crystallographically independent hydrogen bonds are non-linear; the N–D···O angles range between 148° and 163°. The peptide proton and the C_α atom lie 0.145 Å and 0.07 Å respectively out of the plane of the peptide C_α^1–CO–N.

3 Proteins

A. Methods.—One of the more serious problems in protein crystallography is radiation damage of the crystals. All crystals of proteins appear to be susceptible to radiation damage, but the sensitivity varies widely, apparently being dependent on the crystal packing rather than the nature of the protein. In a number of cases, protein crystals have been found to be so sensitive to radiation that structure determination becomes exceedingly difficult or impossible. A promising method of reducing the sensitivity of crystals is to freeze the liquid component of the crystal and thus prevent the migration of free radicals formed in solution to the surface of the protein molecules. Freezing should therefore markedly reduce that part of the radiation damage that arises from free-radical attack. Haas and Rossman [8] have been able to freeze lactate dehydrogenase crystals by equilibrating them with sucrose–ammonium sulphate solutions and then dipping in liquid nitrogen. *X*-Ray examination of the frozen crystals at − 75 °C has shown that a ten-fold reduction in the rate of radiation damage has been achieved. In addition, experiments with heavy-atom derivatives show that satisfactory electron-density maps can probably be obtained from the frozen crystals by using isomorphous replacement in the usual way. As might have been expected, the overall temperature factor of the crystals remains substantially unaltered. Somewhat surprisingly, a report of a study on insulin [9] at − 13 °C (several degrees above the freezing point of the mother liquor) suggests that a large decrease in the temperature factor accompanies cooling. However, in view of the fact that no conclusive evidence for an increase in the average intensity in the high-angle region has been obtained, this observation must be viewed with caution. It is pointed out that the collection of intensity data at low temperatures should be made in the

[7] H. C. Freeman, G. L. Paul, and T. M. Sabine, *Acta Cryst.*, 1970, **B26**, 925.
[8] D. J. Haas and M. G. Rossmann, *Acta Cryst.*, 1970, **B26**, 998.
[9] P. Cucka, L. Singman, F. M. Lovell, and B. W. Low, *Acta Cryst.*, 1970, **B26**, 1756.

range where large temperature-dependent intensity changes are not observed and that the temperature needs to be carefully controlled. It is not clear at this stage whether crystals that are cooled but with the mother liquor still above its freezing point are more sensitive to temperature variation than crystals in which the mother liquor is completely frozen.

B. Results.—*α-Chymotrypsin.* A new map of α-chymotrypsin has been calculated [10] which now shows the orientations of the carbonyl groups of the polypeptide chains and defines the orientations of the planar peptide group. The chain itself is represented by a more continuous electron-density distribution and apart from residues 11—13 and 74—77 the molecule is well defined. An additional short section of α-helix has been found, which involves residues 164—170. The hydrogen-bonding system of the polypeptide chain that has been found in the new map is shown schematically in Figure 1 (see also Figure 4). Adjacent chains in Figure 1 are always

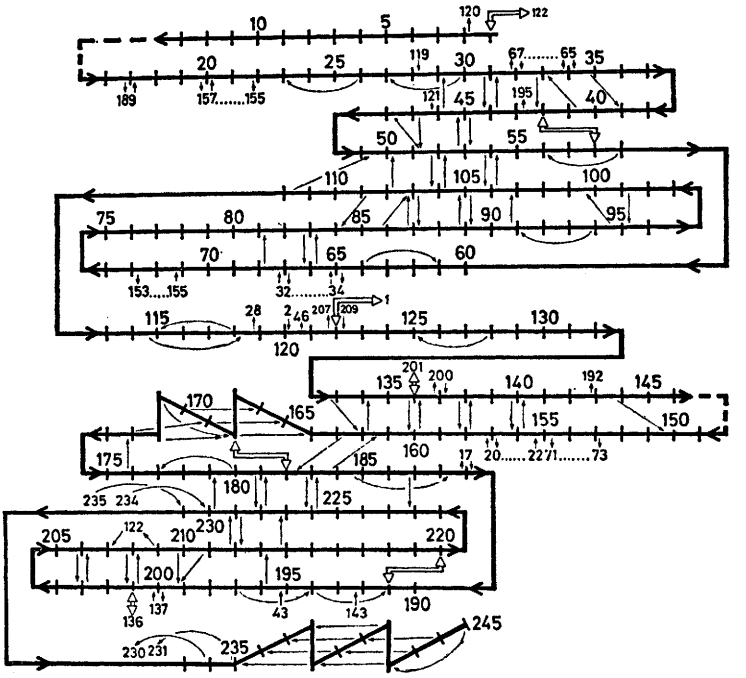
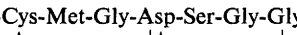


Figure 1 *A schematic diagram of the hydrogen bond system in α-chymotrypsin. The heavy line represents the polypeptide chain, the light arrows represent the hydrogen bonds and point from the amido to the carbonyl groups, and the open-headed arrows represent disulphide bridges. The dotted lines represent residues present in chymotrypsinogen that are removed from the enzyme by autolysis*
(Reproduced by permission from *Phil. Trans. Roy. Soc.*, 1970, **B257**, 67)

[10] J. J. Birktoft, D. M. Blow, R. Henderson, and T. A. Steitz, *Phil. Trans. Roy. Soc.*, 1970, **B257**, 67.

antiparallel and the predominating hydrogen bond structure is the antiparallel pleated sheet. The overall structure of the molecule has the form of two cylinders of antiparallel pleated sheet, which have an interesting relationship. In Figure 1 the top line represents the A-chain, the next six lines are drawn according to a particular pattern and the same pattern is repeated again further down the diagram, or, rather, further along the sequence. Within each of these patterns there is a further series of antiparallel interactions between the last and first lines. In spite of the tendency towards antiparallel interactions only very short sections of classical pleated-sheet structure are found.

Another very interesting feature of the hydrogen-bonding scheme is that hydrogen bonds to the third-nearest neighbour back along the chain are very common. Using the analysis of Venkatachalam,[11] who has defined two types of conformation with this type of hydrogen bond as Type I, in an approximately helical conformation, and Type II, in which the direction of the second of the three peptides is reversed and is permitted only when glycine is in position 3, Blow has found fourteen Type I situations and three Type II. Two of the latter cases are associated with the active site residues 191—197,



in which the hydrogen bonds provide the rigid structure that maintains Ser-195 and Asp-194 in their appropriate positions. It appears that the known invariance of Gly-193 and Gly-196 in vertebrate serine proteases is for the reasons given above.

The overall structure of the molecule is made up of a number of loops in which the polypeptide chain folds back on itself. A number of antiparallel hydrogen bonds stabilize each loop while others are further stabilized by disulphide bridges. These latter include the 'histidine' loop (residues 42—58), the 'methionine' loop (165—182), and the 'serine' loop (190—220). Other loops not stabilized by disulphide bridges are the 'aspartate' loop (84—110) which contains the buried Asp-102, the 'autolysis' loop (133—164), and an unnamed loop composed of residues 66—83.

Further evidence on the mechanism of chymotrypsin has been obtained from a three-dimensional difference map calculated at 2.5 Å resolution of the acyl-enzyme, indoleacryloyl-α-chymotrypsin.[12] This acyl-enzyme is stable at the non-physiological pH value of 4 for a sufficient period for *X*-ray data to be collected. The difference map revealed the position of the acyl moiety clearly. The indolyl part of the substituent was found to bind in the hydrophobic pocket that accommodates the aromatic side-chains of *N*-formyl-L-tryptophan and *N*-formyl-L-phenylalanine. The carbonyl part of the indoleacryloyl group is close to the sulphonyl group of

[11] C. M. Venkatachalam, *Biopolymers*, 1968, **6**, 1425.
[12] R. Henderson, *J. Mol. Biol.*, 1970, **54**, 341.

tosyl-chymotrypsin and the ester link to Ser-195 can be seen. There is some evidence of movement of the protein on acylation and in particular the O^{γ} of Ser-195 has moved about 2.5 Å from its position in the native enzyme, apparently by a rotation of 120° about its C^{α}–C^{β} bond, and His-57 moves by about 0.3 Å towards the solvent region. In addition, there are small movements in the backbone chain in the vicinity of residues 191—192. There are also clear indications of two water molecules hydrogen bonded to the carbonyl oxygen of the acyl substituent, one of which forms an additional strong hydrogen bond with the N^{ϵ_2} of His-57 while the other is hydrogen bonded to the carbonyl of Phe-41. A re-examination of tosyl-chymotrypsin indicates that the tosyl group is bound in the same site as the indoleacryloyl group and that the side-chain movements and binding of water molecules are also essentially the same.

The accepted mechanism of chymotrypsin involves two identifiable steps. The leaving group that is eliminated in the acylation of the enzyme is replaced by a water molecule in the deacylation step that attacks the acyl linkage. The deacylation of these fairly stable acyl- and sulphonyl-chymotrypsins can be deduced from their structures. The movement of the O^{γ} of Ser-195 on acylation breaks the hydrogen bond that links this atom to its partners, His-57 and Asp-102, in the charge relay system. The N^{ϵ_2} of His-57 appears to make a hydrogen bond with the water molecule bound to the carbonyl of the acyl substituent in the acyl-enzyme. This water molecule has a proton partially removed by the combined effects of residues 102 and 57. Thus, this water molecule is identified as the one responsible for attacking the acyl linkage and is activated by the charge relay system so that it is able to deacylate the enzyme.

In order to extrapolate these results to the deacylation of real substrates, two assumptions are required. The first is that the acylamido function on the α-carbon of the L-amino-acid substrate hydrogen bonds to Ser-214, and the second is that the substrate carbonyl is rotated by 45°, from its position in the indoleacryloyl moiety, to allow the activated water molecule reasonable access to the carbonyl carbon. The resultant model of the reactive acyl-enzyme is shown in Figure 2. When the carbonyl carbon is positioned relative to the water molecule as shown in Figure 2, deacylation can proceed by general-base-catalysed removal of the water proton, formation of a tetrahedral intermediate, and breakdown of products.

By applying the principle of microscopic reversibility, the acylation step may be supposed to occur by an exact reversal of the movements of the atoms and electrons that occurred in the deacylation step. During acylation the O^{γ} of Ser-195 would move 2 Å by means of a rotation of 120° of its C^{β}–C^{γ} bond, which would move the oxygen directly towards the carbonyl-carbon of the substrate and allow a covalent bond to form. At the same time, the leaving group must occupy a position near His-57 which would allow the imidazole to protonate it. Thus the structure of the native and acylated enzymes indicate a stereochemically plausible mechanism of the

hydrolytic function of the enzyme. It is interesting to note that the mechanism involves only slight movements in the protein itself and suggests that induced fit plays little part in the functioning of the serine proteases.

Chymotrypsinogen. The determination of the structure of chymotrypsinogen A at 2.5 Å resolution [13] has provided a fascinating view of the structural relationships between an enzyme and its zymogen precursor. Chymotrypsinogen is activated by the tryptic cleavage of the Arg-15 to Ile-16

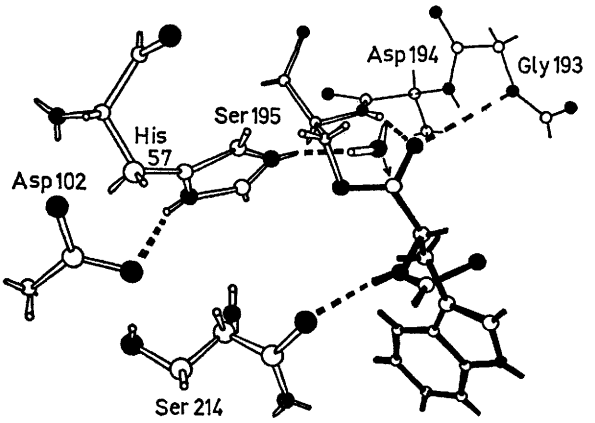


Figure 2 *The proposed model for the active site region of the reactive acylchymotrypsin. The acyl group is shown in the dark line, and hydrogen bonds by dotted lines. The activated water molecule has reasonable access to the carbonyl carbon as shown by the arrow*
(Reproduced by permission from *J. Mol. Biol.*, 1970, **54**, 341)

peptide bond to give π-chymotrypsin in the essential activation step; autolytic hydrolysis of bonds between residues 13—14, 146—147, and 148—149 gives rise to the δ, γ, and α forms of the enzyme, or to various neochymotrypsinogens, depending on the order in which the four bonds are split. α-Chymotrypsin differs covalently from chymotrypsinogen by the deletion of the two dipeptides Ser-14-Arg-15 and Thr-147-Asn-148. The comparison of their crystal structures shows the extent to which they differ structurally.

Chymotrypsinogen type F crystals are orthorhombic, space group $P2_12_12_1$ with $a = 52.0$, $b = 63.9$, and $c = 77.1$ Å, and contain one molecule in the asymmetric unit. Four heavy-atom derivatives were found to be suitable for phasing the high-resolution data. They were prepared from mercuric bromide, uranyl fluoride, uranyl pyrophosphate, and mersalyl.

[13] S. T. Freer, J. Kraut, J. D. Robertus, H. T. Wright, and Ng. T. Xuong, *Biochemistry*, 1970, **9**, 1997.

The *X*-ray intensity data were collected by parallel operation of an automatic diffractometer and screenless precession photography. The resulting map was calculated using the phases which decreased from 0.87 for ∞—5 Å shell to 0.55 for the 2.6—2.5 Å shell.

The folding of the polypeptide chain of chymotrypsinogen is shown in Figure 3 and may be compared with the equivalent view of α-chymotrypsin shown in Figure 4. The two molecules were compared analytically by transforming the atomic co-ordinates of Blow's α-chymotrypsin model into the crystallographic co-ordinate system of chymotrypsinogen and minimizing the distances between the α-carbon atoms of equivalent residues by a least-squares procedure. That the overall structures of the two molecules are essentially the same can be seen by noting that the mean displacement of *all* equivalent α-carbons is only 1.8 Å. However, there are six chain segments which exhibit relatively large conformational differences between the zymogen and the enzyme. Four of these segments involve only a repositioning on the surface of the molecule and of these, three, involving residues 7—8, 37—38, and 72—77, appear not to be essential to the activation process and may simply occur because of the differences in crystal packing forces. However, the remaining surface section of chain, residues 144—152, and the two parts of the chain in which residues move from the surface of the molecule to the interior or *vice versa*, which involve residues 16—17 and 192—193, do seem to be involved in the activation process.

The changes involved in the repositioning of Ile-16-Val-17 result in the formation of the internal ion-pair between the α-amino-group of Ile-16 and the buried carboxy-group of Asp-194. Apparently, upon tryptic scission of the Arg-15-Ile-16 peptide bond, the *N*-terminal residues of the newly-formed B-chain swing out through the solvent as the main chain rotates through 180° and the formerly exposed side-chains of Ile-16 and Val-17 become buried. The side-chain carboxy-group of Asp-194 is buried in the chymotrypsinogen as well as in α-chymotrypsin, but in the zymogen it is hydrogen bonded to His-40. During activation this group undergoes a 4 Å displacement that takes it towards the new position of Ile-16.

The segment containing residues 144—152 changes its conformation with dramatic results to the side-chain of Arg-145. In the zymogen its guanidinium group may be close enough to the buried carboxy-group of Asp-194 to permit appreciable electrostatic interaction between the two charged groups. In the enzyme, however, the backbone carrying Arg-145 has moved by 9 Å and the side-chain has swung out of its previous site to become fully extended into the solvent.

The chain carrying residues 192 and 193 moves about 7—8 Å carrying the side-chain of Met-192 from the interior of the molecule to the surface. The movement of the backbone causing this change is also responsible for the movement of Asp-194 that permits it to form the ion-pair bond, this in turn causes His-40 to move and to replace its previous hydrogen bond with a bond to the carbonyl oxygen of Gly-193. These changes all occur

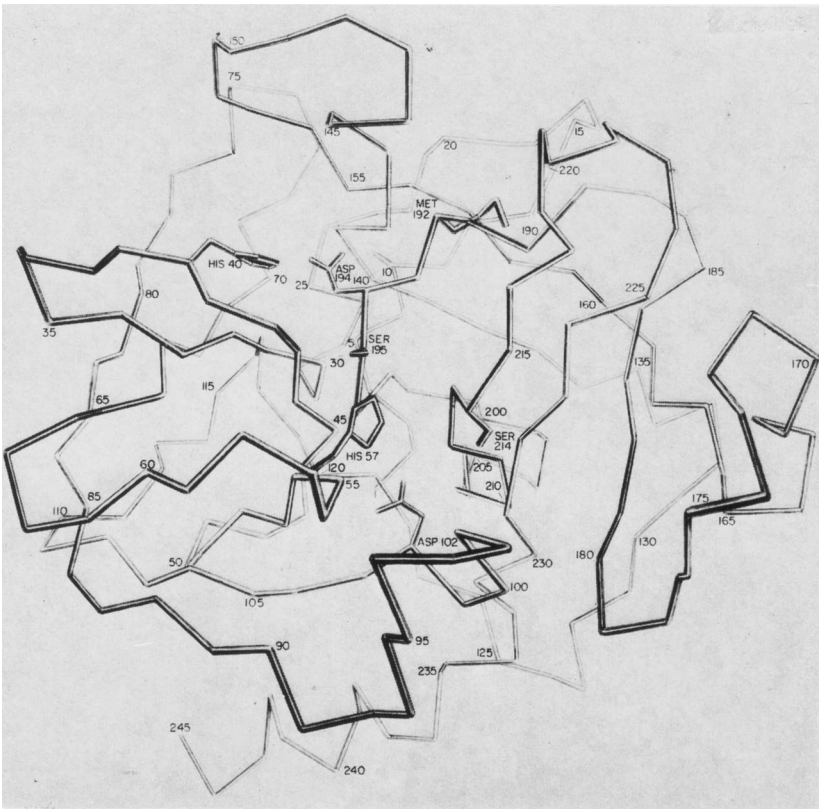


Figure 3 *A line drawing of the tertiary structure of chymotrypsinogen. Only the more important side-chains are shown*
(Reproduced by permission from *Biochemistry*, 1970, **9**, 1997)

[*Facing p. 146*

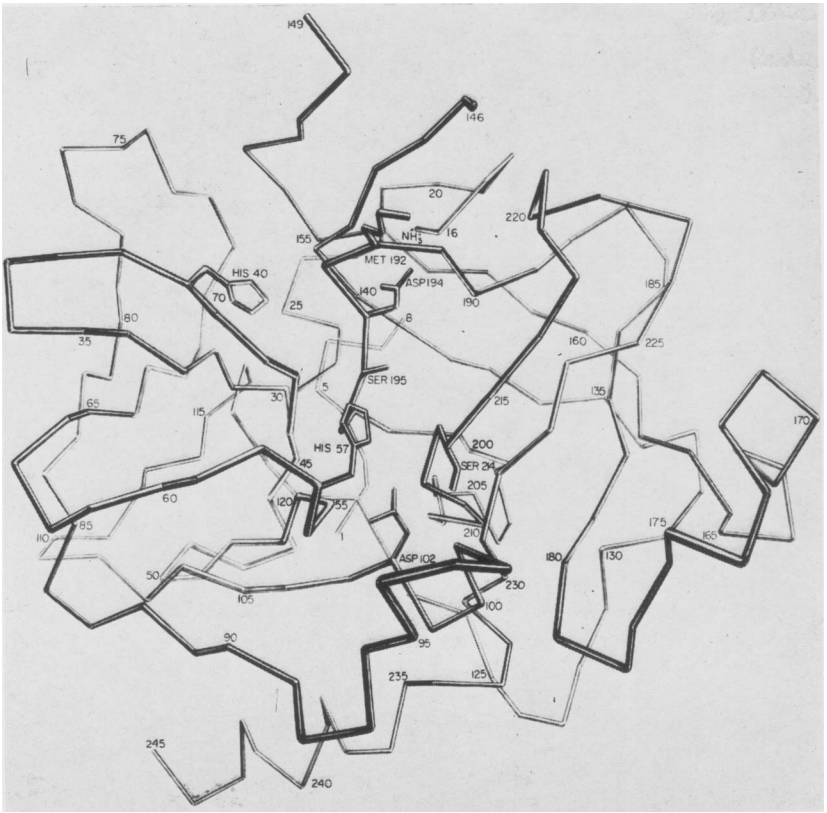


Figure 4 *A line drawing of α-chymotrypsin, shown in the same orientation as chymotrypsinogen*
(Reproduced by permission from *Biochemistry*, 1970, **9**, 1997)

to side-chains in the general vicinity of the active site and have all the signs of occurring in concert, but it is not possible at this stage to assign cause and effect relationships among these molecular reorganizations.

Somewhat surprisingly, these changes in conformation between the zymogen and the enzyme have very little effect in the active site region. In the zymogen the three residues that constitute the charge relay system observed in the enzyme, Ser-195, His-57, and Asp-102, have virtually the same spatial arrangement as they do in the catalytically active enzyme. Two other residues are rearranged in the active site region on activation, Ile-99 moves away from His-57 allowing the solvent greater access to the charge relay system, and Ser-214 moves closer to Asp-102 allowing a hydrogen bond to be formed between the hydroxy- and carboxy-groups. It is not clear what effects these changes have on the catalytic activity of the active site.

At this stage it appears at least possible that the event responsible for the genesis of enzymatic activity upon the transition from inactive zymogen to catalytically active enzyme is the formation of the specificity site. The changes in this region of the molecule on activation are much more pronounced than those occurring to the active site itself. The relatively large movement of residues 187—194 results in an appreciable conformational change which forms the specific side-chain cavity that only partially exists in the zymogen. The particularly large movement that occurs to Met-192 takes it from the interior of the zymogen to the surface of the enzyme to form the lid of the specificity cavity and is of great importance in the formation of the cavity. In addition, the large displacement of Asp-194 away from His-40 and towards Ile-16 completes the formation of the specificity site. It is a reflection of our ignorance of the fundamental basis of enzymatic activity that this study, which allows us to know the detailed changes in the spatial arrangement of nearly all the atoms that take place when chymotrypsinogen is converted to α-chymotrypsin, has not directly lead to an explanation of the genesis of enzymatic activity.

Elastase. The report of the structure determination of another serine protease,[14, 15] elastase, has provided us with yet another insight into the interactions in this group of enzymes. Elastase, which has the ability to digest proteins by breaking peptide bonds at the positions of uncharged non-aromatic side-chains, has a molecular weight of 25 900 and consists of a single polypeptide chain of 240 residues. The complete sequence of the chain has been determined and has been found to be homologous with trypsin and the B and C chains of chymotrypsin.

Elastase crystallises in the orthorhombic space group $P2_12_12_1$ with $a = 51.5$, $b = 58.0$, and $c = 75.5$ Å which corresponds to one molecule in the asymmetric unit. Heavy-atom derivatives were prepared using the tosylated enzyme as the parent form. Crystals of *p*-chloromercuribenzene-

[14] D. M. Shotton and H. C. Watson, *Phil. Trans. Roy. Soc.*, 1970, **B257**, 111.
[15] D. M. Shotton and H. C. Watson, *Nature*, 1970, **225**, 811.

sulphonyl-elastase (PCMBS-elastase) were found to be very highly isomorphous with tosyl-elastase, with the single mercury site almost completely occupied. Two further, but related, derivatives were prepared by soaking both tosyl- and PCMBS-elastase in uranyl nitrate solutions, and again the heavy metal was found to occupy a single site at high occupancy. With the aid of these two highly isomorphous heavy-atom derivatives a Fourier map at 3.5 Å resolution was calculated.

In spite of its limited resolution, the electron-density map of elastase appears to be very clear and easily interpretable. The reasons for this are undoubtedly a combination of factors; very good derivatives which combine the ideals of a high degree of isomorphism with high occupancy; very accurate intensity data collection and a low salt concentration in the crystals that reveals the enzyme in high contrast in the map. However, one disadvantage of the medium resolution is that the orientation of the peptide groups cannot in general be ascertained because of the lack of resolution of the carbonyl groups, and thus the hydrogen-bonding patterns in the molecule can only be determined in general terms.

The polypeptide chain of elastase is folded in an extremely similar manner to the B and C chains of α-chymotrypsin (see Figure 4), and the orientations of corresponding side-chains are also much the same. The elastase polypeptide chain contains twelve more residuces than the B and C chains of α-chymotrypsin and these additions all occur on the molecular surface, usually at the ends of loops of chain which expand to accommodate them without altering the overall structure. The remarkable structural similarity between these two enzymes is explained to some extent by the distribution of homologous residues; the degree of homology in the molecule interior is 81% (see Figure 5) while the overall homology is only 39%.

The most striking feature of the general conformation of elastase is the manner in which the chain is organized into two distinct halves (Figure 5). Within each half the chain is folded into a series of large loops in which the chains run antiparallel. This type of structure has also been noted in the α-chymotrypsin molecule. The two halves of the molecule are similar and each contains six sets of antiparallel interactions. It is interesting to note that the organization within the *N*-terminal half of the chain, traced from the *N*-terminus, resembles that in the *C*-terminal half, but traced in the opposite direction, *i.e.* from the *C*-terminus. Loops possessing similar properties and positions within the molecule occupy corresponding positions in the two halves of the chain. An interesting example of this are the 'histidine' and the 'serine' loops. Within each half of the molecule the antiparallel chains form a twisted sheet which curves round to unite with itself to form a cylinder around the surface of which the chains run obliquely. These cylinders lie one above the other and are orientated approximately at right angles.

The active site of elastase is related to α-chymotrypsin in the way that might have been expected, that is the residues directly involved in bond

cleavage are the same and have the same spatial relation as they do in chymotrypsin but the specificity site is different. In elastase, Ser-195, His-57, and Asp-102 (the numbering is based on chymotrypsinogen A)

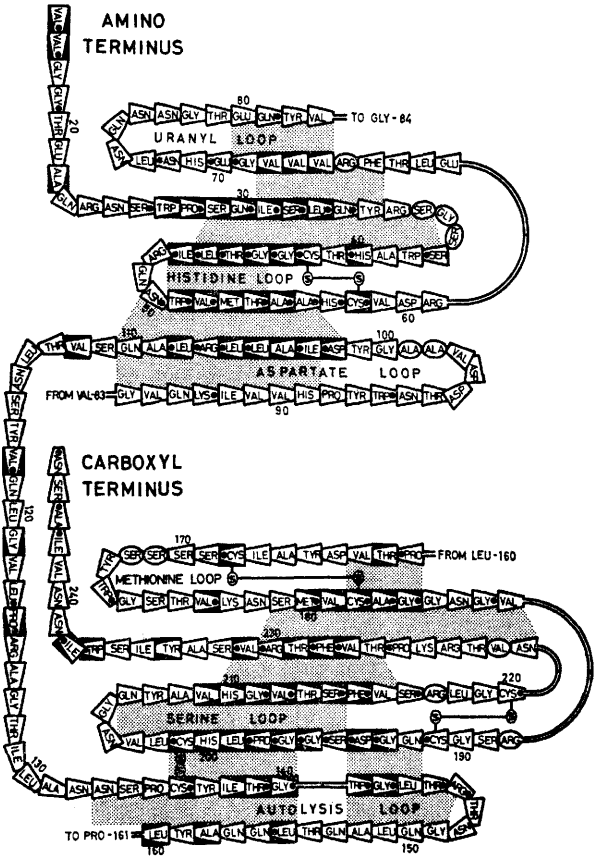


Figure 5 *A schematic representation of the elastase polypeptide chain, folded into its two distinct halves. Neighbouring antiparallel regions are shown shaded. Residues that are homologous in elastase, trypsin, and chymotrypsins A and B have a black dot, while those that are inaccessible to liquid are based in heavy outline*

(Reproduced by permission from *Nature*, 1970, **225**, 811)

occupy the same relative positions as they do in the charge relay system of α-chymotrypsin. In addition, the α-amino-group of Val-6 (the *N*-terminal residue of elastase) forms an internal ion-pair with the carboxy-group of Asp-194. However, there is evidence to suggest that in elastase the ion pair is stable up to pH 10.5, while in α-chymotrypsin the ion pair dissociates

at pH values above 8 with loss of activity and an apparent conformation change that may take the enzyme to a state resembling the zymogen.

The cavity that forms the binding site for aromatic residues in α-chymotrypsin and which appears to give that enzyme its specificity is noticeably changed in the elastase molecule. The most obvious changes are the replacement of Gly-216 at the entrance of the cavity by a valine residue, and Gly-226 at the bottom of the cavity by a threonine. Both changes tend to close the cavity and make it sterically impossible for an aromatic side-chain to be accommodated. The substrate specificity of elastase is directed towards uncharged non-aromatic residues, particularly alanine. It is clear from the elastase model that *N*-formyl-L-alanine is able to bind in an identical fashion to that in which *N*-formyl-L-tryptophan binds to α-chymotrypsin, but with the side-chain methyl group lying above and in contact with Val-216.

Haemoglobin. The structure of human deoxyhaemoglobin has been determined at 3.5 Å resolution,[16] using three derivatives all of which comprise mercury atoms bound to one or more of the three pairs of thiol groups in the molecule. The crystals are monoclinic, space group $P2_1$, with a = 63.2, b = 83.4, and c = 53.8 Å, and β = 99° 15′. The asymmetric unit of the crystal contains a single molecule of 64 000 molecular weight, composed of two α-chains and two β-chains. The heavy-atom derivatives were produced by reacting Cys-112β with mercuric acetate, Cys-93β with *p*-mercuribenzoate and Cys-93β, -112β, and -140α with methyl mercury. The 6900 unique reflexions within the 3.5 Å limit had a mean figure-of-merit of 0.81. In parallel with this study, the resolution of the electron density map of horse deoxyhaemoglobin has now been extended to 2.8 Å.[17] Both structures indicate that $\alpha_1\beta_1$ inter-subunit contact in deoxyhaemoglobin is almost the same as in the oxy form, but that there is a large relative movement at the $\alpha_1\beta_2$ interface with drastic changes in the contacts. The other, more interesting structural differences between oxy- and deoxy-haemoglobin will be discussed in their appropriate places in the following description of the mechanism of haem–haem interaction proposed by Perutz.[18] This proposal is based partly on a comparison of the atomic models of deoxy- and met(oxy)-haemoglobin and partly on unpublished results of an *X*-ray study on a haemoglobin locked into the oxy form by bismaleimido-methyl ether. The extensive arguments presented by Perutz at each stage of his proposed mechanism have been strictly curtailed or omitted here because of the need to discuss the essential structural aspects of the mechanism.

The first step in the mechanism is a consideration of the state of the haem-iron atom. In deoxyhaemoglobin the iron atom is in the ferrous form, five-co-ordinate, and high-spin. The Fourier map of deoxyhaemo-

[16] H. Muirhead and J. Greer, *Nature*, 1970, **228**, 516.
[17] W. Bolton and M. F. Perutz, *Nature*, 1970, **228**, 551.
[18] M. F. Perutz, *Nature*, 1970, **228**, 726.

globin shows that the iron atom is displaced by 0.75 Å out of the plane of the porphyrin ring towards the proximal histidine (F8) to which it is covalently linked (Figure 6). This is in agreement with theoretical predictions that the ionic radius of the high-spin ferrous ion is larger than the

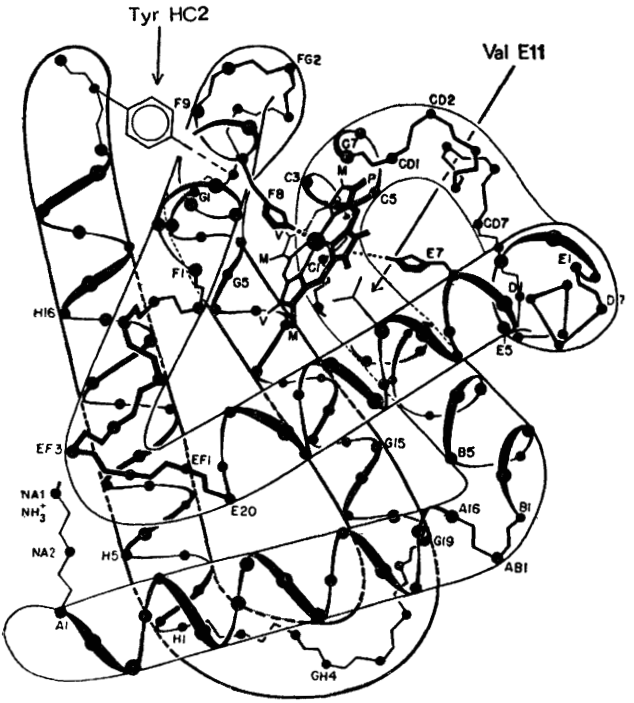


Figure 6 *The course of the polypeptide chain in the β-subunit of haemoglobin. Helical residues are denoted to A to H, while non-helical stretches of chain are denoted by* AB, BC, etc. *The non-helical amino-terminus is* NA *and the non-helical carboxyterminus is* HC. *The β-subunit has essentially the same conformation*
(Reproduced by permission from *Nature*, 1970, **228**, 726)

radius of the 'hole' at the centre of the porphyrin ring. The electron density map of oxyhaemoglobin was obtained from crystals of the acid-met form of the protein, in which the iron atom had been oxidised to the high-spin ferric state and is found to be displaced by 0.3 Å out of the plane of the porphyrin ring. However, it is expected that the true oxy form of haemoglobin has the same overall structure, as far as the globin is concerned, but that the iron atoms, being in the low-spin ferrous state, in which their ionic radii are 0.19 Å less than in the high-spin ferric state, will lie within 0.05 Å of the plane of the porphyrin ring. Thus the transition from the deoxy to the oxy form of haemoglobin is expected to cause the proximal histidine to move by at least 0.75 Å towards the porphyrin ring itself, or

by as much as 0.95 Å if the transition of the central iron atom from high spin to low-spin also causes the bond from the histidine N^{ε} to iron to shorten by 0.2 Å, as has been suggested theoretically but not decided by the *X*-ray studies.

A second important point is the size of the haem pocket that contains the oxygen molecule in the oxy form of the protein. In the α-subunits there is sufficient space for the ligand to occupy the pocket without making short contacts with other side-chains in the deoxy form, and it is found that the transition to the oxy form does not change the geometry of this pocket. On the other hand, the β-subunits do not have a pocket of sufficient size to accommodate even a hydroxide ion. This is largely due to Val-E11(67) whose γ-methyl would be only 2.5 Å from a hypothetical hydroxide ion located on the haem-iron atom. However, comparison of the electron density maps of met- and deoxy-haemoglobin shows that the distance between the porphyrin ring and Val-E11 is increased by 1 Å in going from the deoxy form to the met, or oxy, form of the protein. This shows that in the β-subunits this valine must move relative to the porphyrin ring before the iron atom can react and that in the oxy form there is sufficient room for ligands.

The third important difference between the deoxy and oxy forms of the protein that appears to have an important bearing on the mechanism is concerned with the *C*-termini of the haemoglobin chains (see Figures 9 and 10). In deoxyhaemoglobin the electron density maps show that each of the *C*-terminal residues is doubly anchored by salt-bridges. The α-carboxyl of Arg-HC3(141)α_1 is linked to the α-amino-group of Val-NA1(1)α_2 and its guanidinium groups to Asp-H9(126)α_2. The α-carboxyl of His-HC3(146)β_1 is linked to the ε-amino-group of Lys-C5(40)α_2 and its imidazole ring to Asp-FG1(94)β_1. All four penultimate tyrosines are firmly anchored in pockets between helices F and H. These tyrosines cannot be displaced from these pockets without displacing the *C*-terminal residues and disrupting the salt-bridges. In sharp contrast with these elaborate interactions, in the oxy form the *C*-terminal residues are completely free and the tyrosines appear to spend only a small fraction of their time in the pockets between helices F and H.

The final piece of evidence relates the three previously described effects together and provides the basis for the mechanism. It concerns the changes in the tertiary structure of the haemoglobin molecule that must precede the changes in quaternary structure, and has been obtained from *X*-ray studies of haemoglobin locked in the oxy form by bismaleimidomethyl ether (BME). This is a bifunctional reagent that inhibits all co-operative effects by linking Cys-F9(93) to His-FG4(97) in the same β-chain. The reagent, which lies in the $\alpha_1\beta_2$ contact, also blocks the entry of the penultimate tyrosines HC2(145) of the β-chains into their pockets and causes small changes elsewhere on the β-chains and on the C helix of the α-chains. Crystals of horse met-BME-haemoglobin remain isomorphous

after reduction to the deoxy-BME form, showing that the BME cross-link has locked the protein in its oxy form. An electron density map of the differences between the met- and deoxy-BME-haemoglobins showed striking changes that are in marked contrast to similar maps of myoglobin and erythrocruorin where no significant changes in tertiary structure are seen. The difference map of BME-haemoglobin represents the events that occur during the transition from deoxy- to met- or oxy-haemoglobin.

In the α-chains a positive peak that represents the entry of a water molecule to the previously unoccupied sixth co-ordination site of the haem-iron atom is matched by a negative one of the same magnitude that represents the expulsion of Tyr-HC2(140) from its pocket between helices

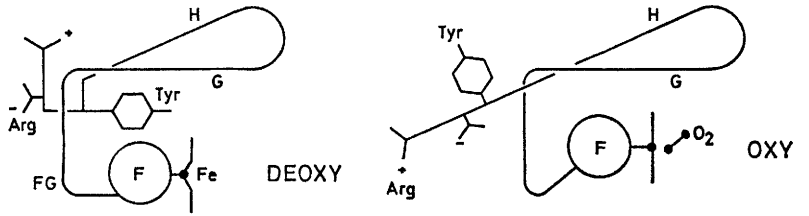


Figure 7 *Schematic diagram of the change in tertiary structure of the haemoglobin subunits on oxygenation*
(Reproduced by permission from *Nature*, 1970, **228**, 726)

F and H. There is a further negative peak in the position of the peptide linking this tyrosine to Arg-HC3(141) that corresponds to the fixed position which the peptide normally takes up in native deoxyhaemoglobin. There is no corresponding positive peak however, showing that it is freely rotating in BME-met, as in native methaemoglobin. There is a pair of positive and negative peaks indicating that helix F has moved inwards towards the centre of the *molecule* narrowing the pocket between it and helix H (Figure 7). This appears to be responsible for the expulsion of the tyrosines. Finally, there are pairs of positive and negative peaks at the positions of the propionic acid side-chains of the haem group showing that its inclination is less steep in the liganded form.

In the β-chains the entry of Tyr-HC2(145) into the pocket between helices F and H is blocked by the cross-link, but movement of helix F nevertheless takes place as if to expel it. The most prominent feature represents the movement of Val-E11(67) showing the widening of the space between the porphyrin and helix E on ligand binding already referred to. There are also peaks near the haem showing that its tilt becomes less steep, as in the α-chains.

These results show that on binding of ligand to the α-subunit, helix F moves so as to expel Tyr-HC2(140) from its pocket. A similar movement of helix F occurs in the β-subunit. On removal of ligand, residues at the

$\alpha_1\beta_2$ contact show signs of strain moving them towards the deoxy conformation, even though this is inhibited by the BME. These incipient changes in quaternary structure appear to be the same as those observed in the oxy-deoxy comparison in the absence of BME. These changes are small shifts of not more than 1 Å at the $\alpha_1\beta_1$ and $\alpha_2\beta_2$ contacts and large shifts of up to 7 Å at the $\alpha_1\beta_2$ and $\alpha_2\beta_1$ contacts.

Using these three pieces of experimental evidence, Perutz has been able to propose a highly plausible mechanism for the haem–haem interaction in haemoglobin. The proposal is discussed in three parts: first, the conformational changes in the individual subunits that follow the binding of ligand, second, the change in quaternary structure caused by them and, finally, the probable order of the steps of the reaction with ligand.

Considering the α-subunit first, oxygenation of its haem moves the iron atom into the plane of the porphyrin ring and causes the haem-linked histidine to move by 0.75—0.95 Å towards the haem. This movement causes helix F to be pulled towards the centre of the molecule, closing the pocket between this helix and helix H and expelling Tyr-HC2(140) from its pocket (Figure 7). The expelled tyrosine pulls Arg-HC3(141) with it, thus breaking its salt-bridges with the opposite α-subunit and releasing Bohr protons. Considering the β-subunits, the ligand cannot reach the iron atom until the haem pocket is prised open by thermal vibrations. When the iron becomes liganded it moves into the plane of the porphyrin ring. The accompanying changes in tertiary structure have one feature in common with those observed in the α-chain: helix F moves towards the centre of the molecule, narrowing the pocket between it and helix H and expelling Tyr-HC2(145). The expelled tyrosine pulls His-HC3(146) with it and breaks its salt-bridge to Asp-FG1(94). The salt-bridge to Lys-C5(40)α may or may not have been broken already before the β-chain could have reacted with the ligand.

The oxygenation of deoxyhaemoglobin is most likely to begin by reaction with the α-subunits, followed by reaction of the β-subunits because the α-subunits have room for ligands while the β-subunits do not. This order of reaction is not, in any case, vital to the proposed mechanism, but is the one shown in the schematic diagram of the mechanism (Figure 8). It is interesting to note that in this case a schematic diagram is necessary because too much is known about the system, not too little. Reaction of $Fe\alpha_1$ of deoxyhaemoglobin with oxygen (1) causes Tyr-HC2(140)α_1 to be expelled from its pocket and the links between Arg-HC3(141)α_1 and the α_2 subunit to be broken with the release of Bohr protons (2). $Fe\alpha_2$ reacts next, followed by the expulsion of its tyrosine and the rupture of the salt-bridges between the *C*-terminal arginine and the α_1 subunit, with further release of Bohr protons (3). The breakage of four of the six salt-bridges that constrain the deoxyhaemoglobin tetramer changes the allosteric equilibrium constant in favour of the quaternary oxy structure. The $\alpha_1\beta_2$ and $\alpha_2\beta_1$ contacts give way and the tetramer clicks into the oxy form,

breaking the remainder of the constraining salt-bridges between Lys-C5(40)α and His-HC3(146)β and the ones between the 2,3-diphosphoglycerate and the β-subunits (see later). This step leads to the release of the diphosphoglycerate but does not release Bohr protons (4).

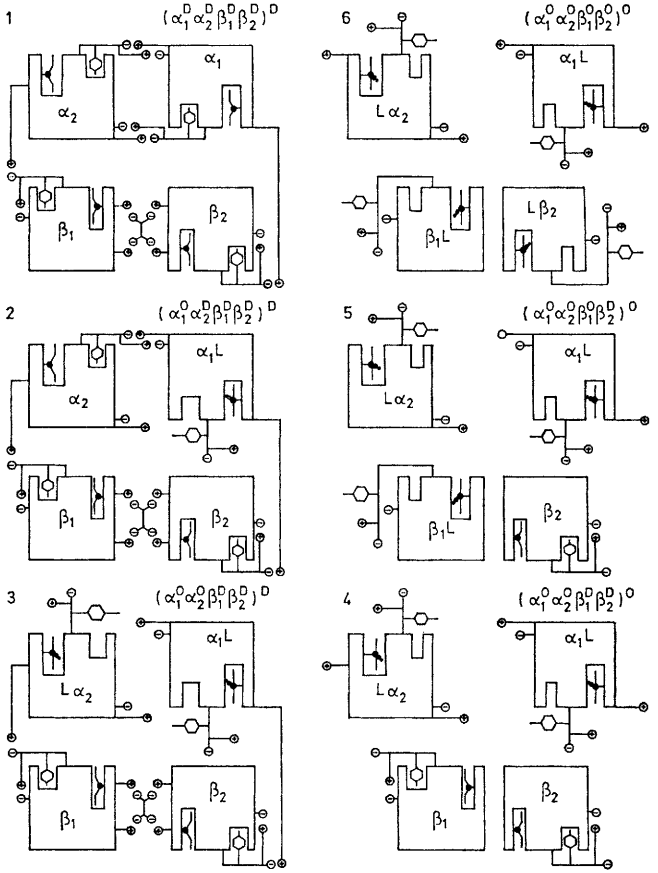


Figure 8 *Diagrammatic sketch showing the sequence of steps in the reaction of haemoglobin with oxygen. Details of the individual steps are given in the text, with the appropriate step number*
(Reproduced by permission from *Nature*, 1970, **228**, 726)

The molecule is now in an intermediate form, with its quaternary structure and the haem and its immediate environment in the α-subunit in the oxy conformation, while the β-subunits have the strained deoxy conformation, that is the tyrosines HC2(145)β are still in their pockets and valines E11(67) still obstruct the haem. The change in quaternary structure has halved the activation energy required to expel the tyrosines because

it has broken the inter-subunit salt-bridges and only internal salt-bridges hold the tyrosines in place, with a resulting increase in oxygen affinity. The iron atoms of the β-chains now react and the tyrosines are expelled rupturing the internal salt-bridges between the histidines HC3(146) and aspartates FG1(94) on the same β-chain and causing the release of the last Bohr protons (5 and 6). It is suggested that all co-operative effects observed on binding of haem ligands take place in the tetrameric molecule while it has the quaternary deoxy structure. Splitting into $\alpha\beta$ dimers may occur at low haemoglobin concentrations or at high salt concentrations, but only after the salt-bridges have been broken, when normally the molecule would take up the oxy conformation.

If haemoglobin is treated as an allosteric enzyme, the mechanism outlined above is in remarkably good agreement with the behaviour of the model allosteric system postulated by Monod, Wyman, and Changeux.[19] Haemoglobin may be regarded as an enzyme made up of two different pairs of subunits, with oxygen as the substrate, haem as the coenzyme, and hydrogen ions and diphosphoglycerate as the allosteric effectors. If α and β denote the subunits in their free conformation then α^s and β^s denote their conformation when substrate is bound, and $(\ )^T$ represents the unreactive and $(\ )^R$ the reactive quaternary structure, then the reaction scheme can be written as

$$(\alpha_2\beta_2)^T \underset{+2s}{\rightleftharpoons} (\alpha_2{}^s\beta_2)^T \rightleftharpoons (\alpha_2{}^s\beta_2)^R \underset{+2s}{\rightleftharpoons} (\alpha_2{}^s\beta_2{}^s)^R$$

The allosteric effectors, H^+ and diphosphoglycerate, both lower the reactivity of haemoglobin; the former strengthen the salt-bridges constraining the molecule in the inactive T-state while the latter may increase the free energy of interaction by introducing additional salt-bridges specific for the T-state, to which it is stereochemically complementary but has either no affinity or a much reduced one for the R-state.

Overall, the proposed mechanism is in agreement with the Monod, Wyman, and Changeux model, but their assumption that all the subunits in the quaternary T-state are unreactive and all those in the R-form are reactive, whether they were liganded or not, has proved too simple, as they themselves suspected. The haemoglobin subunits change their tertiary structure in response, not to the change in quaternary structure, but to the binding of ligand as predicted by Koshland.[20] But Koshland's model implies that the change in tertiary structure of each subunit directly affects the ligand affinity of its neighbours, and there is no evidence for this in haemoglobin. It appears, instead, that the interaction energy arises through step-by-step release of the constraints on the unreactive T-state, which changes the equilibrium in favour of the R-state and diminishes the work required to change the tertiary structure of each subunit from the unreactive to the reactive form.

[19] J. Monod, J. Wyman, and J. P. Changeux, *J. Mol. Biol.*, 1965, **12**, 88.
[20] D. E. Koshland, G. Nemethy, and D. Filmer, *Biochemistry*, 1966, **5**, 365.

In a further paper,[21] Perutz provides more detailed information on the Bohr effect and the binding of organic phosphates to haemoglobin. The Bohr effect is observed only in reactions, such as oxygenation, which involve the change in quaternary structure from the deoxy to the oxy form. At a pH above 6.0 haemoglobin takes up protons on release of haem ligands and reaches a maximum uptake of 0.7 protons per haem ligand at physiological pH. It is clear that the changes in quaternary structure that take place when ligands are bound also alter the environment of ionisable groups away from the haem group. Examination of the high-resolution

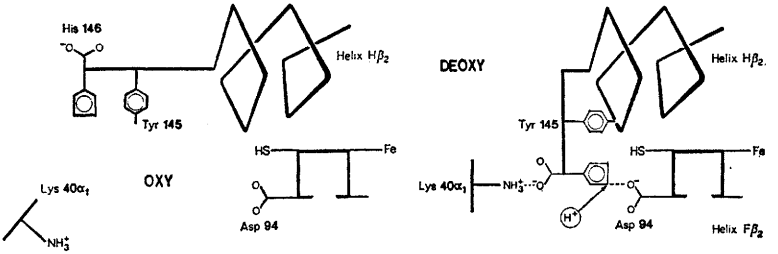


Figure 9 *The change in conformation of the C-terminal residues of the β-chain of haemoglobin on deoxygenation*
(Reproduced by permission from *Nature*, 1970, **228**, 734)

electron density map of oxy- and deoxy-haemoglobin has permitted the exact stereochemical mechanism of the Bohr effect to be defined.

The contribution of the β-chains to the Bohr effect arises from the freedom of the *C*-terminal histidines in oxy- and the linkage of their imidazoles to aspartates 94 of the same β-chains in deoxy-haemoglobin (Figure 9). In oxyhaemoglobin, the side-chain of Lys-C5(40)α_1 is protruding into the liquid, but during deoxygenation the α_1-chains turn relative to the β_2 chain by 13.5° causing the lysine side-chain to shift by 7 Å to form a salt-bridge with the *C*-terminal carboxy-group of the β_2 chain. This salt-bridge keeps Tyr-145β_2 fixed in its pocket and allows a further salt-bridge between the imidazole of the *C*-terminal His-146β_2 and Asp-94β_2 to be formed. This suggests that the *C*-terminal histidine contributes to the Bohr effect by a change in p*K* of its imidazole arising through the shift from its free position in oxy to its fixed position near a carboxy-group in deoxyhaemoglobin.

At least part of the contribution of the α-chains to the Bohr effect arises from the α-amino-groups of the *N*-terminal valine residues. The high-resolution electron-density maps indicate that the *C*-terminal arginines form two salt-bridges with its partner chain in deoxyhaemoglobin (Figure 10). One bridge extends from the guanidinium group α_2 to Asp-H9(126)α_1 and the other from the α-carboxy-group α_2 to a cluster of three ionisable

[21] M. F. Perutz, *Nature*, 1970, **228**, 734.

groups, Val-NA1(1), Asp-A4(6), and Lys-H10(127)α_1. In oxyhaemoglobin the valines-NA1(1)α are free and this suggests that the pKs of the α-amino-groups are normal in oxyhaemoglobin but that they are raised in deoxy-haemoglobin by the salt-bridge to the α-carboxy-group. However, the proximity of the charged ε-amino of lysine 127 may weaken the effect of the carboxy-group and the ionisation of the α-amino-group may not account

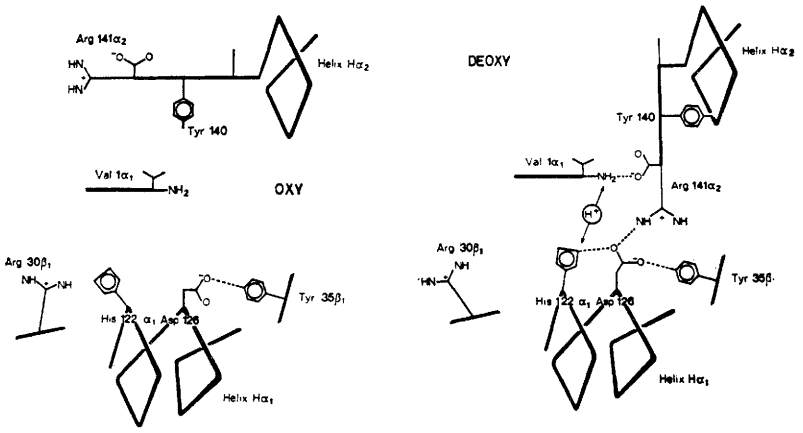


Figure 10 *The change in conformation of the* N- *and* C-*terminal residues of the* α-*chain of haemoglobin on deoxygenation.*
(Reproduced by permission from *Nature*, 1970, **228**, 734)

for the entire contribution of the α-chains to the Bohr effect. There is, however, a second possible contributing group. The link from Arg-HC3(141)α_2 to Asp-H9(126)α_1 causes a rearrangement in a cluster of polar groups that include His-H5(122)α_1 (Figure 10). The rearrangement results in the histidine moving 2 Å closer to a neighbouring Arg-B12(30)β, and 1.5 Å away from the neighbouring Asp-H9(126)α_1 during the transition from deoxy- to oxy-haemoglobin, and its pK is lowered. Thus the expulsion of Tyr-HC2(140) on oxygenation of Feα would rupture the links of the *C*-terminal arginines with Val-1 and Asp-H9. This would liberate the Bohr protons of the α-amino-groups of Val-1, but it is not clear why this would affect the conformation of His-H5, relative to its two neighbours, and their movements may depend on the change in quaternary structure.

2,3-Diphosphoglycerate (DPG) is present in the red blood cells of humans and many other species, and lowers the oxygen affinity of haemoglobin in a physiologically advantageous way. It binds specifically to deoxyhaemoglobin in a 1 : 1 ratio and increases the free energy of haem–haem interaction but does not alter the Bohr effect. The fact that it binds to free β-chains (β_4) in both oxy and deoxy forms suggests that the binding site is on the β-chains. Model-building DPG into the atomic model of deoxyhaemoglobin suggests that it fits into the central cavity with its

phosphate groups hydrogen bonded to Val-NA1(1)β, Lys-EF6(82)β, and His-H21(143)β. On oxygenation the distance between the α-amino-groups increases by 16—20 Å, breaking the hydrogen bonds to DPG, and helices H close up so that DPG is expelled from the central cavity. The influence of DPG may consist simply in stabilising the quaternary structure of the deoxy form where it may form the equivalent of four additional salt-bridges cross-linking the β-chains, but it probably has no direct effect on the tertiary structure of individual subunits.

Erythrocruorin. An *X*-ray investigation of the structural differences between the deoxy and carbonmonoxy forms, and the met form of the monomeric insect haemoglobin erythrocruorin [22] has been carried in terms of the 2.5 Å structure of meterythrocruorin. The deoxy form was prepared from meteryrthrocruorin in the crystal by soaking in solutions containing an excess of sodium dithionite, and *X*-ray data were collected from these crystals sealed in capillaries under nitrogen. The carbonmonoxy form was prepared from the deoxy form by replacing the nitrogen in the capillaries with carbon monoxide. These two forms were isomorphous with the met form apart from some small intensity changes, indicating that no large changes in structure had taken place.

The difference electron-density map calculated for deoxy- *versus* meterythrocruorin had, as its most prominent feature, a negative peak at the position of the haem-bound water molecule, showing that in the deoxy form this ligand has been lost and that the iron has become penta-coordinate. In the difference map of carbonmonoxy *versus* deoxy, the most significant feature is a large peak near the haem iron that represents the binding of the carbon monoxide molecule. This peak is elongated and shows that the bond to the iron is inclined to the C–O direction by 145°. Other features in the map suggest that the binding of carbon monoxide is accompanied by a slight movement of the haem, a movement of the iron atom towards the haem, and shifts in the positions of Ile-E11 and Phe-H14. These changes appear to be caused by a transition of the iron atom's spin state and the relief of overcrowding in the haem pocket as a result of carbon monoxide binding.

Rubredoxin. Rubredoxin is a non-haem iron protein, isolated from *Clostridium pasteurianum*, that appears to function in electron transport. The molecule has a molecular weight of 6300 and contains a single iron atom. Crystallisation [23] was accomplished in 75% saturated ammonium sulphate at pH 4.0. The resulting crystals have rhombohedral symmetry, space group *R*3 with cell dimensions $a = 38.77$ Å and $\alpha = 112.37°$. The structure was solved at 2.5 Å resolution, using two heavy-atom derivatives formed by soaking the crystals in K_2HgI_4 and $UO_2(NO_3)_2$ solutions.

[22] R. Huber, O. Epp, and H. Formanek, *J. Mol. Biol.*, 1970, **52**, 349.
[23] J. R. Herriott, L. C. Sieker, L. H. Jensen, and W. Lovenberg, *J. Mol. Biol.*, 1970, **50**, 391.

The electron density map of rubredoxin was readily interpretable in terms of the main-chain conformation and the arrangement around the iron atom, but the lack of sequence data has prevented definition of the side-chains. The approximately 53-residue polypeptide chain is folded in the fairly simple double hairpin conformation shown in Figure 11. Four cysteine

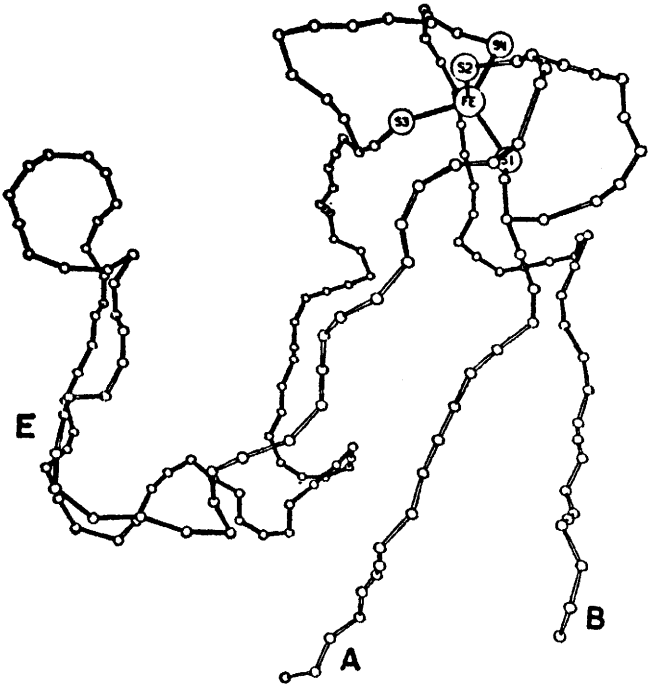


Figure 11 *A drawing of the main-chain conformation of rubredoxin. The iron atom is linked to four cysteine residues arranged tetrahedrally around it*
(Reproduced by permission from *J. Mol. Biol.*, 1970, **50**, 391)

residues that probably occur in closely-spaced pairs in the polypeptide chain form the ligands for the single iron atom. The configuration of the sulphur atoms about the iron does not differ significantly from a regular tetrahedron with mean Fe–S distances of about 2.3 Å.

Lactate Dehydrogenase. The electron-density map of the M_4 isozyme of dogfish lactate dehydrogenase reported last year has now been extended to 2.8 Å resolution [24] and the detailed structure of the molecule is now apparent. Unfortunately, the lack of a complete amino-acid sequence restricts the interpretation of the electron density map to some extent. The following description involves only one of the four exactly equivalent

[24] M. J. Adams, G. C. Ford, R. Koekoek, R. J. Lenz, A. McPherson, M. G. Rossmann, I. E. Smiley, R. W. Schevitz, and A. J. Wonacott, *Nature*, 1970, **227**, 1098.

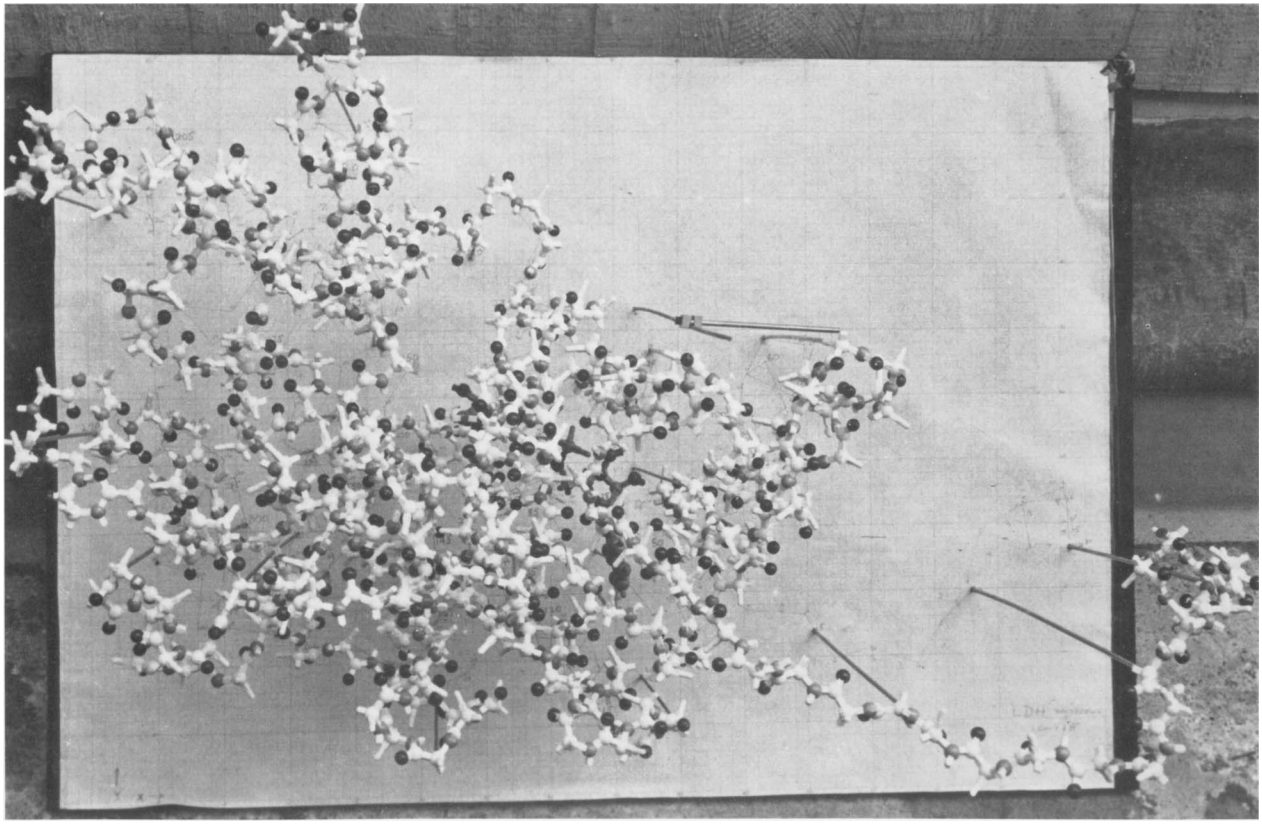


Figure 12 *A view of a single LDH subunit showing the main-chain conformations. The coenzyme is in grey, approximately in the centre of view*

[*Facing p. 160*

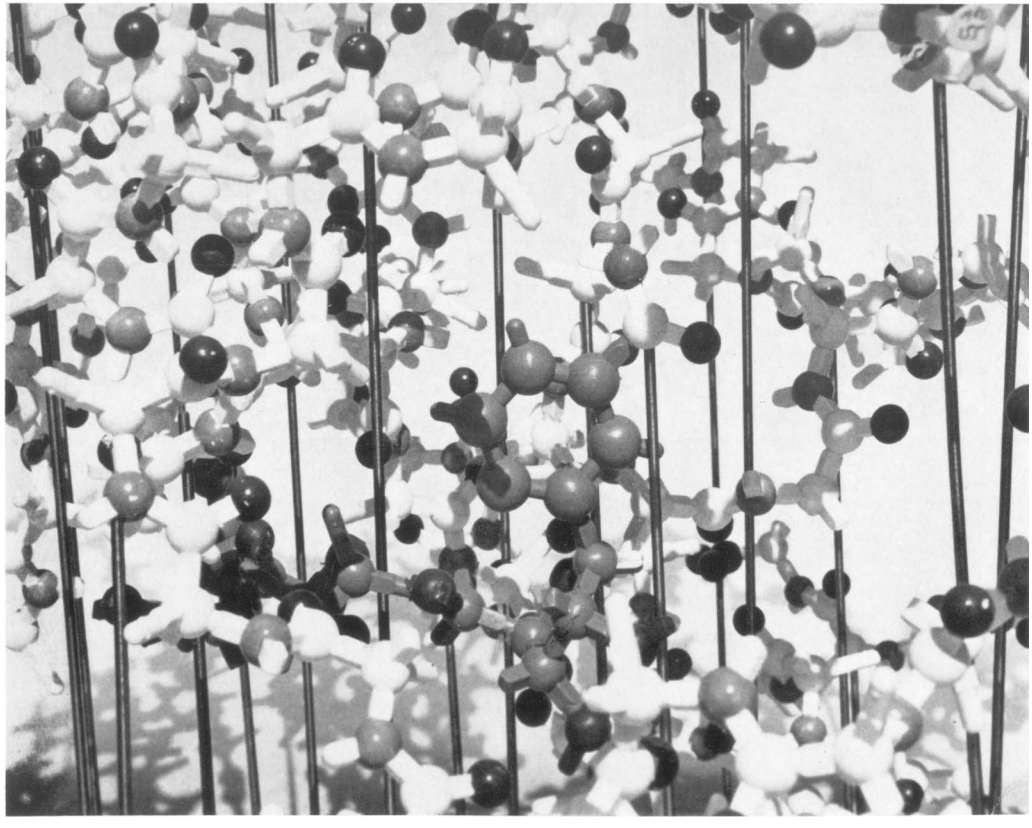


Figure 13 *A close-up view of the coenzyme (shown grey) bound to LDH. The nicotinamide end of the coenzyme is in the centre, buried deeply in the molecule, while the phosphate chain leads to the adenine ring on the left*

subunits that make up the whole molecule. About 24% of the 311—326 residues in the subunit are located in eight helices, while a further 10—15% are found in a well-defined parallel-pleated sheet and a less well-defined antiparallel-pleated ribbon. The overall shape of the molecule is globular with a considerable cleft dividing it into two roughly equal halves, but the first twenty residues form a long tail that extends some distance away from the main body of the subunit (Figure 12). One half of the globular part of the subunit consists mainly of the *N*-terminal 130 residues and has the greater proportion of secondary structure. It contains four helices (42 residues) and 23 residues in a stretch of parallel pleated sheet. The adenine part of the coenzyme binding site [25] shown in Figure 13 is in this part of the molecule while the nicotinamide end is buried deeply in the cleft. The part of the subunit on the other side of the cleft has a less obvious secondary structure and contains the 'essential' thiol that borders on the cleft. These two halves of the molecule are joined by an imperfect antiparallel pleated ribbon. Helix H, residues 289—300, is slightly separated from the remainder of the subunit.

There is some ambiguity in the placing of the first 20 residues and the choice of the unusual extended conformation described above was made because the density is more continuous. In this position the *N*-terminal chain interacts with a second subunit, with the helix formed by the first six residues wrapped around the far side of the second subunit. If this conformation is correct, the molecule can be considered as two pairs of subunits and the subunits making up each pair have their 'arms' around one another. In the alternative conformation the *N*-terminal arm is not so extended and makes fewer contacts with other subunits than does the former conformation. Residues 22—26 run down to form the central strand of the parallel-pleated sheet, the chain then runs up to form helix B(32—40) which lies parallel to the sheet and runs down again to form a second strand of the sheet (45—48), and then goes on to another helix (C; 55—72). The chain then passes above the sheet and a fourth strand forms next to the central one (78—86). A loop across the back of the subunit, which includes helix D (95—106), forms the lip of the central cleft. Helix E (109—118) follows immediately and the fifth strand of the pleated sheet (121—123) is formed next to the cleft. The residues immediately following the three central strands of the sheet are involved in contacts to adenine, adenine ribose, phosphates, and nicotinamide ribose, while the *C*-terminal residue of helix D is in contact with adenine. The essential thiol peptide bounds the central cleft opposite the coenzyme binding site. The thiol itself is separated from the cavity by its own main-chain. The opening of the cleft to the solvent is made by the loop (86—109) joining the fourth strand of parallel sheet and helix D. Examination of the abortive ternary

[25] M. J. Adams, A. McPherson, M. G. Rossmann, R. W. Schevitz, and A. J. Wonacott, *J. Mol. Biol.*, 1970, **51**, 31.

complex [26] at 5 Å resolution shows that this loop has moved by about 12 Å to close the entrance to the cleft and cover the nicotinamide end of the coenzyme. A rather isolated peak in the 5 Å map of the apo-enzyme is also present in the ternary complex; it is located between the essential thiol and the nicotinamide and may represent substrate binding.

The interactions between the four subunits cannot be precisely defined until more sequence information becomes available, but some general indications have been found. Subunits related by the two-fold axis parallel to *c* have helices C, B, and G of one in contact with G, B, and C respectively of the other, and the equivalent *N*-terminal helices A make contacts with wing-like features formed by residues 194—209. The subunits related by the two-fold axes parallel to *a* and *b* make contacts through the pleated sheet and some random coil at residues 160—170. The coenzyme binding sites are wholly within each subunit, with closest approaches between adjacent coenzymes of about 19 Å while nicotinamide ends are as much as 26 Å apart.

Ribonuclease S. The resolution of the electron-density map has been increased from 3.5 to 2 Å.[27] The new map has been interpreted in terms of the known sequence and the original interpretation of the 3.5 Å resolution map. It is found that the level of confidence of the proposed structure varies markedly in different parts of the protein. A preliminary description of the structure indicates that the molecule is roughly divided into two parts, with the three α-helices, residues 3—13, 24—34, and 50—59, and a piece of 3_{10} helix involving residues 15—20, in one half, and the bulk of the extensive, but irregular β-structure in the second half. More detailed structural analysis of the map is to be published later and consideration of the molecular structure of ribonuclease S will be left until these later papers appear, even though some fragmentary information is available at present.

Preliminary Data for Other Proteins. (*a*) *Lysyl-tRNA synthetase.*[28] This important and interesting enzyme from yeast has been crystallised in a form that may be suitable for further *X*-ray analysis. The crystals were grown from 5—10% protein solutions in 0.2M phosphate at pH 7.0 by dialysis against 0.1M phosphate. The crystals are trigonal, space group $P3_121$, with $a = 118$ and $c = 190$ Å, and contain a single molecule of 100 000 molecular weight in the asymmetric unit. Unfortunately, the crystals are very sensitive to *X*-rays, and this could be related to the high solvent content of the crystals, 68%.

(*b*) *Alkaline phosphatase.*[29] Large single crystals of alkaline phosphatase from *E. coli* have been prepared by warming solutions of the enzyme in

[26] I. E. Smiley, R. Koekoek, M. J. Adams, and M. G. Rossmann, *J. Mol. Biol.*, 1971, **55**, 467.
[27] H. W. Wyckoff, D. Tsernoglou, A. W. Hanson, J. R. Knox, B. Lee, and F. M. Richards, *J. Biol. Chem.*, 1970, **245**, 305.
[28] L. Rymo, V. Lagerqvist, and A. Wonacott, *J. Biol. Chem.*, 1970, **245**, 4308; L. Rymo and V. Lagerqvist, *Nature*, 1970, **226**, 77.
[29] A. W. Hanson, M. L. Applebury, J. E. Coleman, and H. W. Wyckoff, *J. Biol. Chem.*, 1970, **245**, 4975.

56% saturated ammonium sulphate at pH 8.0, from 4 to 25 °C. The space group is $P3_121$ with $a = 70.5$ and $c = 155.6$ Å. The asymmetric unit contains one subunit of the 80 000 molecular weight dimeric molecule.

(*c*) *Concanavalin A.*[30] Concanavalin A is a protein from the Jack bean that has a molecular weight of 55 000. It binds transition metals, calcium ions, and α-methyl-D-glucosopyranoside at three pairs of mutually interacting sites. The protein has been crystallised from 0.1M sodium nitrate buffered at pH 6.5 with 0.05M tris-acetate buffer. The crystals are orthorhombic, space group either $I222$ or $I2_12_12_1$, with $a = 87.2$, $b = 89.2$, and $c = 62.9$ Å, indicating a half molecule in the asymmetric unit.

(*d*) *Human Bence Jones protein.* An L-type Bence Jones protein from a single patient has been crystallised from salt-free solutions.[31] The crystals are orthorhombic space group $P2_12_12$ with $a = 72.6$, $b = 81.9$, and $c = 71.0$ Å. The asymmetric unit of the crystals contains one disulphide-linked dimer.

In a second preliminary report[32] the crystallographic parameters of the variant half of a human κ Bence Jones protein are given. The monomeric form of the Bence Jones protein was cleaved by pepsin into its variant half and constant half. Crystallisation was performed either directly from the reaction mixture or after purification of the variant half. The crystals are monoclinic, space group $P2$ with $a = 65.6$, $b = 37.9$, and $c = 43.5$ Å, and $\beta = 90.0°$, indicating that there are two molecules of 11 500 molecular weight in the asymmetric unit.

(*e*) *Human IgG.*[33] A serum IgG immunoglobulin, that originated from the same myeloma patient who was the source of the L-type Bence Jones protein described above, has been crystallised. This ensures that the light chains of the immunoglobin have the same amino-acid sequence as the Bence Jones protein. Crystallisation of the immunoglobulin took place after prolonged dialysis against salt-free water at room temperature. The crystals are orthorhombic, space group $C222_1$, with $a = 87.8$, $b = 111.3$, and $c = 186.3$ Å. The asymmetric unit is composed of a half molecule, that is one light chain and one heavy chain, showing that the two halves of the molecule are related by a two-fold rotation axis. These crystals appear to be much less sensitive to radiation damage than those previously described.

(*f*) *Rennin-like enzyme of* Endothia parasitica.[34] An enzyme that has a similar catalytic property to calf rennin has been isolated from the fungus

[30] J. Greer, H. W. Kaufman, and A. J. Kalb, *J. Mol. Biol.*, 1970, **48**, 365.
[31] M. Schiffer, K. D. Hardman, M. K. Wood, A. B. Edmundson, M. E. Hook, K. R. Ely, and H. F. Deutsch, *J. Biol. Chem.*, 1970, **245**, 728.
[32] A. Solomon, C. L. McLaughlin, C. H. Wei, and J. R. Einstein, *J. Biol. Chem.*, 1970, **245**, 5289.
[33] A. B. Edmundson, M. K. Wood, M. Schiffer, K. D. Hardman, C. F. Ainsworth, K. R. Ely, and H. F. Deutsch, *J. Biol. Chem.*, 1970, **245**, 2763.
[34] P. C. Moews and C. W. Bunn, *J. Mol. Biol.*, 1970, **54**, 395.

Endothia parasitica and crystallised. Two crystal forms were grown from 2.2M ammonium sulphate in the pH range 4.5—6.3 with the addition of a small amount of acetone or dimethylformamide. Both forms are monoclinic, space group $P2_1$, and contain one molecule of 36 000 molecular weight in the asymmetric unit.

(*g*) *Thioredoxin from* E. coli B.[35] Thioredoxin, the hydrogen donor in the reduction of ribonucleotides to deoxyribonucleotides, has been isolated from *E. coli B* and crystallised from 10^{-3}M cupric acetate in 48—50% 2-methyl-2,4-pentanediol buffered at pH 3.7—4.1 by 0.1M acetate. The crystals are monoclinic, space group $C2$ with $a = 89.5$, $b = 50.8$, and $c = 60.2$ Å, and $\beta = 113.5°$. The asymmetric unit of the crystals contains two thioredoxin molecules of 11 700 molecular weight.

(*h*) *Proinsulin.*[36] The single-chain insulin precursor proinsulin has been crystallised[36] at pH 3 from ammonium sulphate. The crystals are tetragonal, space group $P4_12_12$ or $P4_32_12$ with $a = 50.8$ and $c = 148.0$ Å, which suggests that the asymmetric unit is composed of two molecules of 8700 molecular weight.

PART III: Spectroscopic and Solution Studies on the Conformation and Interactions of Polypeptides and Proteins

edited by A. R. Peacocke, with contributions by J. R. Brocklehurst, D. G. Dalgleish, R. Henson, P. H. Lloyd, N. C. Price, and R. M. Stephens

1 Introduction

by A. R. Peacocke

This year's Report follows the same form as that in Volume 2, since the major developments in these areas continue to result chiefly from the application of the techniques of i.r., o.r.d., c.d., n.m.r., fluorescence spectroscopy, and e.s.r., and from the use of the various equilibrium and hydrodynamic methods for studying the dissociation–association behaviour of proteins. In addition to the papers and reviews referred to in the sections following, attention should also be drawn to a survey by Timasheff[1] of some physical probes of enzyme structure in solution and to a monograph on interacting macromolecules by Cann[2] which describes the developments in the theory and analysis of moving boundaries of such systems in electrophoresis, ultracentrifugation, and chromatography.

[35] A. Holmgren abd B.-O. Söderberg, *J. Mol. Biol.*, 1970, **54**, 387.
[36] W. W. Fullerton, R. Potter, and B. W. Low, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 1213.
[1] S. N. Timasheff, 'Some Physical Probes of Enzyme Structure in Solution', in 'The Enzymes', ed. P. D. Boyer, Academic Press, New York and London, 1970, Vol. II, p. 371.
[2] J. R. Cann, 'Interacting Macromolecules,' Academic Press, New York and London, 1970.

2 Infrared

contributed by R. M. Stephens

The helical sense of polypeptides is not usually found from i.r. spectroscopic studies. However, in the case of polyaspartate esters it has been shown that it is possible to predict the helix sense from the amide vibrational frequencies. Using this technique, Malcolm [3] has shown that the amide I and II vibrations obtained from monolayer films of poly-β-benzyl-L-aspartate cast on a water surface occur at 1658 cm^{-1} and 1552 cm^{-1} respectively. These frequencies are characteristic of the polymer forming a right-handed α-helix, which is contrary to the normal left-handed α-helical conformation of poly-β-benzyl-L-aspartate, and the conclusion is drawn that a left-handed to right-handed α-helix transition is effected by the interactions of the water. I.r. studies on monolayer films of poly-ε-benzyloxycarboxyl-L-lysine cast on water surfaces have shown that its conformation was α-helical and was not affected by the compression of the monolayer.[4] For polypeptide films obtained from water surfaces it might be expected that some water would be absorbed, and polarised i.r. studies on partially *N*-deuteriated polypeptides have allowed observations to be made on the symmetrical and antisymmetrical OH stretching vibrational bands of water around 3500 cm^{-1}. These suggest that the water molecules are adsorbed at specific sites and orientations with respect to the substrate molecules. In poly-L-alanine, water orientates with a hydrogen atom directed towards the peptide oxygen and forming a weak hydrogen bond to it. However, with poly-methyl- or -ethyl-glutamate an interaction may occur with the side-chain ester group.[5]

The helical conformation of poly-L-tyrosine in solution has not been determined because the presence of the side-chain chromophore prevents unambiguous interpretation of the results from o.r.d. studies. Consequently, i.r. studies of solutions and films have been made. For films of poly-L-tyrosine cast from aqueous solutions at pH 11.65 and pH 11.25 the results showed that the conformations were random coil and antiparallel β-sheet respectively; in D_2O solutions at pD 12.15 and pD 11.66 the polymer had, respectively, a random and antiparallel β-sheet conformation.[6] The amide I absorption obtained from films of poly-L-tyrosine cast from $(MeO)_3PO$ and solutions in $(MeO)_3PO$ were characteristic of an α-helical conformation, but the helix sense was not determined.[7]

The interpretation of the far-i.r. vibrational frequencies obtained from poly-L-alanine and poly-L-α-amino-n-butyric acid between 700 cm^{-1} and 100 cm^{-1} has been aided by the calculation of the normal co-ordinate vibrational frequencies of the helix. A comparison between the observed

[3] B. R. Malcolm, *Biopolymers*, 1970, **9**, 911.
[4] B. R. Malcolm, *Biochem. J.*, 1970, **110**, 733.
[5] B. R. Malcolm, *Nature*, 1970, **227**, 1358.
[6] E. Patrone, G. Conio, and S. Brighetti, *Biopolymers*, 1970, **9**, 897.
[7] F. Quadrifoglio, A. Ius, and V. Crescenzi, *Makromol. Chem.*, 1970, **135**, 241.

absorption frequencies and the calculated values has shown that the majority of the bands have come from either helical backbone deformation modes or out-of-plane bending modes [8] of the C=O group. A mixture of glutamic acid and α-keto-glutaric acid has been analysed by using the i.r. absorption bands characteristic of each one. Using successive approximations the absorbance measurements gave the relative weights of each acid present.[9] The structural analysis of poly-γ-L-menthyl-L- and -D-glutamates by i.r., *X*-ray, o.r.d., and c.d. techniques has shown that the L-menthyl-L-glutamate is right-handed and has an amide I frequency at 1665 cm^{-1}, which is higher for an L- polymer than has previously been found with right-handed α-helical polypeptides. The D-polymer had an amide I frequency of 1660 cm^{-1}.[10] Thin polypeptide films of poly-γ-methyl-L-glutamate in different conformations have been studied by internal reflection i.r. spectroscopy. The ability of these films to be wetted correlated well with structural assignments deduced from the surface spectra. Coiled polymer configurations correlated with the wetting behaviour typical of an accessible polyamide backbone capable of hydrogen-bonding interactions with contacting liquids. Extended-chain configurations exhibited values of contact angles which were indicative that only the methyl ester side-chain was in the outermost surface layer and that there was no accessibility to the polypeptide backbone for hydrogen-bonding liquids. Since structural transformations similar to these are possible with proteins in various environments, the relative importance of side-chain backbone contributions to the wettability of the proteins–air interfaces is markedly dependent on protein conformation. For poly-γ-benzyl-L-glutamate the hydrogen-bonding capacity of the backbone was masked by the bulky benzyl side-chain without requiring conformational changes of the polymer chain. The influence of adsorbed water was apparently minor on these insoluble polypeptides, but may be important with soluble biopolymers or biopolymers in aqueous environments.[11] Attenuated total reflection has been used to determine the skin and case structure of nylon-66 fibres.[12]

The *cis–trans* isomerism of amino-acid esters and their hydroxamic acid hydrochloride derivatives has been discussed from the results of their i.r. spectra [13] and a review of the spectroscopic and related evidence for *cis*, *trans*, and *skew* conformations of the amide group has been undertaken by Hallam and Jones.[14] The i.r. spectra of eight known depsipeptides and their dipole moments have shown that these 10-, 11-, and 12-membered ring systems exhibit variation of *cis–trans* forms that depend on ring size and substituent structure. In the 10-membered ring the conformations realised

[8] K. Itoh and T. Shimanouchi, *Biopolymers*, 1970, **9**, 383.
[9] F. L. Estes, A. L. Myers, and S. Briney, *Appl. Spectrosc.*, 1970, **24**, 131.
[10] H. Yamamoto, Y. Kando, and T. Hayakawa, *Biopolymers*, 1970, **9**, 41.
[11] R. E. Bairer and W. A. Zisman, *Macromolecules*, 1970, **3**, 70.
[12] G. Heidemann, *Chemiefasern Text. Anwendungstech.*, 1970, **20**, 204.
[13] J. Rosochacka, *Roczniki Chem.*, 1970, **44**, 321.
[14] H. A. Hallam and C. M. Jones, *J. Mol. Struct.*, 1970, **5**, 1.

are those favouring intramolecular interaction of the amide and ester group. Although transannular interactions are also possible they are reduced in larger ring systems. The dipole moments of these peptides ranged from 3.4 D to 4.9 D.[15] The influence of solvent effects and hydrogen-bond formation on the anharmonicity of the N–H stretching vibrations of simple secondary amides and backbones has been studied with the conclusion that the N–H stretching mode always exhibits the same anharmonic behaviour but is different from the anharmonicity of the O–H stretching vibrations.[16] The analysis of the i.r. spectra of the dipeptides $MeCO \cdot NH \cdot CHR^1 \cdot CO \cdot NH \cdot CHR^2$ and some of their *N*-deuteriated and *N*-alkylated derivatives in dilute chloroform solutions indicates that these compounds may have two conformations: an extended form consisting of a five-membered chelation ring and a bent form possessing a seven-membered chelation ring. A semi-extended form without chelation is sometimes identified in very small proportions. The equilibrium of these rotational isomers is solvent dependent. The results obtained suggest that solvation by proton-donating or proton-accepting solvents tends to open the chelation ring in favour of the semi-extended form. The study of the effects of the side and terminal groups on the N–H stretching frequency has indicated the position of these groups with respect to the peptide skeleton. For the bent form, the side-chain is in the equatorial position relative to the mean plane of the chelation ring.[17] Similar studies have shown that the amount of the bent form increases when the side-chain is large.[18]

The principal characteristics of the i.r. spectra of membranes from the bacterium *M. lysodeikticus* originate from the amide I and II bands from the proteins and from the vibrations of the O–H, C–H, C=O, P=O, C–O–C, and P–O–C groups present in the lipids. In addition, the asymmetrical and symmetrical methylene stretching vibrations at 2930 cm^{-1} and 2855 cm^{-1} respectively have provided a parameter interpretable in terms of protein–lipid interactions. Fractions of the protein of lipid-depleted membranes have substantially higher CH_2 asymmetrical/symmetrical absorbance ratios than those observed for whole membranes.[19] Analysis for the far-i.r. absorption spectra of bovine serum albumin, ribonuclease, lysozyme, and myoglobin has shown that despite differences in primary, secondary, and tertiary structures, all the proteins had similar spectra in this region, with an intense band at approximately 150 cm^{-1}. This band probably arises from anharmonic couplings through hydrogen-bonded groups.[20] A quantitative study of the amide I vibrations from

[15] T. M. Ivanova, E. P. Efremox, V. K. Antonov, and M. M. Shemyakin, *Zhur. obshchei Khim.*, 1970, **40**, 475.
[16] A. Foldes and C. Sandorfy, *Canad. J. Chem.*, 1970, **48**, 2197.
[17] M. Marraud, J. Neel, and M. Avignon, *J. Chim. Phys. Physiochem. Biol.*, 1970, **67**, 959.
[18] M. Avignon and P. W. Huong, *Biopolymers*, 1970, **9**, 427.
[19] D. H. Green and M. R. J. Sutton, *Biochim. Biophys. Acta*, 1970, **211**, 139.
[20] V. Buontempo, G. Careri, and P. Fasella, *Phys. Letters*, 1970, **A31**, 543.

β-keratin using polarised i.r. radiation has allowed all three active components of the amide I vibrations associated with the antiparallel-chain pleated sheet to be detected. Previously these had only been obtained from synthetic polypeptides having this conformation.[21] By measuring the ratio of the absorbance of the amide I bands present at 1655 cm^{-1} and 1633 cm^{-1} in solutions of human serum albumin, when the pH and temperature were changed, an indication of the amount of α-helical and random conformation relative to the amount of β-structure present was obtained. The amount of β-structure present increased with heating.[22]

3 Optical Rotatory Dispersion and Circular Dichroism

contributed by D. G. Dalgleish

A. General.—Calculations of the n–π^* and π–π^* rotational strengths of amide groups in the α-helical, β-pleated sheet and polyproline I and II conformations have been successful in predicting the correct c.d. spectra for these forms, apart from the case of polyproline II. The introduction of a further n–π^* transition at 165 nm for the α- and β-conformations also gave good results.[23] Random or 'amorphous' polypeptides were also treated, but only the π–π^* transition of the amide groups was considered in a randomly generated polypeptide, and a partition function was constructed to determine averaged rotational strengths. Such a calculation successfully predicted the two positive and negative bands in the c.d. spectra of unordered polypeptides.[24]

The c.d. spectra of several proteins of known conformations have been compared, and it appears that quantitation of the amounts of α-helix, β-sheet, and unordered conformation is not possible from a simple three-parameter fitting procedure.[25] An analysis of the visible and near-u.v. o.r.d. spectra of proteins to determine these parameters has been proposed.[26] This is basically a three-term equation analogous to the Moffit–Yang plot, and is not likely to offer much advantage over methods used at present.

B. Amino-acids and Synthetic Polypeptides. (See also Chapter 1, Section 3C.) The optical activity of the side-chain chromophores of the aromatic amino-acids and cystine continues to be an important field of study. The disulphide group of cystine is known to show optical activity, but investigations have shown a variability in the positions of the bands and their signs. The positions of the bands in the spectrum are known to be related to the dihedral angle of the disulphide, and recent calculations have shown that, at angles of about 90°, the intrinsic optical activity of the group in its lowest energy absorption band is small since two transitions of

[21] R. D. B. Fraser and E. Suzuki, *Spectrochim. Acta*, 1970, **A26**, 423.
[22] V. Palm, *Z. Chem.*, 1970, **10**, 31.
[23] E. S. Pysh, *J. Chem. Phys.*, 1970, **52**, 4723.
[24] D. Aebersold and E. S. Pysh, *J. Chem. Phys.*, 1970, **53**, 2156.
[25] B. Jirgensons, *Biochim. Biophys. Acta*, 1970, **200**, 9.
[26] S. Sugai, K. Nitta, and M. Ishikawa, *Biophysik.*, 1970, **7**, 8.

opposite rotational strength almost cancel. In such a situation, environmental effects may be operative in determining the sign of the c.d. band, and for dihedral angles greater than 90°, a left-handed chirality will give rise to a positive c.d. band, rather than negative.[27] Evidence for disulphide optical activity between 260 and 320 nm has been provided by Horowitz and co-workers [28] in a study of the near-u.v. bands of ribonuclease-A.

This latter study, however, was primarily concerned with the tyrosyl residues of ribonuclease. From a study of tyrosyl model compounds at 298 and 77 K, it was shown that the absorption and c.d. bands arise from a O–O transition at 282—289 nm, with a series of vibronic bands with 800 and 1250 cm^{-1} spacing. Comparison of the c.d. spectra of the model compounds with that of ribonuclease allowed the conclusion that the protein contained tyrosyl residues of three types, having their O–O bands at different wavelengths: one class contained the three solvent-accessible groups, and of the remaining three residues, two different types could be distinguished. A study of tyrosyl dioxopiperazines has confirmed that the c.d. spectrum of tyrosine is similar to its absorption spectrum.[29] Calculations of the optical activity of *ortho*-, *meta*-, and *para*-tyrosines have been made and, from these, certain conformations can be shown to be favoured by each of the three separate amino-acids.[30] An interesting feature of this work was the use of conformational-energy rotational-strength maps, allowing the direct comparison of rotational strengths in allowed regions of conformational space.

The ^1L_a and ^1L_b bands of tryptophan have also been studied in tryptophan-containing dioxopiperazines,[29] and an attempt has been made to separate the contributions from tyrosine and tryptophan in the near-u.v. c.d. spectrum of lysozyme.[31] By iodine oxidation of lysozyme, tryptophan-108 is destroyed, and changes resulting from this in the c.d. spectrum may be measured. In addition, tyrosine contributions may be found by comparing the c.d. spectra of native and oxidised lysozyme at pH 11.2.

The conformation of proline derivatives and their optical activity have been extensively discussed.[32, 33] The c.d. spectrum of 3-methyl-pyrrolidine-2-one has been shown to possess a positive n–π^* band at 210—220 nm, which does not show the sign expected from the quadrant rule of Litman and Schellman.[34] A second, negative, band at 189—196 nm may arise from the π–π^* transition, and aggregation produces a band at about 202 nm.[35] *N*-Acetyl-L-proline-*NN*-dimethylamide in water has a c.d.

[27] J. Linderberg and J. Michl, *J. Amer. Chem. Soc.*, 1970, **92**, 2619.
[28] J. Horowitz, E. H. Strickland, and C. Billups, *J. Amer. Chem. Soc.*, 1970, **92**, 2119.
[29] E. H. Strickland, M. Wilchek, J. Horowitz, and C. Billups, *J. Biol. Chem.*, 1970, **245**, 4168.
[30] T. M. Hooker and J. A. Schellman, *Biopolymers*, 1970, **9**, 1319.
[31] V. I. Teichberg, C. M. Kay, and N. Sharon, *European J. Biochem.*, 1970, **16**, 55.
[32] V. Madison and J. A. Schellman, *Biopolymers*, 1970, **9**, 511.
[33] V. Madison and J. A. Schellman, *Biopolymers*, 1970, **9**, 569.
[34] B. J. Litman and J. A. Schellman, *J. Phys. Chem.*, 1965, **69**, 678.
[35] N. J. Greenfield and G. D. Fasman, *J. Amer. Chem. Soc.*, 1970, **92**, 177.

spectrum similar to those of polyproline II and fibrous collagen, but in cyclohexane the c.d. spectrum resembles that of polyproline I. Conformational analysis shows that the molecule is confined to two small regions of conformational space: the *trans* isomer similar to polyproline II and the *cis* isomer to polyproline I. N.m.r. spectra show that a conformational equilibrium exists, biased to the *trans* form in aqueous solution and to the *cis* form in non-polar solvents. Analysis of the o.r.d. spectra shows this effect, and the optical activity of the pure isomers can be derived.[36] A conformational transition induced in poly-L-proline by high molarities of calcium chloride has been studied.[37] A similar transition occurs with poly-(4-hydroxy-L-proline), but the latter polymer is more stable than polyproline II to the effects of the high salt concentration.[38] A similar polymer to polyproline is poly-[(*S*)-thiazolidine-4-carboxylic] acid. This polymer was estimated to be helical by calculation[39] and showed a c.d. spectrum similar to that of polyproline II.[40] No mutarotation of the polymer appears to occur.

Various conformations of poly-L-lysine have been studied. It is of interest to note the differences between the c.d. spectra of solutions and cast films of the polymer.[41] The α-helical polylysine gave varying c.d. spectra dependent upon the relative humidity of the film, and these may represent distortions of the helices. However, more importantly, the films of unordered polylysine showed c.d. spectra completely different from those of the unordered conformation in solution, in which the positive extremum at 220 nm appearing in the solution studies is replaced by a negative extremum in the same position. From the point of view of the analysis of unordered conformations of proteins in solution, this may be a highly relevant finding, since it represents the c.d. spectrum of a constrained unordered structure. It has been observed that the c.d. spectrum of fully protonated polylysine is altered by various salts.[42] The altered spectra can be quantitatively accounted for by assuming that a small amount of coil → helix transition occurs. This work also confirms that the c.d. spectra of unordered proteins are different from those of unordered polylysine. The unordered protein c.d. spectra cannot be accounted for by a mixture of random and helical polylysine. In concentrated sulphuric acid, the unordered polylysine c.d. spectrum is of smaller magnitude than it is in dilute acid, and the π–π^* transition is slightly red-shifted.[43] This alteration may be caused by protonation of the peptide backbone, and may explain the altered c.d. spectra of unordered polypeptides in salt

[36] V. Madison and J. A. Schellman, *Biopolymers*, 1970, **9**, 65.
[37] W. L. Mattice and L. Mandelkern, *Biochemistry*, 1970, **9**, 1049.
[38] W. L. Mattice and L. Mandelkern, *Macromolecules*, 1970, **3**, 199.
[39] M. Goodman, G. C.-C. Niu, and K.-C. Su, *J. Amer. Chem. Soc.*, 1970, **92**, 5219.
[40] M. Goodman, K.-C. Su, and G. C.-C. Niu, *J. Amer. Chem. Soc.*, 1970, **92**, 5220.
[41] G. D. Fasman, H. Hoving, and S. N. Timasheff, *Biochemistry*, 1970, **9**, 3316.
[42] D. G. Dearborn and D. B. Wetlaufer, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 314.
[43] E. Peggion, A. Cosani, M. Terbojevich, and A. S. Verdini, *Macromolecules*, 1970, **3**, 318.

solutions.[43, 44] Polylysine also forms inter- and intra-molecular β-pleated sheet, the former being favoured by a low degree of polymerisation or by high concentration, while the latter is formed with long strands of the polymer at low concentration.[45] Polylysine can be induced to form α-helices by increasing the pH or the ionic strength, but polyarginine cannot. The polyarginine does form helices in 60% trifluoroethanol: histones form helices when 2-chloroethanol is added to aqueous solution.[46] The stabilities of the α-helices of poly-L-lysine, poly-L-ornithine, and poly-(L-αγ-diaminobutyric acid) differ. The last of these polymers is almost unordered in aqueous solution even at pH values above 10, where polylysine is completely helical and polyornithine partially so.[47] However, the polymers of the ε-, δ-, and γ-*N*-benzyloxycarbonyl- derivatives show a reversed order of stability in chloroform–dichloroacetic acid solutions, and this is explained by hydrogen-bonded interactions between their side-chains. The helix of poly-L-ornithine may be stabilised by aliphatic alcohols, the efficiency of the stabilisation increasing with the chain length of the alcohol,[48] and stabilisation of the α-helix of polylysine by 95% methanol has been used to determine the electrostatic free energy of the polymer by comparing its titration curve with that of n-butylamine.[49]

The conformation of poly-L-tyrosine in solution at low pH appears to be an intramolecular β-sheet structure, in which aggregation occurs.[50] This is at variance with previous c.d. results, which had estimated that helices were present, and the discrepancy is probably attributable to the appearance of two bands in the region 220—230 nm, representing the n–π^* transitions of the amide and side-chain chromophores. In trimethyl phosphate, pyridine, and dimethylformamide, the polymer appears to be helical, and the spectrum shows evidence for interactions between side-chains.[51] Indirect evidence suggests that poly-L-phenylalanine may also be α-helical, since it has been shown that poly-L-β-cyclohexylalanine adopts this conformation in solvents containing less than 86% methanesulphonic acid.[52] The ordered forms of the two polymers show different stabilities, however, since side-chain interactions will obviously differ between the two. Random copolymers of lysine and phenylalanine undergo an unordered → β transition, as shown by c.d. spectra; the conformational stability of the β-structure appears to increase with the content of phenylalanine in the polymer.[53]

[44] M. L. Tiffany and S. Krimm, *Biopolymers*, 1969, **8**, 347.
[45] S.-Y. C. Wooley and G. Holzwarth, *Biochemistry*, 1970, **9**, 3604.
[46] M. Boublik, E. M. Bradbury, C. Crane-Robinson, and H. W. E. Rattle, *European J. Biochem.*, 1970, **12**, 258.
[47] M. Hatano and M. Yoneyama, *J. Amer. Chem. Soc.*, 1970, **92**, 1392.
[48] G. Conio, E. Patrone, and S. Brighetti, *J. Biol. Chem.*, 1970, **245**, 3335.
[49] R. K. H. Liem, D. Poland, and H. A. Scheraga, *J. Amer. Chem. Soc.*, 1970, **92**, 5717.
[50] E. Patrone, G. Conio, and S. Brighetti, *Biopolymers*, 1970, **9**, 897.
[51] V. N. Damle, *Biopolymers*, 1970, **9**, 937.
[52] E. Peggion, L. Strasorier, and A. Cosani, *J. Amer. Chem. Soc.*, 1970, **92**, 381.
[53] E. Peggion, A. S. Verdini, A. Cosani, and E. Scoffone, *Macromolecules*, 1970, **3**, 194.

The effect of non-polar side-chains is to increase the stability of the α-helix as the size of the side-chain increases. This has been shown by considering the α-helix formation in block copolymers of L-leucine with L-glutamic acid or L-lysine,[54] and comparing the results with those observed for poly-L-alanine. Some calculations on the preferred conformations of *o*-, *m*-, and *p*-chlorobenzyl esters of poly-L-aspartic acid have shown that the preferred form of the first two is a right-handed α-helix, and of the third a left-handed α-helix.[55] This has been proved from measurements of their c.d. spectra.[56] The c.d. spectrum of the left-handed α-helix is not the mirror image of the right-handed helix, since both are composed of L-amino-acid residues.

The c.d. spectrum of poly-(1-benzyl-L-histidine) in trifluoroethanol is similar to that of α-helical polypeptides in shape, but has a much lower magnitude.[57] Random copolymers of the amino-acid with N^{ε}-benzyloxycarbonyl-L-lysine show a linear increase of the magnitude of the far-u.v. c.d. spectrum with increasing lysine content, and it is therefore supposed that the original poly-(1-benzyl-L-histidine) is indeed helical in trifluoroethanol. In aqueous solution, the polymer undergoes an unordered → α-helix transition dependent on the protonation of the histidine residues.

C. Small Peptides.—The peptide antibiotic, stendomycin, contains a number of D-amino-acid residues, and these form a conformation with a c.d. spectrum characteristic of an α-helix in trifluoroethanol and acetonitrile.[58] In water, the c.d. spectrum is very temperature-dependent, showing a random coil type at 40 °C, but a helical type at 5 °C. I.r. spectra also confirm the presence of the α-helix, in which the hydrophobic residues are directed towards the solution.

The problem of protein folding has been studied by investigating the conformations of the peptides formed by cleaving myoglobin at either arginine or proline residues.[59, 60] Cleavage at arginine yields the peptides 1—31, 46—118, and 119—139, all of which show much smaller amounts of helical character in solution than would be expected from their conformations in the intact protein: in 62% methanol, the helix contents increase, but still appear to be less than their expected values.[59] Similarly, peptides 1—36, 37—87, 1—87, 37—119, and 120—153 show less helical character than is to be expected. Therefore, even the large sections of the protein do not fold correctly in isolation from the rest of the molecule. The implication of this work is clearly that proteins may not fold into their

[54] S. E. Ostroy, N. Lotan, R. T. Ingwall, and H. A. Scheraga, *Biopolymers*, 1970, **9**, 749.
[55] J. F. Yan, F. A. Momany, and H. A. Scheraga, *J. Amer. Chem. Soc.*, 1970, **92**, 1109.
[56] E. H. Erenrich, R. H. Andreatta, and H. A. Scheraga, *J. Amer. Chem. Soc.*, 1970, **92**, 1116.
[57] M. Terbojevich, M. Acampora, A. Cosani, E. Peggion, and E. Scoffone, *Macromolecules*, 1970, **3**, 618.
[58] D. W. Urry and A. Ruiter, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 800.
[59] M. Z. Atassi and R. P. Singhal, *J. Biol. Chem.*, 1970, **245**, 5122.
[60] R. P. Singhal and M. Z. Atassi, *Biochemistry*, 1970, **9**, 4252.

final conformations until synthesis is almost or wholly complete (see also ref. 96, below).

A 36-peptide unit from the cyanogen bromide cleavage of the α1 chain from collagen forms a trimer with a collagen-like c.d. spectrum at 2 °C. This structure is about 90% helical.[61] The trimer is the result of a monomer association, and studies of the equilibrium between monomer and trimer allow explanation of the variation of the c.d. results.[62] It has been suggested that a rate-limiting interaction of three chains is followed by a rapid propagation of helix, whose stability depends on the initial alignment of the chains. An analogue of collagen, but without proline, is poly-(L-Ala-L-Ala-Gly), which shows no evidence to indicate the formation of a collagen-like structure, although β-sheet is found in hexafluoroisopropanol–ethylene glycol mixtures, and α-helix may be formed in hexafluoroisopropanol alone.[63] This clearly demonstrates the importance of proline in forming the triple helix of collagen.

Thyrocalcitonin, a hormone with 32 amino-acid residues, shows a small amount of α-helical character in solution, and is completely randomised in 6M guanidine hydrochloride. In 2-chloroethanol, about 50% of the residues adopt the helical conformation.[64] The terminal heptapeptide ring disulphide may be reduced and alkylated without altering the o.r.d. or c.d. properties, and it is suggested that the potential for forming α-helix may be realised at receptor sites in lipid layers, where the water concentration is significantly reduced.

D. Haem Proteins.—The c.d. spectra of human haemoglobin and its subunits have been measured in the oxy-, deoxy-, and CO-forms.[65] These investigations showed that the c.d. spectra of the two types of haem group, and therefore their environments, differ with the spectrum of the intact haemoglobin being a mean of the two spectra. Studies on CO-haemoglobin have shown that, when the protein is acid-denatured, there is evidence that some unfolding occurs without haem detachment,[66] since the Soret Cotton effect is increased at pH 3—3.3. The main unfolding may be followed by the decrease in the 233 nm trough in the o.r.d. spectrum; as the pH is decreased from 3.0, the Soret Cotton effect decreases. Horse haemoglobin(III) forms two pH-dependent forms, related to the existence of two conformational isomers of the protein; these isomers are of the same helical structure, and differ in tertiary, rather than secondary structure.[67] Alkyl isocyanide complexes of human, bovine, and carp haemoglobin show different Soret Cotton effects, but the variations of each with different

[61] K. A. Piez and M. R. Sherman, *Biochemistry*, 1970, **9**, 4129.
[62] K. A. Piez and M. R. Sherman, *Biochemistry*, 1970, **9**, 4134.
[63] B. B. Doyle, W. Traub, G. P. Lorenzi, F. R. Brown tert., and E. R. Blout, *J. Mol. Biol.*, 1970, **51**, 47.
[64] H. B. Brewer jun., and H. Edelhoch, *J. Biol. Chem.*, 1970, **245**, 2402.
[65] Y. Ueda, T. Shiga, and I. Tyuma, *Biochim. Biophys. Acta*, 1970, **207**, 18.
[66] J. W. Allis and J. Steinhardt, *Biochemistry*, 1970, **9**, 2286.
[67] K. Gersonde and A. Wollmer, *European J. Biochem.*, 1970, **15**, 226.

ligands are similar.[68] The 260 nm bands in the c.d. spectra also show species- and ligand-dependent differences. Haemoglobin from *Chironomus thummi* shows conformational changes when oxygen binds, although the helix content is not altered.[69] This haemoglobin is a mixture of components, which alter differently on oxidation.

The effects of different metalloporphyrins (with Cu^{2+} or Zn^{2+}) on the conformation of myoglobin vary, depending on the metals used.[70] Whereas the conformations of native and Cu-haem myoglobin are similar, Zn-haem myoglobin shows considerable unfolding, and denatures at a higher pH than the other two in acid solution. A further derivative, in which the haem group is nitrated at the vinyl side-chain, showed some unfolding, and was intermediate between the native and zinc proteins in its denaturation behaviour.

C.d. difference spectra of cytochrome a_3, obtained in the form $a_3{}^2CO$-$a_3{}^3CN$ were not identical, and the difference was found to be related to the oxidation state of the copper atom of the molecule. The results suggested that, in cytochrome oxidase, the Cu which interacts with the a_3 is probably sandwiched between the a and a_3 moieties.[71] Bands in the c.d. spectrum of cytochrome oxidase appearing between 500 and 600 nm vary with the oxidation state and arise from the haem-iron and copper chromophores being perturbed by the protein.[72] The spectroscopic properties of the protein appear to be dependent on aggregation, as has been shown by a study of the CO-complexes.[73]

Several studies of various cytochromes have recently appeared. Three cytochromes c_3 from species of *Desulfovibiro* show similar rotatory properties, suggesting a similar haem environment despite differences in amino-acid composition.[74] The Soret bands show evidence of two components in the c.d. spectra of both oxidised and reduced forms, and the acid denaturation of one form showed a sharp transition at pH 2.6. In addition, the c.d. spectra of *c'* cytochrome from *Rhodopseudomonas palustris* and the *cc'* cytochromes from *Rhodospirillum rubrum*, *Pseudomonas dinitrificans*, and *Chromatium* have been reported.[75] Two *c*-type cytochromes from *Rhodospirillum molischianum* have also been studied.[76]

The o.r.d. spectrum of *Chromatium* cytochrome *C*-552 which contains two covalently-linked haem groups and one flavin, shows Cotton effects in the Soret region.[77] These effects show four contributions in the oxidised form and three and four contributions in the reduced and CO-forms

[68] K. Ruckpaul, H. Rein, O. Ristau, and F. Jung, *Biochim. Biophys. Acta*, 1970, **221**, 9.
[69] A. Wollmer, H. Sick, and K. Gersonde, *European J. Biochem.*, 1970, **16**, 303.
[70] S. F. Andres and M. Z. Atassi, *Biochemistry*, 1970, **9**, 2268.
[71] F. C. Yong and T. E. King, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 940.
[72] F. C. Yong and T. E. King, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 1445.
[73] B. Love and H. E. Auer, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 1437.
[74] H. Drucker, L. L. Campbell, and R. W. Woody, *Biochemistry*, 1970, **9**, 1519.
[75] F. C. Yong and T. E. King, *J. Biol. Chem.*, 1970, **245**, 2457.
[76] T. Flatmark, K. Dus, H. de Klerk, and M. D. Kamen, *Biochemistry*, 1970, **9**, 1991.
[77] F. C. Yong and T. E. King, *J. Biol. Chem.*, 1970, **245**, 1331.

respectively. These may arise from coupled haem interactions, or the immediate environment of the prosthetic groups could be affected by the oxidation state of the haem, or sharing of the CO between the haems may not be equal. A redox-dependent conformational change of rabbit-liver cytochrome P-450 has been demonstrated by comparing the o.r.d. spectra of the oxidised and reduced forms.[78] The former state has an extrinsic Cotton effect from the haem, which is lacking in the reduced form, and also in the inactivated form P-420. The cytochromes *b*-555 and *b*-563 from housefly larvae show optically active haems in both oxidised and reduced states, although changes in the c.d. spectra occur on oxidation–reduction.[79, 80]

Phenylalanyl fine structure has been studied in the near-u.v. c.d. spectra of the isoenzymes of horseradish peroxidase. Isoenzymes A1 and C with their apoenzymes were examined in guanidine hydrochloride solution, and the phenylalanyl bands in peroxidase-C were found to be comparable in magnitude with those of an equivalent concentration of phenylalanine in solution. This suggests that many different environments for the side-chain exist in the protein.[81]

E. Binding of Small Molecules to Proteins.—Bilirubin interacts with both human and bovine serum albumins. When binding occurs, strong Cotton effects are generated between 400 and 500 nm. These large effects are caused by the coupling of the two dipyrrylmethylene groups, which may be altered by conformational changes when the molecule binds, or by pH alterations, or when conformational changes are induced in the proteins.[82] The amplitude of the induced Cotton effect is maximal at pH 5, and is dependent on the buffer used. Complexes of up to 4 mol of bilirubin per mole of serum albumin may be made, and the amount of bilirubin bound is reduced by oleate at between five and eight times the molarity of the complex; salicylate is a poor competitor.[83] The complexes with human (HSA) and bovine (BSA) albumins show different c.d. spectra and the spectra themselves vary with the amount of ligand bound. While both c.d. spectra show two components between 400 and 500 nm, the HSA-bilirubin complexes show a doublet with the longer-wavelength contribution positive and the other negative,[83] although this may be altered by changing the pH.[84] The BSA complex shows a variation with the amount of ligand bound, in which both contributions vary in magnitude and sign. The differences between the two albumins probably reflect differences in the conformations of the two proteins.

[78] F. C. Yong, T. E. King, S. Oldham, M. R. Waterman, and H. S. Mason, *Arch. Biochem. Biophys.*, 1970, **138**, 96.
[79] Y. Okada and K. Okonuki, *J. Biochem.* (*Japan*), 1970, **67**, 603.
[80] Y. Okada and K. Okonuki, *J. Biochem.* (*Japan*), 1970, **67**, 487.
[81] E. H. Strickland, E. Kay, and L. M. Shannon, *J. Biol. Chem.*, 1970, **245**, 1233.
[82] G. Blauer and T. E. King, *J. Biol. Chem.*, 1970, **245**, 372.
[83] P. V. Wooley tert., and M. J. Hunter, *Arch. Biochem. Biophys.*, 1970, **140**, 197.
[84] G. Blauer, D. Harmatz, and A. Naparstek, *F.E.B.S. Letters*, 1970, **9**, 53.

Bovine neurophysins show a strong negative band at 280 nm in the c.d. spectrum, and a positive extremum at 240 nm, which are attributable to disulphides, since no shift of these contributions occurs when the tyrosyl residues are titrated.[85] Binding of oxytocin, vasopressin, or the peptide L-Cys-L-Tyr-L-Phe-NH_2 produces large alterations in the c.d. spectrum, which are primarily caused by the disulphides of the neurophysin, although changes in the backbone may also occur. Ca^{2+} has little effect on the affinity of the protein for oxytocin.

A small local conformational change appears to occur when threonine-inhibited aspartokinase-homoserine dehydrogenase is complexed with aspartic acid or threonine, when the binding is non-co-operative.[86] On the other hand, when the binding is co-operative, binding of aspartic acid, threonine, or K^+ gives a larger perturbation, accompanied by changes in the far-u.v. c.d. spectrum. An extrinsic Cotton effect is generated by NADH binding, the magnitude being dependent on alterations brought about by allosteric ligands; titrations may be performed by using the c.d. spectra to detect complex formation.

C.d. and absorption spectra have been used to show that 1,10-phenanthroline and 2,2′-bipyridyl bind to only two of the zinc atoms of liver alcohol dehydrogenase, showing that not all of the zinc atoms in the protein are catalytically active.[87] Further ligand-induced conformational changes reported are small backbone changes induced in Na^+, K^+-ATPase induced by ouabain,[88] and complex alterations in the side-chain c.d. spectrum of trypsin complexed with native and modified soybean inhibitor.[89]

A complex of bacteriochlorophyll with protein shows strong interactions between adjacent pigment molecules only at neutral pH,[90] and it is clear that only at this pH is the complex stable.

Thiamine pyrophosphate (TPP) and Ca^{2+} are co-factors for yeast transketolase. Addition of Ca^{2+} does not affect the c.d. spectrum of the enzyme, but bands at 280 and 320 nm are introduced when TPP is added.[91] The latter of these is ascribed to a bond formed between TPP and the protein.

F. Modified Proteins.—Nitration of one tyrosyl residue of aspartate amino-transferase, in the presence of the substrates glutamate and α-keto-glutarate, yields an inactive enzyme, but in the absence of the substrates no inactivation occurs.[92] The inactivated enzyme is in the pyridoxamine-5′-phosphate form, as shown by c.d., indicating that the nitration only

[85] E. Breslow, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 493.
[86] H. d'A. Heck and P. Truffa-Bachi, *Biochemistry*, 1970, **9**, 2776.
[87] D. E. Drum and B. L. Vallee, *Biochemistry*, 1970, **9**, 4078.
[88] G. E. Lindenmayer and A. Schwartz, *Arch. Biochem. Biophys.*, 1970, **140**, 371.
[89] M. Ishida, K. Hamaguchi, and T. Ikenaka, *J. Biochem. (Japan)*, 1970, **67**, 363.
[90] Y. D. Kim and B. Ke, *Arch. Biochem. Biophys.*, 1970, **140**, 341, 354.
[91] G. A. Kochetov, R. A. Usmanov, and V. P. Merzlov, *F.E.B.S. Letters*, 1970, **9**, 265.
[92] P. Christen and J. F. Riordan, *Biochemistry*, 1970, **9**, 3025.

occurs after the transition from the aldimine to the ketimine state of the enzyme–substrate complex. A functional tyrosyl residue in arginine kinase is also susceptible to nitration, giving an apparent conformational change and reduced substrate binding.[93] Tyrosyl residues appear to be involved in the liganding of iron in hemerythrin, since nitration of five residues per subunit releases one of the two bound iron atoms.[94] The visible c.d. spectrum is destroyed by this, in parallel with the loss of iron, and a helix → coil transition occurs as nitration proceeds.

Cyanuric fluoride has been reacted with the tyrosyl residues of apomyoglobin, bovine serum albumin, and carbonic anhydrase B.[95] No conformational change was detected in apomyoglobin as a result of this, but small effects were found in bovine serum albumin.

A study of the refolding of reduced–reoxidised des(121—124)-ribonuclease-A has shown that the last three residues are necessary for correct folding of the protein, since the c.d. spectrum of the modified protein is featureless[96] (*cf.* also refs. 59 and 60). This is to be compared with the correct folding of the native protein after reduction and reoxidation.[97] Reduction of the three disulphides of pepsin and pepsinogen destroys the activity of the protein and the potential activity of the zymogen, which may be restored by reoxidising two of the disulphides.[98] The reoxidised zymogen has a similar c.d. spectrum to the native form, but reoxidised pepsin differs from its native structure, suggesting that refolding is impaired. Modification of the disulphides of histidine ammonia lyase gives little alteration to its far-u.v. c.d. spectrum, suggesting that no important conformational changes occur.[99]

Mono- and di-sulphoxy derivatives of lysozyme may be made by irradiation of the protein in the presence of porphyrin. The mono-derivative (modified at Met-12) is partially active, and has small changes in the near-u.v. c.d. spectrum and a deepening of the 210 nm trough.[100] The disulphoxy derivative (modified at Met-12 and Met-105), is inactive, lacks any contributions to the near-u.v. c.d. spectrum, and shows a greatly reduced far-u.v. effect.

Staphylococcal nuclease oxidised by performic acid is denatured, but still retains about 8% activity. Acetylated and trifluoroacetylated derivatives of the enzyme, with up to 23 residues modified, are less denatured.[101] It has been suggested that the molecule is highly flexible, since it has only a low helical content and lacks disulphides. Modification of the single

[93] A. Kassab, A. Fattoum, and L. A. Pradel, *European J. Biochem.*, 1970, **12**, 264.
[94] R. I. Rill and I. M. Klotz, *Arch. Biochem. Biophys.*, 1970, **136**, 507.
[95] M. J. Gorbunoff, *Arch. Biochem. Biophys.*, 1970, **138**, 684.
[96] H. Taniuchi, *J. Biol. Chem.*, 1970, **245**, 5459.
[97] M. N. Pflumm and S. Beychok, *J. Biol. Chem.*, 1969, **244**, 3982.
[98] Y. Nakagawa and G. E. Perlmann, *Arch. Biochem. Biophys.*, 1970, **140**, 464.
[99] C. B. Klee, *J. Biol. Chem.*, 1970, **245**, 3413.
[100] G. Galiazzo, A. M. Tamburro, and G. Jori, *European J. Biochem.*, 1970, **12**, 362.
[101] G. S. Omenn, D. A. Ontjes, and C. B. Anfinsen, *Biochemistry*, 1970, **9**, 304.

tryptophan residue of the protein gives a disordered protein which is still, however, 15% active.[102]

Proteins modified with diazotised *p*-arsanilic acid give extrinsic Cotton effects between 300 and 450 nm.[103] Modification of the same protein with different diazonium salts yields c.d. spectra which differ only in magnitude, but modifications of different proteins with the same diazonium salt give different c.d. spectra. The extrinsic Cotton effects are structure-dependent and may be used to detect conformational changes.[104]

G. Denaturation and Unfolding.—A study of the denaturation of myoglobin and α-chymotrypsinogen by alcohols shows that the denaturing capability of the alcohol increases with increasing chain length, as would be expected for the disorganisation of hydrophobic regions of proteins.[105] Chain branching in the alcohol reduces its effectiveness, and glycols are less effective than alcohols. A similar study has shown that chain length is important in stabilising the α-helix of poly-L-ornithine.[48] Alkyl-substituted ureas also show increased denaturing capacity with increasing chain length and alkyl substitution, and a series of calculations on model compounds and their free energies of transfer show essential agreement with the experimental results.[106] This permits the deduction that the higher ureas (ethyl, propyl, and n-butyl) have a predominantly hydrophobic interaction with protein side-chains.

Denaturation of bovine serum albumin by aqueous lithium salts gives an incompletely unfolded state, with a more ordered structure than that produced by 6M guanidine hydrochloride.[107] That this ordering is not only caused by disulphide bridges is shown by the observation that *S*-carboxymethyl bovine serum albumin has different conformations in guanidine and Li^+ solutions. A further instance of 'incomplete' denaturation is that given by dodecyl sulphate interacting with proteins, where the o.r.d. spectra of the complexes are not those of typical unordered proteins, but show some evidence of helical character, which varies depending on the proteins used.[108]

In the absence of a supporting electrolyte, a progressively greater interaction of urea with lysozyme (3—9M urea) is not accompanied by a change of o.r.d., but addition of electrolyte causes unfolding in the high molarity of urea.[109] However, the rate of disulphide exchange with 2-mercaptoethanol is dependent on the urea concentration, so that some degree of protein perturbation must occur. Some small changes, which may be of this type, are known to occur in pyridine nucleotide transhydrogenase in 1M urea.[110]

[102] G. S. Omenn, A. Fontana, and C. B. Anfinsen, *J. Biol. Chem.*, 1970, **245**, 1895.
[103] G. F. Fairclough jun., and B. L. Vallee, *Biochemistry*, 1970, **9**, 4087.
[104] H. M. Kagan and B. L. Vallee, *Biochemistry*, 1969, **8**, 4223.
[105] T. T. Herskovits, B. Gadegbeku, and H. Jaillet, *J. Biol. Chem.*, 1970, **245**, 2588.
[106] T. T. Herskovits, H. Jaillet, and B. Gadegbeku, *J. Biol. Chem.*, 1970, **245**, 4544.
[107] M. E. Noelken, *Biochemistry*, 1970, **9**, 4117.
[108] J. A. Reynolds and C. Tanford, *J. Biol. Chem.*, 1970, **245**, 5161.
[109] J. R. Warren and J. A. Gordon, *J. Biol. Chem.*, 1970, **245**, 4097.
[110] D. D. Louie and N. O. Kaplan, *J. Biol. Chem.*, 1970, **245**, 5691.

These changes do not appear to be related to any loss of regular structure within the protein.

A two-step denaturation occurs when muscle aldolase interacts with increasing concentrations of guanidine hydrochloride.[111] Both the native and mixed disulphide proteins showed this process, although the modified derivatives were seen to be destabilised. A further two-step process has been found in the guanidine denaturation of adrenodoxin.[112] This reaction gives an intermediate in which a Cotton effect observed at 470 nm is seen, compared with 440 nm in the native protein, and further unfolding destroys both Cotton effects. Spinach ferredoxin shows similar behaviour.

A partly reversible denaturation of L-asparaginase has been reported.[113] The protein dissociates in 7M urea or 5M guanidine hydrochloride, with reduction of the disulphides, to give inactive monomers. Of the enzyme activity 85% may be regained by oxidation and reconstitution. Some α-helix is present, which is destroyed in 7M urea; the side-chain c.d. spectrum is also lost under these conditions. Titration of all of the tyrosyl residues of the protein also gives denaturation. A reversible conformational change is also induced by guanidine on a leucine-binding protein from *E. coli*.[114]

Myosin shows some 60% of α-helix, which is decreased in alkaline solution or in guanidine solution and drops to about 40% with ageing, while a similar final helix content may be found in aged tropomyosin, whose initial helical content is greater.[115] 2M Potassium chloride reduces the helix content of myosin only.

Thermophilic α-amylase from *Bacillus stearothermophilus* shows an o.r.d. spectrum suggestive of 30% α-helix, similar to the enzyme from *B. subtilis*, although the two differ in amino-acid composition and near-u.v. c.d. spectra.[116] The enzyme is unstable in the absence of Ca^{2+}, and stability is affected more by 0.01M EDTA than 8M urea, which hardly alters the c.d. spectrum at room temperature, although it causes denaturation in acid solution or when the temperature is raised.[117]

Acid denaturation of a Bence Jones protein gives incomplete disorganisation of the molecules, since the magnitude of the c.d. spectrum at 215 nm increases when the pH is decreased; alkali denaturation gives a more typical unordered type of c.d. spectrum.[118]

H. Lipoproteins and Membranes.—The problem of analysing the c.d. spectra of membranes is complicated by the unknown effects of absorption

[111] H. M. Steinman and F. M. Richards, *Biochemistry*, 1970, **9**, 4360.
[112] R. Padmanabhan and T. Kimura, *J. Biol. Chem.*, 1970, **245**, 2469.
[113] B. H. Frank, A. H. Pekar, A. J. Veros, and P. P. K. Ho, *J. Biol. Chem.*, 1970, **245**, 3716.
[114] W. R. Penrose, R. Zand, and D. L. Oxender, *J. Biol. Chem.*, 1970, **245**, 1432.
[115] P. Rainford and R. V. Rice, *Biopolymers*, 1970, **9**, 1.
[116] K. Ogasahara, A. Imanishi, and T. Isemura, *J. Biochem.* (*Japan*), 1970, **67**, 65.
[117] K. Ogasahara, A. Imanishi, and T. Isemura, *J. Biochem.* (*Japan*), 1970, **67**, 77.
[118] H. Takahashi, T. Hirai, T. Azuma, K. Hamaguchi, and S. Migita, *J. Biochem.* (*Japan*), 1970, **67**, 795.

flattening and light scattering caused by the large size of the particles involved. Thus, although it has been known for some time that membranes show anomalous c.d. spectra in the far-u.v., it has not been possible to interpret these spectra unambiguously. Two possibilities exist, since the anomalies may arise from genuine structural factors within the membranes, or from the purely physical distortions mentioned above.[119]

A consideration of the Rayleigh scattering phenomenon, taking into account the asymmetry of the chromophore and the polarisation of the incident light, has shown that distortion of the optical activity should be small for particles less than 1 μ in diameter.[120] This then suggests that the red-shifted c.d. spectrum of membranes arises from structural factors within the membranes. However, it has also been shown that distortions of the c.d. spectrum of haemoglobin may be caused by adding scattering material, and that membranes may be induced to give a more normal type of c.d. spectrum by sonicating the samples and reducing the particle size.[121] The use of poly-L-glutamic acid solutions and suspensions, using simultaneous measurement of absorption and c.d. spectra, has allowed the determination of the scattering and flattening components of membranes,[122] and corrections have been made on the c.d. spectra of several membranes.[123] This last process has shown differences between different membranes, and shows that they are not similar structures with differently damped and red-shifted spectra.

An investigation of the c.d. spectra of membranes of *Mycoplasma laidlawii*, together with results from i.r. spectroscopy, have shown that β-structure is present to a considerable extent.[124] Various curve-fitting procedures were used in this study, including the Moffit parameters, a_0 and b_0, and an analysis using the method of Greenfield and Fasman.[125] A final type of curve fit involved Gaussian fitting of the c.d. spectra of helical, unordered, and β-sheet polylysine and subsequent analysis of the membrane spectra. The final values given were 56% β, 31% α, and 13% unordered.

The high-density lipoproteins from serum have been shown by c.d. spectra to have high helical contents.[126] Removal of lipid produces only small changes in the c.d. spectra, but the change in the e.s.r. spectrum of a spin-labelled lipoprotein was marked, suggesting that this process causes only small changes in the protein structure. Serum high-density lipoprotein may be reconstituted to give a c.d. spectrum similar to that of the native structure only when polar and non-polar lipids (phospholipids and

[119] T. H. Ji and D. W. Urry, *Biochem. Biophys. Res. Comm.*, 1969, **34**, 404.
[120] C. A. Ottaway and D. B. Wetlaufer, *Arch. Biochem. Biophys.*, 1970, **139**, 257.
[121] A. S. Schneider, M.-J. T. Schneider, and K. Rosenheck, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 793.
[122] D. W. Urry, T. A. Hinners, and L. Masotti, *Arch. Biochem. Biophys.*, 1970, **137**, 215.
[123] D. W. Urry, L. Masotti, and J. Krivacic, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 521.
[124] G. L. Choules and R. F. Bjorklund, *Biochemistry*, 1970, **9**, 4759.
[125] N. J. Greenfield and G. D. Fasman, *Biochemistry*, 1969, **8**, 4108.
[126] A. M. Gotto and H. Kon, *Biochemistry*, 1970, **9**, 4276.

cholesterol esters) are present. When only phospholipids are used, both the c.d. spectrum and its temperature dependence differ from those of the native system.[127] The protein fraction of the lipoprotein contains several polypeptide chains, separable by gel filtration, and the c.d. spectra of two of these have been reported and compared.[128]

I. Protein–Nucleic Acid Interactions.—The histone proteins, when bound to DNA, appear to be helical to the extent of 35—45%,[129] and the helix content decreases when the proteins are dissociated from the nucleic acid.[130] Alternative estimates give the helix content in the intact nucleohistone as 27—44%.[131] The DNA c.d. spectra are altered in the complex,[129, 132] and the main cause of the alteration may be the slightly lysine-rich fraction of the histone, although others also have some effect.[133]

The interaction of nucleotides with basic polypeptides is governed by the nature of both polypeptide and nucleotide.[134] Polypeptide specificity depends on the nature of the side-chain, and the helical character of poly-L-arginine is stabilised more than that of poly-L-lysine since the guanidinium group of the arginine interacts with the two oxygen atoms of the phosphate group to give a stable ring. Binding of guanosine monophosphates to poly-L-lysine has also been studied.[135]

The histone proteins of nucleohistone provide a site for the attachment of hydrocortisone to calf thymus nuclei. As the histones are selectively hydrolysed by trypsin, hydrocortisone sensitivity is lost.[136] The sites for attachment of histones to the nucleic acid appear not to be in the small groove, since it has been shown that the binding and c.d. properties of a 'reporter' molecule[137] are not altered between DNA and chromatin.[138] Since the 'reporter' molecule is thought to occupy the small groove of the DNA when bound, it would appear that the histones do not occupy this site. The interaction of the basic polypeptide clupeine with DNA has also been studied,[139, 140] and it has been found that the o.r.d. spectra of the complexes are similar to those of DNA complexed with $(Arg)_{20}$.

Metaphase chromosomes undergo an irreversible conformational change on going from pH 5.6 to 7.0 at low ionic strength;[141] the change may be

[127] R. Hirz and A. M. Scanu, *Biochim. Biophys. Acta*, 1970, **207**, 367.
[128] A. M. Scanu, *Biochim. Biophys. Acta*, 1970, **200**, 570.
[129] R. T. Simpson and H. A. Sober, *Biochemistry*, 1970, **9**, 3103.
[130] T. Y. Shih and G. D. Fasman, *J. Mol. Biol.*, 1970, **52**, 125.
[131] V. I. Permogorov, V. G. Debadov, I. A. Sladkova, and B. A. Rebentish, *Biochim. Biophys. Acta*, 1970, **199**, 556.
[132] G. D. Fasman, B. Schaffhausen, L. Goldsmith, and A. Adler, *Biochemistry*, 1970, **9**, 2814.
[133] P. Henson and I. O. Walker, *European J. Biochem.*, 1970, **16**, 524.
[134] J. M. Rifkind and G. L. Eichhorn, *Biopolymers*, 1970, **9**, 1753.
[135] K. G. Wagner and K. Wulff, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 813.
[136] T. E. Wagner, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 890.
[137] E. J. Gabbay, *J. Amer. Chem. Soc.*, 1968, **90**, 6574.
[138] R. T. Simpson, *Biochemistry*, 1970, **9**, 4814.
[139] S. Inoue and T. Ando, *Biochemistry*, 1970, **9**, 395.
[140] S. Inoue and T. Ando, *Biochemistry*, 1970, **9**, 388.
[141] K. P. Cantor and J. E. Hearst, *J. Mol. Biol.*, 1970, **49**, 913.

prevented by the presence of 1mM-Mg^{2+}. Magnesium is also important in preserving the conformation of ribosomes, since chelation with EDTA causes irreversible changes in both the RNA and protein components of the c.d. spectra.[142] The c.d. spectrum of RNA alone is not dependent on prior treatment, however. Some 20 ribosomal proteins, most of them basic, showed that helix content in the ribosome is approximately that of the free proteins.[143] Most of the proteins were estimated to contain 20—40% helix, but two acidic proteins showed an estimated 50% helical content.

The substitution of leucine for proline-20 or proline-156 in TMV protein causes a different conformation of the virus, with an increased helical content of the protein. The first of these mutants is heat-labile, showing marked alterations in its near-u.v. c.d. spectrum.[144]

J. Proteins.—Freezing and thawing appear to induce an unordered → β conformational change in phosvitin, probably by aggregation. The ordered structure is heat-labile, and the effect may also be reversed by lowering the acidity if aggregation has not proceeded too far.[145] The protein is probably unordered at neutral pH, but may form α-helix and β-structure at pH 3—3.6.[146] That the regular structure is β has already been postulated.[147]

A similarity in conformation between hen-egg-white lysozyme and bovine α-lactalbumin has been suspected for some time, and the two proteins have been compared by c.d. and n.m.r. studies, which confirm that some similarity exists.[148] An analysis of the far-u.v. c.d. spectrum of α-lactalbumin has given 25% α, 15% β, and 60% unordered, figures which are not dissimilar to those found in lysozyme.[149] A further pair of similar proteins are ovine prolactin and bovine growth hormone, which exhibit similar helical contents, and whose c.d. spectra vary in a similar manner with variation in pH.[150] The bovine growth hormone shows a complex mixture of Cotton effects in the near-u.v., which are attributable to tyrosine, tryptophan, and phenylalanine.[151] That all three aromatic side-chains contribute is shown by the behaviour of the c.d. spectrum when the pH is varied or when the protein is subjected to treatment with urea.

In a study of several equine immunoglobulins, it has been shown that all of the proteins studied gave minima in the o.r.d. at 225 and 230 nm. The depth of the minimum varies between immunoglobulins of different classes,

[142] A. J. Adler, G. D. Fasman, and M. Tal, *Biochim. Biophys. Acta*, 1970, **213**, 424.
[143] M. Dzionara, *F.E.B.S. Letters*, 1970, **8**, 197.
[144] A. J. Budzynski, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 841.
[145] G. Taborsky, *J. Biol. Chem.*, 1970, **245**, 1054.
[146] K. Grizzuti and G. E. Perlmann, *J. Biol. Chem.*, 1970, **245**, 2573.
[147] G. Taborsky, *J. Biol. Chem.*, 1968, **243**, 6014.
[148] D. A. Cowburn, E. M. Bradbury, C. Crane-Robinson, and W. B. Gratzer, *European J. Biochem.*, 1970, **14**, 83.
[149] F. M. Robbins and L. G. Holmes, *Biochim. Biophys. Acta*, 1970, **221**, 234.
[150] S. M. Aloj and H. Edelhoch, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 830.
[151] H. Edelhoch and R. E. Lippoldt, *J. Biol. Chem.*, 1970, **245**, 4199.

and between individuals of the same class.[152] A small Cotton effect at 240 nm is characteristic of γG-globulins.[152] γG-Globulin shows two bands in the far-u.v. c.d. spectrum at 217 (negative) and 202 nm (positive), the former of which may be indicative of β-structure.[153] At pH 2.2, a negative contribution at 200 nm replaces the 202 nm maximum, but the 217 nm extremum is unaltered. Denaturation does not give complete disorganisation of the protein structure, and the c.d. spectrum appears to result from unspecified backbone conformations. The possibility that immunoglobulins contain β-structure has been supported by studies on Waldenström macroglobulins.[154] These studies showed that no α-helix was present but that there was some evidence for ordered structure. Subunits prepared from the parent molecule by reduction and alkylation of disulphides showed similar optical rotatory properties to the parent molecule. Myeloma immunoglobulin G proteins showed similar far-u.v. c.d. spectra to one another, which differed considerably from that of normal IgG, being of greater magnitude.[155] The difference suggested a more variable distribution of conformations in normal IgG.

Ribonuclease-A may be induced to aggregate by the presence of substrate, and the aggregate shows a much decreased magnitude of the far-u.v. c.d. spectrum. The aggregation is not reversed by dodecyl sulphate, and the decrease in the far-u.v. spectrum does not appear to be caused by loss of regular structure.[156] Rat and bovine ribonucleases differ in about 30% of their residues, but are similar in their overall conformations, although local differences cause different sensitivities to proteolytic attack.[157] From a comparison of the near-u.v. c.d. spectra it has been concluded that the surface tyrosyl residues of bovine ribonuclease are optically active, since the absence of two of them from the rat enzyme gives a much reduced c.d. spectrum in the near-u.v. Such an assessment is supported by the work of Horowitz and co-workers described above,[28] although the surface tyrosines of the bovine enzyme do not contribute as much to the rotational strength as is estimated from the later study.[157] Ribonuclease-T_1 has been estimated to possess 33% α-helix, 24% β-structure, and 43% unordered structure; the protein denatures on titration to pH 11.5, giving alterations to the near-u.v. c.d. spectrum.[158]

Two different forms of aspartate aminotransferase, the soluble and mitochondrial enzymes, differ in amino-acid composition, kinetics, and heat stability.[159] The c.d. spectrum of the mitochondrial enzyme suggests that some α-helix is present, whereas the soluble enzyme appears to

[152] J. H. Rockey, P. C. Montgomery, and K. J. Dorrington, *Biochemistry*, 1970, **9**, 4310.
[153] E. Doi and B. Jirgensons, *Biochemistry*, 1970, **9**, 1066.
[154] P. Johnson and J. N. Miller, *Biochim. Biophys. Acta*, 1970, **207**, 308.
[155] D. Frommel, G. W. Litman, W. D. Terry, R. A. Good, and A. Rosenberg, *Biochim. Biophys. Acta*, 1970, **221**, 399.
[156] A. Marzotto and L. Galzigna, *Arch. Biochem. Biophys.*, 1970, **137**, 373.
[157] W. A. Klee and R. A. Streaty, *J. Biol. Chem.*, 1970, **245**, 1227.
[158] Y. Yamamoto and J. Tanaka, *Biochim. Biophys. Acta*, 1970, **207**, 522.
[159] L. H. Bertland and N. O. Kaplan, *Biochemistry*, 1970, **9**, 2653.

possess β-structure. The soluble enzyme consists of three differently charged proteins of identical amino-acid composition, although they are of different conformations, as shown by differences in their c.d. spectra. Pig-heart lipoamide dehydrogenase also exists in multiple forms, which appear to result from conformational alterations about the active site.[160] Four different forms of neurotoxin II from *Androctonus australis hector* exist in rapid equilibrium, the proportions of each form depending on pH and temperature, and o.r.d. spectra have been used to allow the drawing of a state diagram.[161] Two different types of 6-phosphogluconate dehydrogenase from *Candida utilis* differ in their primary structure but have apparently similar tertiary structures; neither shows any change in the o.r.d. spectrum when coenzymes and substrates are bound.[162]

A comparison of β-lactoglobulins from cow, sheep, and buffalo has shown that the side-chains and their conformations differ between the three proteins; differences also exist between the α-lactalbumins of the three species, although they are less pronounced than for the lactoglobulins.[163] Creatine kinase has been compared with kinases from invertebrates, and it has been found that they possess similar backbone conformations, but show differences in their near-u.v. c.d. spectra.[164]

The c.d. spectra of several glycoproteins have been reported. A 4*S*-$\alpha_2\beta_1$-glycoprotein from human plasma shows a minimum at 218 nm in the o.r.d. spectrum, shifting to 232 nm in 50% 2-chloroethanol, and to 211 nm in 5M urea.[165] A freezing-point-depressing glycoprotein from an Antarctic fish shows a positive maximum at 218 nm in its c.d. spectrum, with a negative minimum at 197 nm.[166] Bromelain shows a conformational change between pH 10 and 12, shown by a decrease in the magnitude of the o.r.d. trough at 233 nm. Between pH 12 and 13, only small further changes occur, but the sedimentation coefficient shows a sharp decrease, indicative of a further conformational change.[167] The far-u.v. c.d. spectrum suggests that regular structures exist in the protein, and a band at 250 nm shows a linear increase with the number of ionised tyrosyl residues.[168]

Concanavalin-A shows a far-u.v. c.d. spectrum not unlike that of poly-L-lysine in the β-conformation, but the spectrum appears to be red-shifted by some 5 nm.[169] If this is a true β-structure spectrum, then its magnitude indicates about 50% β-structure in the protein. In 5mM dodecyl sulphate, the protein appears to adopt an α-helical conformation in part, with loss

[160] Y. Sakurai, Y. Fukoyoshi, H. Hamada, T. Hayakawa, and M. Koiki, *J. Biol. Chem.*, 1970, **245**, 4453.
[161] R. Chichiportiche and M. Lazdunski, *European J. Biochem.*, 1970, **14**, 549.
[162] M. Rippa, S. Pontremoli, and P. Fasella, *European J. Biochem.*, 1970, **16**, 461.
[163] F. de Felice and M. Vitagliano, *Biokhimya* (English edition), 1970, **35**, 187.
[164] C. Oriol and M. F. Landon, *Biochim. Biophys. Acta*, 1970, **214**, 455.
[165] T. Iwasaki and K. Schmid, *J. Biol. Chem.*, 1970, **245**, 1814.
[166] A. L. De Vries, S. K. Komatsu, and R. E. Feeney, *J. Biol. Chem.*, 1970, **245**, 2901.
[167] T. Murachi and M. Yamazake, *Biochemistry*, 1970, **9**, 1935.
[168] T. Sakai, K. Ikeda, K. Hamaguchi, and T. Murachi, *Biochemistry*, 1970, **9**, 1939.
[169] C. M. Kay, *F.E.B.S. Letters*, 1970, **9**, 78.

of the original β-structure. Further evidence for the formation of ordered structures in dodecyl sulphate has been obtained in a study of F1 and F3 histones.[170]

A study of insulin, des-octapeptide-insulin, and the *C*-terminal B-chain heptapeptide has shown that the helical content of the molecule is left largely unchanged by the loss of the octapeptide.[171] Analysis by the method of Greenfield and Fasman[125] gave 27—30% and 30—36% of α-helix for the native and modified molecules respectively. Some β-structure is lost during the modification.

RNA polymerase possesses about 30% of α-helix, and dissociation in 2M urea gives a 10% decrease in helix content; the helix content is also reduced in 1M urea, and 30% of the enzyme activity remains. No change in the c.d. at 220 nm is induced by binding to DNA, but heating from 25 to 60 °C shows a lack of stability of the enzyme in the complex form; this instability is not found when the enzyme is complexed to RNA or in the DNA complex after the start of RNA synthesis.[172]

Hemerythrin from *Sipunculus nudus* shows no visible c.d. spectrum when deoxygenated, but acquires bands at 340, 430, and 520 when in the oxy form.[173] Comparison of the near-u.v. c.d. bands of the two forms suggests that a reorientation of aromatic side-chains occurs when oxygen binds, although there is a possibility that the observed changes may be caused by altered iron transitions. There is no overall conformational change when oxygen binds.

A low helical content has been reported for mushroom tyrosinase, and the protein suffers an irreversible conformational change at pH 2—3.[174] A difference has been observed between the c.d. spectra of two phytochrome forms, in the visible, and the backbone seems to contain little helix in either case.[175] Mercurial-induced effects in phycoerythrin depend on whether or not the molecule is aggregated, and the c.d. spectrum of the aggregated species suggests that bound pigment molecules interact with one another.[176]

4 Nuclear Magnetic Resonance

contributed by R. Henson

A. General.—During 1970 there has been almost a three-fold increase in the number of papers published on the applications of n.m.r. to the study of amino-acids, peptides, and proteins in solution: a growth reflected by

[170] V. I. Okulov and G. V. Troitskii, *Biokhimya* (English edition), 1970, **35**, 196.
[171] C. J. Menendez and T. T. Herskovits, *Arch. Biochem. Biophys.*, 1970, **140**, 286.
[172] R. L. Novak and P. Doty, *Biochemistry*, 1970, **9**, 1739.
[173] F. Bossa, M. Brunori, G. W. Bates, E. Antonini, and P. Fasella, *Biochim. Biophys. Acta*, 1970, **207**, 41.
[174] H. W. Duckworth and J. E. Coleman, *J. Biol. Chem.*, 1970, **245**, 1613.
[175] G. R. Anderson, E. L. Jenner, and F. E. Mumford, *Biochim. Biophys. Acta*, 1970, **221**, 69.
[176] F. Fujimori and J. Pecci, *Biochim. Biophys. Acta*, 1970, **221**, 132.

the appearance of an abundance of reviews on the biological applications of n.m.r.[177–184] and a more diversified approach to the subject. As expected, there has been more extensive interest in obtaining spectra by Fourier transform spectroscopy. In this technique the free induction decay following a 90° pulse is converted to the frequency-scan spectrum preferably by an on-line digital computer, and offers enhanced signal-to-noise ratio per unit time relative to the corresponding continuous wave (c.w.) spectrum.[185] This is of particular relevance to the study of naturally-occurring, unstable compounds, especially in view of the expected popularity of ^{13}C n.m.r. spectroscopy. The other main advantage is that a high-resolution spectrum can be obtained in a time which is short compared with nuclear relaxation times, and therefore the Fourier transform of the free induction decay following a 180°–τ–90° pulse sequence, for various values of the delay, τ, makes available high resolution T_1 values.[186] Richards and co-workers[187] have recently illustrated the use of this technique to measure exchange rates and Mn^{2+}–P distances in manganese–ATP complexes by ^{31}P n.m.r.

B. Amino-acids.—(See also Chapter 1, Section 3D.) In general, ^{14}N magnetic resonance spectra have very broad lines. Asymmetry of the electron-charge distribution about the nucleus produces electric-field gradients which interact with the quadrupole moment of the nitrogen. This quadrupolar term, however, does not contribute to the relaxation of $^{15}N(I = \frac{1}{2})$ but use of this nucleus is limited by its low natural abundance (0.365%). Therefore, ^{15}N-enriched amino-acids have been used to investigate the dependence of vicinal N–H coupling constants, J^{NH}, on the dihedral angle of the C_α–C_β bonds.[188] The conformations of the amino-acids were defined in terms of the ^1H–^1H vicinal coupling constants and, from the ^{15}N spectra, J^{NH}_{gauche} and J^{NH}_{trans} were calculated. Alanine has no preferred conformation about the C_α–C_β bond, and measured J^{NH} values were compatible with a statistical average over the staggered rotamers. However, the temperature dependence of the vicinal H–H coupling constants of the α–β protons in phenylalanine, tyrosine, tryptophan, and histidine is anomalous. Concentration affects the magnitudes of J^{HH}, and

[177] G. C. K. Roberts and O. Jardetzky, *Adv. Protein Chem.*, 1970, **24**, 447.
[178] B. Sheard and E. M. Bradbury, *Progr. Biophys. and Mol. Biol.*, 1970, **20**, 187.
[179] M. B. Robin, F. A. Bovey, and H. Basch, 'Chemistry of Amides,' ed., J. Zabicky, Interscience, 1970, p. 1.
[180] J. J. M. Rowe, J. Hinton, and K. L. Rowe, *Chem. Rev*,, 1970, **70**, 1.
[181] M. Cohn, *Quart. Rev. Biophys.*, 1970, **3**, 61.
[182] A. S. Mildvan and M. Cohn, *Adv. Enzymol.*, 1970, **33**, 1.
[183] R. Henson, in 'Amino-acids, Peptides, and Proteins,' ed. G. T. Young, (Specialist Periodical Reports), The Chemical Society, London, 1970, Vol. 2, p. 117.
[184] A. Allerhand and E. A. Trull, *Ann. Rev. Phys. Chem.*, 1970, **21**, 317.
[185] R. R. Ernst and W. A. Anderson, *Rev. Sci. Instr.*, 1966, **37**, 93.
[186] R. L. Vold, J. S. Waugh, M. P. Klein, and D. E. Phelps, *J. Chem. Phys.*, 1968, **48**, 3831.
[187] R. E. Richards, I. D. Campbell, F. F. Brown, C. W. J. Hirst, and R. Henson, 4th International Conference on Magnetic Resonance in Biological Systems, Oxford, 1970. Unpublished results.
[188] R. L. Lichter and J. D. Roberts, *J. Org. Chem.*, 1970, **35**, 2806.

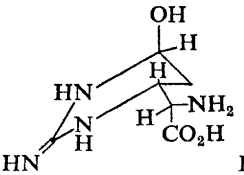
intermolecular interaction was thought to cause variations in the relative energies of the staggered conformers.[189] On the other hand, intramolecular interaction between the aromatic and isoalloxazine rings was thought to be responsible for the preferred conformation of some flavinyl aromatic amino-acids.[190]

Most of the interaction studies of amino-acids were of the spectral perturbations produced by metal ions.[191–194] In particular, Harrison and Rossotti [194] have shown that, in solution, methionine binds to copper(II) in the same way as it does in the solid state: through nitrogen and oxygen, and not through sulphur as do many other thio-carboxylates. Another study was directed towards helping to elucidate the mechanism of the aspartate transaminase reaction.[195] The 100 MHz spectrum of monosodium L-glutamate was recorded in D_2O. In the presence of the enzyme at pH 8 the α-H resonance disappeared, because of exchange of the proton with the solvent. Similar results were obtained with L-aspartate, but not with other amino-acids. Furthermore, no exchange was observed in the spectra of L-glutamate and L-aspartate in the presence of the apo-enzyme. Of related interest is the characterisation of the condensation products of the polyfunctional amino-acids with pyridoxal.[196] Serine, homoserine, and threonine exhibited no interaction between the amino-acid hydroxy-group and the azomethine link; whereas cysteine and homocysteine formed thiazolidines, homocystine formed two Schiff bases and histidine condensed to a tetrahydropyridine.

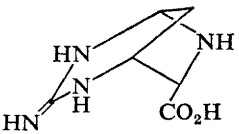
Coupled gas–liquid chromatography and mass spectrometry [197] have been used to characterise the hydrolysis products of the lyophilised algae *Chlorella vulgaris* grown in a medium in which 15% ^{13}C-enriched CO_2 was the sole carbon source.[198] The ^{13}C n.m.r. spectra of the amino-acids were recorded at 15 MHz in D_2O solution by continuous wave and Fourier transform n.m.r. utilising noise decoupling of the protons.[199] Resonances were assigned on the basis of the empirical rules of Grant and Paul [200] and, in cases of doubt, confirmed by coherent decoupling. No tryptophan, cysteine, or cystine was isolated because of their low abundance in the algae proteins, and glutamine and asparagine were hydrolysed to the corresponding acids. The isolation and separation of the analogous

[189] J. R. Cavanaugh, *J. Amer. Chem. Soc.*, 1970, **92**, 1488.
[190] W. Föry, R. E. Mackenzie, F. Ying-Hsiueh Wu, and D. B. McCormick, *Biochemistry*, 1970, **9**, 515.
[191] L. S. Kan and N. C. Li, *J. Amer. Chem. Soc.*, 1970, **92**, 281.
[192] K. C. Tewari, J. Lee, and N. C. Li, *Trans. Faraday Soc.*, 1970, **66**, 2069.
[193] C. E. Strouse and N. A. Matwiyoff, *Chem. Comm.*, 1970, 439.
[194] M. R. Harrison and J. F. C. Rossotti, *Chem. Comm.*, 1970, 175.
[195] Y. Kobayashi and K. Makino, *Biochim. Biophys. Acta*, 1969, **191**, 738.
[196] E. H. Abbott and A. E. Martell, *J. Amer. Chem. Soc.*, 1970, **92**, 1754.
[197] W. Vanderheuval and J. S. Cohen, *Biochim. Biophys. Acta*, 1970, **208**, 251.
[198] J. S. Cohen, W. Horsley, and H. Sternlicht, *Biochim. Biophys. Acta*, 1970, **222**, 521.
[199] W. Horsley, H. Sternlicht, and J. S. Cohen, *J. Amer. Chem. Soc.*, 1970, **92**, 680.
[200] D. M. Grant and E. M. Paul, *J. Amer. Chem. Soc.*, 1964, **86**, 2984.

deuterium-substituted amino-acids has also been described.[201, 202] In these, all the protons were fully replaced except for those attached to β-C of aspartic acid, γ-C of glutamic acid (both $\frac{2}{3}$ substituted), C-2 of histidine imidazole (rapidly exchanged), and the protons *ortho* to the hydroxy-group in tyrosine. Also, proton magnetic resonance, in conjunction with chemical methods, has led to the elucidation of the structure of a guanidino amino-acid isolated from a peptide antibiotic, tuberactinomycin I. The amino-acid, tuberactidine (1), was also thought to be a component of



(1)



(2)

the analogous antibiotic viomycin, and viomycidine (2) was shown to be merely an artefact produced from tuberactidine during the hydrolysis of the peptide.[203]

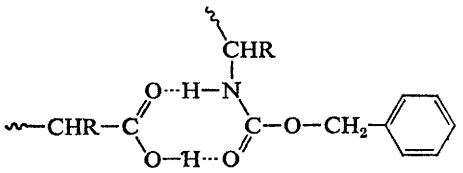
C. Peptides.—Some of the numerous factors which influence the conformation of peptides and proteins have been highlighted again by contemporary n.m.r. studies of polypeptides.[204–212] The usefulness of these compounds as models for proteins relies upon the realisation of all the possible effectors of conformational change and stabilisation, and upon the assignment of the various protons to the peaks in the spectra.

The arylmethyl side-chains of some cyclic dipeptides interact with the oxopiperazine ring; however, electronic perturbations in the benzene ring from *para*-substituents cannot be correlated with the shifts of the C-6 proton.[213] It is thought, therefore, that a charge-transfer mechanism does not make a major contribution to the interaction. Preferred side-chain orientation is also observed in poly-(γ-benzyl-L-aspartate) in chloroform,

[201] H. L. Crespi, H. F. Daboll, and J. J. Katz, *Biochim. Biophys. Acta*, 1970, **200**, 26.
[202] J. S. Cohen and I. Putter, *Biochim. Biophys. Acta*, 1970, **222**, 515.
[203] T. Wakamiya, T. Shiba, T. Kaneko, H. Sakakibara, T. Take, and J. Abe, *Tetrahedron Letters*, 1970, **40**, 3497.
[204] F. J. Joubert, N. Lotan, and R. A. Scheraga, *Biochemistry*, 1970, **9**, 2197.
[205] J. A. Ferretti and B. W. Ninham, *Macromolecules*, 1970, **3**, 30.
[206] J. A. Glasel, *J. Amer. Chem. Soc.*, 1970, **92**, 375.
[207] M. Goodman, A. S. Verdini, C. Toniolo, W. D. Phillips, and F. A. Bovey, *Proc. Nat. Acad. Sci. U.S.A.*, 1969, **64**, 444.
[208] A. E. Tonelli and F. A. Bovey, *Macromolecules*, 1970, **3**, 410.
[209] W. A. Gibbons, G. Nemethy, A. Stern, and L. C. Craigh, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 239.
[210] V. Madison and J. Schellman, *Biopolymers*, 1970, **9**, 511.
[211] A. E. Tonelli, A. I. Brewster, and F. A. Bovey, *Macromolecules*, 1970, **3**, 412.
[212] M. Ihnat and R. Bersohn, *Biochemistry*, 1970, **9**, 4555.
[213] Z. Kopple and K. D. Kopple, *J. Org. Chem.*, 1970, **35**, 253.

where non-equivalence of the benzyl CH_2 and β-CH_2 resonances gives rise to AB quartets.[214]

Assignment of α-CH resonances in polypeptides usually have been made with confidence.[215] The difficulty has been the correlation of the shifts with those in free amino-acids [216] and in the interpretation of the 'fast' and 'slow' rates of helix to random-coil transitions.[217–219] Difficulty, of course, also arises in concentrated solutions of polypeptides and in solutions of high molecular weight polymers, both of which yield uninformative spectra because of incomplete averaging of dipolar interactions. Nevertheless, helix to random-coil transitions have been followed in these systems by investigating the n.m.r. spectra of the solvent molecules. The initial experiments were conducted on solutions of poly-(γ-benzyl-L-glutamate) (PBLG) in chloroform and trifluoroacetic acid mixtures.[220] The proton shift of the helix-breaking component in these solutions relative to the corresponding shifts in the pure solvent were plotted against trifluoroacetic acid composition. The discontinuities in such plots occurred at the helix to random-coil transition (observed by optical rotation measurements) and arose from differences in the solvation behaviour of the two PBLG conformations. Dipeptides, however, have relatively simple n.m.r. spectra, and assignments would be expected to be straightforward. In spite of this, the assignment of the protons in protected dipeptides in concentrated organic solutions [221] has been found not to corroborate those of Bystrov *et al.*[222] Below 289 K additional lines appeared in the spectra, and the *N*-terminal α-CH resonances in Boc- and Z-L-Ala-L-Ala-OMe were always broader than those at the *C*-terminus. Indeed, this relative broadening was enhanced in less polar solvents and the corresponding *N*-protected amino-acids showed similar spectral characteristics, but more distinctly. These results were interpreted in terms of a hydrogen-bonded structure involving the carbonyl group of urethane, as in structure (3).



(3)

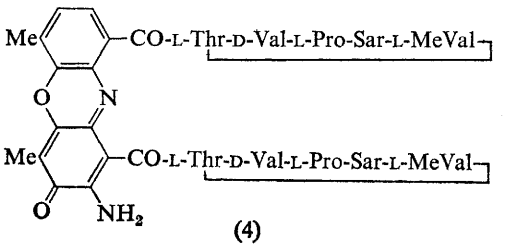
[214] E. M. Bradbury, B. G. Carpenter, C. Crane-Robinson, and H. Goldman, *Nature*, 1970, **225**, 258.
[215] M. Boublik, E. M. Bradbury, C. Crane-Robinson, and H. W. E. Rattle, *European J. Biochem.*, 1970, **12**, 258.
[216] F. Conti, C. Pietronero, and P. Viglino, *Org. Magnetic Resonance*, 1970, **2**, 131.
[217] R. Ullman, *Biopolymers*, 1970, **9**, 471.
[218] R. Cerf, *Compt. rend.*, 1969, **269**, *D*, 2140.
[219] J. A. Ferretti, B. W. Ninham, and V. A. Parsegian, *Macromolecules*, 1970, **3**, 34.
[220] Kang-Jen Liu and S. J. Lignowski, *Biopolymers*, 1970, **9**, 739.
[221] J. L. Dimicoli and M. Ptak, *Tetrahedron Letters*, 1970, **23**, 2013.
[222] V. F. Bystrov, S. L. Portnova, V. I. Tsetlin, V. T. Ivanov, and Y. A. Ovchinnikov, *Tetrahedron*, 1969, **25**, 493.

D. Histones.—The electrostatic binding exhibited by dipeptides for sRNA [223] is probably responsible for the intimate association of basic proteins, histones, with DNA. Histones from calf thymus can be separated into five fractions: *F*1, *F*2a$_1$, *F*2a$_2$, *F*2b, and *F*3, although chicken erythrocyte nucleohistones, for example, contain an additional lysine-rich fraction. These same histone fractions are present in a wide variety of tissues and in some cases the amino-acid sequence is almost completely conserved from one organism to another. This can be reconciled with a specific conformation or conformational change being a functional property of histones.[224] Bradbury and co-workers [225–227] have studied the n.m.r. spectra at 100 and 220 MHz of those fractions whose primary sequence is known: *F*2a$_1$, *F*2b, and *F*1. These primary sequences show that *F*1 (M.W. = 21000) is rich in basic residues in the *C*-terminal half of the molecule (100—216), whereas *F*2a$_1$ (M.W. = 11300) is rich in basic residues in the *N*-terminal half (1—51 of 101 residues). On the other hand, the *N*-terminus (1—51) of *F*1 and the *C*-terminus (90—101) of *F*2a$_1$ are rich in proline and glycine respectively, and are therefore unlikely to exhibit helical structure. The high-resolution spectra of *F*2a$_1$ indicate [226] reduced mobility over a large part of the carboxy half of the molecule, and a localised aggregation, different from that observed in lyphilised histones, has been invoked to account for part of the observed broadenings. Helix formation in the carboxy half of the molecule, probably over residues 55—72, is induced on addition of salts to aqueous solutions of *F*2a$_1$, and subsequently histone–histone interactions take place by hydrophobic bonding. This is vindicated by broadening and loss of area of the proton resonances from the CH_3 of Leu, Ile, Val, Thr, the CH_2 adjacent to CO_2H in Glu and Asp, and the aromatic protons in Tyr and Phe, whereas peaks from the basic residues and glycine appear unaffected. In only about one-quarter of the polypeptide chain of *F*1 (51—100) is the mobility reduced on increasing the salt molarity, and this is caused by the formation of a small amount of intra- and inter-chain β-conformation, although the exclusion of the α-helix is not complete. *F*2b (M.W. = 13770) similarly shows conformational changes and interactions in the centre of the molecule (226—281) with the mobility of the termini unrestricted.[227] Previous studies [228] of *F*2a$_2$ and *F*3 at 60 MHz gave similar results and these molecules may be expected to have sections of polypeptide chain rich in basic and helix-destabilising residues, while other regions have sequences

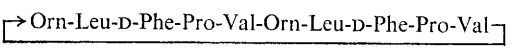
[223] J. M. Backer, Y. N. Molin, and A. S. Sorokin, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 871.
[224] E. M. Bradbury, *Biochem. J.*, 1969, **114**, 49P.
[225] E. M. Bradbury and C. Crane-Robinson (in press).
[226] M. Boublik, E. M. Bradbury, and C. Crane-Robinson, *European J. Biochem.*, 1970, **14**, 486.
[227] M. Boublik, E. M. Bradbury, C. Crane-Robinson, and E. W. Johns, *European J. Biochem.*, 1970, **17**, 151.
[228] E. M. Bradbury, C. Crane-Robinson, H. Goldman, and H. W. E. Rattle, *J. Mol. Biol.*, 1967, **29**, 507.

resembling those of globular proteins. Studies directed towards demonstrating the mode of interaction between histones and DNA are in progress at present.

E. Hormones and Peptide Antibiotics.—Actinomycin D (4) inhibits various DNA-dependent reactions. A study of the model system of actinomycin D and 5-deoxyguanylic acid sheds some light on the mechanism of this inhibition, perhaps, since stacking of the phenoxazine and guanine rings is observed.[229] The nitrogen atom of the phenoxazine ring is also involved in hydrogen bonding to one of the threonine NH groups and, consequently, this explains the non-equivalence of the two peptide lactone rings.[230] Also, hydrogen bonding is in evidence between the two rings,[231, 232] in contrast to gramicidin S (5) [233, 234] and some ferrichrome [235] (6) and oxytocin (7) [236, 237]



(4)

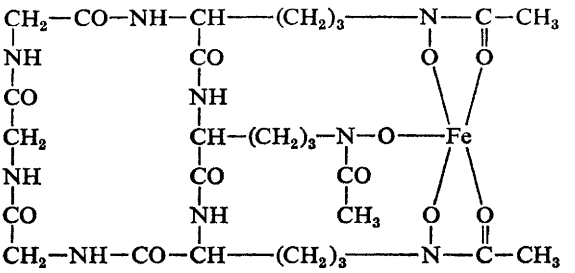


(5)

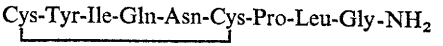
derivatives. In deamino-oxytocin, for example, the hydrogen bonding occurs between asparagine and tyrosine, and establishes a β-turn between these residues. Oxytocin also exhibits this feature, but not 5-valine oxytocin. Furthermore, the latter has low biological activity, a property common to those oxytocin analogues whose asparagine residue has been replaced.

In the remaining antibiotics studied, metal ions play an important rôle. The binding of zinc to bacitracin (8) has been demonstrated by perturbation of the spectrum of the antibiotic by the metal; and, in conjunction with o.r.d. and spectrophotometric titrations, the n.m.r. results suggest that

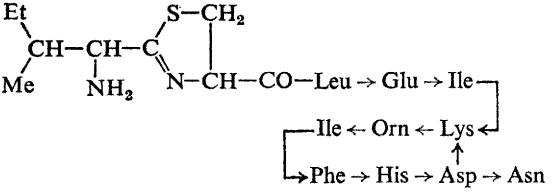
[229] B. H, Arisa and K. Hoogsteen, *Biochemistry*, 1970, **9**, 3976.
[230] T. A. Victor, F. E. Hruska, C. L. Bell, and S. S. Danyluk, *Tetrahedron Letters*, 1969, 4721.
[231] F. Conti and P. De Santis, *Nature*, 1970, **227**, 1239.
[232] H. Lackner, *Tetrahedron Letters*, 1970, **36**, 3184.
[233] Y. A. Ovchinnikov, V. T. Ivanov, V. F. Bystrov, A. I. Miroshnikov, E. N. Shepel, N. D. Abdullaev, E. S. Efremov, and L. B. Senyavina, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 217.
[234] W. A. Gibbons, J. A. Sogn, A. Stern, L. C. Craig, and L. F. Johnson, *Nature*, 1970, **227**, 840.
[235] M. Llinas, M. P. Klein, and J. B. Neilands, *J. Mol. Biol.*, 1970, **52**, 399.
[236] D. W. Urry, M. Ohnishi, and R. Walter, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 111
[237] R. Deslauriers and I. C. P. Smith, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 179.



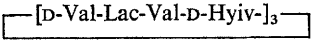
(6)



(7)

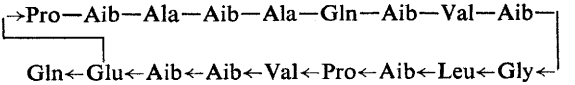


(8)



Lac = lactic acid; Hyiv = α-hydroxyisovaleric acid

(9)



(10)

His-Ser-Gln-Gly-Thr-Phe-Thr-Ser-Asp-Tyr-Ser-Lys-Tyr-Leu-Asp-Ser-Arg-Arg-
5 **10** **15**

Ala-Gln-Asp-Phe-Val-Gln-Trp-Leu-Met-Asn-Thr
20 **25**

(11)

His-Ser-Asp-Gly-Thr-Phe-Thr-Ser-Glu-Leu-Ser-Arg-Leu-Arg-Asp-Ser-Ala-Arg-
5 **10** **15**

Leu-Gln-Arg-Leu-Leu-Gln-Gly-Leu-Val
20 **25**

(12)

zinc was bound to the thiazoline ring and the histidine *N*-3 position.[238] Also the finer details of the conformation of valinomycin (9) have been proposed[239] but, in contrast with valinomycin, the K^+-bound and free alamethicin (10) spectra were indistinguishable.[240] Nevertheless, the spectra of sonicated, aqueous dispersions of phosphatidylcholine and phosphatidylserine were appreciably altered by addition of alamethicin, and were interpreted in terms of different states of aggregation.

The 220 MHz n.m.r. spectra of the hormones glucagon (11) and secretin (12) have also been investigated.[241–243] Evidence from many sources had indicated that glucagon has a helical segment at the *C*-terminus, even in its aggregated forms. However, Patel proposed that the line broadenings and shifts in the ^1H n.m.r. spectra were attributable to aggregation, rather than to secondary structure, since this interpretation was consistent with the concentration dependence of the spectra; moreover, some of the amino-acids in the intermolecular-contact region were identified. Similarly, o.r.d. and c.d. studies had indicated that a helical segment was present at the *N*-terminus of secretin. In spite of this, a comparison of the n.m.r. spectra of secretin and various of its peptide fragments revealed that secretin does exhibit a preferred conformation in solution, but that it is a property of the entire molecule.

F. Non-iron-containing Proteins.—The two papers which dealt with the ^{13}C spectra of proteins illustrated the advantage of using Fourier transform spectroscopy to aid signal-to-noise ratio enhancement per unit time. Lauterbur[244] measured the natural abundance, continuous wave spectrum of hen-egg-white lysozyme (0.025M) at 25 MHz using noise decoupling of the protons. The signal-to-noise ratio after 87 h of accumulation (6260 scans at 50 s each) was about 5, and he suggested that the Fourier transform method should produce better spectra; in fact, a signal-to-noise ratio of 35 to 40 (65000 scans in 10 h) was obtained from a 0.02M solution of ribonuclease-A at 15 MHz:[245] a 100-fold increase in sensitivity. Also, Allerhand and co-workers[245] claimed that most of the peaks were observable after only

[238] N. W. Cornell and D. G. Guiney, jun., *Biochem. Biophys. Res. Comm.*, 1970, **40**, 530.
[239] M. Ohnishi and D. W. Urry, *Science*, 1970, **168**, 1091.
[240] H. Hauser, E. G. Finer, and D. Chapman, *J. Mol. Biol.*, 1970, **53**, 419.
[241] B. Bak, J. J. Led, and E. J. Pederson, *J. Mol. Spectroscopy*, 1969, **32**, 151.
[242] D. J. Patel, *Macromolecules*, 1970, **3**, 448.
[243] D. J. Patel, M. Bodanszky, and M. A. Ondetti, *Macromolecules*, 1970, **3**, 694.
[244] P. C. Lauterbur, *Appl. Spectroscopy*, 1970, **24**, 450.
[245] A. Allerhand, D. W. Cochran, and D. Doddrell, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 1093.

1 h. The regions of the spectrum were tentatively assigned, and the intensities were found to agree quite well with those expected after correction for nuclear Overhauser effects.[246]

Among the other papers concerned with ribonuclease,[247–250] that by Cohen-Addad [249] gave a method for following the denaturation processes by means of a parameter derived from the first moment of the aromatic region of the spectrum; and that by Cohen *et al.*[250] described a curve-fitting programme which was used to interpret histidine pH-titration data. A similar study [251] was also made of staphylococcal nuclease, but the results appeared to be at variance with those of Jardetzky *et al.*[252] who thought that one of the four histidine residues, but none of the other aromatic amino-acid residues, was involved in a conformational equilibrium of the enzyme. The same group also studied the binding of 3′,5′-thymidine diphosphate to selectively-deuteriated [253] and fully active [254] Staphylococcal nuclease,[255] and it was shown that the inhibitor utilised three tyrosine residues in its binding.

Binding of tyrosine residues has also been invoked as the means whereby Ga^{3+} and Fe^{3+} attach themselves to conalbumin and transferrin.[256] Plots of absorbance (at 243 and 294 nm) against added Ga^{3+} concentration indicated that two moles of gallium bound per mole of conalbumin, and that four tyrosine residues were ionised during the binding. The studies were repeated with transferrin, and Fe^{3+}, which inhibits gallium binding, and Ga^{3+} were thought to interact with the proteins analogously. The aromatic region of the ^1H n.m.r. spectra at 220 MHz supported these conclusions, and simulated spectra also suggested that Fe^{3+} binding produced changes in the environment of the tryptophan residues. A comparable n.m.r. experiment was carried out on the copper protein, azurin, and the results were consistent with fluorescence and e.s.r. data.[257]

From the similar amino-acid sequences of bovine α-lactalbumin and lysozyme one would expect some degree of conformational similarity between the two proteins, especially as the four disulphide bonds are preserved. A comparison of the upfield region of the ^1H n.m.r. spectra shows discrepancies that can nearly all be accounted for in terms of

[246] K. F. Kuhlmann, D. M. Grant, and R. K. Harris, *J. Chem. Phys.*, 1970, **52**, 3439.
[247] T. Schleich and P. N. von Hippel, *Biochemistry*, 1970, **9**, 1059.
[248] T. D. Alger, *Biochemistry*, 1970, **9**, 1059.
[249] J. P. Cohen-Addad, *J. Mol. Biol.*, 1970, **50**, 595.
[250] J. S. Cohen, R. I. Shrager, M. McNeel, and A. N. Schechter, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 144.
[251] J. S. Cohen, R. I. Shrager, M. McNeel, and A. N. Schechter, *Nature*, 1970, **228**, 642.
[252] J. L. Markley, M. N. Williams, and O. Jardetzky, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **65**, 645.
[253] I. Putter, A. Barreto, J. L. Markley, and O. Jardetzky, *Proc. Nat. Acad. Sci. U.S.A.*, 1969, **64**, 1396.
[254] I. Putter, J. L. Markley, and O. Jardetzky, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **65**, 395.
[255] J. L. Markley and O. Jardetzky, *J. Mol. Biol.*, 1970, **50**, 223.
[256] R. C. Woodworth, K. G. Morallee, and R. J. P. Williams, *Biochemistry*, 1970, **9**, 839.
[257] G. Rotilio, A. Finazzi Agro, L. Avigliano, A. Lai, F. Conti, C. Franconi, and B. Mondovi, *F.E.B.S. Letters*, 1970, **12**, 114.

sequence differences.[258] From a tentative *X*-ray structure, ring-current shifts can be calculated for the various aliphatic side-chains which are responsible for these high-field shifts, and the results are in semi-quantitative agreement with the observed spectrum; moreover, similar predictions from the *X*-ray structure of lysozyme yield n.m.r. spectra consistent with the assignments of McDonald and Phillips.

Finally, proton relaxation enhancement has continued to be used to probe the metal binding sites of proteins, furnishing their number, their dissociation constant, and in some cases structural information about the complexes. The systems studied this year have been: phosphoenolpyruvate synthetase,[259] histidine ammonia lyase,[260] phosphoglucomutase,[261] creatine kinase,[262] and β-methyl aspartase;[263] however, one ventures to point out that computer, rather than graphical, data analysis is more reliable.[182] Although all of the above experiments were performed using manganese(II) as the paramagnetic probe, gadolinium(III) has also been considered suitable.[264–266] In view of this, the study of the binding of Gd^{3+} to bovine serum albumin is of interest.[266] Four binding sites are found, of dissociation constant 1.3×10^{-4} mol l^{-1} (*cf.* Ca^{2+}), and the mechanism of the enhancement is discussed in the light of temperature and frequency dependencies, and electron spin relaxation rates[267] of the Gd^{3+} ion. Related experiments are those of Koenig *et al.*[268, 269] who have investigated the metal sites of cobalt carbonic anhydrase and copper transferrin; investigations of globular and fibrous proteins have also been made by the field dependencies of relaxation times.[270]

G. Iron Proteins.—The rubredoxins are one of the simplest types of non-haem iron proteins. That from *Clostridium pasteurianium* has a single polypeptide chain (55 amino-acids, 4 cysteine residues, M.W. = 6380) and contains one iron atom per molecule which undergoes a one-electron redox reaction. Both the oxidised and reduced forms are stable, having effective magnetic moments, μ_{eff}, 5.85 and 5.05 Bohr magnetons respectively, corresponding to high-spin Fe^{III} and high-spin Fe^{II}.[271] In reduced rubre-

[258] D. A. Cowburn, E. M. Bradbury, C. Crane-Robinson, and W. B. Gratzer, *European J. Biochem.*, 1970, **14**, 83.
[259] K. M. Berman and M. Cohn, *J. Biol. Chem.*, 1970, **245**, 5309.
[260] I. L. Givot, A. S. Mildvan, and R. H. Abeles, *Fed. Proc.*, 1970, **29**, 531.
[261] W. J. Ray and A. S. Mildvan, *Biochemistry*, 1970, **9**, 3886.
[262] J. S. Taylor, J. S. Leigh, and M. Cohn, *Proc. Nat. Acad. Sci. U.S.A.*, 1969, **64**, 219.
[263] G. A. Field and H. J. Bright, *Biochemistry*, 1970, **9**, 3801.
[264] K. G. Morallee, E. Nieboer, F. J. C. Rossotti, R. J. P. Williams, A. V. Zavier, and R. A. Dwek, *Chem. Comm.*, 1970, 1132.
[265] D. W. Darrell, J. E. Gomez, and E. R. Birnbaum, *J. Amer. Chem. Soc.*, 1970, **92**, 5287.
[266] J. Reuben (in press).
[267] J. Reuben, G. H. Reed, and M. Cohn, *J. Chem. Phys.*, 1970, **52**, 1617.
[268] B. P. Gaber, W. E. Schillinger, S. H. Koenig, and P. Aisen, *J. Biol. Chem.*, 1970, **245**, 4251.
[269] M. E. Fabry, S. H. Koenig, and W. E. Schillinger, *J. Biol. Chem.*, 1970, **245**, 4256.
[270] R. Kimmich and F. Noack, *Z. Naturforsch.*, 9170, **24a**, 299.
[271] W. D. Phillips, M. Poe, J. F. Weiher, C. C. McDonald, and W. Lovenberg, *Nature*, 1970, **227**, 574.

doxin, peaks observed in the n.m.r. spectrum at 220 MHz between − 2 and − 5 p.p.m. exhibit a Curie law temperature dependence and their shifts are attributed to contact interaction. Oxidised rubredoxin has a more indistinct spectrum and no peaks can be unequivocally identified as arising from contact interactions. Phillips *et al.*[272] have also observed sharp contact-shifted resonances in ferredoxin (8 iron atoms) from *C. pasteurianium* and have assigned them to the β-CH_2 group of cysteine residues which are thought to bind the eight iron atoms to the polypeptide chain. For oxidised and reduced ferredoxin the contact-shifted resonances are mainly to lower field than those of reduced rubredoxin, which also exhibits high-field resonances, and these shifted peaks increase in width as the shift becomes more extreme. This has led to pseudocontact interaction being proposed as the mechanism of the shifts, since this affects the line-widths by electron–nuclear dipolar interaction, which diminishes as the sixth power of the distance and increases as μ_{eff}^2. Also, anisotropy in the e.s.r. of oxidised rubredoxin [273, 274] indicates sufficient asymmetry of the ligand field to allow pseudocontact interaction. *X*-Ray and n.m.r. results [275] were the basis of a discussion of a cluster model for the redox centre of *Chromatium* high potential iron protein. Contact shifts in the formally diamagnetic reduced state increase with increase in temperature, indicating the presence of antiferromagnetic coupling as observed in ferredoxin (μ_{eff} = 1.05 Bohr magnetons). In the oxidised form, formally $S = \frac{1}{2}$, the β-CH_2 group of the two cysteines display a Curie law temperature dependence while the two others display antiferromagnetic behaviour.

Details of two very elegant experiments on cytochrome *c*, utilising Fourier transformation, have been published by Gupta and Redfield.[276, 277] In the first,[276] the 100 MHz spectrum of approximately 50%-reduced cytochrome *c* is characterised as the superposition of the oxidised and reduced proteins and, when one of the well-separated, hyperfine-shifted lines in the oxidised state is saturated, cross-saturation occurs. The transfer of saturation is to the resonance of the same proton in the reduced state and occurs because interconversion of the oxidised and reduced molecules is at a rate comparable with the measured spin–lattice relaxation rate. The assignment of the methyl group co-ordinated to the face of the haem [278] was checked in this way. In their other study,[277] Gupta and Redfield observed the interaction of azide anion with ferro- and ferri-cytochrome *c*. The former interaction was unobservable, whereas the latter was weak,

[272] M. Poe, W. D. Phillips, C. C. McDonald, and W. Lovenberg, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **65**, 797.
[273] I. A. Peterson and M. J. Coon, *J. Biol. Chem.*, 1968, **243**, 320.
[274] H. Backmayer, L. H. Piette, K. T. Yasanobin, and H. R. Whitely, *Proc. Nat. Acad. Sci. U.S.A.*, 1967, **57**, 122.
[275] W. D. Phillips, M. Poe, C. C. McDonald, and R. G. Bartsch, *Proc. Nat. Acad. Sci., U.S.A.*, 1970, **67**, 682.
[276] R. K. Gupta and A. G. Redfield, *Science*, 1970, **169**, 1204.
[277] R. K. Gupta and A. G. Redfield, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 273.
[278] K. Wuthrich, *Proc. Nat.Acad. Sci. U.S.A.*, 1969, **63**, 1071.

since in excess of azide the superposition spectrum of bound and unbound ferricytochrome *c* showed quite strong unbound peaks. A difference spectrum of the equilibrium mixture was recorded in the presence and absence of a long radio-frequency pulse (0.1 s) applied to saturate one of the well-resolved, hyperfine-shifted resonances of azidoferricytochrome *c*. The relative intensities of the peaks in the absence and presence of the pulse, together with the T_1 of the resonance in the non-bound state, yielded values of the lifetime of the ligand in the bound state.

Progress in assigning resonances to the residues of myoglobin and haemoglobin continues,[279–281] but not without setbacks,[282, 283] although for cyanoferrimyoglobin[282] this has led to the realisation of the probable importance of pseudocontact interactions. If there are changes in interactions caused by amino-acid sequence variations in this molecule,[284] then they do not influence the haem groups, but there are small changes observed in the spectra of cyanoferrihaemoglobins.[285] Extensive small changes in the ring-current-shifted resonances of myoglobin also occur on oxygen binding,[286] and enhanced oxygen-binding affinity has been observed in haemoglobin when arginine α92 is replaced by glutamine (*i.e.* haemoglobin J Capetown) or leucine (*i.e.* haemoglobin Chesapeake), a residue on the $\alpha\beta$ interface. Studies of these [287, 288] and other systems [289, 290] have promoted discussion of the influence on the haem group of changes at such interfaces.

H. Nuclei other than Hydrogen.—Halogen n.m.r. is now well established and papers appear regularly.[291–295] The enzymic activity of manganese carboxy-peptidase and the interaction of the manganese with water are little affected by the presence of fluoride. However, observed ^{19}F broadening is removed by the addition of the inhibitor β-phenylpropionate.[292] In

[279] D. J. Patel, L. Kampa, R. G. Shulman, T. Yamane, and B. J. Wyluda, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 1109; D. J. Patel, L. Kampa, R. G. Shulman, T. Yamane, and M. Fujiwara, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 1224.
[280] R. G. Schulman, J. Peisach, and B. J. Wyluda, *J. Mol. Biol.*, 1970, **48**, 517.
[281] R. M. Lynden-Bell and A. Velenik, *Croat. Chem. Acta*, 1969, **41**, 205.
[282] B. Sheard, T. Yamane, and R. G. Shulman, *J. Mol. Biol.*, 1970, **53**, 35.
[283] G. Pifat, S. Maricic, M. Petrinovic, V, Kramer, J. Marsel, and K. Bonhard, *Croat. Chem. Acta*, 1969, **41**, 195.
[284] K. Wüthrich, R. G. Shulman, T. Yamane, B. J. Wyluda, T. E. Hugli, and F. R. N. Gurd, *J. Biol. Chem.*, 1970, **245**, 1947.
[285] T. Yamane, K. Wüthrich, R. G. Shulman, and S. Ogawa, *J. Mol. Biol.*, 1970, **49**, 197.
[286] R. G. Shulman, K. Wüthrich, T. Yamane, D. J. Patel, and W. E. Blumberg, *J. Mol. Biol.*, 1970, **53**, 143.
[287] S. Ogawa, R. G. Shulman, P. A. M. Kynoch, and H. Lehmann, *Nature*, 1970, **225**, 1042.
[288] D. G. Davis, N. H. Mock, T. R. Lindstrom, S. Charache, and C. Ho, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 343.
[289] D. G. Davis, S. Charache, and C. Ho, *Proc. Nat. Acad. Sci. U.S.A.*, 1969, **63**, 1403.
[290] C. Ho, D. G. Davies, N. H. Mock, T. R. Lindstrom, and S. Charache, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 779.
[291] T. W. Johnson and N. Muller, *Biochemistry*, 1970, **9**, 1943.
[292] G. Navon, R. G. Shulman, B. J. Wyluda, and T. Yamane, *J. Mol. Biol.*, 1970, **51**, 15.
[293] R. L. Ward, *Biochemistry*, 1970, **9**, 2447.
[294] H. Csopak, B. Lindman, and H. Lilja, *F.E.B.S. Letters*, 1970, **9**, 189.
[295] R. L. Ward and K. J. Fritz, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 707.

other studies, ^{35}Cl line-widths were used to indicate changes in the zinc environment of carbonic anhydrase[293] and alkaline phosphatase.[294] For human C carbonic anhydrase, changes are only visible when the enzyme begins to unfold at pH 5, but in the B isozyme a reversible change in the zinc environment occurs before unfolding is possible. Preliminary studies have also been carried out on cobalt-substituted carbonic anhydrase, with a view to an investigation of this system by halogen n.m.r.[295] The observation of n.m.r. spectra of nuclei other than those mentioned above is usually more difficult, because of sensitivity problems, but for the alkali metals n.m.r. is one of the few probes available.[296] ^{23}Na N.m.r.[297] has been used in biological systems in the past, but now preliminary results [298, 299] have been acquired from ^{39}K and ^{205}Tl n.m.r. in solutions of pyruvate kinase, and the proximity of the bivalent and univalent sites is indicated by both experiments. Limited use of ^{25}Mg n.m.r.[300] may also prove to be possible, but ^{33}S n.m.r.[301] seems likely to remain impracticable for some time to come.

Table 1

Reactant	*Protein*	*Ref.*	*Nucleus*
Oxalacetate	Pyruvate Carboxylase	302	^{1}H
Atropine, eserine	Acetylcholinesterase	303	^{1}H
Acetylsalicylic acid	Human serum albumin	304	^{1}H
Succinate	Aspartate transcarbamylase	305	^{1}H
DPNH	Lactate dehydrogenase	306	^{1}H
Phosphoenolpyruvate	Enolase	307	^{1}H, ^{31}P, ^{13}C
trans-Cinnamate	α-Chymotrypsin	308	^{1}H
Trifluoromethyl-substituted α-bromoacetanilides	α-Chymotrypsin	309	^{19}F
Glucose	Phosphoglucose isomerase	310	^{1}H
Glucose	Glucose oxidase	311	^{1}H

DPNH = diphosphopyridine nucleotide (reduced form).

[296] R. J. P. Williams, *Quart. Rev.*, 1970, **24**, 331.
[297] J. A. Magnuson, D. S. Shelton, and N. S. Magnuson, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 279.
[298] R. G. Bryant, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 1162.
[299] F. J. Kayne and J. Reuben, *J. Amer. Chem. Soc.*, 1970, **92**, 220.
[300] R. G. Bryant (in press).
[301] H. L. Retcofsky and R. A. Friedel, *Appl. Spectroscopy*, 1970, **24**, 379.
[302] M. C. Scrutton and A. S. Mildvan, *Arch. Biochem. Biophys.*, 1970, **140**, 131.
[303] G. Kato, J. Yung, and M. Ihnat, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 15.
[304] B. D. Sykes, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 508.
[305] B. D. Sykes, P. G. Schmidt, and G. R. Stark, *J. Biol. Chem.*, 1970, **245**, 1180.
[306] R. H. Sarma and N. O. Kaplan, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 1375.
[307] M. Cohn, J. E. Pearson, E. L. O'Connell, and I. A. Rose, *J. Amer. Chem. Soc.*, 1970, **92**, 4095.
[308] J. T. Gerig and J. D. Reinheimer, *J. Amer. Chem. Soc.*, 1970, **92**, 3146.
[309] E. W. Bittner and J. T. Gerig, *J. Amer. Chem. Soc.*, 1970, **92**, 5001.
[310] M. S. Feather, V. Deshpande, and M. J. Lybyer, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 859.
[311] M. S. Feather, *Biochim. Biophys. Acta*, 1970, **220**, 127.

I. Conclusion.—The major advances in the biological applications of n.m.r. will usually result from advances in instrumentation and technique. This was demonstrated during 1970 by the breakthrough of Fourier transform spectroscopy and ^{13}C n.m.r. but, during the next year or so, the development of higher field superconducting solenoids should give rise to proton spectra at 330 MHz and ^{13}C spectra at 80 MHz.

Interaction studies not explicitly mentioned in this review are listed in Table 1, and various miscellaneous papers are noted as references 312—321.

5 Fluorescence

contributed by J. R. Brocklehurst

The use of various fluorescence techniques to obtain structural information in biological systems is widespread. Several reviews of the field have appeared during the year,[322–324] one dealing with the application of fluorescence polarisation measurements to immunochemistry;[322] another is an extensive review of the relevant theory of fluorescence, and its applications to proteins and membranes,[323] and last year's Report in this series.[324] Two papers of general interest and importance should be mentioned at this point. Teale [325] has examined the effect of light-scattering on the polarisation of fluorescence of turbid solutions (*i.e.* most solutions of biological interest). He first examined the theory, and then drew up various criteria of recognition of this type of depolarisation, and of differentiation of it from other types of depolarisation. By application of these criteria, Teale showed that the change in depolarisation of the dansyl group attached to actomyosin is the result of a change in light-scattering, not of a change in relaxation time. Weber and Shinitzky [326] have found anomalies in the energy transfer between aromatic dyes bound to protein surfaces. The energy transfer disappears at the red edge of the excitation spectrum, a fact which cannot be explained by decreased overlap, or by rapid transfer.

[312] J. M. T. Hamilton-Miller, E. Richards, and E. P. Abraham, *Biochem. J.*, 1970, **116**, 385.
[313] R. Griesser, B. Prijs, H. Sigel, W. Föry, L. D. Wright, and D. B. McCormick, *Biochemistry*, 1970, **9**, 3285.
[314] D. P. L. Sachs, E. Jellum, and B. Halpern, *Biochim. Biophys. Acta*, 1970, **198**, 88.
[315] G. H. Rist, J. S. Hyde, and T. Vänngard, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 79.
[316] L. E. G. Eriksson, A. Ehrenberg, and J. S. Hyde, *European J. Biochem.*, 1970, **17**, 539.
[317] E. P. Katz, *Biopolymers*, 1970, **9**, 745.
[318] G. Chapman, I. D. Campbell, and K. A. McLauchlan, *Nature*, 1970, **225**, 639.
[319] M. Fridkin, M. Wilchek, and M. Sheinblatt, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 458.
[320] R. G. Parker and J. D. Roberts, *J. Org. Chem.*, 1970, **35**, 996.
[321] F. W. Cope and R. Damadian, *Nature*, 1970, **228**, 76.
[322] W. B. Dandliker and V. A. de Saussure, *Immunochem.*, 1970, **7**, 799.
[323] G. K. Radda, in 'Current Topics in Bioenergetics,' vol. 4, ed. D. R. Sanadi, Academic Press (in press).
[324] J. R. Brocklehurst, in 'Amino Acids, Peptides, and Proteins,' ed. G. T. Young, (Specialist Periodical Reports) The Chemical Society, London, 1970, Vol. 2, p. 123.
[325] F. W. J. Teale, *Photochem. Photobiol.*, 1969, **10**, 363.
[326] G. Weber and M. Shinitzky, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **65**, 823.

Fluorescent probes (particularly the naphthalenesulphonate derivatives) continue to attract a great deal of attention. These dyes have been used to detect hydrophobic binding sites in a variety of proteins, as well as to detect structural changes. 1,8-ANS (8-anilinonaphthalene-1-sulphonate) binds to hydrophobic sites on antibodies, the quantum yield being highest when it is bound to antibodies to which it is a hapten analogue.[327] There are two of these binding sites per antibody, both being on the *F*ab fragment which contains the active site. 1,4-ANS and 2,6-ANS will bind to their respective azo-antibodies indicating specific hydrophobic sites. They also bind non-specifically to bovine serum albumin.[328]

1,8-ANS binds in the haem crevice of apo-horseradish-peroxidase with a typical blue shift and enhancement of fluorescence. This binding quenches the small amount of tryptophan fluorescence, and there is only very inefficient energy transfer from tyrosine to ANS. (Nothing is said about tyrosine fluorescence, however.) The transfer from tyrosine to the native haem group is also inefficient.[329]

ANS has been widely used to demonstrate conformational changes in proteins, though it is not always clear whether the change in ANS fluorescence is due to a change in binding constant and/or occupancy, or to a change in quantum yield. The fluorescence of ANS bound to transaldolase is quenched by fructose-6-phosphate. Since there are two ANS binding sites, and only one fructose-6-phosphate binding site per molecule of transaldolase, it is assumed that the quenching is due to a substrate-induced conformational change.[330] The ligand-induced structural changes in glutamate dehydrogenase detected by ANS have become a matter for some debate. Iwatsubo *et al.* repeated the work of Dodd and Radda,[331] but disagreed with the latters' interpretation.[332] Basing their argument on unpublished observations, they claim that the observed structural changes are not rapid enough to account for the ligand-induced inhibition of the enzyme.

Anisotropy of ANS binding has been detected by studying the variation of fluorescence polarisation with temperature and viscosity.[333] ANS bound to various proteins has both specific and random orientations. The anisotropy varies with the number of ANS molecules bound. Brand and Witholt claim[333] that this method is superior to that of Weber and Anderson[334] since the latter depends on energy transfer, and therefore needs at least two bound dye molecules.

[327] C. W. Parker and C. K. Osterland, *Biochemistry*, 1970, **9**, 1074.
[328] T. J. Yoo, H. Nakamura, A. L. Grossberg, and D. Pressman, *Immunochem.*, 1970, **7**, 627.
[329] C.-G. Rosen, *F.E.B.S. Letters*, 1970, **6**, 158.
[330] K. Brand, *F.E.B.S. Letters*, 1970, **7**, 235.
[331] G. H. Dodd and G. K. Radda, *Biochem. J.*, 1969, **114**, 407.
[332] J. M. Jallon, A. di Franco, and M. Iwatsubo, *European J. Biochem.*, 1970, **13**, 428.
[333] B. Witholt and L. Brand, *Biochemistry*, 1970, **9**, 1948.
[334] G. Weber and S. R. Anderson, *Biochemistry*, 1969, **8**, 361.

Several probes related to ANS have also detected conformational changes. 2-(*N*-Methylaminoanilino)-naphthalene-6-sulphonate (MNS) specifically detects the dimer–tetramer conversion of glycogen phosphorylases-*a* and -*b*.[335] The slow enhancement of MNS-phosphorylase-*b* fluorescence caused by AMP correlates with an increase in sedimentation coefficient, the process being reversed by glucose. In the absence of AMP the fluorescence increase correlates with the appearance of phosphorylase-*a* activity. 2-Toluidinonaphthalene-6-sulphonate (TNS) binds to pepsinogen and pepsin with similar binding constants, but with a higher quantum yield on pepsinogen. This can be used to follow the pepsin–pepsinogen interconversion.[336]

5-Dimethylaminonaphthalene-1-sulphonyl (Dns) derivatives of amino-compounds have also been used as probes. Dns-n-propylamine is a reversible competitive inhibitor of serum cholinesterase and indicates a low polarity environment of the active site. It is claimed to respond to conformational changes in the enzyme.[337] Dns-cystine binds to actin with a blue shift and enhancement of fluorescence which is further enhanced on polymerisation of the actin.[338]

A novel fluorescence study makes use of the oxygen-quenching of pyrene-1-butyrate (PB) fluorescence. This probe has a long lifetime which renders it susceptible to quenching, and hence to measurement of the accessibility of its binding sites to oxygen. When bound to bovine serum albumin, polylysine, or apo-haemoglobin its fluorescence is independent of oxygen concentration, but there is an increase in quenching following denaturation with urea.[339]

Several other types of organic molecule have been investigated as possible probes for structural parameters in proteins. 4-Methyl-7-diethylamino-coumarin is a possible probe since its fluorescence is enhanced and blue-shifted as solvent polarity is lowered, while increasing viscosity merely enhances the fluorescence. It has been used to indicate the presence of hydrophobic binding sites on bovine serum albumin A, β-lactoglobulin, and insulin, and their absence on ovalbumin and lysozyme.[340] Auramine O (a cationic probe) is non-fluorescent in water and ethanol, but fluoresces in viscous solutions and when bound to DNA. Horse-liver alcohol dehydrogenase also renders it fluorescent, unlike 16 similar proteins which have no effect.[341] Polarisation measurements indicate that the probe is firmly bound. Binding data show that while the probe is not competitive with alcohol and NAD^+, the latter increases its binding constant.[341]

[335] D. J. Birkett, G. K. Radda, and A. G. Salmon, *F.E.B.S. Letters*, 1970, **11**, 295.
[336] J. L. Wang and G. M. Edelman, *Fed. Proc.*, 1970, **29**, 402.
[337] C. M. Himel, R. T. Mayer, and L. L. Cook, *J. Poiymer. Sci., A*, 1970, **8**, 2219.
[338] H. C. Cheung, L. Smith, and R. Cooke, *Fed. Proc.*, 1970, **29**, 667.
[339] W. M. Vaughan and G. Weber, *Biochemistry*, 1970, **9**, 464.
[340] O. Takenaka, Y. Nishimura, A. Takenaka, and K. Shibata, *Biochim. Biophys. Acta*, 1970, **207**, 1.
[341] R. H. Conrad, J. R. Heitz, and L. Brand, *Biochemistry*, 1970, **9**, 1540.

Covalent labelling avoids some of the problems inherent in using non-covalent probes (*e.g.* uncertainty in interpretation due to changes in binding). Dansyl chloride continues to be the most popular fluorescent labelling agent. It reacts with heavy meromyosin in a 1 : 1 ratio at the active site lysine, and leads to a complete loss of ATPase activity.[342] However, calcium chloride plus potassium chloride gave 80% protection, and non-covalently bound dansyl acid also protected. In labelling the lysine, dansyl chloride also masks one –SH group near the active site, possibly as a result of its rather large hydration sphere.[342] Lysozyme is very unreactive[343] towards dansyl chloride, and only one Dns group per molecule is incorporated, leaving the enzyme fully active. The Dns group extinction coefficient is higher than in other conjugates, and the fluorescence is red-shifted, indicating that the Dns group is very exposed. Unlike similar conjugates of other proteins, the Dns group on lysozyme is labile.[343] Fluorescence polarisation measurements of Dns-labelled immunoglobulin G have indicated rotational freedom of the subunits.[344] Iodide quenching of fluorescein-isothiocyanate-labelled trypsin has been used to follow conformational changes.[345] Pyridoxamine-5′-phosphate bound to phosphorylase-*b* and subsequently reduced by borohydride acts as a 'built-in' probe for pH-induced structural changes.[346]

An important new probe is 7-chloro-4-nitrobenz-2-oxa-1,3-diazole (NBD-Cl).[347] This reacts with $-NH_2$ and –SH groups to give a fluorescent product. The time course of the reaction may be followed spectrophotometrically. The type of group labelled (*i.e.* $-NH_2$ or –SH) may be determined from the absorption and fluorescence properties of the conjugate.[347] Model studies with cystine have been used to show that the –SH group is attacked first, but under favourable conditions of geometry the label is rapidly transferred to an $-NH_2$ group.[347] The conjugate fluorescence is sensitive to environmental changes and has been used to follow actin polymerisation.[347] Studies with NBD-trypsin have shown that this label may be used like Dns, but that it has the advantages of greater reactivity, greater conjugate stability, and ease of following the reaction (by absorption).[348]

Of great importance is the use of the natural fluorescence of biological systems, since this avoids perturbation of the system by an extrinsic probe. The natural fluorescence may arise from a coenzyme or from the protein (or peptide) itself. NADH binding to numerous dehydrogenases has been

[342] H. Takashina, *Biochim. Biophys. Acta*, 1970, **200**, 319.
[343] R. F. Chen, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 1117.
[344] R. S. Nezlin, Y. A. Zagyansky, and L. A. Tumerman, *J. Mol. Biol.*, 1970, **50**, 569.
[345] B. Arrio, F. Rodier, C. Boisson, and C. Vernolte, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 589.
[346] M. Cortijo and S. Shaltiel, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 212.
[347] D. J. Birkett, N. C. Price, G. K. Radda, and A. G. Salmon, *F.E.B.S. Letters*, 1970, **6**, 346.
[348] F. Rodier, M. Hill, B. Arrio, and C. Parquet, *F.E.B.S. Letters*, 1970, **11**, 246.

studied by fluorescence enhancement.[349] The blue shift in the NADH fluorescence has also been used to estimate binding site polarity, though the values obtained differ from those obtained with ANS.[350] This is not surprising since ANS presumably does not bind at the active site.

Much of the work on protein fluorescence concerns tryptophan emission, particularly the way in which it is quenched by substrates and inhibitors. The fluorescence spectra of sperm whale apomyoglobin indicate that its two tryptophans are in a non-polar environment. However, they can be quenched by IO_3^- and this quenching increases as they become more exposed following unfolding of the protein.[351] The variation of the quantum yield with pH in the region pH 5.5—8.3, and the constancy of the lifetime over this pH range, suggest that one Trp is quenched by a neighbouring His when the latter is protonated.[351] Trp-108 of lysozyme is susceptible to oxidation by iodine, and this has been used to show that this residue is responsible for 60% of lysozyme fluorescence.[352] Trisaccharides quench Trp-108, but bind as well to oxidised lysozyme as to the native enzyme. This suggests a small conformational change in the region of Trp-108.[352] This quenching has been used to follow the dynamics of polysaccharide binding. The observed monophasic quenching is consistent with a stepwise mechanism of binding.[353]

The fluorescence spectrum of *Staphylococcus aureus* endonuclease has an iso-emissive point at 336 nm in the temperature range 140—230 K. This is interpreted in terms of two types of fluorescent species – high- and low-temperature forms – the high temperature form being a Trp exciplex.[354] The fluorescence spectra of glyceraldehyde-3-phosphate dehydrogenase at high concentrations have long wavelength peaks which have been assigned to excimer formation between tryptophans in different protein molecules.[355] However, since the absorption spectrum also changes with concentration, this effect cannot be due to excimer formation. It is quite possible that the observed peaks are due to impurity (*e.g.* NADH) which is only detected at high enzyme concentrations.

Measurement of the polarisation of Trp fluorescence yields information about the shape of the protein molecule. The variation of polarisation of pepsin and pepsinogen with temperature, viscosity, and quenching by iodide has given information on the rotational relaxation of the enzymes, and has indicated anisotropy in the excitation due to energy transfer.[356] Changes in the polarisation of horse-liver alcohol dehydrogenase have been

[349] J.-C. Mani, M. Mousseron-Canet, J. Dornand, and M. Arnaud, *Compt. rend.*, 1970, **270**, *D*, 2717.
[350] J.-C. Mani, M. Mousseron-Cadet and J. Dornaud, *Compt. rend.*, 1970, **270**, *D*, 2373.
[351] E. P. Kirby and R. F. Steiner, *J. Biol. Chem.*, 1970, **245**, 6300.
[352] V. I. Teichberg and N. Sharon, *F.E.B.S. Letters*, 1970, **7**, 171.
[353] I. Pecht, V. I. Teichberg, and N. Sharon, *F.E.B.S. Letters*, 1970, **10**, 241.
[354] J. W. Longworth and M. D. C. Battista, *Photochem. Photobiol.*, 1970, **12**, 29.
[355] T. Keleti, *F.E.B.S. Letters*, 1970, **7**, 280.
[356] F. W. J. Teale and R. A. Badley, *Biochem. J.*, 1970, **116**, 341.

correlated with the enzymic activity, and were used to follow conformational changes of the enzyme. Low pH caused dissociation without unfolding, while urea caused a flattening of the polarisation spectrum characteristic of unfolding.[357]

Tyrosine fluorescence is not as intense as that of tryptophan, and so is more difficult to study. Cowgill has examined the various possible mechanisms of peptide quenching of Tyr and Trp in proteins, and suggests that quenching by the carbonyl group of the peptide is the main mechanism.[358] A great deal of work has been done on the various forms of ribonuclease (RNase). Acetylation of the three exposed tyrosines of RNase-A removes the fluorescence, leading to the conclusion that the other three (non-fluorescent) tyrosines are buried. These three become fluorescent when the protein is denatured.[359] In RNase-T_1 fluorescence spectra indicate that the one Trp is partially buried, and that most of the tyrosines are buried. Urea exposes the Trp, while sodium dodecyl sulphate exposes the tyrosines. About a third of the tyrosines are quenched by the Trp and are therefore close to it. The others respond to changes in solvent polarity and are apparently quenched by acidic amino-acid side-chains in the native enzyme.[360] Substrate analogues (which bind at the active site) also quench the Tyr and Trp. From the pH dependence it was possible to deduce that the Trp and one or two of the tyrosines are very close to the active site.[361]

The phycobiliproteins are highly fluorescent. Dale and Teale have purified several of them and determined their fluorescence characteristics.[362] The variation of fluorescence with ionic strength and protein concentration could be correlated with energy transfer and subunit dissociation. The complex polarisation of fluorescence spectra indicated various chromophores in single protein.[362] By studying the life-times and quantum yields it was possible to calculate the number and distribution of fluorescent species in each protein.[363]

An excellent example of the way in which the various fluorescence techniques may be used complementarily is the work of Malcolm and Radda on the inactivation of glutamate dehydrogenase by iodoacetamidosalicylic acid (isa)[364] (isa is an irreversible active-site-directed inhibitor). Progressive modification causes increasing quenching of protein fluorescence together with an increase in isa fluorescence, showing that a Trp is close to the active site. The modified enzyme still binds NADH as indicated by enhancement of NADH fluorescence. The fluorescence of the isa-GDH-NADH complex is further enhanced by glutamate, as in the native enzyme.

[357] R. H. McKay, *Arch. Biochem. Biophys.*, 1970, **135**, 218.
[358] R. W. Cowgill, *Biochim. Biophys. Acta*, 1970, **200**, 18.
[359] R. W. Cowgill and N. K. Lang, *Biochim. Biophys. Acta*, 1970, **214**, 228.
[360] O. Pongs, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 431.
[361] O. Pongs, *Biochemistry*, 1970, **9**, 2316.
[362] F. W. J. Teale and R. E. Dale, *Biochem. J.*, 1970, **116**, 161.
[363] R. E. Dale and F. W. J. Teale, *Photochem. Photobiol.*, 1970, **12**, 99.
[364] A. D. B. Malcolm and G. K. Radda, *European J. Biochem.*, 1970, **15**, 555.

However, modification with isa removes the sensitivity to GTP inhibition and this is reflected in a lack of response of ANS fluorescence to the addition of NADH and GTP.

6 Spin Labels

contributed by N. C. Price

Since last year's Report,[365] a detailed review of the chemistry and physics of spin labels has been published.[366] Spin labels have continued to be used to examine a wide variety of biochemical systems, including proteins, enzymes, transfer RNA, phospholipids, and membranes. As in last year's Report, it is convenient to discuss the information gained from spin label studies on proteins under three main headings.

A. Measurements of Correlation Time.—The shape of the e.s.r. spectrum of a spin label depends on the rotational mobility of the label. Kivelson [367] and McConnell *et al.*[368] have shown how the rotational correlation time (τ_c) of the label can be derived from an e.s.r. spectrum.

Paton and Kaiser used a spin-labelled derivative of *m*-nitrobenzoic acid to detect the formation of a non-covalent 'Michaelis' complex between a model enzyme (cyclohepta-amylose, CA) and its substrate, S.[369] CA is known to catalyse the hydrolysis of phosphate [370] and carboxylic [371] esters *via* the intermediate formation of inclusion complexes. Kinetic studies have shown that this catalysis occurs by the scheme outlined below, which is analogous to that demonstrated for catalysis by α-chymotrypsin.[372]

$$S + CA \underset{K_d}{\rightleftharpoons} S\cdot CA \xrightarrow{k_2} S^1\cdot CA + P_1 \xrightarrow{k_3} CA + P_2$$

A solution of the spin-labelled substrate and CA incubated at pH 5.75 gave rise to an e.s.r. spectrum corresponding to a partially immobilised label whose τ_c (3.34×10^{-10} s) was approximately 10 times that of the free label. No production of *m*-nitrophenolate ion (P_1 in the above scheme) occurred during this experiment. The acyl-CA derivative $S^1\cdot CA$ showed a more immobilised type of e.s.r. spectrum ($\tau_c = 5.04 \times 10^{-10}$ s). Deacylation of $S^1\cdot CA$ could be followed by the change in the e.s.r. spectrum on raising the pH to 9.6.

A number of papers have dealt with the study of ligand-induced structural changes in enzymes and proteins. Timofeev and co-workers reacted

[365] N. C. Price, in 'Amino-acids, Peptides, and Proteins', ed. G. T. Young (Specialist Periodical Report) The Chemical Society, London, 1970, Vol. 2, p. 127.
[366] H. M. McConnell and B. G. McFarland, *Quart. Rev. Biophys.*, 1970, **3**, 91.
[367] D. Kivelson, *J. Chem. Phys.*, 1960, **33**, 1094.
[368] T. J. Stone, T. Buckman, P. L. Nordio, and H. M. McConnell, *Proc. Nat. Acad. Sci. U.S.A.*, 1965, **54**, 1010.
[369] R. M. Paton and E. T. Kaiser, *J. Amer. Chem. Soc.*, 1970, **92**, 4723.
[370] N. Hennrich and F. Cramer, *J. Amer. Chem. Soc.*, 1965, **87**, 1121.
[371] M. L. Bender, R. L. Van Etten, G. A. Clowes, and J. F. Sebastian, *J. Amer. Chem. Soc.*, 1966, **88**, 2318.
[372] M. L. Bender and F. J. Kézdy, *Ann. Rev. Biochem.*, 1965, **34**, 49.

aspartate aminotransferase from pig heart with iodoacetamide and maleimide spin-label derivatives.[373] The labelled enzyme (with $\leqslant 4$ labels incorporated per dimer) was 85—90% active and had the same absorption and c.d. spectra as the native enzyme. Addition of substrates or pseudo-substrates (glutamate + 2-oxoglutarate, 2-oxoglutarate, glutarate, *β-erythro*-oxyaspartate, or alanine) caused pronounced changes in the e.s.r. spectrum of the labelled enzyme. These changes indicated an increase in the mobility of the label: τ_c was 8.0×10^{-10} and 5.1×10^{-10} s in the absence and presence of ligands respectively. Since the labels were not located at the active site of the enzyme (they are estimated to be $\geqslant 15$ Å from the pyridoxal-5′-phosphate cofactor [374]), this work demonstrates the occurrence of substrate-induced changes in the enzyme, and provides evidence for the hypothesis of 'induced fit' of enzyme and substrate.[375]

Buckman has used the spin-labelling technique to study the structural changes in aspartate transcarbamylase (ATCase) from *E. coli* concomitant with binding of the inhibitor, cytidine-5′-triphosphate (CTP).[376] This enzyme contains two types of subunits. One type is purely catalytic and is not subject to CTP inhibition; the other type has a strong affinity for CTP but no catalytic activity. Using a spin-label derivative of bromoacetamide, the catalytic subunits could be labelled with no change in the enzyme activity, CTP inhibition, or overall physical properties of the enzyme. Addition of substrates had no effect on the e.s.r. spectrum of the labelled enzyme; addition of CTP, however, led to pronounced broadening of the e.s.r. spectrum. This CTP-induced change could be reversed by addition of substrates. The changes in the e.s.r. spectrum could be taken as a measure of the function of state of the labelled enzyme (*i.e.* as a measure of the extent of the CTP-induced conformational change). Comparison of this function of state with the binding data for CTP indicated that the behaviour of ATCase did not correspond with the predictions of either the 'concerted' [377] or 'sequential' [378] models for allosteric behaviour.

More studies on the ligand-induced conformational changes in haemoglobin (Hb) have been reported.[379, 380] Those haemoglobins (Hb A, Hb F, and Hb Zürich) which show fully or largely co-operative ligand binding, exhibit a lack of isosbesticity in the e.s.r. spectra of the labelled Hb as a function of CO saturation. This is taken to be consistent with the conclusion that the co-operative ligand binding arises from the occurrence of intermediate structures in the conversion of unligated Hb to fully ligated

[373] V. P. Timofeev, O. L. Polianovsky, V. M. Volkenstein, and G. I. Lichtenstein, *Biochim. Biophys. Acta*, 1970, **220**, 357.
[374] V. P. Timofeev, O. L. Polianovsky, M. V. Volkenstein, O. V. Preobrazhenskaya, U. V. Kochanov, and G. I. Lichtenstein, *Molek. Biol.* (*S.S.S.R.*), (in press).
[375] D. E. Koshland, jun., *Cold Spr. Hab. Symp. Quant. Biol.*, 1963, **28**, 473.
[376] T. Buckman, *Biochemistry*, 1970, **9**, 3255.
[377] J. Monod, J. Wyman, and J.-P. Changeux, *J. Mol. Biol.*, 1965, **12**, 88.
[378] D. E. Koshland, jun., G. Némethy, and D. Filmer, *Biochemistry*, 1966, **5**, 365.
[379] C. Ho, J. J. Baldassare, and S. Charache, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 722.
[380] J. J. Baldassare, S. Charache, R. T. Jones, and C. Ho, *Biochemistry*, 1970, **9**, 4707.

Hb (the 'sequential' model).[378] Those haemoglobins (Hb Yakima, Hb J Capetown, and Hb Chesapeake), which show intermediate or non-co-operative ligand binding, exhibit isosbesticity in the e.s.r. spectra as a function of CO saturation, suggesting that a two-state model is more appropriate in these cases. It is to be noted that the latter Hb's have amino-acid substitutions (at position $\alpha 92$) in the $\alpha_1\beta_2$ contact region. The importance of this contact region in transmitting structural effects between the subunits of Hb has been emphasised previously.[381]

Interpretation of the spin-labelling experiments carried out on Hb has been complicated by the finding that introduction of the bulky spin label (at the $\beta 93$ cysteine residues) causes severe perturbations of the Hb structure.[382] The *X*-ray diffraction data show that the largest changes are in the environment of the haem groups of both the α and β chains, but that there are also significant differences in the $\alpha_1\beta_2$ contact region. It has been claimed that these structural perturbations caused by introduction of the spin label do not drastically affect the ligand-binding characteristics of Hb.[380] However, the value of the Hill coefficient is reduced from 2.9 to 2.3 by the introduction of the spin label [383, 384] and this latter value is not much larger than that for Hb J Capetown (2.2), for which the two-state model is considered appropriate.[380]

Clearly it is essential to consider the possibility of structural perturbations in macromolecules caused by the introduction of spin labels. Thus in the studies on aspartate aminotransferase [373] and ATCase,[376] it was shown that the spin-labelled enzyme possessed substantially the same activity as the native enzyme, suggesting that the labelling reaction has not led to any significant structural changes at the active site.

Spin-labelling experiments have been carried out on the actin and myosin components of muscle. A maleimide spin label (incorporated to the extent of one mole per mole of actin) was found to be an excellent probe of actin–actin interactions.[385] Polymerisation of the labelled actin (which occurred as readily as in the native protein) resulted in a pronounced immobilisation of the label. No change in the e.s.r. spectrum was observed on adding myosin to the labelled actin. Reaction of myosin with a spin-label derivative of iodoacetamide led to labelling at two types of sites; one strongly immobilised (S_1), the other weakly immobilised (S_2).[386, 387] Addition of ADP, ATP, or pyrophosphate to the labelled myosin caused a significant decrease in the proportion of S_1 label; these changes were attributed to localised deformations in the myosin, leading to an opening up of the

[381] M. F. Perutz and H. Lehmann, *Nature*, 1968, **219**, 902.
[382] J. K. Moffatt, *J. Mol. Biol.*, (in press).
[383] S. Ogawa and H. M. McConnell, *Proc. Nat. Acad. Sci. U.S.A.*, 1957, **58**, 19.
[384] S. Ogawa, H. M. McConnell, and A. Horwitz, *Proc. Nat. Acad. Sci. U.S.A.*, 1968 **61**, 401.
[385] D. B. Stone, S. C. Prevost, and J. Botts, *Biochemistry*, 1970, **9**, 3937.
[386] D. B. Stone, *Arch. Biochem. Biophys.*, 1970, **141**, 378.
[387] J. C. Seidel, M. Chopek, and J. Gergely, *Biochemistry*, 1970, **9**, 3265.

structure around the label. These structural changes must be relatively localised, since no gross changes in the o.r.d. of the myosin can be detected on addition of ATP.[388]

Schneider and Smith proposed the use of a spin-labelling procedure to establish whether the structural integrity of a membrane protein is retained during isolation.[389] Intact erythrocyte ghost cells were reacted with a spin-label derivative of maleimide, and the lipid component was removed in a number of different ways. After removal of the lipid-solubilising reagents by dialysis, only some of these isolated labelled proteins showed e.s.r. spectra similar to that of the labelled intact ghosts; these samples also retained enzymic activity. This technique could provide useful information on the structural integrity of proteins isolated from membranes, before more detailed studies on these proteins are undertaken.

Dodd and co-workers examined the e.s.r. spectra of two spin-label probes in a variety of solvents and established a linear correlation between the hyperfine coupling constant and the polarity ('Z-value') of the solvent.[390] This correlation was then used to estimate the polarity of various phospholipid systems. Measurements of the rotational mobility of the labels were also of use in locating the spin probes.

B. Interaction of the Unpaired Electron with Neighbouring Paramagnetic Ions or Nuclei.—The study of the active sites of dehydrogenases using a spin-labelled analogue of β-nicotinamide adenine dinucleotide (NAD^+),[391–393] described in last year's Report, has now been extended to include malate dehydrogenase (MDH).[394] The spin-label analogue (ADP-R•) was a competitive inhibitor of the enzyme; the binary complex MDH $(ADP\text{-}R^\bullet)_2$ had a dissociation constant of 0.2mmol l^{-1}. The unpaired electron of ADP-R• bound to MDH enhanced the relaxation rate $(1/T_1)$ of water protons 50 times as efficiently as did free ADP-R•. This factor decreased as the occupancy of the ADP-R• sites increased from 0 to 2, presumably due to an interaction between these sites. Oxaloacetate formed a ternary complex with $MDH(ADP\text{-}R^\bullet)_2$, and lowered the enhancement factor for water protons by 70%, due to a shielding of the unpaired electron from the water. By contrast, neither L- nor D-malate changed the enhancement factor on formation of the respective ternary complexes. The unpaired electron in ADP-R• increased the relaxation rate $(1/T_2)$ of the methylene protons of bound oxaloacetate. This enabled the rate constant of the binding of oxaloacetate to the $MDH(ADP\text{-}R^\bullet)_2$ complex ($\geqslant 3.7 \times$

[388] W. B. Gratzer and S. Lowy, *J. Biol. Chem.*, 1969, **244**, 22.
[389] H. Schneider and I. C. P. Smith, *Biochim. Biophys. Acta*, 1970, **219**, 73.
[390] G. H. Dodd, M. D. Barrett, and L. Rayner, *F.E.B.S. Letters*, 1970, **8**, 286.
[391] H. Weiner, *Biochemistry*, 1969, **8**, 526.
[392] A. S. Mildvan and H. Weiner, *Biochemistry*, 1969, **8**, 552.
[393] A. S. Mildvan and H. Weiner, *J. Biol. Chem.*, 1969, **244**, 2465.
[394] A. S. Mildvan, L. Waber, J. J. Villafranca, and H. Weiner, Wenner-Gren Symposium on Structure and Function of Oxidation Reduction Enzymes, Stockholm, 1970, Abstracts, p. 94.

10^6 l mol^{-1} s^{-1}) and the distance between the methylene protons and the unpaired electron of ADP-R$^{\bullet}$ in the ternary complex (3.6 ± 1.2 Å) to be determined. The importance of this distance lies in the fact that the unpaired electron in ADP-R$^{\bullet}$ is in a position sterically equivalent to the bond between the ring nitrogen of the pyridine and the C-1 atom of the ribose of NAD^+.[392] The ADP-R$^{\bullet}$ did not cause any enhancement of the relaxation rate of the methylene protons of bound L-malate; either due to slow exchange of substrate in this ternary complex or to binding of L-malate at a site remote ($\geqslant$ 6.2 Å) from the bound ADP-R$^{\bullet}$. These possibilities could be distinguished by changing the temperature.

Roberts and co-workers have investigated the interaction of ribonuclease with a spin-labelled phosphate ester, which should behave as a competitive inhibitor of the enzyme.[395] Addition of the spin label (2mmol l^{-1}) to ribonuclease (6.5mmol l^{-1}) in D_2O caused broadening of the n.m.r. lines due to residues His-12 and His-119 in the enzyme; the changes in the His-12 line were more pronounced. No changes in the n.m.r. lines due to the other aromatic residues in the enzyme were detected. The conclusion that the spin label was found relatively near His-12 and His-119 was consistent with previous *X*-ray diffraction [396] and n.m.r.[397] data which suggested that the active site of ribonuclease is in a cleft with residues His-12 and His-119 close together.

Leigh has developed a theory to describe the line shape of an e.s.r. signal influenced by dipolar coupling to a second spin (*e.g.* Mn^{2+}).[398] The theory allows the calculation of the spin–spin distances. The experimental results on the interaction of Mn^{2+} with spin-labelled creatine kinase [399] were compared with the theoretical curves.

C. Interaction of the Unpaired Electrons in a Biradical.—The interaction between the unpaired electrons in a biradical causes changes in the e.s.r. spectrum. Ferruti *et al.* have described the synthesis of a number of mono-, di-, and poly-nitroxides and examined their e.s.r. spectra as a function of temperature.[400] The various radicals were classified in terms of the conformation and mobility of the radical backbone, and the exchange between the electrons. Ferruti's interpretation of the changes in the height (and hence width) of the lines of the e.s.r. spectra in terms of the exchange between the spins [401] has been criticised by Luckhurst and Pedulli.[402] These authors pointed out that Ferruti's results could be explained either by

[395] G. C. K. Roberts, J. Hannah, and O. Jardetzky, *Science*, 1969, **165**, 505.
[396] G. Kartha, J. Bello, and D. Harkler, *Nature*, 1967, **213**, 862.
[397] D. H. Meadows and O. Jardetzky, *Proc. Nat. Acad. Sci. U.S.A.*, 1968, **61**, 406.
[398] J. S. Leigh jun., *J. Chem. Phys.*, 1970, **52**, 2608.
[399] J. S. Taylor, J. S. Leigh jun., and M. Cohn, *Proc. Nat. Acad. Sci. U.S.A.*, 1969, **64**, 219.
[400] P. Ferruti, D. Gill, M. P. Klein, H. H. Wang, G. Entine, and M. Calvin, *J. Amer. Chem. Soc.*, 1970, **92**, 3704.
[401] P. Ferruti, D. Gill, M. P. Klein, and M. Calvin, *J. Amer. Chem. Soc.*, 1969, **91**, 7765.
[402] G. R. Luckhurst and G. F. Pedulli, *J. Amer. Chem. Soc.*, 1970, **92**, 4738.

changes in the spin exchange or by changes in the spectral density, and suggested a way of distinguishing between these possibilities. The interpretation of the biradical experiments with nerve membranes [403] may also be in error.

Among other papers on spin labels which appeared in 1970 were studies on transfer RNA's,[404, 405] phospholipids,[406–408] lipid–protein interactions,[409–410] membrane structure,[411–415] and the interaction of anaesthetics with membranes.[416]

Spin labels have thus continued to provide valuable information on the structure and structural changes (often very localised) of a variety of enzymes and proteins. However, the detailed studies of the perturbations of the structure of haemoglobin caused by introduction of a spin label [382] sound a necessary note of caution in the interpretation of some of these experiments.

7 Dissociation and Association of Proteins

contributed by P. H. Lloyd

A very valuable review of the quaternary structure of proteins has appeared this year.[417] The authors reviewed the methods for dissociation and hybridisation of proteins and list 110 proteins whose subunit structures have been determined. The thermodynamic constants for the association of 14 proteins are given and the factors modifying subunit interactions are reviewed. Discussion of the structures of interface contacts had to be restricted to haemoglobin, a protein which is itself the subject of a review in the same volume.[418] Indeed, a number of papers have been published

[403] M. Calvin, H. H. Wang, G. Entine, D. Gill, P. Ferruti, M. A. Harpold, and M. P. Klein, *Proc. Nat. Acad. Sci. U.S.A.*, 1969, **63**, 1.
[404] P. Schofield, B. M. Hoffman, and A. Rich, *Biochemistry*, 1970, **9**, 2525.
[405] H, Hara, T. Horiuchi, M. Saneyoshi, and S. Nishmura, *Biochem.. Biophys Res. Comm.*, 1970, **38**, 305.
[406] J.-C. Hsia, H. Schneider, and I. C. P. Smith, *Biochim. Biophys. Acta*, 1970, **202**, 399.
[407] C. Huang, J. P. Charlton, C. I. Shyr, and T. E. Thompson, *Biochemistry*, 1970, **9**, 3422.
[408] G. L. Jendrasiak and R. Hayes, *Nature*, 1970, **225**, 278.
[409] R. Aneja and A. P. Davies, *Chem. Phys. Lipids*, 1970, **4**, 60.
[410] A. M. Gotto, H. Kon, and M. E. Birnbaumer, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **65**, 145.
[411] R. A. Long, F. Hruska, H. D. Gesser, J. C. Hsia, and R. Williams, *Biochem. Biophys., Res. Comm.*, 1970, **41**, 321.
[412] A. S. Waggoner, T. J. Kingzett, S. Rotschaefer, O. H. Griffith, and A. D. Keith, *Chem. Phys. Lipids*, 1969, **3**, 245.
[413] A. Keith, G. Bulfieid, and W. Snipes, *Biophys. J.*, 1970, **10**, 618.
[414] M. E. Tourtellotte, D. Branton, and A. Keith, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 909.
[415] S. Rottem, W. L. Hubbell, L. Hayflick, and H. M. McConnell, *Biochim. Biophys. Acta*, 1970, **219**, 104.
[416] W. L. Hubbell, J. C. Metcalfe, S. M. Metcalfe, and H. M. McConnell, *Biochim. Biophys. Acta*, 1970, **219**, 415.
[417] I. M. Klotz and N. R. Langerman, *Ann. Rev. Biochem.*, 1970, **39**, 25.
[418] E. Antonini and M. Brunori, *Ann. Rev. Biochem.*, 1970, **39**, 977.

this year on haemoglobin [419–423] of which perhaps the most fascinating is that by Perutz [424] in which he explains how his *X*-ray diffraction studies have thrown light on the co-operative binding of oxygen, the Bohr effect, and the role of 2,3-diphosphoglycerate. The mechanism he proposes for the transition between the deoxy- and oxy-structures of haemoglobin fits neither the Monod [377] nor the Koshland [378] models of allosteric behaviour in proteins, but is a most satisfying compromise between them. His results also explain the difference in the dissociation constants for the two forms separating into dimers [420] (see also Table 3).

The subunit structures of many more proteins have been determined by measurements of molecular weights in dissociating and non-dissociating solvents, and by other means. These results are summarised in Table 2 (see p. 216). 6M Guanidine hydrochloride continues to be the most popular dissociating solvent (refs. 425—427 and others given in Table 2), but sodium dodecyl sulphate (SDS) is now used almost as much because of the increasing popularity of electrophoresis on SDS-polyacrylamide gels as a method of determining the molecular weights of subunits (refs. 428—430 and others in Table 2). 8M Urea is still popular also (refs. in Table 2). Again some proteins are dissociated by smaller concentrations [431–441] and others require higher concentrations for complete dissociation.[427, 442, 443] In most cases a sulphydryl reducing agent is also present, although this is

[419] R. Briehl, *J. Biol. Chem.*, 1970, **245**, 538.
[420] S. J. Edelstein, M. J. Rehmar, J. S. Olsen, and Q. H. Gibson, *J. Biol. Chem.*, 1970, **245**, 4372.
[421] C. M. Park, *J. Biol. Chem.*, 1970, **245**, 5390.
[422] A. C. Peacock, J. V. Pastewka, R. A. Reed, and A. T. Ness, *Biochemistry*, 1970, **9**, 2275.
[423] C. Ho, J. J. Baldassare, and S. Charache, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 722.
[424] M. F. Perutz, *Nature*, 1970, **288**, 726.
[425] S. L. Marchesi, E. Steers, V. T. Marchesi, and T. W. Tillack, *Biochemistry*, 1970, **9**, 50.
[426] F. K. Stevenson and P. W. Kent, *Biochem. J.*, 1970, **116**, 791.
[427] F. C. Stevens, D. M. Brown, and E. L. Smith, *Arch. Biochem. Biophys.*, 1970, **136**, 413.
[428] B. Paterson and R. C. Strohman, *Biochemistry*, 1970, **9**, 4094.
[429] A. P. Fletcher, A. Neuberger, and W. A. Ratcliffe, *Biochem. J.*, 1970, **120**, 425.
[430] R. Koenig, H. Stegemann, H. Francksen, and H. L. Paul, *Biochim. Biophys. Acta*, 1970, **207**, 184.
[431] J. Gazith, S. Himmelfarb, and W. Harrington, *J. Biol. Chem.*, 1970, **245**, 15.
[432] A. Raz, T. Shiratori, and D. S. Goodman, *J. Biol. Chem.*, 1970, **245**, 1903.
[433] V. M. Wickson and R. E. Huber, *Biochim. Biophys. Acta*, 1970, **207**, 150.
[434] W. D. Riley and E. E. Snell, *Biochemistry*, 1970, **9**, 1485.
[435] R. M. Metrione, Y. Okuda, and G. F. Fairclough, *Biochemistry*, 1970, **9**, 2427.
[436] T. Kakimoto, J. Kato, T. Shibatani, N. Nishimura, and I. Chibata, *J. Biol. Chem.*, 1970, **245**, 3369.
[437] B. H. Frank, A. H. Pekar, A. J. Veros, and P. P. K. Ho, *J. Biol. Chem.*, 1970, **245**, 3716.
[438] S. C. K. Wong, D. C. Hall, and J. Josse, *J. Biol. Chem.*, 1970, **245**, 4335.
[439] J. M. Brewer, L. Ljungdahl, T. E. Spencer, and S. H Neece, *J. Biol. Chem.*, 1970, **245**, 4798.
[440] K. G. Mann and C. S. Vestling, *Biochemistry*, 1970, **9**, 3020.
[441] K. B. Jacobron and P. Pfuderer, *J. Biol. Chem.*, 1970, **245**, 3938.
[442] V. L. Seery, E. H. Fischer, and D. C. Teller, *Biochemistry*, 1970, **9**, 3591.
[443] M. Singh, G. C. Brooks, and P. A. Srere, *J. Biol. Chem.*, 1970, **245**, 4636.

not always necessary.[444] Such measurements are not without their uncertainties, however. Aune and Timasheff[445] found a decrease in the apparent molecular weight of glyceraldehyde-3-phosphate dehydrogenase when it was transferred from dilute salt solution to 1.3M potassium phosphate, but they concluded that preferential hydration was responsible for the observation and not dissociation of the protein. The role of monomers in the inactivation of a tetrameric enzyme has been shown.[446]

Although sedimentation velocity experiments are probably in many cases the initial indication that a protein has dissociated, few workers now rely on them for quantitative measurements[447–453] although they are quite suitable for studying the action of effectors on the aggregation of an 'allosteric' enzyme.[454] Of the methods available for the determination of molecular weights the high-speed sedimentation equilibrium method of Yphantis[455] remains the most popular. One of its major disadvantages has been overcome by computational techniques[442, 456] which yield the concentration at the top meniscus when this is small but not insignificant. However, the assumptions implicit in these computations should be carefully evaluated. In one case[457] the computer programme used is only published in a PhD thesis. A valuable review of computational techniques for use with sedimentation equilibrium results has appeared.[458] An example has been published[439] in which the molecular weight of a protein determined by the high-speed technique[455] appeared to decrease as the speed of the rotor was increased. This effect is to be expected if the sample contained a significant amount of a component of small molecular weight. Absorption optical systems combined with electronic scanners are beginning to be used for sedimentation equilibrium experiments[459–461] with

[444] S. Kashiwamata, Y. Kotake, and D. M. Greenberg, *Biochim. Biophys. Acta*, 1970, **212**, 501.
[445] K. C. Aune and S. N. Timasheff, *Biochemistry*, 1970, **9**, 1481.
[446] J. H. Griffin and R. S. Criddle, *Biochemistry*, 1970, **9**, 1195.
[447] H. D. Ellerton, D. E. Carpenter, and K. E. van Holde, *Biochemistry*, 1970, **9**, 2225.
[448] A. B. Schneider and H. Edelhoch, *J. Biol. Chem.*, 1970, **245**, 885.
[449] E. Kvamme, B. Tveit, and G. Svenneby, *J. Biol. Chem.*, 1970, **245**, 1871.
[450] I. T. Schulze, C. J. Lusty, and S. Ratner, *J. Biol. Chem.*, 1970, **245**, 4534.
[451] H. A. De Phillips, jun., K. W. Nickerson, and K. E. van Holde, *J. Mol. Biol.*, 1970, **50**, 471.
[452] E. Fujimori and J. Pecci, *Biochim. Biophys. Acta*, 1970, **207**, 259.
[453] S. Shifrin, S. W. Luborsky, and B. J. Grochowski, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 494.
[454] P. M. Anderson and S. V. Martin, *Biochemistry*, 1970, **9**, 171.
[455] D. A. Yphantis, *Biochemistry*, 1964, **3**, 297.
[456] D. C. Teller, T. A. Horbett, E. G. Richards, and H. K. Schachman, *Ann. N.Y. Acad. Sci.*, 1969, **164**, 66.
[457] J. K. Zimmerman, G. H. Barlow, and I. M. Klotz, *Arch. Biochem. Biophys.*, 1970, **138**, 101.
[458] 'Advances in Ultracentrifugal Analysis,' ed. D. A. Yphantis, *Ann. N.Y. Acad. Sci.*, 1969, **164**, Art. 1.
[459] G. M. Hathaway and I. P. Crawford, *Biochemistry*, 1970, **9**, 1801.
[460] E. Reisler, J. Pouget, and H. Eisenberg, *Biochemistry*, 1970, **9**, 3095.
[461] D. E. Wampler, M. Takahashi, and E. W. Westhead, *Biochemistry*, 1970, **9**, 4210.

certain advantages over the interferometer which include sensitivity and absolute concentration determination at all points. The low-speed sedimentation equilibrium technique [431, 434, 437, 438, 457, 459, 460, 462–468] is still used quite extensively. However, the authors have not in all cases indicated how they determined the absolute concentrations from the interference pictures. Some [435, 468] used schlieren optics, others used the scanner,[460, 466] and, in spite of its limitations in this context, the principle of conservation of mass.[464] Although simple in application, the achromatic fringe [469] is still seldom used.[434, 438, 470] The use of very short columns (0.8 mm) has now been abandoned by all but a few workers.[471]

Despite its obvious advantages for this type of work osmometry has been little used.[472–475] However, the development [476] of a method of measuring osmotic pressures with a single bead of Sephadex may revive its use because such small volumes of solution are needed, the measurements are very quick and the apparatus relatively simple. The use of gel filtration [477–484] and electrophoresis on polyacrylamide gels containing sodium dodecyl sulphate (SDS) have become less empirical as a result of some work by Tanford *et al.*[485, 486] They have shown by direct binding studies that 13 different proteins all bind the same amount of dodecyl sulphate when the free concentration of the detergent is greater than 5×10^{-4} mol l^{-1}. However, the figures they derive (0.4 and 1.4 g/g) [487] differ from that found by de Groot *et al.*[488] for α-crystallin (0.19 ± 0.09 g/g) in 2% dodecyl sulphate.

[462] J. E. Godfrey and W. F. Harrington, *Biochemistry*, 1970, **9**, 886.
[463] L. C. Gershman and P. Dreizen, *Biochemistry*, 1970, **9**, 1677.
[464] K. G. Mann, F. J. Castellino, and P. A. Hargrave, *Biochemistry*, 1970, **9**, 4002.
[465] D. C. Teller, *Biochemistry*, 1970, **9**, 4201.
[466] K. S. King and C. Frieden, *J. Biol. Chem.*, 1970, **245**, 4391.
[467] R. C. Deonier and J. W. Williams, *Biochemistry*, 1970, **9**, 4260.
[468] D. Snary, A. Allen, and R. H. Pain, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 844.
[469] E. G. Richards and H. K. Schachman, *J. Phys. Chem.*, 1959, **63**, 1578.
[470] P. H. Lloyd and A. R. Peacocke, *Biochem. J.*, 1970, **118**, 467.
[471] T. Sugiyama and T. Akazawa, *Biochemistry*, 1970, **9**, 4499.
[472] H. Stauffer, S. Srinivasan, and M. A. Lauffer, *Biochemistry*, 1970, **9**, 193.
[473] N. Lee, J. W. Patrick, and N. B. Barnes, *J. Biol. Chem.*, 1970, **245**, 1357.
[474] N. Feliss and M. Martinez-Carrion, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 932.
[475] H. Papkoff and J. Gan, *Arch. Biochem. Biophys.*, 1970, **136**, 522.
[476] A. G. Ogston and J. D. Wells, *Biochem. J.*, 1970, **119**, 67.
[477] G. Philips, J. Moran, and R. W. Colman, *Biochemistry*, 1970, **9**, 2212.
[478] E. Y. C. Lee, J. H. Carter, L. D. Nielson, and E. H. Fischer, *Biochemistry*, 1970, **9**, 2347.
[479] H. P. Schnebli, A. E. Valter, and A. Abrams, *J. Biol. Chem.*, 1970, **245**, 1122.
[480] D. Bannerjee, J. E. Sanders, and J. R. Sokatch, *J. Biol. Chem.*, 1970, **245**, 1828.
[481] T.-H. Liao and J. G. Pierce, *J. Biol. Chem.*, 1970, **245**, 3275.
[482] S. Kaufman and D. B. Fisher, *J. Biol. Chem.*, 1970, **245**, 4745.
[483] H. Papkoff and M. Ekbald, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 614.
[484] A. Marzotto and L. Galzigna, *Arch. Biochem. Biophys.*, 1970, **137**, 373.
[485] J. A. Reynolds and C. Tanford, *J. Biol. Chem.*, 1970, **245**, 5161.
[486] W. W. Fish, J. A. Reynolds, and C. Tanford, *J. Biol. Chem.*, 1970, **245**, 5166.
[487] J. A. Reynolds and C. Tanford, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 1002.
[488] K. De Groot, H. J. Hoenders, J. J. T. Gerding, and H. Bloemendal, *Biochim. Biophys. Acta*, 1970, **207**, 202.

Many authors still derive molecular weights by combining sedimentation and diffusion coefficients in the Svedberg equation.[438, 468, 473, 487–494]

Few papers this year report association constants for proteins aggregating with themselves and the results of these are summarised in Table 3 (see p. 218). One paper reports measurements of the association constant for the association of two different proteins,[495] carboxypeptidase A and Fraction II of procarboxypeptidase A (2.9×10^4 l mol^{-1}). The self-association of procarboxypeptidase A has also been studied,[465] as has the stoicheiometry of the interaction of haptoglobin and haemoglobin [422] and that of retinol-binding protein and prealbumin.[432] The enthalpy change for the aggregation of tobacco mosaic virus protein has been determined [472] as 25—30 kcal mol^{-1} of bonds. Agents whose effects on the state of aggregation of proteins have been determined include allosteric effectors,[454, 461, 496] pH, ionic strength, concentration, and temperature,[472] calcium and magnesium ions,[497] deuterium oxide,[492] oxygen,[451] and toluene.[498]

Another field of interest is the kinetics of recombination of subunits or refolding of tertiary structure. Collagen is the subject of a series of five papers by Harrington and others on the refolding kinetics of long-chain polymers.[499–501] The rates of recombination of the subunits of thyroglobin [448] and glycogen phosphorylase [502] have been studied by light-scattering and of haemoglobin by stop-flow and flash photolysis.[420] The combination of haemoglobin with haptoglobin has been followed by fluorescence quenching in a stop-flow apparatus.[503]

Some theoretical papers have appeared which are of interest to workers in this field. Steiner [504] has extended his work on colligative data to include any number of components which may or may not be associating with themselves, and non-ideal systems of two components.[505] A method of determining association constants from measurements of weight-average molecular weights as a function of concentration for both discrete and indefinite association has been published by Chun and Kim.[506] Light

[489] P. H. Strausbauch and E. H. Fischer, *Biochemistry*, 1970, **9**, 226.
[490] R. T. Kuczenski and C. H. Suelter, *Biochemistry*, 1970, **9**, 2043.
[491] J. L. Denburg and W. D. McElroy, *Biochemistry*, 1970, **9**, 4619.
[492] R. F. Henderson and T. R. Henderson, *J. Biol. Chem.*, 1970, **245**, 3733.
[493] L. D. Kohn, W. A. Warren, and W. R. Carroll, *J. Biol. Chem.*, 1970, **245**, 3821.
[494] N. R. Gould and S. L. Scheinberg, *Arch. Biochem. Biophys.*, 1970, **137**, 1.
[495] W. D. Behnke, D. C. Teller, R. D. Wade, and H. Neurath, *Biochemistry*, 1970, **9**, 4189.
[496] T. R. Leary and G. Kohlhaw, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 494.
[497] R. C. Weisenberg and S. N. Timacheff, *Biochemistry*, 1970, **9**, 4110.
[498] E. Reisler and H. Eisenberg, *Biopolymers*, 1970, **9**, 877.
[499] W. F. Harrington and N. V. Rao, *Biochemistry*, 1970, **9**, 3714.
[500] W. F. Harrington and G. M. Karr, *Biochemistry*, 1970, **9**, 3725.
[501] P. V. Hauschka and W. F. Harrington, *Biochemistry*, 1970, **9**, 3734; 3745; 3754.
[502] D. L. De Vincenzi and J. L. Hendrick, *Biochemistry*, 1970, **9**, 2048.
[503] A. Alfson, E. Chiancone, J. Wyman, and E. Antonini, *Biochim. Biophys. Acta*, 1970, **200**, 76.
[504] R. F. Steiner, *Biochemistry*, 1970, **9**, 1375.
[505] R. F. Steiner, *Biochemistry*, 1970, **9**, 4268.
[506] P. W. Chun and S. S. Kim, *Biochemistry*, 1970, **9**, 1957.

scattering and low-angle *X*-ray scattering may be used to determine the molecular-weight distributions and particle scattering factors of the various species in an aggregating system.[507] Other authors have used theoretical analysis in attempts to explain experimental results.[465, 467, 470] A unified theory of gel electrophoresis and gel filtration has been published.[508]

Several new techniques have appeared this year which may also be of interest to workers in this field. A way of determining the number of subunits in a protein has been described.[509] All the subunits are modified chemically, mixed with native subunits, and the protein reconstituted. From the number of different forms so produced the number of subunits can be deduced.[510, 511] Another technique uses a cross-linking reagent and SDS-gel electrophoresis in a similar way.[512] It is frequently of interest to know if the subunits of a multi-subunit protein are biologically active in their native form when not combined into the normal complex. Removal of the denaturing conditions frequently results in re-association to the aggregated form. Chan [513] describes a technique of combining the subunits with an insoluble matrix to prevent this re-association. A new quick method for isoelectric focusing in a pH gradient makes use of the variation with temperature of the pH of a tris buffer to achieve in 15 min what by other means takes many hours.[514]

When using the interferometer for analysis of sedimentation–equilibrium experiments the problem arises of determining the concentration at any one level from which that at all others may be determined by means of the interferograms. A classical method has been to take a large number of photographs of only a small part of the cell during the approach to equilibrium and to count the number of fringes which pass a reference line during this period. An improved version of this technique has been described by Bethune.[515] The normal camera is replaced by a time-lapse ciné camera. This takes a large number of full-frame photographs of the interferograms during the approach to equilibrium, and when these are projected at normal speed the movement of the fringes during this period can be seen easily. The final frames can be used in the normal way and the earlier ones used to deduce the sedimentation and diffusion coefficients without the need for ancillary measurements (apart presumably from the measurement of the partial specific volume).

[507] M. E. Magar, *Biopolymers*, 1970, **9**, 307.
[508] D. Rodband and A. Chrambach, *Proc. Nat. Acad. Sci.*, 1970, **65**, 970.
[509] E. A. Meighen and H. K. Schachman, *Biochemistry*, 1970, **9**, 1163.
[510] E. A. Meighen and H. K. Schachman, *Biochemistry*, 1970, **9**, 1177.
[511] E. A. Meighen, V. Pigiet, and H. K. Schachman, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **65**, 234.
[512] G. E. Davies and G. R. Stark, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 651.
[513] W. W.-C. Chan, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 1198.
[514] S. J. Luner and A. Kolin, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **66**, 898.
[515] J. L. Bethune, *Biochemistry*, 1970, **9**, 2737; R. T. Simpson and J. L. Bethune, *ibid.*, 2745.

Ogston and Wells [476] have improved their osmometer using a single Sephadex bead to make it a quick and precise instrument. A new method of measuring diffusion coefficients in homogeneous solutions has been described.[516] This uses the fluctuations of the intensity of laser light scattered by the solution of a macromolecule. This technique uses very small volumes of solution (as little as 15 μl has been used) and quite dilute solutions (0.08 mg per ml for haemocyanin) to determine diffusion coefficients to $\pm$ 1% in under 3 min (apart from the time taken to clarify the solution).[516] The basis for the technique and the technology come from the field of microwave radar and this shows what advances can be made when scientists of very different backgrounds and fields of specialisation co-operate (in this case physicists, radar technologists, and biochemists).

The structure of the M_4 isoenzyme of lactate dehydrogenase has been determined by *X*-ray diffraction to 2.8 Å resolution.[517] The four subunits can now be seen and the nature of the subunit contacts examined. The work on haemoglobin [424] has shown how important these subunit contacts can be to the properties of allosteric proteins.

Table 2

Protein	*Number of Subunits*	*Dissociating agent*	*Ref.*
Glyceraldehyde 3-phosphate dehydrogenase	4	Hybridisation	510
Lactate dehydrogenase	4	—	446
Histidine decarboxylase from *Lactobacillus* 30*a*	10[a]	2M-GuCl + 5,5′-DTBN	434
Pyruvate kinase from yeast	4	6M-GuCl	490
Glycogen phosphorylase	4	—	502
Haemocyanin from cancer magister	—	6M-GuCl	447
Amylo-1,6-glucosidase-oligo-1,4 → 1,4-glucan-transferase from yeast	2 or 3[a]	SDS	478
Dipeptidyl transferase	8[a]	8M urea, GuCl, pH 3	435
Aspartate-β-decarboxylase from *Pseudomonas dacunhae*	12	Recombination	518
	16	5M-GuCl	436
Malate dehydrogenase from rat	2[a]	6.25M urea at pH 3·2	440
L-Asparaginase from *E. coli*	4[b]	0.1% SDS	519
	4	7M urea, 5M-GuCl	437
L-Malic enzyme from *E. coli*	8	6M-GuCl	520
Enolase from yeast	2	6M-GuCl	464
Procarboxypeptidase A-S6 (bovine)	2[b]	—	521
Luciferase from firefly	2[a]	—	491
Rhodanase	2[a]	—	522

[516] R. Foord, E. Jakeman, C. J. Oliver, E. R. Pike, R. J. Blagrove, E. Wood, and A. R. Peacocke, *Nature*, 1970, **227**, 242.
[517] M. J. Adams, G. C. Ford, R. Koekoek, P. S. Lentz, jun., A. McPherson, jun., M. G. Rossmann, I. E. Smiley, R. W. Schevitz, and A. J. Wonacott, *Nature*, 1970, **227**, 1098.
[518] S. S. Tate and A. Meister, *Biochemistry*, 1970, **9**, 2626.
[519] R. C. Jackman and R. E. Handschumaker, *Biochemistry*, 1970, **9**, 3585.
[520] J. Spina, H. J. Bright, and J. Rosenbloom, *Biochemistry*, 1970, **9**, 3794.
[521] W.-D. Behnke, R. D. Wade, and H. Neurath, *Biochemistry*, 1970, **9**, 4179.
[522] P. Horowitz and J. Westley, *J. Biol. Chem.*, 1970, **245**, 986.

Table 2 (*cont.*)

Protein	*Number of Subunits*	*Dissociating agent*	*Ref.*
Seryl transfer ribonucleic acid synthetase from *E. coli*	2	6M-GuCl	523
Adenosine triphosphatase from *Streptococcus faecalis*	12[a]	6M-GuCl	479
L-Ribulokinase from *E. coli*	2	8M urea	473
Methylmalonate semialdehyde dehydrogenase from *Pseudomonas aeruginosa*	2	Urea, SDS	480
Thyroid-stimulating hormone from ox	2	1M propionic acid	481
Alcohol dehydrogenase from *Drosophila*	8	7.5M urea, SDS	441
Glyoxylic acid reductase from spinach	4	6M-GuCl, 8M urea	493
Glutamate dehydrogenase from rat	6–8	6M-GuCl	466
Citrate synthase from pig	2	7M-GuCl	443
Phenylalanine hydroxylase from rat	4	SDS	482
Formyltetrahydrofolate synthetase from *Clostridium thermoaceticum*	4	4M-GuCl	439
Turnip crinkle and tomato bushy stunt viruses	192[a]	SDS	524
α-Isopropylmalate synthase from *Salmonella typhimurium*	3[a]	Leucine	496
Ribulose diphosphate carboxylase	16–18	SDS	525
Chorionic gonadotropin from man	2[a]	8M urea	526
Fatty acid synthetase from pigeon	2	Low ionic strength	527
Glutamic aspartictransaminase	2	6M-GuCl	474
Glucose-6-phosphate dehydrogenase	2	—	528
Anthranilate synthase from *Bacillus subtilis*	2	—	529
α-Crystallin from ox	4	7M urea	530
Cell-stimulating hormone from ox	—	—	475
Two anti-A haemagglutinins from *Phaseolus lunatus*	8, 4	SDS, 8M urea	494
Succinyl coenzyme A synthetase	4	6M-GuCl, 8M urea PMB, SDS	531
Pyrophosphatase from *E. coli*	6	5M-GuCl	438
Cystathionine synthetase from rat	4[a]	SDS	444

[a] Subunits stated to be of two or more types.
[b] Subunits stated to be similar or identical.
SDS = sodium dodecyl sulphate.
PMB = *p*-chloro-mercuri-benzoate.
5,5′-DTBN = 5,5′-dithiobis-2-(nitrobenzoic acid).
GuCl = guanidine hydrochloride.
In addition to the dissociating agents listed, reducing agents such as 2-mercaptoethanol or dithiothreitol are usually (though not always) needed.

523 J. R. Katze and W. Konigsberg, *J. Biol. Chem.*, 1970, **245**, 923.
524 P. J. G. Butler, *J. Mol. Biol.*, 1970, **52**, 589.
525 A. C. Rutner, *Biochem. Biophys. Res. Conm.*, 1970, **39**, 923.
526 N. Swaminathan and O. P. Bahl, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 422.
527 S. Kumar, J. K. Dorsey, and J. W. Porter, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 825.
528 A. Yoshida and V. D. Hoagland, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 1167.
529 J. F. Kane and R. A. Jenson, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 328.
530 J. Delcour and J. Papaconstantinou, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 401.
531 C. Leitzmann, J.-Y. Wu, and P. D. Boyer, *Biochemistry*, 1970, **9**, 2338.

Table 3

Protein	*Association Reaction*	*Association constant*	*Ref.*
Myosin	M ⇌ D	$K_2 = 9.98$ dl g^{-1}	532
Glutamate dehydrogenase from ox	Infinite	$K = 2$ ml mg^{-1}	460
Haemoglobin from man, oxy-	D ⇌ T	$K = 5 \times 10^5$ l mol^{-1} (heme)	420
deoxy-	D ⇌ T	$K = 10^7$ l mol^{-1} (heme)	420
β-Lactoglobulin A subunits	M ⇌ D	$K_2 = 4.88 \times 10^4$ l mol^{-1}	457
B subunits	M ⇌ D	$K_2 = 14.2 \times 10^4$ l mol^{-1}	457
Muramidase	M ⇌ D	$K_2 = 0.347$ dl g^{-1}	
	Isodesmic	$k = 0.2$ dl g^{-1}	467
β-Chain monomers of tryptophan synthetase from *E. coli*	M ⇌ D	$K_2 = 10.4$ l g^{-1}	459
	M + D ⇌ Tr	$K_3 = 1.4$ l g^{-1}	
Carboxypeptidase A with Fraction II of procarboxypeptidase A	—	$K = 2.9 \times 10^4$ l mol^{-1}	495

M = monomer, D = dimer, Tr = trimer, T = tetramer.

[532] J. E. Godfrey and W. F. Harrington, *Biochemistry*, 1970, **9**, 894.

3
Peptide Synthesis

BY J. H. JONES

1 Introduction

The arrangement of this chapter follows essentially that established in the two previous volumes, except that cyclic peptides are covered only in Chapter 4 this year, and two new sections which seemed timely [on Purification of Synthetic Intermediates (Sections 2C and 6) and on Partial Synthesis (Section 2G)] have been introduced. In one important respect, however, this review differs from its predecessors: the syllabus for Volumes 1 and 2 was defined by *Current Chemical Papers*, but the great usefulness and convenience of this periodical to limited groups of the chemical community (such as peptide chemists) were insufficient to make its publication an economically viable operation beyond December 1969. A more laborious procedure has therefore been devised by the Reporter for gathering the harvest of references for the present volume. The gap left by the disappearance of *Current Chemical Papers* will no doubt have forced most practitioners of peptide synthesis to reconsider their techniques for keeping pace with current research in the area: a short account of the Reporter's approach may therefore be of some interest.

Computerised key-word searching of *Chemical Titles* proved to be both impractical and inefficient. There is an insufficiently small number of words which are commonly used in titles of papers relating to peptide synthesis which are peculiar to the subject. A reasonably comprehensive coverage could only be achieved by use of a profile containing a large and therefore uneconomic number of key-words, and much extraneous material was also unavoidably collected. In any case, important papers sometimes appear with titles containing no key-words which are necessarily suggestive of peptides at all* (such titles can also escape a screening procedure based on scanning journal title pages unless an author's name arouses attention). Some universally observed convention on the formulation of titles – such as the presence of the phrase 'Peptide Synthesis' in all titles – would be to the convenience and advantage of all concerned. However, even if such a convention were observed by the majority, their

* A recent example is provided by ref. 1, which is entitled 'Microanalysis by successive isotopic dilution. A new assay for racemic content.'

[1] D. S. Kemp, S. W. Wang, G. Busby, and G. Hugel, *J. Amer. Chem. Soc.*, 1970, **92**, 1043.

efforts would probably be confounded by a minority if experience with abbreviated formulae is any guide to the likely response to a plea for conformity. For the time being, therefore, a system based on personal inspection of journal title pages has been adopted. In order to provide a rational basis for the selection of journals for individual examination, a survey of the distribution of nearly 500 recent publications among different journals was conducted (Table 1). The 487 papers included in the survey

Table 1 *The distribution of papers relevant to peptide synthesis which were published during the period January* 1969—*October* 1970

Journal	*Total No. of references*	*No. of references of general interest* (i.e. '*Important*' *papers*)
Annalen	19	11
Austral. J. Chem.	4	3
Biochemistry	12	10
Bull. Chem. Soc. Japan	26	13
Chem. Ber.	13	11
Chem. Comm.	8	6
Chem. and Ind.	3	2
Chem. and Pharm. Bull. (Japan)	23	20
Coll. Czech. Chem. Comm.	6	6
Experientia	15	10
Helv. Chim. Acta	13	11
Internat. J. Protein Res.	2	2
J. Amer. Chem. Soc.	43	38
J. Biol. Chem.	9	6
J. Chem. Soc. (C)	12	9
J. Medicin. Chem.	17	14
J. Org. Chem.	31	14
Nature	6	5
Proc. Nat. Acad. Sci., U.S.A.	3	3
Rec. Trav. chim.	7	6
Tetrahedron	10	4
Tetrahedron Letters	23	14
Zhur. obshchei Khim.	30	4
Z. Naturforsch. (b)	12	8
Z. physiol. Chem.	6	6
48 Other journals	134	14
	Total 487	Total 250

(which comprise all relevant primary references which came to the Reporter's attention in the period January 1969—October 1970) were scattered through 73 journals, but only the 25 periodicals listed contained more than one paper of general interest. Further inspection showed that scanning the title pages of the journals named in the Table would have provided *ca.* 70% coverage of all relevant material and *ca.* 95% of all 'important' papers: if the reader will permit a yet more subjective judgement, it can be said that every paper which in the Reporter's opinion could

be called 'very important' appeared in the pages of one or other of these journals. Assuming that the distribution of the literature shown in Table 1 will be roughly maintained at least in the short term, it appears that the working peptide chemist can be confident of keeping well up to date with major developments by regularly perusing the title pages of current issues of the journals represented in the Table. The literature cited in this article was mainly located in this way, with use of *Chemical Abstracts*, Section 34, as a less up-to-date but more comprehensive net for other papers and secondary publications.

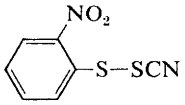
A number of books,[2–5] general reviews,[6–10] and the proceedings of three symposia [11–13] relevant to synthetic work with peptides have been published. Law's book [4] is essentially pedagogical in its aims, at an advanced undergraduate–graduate student level and contains many instructive problems. Jakubke and Jeschkeit's monograph [5] will be useful to beginners in the field (not only as a text but also, for English-speaking students, as a salutary exercise in reading Emil Fischer's native tongue) and established workers will find it a valuable concise compendium – critical original references on all aspects of peptide chemistry are given up to mid-1968.

2 Methods

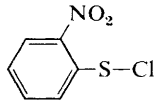
A. Protective Groups.—*Established Methods of Amino-group Protection.* Reviews on the synthesis of sulphenyl halides,[14] and their applications in peptide and protein chemistry [15] and experiments on the chemistry of sulphenamides [16] will be of interest to workers engaged in the use of sulphenyl protecting groups. *o*-Nitrophenylsulphenylamino-acids can be

[2] A. N. Shamin, 'Khimicheskii Sintez Belka,' Nauka, Moscow, 1969.
[3] M. M. Shemyakin and Yu. A. Ovchinnikov (eds.) 'Sovremennye Problemy Khimii Peptidov i Belkov,' Nauka, Moscow, 1969.
[4] H. D. Law, 'The Organic Chemistry of Peptides,' Wiley–Interscience, London, 1970.
[5] H. D. Jakubke and H. Jeschkeit, 'Aminosäuren, Peptide, Proteine: Eine Einführung,' Akademie-Verlag, Berlin, 1969.
[6] A. Kapoor, *J. Pharm. Sci.*, 1970, **59**, 1.
[7] N. Yanaihara, *Kagaku No Ryoiki*, 1970, **24**, 11 (*Chem. Abs.*, 1970, **73**, 15186w).
[8] M. Bodanszky, ref. 11, p. 1.
[9] A. Marglin and R. B. Merrifield, *Ann. Rev. Biochem.*, 1970, **39**, 841.
[10] D. Gish, in 'Protein Sequence Determination,' ed. S. B. Needham, Chapman and Hall, London, 1970, p. 276.
[11] Proceedings of the First American Peptide Symposium held at Yale, 1968: 'Peptides: Chemistry and Biochemistry,' ed. S. Lande and B. Weinstein, Dekker, New York, 1970.
[12] Proceedings of the Tenth European Peptide Symposium held at Albano, 1969, 'Peptides 1969,' ed. E. Scoffone, North Holland Publishing Co., 1971. (Not available at time of writing: see Vol. 4 of these Reports for reference to individual papers.)
[13] Proceedings of a symposium held at Paris, 1968, entitled 'La specificité zoologique des Hormones hyphophysaires et de leurs Activités', *Colloq. Int. Cent. Nat. Rech. Sci.*, 1969, **177**.
[14] E. Kühle, *Synthesis*, 1970, 561.
[15] A. Fontana and E. Scoffone in 'Mechanisms of Reactions of Sulphur Compounds,' Vol. IV, ed. N. Kharasch, Intrascience Research Foundation, Santa Monica, California, 1969, p. 15.
[16] N. H. Heimer and L. Field, *J. Org. Chem.*, 1970, **35**, 3012.

prepared in good yield from the amino-acid on treatment with *o*-nitrophenylsulphenyl thiocyanate (1) in the presence of silver nitrate at pH 6—9.[17] The reagent (1) is claimed [17] to suffer less decomposition on storage than the corresponding chloride (2). In the Reporter's experience the bad



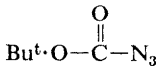
(1)



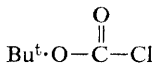
(2)

reputation which *o*-nitrophenylsulphenyl chloride (2) seems to have acquired in this respect is exaggerated. Provided that the chloride (2) is well purified and kept in a well-closed bottle at 0 °C, it can usually be stored for months before use without detriment.

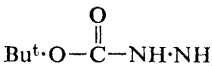
Explosions have occurred during the direct preparation of t-butyl azidoformate (3) from the chloroformate (4) and sodium azide.[18] The azidoformate is usually obtained from the carbazate (5), which can be



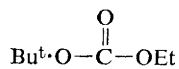
(3)



(4)



(5)



(6)

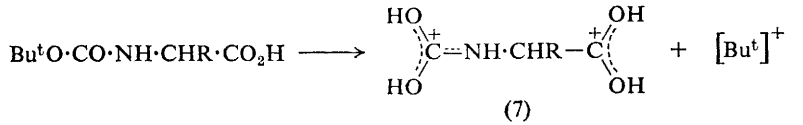
made in a number of ways. A further preparation of (5) by hydrazinolysis of the mixed carbonate (6) has been described.[19] Numerous reagents for the introduction of t-butoxycarbonyl groups have been recommended from time to time: a comprehensive bibliography is given in ref. 20. Explorations to find alternatives to the well-established azidoformate (3) are continuing.[20] A partially automated procedure for the preparation of t-butoxycarbonylamino-acids from the azidoformate (3) at controlled pH has been published.[21]

One of the difficulties with protecting groups which must be removed by acidolysis is that the electrophilic species generated on cleavage can attack sensitive side-chains. For example, treatment of t-butoxycarbonyltryptophan-containing peptides with trifluoroacetic acid results in butylation of the indole ring.[22] It is reported [22] that this side-reaction occurs to a

[17] J. Šavrda and D. H. Veyrat, *J. Chem. Soc.* (*C*), 1970, 2180.
[18] H. Yajima, H. Kawatani, and Y. Kiso, *Chem. and Pharm. Bull.* (*Japan*), 1970, **18**, 850.
[19] M. Muraki and T. Mizoguchi, *Chem. and Pharm. Bull.* (*Japan*), 1970, **18**, 217.
[20] L. A. Carpino, K. N. Parameswaran, R. K. Kirkley, J. W. Spiewak, and E. Schmitz, *J. Org. Chem.*, 1970, **35**, 3291.
[21] C. Birr and R. Frode, *Synthesis*, 1970, 474.
[22] Yu. B. Alakhov, A. A. Kiryushkin, V. M. Lipkin, and G. W. A. Milne, *Chem. Comm.*, 1970, 406.

'high degree', and is dependent on the position of tryptophan in the chain, these conclusions being based on the relative intensities of ions in the mass spectra of the mixtures of by-products and expected peptides. A caveat is in order here, however, since deductions about the proportion of a component in a mixture based on such observations are likely to be misleading if made in the absence of knowledge about the relative volatilities, thermal labilities, and molecular ion stabilities of the constituents. N.m.r. spectroscopy would appear to be a less equivocal analytical tool for the determination of the extent of this side-reaction. Nevertheless, this quibble about quantitative aspects does not detract from the fact that a serious (although not unexpected) problem has been revealed. No doubt the use of competing scavengers for electrophiles would ease the problem, but more elegant solutions probably lie in the use of protective groups which can be cleaved under milder conditions.

Olah and his colleagues have extended their n.m.r. studies of carbonium ions in superacidic solutions to amino-acids,[23] simple peptides,[23] and alkoxycarbonylamino-acids.[24] Complete alkyl–oxygen fission occurs with benzyloxycarbonyl- and t-butoxycarbonyl-amino-acids in antimony pentafluoride–fluorosulphonic acid–sulphur dioxide mixtures at − 76 °C giving, in the latter case, solutions in which both the t-butyl carbonium ion and the doubly-protonated carbamic acid (7) were observed (Scheme 1).[24] Benzyl-



(7)

Conditions: SbF_5–FSO_3H–SO_2, −76 °C

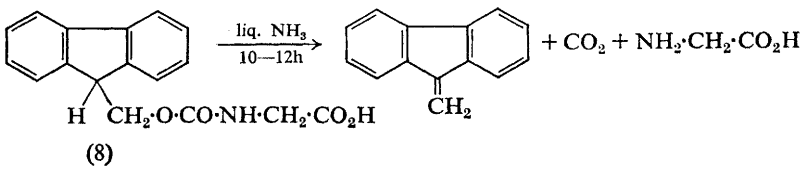
Scheme 1

oxycarbonylamino-acids also gave solutions containing species (7) but benzyl cations could not be detected, apparently because of their rapid polymerisation.[24]

Papers have appeared on the following: the benzyloxycarbonylation of amino-acid mixtures and separation of the products by preparative t.l.c.;[25] the preparation of *p*-methoxybenzyloxycarbonylamino-acids;[26, 27] the preparation of benzyloxycarbonylamino-acids by use of *N*-benzyloxycarbonyl-*N'*-methylimidazolium chloride;[28] and the use of modified benzyloxycarbonyl groups for lysine side-chain protection.[29]

[23] G. A. Olah, D. L. Brydon, and R. D. Porter, *J. Org. Chem.*, 1970, **35**, 317.
[24] G. A. Olah and D. L. Brydon, *J. Org. Chem.*, 1970, **35**, 313.
[25] V. Kalis, A. S. Karsakevich, V. Zuhovska, and A. Lesins, *Latv. PSR. Zinat. Akad. Vestis, Kim. Ser.*, 1970, 338 (*Chem. Abs.*, 1970, **73**, 77574e).
[26] S. Sofuku, M. Mizumura, and A. Hagitani, *Bull. Chem. Soc. Japan*, 1970, **43**, 177.
[27] F. Vandesande, *Bull. Soc. chim. belges*, 1970, **79**, 397.
[28] E. Guibé-Jampel, G. Bram, and M. Vilkas, *Tetrahedron Letters*, 1969, 3541.
[29] K. Noda, S. Terada, and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 1883.

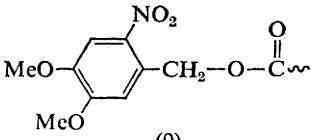
New Methods of Amino-group Protection. Limitations imposed by the need to conserve optical integrity during peptide synthesis severely restrict the development of new methods for the protection of α-amino-groups. There is, however, considerable scope for useful variations on the carbamate ester theme: any protecting group of this type can be relied on to restrain itself from entering into oxazolone formation and so allowing racemisation. Fission of the ester–oxygen bond (which results ultimately in exposure of the original amino-function because decarboxylation of the carbamic acid rapidly ensues) can in principle be performed by a number of selective methods. In practice, carbamate ester-protecting groups labile to acidolysis of graded severity or to catalytic hydrogenolysis have been most widely used. Cleavage by other reductive conditions (as with the piperidinooxycarbonyl group – see Vol. 2, p. 149, of these Reports) may yet become popular, and selective deprotection by β-elimination or by photolysis has also been considered from time to time. It has recently been reported [30] that 9-fluorenylmethoxycarbonylglycine (8) can be cleaved by β-elimination under very mild basic conditions (Scheme 2) but is not affected by hydrogen



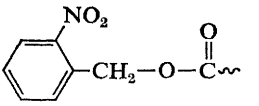
(8)

Scheme 2

bromide or chloride in organic solvents, by trifluoroacetic acid, or by the conditions used for catalytic hydrogenolysis. However, no other derivatives bearing this new protecting group have yet been described and the use of (8) in coupling reactions was not mentioned in the preliminary note,[30] so that an assessment of the potential of the 9-fluorenylmethoxycarbonyl group must await further details. A similar shortage of information on practical applications inhibits full discussion of recently-described [31] photochemically-labile *N*-protecting groups such as (9) and (10).



(9)



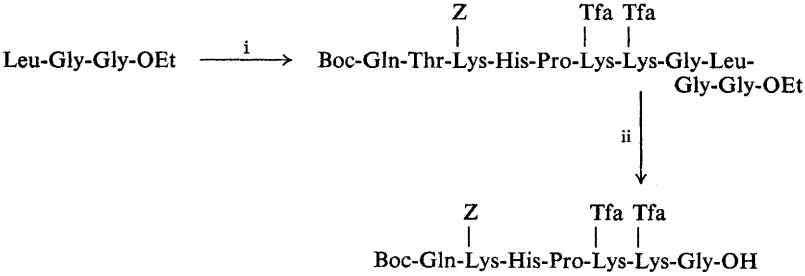
(10)

Miscellaneous Publications on Amino-group Protection. Reservations about the usefulness of tritylvaline and *o*-nitrophenylsulphenyl-*S*-tritylcysteine have been answered.[32] Other relevant references include: work on the

[30] L. A. Carpino and G. Y. Han, *J. Amer. Chem. Soc.*, 1970, **92**, 5748.
[31] A. Patchornik, B. Amit, and R. B. Woodward, *J. Amer. Chem. Soc.*, 1970, **92**, 6333.
[32] L. Zervas, *Z. Naturforsch.*, 1970, **25b**, 322.

reductive cleavage of *N*-trifluoroacetyl and *N*-trichloroacetyl groups with sodium borohydride;[33] the trichloroacetylation of peptides with hexachloroacetone in dimethyl sulphoxide;[34] the removal of *N*-chloroacetyl groups with thiourea;[35] the preparation of *N*-protected amino-acids *via* their trimethylsilyl esters;[36] and the *N*-methylation of alkoxycarbonyl-amino-acids on a preparative scale.[37]

Protection of Carboxy-groups. The synthesis of partially-protected oligopeptides which have a free *C*-terminal carboxy-group but differential protection between the *N*-terminal amino-group and the side-chains is often quite tricky. It has recently been shown[38] that enzymic removal of part of the peptide chain (which can be regarded as the *C*-terminal protecting group of the desired intermediate) can be used for this purpose, and an abbreviated outline of the example described[38] is shown in Scheme 3.



Conditions: i, standard stepwise and fragment condensation methods;
ii, thermolysin at pH 7·4

Scheme 3

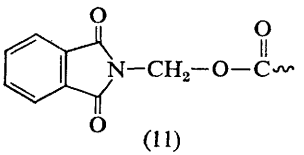
This particular approach is clearly limited in its applicability, since the desired intermediate must obviously contain no thermolysin-sensitive bonds and, furthermore, the fully-blocked peptide must be sufficiently soluble in aqueous media for enzymic cleavage. However, this is an intriguing approach and this report[38] may prove to be the nucleation point for a fresh direction of development.

Phthalimidomethyl esters (11) were recommended[39] for use in peptide synthesis some seven years ago. Acidolysis, saponification, and hydrazinolysis, but not hydrogenolysis may be used for fission of (11)[39] but these conditions offer no advantage over more thoroughly tried protective groups. This situation is altered by a brief note[40] to the effect that (11) is cleaved

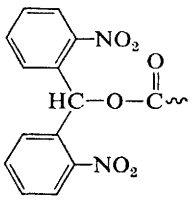
[33] F. Weygand and E. Frauendorfer, *Chem. Ber.*, 1970, **103**, 2437.
[34] C. A. Panetta and T. G. Casanova, *J. Org. Chem.*, 1970, **35**, 2423.
[35] C. Toniolo and A. Fontana, *Gazzetta*, 1969, **99**, 1017.
[36] H. R. Kricheldorf, *Synthesis*, 1970, 592.
[37] R. K. Olsen, *J. Org. Chem.*, 1970, **35**, 1912.
[38] M. Ohno and C. B. Anfinsen, *J. Amer. Chem. Soc.*, 1970, **92**, 4098.
[39] G. H. L. Nefkens, G. I. Tesser, and R. J. F. Nivard, *Rec. Trav. chim.*, 1963, **82**, 941.
[40] D. L. Turner and E. Baczynski, *Chem. and Ind.*, 1970, 1204.

by zinc in acetic acid at room temperature – a very selective reagent which disturbs neither benzyl- nor t-butyl-based protection.

The carboxy-protecting group (12) can be removed in very high yield from simple amino-acid derivatives[31] by photolysis, but practical details of its use are not yet published.



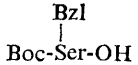
(11)



(12)

ε-*N*-Benzyloxycarbonyl-lysine can be esterified using boron trifluoride–alcohol mixtures.[41] Amino-acid *p*-methoxybenzyl ester hydrochlorides can be obtained *via* the corresponding *o*-nitrophenylsulphenyl derivative.[42] 2,2-Diphenyl-2-hydroxyethyl esters can be used as a 'safety-catch' method of carboxy-protection[43] (see below under Activated Esters). Preparations of benzyloxycarbonylasparagine *p*-nitrobenzyl ester and of benzyloxycarbonylaspartic acid *p*-nitrobenzyl ester have been described.[44]

Protection of Hydroxy-groups. *O*-Benzyl-t-butoxycarbonylserine (13) is an important intermediate for the introduction of protected serine residues, but until recently has only been accessible by circuitous routes. Experimental details of a direct conversion of t-butoxycarbonylserine to (13) in



(13)

45% yield are now available.[45] The procedure, which was much less satisfactory when applied to t-butoxycarbonylthreonine, involves dissolution in liquid ammonia followed by addition of sodium and benzyl bromide. The fact that column chromatography seems to be necessary for isolation and purification of (13) when prepared by this method may prove to be a nuisance in laboratories where large amounts are required, but perhaps a more conveniently scaled-up method will be developed in due course.

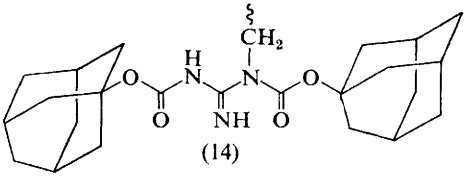
Protection of Primary Amide Groups. Details[46] of the protection of amide side-chain and *C*-terminal amide groups as their acid-labile 4,4′-di-

[41] J. Coggins, R. Demayo, and N. L. Benoiton, *Canad. J. Chem.*, 1970, **48**, 385.
[42] G. C. Stelakatos and N. Argyropoulos, *J. Chem. Soc.* (*C*), 1970, 964.
[43] T. Wieland, J. Lewalter, and C. Birr, *Annalen*, 1970, **740**, 31.
[44] L. K. Polevaya, O. V. Smirnov, and G. Cipens, *Latv. P.S.R. Zinat. Akad. Vestis, Kim. Ser.*, 1969, 711 (*Chem. Abs.*, 1970, **72**, 133160k).
[45] V. J. Hruby and K. W. Ehler, *J. Org. Chem.*, 1970, **35**, 1690.
[46] W. König and R. Geiger, *Chem. Ber.*, 1970, **103**, 2041.

methoxybenzhydryl derivatives (see Vol. 1, p. 182, of these Reports) have been published, and a preliminary notice [47] of the use of the less acid-sensitive *p*-methoxybenzylamides has also appeared. The new derivatives which have been described are incorporated in Appendix B.

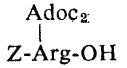
Protection of Arginine Side-chains. Nitration has been used more than any other means of protecting arginine side-chains, but ω-nitroarginine derivatives are notoriously insoluble and are subject to several side-reactions. Tosylation also provides adequate protection but removal of this blocking group requires undesirably vigorous conditions (sodium in ammonia, or liquid hydrogen fluoride). Various other methods have been examined; for leading references see the introductory section of ref. 48.

Extensive practical details of a new method of protecting guanidino-functions with two 1-adamantyloxycarbonyl groups (14) have now



(14)

appeared.[48] These details were expected following the use of this method in the first synthesis of porcine proinsulin connecting peptide which was discussed in last year's Report (see also the footnote on p. 267). Treatment of N^{α}-benzyloxycarbonylarginine with 1-adamantyl chloroformate under alkaline conditions gives (15) which has been used in standard coupling

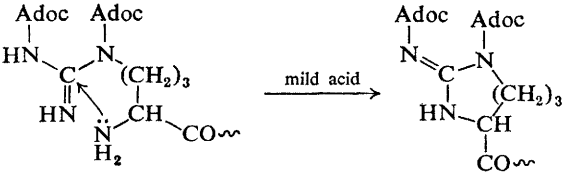


(15)

reactions. The guanidino-group remains blocked if hydrogenolysis is used for the removal of other groups, but is regenerated on treatment with trifluoroacetic acid. Of practical value is the fact that the two adamantyl groups provide so effective a hydrophobic shell for the side-chain that intermediates protected in this way are soluble in organic solvents. No side-reactions have been reported so far except when α-amino-groups of *N*-terminal bis-ω-(1-adamantyloxycarbonyl)-arginine residues were liberated under weakly acidic conditions, which gave rise to cyclisation as shown in Scheme 4.

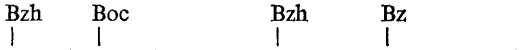
The products of electrochemical reduction of ω-nitroarginine under various conditions have been studied.[49]

[47] P. G. Pietta and G. R. Marshall, *Chem. Comm.*, 1970, 650.
[48] G. Jäger and R. Geiger, *Chem. Ber.*, 1970, **103**, 1727.
[49] G. Capobianco, S. Zecchin, and G. Vidali, *Ann. Chim. (Italy)*, 1970, **60**, 37.



Scheme 4

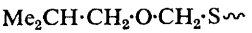
Protection of Thiol Groups and Synthesis of Cystine Peptides. There have been several further publications [50–52] on the scope and limitations of the sulphenyl isothiocyanate method of forming disulphide bridges specifically (see Vol. 1, p. 185, of these Reports) showing *inter alia* that this procedure can be made compatible with all the functional side-chains which occur in proteins.[51] In the course of this work it was discovered [53] that attempted saponification of esters such as (16) resulted in significant amounts of β-elimination of diphenylmethanethiol. Similarly, attempted selective



(16) (17)

exposure of the benzoylated sulphur atom in (17) by treatment with sodium methoxide caused serious amounts of β-elimination in solvent mixtures containing dimethylacetamide.[53]

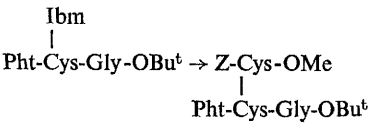
The *S*-isobutoxymethyl group (18) [54] is stable to hydrazine in boiling ethanol and to hydrochloric acid in acetone but decomposes in aqueous



(18)

alkali, hydrogen bromide in acetic acid, and trifluoroacetic acid.[52] These properties limit its use somewhat, but it is compatible with trityl or phthaloyl *N*-protection. The reaction of (18) with sulphenyl isothiocyanates can be used for the synthesis of unsymmetrical disulphides (*e.g.* Scheme 5) but the reactivities of (18) and the *S*-trityl group are not sufficiently different to permit selective attack of one in the presence of the other. However, *S*-isobutoxymethylcysteine peptides have the important practical advantage of being in general more easily isolated as crystalline solids than are the corresponding *S*-trityl derivatives.[52]

[50] R. G. Hiskey, R. L. Smith, A. M. Thomas, J. T. Sparrow, and W. C. Jones, jun., ref. 11, p. 411.
[51] R. G. Hiskey and B. J. Ward, jun., *J. Org. Chem.*, 1970, **35**, 1118.
[52] R. G. Hiskey and J. T. Sparrow, *J. Org. Chem.*, 1970, **35**, 215.
[53] R. G. Hiskey, R. A. Upham, G. M. Beverly, and W. C. Jones, jun., *J. Org. Chem.*, 1970, **35**, 513.
[54] P. J. E. Brownlee, M. E. Cox, B. O. Handford, J. C. Marsden, and G. T. Young, *J. Chem. Soc.*, 1964, 3832.

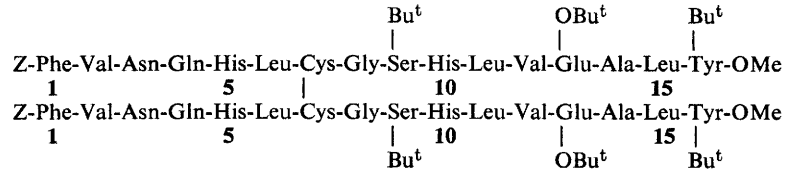


Ibm = isobutoxymethyl

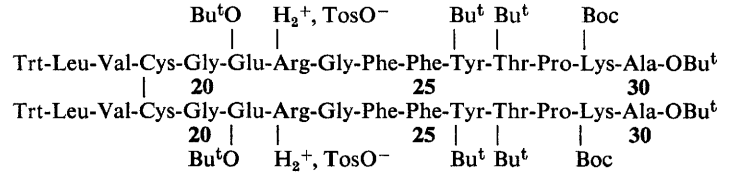
Conditions: Z-Cys(SCN)-OMe,AcOH (77% yield)

Scheme 5

The usefulness of a series of ten *S*-protecting groups in solid-phase synthesis has been evaluated.[55] One approach to the problem of *S*-protection is to use the cysteine peptide as its own *S*-blocking group, *i.e.* to construct symmetrical cystine peptides as interim objectives. Thus a combination of stepwise and fragment condensations, starting in both cases from bis-benzyloxycarbonylcystine, has been used for the preparation of the two symmetrical disulphides (19) and (20) which correspond to sequences



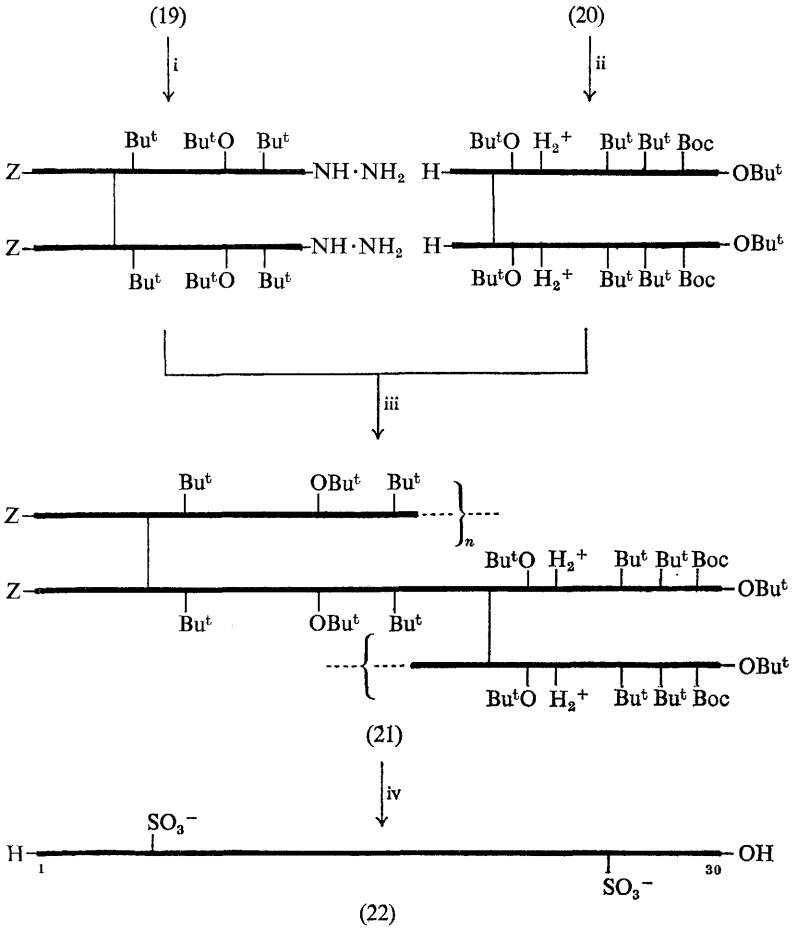
(19)



(20)

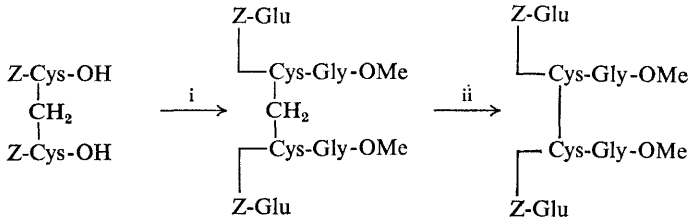
1—16[56] and 17—30[57] respectively of the insulin B chain. The two fragments were combined as shown in Scheme 6 to give a polymer (21) which, on deprotection and oxidative sulphitolysis, gave insulin B chain di-*S*-sulphonate (22).[57] The amount of insulin activity regenerated when the synthetic B chain was combined with natural A chain was almost as great as when natural B chain was used. Synthetic B chains which have been exposed to sodium in ammonia are considerably less effective than natural B chain in regenerating insulin activity when recombined with natural A chain. The remarkable success of this demanding example suggests that

[55] K. Hammerstroem, W. Lunkenheimer, and H. Zahn, *Makromol. Chem.*, 1970, **133**, 41.
[56] H. Zahn and G. Schmidt, *Annalen*, 1970, **731**, 91.
[57] H. Zhan and G. Schmidt, *Annalen*, 1970, **731**, 101.



Conditions: i, $NH_2 \cdot NH_2$–DMF; ii, 80% AcOH, 14 h, 35 °C; iii, azide coupling in DMF; iv, deprotection and oxidative sulphitolysis

Scheme 6

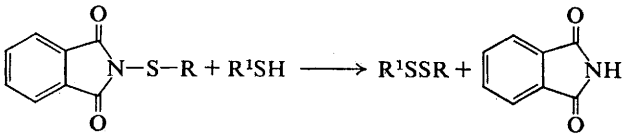


Conditions: i, standard procedures; ii, Br_2–$CHCl_3$, then $NaHCO_3$

Scheme 7

the symmetrical disulphide route to cysteine peptides may prove generally useful. Symmetrical cystine peptides can also be obtained by solid-phase synthesis,[58, 59] and by oxidation of symmetrical peptides of djenkolic acid with bromine, as illustrated by a recent synthesis of a protected oxidised glutathione (Scheme 7).[60]

It has been suggested that the reaction of thiols with *S*-phthalimido derivatives (Scheme 8) could be elaborated into a practicable method for the specific construction of unsymmetrical disulphide bridges.[61]



Conditions: equivalent amounts of reactants refluxed in benzene for up to 20 h

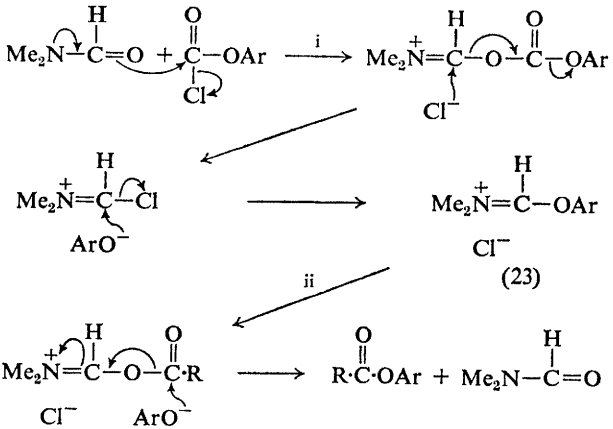
Scheme 8

Simplified preparations of *S*-trityl- and *S*-benzhydryl-cysteine have been elaborated, and the optimal conditions for acidolysis of these protecting groups have been carefully investigated and defined.[62] The use of the acid-labile *S*-*p*-methoxybenzyloxycarbonyl group in a simple tripeptide synthesis has been described.[63]

Several publications on problems associated with the synthesis of selenium-containing peptides have appeared.[64–68] Interest in these stems largely from the fact that synthetic analogues such as 1-deamino-[1,6-diseleno]-oxytocin [69] have high biological activity.

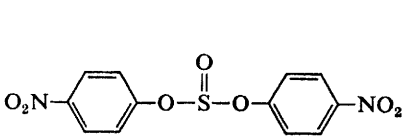
B. Formation of the Peptide Bonds.—*Activated Esters.* The use of activated esters in peptide synthesis has been reviewed.[70] Aryl chloroformates react with dry dimethylformamide to give imonium salts of general formula (23) which react with acylamino-acids yielding active esters in moderate to good yield (Scheme 9),[71] but the examples described comprise only active esters which can be more conveniently prepared by simpler means. Application of the method shown in Scheme 9 in the Anderson test (*i.e.*

[58] W. Lunkenheimer and H. Zahn, *Angew. Makromol. Chem.*, 1970, **10**, 69.
[59] W. Lunkenheimer and H. Zahn, *Annalen*, 1970, **740**, 1.
[60] B. Marinier and M. Berube, *Canad. J. Chem.*, 1960, **47**, 4507.
[61] D. N. Harpp, D. K. Ash, T. G. Back, J. G. Gleason, B. A. Orwig, W. F. Van Horn, and J. P. Snyder, *Tetrahedron Letters*, 1970, 3551.
[62] I. Photaki, J. Taylor-Papadimitriou, C. Sakarellos, P. Mazarakis, and L. Zervas, *J. Chem. Soc.* (*C*), 1970, 2683.
[63] I. Photaki, *J. Chem. Soc.* (*C*), 1970, 2687.
[64] R. Walter, ref. 11, p. 467.
[65] J. Roy, W. Gordon, I. L. Schwartz, and R. Walter, *J. Org. Chem.*, 1970, **35**, 510.
[66] W. Gordon, R. Walter, I. L. Schwarz, and D. Theodoropoulos, *Chem. Chron., Epistem. Ekdosis*, 1969, **34**, 75 (*Chem. Abs.*, 1970, 72, 55857b).
[67] J. Roy, I. L. Schwartz, and R. Walter, *J. Org. Chem.*, 1970, **35**, 2840.
[68] C. S. Pande, J. Rudick, and R. Walter, *J. Org. Chem.*, 1970, **35**, 1440.
[69] R. Walter and V. du Vigneaud, *J. Amer. Chem. Soc.*, 1966, **88**, 1331.
[70] H. G. Garg, *J. Sci. Ind. Res. India*, 1970, **29**, 236.
[71] M. Itoh, *Chem. and Pharm. Bull.* (*Japan*), 1970, **18**, 784.

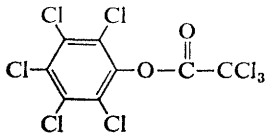


Conditions: i, dry conditions; ii, RCO_2H and a tertiary amine in a suitable solvent

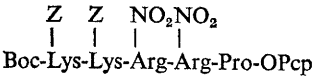
Scheme 9



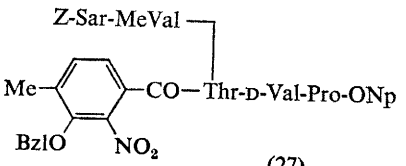
(24)



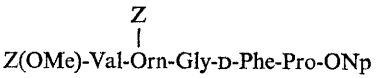
(25)



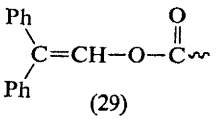
(26)



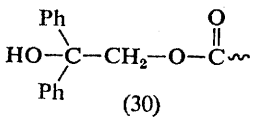
(27)



(28)



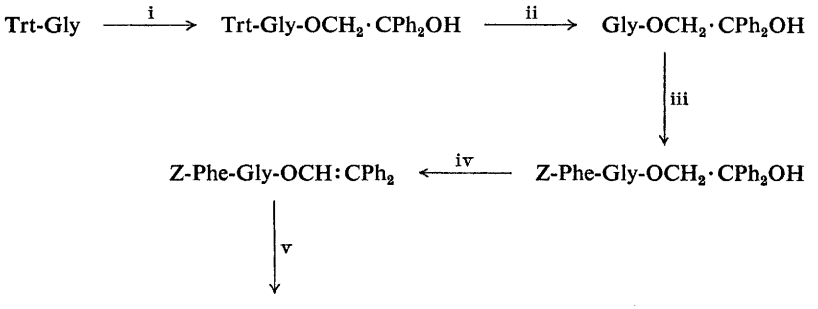
(29)



(30)

activation of benzyloxycarbonylglycylphenylalanine) caused complete racemisation, but this route may prove useful for the direct activation of partially protected oligopeptide acids which terminate in glycine or proline. The dicyclohexylcarbodi-imide procedure which is routinely employed for the preparation of acylamino-acid active esters is often dogged with isolation and purification difficulties when applied to acylpeptides. In this context it is worth noting that although only a few laboratories have shown interest in bis-*p*-nitrophenyl sulphite (24),[72] pentachlorophenyl trichloroacetate (25)[73] and related reagents, some quite complex protected peptide active esters have been prepared by use of such compounds. Recent examples include intermediates in syntheses of corticotropic peptides [*e.g.* (26)[74]], actinomycin D (27),[75, 76] and cyclopeptide antibiotics [*e.g.* (28)[77]].

2,2-Diphenylvinyl esters [*e.g.* (29)], which are reactive towards aminolysis, can be generated from 2,2-diphenyl-2-hydroxyethyl esters [*e.g.* (30)] by acid-catalysed dehydration under mild conditions.[43] The feasibility of using this fact as a 'safety-catch' device has been illustrated by a simple example (Scheme 10).[43]



Conditions: i, $CH_2OH \cdot CPh_2OH$–$SOCl_2$–py; ii, 50% AcOH; iii, Z-Phe-O·CO_2Et; iv, CF_3CO_2H; v, Bzl–NH_2

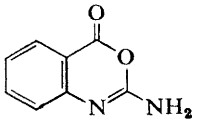
Scheme 10

Other papers on active esters have covered the following: the synthesis of pentafluorophenyl esters;[78] the use of esters of *N*-hydroxyphthalimide;[79] the use of acetylglycine cyanomethyl ester;[80] esters of benzohydroxamic

[72] B. Iselin, W. Rittel, P. Sieber, and R. Schwyzer, *Helv. Chim. Acta*, 1957, **40**, 373.
[73] M. Fujino and C. Hatanaka, *Chem. and Pharm. Bull. (Japan)*, 1968, **16**, 929.
[74] M. Fujino, C. Hatanaka, and O. Nishimura, *Chem. and Pharm. Bull. (Japan)*, 1970, **18**, 771.
[75] J. Meienhofer, Y. Sano, and R. P. Patel, ref. 11, p. 419.
[76] J. Meienhofer, *J. Amer. Chem. Soc.*, 1970, **92**, 3771.
[77] O. Abe and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 1202.
[78] L. Kisfaludy, J. E. Roberts, R. H. Johnson, G. L. Mayers, and J. Kovacs, *J. Org. Chem.*, 1970, **35**, 3563.
[79] O. A. Kaurov, V. F. Martynov, and V. B. Morozov, *Zhur. obshchei. Khim.*, 1970, **40**, 908 (*Chem. Abs.*, 1970, **73**, 35748j).
[80] F. H. C. Stewart, *Austral. J. Chem.*, 1970, **23**, 2147.

acid;[81] the preparation of 1-acyloxy-2-pyridones;[82] esters of 1-hydroxybenzotriazole (see below);[83] the non-aqueous titrimetric determination of active esters;[84] further studies with esters of 4-(methylthio)-phenol;[85, 86] polymeric active esters;[87] applications of the di- and tri-chloroacetates of *p*-nitrophenol, 2,4,5-trichlorophenol, and *N*-hydroxysuccinimide;[88] and pentachlorophenyl esters.[89]

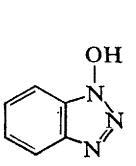
Coupling Methods Involving Dicyclohexylcarbodi-imide. The mechanism of aminolysis of 2-amino-4,5-benzo-6-oxo-1,3-oxazine (31) has been investigated:[90] this compound is a model for the *O*-acylisourea intermediate



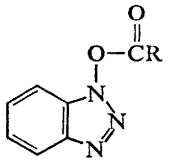
(31)

formed during direct carbodi-imide coupling reactions. 1-Hydroxybenzotriazole (32) has been shown [83] to be a very suitable 'ester component' for use as an additive in dicyclohexylcarbodi-imide coupling. Racemisation in the Weygand test was greatly suppressed, acylurea formation was eliminated and yields were high in a large number of examples. The reaction can be performed in stages with separate formation and isolation of a highly reactive 'active ester' of (32): in the crystalline state the 'active ester' appears in most cases to be mainly the *O*-acyl derivative (33) but is an *N*-acyl derivative (34) and/or (35) in solution.[83] Whereas side-reactions have been observed with the *N*-hydroxysuccinimide–dicyclohexylcarbodi-imide procedure,[91] none were reported with 1-hydroxybenzotriazole as additive. 3-Hydroxy-4-oxo-3,4-dihydro-1,2,3-benzotriazine (36) can be used in a similar fashion,[92] and is the most efficient of a large number of *N*-hydroxy-compounds screened as additives for suppressing racemisation,[93] but a by-product (37) which is an acylating agent is formed by opening of the heterocyclic ring.[92]

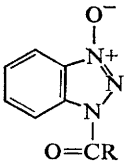
[81] L. Lubiewska-Nakonieczna, B. Rzeszotarska, and E. Taschner, *Annalen*, 1970, **741**, 157.
[82] E. C. Taylor, F. Kienzle, and A. McKillop, *J. Org. Chem.*, 1970, **35**, 1672.
[83] W. König and R. Geiger, *Chem. Ber.*, 1970, **103**, 788.
[84] M. Wilchek, M. Fridkin, and A. Patchornik, *Analyt. Chem.*, 1970, **42**, 275.
[85] B. J. Johnson and T. A. Ruettinger, *J. Org. Chem.*, 1970, **35**, 255.
[86] B. J. Johnson and D. E. Tracey, *J. Pharm. Sci.*, 1969, **59**, 1299.
[87] G. T. Panse and D. A. Laufer, *Tetrahedron Letters*, 1970, 4181.
[88] M. Fujino and C. Hatanaka, *Takeda Kenkyusho Nempo*, 1969, **28**, 12 (*Chem. Abs.*, 1970, **72**, 55871b).
[89] A. Kapoor, ref. 11, p. 17.
[90] A. F. Hegarty and T. C. Bruice, *J. Amer. Chem. Soc.*, 1970, **92**, 6568.
[91] *E.g.* F. Weygand, W. Steglich, and N. Chytil, *Z. Naturforsch.*, 1968, **23b**, 1391.
[92] W. König and R. Geiger, *Chem. Ber.*, 1970, **103**, 2034.
[93] W. König and R. Geiger, *Chem. Ber.*, 1970, **103**, 2024.



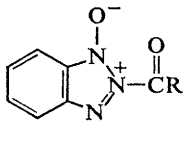
(32)



(33)

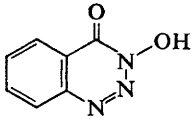


(34)

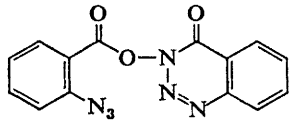


(35)

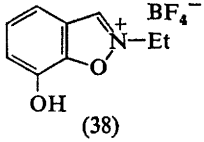
Other Methods. A discussion [94] of the scope and limitations of 7-hydroxybenzisoxazolium fluoroborate (38) as a reagent for fragment condensation has appeared, and a detailed progress report [95] on a rational search for other isoxazolium salt reagents has been published.



(36)

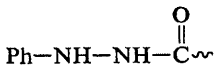


(37)



(38)

Mention was made in Volume 1 of these Reports (p. 191) of the oxidation of phenylhydrazides (39) to phenyldi-imides (40) which are reactive towards aminolysis. The *N*-bromosuccinimide technique for oxidative activation of (39) has been used in a synthesis of antamanide,[43] and as the basis for a

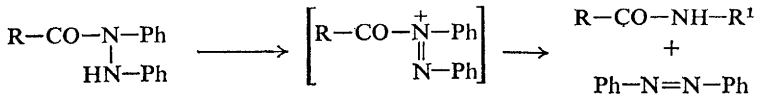


(39)

(40)

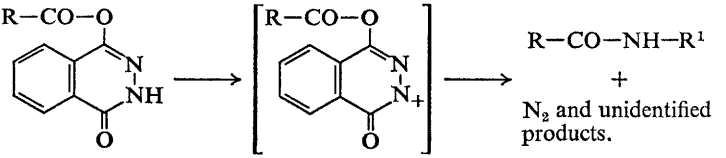
new mild method of cleaving from solid-phase resins peptides which are attached through a hydrazide link.[43] Further variations on this method are outlined in Schemes 11,[43] 12,[43] and 13.[96]

[94] D. S. Kemp, ref. 11, p. 33.
[95] R. A. Olofson and Y. L. Marino, *Tetrahedron*, 1970, **26**, 1779.
[96] J. Lewalter and C. Birr, *Annalen*, 1970, **740**, 48.



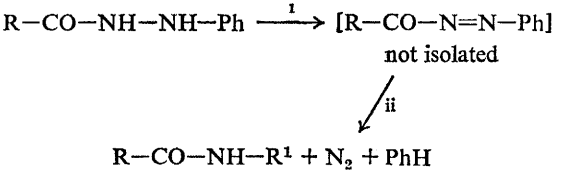
Conditions: *N*-Bromosuccinimide–R^1NH_2

Scheme 11



Conditions: *N*-Bromosuccinimide–R^1NH_2

Scheme 12

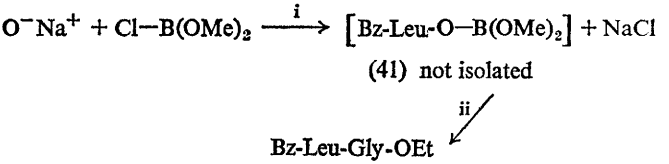


Conditions: i, anodic oxidation in CH_2Cl_2 in the presence of $Bu_4N^+ClO_4^-$; ii, R^1NH_2

Scheme 13

Acyloxyborane derivatives [*e.g.* (41)] can be obtained by reaction of carboxylic acids with variously substituted boranes and are easily aminolysed.[97, 98] However, the yield in a Young test (Scheme 14) was low and some racemisation was observed,[97] and the authors recognise [97] that this is not at present a viable method of peptide synthesis.

Other papers have been concerned with: the development of a practical method of synthesising optically active peptides based on four-component



Conditions: i, C_6H_6, room temp., 18 h; ii, Gly-OEt

Scheme 14

[97] A. Pelter, T. E. Levitt, and P. Nelson, *Tetrahedron*, 1970, **26**, 1539.
[98] A. Pelter and T. E. Levitt, *Tetrahedron*, 1970, **26**, 1545.

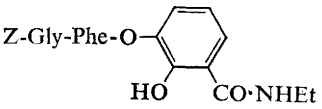
condensations;[99] further studies of the so-called oxidation–reduction method of peptide synthesis;[100, 101] a brief review of acylating agents suitable for peptide synthesis, including acyloxyphosphonium salts;[102] cupric ion catalysis of the oligomerisation of amino-acid esters;[103] peptide synthesis with triphenylphosphine and imidazole;[104] side-reactions observed on treatment of aspartic acid-containing peptides with hydrazine;[105] comparison of some coupling methods in a sterically hindered situation;[106] and the detection of acylamino-acid and acylpeptide hydrazides on t.l.c.[107]

C. Purification of Synthetic Intermediates.—Some workers, notably Hofmann and his colleagues,[108] have stressed that minimal protection of ionisable side-chains carries the advantage that ion-exchange chromatography can be used for the purification of synthetic intermediates. Most research groups, on the other hand, have been of the opinion that in most cases the disadvantages of this approach (poor solubility in organic media and side-reactions) argue in favour of maximal protection. Since fully-protected synthetic peptide intermediates are usually soluble only in organic solvents their purification has sometimes presented considerable difficulty, and several laboratories have been quick to appreciate the suitability of Sephadex LH-20 [109] for this purpose. LH-20 is a hydroxypropylated cross-linked dextran which swells not only in water but also in organic solvents such as methanol, dimethylformamide, chloroform, and dioxan, giving gels with exclusion limits of some 5000, depending on the solvent.[109] LH-20 gels are thus ideally suited for gel filtration of fully-blocked oligopeptides. Furthermore, if a mixture of solvents is used, the most polar solvent is preferentially adsorbed, making LH-20 suitable for partition chromatography. A fair number of applications of LH-20 in peptide chemistry have been described in the last year or so, but most of these are hidden away in experimental sections of papers which are primarily concerned with other matters. For this reason, the Reporter felt that a list of recent examples (Appendix C) might be of some value to those who are considering use of LH-20.

[99] I. Ugi, *Rec. Chem. Progr.*, 1969, **30**, 289; D. Marquarding, P. Hoffmann, H. Heitzer, and I. Ugi, *J. Amer. Chem. Soc.*, 1970, **92**, 1969.
[100] T. Mukaiyama, R. Matsueda, and M. Suzuki, *Tetrahedron Letters*, 1970, 1901.
[101] T. Mukaiyama, R. Matsueda, and H. Maruyama, *Bull. Chem. Soc. Japan*, 1970, **43**, 1271.
[102] Y. Ueno and M. Okawura, *Yuki Gosei Kagaku Kyokai Shi*, 1970, **28**, 351 (*Chem. Abs.*, 1970, **72**, 133151h).
[103] S. Yamada, S. Terashima, and M. Wagatsuma, *Tetrahedron Letters*, 1970, 1501.
[104] U. V. Mitin and O. V. Glinskaya, *Tetrahedron Letters*, 1969, 5267.
[105] N. Mitsuyasu, M. Waki, T. Kato, and N. Izumiya, *Mem. Fac. Sci., Kyushu Univ., Ser. C.*, 1970, **7**, 97 (*Chem. Abs.*, 1970, **73**, 88154u).
[106] C. Birr, W. Lochinger, and T. Wieland, *Annalen*, 1969, **729**, 213.
[107] H. J. Goren and M. Fridkin, *J. Chromatog.*, 1970, **47**, 519.
[108] *E.g.* ref. 233 and previous papers in the series.
[109] 'Sephadex LH-20: Chromatography in Organic Solvents,' Pharmacia Fine Chemicals, Uppsala, 1970.

D. Racemisation.—The widespread practice in organic chemistry of referring to important well-known reactions and procedures by means of the names of the principal originators has not been excessively applied in peptide chemistry, but the convention of making reference to the main racemisation tests in this way seems to be firmly established. The name of Kemp must now be added to those of Anderson,[110] Izumiya,[111] Weygand,[112] and Young [113] (Table 2).

Kemp's racemisation tests [1] are in essence extensions of those of Anderson and Young, but include the important original feature of using isotopic dilution for assay of racemate. The sensitivity is thereby increased by some orders of magnitude, making the method suitable for determining racemic contents in the range 1.0—0.001%, which is below that accessible by other tests. The details of the procedure are somewhat involved, so we shall content ourselves here with mention of selected results only.[1, 114] In the modified Anderson test, Woodward's Reagent K gave 2—4% racemisation,[1] and the *p*-nitrophenyl ester gave amounts depending sharply on the solvent (0.2% in acetonitrile, 0.53% in dimethylformamide, 1—2% in dimethyl sulphoxide or hexamethylphosphoramide).[1] Only results for chloroform solution were reported for coupling of the optically pure succinimido ester (0.86% racemate) [114] – since the *p*-nitrophenyl ester [1] does not appear to have been examined in this solvent a direct comparison here seems unjustified. Data presented [114] on the oxine ester (0.16% racemate in dimethylformamide) seem to suggest its superiority over the *p*-nitrophenyl ester. The summarising statement that 'The optically pure *N*-hydroxysuccinimide and 8-hydroxyquinoline esters have been found to be roughly comparable to the *p*-nitrophenyl esters in the extent of racemisation during coupling [114] therefore seems to call for some clarification. The ester (42),



(42)

which was prepared from 7-hydroxybenzisoxazolium fluoroborate (38) and recrystallised, gave slightly less racemate (0.13%) than the oxine ester when coupled with glycine ethyl ester in DMF at room temperature, but the optimal conditions for peptide bond formation (coupling with glycine tetramethylammonium salt in 50 : 50 DMF–DMSO at 0 °C or coupling with glycine ethyl ester in DMF at 0 °C) gave only 0.01% racemate.[1] The acyl azide procedure, on which so much confidence in the design of fragment condensation strategies has previously rested, also gave under the

[110] G. W. Anderson and F. M. Callahan, *J. Amer. Chem. Soc.*, 1958, **80**, 2902.
[111] N. Izumiya and M. Muraoka, *J. Amer. Chem. Soc.*, 1969, **91**, 2391.
[112] F. Weygand, A. Prox, L. Schmidhammer, and W. König, *Angew. Chem.*, 1963, **75**, 282.
[113] M. W. Williams and G. T. Young, *J. Chem. Soc.*, 1963, 881.
[114] D. S. Kemp, Z. Bernstein, and J. Rebek, jun., *J. Amer. Chem. Soc.*, 1970, **92**, 4756.

Table 2 *The principal racemisation tests*

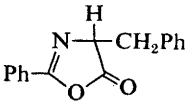
Test	*Model reaction*	*Method of determining racemisation*	*Sensitivity limit*
1. Anderson	Z-Gly-Phe + Gly-OEt	Specific rotation of crude material and isolation of racemate by fractional crystallisation	*ca.* 1—2% racemisation
2. Izumiya	Z-Gly-Ala + Leu-OBzl	Deprotection and determination of Gly-D-Ala-Leu on an amino-acid analyser	*ca.* 0.1—1.0% racemisation
3. Kemp	As in the Anderson or Young tests, but using [^{14}C]-labelled carboxy-components	Isolation of radioactive racemate after dilution with unlabelled racemate	*ca.* 0.001—0.01% racemisation
4. Weygand	Z-Leu-Phe + Val-OBu^{t}[a]	Deprotection and partial hydrolysis followed by determination of D-Phe-Val by g.l.c.	*ca.* 0.1—1.0% racemisation
5. Young	Bz-Leu + Gly-OEt	Specific rotation of crude material: presence of racemate can be confirmed by isolation after saponification	*ca.* 1—2% racemisation[b]

[a] And others. [b] The model reaction is some ten times as susceptible to racemisation as most typical peptide coupling reactions.

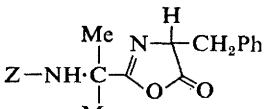
best conditions only 0.01—0.04% racemate.[1] However, the need for care in azide couplings is apparent because the amount of racemate found depended on the technique employed, exposure to excess base at any stage being detrimental [114] (see also Section 3A). Under very carefully defined conditions the mixed carbonic anhydride procedure could be made to yield as little as 0.01% racemate, but slight deviations from exact stoicheiometry raised this by an order of magnitude.[114]

Kemp has also applied his modified racemisation tests to α-deuteriated carboxy components and obtained kinetic isotope effects for racemisation during coupling.[115] Isotope effects of unity were observed in the modified Young tests, which is consistent with a mechanism in which oxazolone formation is rate-determining. In contrast, modified Anderson tests on the mixed carbonic anhydride and azide methods gave isotope effects between 1.5 and 2.9, indicative of mechanisms in which at least some racemisation occurs by direct ionisation of the peptide α-hydrogen.

General experience over the last decade has shown that the Young test is about an order of magnitude more sensitive than the Anderson test, and this is confirmed by Kemp's [1] recent precise work. This is partly because oxazolones are formed with greater ease from benzoylamino-acids than from acylpeptides. Another factor is also important, however, since kinetic measurements [116] have demonstrated that the competition between racemisation and ring-opening favours racemisation to a greater extent with the oxazolone (43) than with (44). The kinetics and mechanisms of the reactions of (43) with α-nucleophiles (*i.e.* which have contiguous nucleophilic centres such as in hydrazine and hydroxylamine derivatives) have been studied in detail.[117]



(43)

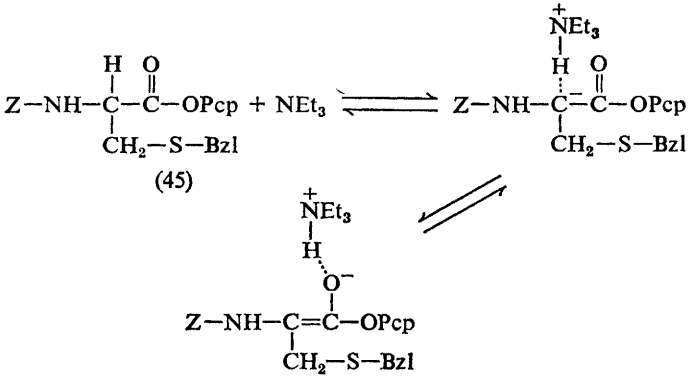


(44)

Further publications on the vexing question of the facile base-catalysed racemisation of active esters of benzyloxycarbonyl-*S*-benzyl cysteine have

[115] D. S. Kemp and J. Rebek, jun., *J. Amer. Chem. Soc.*, 1970, **92**, 5792.
[116] M. Goodman and C. B. Glaser, *Tetrahedron Letters*, 1969, 3473.
[117] M. Goodman and C. B. Glaser, *J. Org. Chem.*, 1970, **35**, 1954.

appeared.[118–121] The once favoured 'β-elimination–readdition mechanism' having been ruled out [119, 122] (see also Vol. 1, p. 192, of these Reports), a direct exchange mechanism appears to be implicated, but even this is not as simple as it might seem since the racemisation of (45) (in chloroform containing 7 equivalents of triethylamine and 50 equivalents of deuteriomethanol) is much faster than exchange of the α-hydrogen for deuterium under the same conditions. A 'conducted tour' mechanism was therefore suggested (Scheme 15).[121] A clear explanation of the rôle of the sulphur side-chain in the racemisation of esters such as (45) is still lacking.



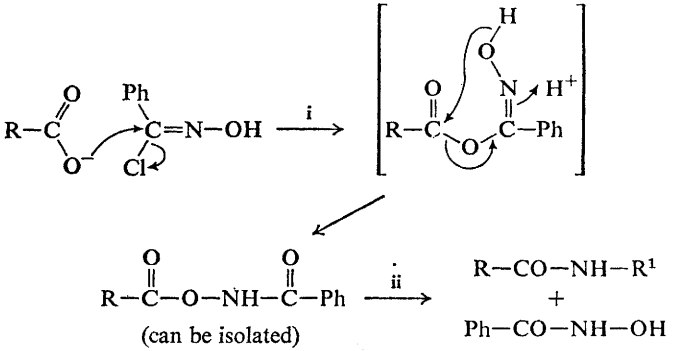
Scheme 15

The formation and coupling of esters of benzohydroxamic acid as shown in Scheme 16 have been claimed to give less than 0.1% racemisation in '2 + 1' tripeptide preparations.[81] However, this claim is based on examination for racemate by paper chromatography, and clearly needs to be corroborated by more reliable quantitative tests before further comment is in order.

Symposium lectures on the detection of racemisation by nuclear magnetic resonance [123] and by use of an amino-acid analyser [124] have appeared in print. Other recent papers relating to the racemisation question include coverage of the following: the determination of the optical purity of *N*-methylamino-acids;[125] the epimerisation by base of threonine deriva-

[118] J. Kovacs, G. L. Mayers, R. H. Johnson, and U. R. Ghatak, ref. 11, p. 337.
[119] J. Kovacs, G. L. Mayers, R. H. Johnson, R. E. Cover, and U. R. Ghatak, *J. Org. Chem.*, 1970, **35**, 1810.
[120] J. Kovacs, G. L. Mayers, R. H. Johnson, R. E. Cover, and U. R. Ghatak, *Chem. Comm.*, 1970, 53.
[121] G. L. Mayers and J. Kovacs, *Chem. Comm.*, 1970, 1145.
[122] J. Kovacs, G. L. Mayers, R. H. Johnson, and U. R. Ghatak, *Chem. Comm.*, 1968, 1066.
[123] B. Weinstein, ref. 11, p. 371.
[124] N. Izumiya and M. Muraoka, ref. 11, p. 347.
[125] J. R. Coggins and N. L. Benoiton, *J. Chromatog.*, 1970, **52**, 251.

tives;[126] asymmetric synthesis during peptide formation;[127, 128] the racemisation of amino-acid derivatives in acetic acid;[129] the effect of variations in a hydrocarbon side-chain on the base-catalysed racemisation of *N*-benzoyl-amino-acid anilides;[130] a modified Izumiya racemisation test;[131] racemisation control in the synthesis of peptides by the mixed carbonic anhydride



Conditions: i, the acid is used as its Ag or tertiary amine salt in an inert solvent; ii, R^1NH_2

Scheme 16

and dicyclohexylcarbodi-imide methods;[132] determination of the extent of racemisation in the mixed anhydride method by paper chromatography;[133] the use of the Weygand test for comparing the efficacy of various *N*-hydroxy-compound additives for suppressing racemisation in the dicyclohexyl-carbodi-imide method;[93] the g.l.c. of trifluoroacetylamino-acid alkyl esters on an optically active stationary phase,[134] and the use of this technique for proving the absence of racemisation in a solid-phase synthesis.[134] The racemisation question has been reviewed in Russian.[135]

The use of ephedrine for the resolution of benzyloxycarbonyl-DL-amino-acids has been described.[136] Further applications of the convenient tyrosine hydrazide method of Vogler *et al.*[137] include resolutions (all in

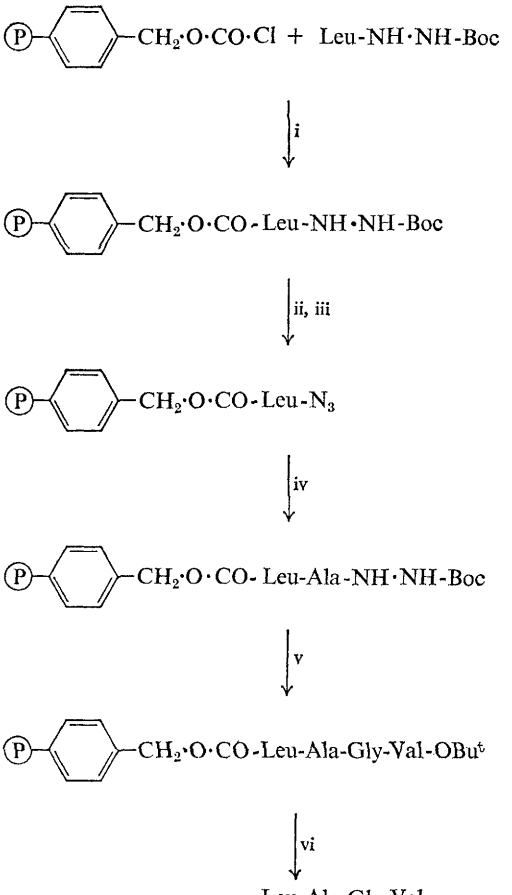
[126] T. Inui, *Sci. Rep. Osaka Univ.*, 1969, **18**, 19 (*Chem. Abs.*, 1970, **73**, 77582f).
[127] L. Ötvös, T. Tömösközi, and T. Mohácsi, *Tetrahedron Letters*, 1970, 1995.
[128] L. Ötvös, T. Mohácsi, I. Tömösközi, and É. Boromissza, *Radiochem. Radioanal. Letters*, 1970, **3**, 169.
[129] H. Matsuo, Y. Kawazoe, M. Sato, M. Ohnishi, and T. Tatsuno, *Chem. and Pharm. Bull. (Japan)*, 1970, **18**, 1788.
[130] M. Sato, T. Tatsuno, and H. Matsuo, *Chem. and Pharm. Bull. (Japan)*, 1970, **18**, 1794.
[131] F. H. C. Stewart, *Austral. J. Chem.*, 1970, **23**, 1073.
[132] G. W. Anderson, ref. 11, p. 255.
[133] E. Taschner, M. Smulkowski, and L. Lubiewska-Nakonieczna, *Annalen*, 1970, **739**, 228.
[134] E. Bayer, E. Gil-Av, W. A. König, S. Nakaparksin, J. Oró, and W. Parr, *J. Amer. Chem. Soc.*, 1970, **92**, 1738.
[135] L. A. Shchukina, ref. 3, p. 43.
[136] K. Oki, K. Suzuki, S. Tuchida, T. Saito, and H. Kotake, *Bull. Chem. Soc. Japan*, 1970, **43**, 2554.
[137] K. Vogler and P. Lanz, *Helv. Chim. Acta*, 1966, **49**, 1348.

excellent yield) of benzyloxycarbonyl-*S*-benzyl-DL-cysteine,[138] benzyloxycarbonyl-DL-pipecolic acid,[139] and benzyloxycarbonyl-DL-azetidine-2-carboxylic acid,[140] but an attempt to resolve benzyloxycarbonyl-DL-methionine in this way failed.[138]

E. Repetitive Methods of Peptide Synthesis.—*Solid-phase Synthesis.* The technique and its pitfalls have been reviewed.[141–143] Until recently the only tactic used for solid-phase peptide synthesis was that of stepwise construction working away from the *C*-terminal. The complementary reverse approach necessitates the activation of *C*-terminal carboxy-groups belonging to resin-bound peptides. Severe restrictions on the choice of coupling methods are therefore incurred, but Merrifield has now shown [144] that the use of the azide method in this strategy is feasible (Scheme 17). Another tactical variation is the fragment condensation technique, in which one fragment is in solution and the amino-component fragment on the resin. Such an approach carries the important advantage over the simple stepwise method that the end product is much easier to separate from erroneous sequences. The first report [145] of the use of an acylpeptide azide for coupling to an insolubilised peptide nucleophile was not encouraging, and so it appeared that the flexibility of synthetic design would be hampered. However, a successful coupling of a peptide azide with an amino-component anchored to a resin has now been described (Scheme 18).[146] Of course, if junction between fragments can be made at glycine or proline, dicyclohexylcarbodi-imide can be used (*e.g.* Scheme 19).[147] It has been pointed out [147] that the fragment condensation strategy provides a means of minimising the number of acidolytic deprotection stages and therefore reducing the problems [148] which attend the solid-phase synthesis of tryptophan-containing peptides (Scheme 19). On the debit side, however, there are economic arguments against the use of acylpeptides as carboxy-components in solid-phase procedures, since a generous excess is commonly used to maximise coupling, and recovery is difficult or impossible.

Bayer has discussed [148, 149] the problems which arise because of incomplete coupling in solid-phase work. He distinguishes two types of contaminants in the final products: 'truncated sequences' (*i.e.* peptides in

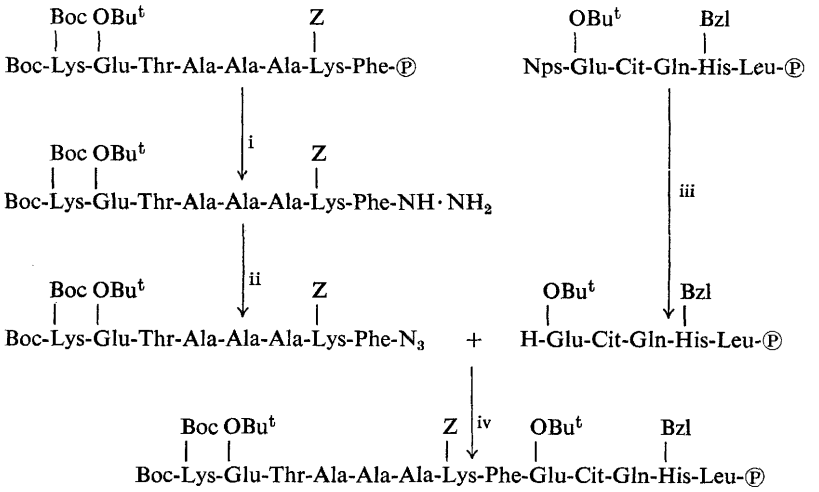
[138] B. Liberek and S. Dziala, *Zeszyty Nauk. Wyzszej Szkoly Pedagogicznej Gdansku Mat., Fiz., Chem.*, 1969, **9**, 165 (*Chem. Abs.*, 1970, **72**, 32184r).
[139] L. Balaspiri, B. Penke, J. Petres, and K. Kovacs, *Monatsh.*, 1970, **101**, 1177.
[140] R. M. Rodebaugh and N. H. Cromwell, *J. Heterocyclic Chem.*, 1969, **6**, 993.
[141] G. Losse and K. Neubert, *Z. Chem.*, 1970, **10**, 48.
[142] E. Bayer, *Chemie für Labor und Betrieb*, 1969, 193.
[143] E. Bayer, ref. 11, p. 99.
[144] A. M. Felix and R. B. Merrifield, *J. Amer. Chem. Soc.*, 1970, **92**, 1385.
[145] G. S. Omenn and C. B. Anfinsen, *J. Amer. Chem. Soc.*, 1968, **90**, 6571.
[146] S. Visser and K. E. T. Kerling, *Rev. Trav. chim.*, 1970, **89**, 880.
[147] H. Yajima, H. Kawatani, and H. Watanabe, *Chem. and Pharm. Bull.* (*Japan*), 1970, **18**, 1333.
[148] G. R. Marshall, *Adv. Exptl Med. and Biol.*, 1967, **2**, 48.
[149] E. Bayer, H. Eckstein, K. Hägele, W. A. König, W. Brüning, H. Hagenmaier, and W. Parr, *J. Amer. Chem. Soc.*, 1970, **92**, 1735.



Conditions: i, $CHCl_3$–NEt_3; ii, 4M-HCl–dioxan; iii, Bu^nONO; iv, Ala–NH·NH–Boc; v, similarly – either stepwise or by addition of a dipeptide fragment; vi, HBr–CF_3CO_2H

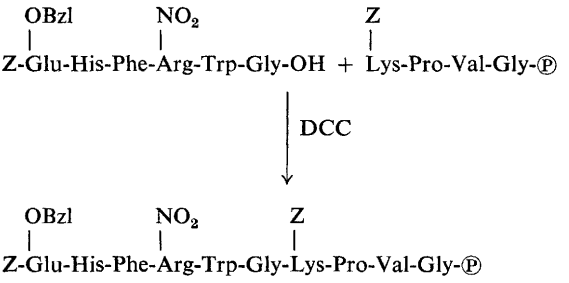
Scheme 17

which growth ceased at some stage) and 'failure sequences' (*i.e.* in which one coupling failed but some subsequent steps succeeded). The repeating dipeptides $(Leu\text{-}Ala)_6$ and $(Ala\text{-}Phe)_6$ were synthesised in a stepwise manner, and their partial hydrolysates were examined for the presence of dipeptides corresponding to 'failure sequences' (*i.e.* Ala-Ala, Leu-Leu and Ala-Ala, Phe-Phe respectively).[149] In both cases definite qualitative evidence for the formation of 'failure sequences' was obtained, and it was shown that acetylation of residual amino-groups after each coupling step (*i.e.* in Bayer's terminology, preventing the development of 'failure sequences' from 'truncated sequences') reduced greatly the proportion of 'failure



Conditions: i, $NH_2 \cdot NH_2$; ii, Bu^{t}ONO; iii, 0.1N-HCl in AcOH–$CHCl_3$; iv, DMF, −20 °C; Cit = citrulline

Scheme 18



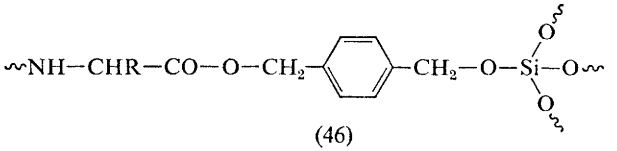
Scheme 19

sequences'. The use of glass beads covered with a thin layer of resin rather than the conventional uniform resin beads was also advantageous in this respect.[149] It seems that elimination of the need for penetration of a gel matrix by the reactant in solution permits more complete coupling. One other system in which the growing peptide is linked to the surface of a polymer [through a benzyl ester attached to silica (46)] has been briefly described.[150] The amounts of 'failure sequences' consequent upon using (46) as support in a stepwise synthesis of $(Leu\text{-}Ala)_6$ were reported ambiguously [150] as being 'well beyond the figures found for resin synthesis'. In principle, the use of a solid-phase support in which all the points of attachment protrude from an insoluble particle would have several highly desirable features, despite the lower capacity of such supports *viz.* (i) more

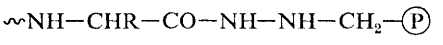
[150] E. Bayer, G. Jung, I. Halasz, and I. Sebastian, *Tetrahedron Letters*, 1970, 4503.

nearly complete coupling, (ii) supports such as (46) which are not subject to variable swelling in different solvents could be used in columns [150] – this would eliminate the need for shaking with concomitant pulverisation of resin particles into fines which are lost through sintered filters. These preliminary investigations [149, 150] are encouraging but further work and the publication of experimental detail must precede a full appraisal.

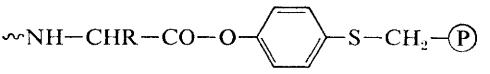
The potential of solid-phase synthesis as a means of obtaining protected oligopeptide intermediates suitable for further classical application has been discussed in earlier Reports (Vol. 2, p. 162). Hydrazinolytic removal of peptides from conventional Merrifield resins has now been used several times (see refs. 151 and 152 for some recent examples) and ammonolysis has



(46)



(47)



(48)

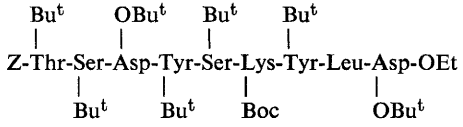
proved valuable for the preparation of synthetic precursors of oxytocin and its analogues. More subtle requirements such as a free *C*-terminal carboxy-group in an otherwise fully-protected peptide can only be met by use of a solid-phase support from which the peptide can be removed under selective conditions. Two recently introduced modified supports suitable for this application are both of the 'safety-catch' type: (47) can be oxidatively activated towards nucleophiles by treatment with *N*-bromosuccinimide [43] and the same end is achieved by treating (48) with hydrogen peroxide in acetic acid [153] or *p*-chloroperoxybenzoic acid in dichloromethane [154] (*cf.* ref. 85).

Other relevant papers cover the following topics: the choice of terminating agents;[155] choice of solvent for coupling;[156] a comparative evaluation

[151] S. Visser, J. Raap, K. E. T. Kerling, and E. Havinga, *Rec. Trav. chim.*, 1970, **89**, 865.
[152] S. Visser, K. E. T. Kerling, and E. Havinga, *Rec. Trav. chim.*, 1970, **89**, 876.
[153] D. L. Marshall and I. E. Liener, *J. Org. Chem.*, 1970, **35**, 867.
[154] D. L. Marshall, personal communication.
[155] L. D. Markley and L. C. Dorman, *Tetrahedron Letters*, 1970, 1787.
[156] F. C. Westall and A. B. Robinson, *J. Org. Chem.*, 1970, **35**, 2842.

of ten *S*-protecting groups for solid-phase applications;[55] side-reactions occurring on acidolysis with hydrogen chloride in dioxan;[157] proof of the absence of racemisation in conventional solid-phase procedures;[134] halogeno-acylpolystyrene supports;[158] amino-acyl insertion reactions in solid-phase synthesis;[159] the reactivity of *p*-nitrophenyl esters;[160] the cleavage of t-butoxycarbonyl groups;[161] the influence of chain length on the completeness of coupling;[162] technical details of an automated solid-phase synthesis system controlled by punched tape;[163, 164] process control by determination of free *N*-terminal amino-groups;[165–167] the attachment of glycine residues to dextran[168] and to a modified styrene-divinylbenzene copolymer;[169] the coupling of adsorbed carboxy-components;[167] and the optimum conditions for acidolytic scission from the resin support.[170]

Other Methods. A rapid repetitive stepwise procedure using the mixed carbonic anhydride method has been outlined.[171] Good yields (> 96%) were obtained at each age of the synthesis of a complex protected nona-peptide glucagon fragment (49), residues being added at an average rate



(49)

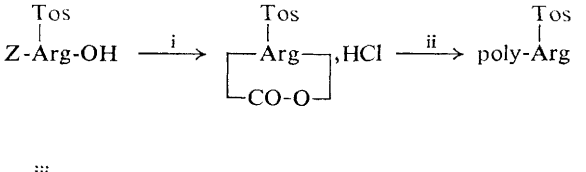
of one per day. Whilst the claim that 'No other method of peptide synthesis has shown the results which are being reported here' seems a trifle extravagant, sufficient experimental detail is given to show that this may prove to be a useful additional technique for constructing protected fragments of medium size. The repetitive stepwise application of active esters also gives near-quantitative yields and can be performed so as to add a residue every two days or so in the synthesis of complex peptides at the

[157] K. P. Polzhofer and K. H. Ney, *Tetrahedron*, 1970, **26**, 3787.
[158] T. Mizoguchi, K. Shigezane, and N. Takamura, *Chem. and Pharm. Bull. (Japan)*, 1970, **18**, 1465.
[159] A. R. Mitchell and R. W. Roeske, *J. Org. Chem.*, 1970, **35**, 1171.
[160] S. Karlsson, G. Lindeberg, and U. Ragnarsson, *Acta Chem. Scand.*, 1970, **24**, 337.
[161] S. Karlsson, G. Lindeberg, J. Porath, and U. Ragnarsson, *Acta Chem. Scand.*, 1970, **24**, 1010.
[162] H. Hagenmaier, *Tetrahedron Letters*, 1970, 283.
[163] K. Brunfeldt, J. Halstrøm, and P. Roepstorff, *Acta Chem. Scand.*, 1969, **23**, 2830.
[164] K. Brunfeldt, *Dan. Kemi*, 1969, **50**, 169 (*Chem. Abs.*, 1970, **72**, 79458y).
[165] K. Brunfeldt, P. Roepstorff, and J. Thomsen, *Acta Chem. Scand.*, 1969, **23**, 2906.
[166] E. Kaiser, R. L. Colescott, C. D. Bossinger, and P. I. Cook, *Analyt. Biochem.*, 1970, **34**, 595.
[167] K. Esko and S. Karlsson, *Acta Chem. Scand.*, 1970, **24**, 1415.
[168] N. K. Kochetkov, A. A. Khachatur'yan, A. E. Vasil'ev, and G. Ya. Rozenberg, *Khim. prirod Soedinenii*, 1969, **5**, 427 (*Chem. Abs.*, 1970, **72**, 90850r).
[169] M. Buka and R. Zagats, *Latv. P.S.R. Zinat Akad. Vestis, Kim. Ser.*, 1969, 503 (*Chem. Abs.*, 1970, **72**, 90857y).
[170] J. Scotchler, R. Lozier, and A. B. Robinson, *J. Org. Chem.*, 1970, **35**, 3151.
[171] M. A. Tilak, *Tetrahedron Letters*, 1970, 849.

decapeptide level and above.[172] Although time is money it is not necessarily the essence of the matter, and the repetitive active ester method has substantial advantages over the rapid anhydride procedure in other respects. Thus active esters can be prepared very simply, purified with ease, and stored practically indefinitely, whereas mixed anhydrides must be prepared with care and used immediately without purification. Furthermore, active esters are less subject to side-reactions and can be used for acylation under a more versatile range of conditions than can mixed anhydrides.

The 4-picolyl ester repetitive method has been further evaluated by its use in a synthesis of bradykinin.[173] The protected precursor was obtained in *ca.* 40% overall yield based on the first residue introduced, and hydrogenation gave bradykinin of satisfactory amino-acid analysis, having on the guinea-pig ileum 87% as much biological activity as the Sigma Research Standard. A number of variations on the original procedure have been examined using 5-valine-angiotensin II as a test case.

F. Synthesis of Polymeric Models for Studies in Protein Chemistry.—*Polyamino-acids.* A further synthesis (see Vol. 2, p. 168, of these Reports) of polyarginine has been described (Scheme 20).[174] The molecular weight



Conditions: i, $SOCl_2$; ii, Ag_2O–acetone, then triethylene diamine in dimethylacetamide; iii, HF then IRA–400 resin (chloride form)

Scheme 20

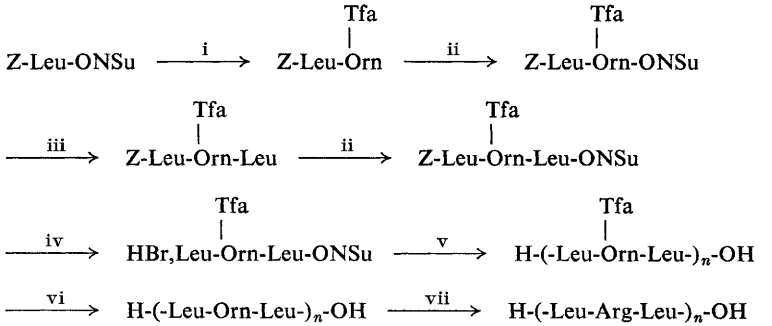
was greatest when the polymerisation was performed in dimethylacetamide, but the figure of 16 000 must be cautiously assessed in view of the fact that it was determined by carboxyl end-group titration.[174]

Other topics covered include: the kinetics of *N*-carboxy-anhydride polymerisation in DMSO;[175] the polymerisation of *N*-carboxy-anhydrides by organo-zinc compounds;[176] copolyamino-acid preparation from *N*-carboxy-anhydrides;[177] copoly-glutamic acid–phenylalanine with apparent lysozyme activity;[178] the preparation of copolymers containing 18

[172] *E.g.*, J. S. Morley, *J. Chem. Soc.* (*C*), 1967, 2410.
[173] R. Garner, D. J. Schafer, and G. T. Young, ref. 12, p. 102.
[174] J. Noguchi and Y. Fujiwara, *Bull. Chem. Soc. Japan*, 1970, **43**, 2515.
[175] M. Oya, K. Uno, and Y. Iwakura, *J. Polym. Sci., Part A-1, Polymer Chem.*, 1970, **8**, 1851.
[176] T. Makino, S. Inoue, and T. Tsurata, *Makromol. Chem.*, 1970, **131**, 147.
[177] M. Oya, M. Ohta, and H. Kato, *Bull. Chem. Soc. Japan*, 1970, **43**, 1788.
[178] S. Srivastava, K. B. Mathur, and M. M. Dhar, *Experientia*, 1970, **26**, 11.

protein amino-acids and their possible use as nutrients;[179] chain branching in poly-β-alanine obtained by polymerisation of acrylamide;[180] the preparation of phenyl-substituted poly-β-amino-acids,[181] the base-catalysed polymerisation of *p*-nitrophenoxycarbonylamino-acids;[182] syntheses of stereoisomers of poly-γ-glutamic acid;[183, 184] chemical modification of polyamino-acid side-chains;[185] the preparation of polyamino-acids *via* isothiocyanato-carboxylic acids;[186] the preparation of *N*-carboxy-anhydrides *via* amino-acid trimethylsilylamino-acids;[187] and the preparation of block copolyamino-acids.[188]

Sequential Polypeptides. A sequential polypeptide containing arginine has been prepared *via* the corresponding ornithine-containing polymer as shown in Scheme 21.[189] Johnson and his co-workers have described



Conditions: i, Orn(Tfa); ii, DCCI–HONSu; iii, Leu; iv, HBr–AcOH; v, Et_3N–DMF; vi, piperidine; vii, 1-guanidino-3,5-dimethyl pyrazole; DCCI = dicyclohexylcarbodiimide

Scheme 21

further examples[190, 191] of their high-dilution polymerisation technique (see Vol. 2, p. 171, of these Reports), and Blout and his colleagues have reported on a detailed study[192] of the synthesis (through a succinimido

[179] G. Krampitz, S. Baars-Diehl, W. Haas, and P. Anda, *Zeitschrift für Tierphysiologie, Tierernährung und Futtermittelk*, 1970, **26**, 33 (*Chem. Abs.*, 1970, **73**, 15214d).
[180] J. D. Glickson and J. Applequist, *Macromolecules*, 1969, **2**, 628.
[181] H. Yuki, Y. Taketani, S. Yamashita, H. Ohuno, and H. Tanaka, *Bull. Chem. Soc. Japan*, 1970, **43**, 1855.
[182] P. Baudet, C. Otten, and D. Rao, *Helv. Chim. Acta*, 1970, **53**, 859.
[183] M. Hollosi, M. Kajtar, and V. Bruckner, *Acta Chim. Acad. Sci. Hung.*, 1969, **62**, 305.
[184] M. Kajtar and V. Bruckner, *Acta Chim. Acad. Sci. Hung.*, 1969, **62**, 191.
[185] A. Kotai, G. Szokan, I. Ferencz, and M. Almas, *Acta Chim. Acad. Sci. Hung.*, 1969, **62**, 293.
[186] H. R. Kricheldorf, *Angew. Chem. Internat. Edn.*, 1970, **9**, 526.
[187] H. R. Kricheldorf, *Chem. Ber.*, 1970, **103**, 3353.
[188] M. Oya, *Bull. Chem. Soc. Japan*, 1970, **43**, 1788.
[189] M. Fridkin, A. Frenkel, and S. Ariely, *Biopolymers*, 1969, **8**, 661.
[190] B. J. Johnson and D. S. Rea, *Canad. J. Chem.*, 1970, **48**, 2509.
[191] B. J. Johnson and E. G. Trask, *J. Chem. Soc.* (*C*), 1970, 2247.
[192] B. B. Doyle, W. Traub, G. P. Lorenzi, F. R. Brown, tert., and E. R. Blout, *J. Mol. Biol.*, 1970, **51**, 47.

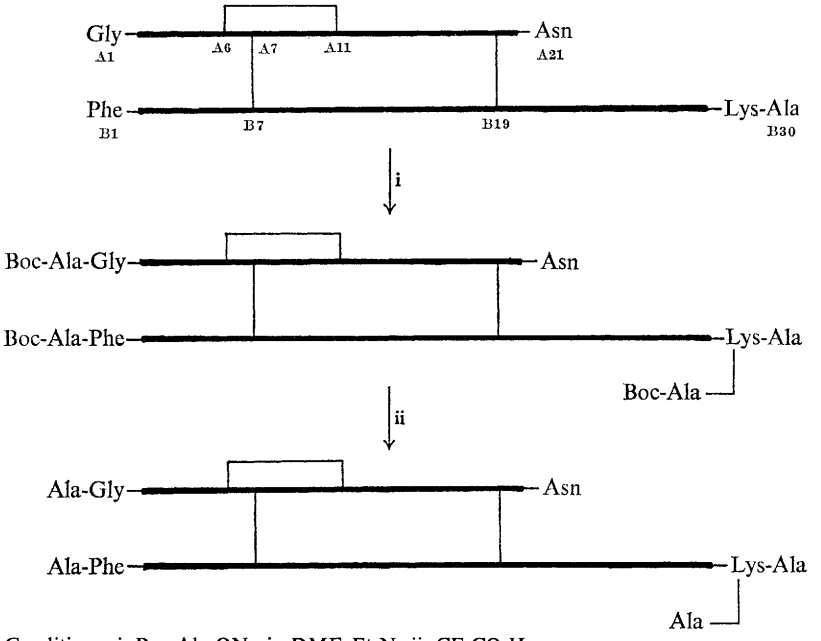
ester) and properties of poly-alanylalanylglycine. Poly-alanylalanylglycine has glycine at every third position, and is therefore a collagen model in some respects, but collagen-like conformations were not detected in the solid state or in solution, underlining the essentiality of imino-acids for triple helical folding. The gel-filtration behaviour of poly-D-aspartyl-lysine has been discovered to be critically dependent on ionic strength.[193]

A large volume of literature [194–208] has issued from the laboratories of Shibnev, Poroshin, and their colleagues.

G. Synthetic Operations with Peptides of Biological Origin.—In this section, only work involving controlled formation of new peptide bonds will be included.

With Naked Natural Peptides. Several insulin analogues with additional amino-acid residues have been prepared by reaction of the hormone with activated t-butoxycarbonylamino-acids followed by trifluoroacetic acid deprotection. Use of t-butoxycarbonylamino-acid *p*-nitrophenyl esters in DMF gave a series of triaminoacylinsulins (*e.g.* Scheme 22) in which the same new amino-acid residue was attached to each of the amino-groups of the original hormone (*viz.* glycine-A1, phenylalanine-B1, and the side-chain of lysine-B29).[209] By use of t-butoxycarbonylamino-acid phenyldiimides as acylating agents, diaminoacylinsulins with largely unattacked

[193] A. S. Akerkar, *J. Chromatog.*, 1970, **51**, 543.
[194] V. A. Shibnev, Sh. Kh. Khalikov, M. P. Finogenova, and K. T. Poroshin, *Izvest. Akad. Nauk. S.S.S.R., Ser. khim.*, 1970, 880 (*Chem. Abs.*, 1970, **73**, 35728c).
[195] V. A. Shibnev, T. P. Chuvaeva, and K. T. Poroshin, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1969, 2527 (*Chem. Abs.*, 1970, **72**, 79465y).
[196] V. A. Shibnev, Sh. Kh. Khalikov, M. P. Finogenova, and K. T. Poroshin, *Izvest. Akad. Nauk. S.S.S.R., Ser. khim.*, 1970, 399 (*Chem. Abs.*, 1970, **73**, 4182r).
[197] V. A. Shibnev, T. P. Chuvaeva, G. A. Martynova, and K. T. Poroshin, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1969, 2532 (*Chem. Abs.*, 1970, **72**, 79467a).
[198] V. A. Shibnev, T. P. Chuvaeva, and K. T. Poroshin, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1970, 121 (*Chem. Abs.*, 1970, **72**, 12191y).
[199] K. T. Poroshin, T. P. Chuvaeva, and V. A. Shibnev, *Doklady Akad. Nauk Tadzh. S.S.R.*, 1969, **12** (11), 15.
[200] K. T. Poroshin, L. I. Mar'yash, V. S. Grechishko, and V. A. Shibnev, *Doklady Akad. Nauk Tadzh. S.S.R.*, 1970, **13** (4), 19.
[201] K. T. Poroshin, T. P. Chuvaeva, and V. A. Shibnev, *Doklady Akad. Nauk Tadzh. S.S.R.*, 1969, **12** (7), 28.
[202] K. T. Poroshin, T. P. Chuvaeva, and V. A. Shibnev, *Doklady Akad. Nauk. Tadhz. S.S.R.*, 1969, **12** (8), 21.
[203] K. T. Poroshin, A. B. Zegel'man, T. U. Yusupov, and G. N. Demyanik, *Doklady Akad. Nauk Tadzh. S.S.R.*, 1969, **12** (10), 28.
[204] T. Yu. Yusupov, A. B. Zegel'man, Z. Sharifora, and K. T. Poroshin, *Doklady Akad. Nauk Tadzh. S.S.R.*, 1969, **12** (9), 32.
[205] T. Yu. Yusupov, A. B. Zegel'man, and K. T. Poroshin, *Doklady Akad. Nauk Tadzh. S.S.R.*, 1969, **12** (4), 27.
[206] A. B. Zegel'man, T. Yu. Yusupov, and K. T. Poroshin, *Doklady Akad. Nauk Tadzh. S.S.R.*, 1970, **13** (3), 22.
[207] A. B. Zegel'man, T. Yu. Yusupov, and K. T. Poroshin, *Doklady Akad. Nauk Tadzh. S.S.R.*, 1970, **13** (2), 25.
[208] K. T. Poroshin, T. P. Chavaeva, and V. A. Shibnev, *Doklady Akad. Nauk Tadzh. S.S.R.*, 1969, **12** (12), 21.
[209] D. Levy and F. H. Carpenter, *Biochemistry*, 1967, **6**, 3559.

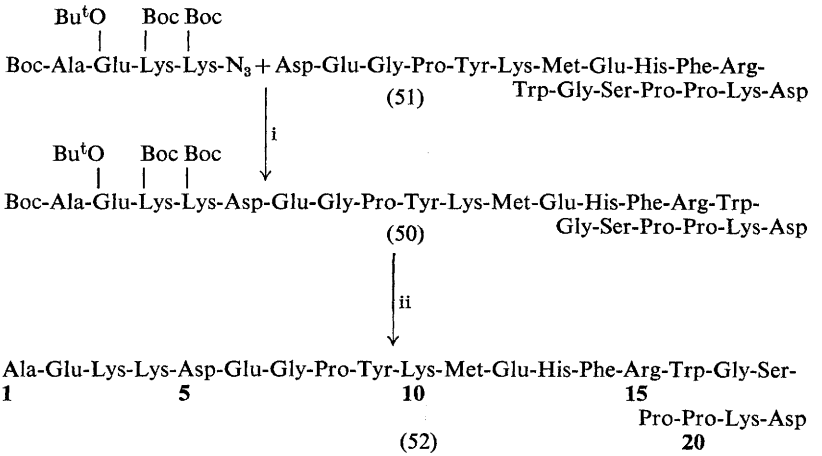


Conditions: i, Boc-Ala-ONp in DMF–Et_3N; ii, CF_3CO_2H

Scheme 22

ε-amino-groups could be obtained.[210] With single-chain peptides which lack lysine, t-butoxycarbonylamino-acid active esters can be employed for the unambiguous introduction of an extra residue at the *N*-terminus: recent examples are provided by a series of aminoacyl-apoferredoxins [211] (which were converted to ferredoxin analogues), and extended insulin A chains.[212] However, all these cases are relatively trivial as far as the synthetic aspects are concerned, and of much greater interest is the conversion of natural porcine β-melanotropin into the [10-lysine]-human hormone (Scheme 23) by Burton and Lande:[213] although the natural starting material contains three amino-groups it was possible to isolate the desired analogue in partially protected form by ion-exchange chromatography. The partially protected product (50) had a good amino-acid analysis [there are obvious large differences between the starting material (51) and product (50)], and was shown by dinitrophenylation to have two free lysine ε-amino-groups in accordance with expectation. Removal of the t-butyl-containing protecting groups gave the chromatographically homogeneous free peptide (52) which had practically the same biological potency as fully synthetic

[210] H. B. Milne and F. H. Carpenter, *J. Org. Chem.*, 1968, **33**, 4476.
[211] J.-S. Hong and J. C. Rabinowitz, *J. Biol. Chem.*, 1970, **245**, 4988.
[212] M. Weinert, D. Brandenburg, and H. Zahn, *Z. physiol. Chem.*, 1969, **350**, 1556.
[213] J. Burton and S. Lande, *J. Amer. Chem. Soc.*, 1970, **92**, 3746.



Conditions: i, 50% py, 5 °C, 4 h; ii, anhydrous HCO_2H

Scheme 23

human β-melanotropin. The overall yield in this partial synthesis of (52) was only about 4%, but material with the tetrapeptide adjunct on one of the lysine side-chains was also isolated, and the unreacted porcine hormone was recovered in 50% yield.

With Partially Blocked Peptides. Offord's reports on the progress of his approach to partial synthesis, which involves protection of natural peptides before coupling, comprise experiments on the differential protection of the amino-groups of insulin [214] and of the two carboxy-groups of a lysozyme fragment.[215]

3 Syntheses Achieved and Structure–Activity Correlations

As in the two previous volumes, only a limited coverage of a small number of syntheses will be given in this section: the Reader is referred to Appendix A for a comprehensive list of syntheses described during the year.

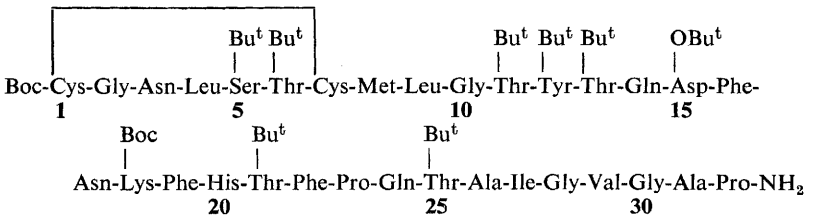
Rudinger [216] has written a detailed and critical review of studies on synthetic polypeptide hormone analogues, with emphasis on the rational design of analogues. A total of 630 references are dealt with in masterly style, and this review will no doubt influence research in this area for a considerable time.

An authoritative review summarising current views on hypothalamic-releasing factors has been published,[217] and useful summaries of important

[214] F. Borrás and R. E. Offord, *Nature*, 1970, **227**, 716.
[215] C. F. Hayward and R. E. Offord, ref. 12, p. 116.
[216] J. Rudinger, in 'Drug Design,' ed. E. J. Ariens, Academic Press, New York, Vol. 2 (in press).
[217] R. Burgus and R. Guillemin, *Ann. Rev. Biochem.*, 1970, **39**, 499.

work on the neurohypophysial peptide hormones,[218] angiotensin II,[219, 220] gastrin,[221] and their analogues have appeared. Studies with synthetic analogues of angiotensin II have led to a working hypothesis [222] about the mechanism of receptor interaction. Synthetic work on insulin has been reviewed.[223]

A. Calcitonin.—Full details of the total synthesis of the human hypocalcaemic hormone 'calcitonin M' have been published:[224] the detailed exposition contains important discussion of the difficulties and by-products encountered which is not to be found in the preliminary communication.[225] The penultimate stage of the synthesis was the fully-protected calcitonin M derivative (53), which was constructed by a fragment condensation strategy. Fragment conjunctions were performed not only at the glycine and proline residues but also, by the azide method, between residues 16 and 17 and



(53)

between positions 20 and 21. Both azide couplings resulted in serious racemisation (20% in one case under the worst conditions) if carried out in the presence of excess base. This was revealed by amino-acid oxidase studies and confirmed by the actual isolation by counter-current distribution of peptides containing D-amino-acid residues. The extent of racemisation was much reduced by use of an exact amount of base for neutralisation of the azide preparation and for liberation of the free nucleophile of the amino-component: ethyldi-isopropylamine was superior to triethylamine. However, even careful attention to these matters did not reduce racemisation to below 1—2%. Kemp's recent emphatic warning [114] about the caution which must be exercised to preserve optical integrity in manipulations of acylpeptide azides was based on studies with simple models. The wisdom of this advice has now been demonstrated in actual experience.

[218] (*a*) V. du Vigneaud, in 'Perspectives in Biological Chemistry,' ed. R. E. Olson, Marcel Dekker, New York, 1970, p. 133; (*b*) K. Jošt and J. Rudinger, ref. 13, p. 13.
[219] F. M. Bumpus, R. R. Smeby, and P. A. Khairallah, ref. 11, p. 127.
[220] E. C. Jorgensen, G. C. Windridge, W. Patton, and T. C. Lee, ref. 11, p. 113.
[221] M. I. Grossman, *Nature*, 1970, **228**, 1147.
[222] P. A. Khairallah, A. Toth, and F. M. Bumpus, *J. Medicin. Chem.*, 1970, **13**, 181.
[223] K. Lübke and H. Klostermeyer, *Adv. Enzymol.*, 1970, **33**, 445.
[224] P. Sieber, B. Riniker, M. Brugger, B. Kamber, and W. Rittel, *Helv. Chim. Acta*, 1970, **53**, 2135.
[225] P. Sieber, M. Brugger, B. Kamber, B. Riniker, and W. Rittel, *Helv. Chim. Acta*, 1968, **51**, 2057.

In view of these results with calcitonin M, it seems possible that many other important syntheses of recent years have incurred undetected racemisation because of confidence in fragment condensation strategies employing the azide method. To cite an important example, the published synthesis [226] of ribonuclease-S′ involved no less than sixteen azide couplings with optically active carboxy-components. Perhaps the cumulative effect of a few percent racemisation at some of these stages is one of the factors responsible for the low activity (*ca.* 30%) of the synthetic ribonuclease S′.

Fortunately, counter-current distribution at various stages sufficed to remove the diastereoisomeric contaminants, and the fully-blocked precursor (53) of calcitonin M was obtained in a pure condition. Even at this juncture all the hurdles had not been cleared, as acidolysis of the protecting groups of (53) by brief exposure to cold concentrated hydrochloric acid caused two side-reactions. One of these was the formation of a sulphonium derivative by reaction of the methionine side-chain with a t-butyl carbonium ion, and the other was migration of the peptide chain from nitrogen to oxygen at one or more of the hydroxyamino-acid residues. Again the power of counter-current distribution came to the rescue, and hormone which was pure except for a small amount of sulphoxide was finally obtained in an amount of no less than 1.6 g. Incubation of a total hydrolysate with L-amino-acid oxidase left unoxidised D-amino-acids in traces, but the amounts were, within experimental error, the same as were formed on total hydrolysis of the natural peptide. The purified synthetic material had the full biological activity of native calcitonin M.

B. Cholecystokinin-pancreozymin (CCK-PZ).*—Extracts of intestinal mucosa contain a tritriacontapeptide amide which induces gall bladder contraction and stimulates the release of pancreatic enzymes. The clumsy name of this substance results from the fact that the two biological activities were discovered on separate occasions and attributed to hormones named 'cholecystokinin' and 'pancreozymin' which were only found to be identical much later. Studies on this hormone have been reviewed.[227, 228]

Partial tryptic digestion of CCK-PZ gives a dodecapeptide amide (54) from the *C*-terminus which has been synthesised,[229] but of more interest is the octapeptide amide (55) produced by complete tryptic digestion, which has also been synthesised.[229] All of the biological activities which can at

[226] R. Hirschmann, R. F. Nutt, D. F. Veber, R. A. Vitali, S. L. Varga, T. A. Jacob, F. W. Holley, and R. G. Denkewalter, *J. Amer. Chem. Soc.*, 1969, **91**, 507, and the four immediately preceding communications.

[227] M. A. Ondetti, J. T. Sheehan, and J. Pluščec, ref. 11, p. 181.

[228] M. A. Ondetti, B. Rubin, S. L. Engel, J. Pluščec, and J. T. Sheehan, *Amer. J. Digest. Diseases*, 1970, **15**, 149.

[229] M. A. Ondetti, J. Pluščec, E. F. Sabo, J. T. Sheehan, and N. Williams, *J. Amer. Chem. Soc.*, 1970, **92**, 195.

* Since it has already been used in several publications, the Reporter feels obliged to adopt the cumbersome and typographically ugly abbreviation CCK-PZ for cholecystokinin-pancreozymin.

$$\begin{array}{c} SO_3H \\ | \\ \text{Ile-Ser-Asp-Arg-Asp-Tyr-Met-Gly-Trp-Met-Asp-Phe-NH}_2 \end{array}$$

(54)

$$\begin{array}{c} SO_3H \\ | \\ \text{Asp-Tyr-Met-Gly-Trp-Met-Asp-Phe-NH}_2 \end{array}$$

(55)

present be ascribed to CCK-PZ reside in the sequence (55), which is very similar to the *C*-terminal octapeptide of gastrin II and to that of caerulin, differing by only one residue in each case.

The synthesis of *O*-sulphated tyrosine peptides presents a number of interesting but difficult problems.[227, 229, 230] The acid-lability of sulphate esters and the presence of sulphur would restrict protective tactics so much that it would be difficult to devise a synthesis in which the sulphate ester was carried through the construction of a complex peptide. At least, this was so at the time the published syntheses[229, 230] were designed, but the fact that exceedingly acid-labile *N*-protecting groups are now available may make this strategy feasible. However, the alternative approach of introducing the sulphate ester at a very late stage was adopted for the syntheses under discussion.

The octapeptide amide (55) was prepared[229] by treating the corresponding unsulphated free peptide with concentrated sulphuric acid at − 5 °C for a short period. The reaction conditions were critical: the low temperature and short reaction time were essential not only to avoid degradation (particularly at the acid-sensitive residues of aspartic acid and tryptophan) but also because the *O*-sulphate is the product of kinetic control. Under conditions of thermodynamic control the tyrosine was sulphonated in the ring. Fortunately, the *O*-sulphate, 3′-sulphonate, and unreacted peptide were easily separated by ion-exchange chromatography, and (55) was obtained in about 25% yield. The complex which sulphur trioxide forms with pyridine is a milder reagent for *O*-sulphation, but since it reacts with free amino-groups to give sulphamic acids it could not be employed on free peptides. However, an alternative route to (55) was found in which sulphur trioxide–pyridine was used at the penultimate stage as shown in Scheme 24. The sulphate ester proved sufficiently resistant to brief exposure

$$\underset{(56)}{\text{Boc-Asp-Tyr-Met-Gly-Trp-Met-Asp-Phe-NH}_2} \xrightarrow[\text{ii}]{\text{i}} (55)$$

Conditions: i, py–SO_3 in py–DMF; ii, CF_3CO_2H

Scheme 24

[230] J. Pluščec, J. T. Sheehan, E. F. Sabo, N. Williams, O. Kocy, and M. A. Ondetti, *J. Medicin. Chem.*, 1970, **13**, 349.

to trifluoroacetic acid to permit selective removal of the t-butoxycarbonyl group. The overall yield of (55) from its *N*-protected precursor (56) was about 30% after purification, and this tactic seems to have been preferred in more recent work [230] on analogues of (55), although rigorous control of solvent purity is essential. The *C*-terminal dodecapeptide amide (54) of CCK-PZ contains a serine residue and so in this case [229] sulphation was necessarily performed at the decapeptide level.

A series of peptides related to (54) and (55) have been synthesised and some of the critical features for cholecystokinic activity are already apparent.[228–230] Large decreases in activity result if the tyrosine residue is not sulphated or if the position of the tyrosine *O*-sulphate residue is changed. The aspartic acid residue adjacent to phenylalanine is very important for biological activity, as in the case of gastrin. On the other hand, in contrast to results with gastrin, substitution of leucine for the methionine residue which is next to tryptophan causes a considerable decrease in potency. The *C*-terminal heptapeptide sequence of CCK-PZ is the smallest peptide with significant activity, and the activity (on a weight basis) increases as the chain is extended, reaching a maximum of about ten times that of the natural hormone at about the decapeptide level. Even on a molar basis this is about two and a half times the activity of the whole hormone. Although various bits and pieces of several peptide hormones and even of some enzymes can be removed without great loss of activity, no other case of potentiation by such truncation has been reported.

C. Ribonuclease Systems.—It has been known for some years that bovine pancreatic ribonuclease gives, on limited proteolysis with subtilisin, fragments corresponding to residues 1—20 ('S-peptide') and 21—124 ('S-protein') respectively. These two fragments are separately inactive but when mixed form an enzymically active complex ('ribonuclease S'). It has recently been shown [231] that a similar operation can be performed at the opposite end of the sequence. Peptic hydrolysis of the intact enzyme followed by carboxypeptidase A treatment gives inactive des-(119—124)-ribonuclease, but most of the activity can be regenerated by adding a synthetic tetradecapeptide corresponding to the sequence 111—124. The complex thus formed has an overlapping sequence of eight residues and is therefore analogous to the enzymically active complex [232] formed by mixing peptides corresponding to the sequences 49—149 and 1—126 of staphylococcal nuclease. Presumably one of the duplicated sequences is superfluous: in the nuclease complex [232] just mentioned the 49—149 peptide binds only to residues 6—48 of the 1—126 peptide, the remainder protruding from the organised structure in a random manner. The scission of peptide bonds in ribonuclease can be taken yet one stage further with retention of high activity. Removal of residues 119—124 from S-protein

[231] M. C. Lin, B. Gutte, S. Moore, and R. B. Merrifield, *J. Biol. Chem.*, 1970, **245**, 5169.
[232] H. Taniuchi and C. B. Anfinsen, *J. Biol. Chem.*, 1969, **244**, 3864.

by controlled enzymic degradation gives des-(119—124)-S-protein which forms a termolecular active complex when mixed with S-peptide and the synthetic (111—124)-tetradecapeptide.[231] The availability of three non-covalently bound active peptide complexes clearly extends greatly the potential of ribonuclease systems as models for exploring by the synthetic approach the relationship between structure and enzymic activity.

Details of a study [233] of a series of S-peptide analogues which are competitive inhibitors have been published. A general picture of the rôle of the various parts of S-peptide has slowly emerged during the last few years from investigations of a large number of analogues, although there are controversies [233] over some important matters of detail (see Vol. 2, p. 183, of these Reports for further leading references). The histidine in position 12 is the only catalytically vital amino-acid in the sequence and this residue is possibly also important in the binding of S-peptide to S-protein. Several other residues such as aspartic acid-14, methionine-13, and phenylalanine-8 are certainly involved in binding but others can be extensively modified or, in the case of residues 15—20, completely discarded without significant loss of the capacity to activate S-protein. It was therefore argued [233] that analogues of S-peptide-(1—14)-tetradecapeptide in which the histidine residue was replaced or modified might be competitive inhibitors. A series of peptides was made to test this prediction, which proved to be correct. For example, [12-(3-carboxymethyl)-histidine]-S-peptide-(1—14)-tetradecapeptide not only fails to activate S-protein but binds more strongly than S-peptide, for which it is therefore a potent antagonist. The synthetic methods [233] involve a number of points of general interest, including minimal side-chain protection to facilitate purification of intermediates by ion-exchange chromatography, protection of methionine as its sulphoxide and ornithine side-chains by formylation (removal of the formyl group was achieved simultaneously with reduction of the sulphoxide by means of 50% thioglycollic acid at 37 °C) and the use of protected hydrazides. Of greater importance, however, is the fact that since the activation of S-protein by S-peptide may be analogous to the interaction of a peptide hormone with its receptor, this paper [233] provides some guidelines for the rational design of peptide hormone antagonists.

D. Staphylococcal Nuclease Systems.—Principal references to work on this enzyme and the structurally ordered and enzymically active complex (nuclease T) formed by non-covalent reassembly from the disordered and inactive peptides corresponding to residues 6—48 [P_2: (57)] and 49—149 (P_3) are to be found in Vol. 2, p. 177, of these Reports.

Since removal of lysine-48 from P_2 does not affect the ability to generate nuclease-T activity on admixture with P_3, synthetic studies have so far been focused on the sequence 6—47. Details are now available [234, 235] of the

[233] K. Hofmann, J. P. Visser, and F. M. Finn, *J. Amer. Chem. Soc.*, 1970, **92**, 2900.
[234] D. A. Ontjes and C. B. Anfinsen, ref. 11, p. 79.
[235] D. A. Ontjes and C. B. Anfinsen, *Proc. Nat. Acad. Sci., U.S.A.*, 1969, **64**, 428.

synthesis of this peptide by the solid-phase method. Trifluoroacetylation was used to protect lysine side-chains, and benzyloxycarbonylation of the imidazole ring for histidine side-chains, but otherwise the tactics were those of the standard (manual) solid-phase technique. Cleavage from the resin was performed with hydrofluoric acid in the presence of anisole as a

-Lys-Leu-His-Lys-Glu-Pro-Ala-Thr-Leu-Ile-Lys-Ala-Ile-Asp-Gly-Asp-Thr-Val-
6 10 15 20
Lys-Leu-Met-Tyr-Lys-Gly-Gln-Pro-Met-Thr-Phe-Arg-Leu-Leu-Leu-Val-Asp-
25 30 35 40
Thr-Pro-Glu-Thr-Lys-His-Pro-Lys
45 48

(57)

scavenger for electrophiles. The trifluoroacetyl groups were removed with piperidine, and preliminary purification by gel filtration was followed by affinity chromatography on a column of P_3 bound to Sepharose. The final peptide produced only about 30% of nuclease-T activity when added to P_3 in equimolecular amount. Further purification of the semi-synthetic nuclease-T has been achieved more recently [236] by limited tryptic digestion in the presence of calcium ions and deoxythymidine 3,5′-diphosphate followed by phosphocellulose chromatography. By this means the specific activity of the semi-synthetic nuclease-T was raised to about half that of material prepared from the native fragments P_2 and P_3. This degree of activity probably represents the maximum level attainable by the methods used, since a control experiment [235] in which native P_2 was treated with hydrofluoric acid followed by piperidine caused loss of rather more than 50% of P_3-activating ability.

There are clearly many shortcomings in this synthesis and therefore uncertainties about the product, and purists might argue that any conclusions based on experiments with impure analogues prepared in this way would rest on shaky foundations. This view is unnecessarily restrictive, since even a very unsatisfactory solid-phase synthesis would presumably lead to a product containing at least some of the desired peptide. If this much can be assumed it follows that, if significance is attached only to the presence or absence of activity in a synthetic product, without detailed attention to any but gross quantitative aspects, then in some cases confident deductions can be made. There is, however, still a little room for argument. For example, a completely inactive product might be taken as an unambiguous indication that the nominal synthetic peptide is inactive, but an inhibitor might be present in the product. The opposite situation is also slightly uncertain, since a sequence which has suffered deletion or damage during synthesis might conceivably have high activity. However, these objections could well be hypercritical, and Anfinsen and his colleagues were therefore not deterred from applying the route used for the synthesis of native P_2 to the preparation of analogues.

A series of modified or truncated sequences (for a list see Table 3) have now been described in detail.[236–238] It has been shown that the synthetic (9—47)-fragment is as active as the synthetic (6—47)-sequence, but further stepwise deletion from the *N*-terminal reduces (on removal of lysine-9) and then (on removal of glutamic acid-10) abolishes activity entirely although even the greatly shortened (18—47)-sequence still binds to P_3 significantly.[237, 238] Crystal structure studies do not implicate residues 9 and 10 in the catalytic site but suggest that they are important for the maintenance of an active conformation. The methionine residues can be replaced by norleucine with full retention of activity but the carboxy-group of glutamic acid-43 is essential, since the [43-glutamine]-analogue does not activate P_3 although it binds to it.[237] Histidine-8 is clearly inessential, it being situated in a section of the chain which can be discarded without inactivation: the other histidine residue (number 46) is also unimportant since the analogues with glycine in position 46 give active semi-synthetic nuclease-T preparations. The results published to date are summarised in Table 3.

Table 3[a] *Synthetic analogues of fragment P_2 of staphylococcal nuclease*

Peptide[b]	*Does the peptide bind to P_3?*[c]	*Does the peptide activate P_3?*
(6—47)-P_2	Yes	Yes
(9—47)-P_2	Yes	Yes
(10—47)-P_2	Yes	Slightly
(11—47)-P_2	Yes	No
(12—47)-P_2	Yes	No
(16—47)-P_2	Yes	No
(18—47)-P_2	Yes	No
(33—47)-P_2	No	No
[$Nle^{26,32}$]-(6—47)-P_2	Yes	Yes
[Gln^{43}]-(6—47)-P_2	Yes	No
[Gly^{46}]-(6—47)-P_2	Yes	Yes
[Gly^{46}]-(9—47)-P_2	Yes	Yes

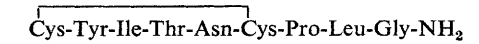
[a] From refs. 235, 236, 237, 238.
[b] The numbering system in the table and throughout this section is that of the intact staphylococcal nuclease.
[c] Various criteria were used.

E. [4-Threonine]-Oxytocin.—There are many reasons for the sustained work of the last decade or so on synthetic analogues of known peptide hormones. These include interest in the missing links of evolutionary processes, the use of analogue investigation as a basis for hypotheses concerning hormone-receptor interaction, and the search for more potent, selective, and longer-acting materials for clinical use. All too often a laborious synthesis leads to a useless analogue which yields but a tiny

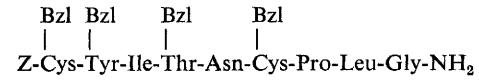
[236] I. M. Chaiken and C. B. Anfinsen, *J. Biol. Chem.*, 1970, **245**, 2337.
[237] D. A. Ontjes and C. B. Anfinsen, *J. Biol. Chem.*, 1969, **244**, 6316.
[238] I. M. Chaiken and C. B. Anfinsen, *J. Biol. Chem.*, 1970, **245**, 4718.

morsel of information, and rarely does a single peptide hit the jackpot. The recent description [239, 240] of an oxytocin analogue with threonine in the 4-position is an exception to this generalisation, because this compound is not only of great theoretical interest but also has promise of clinical value.

All of the naturally occurring neurohypophysial hormones so far characterised have either serine or glutamine at position 4. Since, however, the 4-serine peptides have only been found in certain types of fish these may represent a diversion from the principal evolutionary procession. In any event, mutation from one series to the other is clearly of interest because interchange of serine and glutamine requires two base changes in the codons for these amino-acids, which raises the possibility of the existence in Nature of a missing link which lies undiscovered. Manning and his colleagues [239, 240] therefore embarked on an examination of possible evolutionary stepping-stones, which they prepared by the solid-phase method. One of the peptides prepared [240] as part of this screening was [4-threonine]-oxytocin (58). Standard solid-phase methods were used to assemble the nine protected amino-acid residues on the resin, except that each coupling stage was performed twice in the hope of minimising the



(58)



(59)

formation of erroneous sequences. Ammonolysis of the protected peptide–resin conjugate gave analytically and chromatographically pure fully-protected nonapeptide amide (59) which, after treatment with sodium in liquid ammonia, oxidation, and appropriate purification gave the desired analogue (58). Assay of activity in standard tests showed that this peptide has greater oxytocic, avian vasodepressor, and galactogogic activity than oxytocin, but reduced vasopressor and antidiuretic potency. In fact, the ratio between oxytocic and antidiuretic activity is at least four times more in favour of the former in the case of [4-threonine]-oxytocin than with oxytocin itself. Intravenous infusion of oxytocin to induce labour can, if prolonged, cause water retention to the point of intoxication, and if the relative activities so far reported (determined on the rat) are paralleled in women, [4-threonine]-oxytocin may become a clinically important oxytocic.

Many examples of synthetic analogues which have higher specific activities than the parent hormones have been reported, but (58) is the

[239] M. Manning and W. H. Sawyer, *Nature*, 1970, **227**, 715.
[240] M. Manning, E. Coy, and W. H. Sawyer, *Biochemistry*, 1970, **9**, 3925.

first to be discovered in which potentiation is achieved merely by replacing a single residue of the natural peptide with a different 'ordinary' (*i.e.* which occurs in proteins) L-α-amino-acid. Increased activity has previously only been brought about by an 'unnatural' feature such as the absence of an *N*-terminal amino-group or the presence of an inverted configuration. In all probability the greater activity of these compounds is not intrinsic but can be attributed to reduced susceptibility to enzymic inactivation. Thus β-corticotropin fragment analogues with a D- or β-amino-acid in position 1 have increased and longer-lasting steroidogenic capacity, which can be ascribed to their slower degradation by aminopeptidases (see the Section under Adrenocorticotropins in Appendix A for an entry into the literature on the effects of structural modifications on β-corticotropin).

Since it appears that [4-threonine]-oxytocin could be an advantageous mammalian oxytocic hormone, it may yet be discovered in Nature. In this context it will be recalled that both arginine-vasotocin and mesotocin were synthesised in the laboratory before their identification as naturally-occurring hormones. However, evolution of (58) from the [4-glutamine]-series, which apparently forms the mainstream of neurohypophysial hormone phylogeny,[241] requires two base changes. Stepwise mutation of glutamine to threonine can occur *via* proline or lysine but [4-proline]-[242] and [4-lysine]-[243] analogues have very low oxytocic activities, and would therefore have been unlikely to survive long enough for subsequent emergence of (58).

F. Thyroid-stimulating Hormone Releasing Factor.—The existence of a thyrotropin (thyroid-stimulating hormone) releasing factor (hereafter abbreviated TRF) in extracts of hypothalami has been known for nearly ten years, but it has only been isolated in the pure state and fully characterised recently. Lucid accounts of patient work in several laboratories on this substance have been given in a recent review [217] and a summarising paper [244] and the reader is referred to these for the original references to matters other than those directly related to the synthetic work involved.

Heroic efforts were required for the isolation of TRF, some hundreds of thousands of hypothalami yielding at first but a few milligrams of material heavily contaminated with artefacts. The TRF activity resisted proteolytic enzymes, and acid hydrolysates of early preparations contained amino-acids only sufficient to account for a low proportion of the weight hydrolysed. The hypothesis was therefore formed that TRF, if a peptide, was not homomeric. Later preparations gave on hydrolysis much higher yields of amino-acids, these being mainly glutamic acid, histidine, and proline in equimolecular amounts. Experiments on synthetic tripeptides

[241] I. L. Geschwind, ref. 13, p. 385.
[242] W. H. Sawyer, T. C. Wuu, J. W. M. Baxter, and M. Manning, *Endocrinology*, 1969, **85**, 385.
[243] W. H. Sawyer *et al.*, unpublished data referred to in ref. 240.
[244] R. Burgus, T. F. Dunn, D. Desidero, D. N. Ward, W. Vale, and R. Guillemin, *Nature*, 1970, **226**, 321.

containing these amino-acids were therefore initiated.[245–247] All the possible isomers were synthesised [245–249] but none had biological activity. Since there was evidence that TRF lacked a free *N*-terminal amino-group, various derivatives were examined and pyroglutamylhistidylprolineamide (60) was soon found to have high TRF activity.[246–250] After a short period of residual doubt the availability of much purer natural ovine TRF and its comparison with unambiguously synthesised pyroglutamylhistidyl-prolineamide by a wide range of techniques finally clinched the identity of the ovine hormone as (60).[244, 247, 249] A slightly less rigorous comparison [250]


Glu-His-Pro-NH_2

(60)

places the assignment of the same structure to porcine TRF on an almost equally sure foundation.

Practical details of two classical [249, 251] and of one solid-phase preparations [251] of the hormone have been published. Synthetic (60) is highly potent and specific in its action when intravenously injected in man,[252] and it has also been shown [253] that its resistance to proteolysis by digestive enzymes enables it to be administered orally in experimental animals. The availability of pure synthetic TRF in virtually unlimited amount seems likely to be of great importance for academic, diagnostic, and therapeutic purposes.

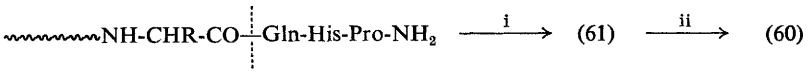
The great simplicity of the structure (60) – it is the smallest and easily the most synthetically accessible peptide with hormonal activity yet discovered – has already made possible the rapid screening of a fair number of analogues. Only relatively minor structural modifications are consistent with retention of activity. The *C*-terminal amide can be replaced by a methyl ester with partial retention of activity, but widely varying figures (*ca.* 2% [254] and *ca.* 50% [248] of the activity of TRF in mice) have been reported

[245] R. Burgus, T. F. Dunn, D. N. Ward, W. Vale, M. Amoss, and R. Guillemin, *Compt. rend.*, 1969, **268**, *D*, 2116.
[246] R. Burgus, T. F. Dunn, D. Desiderio, W. Vale, and R. Guillemin, *Compt. rend.*, 1969, **269**, *D*, 226.
[247] K. Folkers, F. Enzmann, J. Bøler, C. Y. Bowers, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1969, **37**, 123.
[248] R. Burgus, T. F. Dunn, D. M. Desiderio, D. N. Ward, W. Vale, R. Guillemin, A. M. Felix, D. Gillesssen, and R. O. Studer, *Endocrinology*, 1970, **86**, 573.
[249] D. Gillessen, A. M. Felix, W. Lergier, and R. O. Studer, *Helv. Chim. Acta*, 1970, **53**, 63.
[250] J. Bøler, F. Enzmann, K. Folkers, C. Y. Bowers, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1969, **37**, 705.
[251] G. Flouret, *J. Medicin. Chem.*, 1970, **13**, 843.
[252] C. Y. Bowers, A. V. Schally, D. S. Schalch, C. Gual, A. J. Kastin, and K. Folkers, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 352.
[253] W. Vale, R. Burgus, T. F. Dunn, and R. Guillemin, *J. Clin. Endocrinol. Metab.*, 1970, **30**, 148.
[254] C. Y. Bowers, A. Weil, J.-K. Chang, H. Sievertsson, F. Enzmann, and K. Folkers, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 683.

for the quantitative effect of this change. The *N*-methylamide [254] appears to have activity comparable with the methyl ester, but the corresponding free acid [248] has much lower activity. Replacement of the proline residue by glycine or phenylalanine abolishes activity, but the analogues with alanine, leucine, or valine at position 3 are about one-thousandth as active as TRF.[254] However, the interpretation of these results is complicated by the fact that modifications in the *C*-terminal region result in much slower serum inactivation.[254] The introduction of a benzyl group at the histidine residue reduces activity to 0.2% of that of TRF,[254] but replacement of histidine by β-(3-pyrazolyl)-alanine gives an analogue with 5% activity.[255] The glutamine analogue (61) [256] also has an apparent activity of about 5%

Gln-His-Pro-NH_2
(61)

but, since (61) cyclises to (60) under mild conditions, (61) may nevertheless be intrinsically inactive. The ready conversion of (61) to (60) raises the possibility [244, 256] that the pyrrolidine ring of the isolated hormone is artefactual, but this seems unlikely in view of the higher potency of (60). The glutaminyl peptide amide (61) could, however, be a biosynthetic precursor of (60) in which case TRF might be formed *in vivo* from a larger peptide,[256] in a manner (Scheme 25) similar to that which has been demonstrated for some other pharmacologically active peptides.



Conditions: i, proteolysis; ii, *in vivo* cyclisation

Scheme 25

Perhaps other currently elusive hypothalamic-releasing factors [217] will prove to be similar to TRF (they cannot be more than similar since experiments [244, 252] with synthetic TRF show it to be highly specific in its action). It has been pointed out [253] that if luteinising-hormone releasing factor also turns out to be an easily synthesised small peptide which can be administered orally this could have great impact on contraceptive methodology.

4 Appendix A. A List of Syntheses Reported during 1970

A. Naturally Occurring Peptides, Proteins, Analogues, and Partial Sequences.—Analogues and partial sequences are listed under the peptide to which they are related.

[255] K. Hofmann and C. Y. Bowers, *J. Medicin. Chem.*, 1970, **13**, 1099.
[256] K. Folkers, J.-K. Chang, B. C. Currie, C. Y. Bowers, A. Weil, and A. V. Schally, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 110.

Peptide	*Ref.*
Acyl carrier protein	
A pentapeptide fragment of *E. coli* acyl carrier protein apoprotein	257
Adrenocorticotropins	
β-Corticotropin-(1–23)-tricosapeptide amide	74, 258
[β-Ala^{1}]-β-corticotropin-(1–23)-tricosapeptide amide	74
[Gln^{5},Phe^{9}]-β-corticotropin-(1—20)-eicosapeptide amide	259
[Gln^{5}]-6-β-(3-pyrazolyl)-alanine-β-corticotropin-(1—20)-eicosapeptide amide	259
[D-Ser^{1},Lys^{17,8}]-β-corticotropin-(1—18)-octadecapeptide amide	260
[Gly^{1}]-β-corticotropin-(1—18)-octadecapeptide amide	261
[β-Ala^{1}]-β-corticotropin-(1—18)-octadecapeptide amide	262
[β-Ala^{1}]-β-corticotropin-(1—24)-tetracosapeptide	263
[Abu^{1}]-β-corticotropin-(1—24)-tetracosapeptide	263
[Sar^{1}]-β-corticotropin-(1—24)-tetracosapeptide	263
[Pro^{1}]-β-corticotropin-(1—24)-tetracosapeptide	263
[Lys^{1}]-β-corticotropin-(1—24)-tetracosapeptide	263
[Leu^{4}]-β-corticotropin-(1—24)-tetracosapeptide	264
[Ile^{4}]-β-corticotropin-(1—24)-tetracosapeptide	264
Human β-corticotropin-(1—27)-heptacosapeptide	265
[D-Ser^{1},Nva^{17,18}]-β-corticotropin-(1—19)-nonadecapeptide amide	266
[D-Ser^{1},Nle^{17,18}]-β-corticotropin-(1—19)-nonadecapeptide amide	266
[Gln^{5}]-7-[^{14}C]-phenylalanine-β-corticotropin-(1—20)-eicosapeptide amide	267
A series of protected fragments spanning the sequence of β-corticotropin-(1—23)-tricosapeptide amide	268
β-Corticotropin-(5—14)-decapeptide	147
Protected β-corticotropin-(8—23)-hexapeptide amide	48
Angiotensins	
[Val^{5}]-angiotensin II	173, 269
[Ile^{5},Tyr^{8}]-angiotensin II	270
[Ile^{5},Tyr(Me)^{8}]-angiotensin II	270
[Val^{5},Tyr(Me)^{8}]-angiotensin II	270
[Tyr(Me)^{4},Ile^{5}]-angiotensin II	270
[Tyr(Me)^{4},Val^{5}]-angiotensin II	270
[Asn^{1},Tyr(Me)^{4},Val^{5}]-angiotensin II	270

[257] Y. Miura and S. Seto, *Bull. Chem. Soc. Japan*, 1969, **42**, 3592.
[258] M. Fujino, O. Nishamura, and C. Hatanaka, *Chem. and Pharm. Bull.* (*Japan*), 1969, **17**, 2186.
[259] K. Hofmann, R. Andreatta, H. Bohn, and L. Moroder, *J. Medicin. Chem.*, 1970, **13**, 339.
[260] B. Riniker and W. Rittel, *Helv. Chim. Acta*, 1970, **53**, 513.
[261] H. Otsuka, M. Shin, Y. Kinomura, and K. Inouye, *Bull. Chem. Soc. Japan*, 1970, **43**, 196.
[262] K. Inouye, A. Tanaka, and H. Otsuka, *Bull. Chem. Soc. Japan*, 1970, **43**, 1163.
[263] M. Fujino, C. Hatanaka, and O. Nishimura, *Chem. and Pharm. Bull.* (*Japan*), 1970, **188**, 1288.
[264] M. Fujino, O. Nishimura, and C. Hatanaka, *Chem. and Pharm. Bull.* (*Japan*), 1970, **18**, 1291.
[265] H. Otsuka, K. Watanabe, and K. Inouye, *Bull. Chem. Soc. Japan*, 1970, **43**, 2278.
[266] M. Brugger, P. Barthe, and P. A. Desaulles, *Experientia*, 1970, **26**, 1050.
[267] L. Moroder and K. Hofmann, *J. Medicin. Chem.*, 1970, **13**, 839.
[268] M. Fujino, O. Nishimura, and C. Hatanaka, *Chem. and Pharm. Bull.* (*Japan*), 1969, **17**, 2135.
[269] A. Pavars and G. Cipens, *Latv. P.S.R. Zinat. Akad. Vestis, Kim. Ser.*, 1970, 121 (*Chem. Abs.*, 1970, **73**, 15242m).

Peptide	*Ref.*
[Ile^5]-7-pipecolic acid-angiotensin II	270
[Ile^5]-8-(3-amino-4-phenyl)butyric acid-angiotensin II	270
[Ile^5]-8-(3-amino-3′-phenyl)isobutyric acid-angiotensin II	270
[Asn^1,Gly^2,Val^5]-angiotensin II	271
[Ile^5]-angiotensin II-(3—8)-hexapeptide	272
Butyryl-, D- and L-α-aminobutyryl-, ε-aminocaproyl-, δ-aminovaleryl-, glycyl-, and acetylglycyl-[Ile^5]-angiotensin II-(3—8)-hexapeptide	272
Triglycyl-[Ile^5]-angiotensin II-(4—8)-pentapeptide	273
Triglycyl-[Val^5]-angiotensin II-(4—8)-pentapeptide	269
A series of oligoglycyl derivatives of truncated [Val^5]-angiotensins II	269
ω-Amino-octanoyl-[Ile^5]-angiotensin II-(4—8)-pentapeptide	273
The *C*-terminal tetra- and hexa-peptides of [Val^5]-angiotensin II	269
Bradykinins	173
A series of analogues of bradykinin comprising Arg-Gly$_n$-Arg (n = 1—7), Arg-Gly$_3$-Phe-Gly$_3$-Arg, Arg-Gly$_3$-Phe-Gly$_2$-Phe-Arg, and Arg-Gly$_6$-Phe-Arg	274
Calcitonins	
Partially protected calcitonin *N*-terminal nonapeptide	132
The human hormone 'calcitonin M': full details	224, 275
Calcitonin M *N*-terminal decapeptide: full details	275
Casein	
An octadecapeptide related to an α-casein degradation product	276
Cholecystokinin-pancreozymin(CCK-PZ)	
CCK-PZ *C*-terminal dodecapeptide amide	229
CCK-PZ *C*-terminal decapeptide amide	229
CCK-PZ *C*-terminal octapeptide amide	229
Analogues and partial sequences of CCK-PZ *C*-terminal octapeptide amide	228, 230
Cytochrome *c*	
Various protected sequences of cytochrome *c*	277–281
Eledoisin	
Protected [Lys^5]-eledoisin-(5—11)-heptapeptide amide	282

[270] N. C. Chaturvedi, W. K. Park, R. R. Smeby, and F. M. Bumpus, *J. Medicin. Chem.*, 1970, **13**, 177.

[271] A. Pavars and G. Cipens, *Latv. P.S.R. Zinat. Akad. Vestis, Kim. Ser.*, 1970, 250 (*Chem. Abs.*, 1970, **73**, 56406m).

[272] E. C. Jorgensen, G. C. Windridge, and T. C. Lee, *J. Medicin. Chem.*, 1970, **13**, 352.

[273] E. C. Jorgensen, G. C. Windridge, and T. C. Lee, *J. Medicin. Chem.*, 1970, **13**, 744.

[274] F. B. Abramson, D. F. Elliott, D. G. Lindsay, and R. Wade, *J. Chem. Soc.* (*C*), 1970, 1042.

[275] B. Kamber, H. Brückner, B. Riniker, P. Sieber, and W. Rittel, *Helv. Chim. Acta*, 1970, **53**, 556.

[276] K. P. Polzhofer and K. H. Ney, *Tetrahedron*, 1970, **26**, 3221.

[277] R. P. Evstigneeva, A. F. Mironov, G. A. Vasil'eva, L. V. Skibenko, and T. S. Nikitina, *Zhur. obshchei Khim.*, 1970, **40**, 661 (*Chem. Abs.*, 1970, **73**, 15222e).

[278] S. D. L'vova, R. P. Evstigneeva, O. S. Chuprikova, and N. A. Preobrazhenskii, *Khim. prirod. Soedinenii*, 1970, **6**, 114 (*Chem. Abs.*, 1970, **73**, 35745f).

[279] S. D. L'vova, R. P. Evstigneeva, S. A. Mel'nikova, I. N. Ryabova, and N. A. Prebrazhenskii, *Zhur. obshchei Khim.*, 1970, **40**, 242 (*Chem. Abs.*, 1970, **72**, 133182n).

[280] Y. Wolman, M. Schwarzenberg, and M. Frankel, *Israel J. Chem.*, 1970, **8**, 53.

[281] Y. Wolman, D. Ladkany, and M. Frankel, *Israel J. Chem.*, 1970, **8**, 59.

[282] J. Halstrøm, K. Brunfeldt, J. Thomsen, and K. Kovacs, *Acta Chem. Scand.*, 1969, **23**, 2335.

Peptide	*Ref.*
Encephalitogenic protein	
A series of peptides related to the disease-including site of encephalitogenic protein A1 from bovine myelin	283, 284
Ferredoxin	
Glycyl-, phenylalanyl-, lysyl-, glutamyl-, and methionyl-ferredoxins	211
A series of peptides related to the *C*-terminal haptenic octapeptide of *C. pasteurianum* ferredoxin and the *C*-terminal octapeptide of the cross-reacting *C. butyricum* ferredoxin	285, 286
Fibrinopeptide A (human)	287
Gastrins	
Analogues of gastrin *C*-terminal tetrapeptide amide having β-cyanoalanine or γ-cyano-α-amino-butyric acid in place of aspartic acid	288
Analogues of gastrin *C*-terminal tetrapeptide amide having various substitutions for the tryptophan residue	289
The *N*-benzyloxycarbonyl derivative of the *C*-terminal tetrapeptide amide	290
Glucagon	
Fully protected glucagon-(7—15)-nonapeptide	171
Glutathione (oxidised)	60
The analogue of oxidised glutathione with djenkolic acid substituted for cystine	60
Growth hormone	
Fully protected human growth hormone-(180—184)-pentapeptide	153
Insulin	
The B chain of the bovine hormone and the preparation of semi-synthetic insulin from it	57
Protected fragments of the bovine hormone B chain, including symmetrical cystine derivatives	55–57
Arginyl-, lysyl-, and histidyl-bovine insulin A chain	212
A series of 12 bovine insulin B chain analogues with substitutions at two or more of positions 4, 5, 9, 10, 27, and 28 by alanine and the formation of semi-synthetic insulin analogues from these	291
Sheep insulin A chain-(9—13)-pentapeptide	92
Lysozyme	
The *N*-terminal decapeptide of the tryptic peptide T-11 from reduced and carboxymethylated egg-white lysozyme	292
Melanocyte-stimulating hormones	
[Lys^{10}]-human β-MSH	213

283 E. H. Eylar, J. Caccam, J. J. Jackson, F. C. Westall, and A. B. Robinson, *Science*, 1970, **168**, 1220.

284 F. C. Westall, A. B. Robinson, J. Caccam, J. Jackson, and E. H. Eylar, *Nature*, 1971, **229**, 22.

285 B. Mitchell and J. G. Levy, *Biochemistry*, 1970, **9**, 2762.

286 B. Mitchell, J. G. Levy, and R. N. Nitz, *Biochemistry*, 1970, **9**, 1839.

287 B. J. Johnson and W. P. May, *J. Pharm. Sci.*, 1969, **58**, 1568.

288 Z. Grzonka and B. Liberek, *Roczniki Chem.*, 1969, **9**, 1659.

289 G. Cipens, V. Vilka, and Z. Atare, *Khim. prirod. Soedinenii*, 1970, **6**, 117 (*Chem. Abs.*, 1970, **73**, 35741b).

290 Z. Atare, V. Vilka, and G. Cipens, *Khim. prirod. Soedinenii*, 1969, **5**, 283 (*Chem. Abs.*, 1970, **72**, 55867e).

291 G. Weitzel, U. Weber, K. Eisele, H. Zollner, and J. Martin, *Z. physiol. Chem.*, 1970, **351**, 263.

292 K. E. Thompson and J. G. Levy, *Biochemistry*, 1970, **9**, 3463.

Peptide	*Ref.*
Human β-MSH: full details	293
[Gln^5]-α-MSH	294
Oxytocin	
[Thr^4]-oxytocin	239, 240
[Ile^{2,4}]-oxytocin	295, 296
The parallel dimer of oxytocin	297
1-Desamino-oxytocinoic acid	298
1-Desamino-oxytocinoyloxytocin	298
The acetone adduct of oxytocinoic acid	298
Oxytocinoyloxytocin	298
1-Desamino-[Leu^4]-oxytocin	299
1-Desamino-[Ile^{2,4}]-oxytocin	299
1-Desamino-oxytocin specifically deuteriated at the asparagine residue	300
1-Desamino-[Lys^8]-oxytocin (*i.e.* 1-desamino-[Lys^8]-vasotocin)	301
1-Desamino-dicarba-oxytocin	302
[1-(ω-Meraptoundecanoic acid)]-oxytocin	303
[1-(δ-Mercaptovaleric acid)]-oxytocin	304
Various partial sequences and protected derivatives thereof	79, 305
Parathyroid hormone	
Bovine parathyroid hormone *C*-terminal eicosapeptide	306
Phyllocaerulin	
[Thr^3,Tyr(SO$_3$H)^4]-phyllocaerulin	307
Proinsulin*	
[Gln^{35},Lys(For)^{62}]-porcine proinsulin-(31—63)-tricontatripeptide *i.e.* the 'connecting peptide'	308
Protected [Gln^{35}]-porcine proinsulin-(31—39)-nonapeptide [details]	48

[293] H. Yajima, K. Kawasaki, H. Minami, H. Kawatani, N. Mizokami, Y. Kiso, and F. Tamura, *Chem. and Pharm. Bull. (Japan)*, 1970, **18**, 1394.
[294] J. J. Blake, R. W. Crooks, and C. H. Li, *Biochemistry*, 1970, **9**, 2071.
[295] V. J. Hruby, V. du Vigneaud, and W. Y. Chan, *J. Medicin. Chem.*, 1970, **13**, 185.
[296] W. Y. Chan and V. du Vigneaud, *J. Pharmacol. Exptl. Therapeutics*, 1970, **174**, 541.
[297] H. L. Aanning and D. Yamashiro, *J. Amer. Chem. Soc.*, 1970, **92**, 5214.
[298] H. Takashima and V. du Vigneaud, *J. Amer. Chem. Soc.*, 1970, **92**, 2501.
[299] H. Takashima, V. J. Hruby, and V. du Vigneaud, *J. Amer. Chem. Soc.*, 1970, **92**, 677.
[300] A. T. Blomquist, D. H. Rich, B. A. Carlson, G. A. Allen, V. J. Hruby, H. Takashima, L. L. Nangeroni, P. Glose, and V. du Vigneaud, *Proc. Nat. Acad. Sci., U.S.A.*, 1969, **64**, 263.
[301] M. Rimpler and A. Schoeberl, *Naturwiss.*, 1969, **56**, 638.
[302] A. Kozbayashi, S. Hase, R. Kiyoi, and S. Sakakibara, *Bull. Chem. Soc. Japan*, 1969, **42**, 3491.
[303] W. Fraefel and V. du Vigneaud, *J. Amer. Chem. Soc.*, 1970, **92**, 4426.
[304] W. Fraefel and V. du Vigneaud, *J. Amer. Chem. Soc.*, 1970, **92**, 1030.
[305] O. A. Kaurov, V. F. Martynov, and O. Popernatskii, *Zhur. obshchei Khim.*, 1970, **40**, 904 (*Chem. Abs.*, 1970, **73**, 25860a).
[306] H. Yajima, H. Kawatani, and H. Watanabe, *Chem. and Pharm. Bull. (Japan)*, 1970, **18**, 1279.
[307] R. de Castiglione and O. Goffredo, *Farmaco, Ed. Sci.*, 1970, **25**, 40 (*Chem. Abs.*, 1970, **72**, 121919b).
[308] N. Yanaihara, T. Hashimoto, C. Yanaihara, and N. Sakura, *Chem. and Pharm. Bull. (Japan)*, 1970, **18**, 417.

* It has now been found [311] that residue 35 of porcine proinsulin is glutamic acid, not glutamine, so that synthetic efforts based on the sequence as first proposed have produced by this mischance not the connecting peptide but slightly modified analogues.

Peptide	*Ref.*
Ribonuclease	
Bovine pancreatic ribonuclease *C*-terminal tetradecapeptide	231
Protected fragments of ribonuclease T_1	309, 310
Ribonuclease *N*-terminal eicosapeptide (S-peptide)	
A series of analogues of S-peptide-(1—14)-tetradecapeptide with modifications at positions 10 and 12	233
A series of analogues of S-peptide-(1—13)-tridecapeptide with modifications at positions 10 and 13	146, 151, 152
Staphylococcal nuclease	
Analogues and partial sequences of staphylococcal nuclease-(6—48)-peptide (P_2)	236–238
Partially protected staphylococcal nuclease-[Gln[43]]-(43—50)-octapeptide	38
Trypsinogen	
The *N*-terminal nonapeptide of bovine trypsinogen: two further syntheses	312
Thyrotropin-releasing hormone (TRH)	247, 249, 251
Various analogues of TRH	248, 249, 254, 255, 256
Vasopressin	
1-Desamino-[Ile[3]]-lysine-vasopressin (*i.e.* 1-desamino-[Lys[8]]-vasotocin)	301
4-Descarboxyamido-arginine-vasopressin	313
1-Desamino-4-descarboxyamido-arginine-vasopressin	313
Des-Gly[9]-arginine-vasopressin	314
Vasopressin-(1—7)-heptapeptide amide	314
Vasopressin-(1—6)-heptapeptide amide	315
[Dap[8]]-vasopressin	316
[D-Dap[8]]-vasopressin	316
[D-Orn[8]]-vasopressin	316

B. Sequential Polypeptides

Nominally monodisperse repeating sequence polymers:	
$(Leu\text{-}Ala)_6$	149, 162
$(Ala\text{-}Phe)_6$	149
$(Pro\text{-}Pro\text{-}Gly)_n$, $n = 10$, 15, and 20	317
Polydisperse repeating sequence polymers:	
Protected polyglutathione	118
poly-(Leu-Orn-Leu)	189
poly-(Leu-Arg-Leu)	189
poly-(Lys-Gly)-[1-^{14}C]-Gly-OEt	190

[309] N. Mitsuyasu, M. Waki, S. Makisumi, and N. Izumiya, *Mem. Fac. Sci., Kyushu Univ., Ser. C.*, 1970, **7**, 63 (*Chem. Abs.*, 1970, **73**, 88156w).

[310] N. Mitsuyasu, M. Waki, T. Kato, S. Makisumi, and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 1556.

[311] R. E. Chance, personal communication to G. Jäger and R. Geiger quoted in ref. 48.

[312] N. Yoshida, T. Kato, and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 2912.

[313] D. Gillessen and V. du Vigneaud, *J. Medicin. Chem.*, 1970, **13**, 346.

[314] O. S. Papsuevich and G. Cipens, *Zhur. obshchei Khim.*, 1970, **40**, 709 (*Chem. Abs.*, 1970, **73**, 15210z).

[315] O. S. Papsuevich and G. Cipens, *Latv. P.S.R. Zinat. Akad. Vestis, Khim. Ser.*, 1969, 751 (*Chem. Abs.*, 1970, **72**, 111807z).

[316] M. Zaorl, J. Kolc, and F. Sorm, *Coll. Czech. Chem. Comm.*, 1970, **35**, 1716.

[317] Y. Kobashi, R. Sakai, K. Kakiuchi, and T. Isemura, *Biopolymers*, 1970, **9**, 415; see also R. A. Berg, B. R. Olsen, and D. J. Prockop, *J. Biol. Chem.*, 1970, **245**, 5759.

Peptide	*Ref.*
poly-(Tyr-Glu-Val-Gly)-[1-^{14}C]-Gly-OEt	191
poly-(Ala-Gly-Gly)	318
poly-(Ala-Ala-Gly)	192
poly-(Ala-Gly)	201, 319
poly-(D-Asp-Lys)	193
poly-(Met-Ala)	320
Various collagen models containing glutamic acid and/or serine	194, 196
Other collagen models	197, 198, 321
Various *bombyx mori* fibroin models	319
Various sequential polypeptides containing glutamic acid and/or lysine	199, 200, 204, 207, 318

C. Miscellaneous Peptides

Unsymmetrical protected cystine peptides	51
Selenium-containing peptides	64, 68
Acetylglycylpeptides	80
Leu-Ala-Gly-Val	144, 155
An eicosapeptide of unspecified relevance prepared by the solid-phase method: the sequence was Lys-Asp-$(Glu)_4$-Val-Glu-Ser-Phe-Ser-Gly-Pro-Asp-Ala-Pro-Leu-Pro-Ala-Gly (related to casein?)	157
Gly-Asp-Ser-Gly-His	89
Gly-Asp-Ser-Phe	89
His-Gly-Asp-Ser-Phe	322
γ-Glutamyl peptides	323
Lysine- and ornithine-containing peptides	324, 325
Peptides of 3,3,3-trifluoroalanine	326
Oligomers of tyrosine	327
D-Alanine-containing peptides	328
A series of peptide derivatives containing (*inter alia*) Asn and/or Gln: these were prepared for mass spectrometric studies	329
Peptides of 3-amino-3-methylbutanoic acid	330
Oligomers of arginine	331
Peptides of *C*-phenylglycine	332

[318] J. M. Anderson, W. B. Rippon, and A. G. Walton, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 802.
[319] T. Hayakawa, Y. Kondo, S. Yamamoto, and J. Noguchi, *Kobunshi Kagaku*, 1970, **27**, 229 (*Chem. Abs.*, 1970, **73**, 110124h).
[320] G. Boussard, É. Dellacherie, and J. Néel, *Bull. Soc. chim. France*, 1970, 2544.
[321] W. L. Mattice and L. Mandelkern, *J. Amer. Chem. Soc.*, 1970, **92**, 5285.
[322] A. Kapoor, S. M. Kang, and M. A. Trimboli, *J. Pharm. Sci.*, 1970, **59**, 129.
[323] D. E. Nitecki and J. W. Goodman, ref. 11, p. 435.
[324] N. A. Poddubnaya and N. Ya. Krasnobrizhii, *Zhur. obshchei Khim.*, 1970, **40**, 1137, 1142 (*Chem. Abs.*, 1970, **73**, 77585j, 77586k).
[325] V. V. Bondar, V. A. Shibnev, and K. T. Poroshin, *Doklady Akad. Nauk Tadzh. S.S.R.*, 1970, **13** (5), 14 (*Chem. Abs.*, 1970, **73**, 66884e).
[326] F. Weygand, W. Steglich, and W. Oettmeier, *Chem. Ber.*, 1970, **103**, 1655.
[327] C. Sorg, E. Rüde, and O. Westphal, *Annalen*, 1970, **734**, 180.
[328] B. Schechter, I. Schechter, and M. Sela, *J. Biol. Chem.*, 1970, **245**, 1438.
[329] A. I. Miroshnikovs, A. A. Kiryushkin, and Yu. A. Ovchinnikov, *Zhur. obshchei Khim.*, 1970, **40**, 223 (*Chem. Abs.*, 1970, **72**, 101080t).
[330] J. Lowbridge and C. N. C. Drey, *Chem. Comm.*, 1970, 791.
[331] H. Ito, I. Ichikizaki, and T. Ando, *Internat. J. Protein Res.*, 1970, **2**, 59.
[332] A. Balog, E. Vargha, V. Voinescu, M. Herman, E. Ramontian, and L. Moraru, *Stud. Univ. Babes-Bolyai, Ser. Chem.*, 1969, **14**, 137 (*Chem. Abs.*, 1970, **73**, 25843x).

Peptide	*Ref.*
Oligomers of proline	333, 334
Peptides containing sarcolysine and γ-aminobutyric acid	335
Peptides containing fluorinated phenylalanines	336
Radioactive peptides containing one phenylalanine residue and several glycine residues	337
Peptides of β-cyano-alanine	338
Ac-Glu-Gly, Ac-Tyr-Glu, Ac-Tyr-Tyr, and Ac-Phe-Tyr	339
A series of peptides which were screened for antirenin activity	340
A protected nonapeptide with the sequence of a cyclononapeptide obtained from linseed	33
A series of fully protected di- and tri-peptides containing *inter alia* asparagine, glutamine, tyrosine, and glutamic acid	46

5 Appendix B. A List of Some Useful New Synthetic Intermediates Described during 1970

As in Volume 2, apologies are offered for the selective and subjective nature of this compilation. Compounds were taken to be new on the authority of the authors who described them: all are crystalline unless otherwise stated. Newly crystallised intermediates which have been reported previously as oils are also included. Last year only derivatives of amino-acids which commonly occur in proteins were covered, but this year some derivatives of especially interesting protein amino-acid analogues are listed. In this appendix, substitution into protecting groups is indicated by placing the substituent in brackets immediately after the symbol for the protecting group, *e.g.* Bzl(OMe) = methoxybenzyl; Ph(SMe) = methylthio-phenyl. For the position of the substituent, the reference should be consulted.

Compound	*Ref.*
Alanine	
Ala-OBzl(OMe),HCl	44
Z-Ala-OPfp	78
Z-Ala-OPh(SMe)	85
Z-Ala-OPh(SO_2Me)	85
Ala-OPh(SMe),HBr	85
Ala-NH·NH·Boc	144

[333] M. Miyoshi, T. Kimura, and S. Sakakibara, *Bull. Chem. Soc. Japan*, 1970, **43**, 2941; C. M. Deber, F. A. Bovey, J. P. Carver, and E. R. Blout, *J. Amer. Chem. Soc.*, 1970, **92**, 6191.

[334] M. Rothe, R. Theyson, and K.-D. Steffen, *Tetrahedron Letters*, 1970, 4063.

[335] A. Nauliukonis, K. Karpavicius, O. V. Kil'disheva, and I. L. Knunyants, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1970, 161 (*Chem. Abs.*, 1970, **72**, 121915x).

[336] O. A. Kaurov, V. F. Martynov, and M. P. Smirnova, *Zhur. obshchei Khim.*, 1970, **40**, 903 (*Chem. Abs.*, 1970, **72**, 25851y).

[337] R. L. Smith, E. G. Archer, and F. W. Dunn, *J. Biol. Chem.*, 1970, **245**, 2967.

[338] Z. Grzonka and B. Liberek, *Zeszyty Nauk. Wyzszej Szkoly Pedagogicznej Gdansku Mat., Fis., Chem.*, 1969, **9**, 107 (*Chem. Abs.*, 1970, **72**, 101093z).

[339] M. Shiraki, T. Ojima, M. Honjo, H. Usami, and M. Masaru, *Sci. Pap. Coll. Gen. Educ. Univ. Tokyo*, 1969, **19**, 121 (*Chem. Abs.*, 1970, **73**, 66885).

[340] A. Ide, K. Shigezane, S. Shigezane, T. Mizoguchi and S. Saito, *Yakugaku Zasshi*, 1970, **90**, 850 (*Chem. Abs.*, 1970, **73**, 77594m).

Compound	*Ref.*
β-alanine	
Boc-β-Ala	262
Arginine	
Z-Arg($Adoc_2$)	48
Z-Arg($Adoc_2$)-ONp	48
Z-Arg($Adoc_2$)-ONSu	48
Z(OMe)-Arg($Adoc_2$)	48
Z(OMe)-Arg($Adoc_2$)-ONSu	48
Arg($Adoc_2$)	48
Adoc-Arg($Adoc_2$)	48
Adoc-Arg($Adoc_2$)-ONp	48
Adoc-Arg($Adoc_2$)-ONSu	48
Nps-Arg($Adoc_2$)	48
Nps-Arg($Adoc_2$)-ONSu	48
Boc-Arg(NO_2)-ONp	75
Z-Arg(NO_2)-OPfp	78
Z-Arg(NO_2)-OPh(SMe)	85
Z-Arg(NO_2)-OPh(SO_2Me)	85
Arg(Tos) *N*-carboxy-anhydride hydrochloride	174
Asparagine	
Z-Asn[Bzh(OMe_2)]	46
Asn[Bzh(OMe_2)]	46
Asn[Bzh(OMe_2)]-OMe,HCl	46
Nps-Asn[Bzh(OMe_2)],Cha	46
Z-Asn[Bzh(OMe_2)]-OMe	46
Z-Asn[Bzh(OMe_2)]-OBu^t	46
Z-Asn[Bzh(OMe_2)]-ONSu	46
Nps-Asn[Bzh(OMe_2)]-ONSu	46
Nps-Asn[Bzh(OMe_2)]-ONp	46
Boc-Asn[Bzl(OMe)]	47
Boc-Asn[Bzl(OMe)]-OBzl	47
Aspartic acid	
Z-Asp(OPfp)-OBzl	78
Z-Asp(OPfp)-OBu^t	78
Z-Asp(OPic)	173
Z-Asp(OPic)-OTcp	173
Boc-Asp(OPic)	173
Citrulline	
Boc-Cit	146
Boc-Cit-OPcp	146
Cysteine	
Pht-Cys(Ibm), diethylammonium salt	52
Z-Cys(Bzl)-OPfp	78
Tfa-Cys[Bzl(OMe)]-NH·NH·Boc	144
Cys[Bzl(OMe)]-NH·NH·Boc	144
Cys[Z(OMe)]	63
Nps-Cys[Z(OMe)],Dcha	63
αβ-Diaminopropionic acid	
Tos-Dap(Z) [L and D]	316
Dap(Z) [L and D]	316
Nps-Dap(Z),Dcha [L and D]	316
Glutamic acid	
Z-Glu(OBu^t)-OPfp	78

Compound	*Ref.*
Z-Glu(OBu^t)-OPh(SMe)	85
Z-Glu(OBu^t)-OPh(SO_2Me)	85
Boc-Glu(OBu^t)-OPcp	151
Boc-Glu(OPic)	173
Boc-Glu(OBzl)-OQ	341
Glutamine	
Z-Gln[Bzh(OMe_2)]	46
Gln[Bzh(OMe_2)]	46
Z-Gln[Bzh(OMe_2)]-OMe	46
Gln[Bzh(OMe_2)]-OMe,HCl	46
Nps-Gln[Bzh(OMe_2)],Dcha	46
Z-Gln[Bzh(OMe_2)]-ONp	46
Boc-Gln[Bzl(OMe)]	47
Boc-Gln[Bzl(OMe)]-OBzl	47
Z-Gln-OPfp	78
Z-Gln-OPh(SMe)	85
Z-Gln-OPh(SO_2Me)	85
Boc-Gln-OPcp	151
Glycine	
Gly-NH·Bzh($OMe)_2$,HCl	46
Gly-NH·Bzl(OMe),HCl	47
Z-Gly-OPfp	78
Z-Gly-OPh(SMe)	85
Z-Gly-OPh(SO_2Me)	85
Gly-OPh(SMe),HBr	85
Boc-Gly-OPh(SMe)	85
Boc-Gly-OPh(SO_2Me)	85
Gly-OCH_2·$COHPh_2$,AcOH	43
Gly-NH·NH·Boc	144
Histidine	
Z-His(Bzl)-OPh(SMe)	85
Homoarginine	
Boc-Har(NO_2) [amorphous, as was the Dcha salt]	152
Boc-Har(NO_2)-OPcp	152
Homoserine	
Z-Hsr-OBzl(NO_2)	68
Isoleucine	
Z-Ile-OPfp [oil]	78
Z-Ile-OPh(SMe)	85
Z-Ile-OPh(SO_2Me)	85
Ile-OPh(SMe),Hbr	85
Leucine	
Z-Leu-OPfp [oil]	78
Z-Leu-OPh(SMe)	85
Z-Leu-OPh(SO_2Me)	85
Leu-OPh(SMe),Hbr	85
Lysine	
Nps-Lys(Nps),Dcha	17
Lys(Nps)	17
Z-Lys(Z)-OPfp	78

Pfp = Pentafluorophenyl; Ibm = Isobutoxymethyl; For = formyl.

[341] F. H. C. Stewart, *Austral. J. Chem.*, 1969, **22**, 1291.

Compound	*Ref.*
Lys(Z)-OBzl(Me_3),HCl	80
Z-Lys(Tfa)-OPh(SMe)	85
Z-Lys(Tfa)-OPh(SO_2Me)	85
Lys(Tfa)-OPh(SMe),HBr	85
Methionine	
Z-Met-OPfp	78
N-Methyl amino-acids	
Z-MePhe	37
Z-Sar-OPh(SMe)	85
Z-Sar-OPh(SO_2Me)	85
Sar-OPh(SMe),HBr	85
Boc-Sar-ONp	75
Ornithine	
Orn(For)	233
Boc-Orn(For) [Oil]	233
Boc-Orn(For)-ONSu	233
Tos-D-Orn(Pht)	316
Phenylalanine	
Phe-OBzl(OMe),HCl	42
Phe-NH·Bzl(OMe),HCl	47
Z-Phe-OPfp	78
Z-Phe-OPh(SMe)	85
Z-Phe-OPh(SO_2Me)	85
Phe-OPh(SMe),HBr	85
Phe-NH_2,HBr	288
Pipecolic acid	
A series of useful derivatives (Z-, Boc-, Nps-, and also active esters) of both D- and L-pipecolic acid	139
Boc-pipecolic acid	270
Proline	
Z-Pro-OPfp [oil]	78
Z-Pro-OPh(SMe)	85
Z-Pro-OPh(SO_2Me)	85
Pro-OPh(SMe),HBr	85
β-(3-Pyrazolyl)-alanine	
β-(3-Pyrazolyl)-alanine t-butoxycarbonylhydrazide	255
Pyroglutamic acid	
Glp-OPcp	251
Selenium-containing amino-acids	
Various useful derivatives of selenocysteine, selenohomocysteine, selenocystine, and selenomethionine	65, 68
Serine	
Z-Ser-OPfp	78
Threonine	
Boc-Thr-ONp	75
Boc-Thr(Bzl)-OPh(SMe)	85
Boc-Thr(Bzl)-OPh(SO_2Me)	85
Boc-Thr-Bu^t	342
Tyrosine	
Boc-Tyr-ONp	75
Z-Tyr(Bu^t)-OPh(SMe)	85
Z-Tyr(Bu^t)-OPh(SO_2Me)	85

[342] J. W. Moore and M. Szelke, *Tetrahedron Letters*, 1970, 4423.

Compound	*Ref.*
Tryptophan	
Z-Trp-OPfp	78
Z-Trp-OPh(SMe)	85
Z-Trp-OPh(SO_2Me)	85
Valine	
Val-OBzl(OMe),HCl	42
Tos-Val-ONp	75
Z-Val-OPfp	78
Z-Val-OPh(SMe)	85
Z-Val-OPh(SO_2Me)	85
Val-OPh(SMe),HBr	85

6 Appendix C. Some Recent Applications of Hydroxypropylated Dextran (Sephadex LH-20) in Peptide Chemistry

This list is not meant to be comprehensive but merely to give some idea of the range of recent applications of Sephadex LH-20 in peptide chemistry. It should be noted that applications using mixed solvents are essentially cases of partition chromatography, not gel filtration.[109]

Compounds purified	*Solvent*	*Ref.*
A range of fully-protected peptides (from one to ten residues) with *N*-terminal Nps-groups[a]	Dioxan or Me_2NCHO	343
Cyclopenta- and deca-peptides with protected side-chains: separated from each other[a]	MeOH	77, 344, 345, 346
Z-Gly-Phe-NH·Bzl[a]	MeOH	43
Trt-Gly-O·CH_2·$COHPh_2$[a]	MeOH	43
An antamanide (cyclodecapeptide)[a] analogue	MeOH	43
A t-butoxycarbonyldecapeptide[a] hydrazide	MeOH	43
Hexa- and deca-peptide (4-carboxy)-phenylhydrazides[a]	MeOH	43
A complex protected disulphide dimer of a tetradecapeptide-containing cysteine[a]	MeOH	57
A complex fully-protected disulphide dimer of a hexadecapeptide containing cysteine[a]	70% AcOH	56
Various acylamino-acid *p*-nitrophenyl esters prepared by use of bis-(*p*-nitrophenyl)sulphite: some which had previously been obtained as oils crystallised after this purification[b]	EtOAc	75
Complex peptide intermediates in a total synthesis of actinomycin D[c]	EtOAc or MeOH	76

[343] V. Gut and M. Cimrova, *Coll. Czech. Chem. Comm.*, 1969, **34**, 1620.
[344] S. Matsuura, M. Waki, S. Makisumi, and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 1197.
[345] M. Kondo and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 1850.
[346] M. Waki and N. Izumiya, *Bull. Chem. Soc. Japan*, 1967, **40**, 1687.

Compounds purified	*Solvent*	*Ref.*
A fully-protected heptapeptide amide obtained by the solid-phase method[d]	$CHCl_3$: AcOH : MeOH (90 : 5 : 5 by volume)	282
A protected pentapeptide acid formed by saponification of the corresponding ethyl ester: separated from the hydantoin which was the major product[b]	Unspecified	347
A fully-protected precursor of an oxytocin analogue obtained by the solid-phase method[e] (See also ref. 303)	The upper phase of CH_3Ph : AcOH : H_2O (20 : 20 : 3.5)	304
Fully-protected intermediates (nona- and hexa-peptides) in a trypsinogen nonapeptide fragment synthesis[a]	Dioxan or MeOH or Me_2NCHO	312
The product from coupling Z-Ser-N_3 with Thr-OEt: separated from the oxazolidinone by-product	Unspecified in *Chem. Abs.*	348

[a] Experimental detail given.
[b] No experimental detail given.
[c] *ca.* 5 g scale.
[d] Full experimental detail given: product obtained crystalline and sharply melting.
[e] Partition chromatography: column equilibrated with the lower phase of CH_3Ph : AcOH : H_2O (20 : 20 : 35).

[347] N. Izumiya, T. Kato, H. Aoyagi, S. Makisumi, M. Waki, O. Abe, and N. Mitsuyasu, ref. 11, p. 45.
[348] N. Mitsuyasu, M. Waki, S. Makisumi, and N. Izumiya, *Mem. Fac. Sci., Kyushu Univ., Ser. C*, 1969, **6**, 145 (*Chem. Abs.*, 1970, **73**, 77605r)

4

Peptides with Structural Features Not Typical of Proteins

BY J. S. DAVIES

1 Introduction

In the world of microbial peptide antibiotics, the trend of recent years continues with the isolation and characterisation of new structures with novel features. The biological rôle of these molecules is still, however, not well established. Elucidation of structure of peptide antibiotics, which are generally either cyclic peptides or cyclic depsipeptides, is at a highly developed stage. Once the major components of these molecules are identified, the application of techniques such as mass spectrometry to derivatives very often gives a good indication of the sequence of amino- or hydroxy-acid residues. The development of the art is such that sequence determination on a complex mixture of congeners is now possible. ^{1}H N.m.r. spectroscopy, in conjunction with many other physico-chemical techniques, has proved to be an invaluable tool for determining the preferred conformations of cyclic-homodetic and -heterodetic peptides. An increasing number of these peptides exhibit the property of co-ordinating to metal ions, and currently there is a great deal of interest in the ion-complexing and ion-transporting properties of these molecules. There is still some doubt as to whether the ion-complexing properties are responsible for their antibiotic activity. In fact, in one or two cases reported this year there is evidence to suggest that the antibiotic properties of these molecules are quite independent of complexation.

The ribosome-independent pathway to the biosynthesis of many of the microbial peptides is now well accepted and there is a great deal of interest in elucidating the mechanism by which selective peptide bond formation and cyclisation occur under enzyme catalysis. The isolation of enzyme systems capable of synthesising cyclic peptides of the tyrocidin and gramicidin series provides the first of the stepping-stones to what should be a rewarding and exciting field of study.

The increasing amount of information now being accumulated on the structure of bacterial cell wall glycopeptides (peptidoglycans) has initiated a great deal of interest in the exact nature of the interference of penicillin and related antibiotics in the biosynthesis of the cell walls. The β-lactam moiety can no longer be attributed only to penicillins, cephalosporins, and

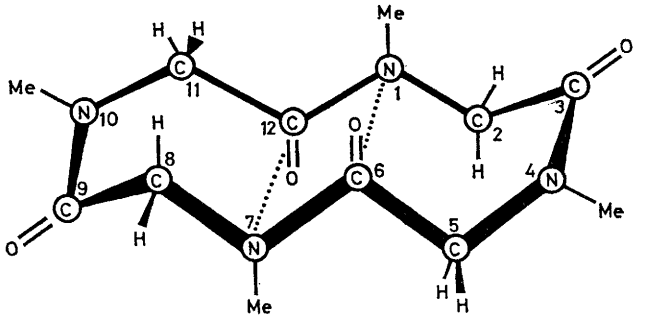
pachystermins since there is now good evidence for its existence in 'Woolley's wild-fire toxin,' which was thought to have a lactone structure.

The Proceedings of the First American Peptide Symposium, Yale University, 1968,[1] contain many papers of relevance to the subject matter of this chapter, although much of the work presented has now been published in various scientific journals. Various aspects of the application of spectroscopy to determine the conformation of peptides have been published [2] and the recent application of physical methods in determining the conformation of cyclic-homodetic and -heterodetic peptides has been reviewed.[3]

2 Cyclic Peptides

Three main themes appear to dominate the current interest in these compounds: (*a*) investigations directed at structure–activity relationships, (*b*) the conformation of cyclic peptides in solution and their ability to complex with suitable metal ions, and (*c*) the biosynthesis of these cyclic systems. Most of these themes obviously relate to molecules isolated from natural sources and are discussed below for specific compounds. 'Purpose-built' model systems also play an important rôle in studies on the application of physical and theoretical methods, since systems can be constructed so as to contain fewer variables. Thus a mathematical calculation has been applied [4] to *cyclo*-(Gly-Gly-Gly-Pro-Pro-) to calculate sets of dihedral angles which lead to exact ring closures, and to determine the local conformational deformation in cyclic peptides. *X*-Ray diffraction analysis [5] of conformation (Figure 1) exists in the crystal structure. Disregarding the *cyclo*-(Gly-Gly-D-Ala-D-Ala-Gly-Gly-),$3H_2O$ has shown that only one alanyl methyl groups, the molecule very nearly has a centre of symmetry, with all peptide units planar and in the *trans* conformation. The 18-membered ring is stabilised by two intramolecular hydrogen bonds and by hydrogen-bonding to water molecules. In an extension [6] of the work previously reported on cyclic *N*-methylglycine oligopeptides (see this Report, Vol. 2, p. 194), n.m.r. and i.r. studies show that *cyclo*-(Sar_3-MeAla-), *cyclo*-(Sar_3-Ala-), *cyclo*-(Sar_3-Gly-), and *cyclo*-(Sar-Gly-Sar-Gly-) have similar conformations to *cyclo*-(Sar_4-), whose crystal structure has also been obtained.[7] The *cyclo*-(Sar_4-) (1) possesses a rigid centrosymmetric conformation with an amide configuration sequence *cis*, *trans*, *cis*, *trans*. The methylene groups in the 2- and 8-positions have one outer 'equatorial'

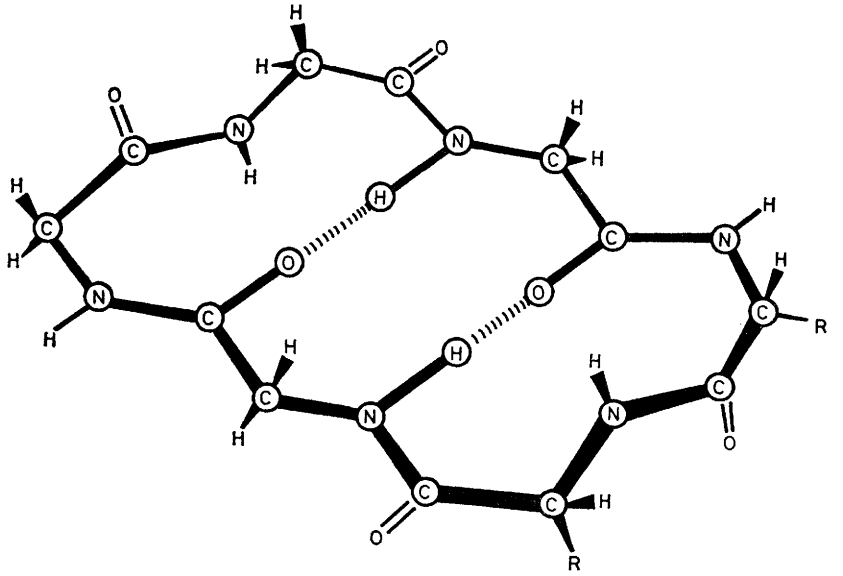
[1] 'Peptides: Chemistry and Biochemistry,' ed. B. Weinstein and S. Lande, Marcel Dekker, New York, 1970.
[2] 'Spectroscopic Approaches to Biomolecular Conformation,' ed. D. W. Urry, American Medical Association, 1970.
[3] C. H. Hassall and W. A. Thomas, *Chem. in Britain*, 1971, 145.
[4] N. Go and H. A. Scheraga, *Macromolecules*, 1970, **3**, 178, 188; N. Go, P. N. Lewis, and H. A. Scheraga, *ibid.*, p. 628.
[5] I. L. Karle, J. W. Gibson, and J. Karle, *J. Amer. Chem. Soc.*, 1970, **92**, 3755.
[6] J. Dale and K. Titlestad, *Chem. Comm.*, 1970, 1403.
[7] P. Groth, *Acta Chem. Scand.*, 1970, **24**, 780.



(1)

hydrogen and one 'axial', while those in the 5- and 11-positions are of the straight-chain type. The α-methyl group of *N*-methylalanine takes up the outer equatorial 2- or 8-position.

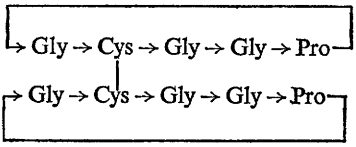
With the current interest in the ion-complexing properties of cyclic peptides from natural sources, it is not surprising that the criteria necessary for complexation have been investigated using a model system. The bicyclic peptide (2) appears to be the first example of this type to be synthesised.[8] E.m.f. measurements using membranes containing cyclopeptide



R = Me

Figure 1

[8] R. Schwyzer, T. K. Aung, M. Caviezel, and P. Moser, *Helv. Chim. Acta*, 1970, **53**, 15; R. Schwyzer, *Experientia*, 1970, **26**, 577.



(2)

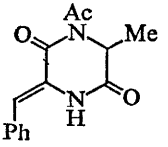
(2) revealed a cation specificity in the order $K^+ > Na^+ > Li^+ > Ca^{2+}$. O.r.d. measurements show that the disulphide bond possesses a 75—90° dihedral angle and a right-handed helicity. The bicyclic system was synthesised by the cyclisation of Gly-(*S*-diphenylmethyl)-L-Cys-Gly-Gly-L-Pro-ONp (see Table 1, p. 291). Deprotection and oxidation of the thiol group furnished the disulphide link.

A. 2,5-Dioxopiperazines.—Recently, n.m.r. studies have shown that the favoured rotamer of dioxopiperazines containing aromatic rings in their side-chains has the aromatic ring and the dioxopiperazine rings in a face-to-face proximity. A study [9] of the near-u.v. circular dichroism bands confirms these results for dioxopiperazines containing a single aromatic residue. In dioxopiperazines having two aromatic amino-acids with L configuration, both aromatic rings share the space [9] over the dioxopiperazine ring. The chemical shifts of the *cis*-6-protons of a series of *para*-substituted 3-benzyl-2,5-dioxopiperazines suggest [10] that there is no major donor–acceptor interaction involved in stabilising the face-to-face arrangement of rings in such systems.

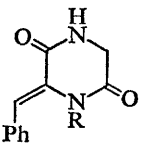
In a parallel study, independent of work reported last year, it has been further confirmed [11] that *cyclo*-(D-Ala-L-Ala-) has a planar ring system while the ring in *cyclo*-(L-Ala-L-Ala-) is appreciably puckered into a skew-boat conformation. The molecular dimensions obtained in this work are different from the previous report but no explanation has been suggested for the puckering of the LL-form. It is suggested, however, that this puckering of the ring system brings about a slight twisting effect in the amide bonds and might account for the difference in hydrolysis rates [12] of the DL- and LL-forms.

The fact that the dioxopiperazine (3), rather than the isomer containing a 4-acetyl group, is formed when 3-methyl-2,5-dioxopiperazine is condensed with aromatic aldehydes and acetic anhydride has been attributed [13] to the steric hindrance of the 4,5-amide bond. Photolysis converts the compound (3) into its geometrical isomer, and both isomers react with base to give tautomeric pyrazines. The steric hindrance of the 4,5-amide bond has been

[9] E. H. Strickland, M. Wilchek, J. Horowitz, and C. Billups, *J. Biol. Chem.*, 1970, **245**, 4168.
[10] Z. Kopple and K. D. Kopple, *J. Org. Chem.*, 1970, **35**, 253.
[11] E. Sletten, *J. Amer. Chem. Soc.*, 1970, **92**, 172.
[12] O. Grahl-Nielsen, *Tetrahedron Letters*, 1969, 2827.
[13] K. W. Blake and P. G. Sammes, *J. Chem. Soc.* (*C*), 1970, 980.

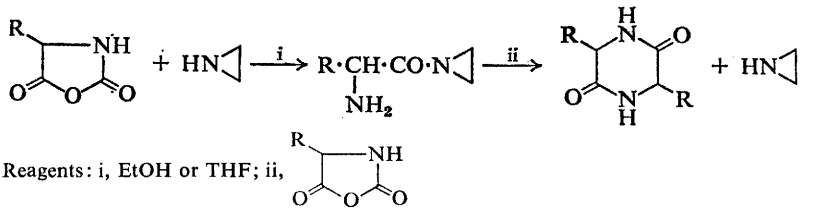


(3)



(4) R = Ac
(5) R = H

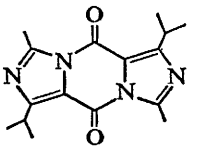
used as evidence [14] to query the structure of the acetyl intermediate (4) obtained in the synthesis [15] of 3-benzylidene-2,5-dioxopiperazine (5). Leuchs anhydrides (1,3-oxazolidine-2,5-diones) react [16] with aziridines to give 2,5-dioxopiperazines in high yield (70—80%). The reaction appears to be specific to aziridine and probably proceeds *via* the route summarised in Scheme 1.



Reagents: i, EtOH or THF; ii,

Scheme 1

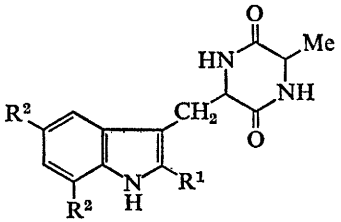
3,6-Di-isobutyl-2,5-dioxopiperazine undergoes [17] an anodically-induced reaction in the presence of methyl cyanide and sodium perchlorate, giving (6) as the only recognisable product. A series of DL- and LL-3,6-disubstituted-2,5-dioxopiperazine diastereoisomers have been synthesised [18] from their corresponding dipeptide ester hydrohalides using methanolic ammonia. The o.r.d. and c.d. properties of such dioxopiperazines have also been investigated [19] in several solvents and Cotton effects have been assigned to



(6)

aromatic $\pi \rightarrow \pi^*$, amide $n \rightarrow \pi^*$, and amide $\pi \rightarrow \pi^*$ transitions. [14] C-Labelling studies have shown [20] that *cyclo*-(L-Ala-L-[^{14}C]Trp-) can be incorporated directly into echinulin (7) in *Aspergillus amstelodami* without prior hydrolysis into tryptophan. A possible rôle for dioxopiperazines in

[14] A. E. A. Porter and P. G. Sammes, *J. Chem. Soc.* (*C*), 1970, 2530.
[15] B. W. Dominy and R. G. Lawton, *J. Org. Chem.*, 1969, **34**, 2013.
[16] P. Rosenmund and K. Kaiser, *Angew. Chem. Internat. Edn.*, 1970, **9**, 162.
[17] L. A. Simonson and C. K. Mann, *Tetrahedron Letters*, 1970, 3303.
[18] K. Bláha, *Coll. Czech. Chem. Comm.*, 1969, **34**, 4000.
[19] K. Bláha and I. Frič, *Coll. Czech. Chem. Comm.*, 1970, **35**, 619.
[20] G. P. Slater, J. C. MacDonald, and R. Nakashima, *Biochemistry*, 1970, **9**, 2886.



$R^1 = -CMe_2-CH=CH_2$
$R^2 = -CH_2-CH=CMe_2$

(7)

the biosynthesis of cyclic peptides has been suggested on more than one occasion, so the successful incorporation of an intact moiety in this case might justify investigations on the more general rôle of these compounds in Nature.

The molecular packing in a dioxopiperazine crystal has been determined [21] using van der Waals and hydrogen-bonding calculations in an attempt to assess the application of the technique to larger molecules of biological interest.

B. Gramicidins.—The topochemical requirements for antibacterial activity have been investigated further with the synthesis [22–28] of a number of analogues of gramicidin S. (The synthetic procedures are discussed below.) The lack of antibacterial activity [22] shown by *cyclo*-(L-Val-L-Orn-L-Leu-D-Phe-L-Pro-Gly-Gly-), which has only seven amino-acid residues in the ring, and by [β-Ala$^{5,5'}$]-gramicidin S,[23] whose ring has two extra carbon–carbon bonds, indicates a critical requirement for a specific ring size. Significant antibacterial activity exists [24] in the analogue where the 3,3′-leucyl residues are replaced by L-alanine, but [Gly$^{3,3'}$]-gramicidin S showed no activity. [Leu$^{1,1'}$]-Gramicidin S is as active as natural gramicidin S, so valyl residues can be replaced by leucyl without loss of activity,[25] but replacement of valyl side-chains by the hydroxymethyl groups of serine destroys [26] the activity. The cyclic pentapeptide analogues (cyclosemigramicidin S analogues) [22–26] have also been tested but in each case no activity was observed. New syntheses using a polymer carrier have been reported for gramicidin S [27] and [Gly$^{5,5'}$]-gramicidin S.[28] A possible biosyn-

[21] P. Giacomello and E. Giglio, *Acta Cryst.*, 1970, **A26**, 324.
[22] O. Abe, K. Kuromizu, M. Kondo, and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 914.
[23] S. Matsuura, M. Waki, S. Makisumi, and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 1197.
[24] O. Abe and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 1202.
[25] M. Kondo and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 1850.
[26] M. Iwai, K. Nakajima, A. Uno, S. Hase, I. Takeuchi, and K. Okawa, *Bull. Chem. Soc. Japan*, 1970, **43**, 3246.
[27] G. Losse and K. Neubert, *Tetrahedron Letters*, 1970, 1267.
[28] L. U. Sklyarov and I. V. Shashkova, *Zhur. obshchei Khim.*, 1969, **39**, 2778.

thetic precursor of gramicidin S, HCO-D-Phe-Pro-Val-Orn-Leu-D-Phe-Pro-Val-Orn-Leu-NH·$(CH_2)_2$·OH, has been synthesised[29] by linear expansion from the *C*-terminal amino-alcohol using conventional methods, but the peptide showed no antibiotic activity against Gram-positive organisms.

O.r.d., i.r., and n.m.r. studies[30] on gramicidin S and *NN'*-diacetylgramicidin S provide further proof for the Hodgkin–Oughton–Schwyzer model for the molecule (Figure 2). The *NN'*-diacetyl derivative contains

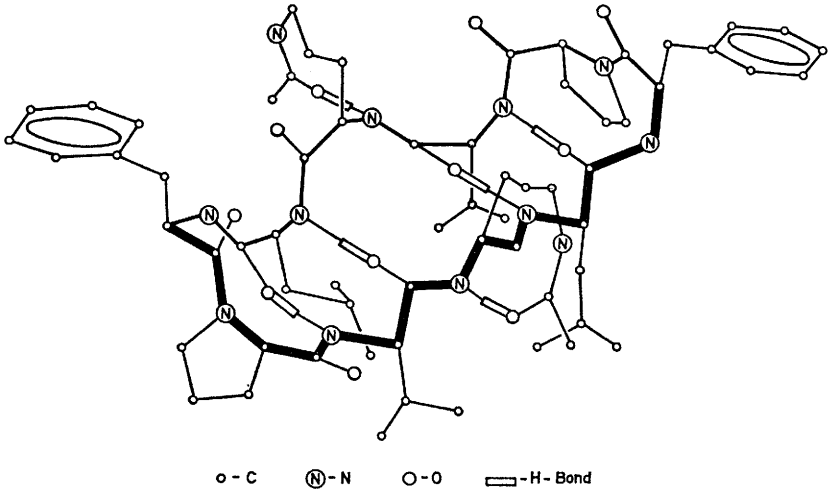


Figure 2 *Conformation of* NN'*-diacetylgramicidin S in non-polar solvent* (Reproduced by permission from *Biochem. Biophys. Res. Comm.*, 1970, **39**, 217)

six hydrogen-bonds, four strong hydrogen bonds being due to valine and leucine NH groups and two weaker hydrogen bonds due to the ornithine α-NH group. From the N*H*–$C_\alpha H$ proton coupling constants it can be shown that these protons are *trans* in the valyl, ornithyl, and leucyl residues and *gauche* in D-phenylalanyl. The ^{13}C n.m.r. spectrum of gramicidin S appears[31] to be more easily interpretable than proton spectra. The interpretation of ^{13}C-spectra is at an early stage of development but it has been possible, using individual amino-acids for comparison, to divide the spectra into regions corresponding to: (*a*) carbonyl carbon atoms (19—24 p.p.m.), (*b*) aromatic and hetero-aromatic carbon atoms (50—70 p.p.m.), (*c*) C_α-carbon atoms (130—140 p.p.m.), (*d*) C_β-carbon atoms (150—165 p.p.m.),

[29] S. Makisumi and N. Izumiya, *Tetrahedron Letters*, 1970, 227.

[30] Yu. A. Ovchinnikov, V. T. Ivanov, V. F. Bystrov, A. I. Miroshnikov, E. N. Shepel, N. D. Abdullaev, E. S. Efremov, and L. B. Senyavina, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 217; K. A. Zykalova, G. N. Tischenko, G. A. Kogan, and V. T. Ivanov, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1970, 1547.

[31] W. A. Gibbons, J. A. Sogn, A. Stern, L. C. Craig, and L. F. Johnson, *Nature*, 1970, **227**, 840.

and (*e*) C_γ and methyl carbon atoms (165—180 p.p.m.). Preliminary *X*-ray studies [32] have shown that a copper complex of the Schiff base formed between salicylaldehyde and the two ornithyl residues of gramicidin S has a similar conformation to gramicidin S. The detailed knowledge now available on the conformation of gramicidin S makes it a useful model for testing new ideas and methods. An example of this is the application of optical activity calculations [33] as a means of screening the low-energy conformations found by minimisation routines from primary sequences.

The enzyme system participating in the biosynthesis of gramicidin S has been fractionated [34] into four protein fractions (F-1 to F-4). Each separate enzyme fraction failed to form gramicidin S, but the antibiotic was formed when F-3 and F-4 were combined, although addition of F-1 was essential for full activity. The only amino-acid-activating enzyme which was separated was the D-phenylalanine-activating enzyme. The mechanism of gramicidin S biosynthesis has been shown [35] to be basically as depicted in Scheme 2, with the last step being a head-to-tail condensation of two

$$E_{SH} + AA + ATP \longrightarrow E_{SH}^{AA\text{-}AMP} + PP_i \longrightarrow E_{S\text{-}AA} + AMP$$

$$E_{S\text{-}AA}^{AA\text{-}AMP} + PP_i \xleftarrow{AA + ATP}$$

AA = amino-acid

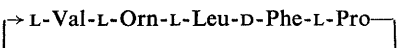
Scheme 2

pentapeptide units. A rôle for phosphopantothenic acid is also envisaged in the process.

C. Tyrocidins.—A great deal of effort has also been expended in the synthesis of the components of the tyrocidin family. In addition to tyrocidin A (8),[36] synthesised in 1966, complete syntheses of tyrocidins B (9),[37] C (10),[38] and E (11) [39] have now been accomplished. The methods used for cyclisation are discussed later. The synthetic tyrocidins B, C, and E showed biological activity very similar to that of tyrocidin A, so replace-

[32] G. Camilletti, P. De Santis, and R. Rizzo, *Chem. Comm.*, 1970, 1073.
[33] E. S. Pysh, *Science*, 1970, **167**, 290.
[34] S. Otani, T. Yamanoi, and Y. Saito, *Biochim. Biophys. Acta*, 1970, **208**, 496.
[35] Ø. Frøshov, T. L. Zimmer, and S. G. Laland, *F.E.B.S. Letters*, 1970, **7**, 68; C. C. Gilhuus-Moe, T. Kristensen, J. E. Bredesen, T. L. Zimmer, and S. G. Laland, *ibid.*, p. 287.
[36] M. Ohno, T. Kato, S. Makisumi, and N. Izumiya, *Bull. Chem. Soc. Japan*, 1966, **39**, 1738.
[37] K. Kuromizu and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 2199; *Experientia*, 1970, **26**, 587.
[38] K. Kuromizu and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 2944; *Tetrahedron Letters*, 1970, 1471.
[39] N. Mitsuyasu, S. Matasuura, M. Waki, M. Ohno, S. Makisumi, and N. Izumiya, *Bull. Chem. Soc. Japan*, 1970, **43**, 1829; N. Mitsuyasu and N. Izumiya, *Experientia*, 1970, **26**, 476.

ment of the D- and L-phenylalanyl and L-tyrosyl residues in tyrocidin A by other very similar amino-acids does not alter the activity.



(8) X = Tyr, Y = Phe, Z = Phe

(9) X = Tyr, Y = Phe, Z = Trp

(10) X = Tyr, Y = Trp, Z = Trp

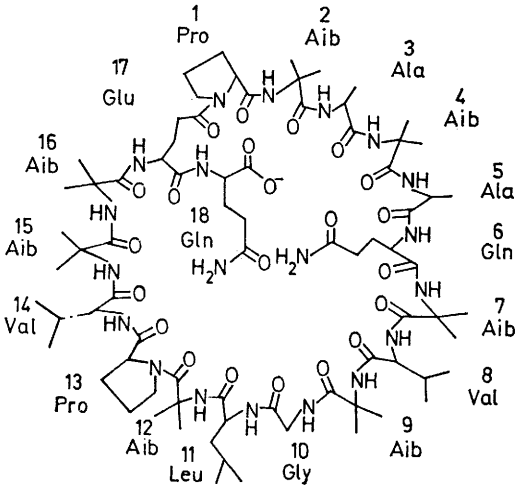
(11) X = Phe, Y = Phe, Z = Phe

More information has been obtained on the nature of the enzyme systems capable of synthesising tyrocidins. Three complementary fractions of an enzyme system isolated [40] from *Bacillus brevis* are capable of activating certain amino-acids for peptide bond synthesis. A light fraction (M.W. 100 000) and an intermediate fraction (M.W. 230 000) activate phenylalanine and proline respectively. A heavy fraction (M.W. 460 000) activates the remaining constituent amino-acids of tyrocidin. The protein-bound amino-acid intermediates can be isolated using Sephadex G-50. Structurally similar amino-acids can be utilised by the enzyme system to produce tyrocidin analogues. The activation of each amino-acid is depicted in Scheme 2. On comparing the enzyme systems [34, 40] capable of synthesising the gramicidins and the tyrocidins there appears to be some relationship between the molecular weight of the enzyme fraction and the number of amino-acids it can activate. This might be an indication of further enzyme subunits specific for each amino-acid within each larger fraction, but so far there is no proof for this. It is also interesting to note [41] that the cyclisation point to form the cyclic peptide for both gramicidin S and tyrocidin occurs between L-leucine and D-phenylalanine. Phosphopantetheine has also been shown [41] to be an enzyme-bound co-factor in tyrocidin biosynthesis, its rôle being in the later stages of peptide chain formation.

Previously published evidence from n.m.r. studies (see last year's Report) that tyrocidins appear to self-associate in solution to form aggregates has been augmented by the results of the thin-film dialysis technique.[42] Tyrocidin B dialyses in acetic acid as an aggregate of five or six monomers.

D. Alamethicin.—This peptide, which transports cations and induces action potentials in synthetic membranes, has been shown [43] to be a cyclic peptide (12). Alamethicin and derived peptides are immune to enzymic

[40] R. Roskoski, H. Kleinkauf, W. Gevers, and F. Lipmann, *Biochemistry*, 1970, **9**, 4839, 4846.
[41] H. Kleinkauf, W. Gevers, R. Roskoski, and F. Lipmann, *Biochem. Biophys. Res. Comm* 1970, **41**, 1218.
[42] M. Burachik, L. C. Craig, and J. Chang, *Biochemistry*, 1970, **9**, 3293.
[43] J. W. Payne, R. Jakes, and B. S. Hartley, *Biochem. J.*, 1970, **117**, 757.



(12)

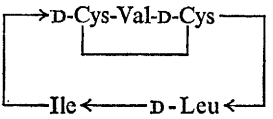
digestion, so identification of fragments from partial acid hydrolysis had to be used extensively in the characterisation procedures. Proof that the α-carboxy-group of glutamine-18 is free was obtained from diborane reduction. The large ring in alamethicin is probably flexible enough to adopt a variety of conformations, and affords a plausible explanation for the poor discrimination that alamethicin shows in its complexing properties with alkali metals. The mode of action of alamethicin is far from clear, but model building does allow the possibility of a structure that could stack to form a tunnel with lypophilic exterior and hydrophilic interior. This could then facilitate the 'tunnelling' of metal ions across membranes. Studies on the ion-gating properties of alamethicin show [44] that micelle formation occurs in dilute solution, the aggregates increasing in size as concentration increases.

E. Peptides from *Amanita phalloides*.—The affinity of antamanide, *cyclo*-(Pro-Phe-Phe-Val-Pro-Pro-Ala-Phe-Phe-Pro-), towards metal ions has been investigated [45] using a number of physico-chemical techniques. It forms a distinct complex with Na^+ ions which is more stable than that formed with K^+ ions. An analogue of antamanide containing D- instead of L-alanine has been synthesised [46] using two novel approaches to cyclisation (see Table 1). The yields (12 and 20%) in the cyclisation step were lower than those from other methods previously used. The labile nature of the γ-hydroxyacylpeptide bond in phalloin has posed difficulties in its synthesis.

[44] A. I. McMullen, *Biochem. J.*, 1970, **119**, 10P.
[45] T. Wieland, H. Faulstich, W. Burgermeister, W. Otting, W. Moehle, M. M. Shemyakin, Yu. A. Ovchinnikov, V. T. Ivanov, and G. G. Malenkov, *F.E.B.S. Letters*, 1970, **9**, 89.
[46] T. Wieland, J. Lewalter, and C. Birr, *Annalen*, 1970, **740**, 31.

The fragment L-Trp-L-γ-Hyleu*-L-Ala-Thr-OMe, however, has now been synthesised [47] by reacting Z-DL-Hyleu-lactone in molten imidazole with the sodium salt of L-alanine to form Z-DL-γ-Hyleu-Ala, which can be utilised for further coupling. The γδ-dihydroxyisoleucine side-chain of *O*-methyl-α-amanitin has been oxidised [48] with periodate to give the aldehyde, which is not toxic. Sodium borohydride reduction of this aldehyde gives a compound with a γ-hydroxy-group, and which is again toxic. A γ-hydroxy-group appears therefore to be one of the requirements necessary for toxicity.

F. Malformins.—The structure (13) for malformin, isolated from *Aspergillus niger*, has been confirmed by synthesis [49] of a series of analogues. The peptide ring was first formed by using dicyclohexylcarbodi-imide–*N*-hydroxysuccinimide and the disulphide ring was closed using 1,2-di-iodoethane. The cyclic peptides were purified by gel filtration on Sephadex LH-20 with a dimethylformamide–ethanol mixture as elution solvent.

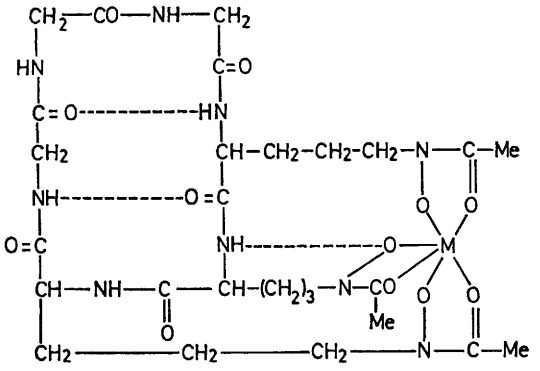


(13)

G. Ferrichrome.—^{1}H N.m.r. studies [50] at 220 MHz on desferri-ferrichrome and its Al^{3+} complex give further support to the structure, which was determined by the *X*-ray method, which contains two hydrogen bonds. The ^{1}H n.m.r. studies favour a further hydrogen bond, as shown in (14), between the carbonyl oxygen of an ornithine residue deeply sheltered between chelated *N*-acetyl-*N*-hydroxyornithyl residues and a glycyl *N*-proton in the peptide cycle itself. Large changes are induced in the ^{1}H n.m.r. spectrum, especially in the amide resonance regions, when the cyclopeptide co-ordinates with Al^{3+} ion. Chemical shifts in the amide region, normally covering a width of 0.6 p.p.m., become spread over 3.6 p.p.m. in the complex. A total synthesis [51] of ferrichrome (14, M = Fe^{3+}), giving rise to a product having full biological activity, has been achieved using δ-nitronorvaline as the precursor of the hydroxy-amino-acid in the final sequence. Cyclisation of the *p*-nitrophenyl ester of (15) gave a 70—80% yield of the cyclohexapeptide. Nitro-groups were then reduced, acetylated, and selectively deacetylated to give the required cyclopeptide.

[47] H. Faulstich and H. Trischmann, *Annalen*, 1970, **741**, 55.
[48] T. Wieland and A. Fahrmeir, *Annalen*, 1970, **736**, 95.
[49] A. Schoeberl, M. Rimpler, and E. Clauss, *Naturwiss.*, 1969, **56**, 516; *Chem. Ber.*, 1970, **103**, 2252; 3159; *Annalen*, 1970, **742**, 68.
[50] M. Llinas, M. P. Klein and J. B. Neilands, *J. Mol. Biol.*, 1970, **52** 399.
[51] W. Keller-Schierlein and B. Maurer, *Helv. Chim. Acta*, 1969, **52**, 603.

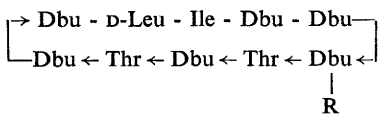
* γ-Hyleu = γ-hydroxyleucine.



(14)

Gly-[Nva(δ-NO_2)]$_3$-Gly-Gly

(15)

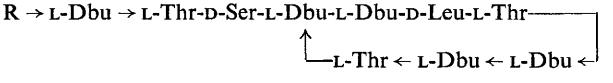


R = (+)-Et(Me)CH·$(CH_2)_4$·CO-

(16)

H. Circulins.—The cyclodecapeptide structure (16) previously proposed for circulin B is now questionable since the synthetic product [52] bearing this sequence showed distinct differences from the naturally occurring antibiotic. An analogue [53] of the proposed structure *cyclo*-[Dbu(R)-*a*Thr-Dbu-*a*Thr-Dbu-Dbu-D-Leu-Ile-Dbu-Dbu-], where R = (+)-6-methyloctanoyl, also shows no biological activity. Cyclisation of the linear decapeptide in each case was carried out using the azide method.

I. Polymyxins.—The cycloheptapeptide (17) has been synthesised [54] and corresponds to the proposed structure for polymyxin D_1. Cyclisation was carried out at the bond between D-leucine and the threonine residue using dicyclohexylcarbodi-imide. A series of analogues of polymyxin M containing L-$\alpha\gamma$-diaminobutyric acid has also been synthesised.[55] Yields of



R = (+)-Et(Me)CH·$(CH_2)_4$·CO—

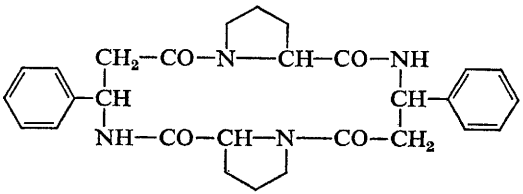
(17)

[52] H. Arold, *Annalen*, 1970, **731**, 152.
[53] H. Arold, *Annalen*, 1970, **739**, 194.
[54] R. O. Studer and W. Lergier, *Helv. Chim. Acta*, 1970, **53**, 929.
[55] E. M. S. Salem, N. V. Fedoseeva, and A. B. Silaev, *Zhur. obshchei Khim.*, 1969, **39**, 2541; *ibid.*, 1970, **40**, 480, 655.

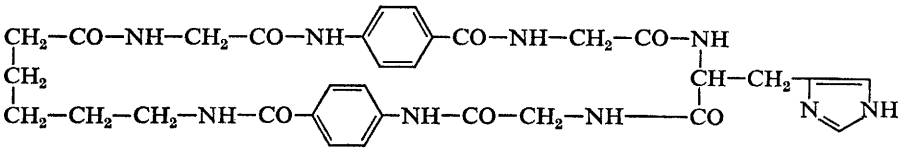
approximately 45% were reported utilising the azide method for cyclisation, although if D-leucine in *cyclo*-[D-Leu-L-Dbu(Z)-L-Dbu(Z)-L-Thr-L-Dbu(Z)-L-Dbu(Z)-L-Thr-] is replaced by L-leucine the yields drop to 15%. The Merrifield technique has been used to synthesise [56] the linear heptapeptide corresponding to the cyclopeptide part of polymyxin B. A biologically active mono-*N*-(acetyl-^{14}C) derivative of polymyxin B has also been prepared.[57]

Further work [58] on the enzyme system capable of activating L-$\alpha\gamma$-diaminobutyric acid (one of the amino-acids found in polymyxins) shows that the mechanism does not invove the intermediacy of tRNA. This brings the polymyxins into line with the gramicidins and tyrocidins. Similar results have also been obtained [59] for the cell-free system capable of synthesising colistin.

J. Other Cyclic Peptides.—A cyclic tetrapeptide (18) has been isolated [60] from the lichen *Roccella canariensis*. N.m.r. studies strongly support structure (18), although the alternative with two similar amino-acids next



(18)



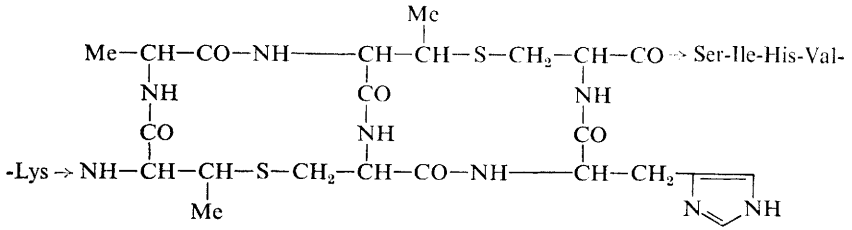
(19)

to one another is theoretically possible. Interest continues in the quest for model enzyme systems showing catalytic activity. The cyclic peptide (19) has been synthesised [61] but the cyclopeptide is less active than imidazole as a catalyst for the hydrolysis of 2,4-dinitrophenyl acetate. A novel bicyclic system (20) has been suggested [62] for a biologically active tridecapeptide isolated from nisin after cleavage with cyanogen bromide.

[56] E. A. Morozova and M. A. Zevail, *Khim. prirod. Soedinenii*, 1970, **6**, 359.
[57] M. Teuber, *Z. Naturforsch.*, 1970, **25b**, 117.
[58] J. Monreal and H. Paules, *Biochim. Biophys. Acta*, 1970, **199**, 280.
[59] M. Ito, Y. Koyama, K. Alda, and T. Uemura, *Biochem. Biophys. Acta*, 1970, **215**, 418.
[60] G. Bohmann, *Tetrahedron Letters*, 1970, 3065.
[61] A. R. Mitchell, S. K. Gupta, and R. W. Roeske, *J. Org. Chem.*, 1970, **35**, 2877.
[62] E. Gross and J. L. Morell, *J. Amer. Chem. Soc.*, 1970, **92**, 2919.

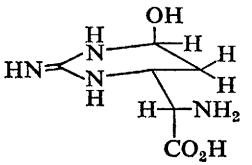
Evidence for the structure was obtained from the results of selective performic acid treatment.

Formation of the bacitracin–Zn^{2+} complex has been studied [63] by spectrophotometric titration, o.r.d., and n.m.r. measurements. The thiazoline



(20)

ring and the nitrogen at the 3-position of the imidazole ring of histidine provide the two co-ordinating sites for the metal ions. Preliminary studies [64] on tuberactinomycin, a peptide antibiotic from *Streptomyces griseoverticillatus var. tuberacticus*, show that the component amino-acids are L-serine, L-$\alpha\beta$-diaminopropionic acid, γ-hydroxy-L-β-lysine, and a guanido amino-acid, tuberactidine (21). Tuberactidine is converted to



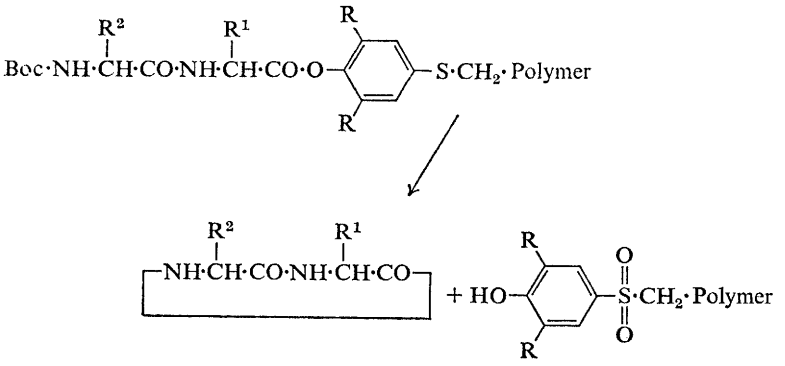
(21)

viomycidine (an artefact in the hydrolysis of viomycin) in the presence of trifluoroacetic acid. The characteristics of the tuberactidine structure suggest that it could well be a part structure of both tuberactinomycin and viomycin.

K. Synthesis of Homodetic Cyclic Peptides.—The major approaches to the cyclisation of linear peptides were critically reviewed last year (this Report, Vol. 2, p. 165). Few changes have occurred over the past year, and it appears that cyclisation *via* active esters is still the method of choice, used by most workers. Details of the methods and the yields obtained have been summarised in Table 1. The yields obtained from the active ester method are consistently good and justify the continued use of the method. Cyclohexapeptides built from D- and L-alanine and glycine have also been

[63] N. W. Cornell and D. H. Guiney, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 530.
[64] T. Shiba, T. Wakamiya, T. Kaneko, H. Sakakibara, T. Take, and J. Abe, *Tetrahedron Letters*, 1970, 3497.

synthesised [65] by this method but the details were not available when this compilation was being collected. However, new methods are being investigated, and an interesting prospect appears to be the possibility of activating the carboxy-end for cyclisation while it is already linked [66] to the polymeric support used in the Merrifield technique. It involves the incorporation of the 4-(methylthio)phenyl (MTP) group [67] into a polymer and this is then activated by oxidation to the sulphone, as shown in Scheme 3.



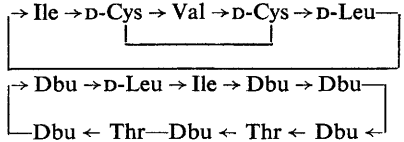
Scheme 3

The yields of a series of simple cyclopeptides ranged from 20 to 60%. Consecutive activation of suitable derivatives [46] such as phenylhydrazides and 2,2-diphenylglycol esters, using *N*-bromosuccinimide and trifluoroacetic acid respectively, gave only moderate yields (12 and 20%) when applied to the synthesis of [D-Ala]-antamanide.

Cyclic peptides containing a repeating sequence of amino-acid residues can theoretically be synthesised by dimerising appropriately activated derivatives of the monomer sequence. Gramicidin S and analogues are ideal examples for assessing the potential of such an approach. In most of the cases studied [22–26] cyclisation to the cyclic monomer predominates over cyclodimer formation, and although the separation of cyclic monomers from cyclic dimers has been greatly assisted by Sephadex LH-20 gel filtration, this approach still appears to be less satisfactory than the cyclisation of the linear dimeric sequence. The degree of monomer formation is, however, influenced [25] by the nature and shape of the *N*-terminal side-chain in the pentapeptide active esters, as shown in Table 2. *N*-Terminal valine gives a very favourable cyclic-dimer : monomer ratio using this method, and with glycine in position 4 no monomer is formed.

[65] V. T. Ivanov, V. V. Shilin, and Yu. A. Ovchinnikov, *Zhur. obshchei Khim.*, 1970, **40**, 924.
[66] E. Flanigan and G. R. Marshall, *Tetrahedron Letters*, 1970, 2403.
[67] B. J. Johnson and T. A. Ruettinger, *J. Org. Chem.*, 1970, **35**, 255.

Table 1 *Syntheses of cyclic peptides reported in* 1970

Peptide	*Bond constructed in cyclisation step*	*Method*	*% yield*	*Ref.*
cyclo-(L-Val-L-Orn-L-Leu-D-Phe-L-Pro-Gly-Gly-)	Gly–Val	*a*	50	22
[β-Ala$^{5,5'}$]-gramicidin S	β-Ala–Val	*a*	42	23
[Gly$^{3,3'}$]-gramicidin S	Pro–Val	*b*	16	24
[Ala$^{3,3'}$]-gramicidin S	Pro–Val	*b*	16	24
[Leu$^{1,1'}$]-gramicidin S	Pro–Leu	*a*	39	25
[Ser$^{1,1',3,3'}$]-gramicidin S	Pro–Ser	*a*	45	26
[Orn(Tos)$^{2,2'}$]-gramicidin S	Pro–Val	*c*	42	27
[Gly$^{5,5'}$]-gramicidin S	Gly–Val	*d*	20	28
tyrocidin B	Pro–Trp	*a*	43	37
tyrocidin C	Pro–Trp	*a*	59	38
tyrocidin E	Pro–Phe	*a*	40	39
[D-Ala]-antamanide	Phe–Phe	*e*	12	46
[D-Ala]-antamanide	Phe–Phe	*f*	20	46
	Cys–D-Leu	*d*	50	49
	Dbu–Dbu	*g*	68	52
cyclo-[Dbu(R)-*a*Thr-Dbu-*a*Thr-Dbu-Dbu-D-Leu-Ile-Dbu-Dbu]	Dbu(R)–Dbu	*g*	72	53
polymyxin D_1	D-Leu–L-Thr	*h*	?	54
S,*S*′-bis*cyclo*-[Gly-L-Cys-(Gly)$_2$-Pro]	Pro–Gly	*a*	68	8
cyclo-[D-Leu-L-Dbu(Z)-L-Dbu(Z)-LThr-L-Dbu(Z)-L-Dbu(Z)-L-Thr]	Thr–D-Leu	*g*	48	55
cyclopeptide (19)	ε-aminocaproyl–Gly	*i*	25	61
cyclo-[Gly$_3$-(δ-nitro-norvalyl)$_3$]	Gly–Gly	*a*	75	51

[a] using the *p*-nitrophenyl ester method with pyridine as base
[b] dimerisation of pentapeptide *p*-nitrophenyl ester using pyridine
[c] DCCI–*p*-nitrophenol in pyridine at 20 °C
[d] DCCI–*N*-hydroxysuccinimide–pyridine
[e] activation of a phenylhydrazide
[f] activation of 2,2-diphenylglycol esters
[g] azide method
[h] DCCI–Me_2NCHO
[i] excess DCCI

Table 2 *Ratios of cyclic-monomer to cyclic-dimer obtained from the cyclisation of linear pentapeptide active esters*

p-*Nitrophenyl ester of*[a] (1 2 3 4 5)	*Cyclic monomer*	*Cyclic dimer*
H-Gly-Orn(Z)-Leu-D-Phe-Pro-OH	100	0
H-Ala———————————OH	91	9
H-Leu———————————OH	78	22
H-Val———————————OH	32	68
H-Gly-Lys(Z) ———————OH	100	0
H-Val-Lys(Z) ———————OH	29	71
H——Orn(Z)-Gly ———————OH	59	41
H————— Ala————————OH	43	47
H ————————Leu-Gly ———OH	0	100
H ———————————D-Phe-Gly-OH	79	21

[a] A horizontal line represents an amino-acid residue unchanged from the residue directly above it in the Table.

3 Depsipeptides (Heterodetic Peptides)

A very comprehensive review on peptide lactones (cyclodepsipeptides) has appeared.[68] The recent success (see last year's Report) in the use of physico-chemical techniques for determining the conformations of valinomycin and enniatin has been discussed [69] in the context of the rôle of these molecules in studying ion-transport through membranes. A review [3] of the applications of physico-chemical techniques to determine conformation also contains very relevant discussions on cyclodepsipeptides.

A. Thiostrepton.—The hard-won results of degradative studies [70] on thiostrepton, a metabolic product of *Streptomyces azureus*, have been confirmed and augmented by *X*-ray crystallography.[71] Thiostrepton is a much-modified peptide containing a number of thiazoline rings. Because no heavy-atom derivative of thiostrepton could be prepared, the *X*-ray structure, shown in Figure 3, was based on the initial identification of the positions of five sulphur atoms using the technique of multiple superpositions on three-dimensional Patterson syntheses. Some doubt still remains as to the nature of the five terminal atoms of the long side-chain.

B. Valinomycin.—Further work on the details of the conformation of valinomycin $-K^+$ ion complex has been reported.[72] The K^+ ion is pictured as being held in a polar core, with the depsipeptide side-chains presenting a non-polar exterior to the surrounding medium. Selective ion permeation

[68] A. Taylor, *Adv. Appl. Microbiol.*, 1970, **12**, 189.
[69] M. M. Shemyakin, Yu. A. Ovchinnikov, V. T. Ivanov, V. K. Antonov, E. I. Vinogradova, A. M. Shkrob, G. G. Malenkov, A. V. Evstratov, I. A. Laine, E. I. Melnik, and I. D. Ryabova, *J. Membrane Biol.*, 1969, **1**, 402.
[70] M. Bodanszky, J. A. Scozzie, and I. Muramatsu, *J. Amer. Chem. Soc.*, 1969, **91**, 4934 and refs. cited therein.
[71] B. Anderson, D. C. Hodgkin, and M. A. Viswamitra, *Nature*, 1970, **225**, 233.
[72] M. Ohnishi and D. W. Urry, *Science*, 1970, **168**, 1091.

of membranes is then envisaged as a carrier-type mechanism. There is currently much speculation as to whether ion-complexing and biological function bear any relationship to one another in peptide antibiotics. Valinomycin (and nonactin) inhibits [73] photophosphorylation and photoreduction in the *Rhodospirillium rubrum* chromatophore, but there appears to be no link between this and the ion-transporting properties of the antibiotics.

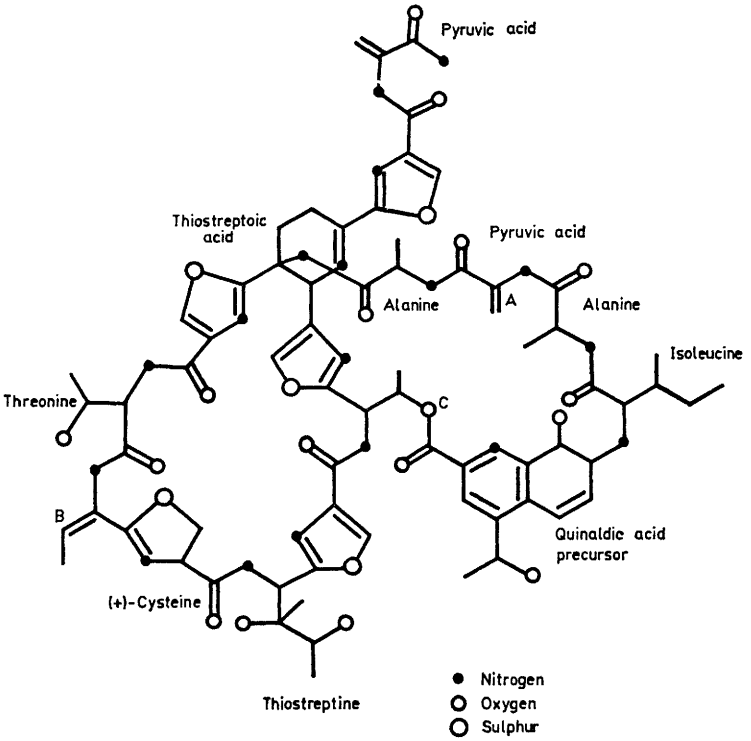
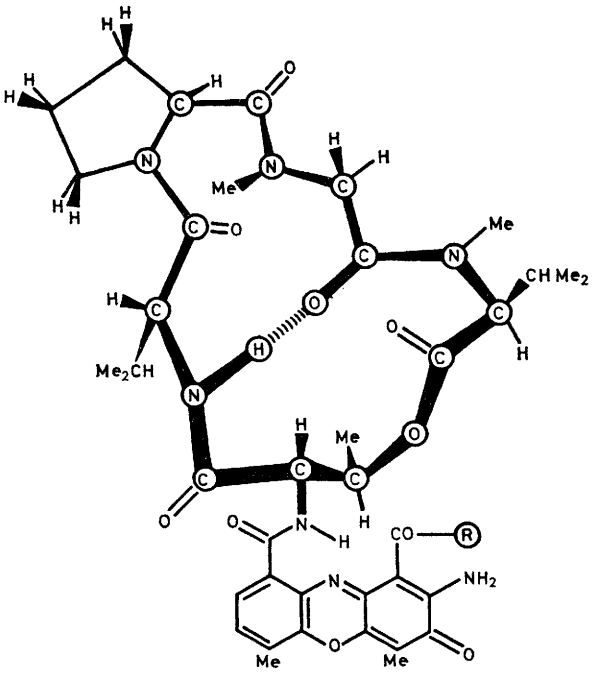


Figure 3 *The structure of thiostrepton*
(Reproduced by permission from *Nature*, 1970, **225**, 233)

C. Actinomycins.—The problems to be solved in the determination of the conformation and rotational isomerism of the actinomycins have been highlighted.[74] From an n.m.r (100 MHz) study [75] made on one of the pentapeptide lactone rings, it can be shown that a stretched (Val)N-H···O=C(Sar) hydrogen bond exists, giving rise to the conformation (22). However, when the two lactone rings are considered in actinomycin

[73] Z. Gromet-Elhanan, *Biochem. Biophys. Acta*, 1970, **223**, 174.
[74] H. Lackner, *Tetrahedron Letters*, 1970, 2807.
[75] H. Lackner, *Tetrahedron Letters*, 1970, 3189.



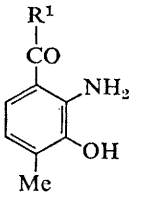
(22)

D itself, then n.m.r. evidence [76] shows that chemically equivalent protons, especially the N-H protons of threonine and valine residues, are not magnetically equivalent. The existence of hydrogen bonds between the N-H of D-valine and the carbonyl group of *N*-methylvaline of the other lactone ring is implied and the differences between the two rings may be ascribed to the non-equivalence of the conformations of L-threonine because of the different attachment of the peptide rings to the chromophoric group. The reversible inversion of sign that takes place in c.d. spectra [77] of actinomycin D in going from methyl cyanide to hexafluoroacetone hydrate has been explained in terms of the two conformations obtained by rotation of the peptide lactone rings around the bond connecting them to the phenoxazone ring. A complete assignment [78] of the 220 MHz n.m.r. spectrum of actinomycin D has been reported, and spectral studies in solution show that actinomycin D forms a complex with 5′-deoxyguanylic acid in which the pyrimidine ring of the acid is located above and below the phenoxazone chromophore.

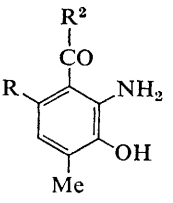
A full account of the total synthesis of actinomycin D (C_1) has appeared.[79] The cyclisation step to form the peptide lactone was made at

[76] F. Conti and P. De-Santis, *Nature*, 1970, **227**, 1239.
[77] F. Ascoli, P. De-Santis, and M. Savino, *Nature*, 1970, **227**, 1237.
[78] B. H. Arison and K. Hoogsteen, *Biochemistry*, 1970, **9**, 3976.
[79] J. Meienhofer, *J. Amer. Chem. Soc.*, 1970, **92**, 3771.

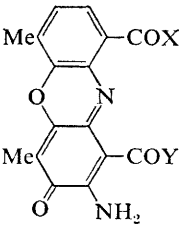
the proline–sarcosine bond using the *p*-nitrophenyl ester for activation. A yield of 26% was obtained and many of the intermediates were purified by Sephadex LH-20 gel filtration. This example, together with others considered previously, emphasises the success of gel filtration as a separation technique in cyclic peptide chemistry. Oxidative coupling of a mixture of actinocinyl derivatives such as (23) and (24, R = H), where R^1 and R^2 represent pentapeptides, can theoretically give four isomers. When the intermediate is deuteriated as in (24, R = D), n.m.r. spectroscopy can be



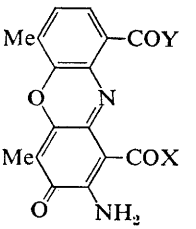
(23)



(24)



(25)



(26)

X = L-Thr-D-Val-L-Pro-Sar-L-MeVal

Y = L-Thr-D-*a*Ile-L-Pro-Sar-L-MeVal

used to distinguish between structural isomers. Using this technique [80] actinocinyl pentapeptide pairs have been synthesised and characterised, and a synthesis of a non-separable pair of *aniso*-actinomycins, actinomycins C_2 (25), and *i*-C_2 (26) has been achieved. Interest continues in simple chloro-actinocinyl analogues of actinomycins, and a number of analogues have been synthesised.[81] An actinomycin lactonase which hydrolyses only actinomycin lactone rings has been purified.[82] The pure enzyme (M.W. $\approx$ 60 000) hydrolyses only one of the lactone rings, while the crude

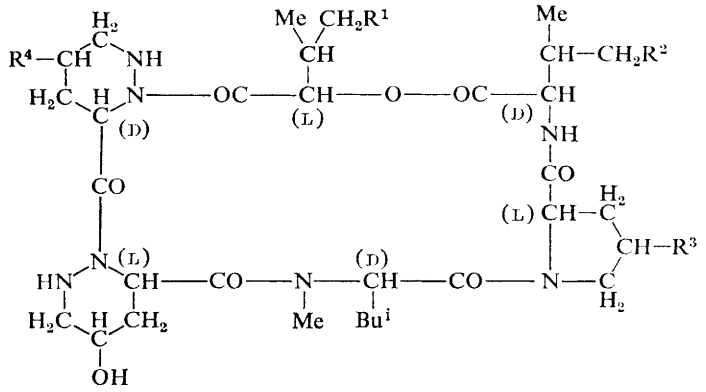
[80] H. Lackner, *Chem. Ber.*, 1970, **103**, 2476.

[81] E. N. Glibin, Z. I. Korshunova, O. F. Ginzburg, and L. F. Larionov, *Khim. geterotsikl. Soedinenii*, 1969, 602; V. G. Sinitsyn, E. N. Glibin, and O. F. Ginzburg, *Zhur. org. Khim.*, 1970, **6**, 500; Z. I. Korshunova, E. R. Zakhs, and O. F. Ginzburg, *ibid.*, p. 504; E. N. Glibin, V. G. Sinitsyn, and O. F. Ginzburg, *ibid.*, p. 1020.

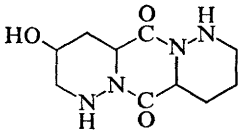
[82] C. T. Hou and D. Perlman, *J. Biol. Chem.*, 1970, **254**, 1289.

enzyme preparation hydrolyses both rings. Two enzyme systems may therefore exist.

D. Monamycins.—The crystalline compound isolated [83] from cultures of *Streptomyces jamaicensis* has been shown [84] to be a mixture of fifteen cyclohexadepsipeptides. The structure of the main component, D_1, is given in (27), together with the substitutions that make up the other congeners in the series. Much of the evidence for the sequence was based on high-resolution mass spectrometry of the methyl ester of the linear peptide



(27) R^1 = Me, R^2 = H, R^3 = Me, R^4 = H.
(Substitutions R^1 = H or Me; R^2 = H or Me; R^3 = H or Me; R^4 = H or Cl)



(28)

formed after cleaving the depside link with mild alkali, and from partial hydrolysis, which yielded the interesting dioxopiperazine (28) as one of the products. The residues have alternating D and L configurations in the ring system, and a detailed i.r. and n.m.r. study indicates [3] that the most stable conformation is that given in Figure 4. The monoamycins complex with K^+, Rb^+, and Cs^+ but not readily with Na^+ or Li^+, thus showing similar specificities to those of the actins. It is likely, however, that the mode of

[83] M. J. Hall and C. H. Hassall, *Appl. Microbiol.*, 1970, **19**, 109.
[84] K. Bevan, J. S. Davies, M. J. Hall, C. H. Hassall, R. B. Morton, D. A. S. Phillips, Y. Ogihara, and W. A. Thomas, *Experientia*, 1970, **26**, 122; K. Bevan, J. S. Davies, C. H. Hassall, R. B. Morton, and D. A. S. Phillips, *J. Chem. Soc.* (*C*), 1971, 514; C. H. Hassall Y. Ogihara, and W. A. Thomas, *ibid.*, p. 522; C. H. Hassall, R. B. Morton, Y. Ogihara, and D. A. S. Phillips, *ibid.*, p. 625.

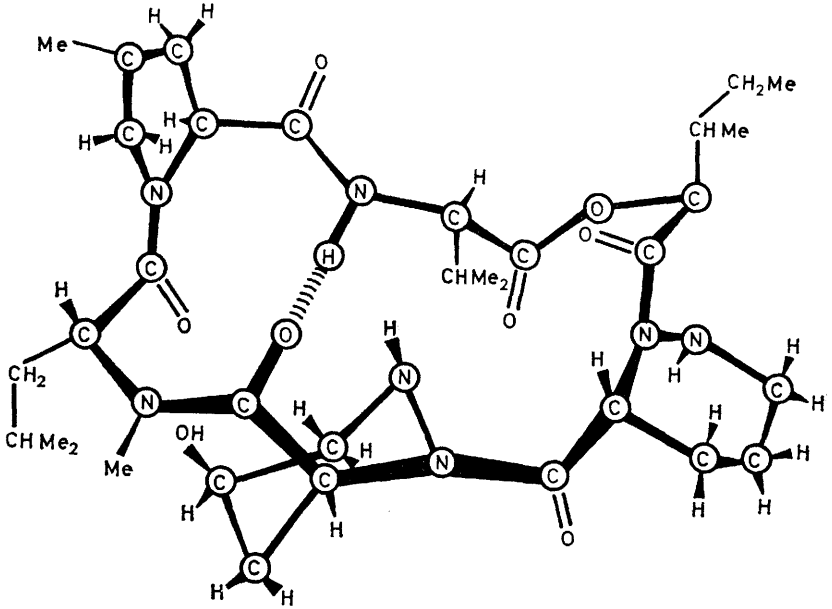


Figure 4 *The conformation of monamycin*

action of these antibiotics is due to their lytic effects on cell membranes rather than to any ion-transport properties.[85]

E. Viscosin.—The revised structure (29) has now been suggested [86] for viscosin in the light of new data from the mass spectra of *N*-permethylated derivatives of the linear peptide (viscosic acid) obtained on hydrolysis of

D-Me$(CH_2)_6$·CH(OH)·CH_2·CO-L-Leu-D-Glu-D-*a*Thr-D-Val-L-Leu-D-Ser

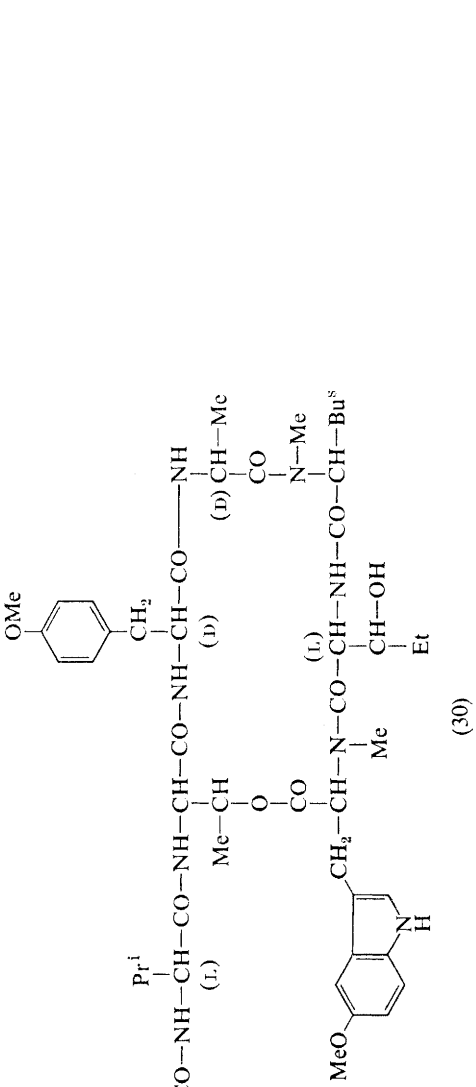
L-Ile ← D-Ser ← L-Leu ←

(29)

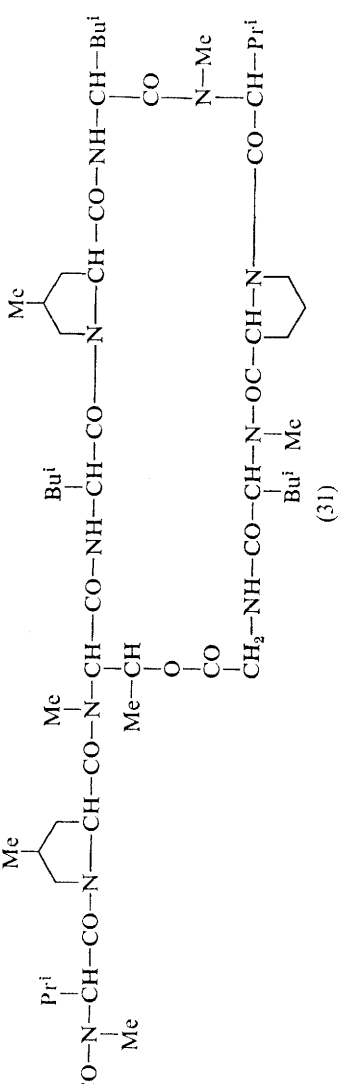
the lactone ring. The position of the lactone ring was confirmed by comparing the products of the chromic acid oxidation of viscosic acid and viscosin.

F. Cycloheptamycin.—The mass-spectral breakdown patterns of *N*-permethylated derivatives have also played an important rôle [87] in the characterisation of cycloheptamycin (30), isolated from an unidentified *Streptomyces* species. It is the first time that *N*-methyl-5-methoxytryptophan,

[85] M. J. Hall, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 590.
[86] M. Hiramoto, K. Okada, and S. Nagai, *Tetrahedron Letters*, 1970, 1087.
[87] W. O. Godtfredsen, S. Vangedal, and D. W. Thomas, *Tetrahedron*, 1970, **26**, 4931.



(30)

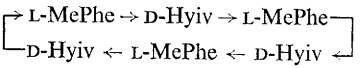


(31)

β-hydroxynorvaline, and the D-form of *O*-methyltyrosine have been found in Nature.

G. Griselimycin.—Physical methods and classical degradative evidence support the structure (31) for this antibiotic.[88] The side-chain sequence was revealed using the Edman procedure after prior removal of the acetyl group, using anhydrous hydrogen chloride. The sequence of the cyclic part was obtained from the mass spectrum of a trifluoroacetyl derivative of the nonapeptide.

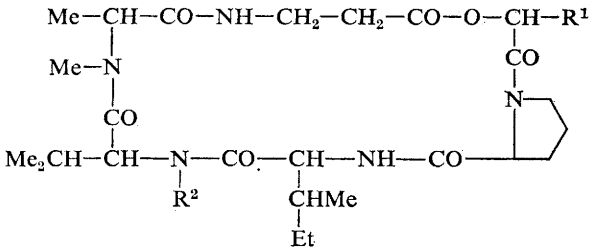
H. Beauvericin.—The ion-complexing and antimicrobial properties[89] of beauvericin (32) bear out the close similarity in structure between the



(32)

antibiotic and the enniatins. As in the enniatins, beauvericin exists as an equilibrium mixture of conformers (a non-polar form and a polar form), the composition of the mixture depending on the polarity of the medium and the presence of cations. Cyclisation[89] of the linear depsipeptide* L-MePhe-D-Hyiv-L-MePhe-D-Hyiv-L-MePhe-D-Hyiv using thionyl chloride–triethylamine gives beauvericin in 25% yield.

I. Destruxins.—Three new insecticidal destruxins, desmethyldestruxin B (33), destruxin C (34), and destruxin D (35), have been isolated[90] from



(33) $R^1 = Me_2CH \cdot CH_2$, $R^2 = H$

(34) $R^1 = HO \cdot CH_2 \cdot CH(Me) \cdot CH_2$, $R^2 = Me$

(35) $R^1 = HO_2C \cdot CH(Me) \cdot CH_2$, $R^2 = Me$

[88] B. Terlain and J. P. Thomas, *Compt. rend.*, 1969, **269**, *C*, 1546.

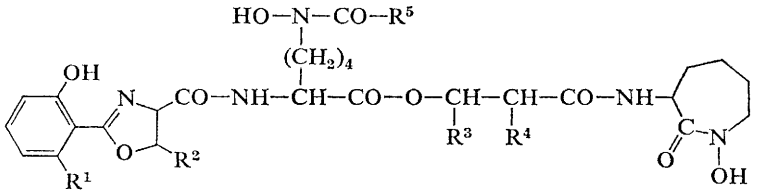
[89] Yu. A. Ovchinnikov, V. T. Ivanov, and I. I. Mikhaleva, *Tetrahedron Letters*, 1971, 159.

[90] A. Suzuki, H. Taguchi, and S. Tamura, *Agric. and Biol. Chem.* (*Japan*), 1970, **34**, 813.

* Hyiv = α-hydroxyisovaleric acid.

culture filtrates of *Metarrhizium anisopliae*. The mass-spectral breakdown of the methyl ester of the linear peptide produced on mild hydrolysis of the depside link was used to determine the sequence. There appears to be an obvious biosynthetic relationship between members of the series although no experimental data to confirm this have yet been published.

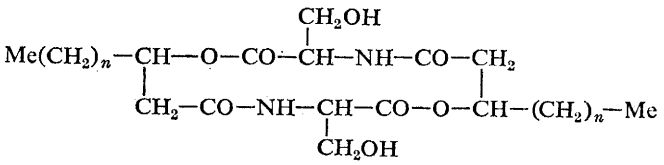
J. Mycobactins.—New mycobactins containing different side-chains have been obtained[91] from various species of Mycobacteria grown in media deficient in iron. Structures (36), (37), (38), and (39) have been given to



(36) $R^1 = Me$, $R^2 = H$, $R^3 = Et$, $R^4 = Me$, $R^5 = C_{17}H_{33}$
(37) $R^1 = H$, $R^2 = H$, $R^3 = Me$, $R^4 = H$, $R^5 = C_{17}H_{33}$
(38) $R^1 = H$, $R^2 = H$, $R^3 = Me$, $R^4 = H$, $R^5 = C_{20}H_{37}$
(39) $R^1 = Me$, $R^2 = Me$, $R^3 = Me$, $R^4 = H$, $R^5 = C_{17}H_{37}$

mycobactins P, S, T, and H respectively. ^{14}C-Labelling studies[92] show conclusively that the phenolic ring in mycobactin S is derived from shikimic acid, and lysine has been shown[93] to be the source of the hydroxamic acid moiety in the molecule.

K. Miscellaneous.—Further investigation of the acetone extracts of *Serratia marcescens* has revealed[94] the existence of a family of 14-membered cyclic depsipeptides related to serratamolide (40). The other



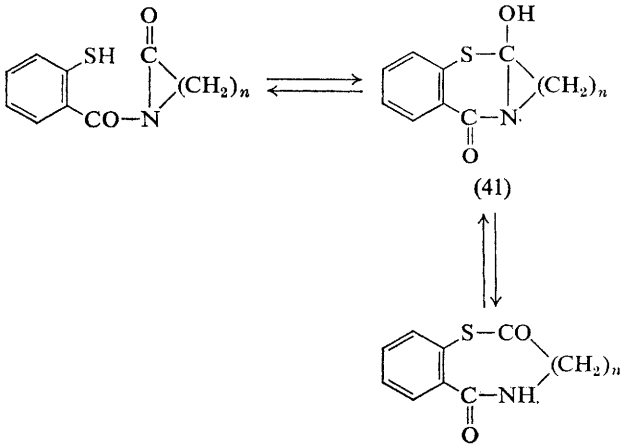
(40, $n = 6$)

members of the family are made up of different combinations of fatty acids, *e.g.* with $n = 8$, with serine. On addition of low concentrations of streptomycin to the culture of *S. marcescens*, production of the serratamolide antibiotics was completely inhibited.[95]

[91] A. J. White and G. A. Snow, *Biochem. J.*, 1969, **111**, 785.
[92] A. T. Hudson and R. Bentley, *Tetrahedron Letters*, 1970, 2077,
[93] J. E. Tateson, *Biochem. J.*, 1970, **118**, 747; M. Allen, A. J. Birch, and A. R. Jones, *Austral. J. Chem.*, 1970, **23**, 427.
[94] M. A. C. Bermingham, B. S. Deol, and J. L. Still, *Biochem. J.*, 1970, **116**, 759.
[95] M. A. C. Bermingham, B. S. Deol, and J. L. Still, *Biochem. J.*, 1970, **119**, 861.

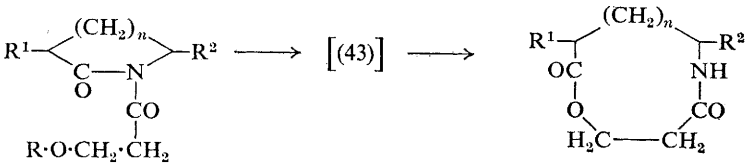
From c.d. data,[96] stendomycin (see this Report, Vol. 1, p. 222) appears to be relatively stable in trifluoroethanol, but in water the c.d. pattern exhibits an inverse temperature transition; at low temperatures the pattern is characteristic of disordered polypeptides, while at temperatures > 40 °C a stable c.d. pattern which is indicative of more order is observed. An assessment of the detailed significance of these effects will have to await comparison of the results with those of conformations derived using other techniques.

The spectral properties of the intermediate and the factors which govern the equilibrium (Scheme 4) and rearrangements of aromatic thiocyclols



Scheme 4

(41) have been investigated.[97] A series of model cyclodepsipeptides have been synthesised[98] from *N*-benzyloxyacyl-lactams (42) *via* the hydroxy-acyl insertion of (43). Yields of between 32 and 88% were recorded for the examples taken. Vibrational spectra and dipole moment measurements



(42) R = $PhCH_2$
(43) R = H

[96] D. W. Urry and A. Ruiter, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 800.
[97] M. Rothe and R. Steinberger, *Tetrahedron Letters*, 1970, 649, 2467.
[98] L. I. Andreeva, A. G. Lyakisheva, V. K. Antonov, and M. M. Shemyakin, *Zhur. obshchei Khim.*, 1969, **39**, 2774.

show that the ratio of *cis*- and *trans*-arrangement of amide groups depends [99] on the size of ring and on the nature of the substituents on the ring. The conformations of the ten-membered cyclodepsipeptide rings are favourable for intramolecular interaction between amide and ester groups but this interaction is much weaker in the 11- and 12-membered rings.

The 2-quinoxalinecarboxylic acid residue, a component of the quinomycins and triostins, has been linked [100] to a peptide or amino-acid attached to a polymer support, using dicyclohexylcarbodi-imide.

The details of a symposium lecture on the synthesis of depsipeptides have appeared.[101]

4 Peptide–Carbohydrate Linkages

A. Glycopeptides.—*Structural Aspects.* As a result of the great deal of information currently available on the nature of the bridges that cross-link the peptide subunits in bacterial cell walls, the existence of four main types [102] of peptidoglycans has been revealed. This differentiation appears to be useful as a criterion which has some considerable taxonomic importance. Of the four types, chemotype II [102] appears to be the most widely distributed peptidoglycan structure. Most of the variations found for this type of subunit are summarised in Figure 5. An example of this type of subunit, bridged by asparagine (as in Figure 5), has been identified [103]

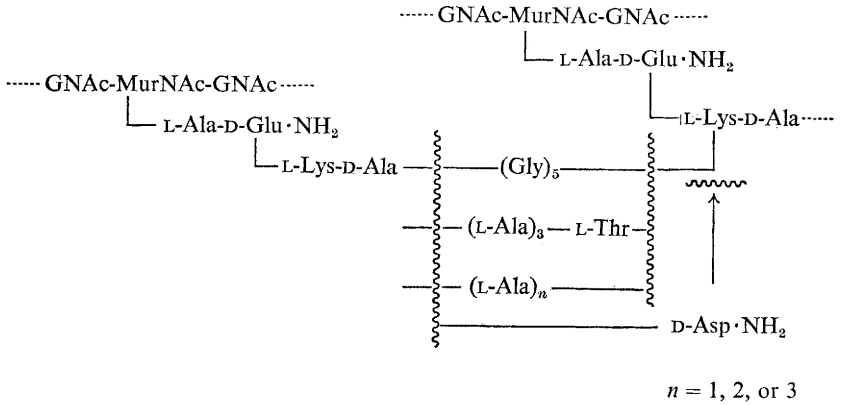
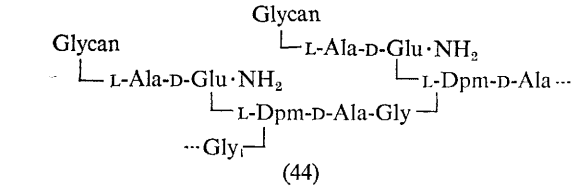


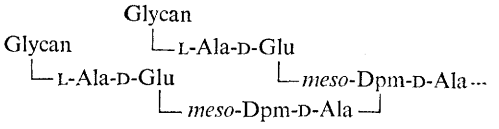
Figure 5 *Summary of typical chemotype II subunit structure.* GNAc = N-*acetyl*-D-*glucosamine;* MurNAc = N-*acetylmuramic acid. Variations in the cross-linking between the subunits are shown between the wavy lines*

[99] L. I. Andreeva, T. M. Ivanova, E. P. Efremov, V. K. Antonov, and M. M. Shemyakin, *Zhur. obshchei Khim.*, 1970, **40**, 475.
[100] R. K. Olsen, *J. Heterocyclic Chem.*, 1970, **7**, 435.
[101] R. B. Merrifield, B. F. Gisin, D. C. Tosteson, and M. Tieffenberg, in a Report of a Symposium on the 'Molecular Basis of Membrane Function,' 1969, p. 211.
[102] J.-M. Ghuysen, J. L. Strominger, and D. J. Tipper, *Comp. Biochem.*, 1968, **26A**, 53; J.-M. Ghuysen, *Bacterial Rev.*, 1968, **32**, 425.
[103] J. Coyette and J.-M. Ghuysen, *Biochemistry*, 1970, **9**, 2935.

in the cell walls of *Lactobacillus acidophilus*. Structure elucidation in this case was again carried out using enzymic breakdown followed by conventional methods of analysis. Yet another variation on the chemotype II structure is that found in glycopeptides isolated[104] from *Streptomyces* species and *Clostridium perfringens*. In confirming the previous work on the nature of the cell walls of these bacteria, this work shows that the part structure for the glycopeptide can be represented as (44). Myxobacter AL-1 endopeptidase hydrolyses both the D-Ala–Gly and Gly–(LL)Dpm linkages in *C. perfringens* but only hydrolyses the former linkage in *Streptomyces* species. No D-Ala-D-Ala units were identified as *C*-termini in the walls of these organisms, thus indicating the presence in the organism



(44)



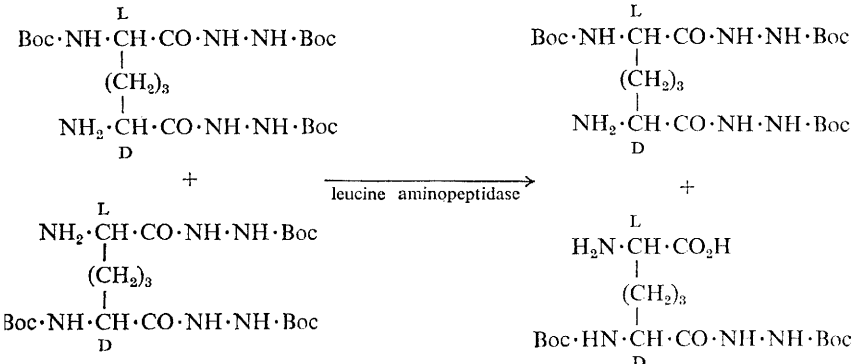
(45)

of carboxypeptidases similar to those of *E. coli*. The structures of the tri- and hepta-peptides from the autolysates[105] of *Bacillus stearothermophilus* show that the peptide network in the cell wall is similar to the chemotype I found in *E. coli*, with direct linking between D-alanine and diaminopimelic acid residues as shown in (45). A similar structure is found[106] in glycopeptides isolated after autolysis of *Bacillus licheniformis* and *B. subtilis*. The autolytic enzyme in these cases hydrolyses the amide bond between the carboxy-function of the *N*-acetylmuramic acid residue and the L-alanyl residue of the peptide unit. A glycopeptide, GNAc-MurNAc-L-Ala-D-Glu-Dpm-Lys-Arg,* has been isolated[107] from the products of enzymic breakdown of the murein lipoprotein complex of *E. coli*. The D-alanyl residue of chemotype I (45) is now replaced by the lysyl-arginyl dipeptide of the lipoprotein. A peptidoglycan containing *meso*-diaminopimelic acid has been identified[108] from studies carried out

[104] M. Leyh-Bouille, R. Bonaly, J.-M. Ghuysen, R. Tinelli, and D. Tipper, *Biochemistry*, 1970, **9**, 2944.
[105] W. D. Grant and A. J. Wicken, *Biochem. J.*, 1970, **118**, 859.
[106] R. C. Hughes, *Biochem. J.*, 1970, **119**, 849.
[107] V. Braun and H. Wolff, *European J. Biochem.*, 1970, **14**, 387.
[108] R. Tinelli, M. Shilo, M. Laurent, and J.-M. Ghuysen, *Compt. rend.*, 1970, **270**, *D*, 2600.

* GNAc = *N*-acetyl-D-glucosamine, MurNAc = *N*-acetylmuramic acid.

on cell walls of *Bdellovibrio bacteriovorus*, a microbial parasite of some Gram-negative bacteria. The unambiguous synthesis [109] of the tetrapeptide L-Ala-γ-D-Glu-(L)-*meso*-Dpm(L)-D-Ala has confirmed the structure of the peptide moiety in *E. coli* and *B. megaterium*. The problems associated with the introduction of a *meso*-diaminopimelic acid residue into the sequence were overcome by using a symmetrical derivative of the acid, either the bis-t-butoxycarbonylhydrazide (46) or its benzyloxycarbonylhydrazide analogue, and allowing these to be stereoselectively cleaved by



(46) (47)

enzymes at the hydrazide bond in the α-position of a free amino-group belonging to the L-moiety to give (47).

The lability towards alkali of *O*-glycoside bonds linked to hydroxy-groups of β-hydroxy-substituted amino-acids continues to play a major rôle in the characterisation of carbohydrate–peptide linkages. The results [110] of treating keratin sulphate from whale cartilage with alkali support the involvement of serine and threonine in such linkages. Earthworm cuticle collagen appears to have [111] similar linkages. It is the first time that such a linkage has been found in collagen; hydroxylysine and hydroxyproline are usually involved. Serine and threonine hydroxy-groups have also been implicated in the carbohydrate–peptide link in the phosphomannan peptide from yeast cell walls.[112] However, the only linkage found [113] in the glycopeptides from Pronase digests of Lorenzini jelly from the elasmobranch fish was that of 2-acetamido-2-deoxy-D-galactosylthreonine.

The *O*-glycosidic–amino-acid bonds are not the only ones degraded by alkali treatment. In contrast to peptidoglycans of several bacilli and lactobacilli, glycopeptides from *Staphylococci* dissolve fairly readily in

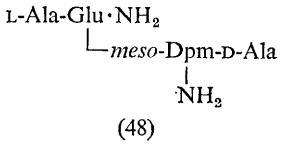
[109] P. Dezélée and E. Bricas, *Biochemistry*, 1970, **9**, 823.
[110] N. Seno and N. Toda, *Biochim. Biophys. Acta*, 1970, **215**, 544.
[111] L. Muir and Y. C. Lee, *J. Biol. Chem.*, 1970, **245**, 502.
[112] T. N. Cawley and R. Letters, *Biochem. J.*, 1969, **115**, 9P.
[113] M. J. How, J. V. S. Jones, and M. Stacey, *Carbohydrate Res.*, 1970, **12**, 171.

dilute alkali,[114] and this has been explained as being due to the labile character of the cross-linking glycyl peptides. Model studies[114] showed that glycylglycylglycine was extensively hydrolysed in molar sodium hydroxide, 1.4 mole of glycine being produced per mole of tripeptide in 60 h. Only 0.14 mole of alanine was formed from a mole of alanylalanylalanine over this time. Steric and probably inductive effects could explain these differences. The authors state that apart from one example, to their knowledge no relevant studies on the action of dilute alkali on the primary structure of proteins have been carried out – so redundant acid-hydrolysis experts should take note!

The alkali-resistant 1-L-β-aspartamido–1,2-dideoxy-β-D-glucose linkage, which has widespread occurrence in animal glycoprotein, has also been found[115] in soybean haemagglutinin. Pronase and tryptic digests of porcine[116] and whale[117] pancreatic ribonuclease give rise to glycopeptides possessing a carbohydrate linkage onto the asparagine amido-nitrogen. Preliminary studies[118] on a glycopeptide from *Metridium dianthus* collagen show that the peptide contains 8% of 3- and 4-hydroxyproline and 27% glycine, yet alkaline borohydride reduction suggests that the main linkage probably involves an asparaginyl residue. Chemical and enzymatic methods have established[119] the sequence Met-Asx(Sugar)-Gly-Thr-Glu-Gly-Pro-Asn-Phe for a glycopeptide from a peptic digest of the membrane protein, bovine visual pigment 500. The stability of the peptide to alkali suggests an *N*-aspartylglycosylamino linkage. In a phytotoxic glycopeptide isolated[120] from *Corynebacterium sepedonicum* the ratio of the amino-acid residues has been found to be Asp_1, Thr_1, Ser_1, Gly_1, Lys_1, Ala_1, Met_1, Glu_2, and His_1. Preliminary work has been reported on glycopeptides obtained from the yeast form of *Cladosporium werneckii*,[121] from human erythrocyte membranes,[122] from human platelets,[123] and from rat brain glycoproteins.[124] The latter peptides contain ester sulphate and appear to be distinct in structure from any known class of sulphated acid mucopolysaccharides. A group of glycopeptides of M.W. $\approx$ 5,000 make up a major fraction[125] of certain biological membranes. Approximately 30% of the total amino-acid residues of orosomucoid have been located[126] in seven lysine-containing and three arginine-containing peptides.

[114] A. R. Archibald, J. Baddiley, and J. Goundry, *Biochem. J.*, 1970, **116**, 313.
[115] H. Lis, N. Sharon, and E. Katchalski, *Biochim. Biophys. Acta*, 1969, **192**, 364.
[116] T. Tsuro, S. Yamashita, T. Terao, and T. Ukita, *Biochim. Biophys. Acta*, 1970, **200**, 544.
[117] T. Tsuro, K. Sudo, T. Terao, and T. Ukita, *Biochim. Biophys. Acta*, 1970, **200**, 560.
[118] R. L. Katzman and R. N. Jeanloz, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 628.
[119] J. Heller and M. A. Lawrence, *Biochemistry*, 1970, **9**, 864.
[120] G. A. Strobel, *J. Biol. Chem.*, 1970, **245**, 32.
[121] K. O. Lloyd, *Biochemistry*, 1970, **9**, 3446.
[122] R. Kornfeld and S. Kornfeld, *J. Biol. Chem.*, 1970, **245**, 2536.
[123] D. S. Pepper and G. A. Jamieson, *Biochemistry*, 1970, **9**, 3706.
[124] R. K. Margolis and R. U. Margolis, *Biochemistry*, 1970, **9**, 4389.
[125] M. T. Laico, E. I. Ruoslahti, D. S. Papermaster, and W. J. Dreyer, *Proc. Nat. Acad. Sci. U.S.A.*, 1970, **67**, 120.
[126] R. Bourrillon, D. Meyer, and A. Parnadeau, *Biochem. Biophys. Acta*, 1970, **200**, 378.

Almost all the structure elucidation work on glycopeptides reviewed in this Report over the last three years has been carried out by conventional degradation and analytical techniques. It was inevitable, therefore, that sooner or later more recent methods of structure determination would be applied to this field. Mass spectrometry using *N*-acetyl- and *N*-permethylated-derivatives of peptide subunits has now been used [127] to identify a peptide fragment (48) from mycobacterial cell walls. The peptide (48)



(48)

shows a very characteristic peak at *m/e* 484 ($M-199$), which has been interpreted as being due to the loss of $\cdot(CH_2)_3CH\cdot CON(Me)$ from the

diaminopimelic acid moiety. This confirmed that the glutamine was linked to the α-position of the diaminopimelic acid. Comparison with synthetic isomers established that the γ-carboxy-group of glutamic acid is linked to diaminopimelic acid.

Biosynthetic Aspects. It is now well established that the incorporation of amino-acids into the inter-peptide bridges of the peptidoglycans of bacterial cell walls is a tRNA-dependent process, although the study of the exact mechanism which determines the sequence of amino-acids is still a field of active interest. The use of radioactive labelling techniques has shown [128] that in the peptidoglycan of *Lactobacillus viridescens* containing the interpeptide D-Ala-L-Ser-L-Ala-L-Lys there is a surprising lack of specificity in the order of amino-acid incorporation. In the initial and later stages either L-alanine or L-serine can be added to the precursor uridine nucleotide in an enzyme-catalysed reaction requiring amino-acyl tRNA. The enzyme which catalyses the transfer of L-alanine from L-Ala-tRNA to UDP-*N*-acetylmuramyl-L-Ala-D-Glu-L-Lys-D-Ala-D-Ala in cell walls has been purified [129] to homogeneity. Aspects of the kinetics and solubilisation of phospho-*N*-acetylmuramyl-pentapeptide translocase, the enzyme involved in the initial stages of peptidoglycan synthesis, have also been reported.[130]

The recent recognition in *E. coli* of a DD-carboxypeptidase capable of hydrolysing the *C*-terminal D-Ala-D-Ala moiety of the nucleotide

[127] J. W.-Falszpan, B. C. Das, I. Azuma, A. Adam, J. F. Petit, and E. Lederer, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 57.
[128] R. Plapp and J. L. Strominger, *J. Biol. Chem.*, 1970, **245**, 3667.
[129] R. Plapp and J. L. Strominger, *J. Biol. Chem.*, 1970, **245**, 3675.
[130] M. G. Heydanek, W. G. Struve, and F. C. Neuhaus, *Biochemistry*, 1969, **8**, 1214; M. G. Heydanek and F. C. Neuhaus, *ibid.*, p. 1474.

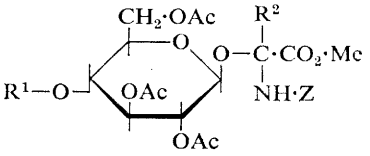
precursor uridine-5′-pyrophosphoryl-*N*-acetylmuramyl-L-Ala-γ-D-Glu-(L)-*meso*-Dpm(L)-D-Ala-D-Ala, and which exerts an endopeptidase action [131] on peptide dimers, has been followed up with the isolation [132] of a soluble DD-carboxypeptidase from *Streptomyces albus G*. This enzyme hydrolyses [133] the *C*-terminal D-Ala-D-Ala, as well as *C*-terminal D-Ala-Gly, which makes it identical in properties with that isolated from *E. coli*. However, in their interaction with penicillin the DD-carboxypeptidases from the two sources show striking differences.[134] Penicillins competitively inhibit *E. coli* DD-carboxypeptidase and this has been explained in terms of the analogous conformations of acyl-D-Ala-D-Ala and the penicillin which is accepted by the transpeptidase. However, the enzyme from *S. albus G.* does not recognise the penicillins as structural analogues of the terminal dipeptide moiety. It is therefore possible that the analogy between the *C*-terminal dipeptide and penicillin is not universal among bacteria. The mode of action of penicillin has also been studied [135] in mureins of *Proteus mirabilis*. In the natural and in a penicillin-induced unstable form of the murein, no differences have been found in the extent of the diaminopimelic acid residues involved in cross-linking.

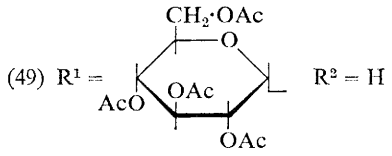
Biosynthetic studies on the synthesis of bacterial cell walls and the mechanism of inhibition of these biosynthetic reactions by penicillin have indicated that the inactivation results from the opening of the β-lactam ring and the penicilloylation of a specific group in the enzyme. Penicillin G has in fact been shown [136] to bind to the particulate enzyme preparation of *Bacillus subtilis* and the binding is reversed by neutral hydroxylamine, ethanethiol, or hydrogen peroxide. It therefore appears that the penicillin reacts with the enzyme to form a thiol ester between penicilloic acid and a thiol group in the enzyme. [^{14}C]Penicillin G also binds in a similar way [137] to the D-alanine carboxypeptidase present in *B. subtilis*.

B. Amino-acid–Carbohydrate Linkages.—Recent interest in the nature of sugar–amino-acid linkages in glycopeptides and glycoproteins has stimulated intensive research on simple models containing these linkages. Glycosyl and glucuronic esters of a number of amino-acids have been synthesised.[138] The best methods found for linking the carboxy-group of a suitably protected amino-acid to the hydroxy-group (at C-1) of the carbohydrate used dicyclohexylcarbodi-imide and active ester coupling, with

[131] D. Bogdanovsky, E. Bricas, and P. Dezéleé, *Compt. rend.*, 1969, **269**, *D*, 390.
[132] J.-M. Ghuysen, M. Leyh-Bouille, R. Bonaly, M. Nieto, H. R. Perkins, K. H. Schleifer, and O. Kandler, *Biochemistry*, 1970, **9**, 2955.
[133] M. Leyh-Bouille, J.-M. Ghuysen, R. Bonaly, M. Nieto, H. R. Perkins, K. H. Schleifer, and O. Kandler, *Biochemistry*, 1970, **9**, 2961.
[134] M. Leyh-Bouille, J.-M. Ghuysen, M. Nieto, H. R. Perkins, K. H. Schleifer, and O. Kandler, *Biochemistry*, 1970, **9**, 2971.
[135] W. Katz and H. H. Martin, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 744.
[136] P. J. Lawrence and J. L. Strominger, *J. Biol. Chem.*, 1970, **245**, 3653.
[137] P. J. Lawrence and J. L. Strominger, *J. Biol. Chem.*, 1970, **245**, 3660.
[138] D. Keglević, A. Kornhauser, G. Roglic, and T. Kovač, *Tetrahedron Letters*, 1970, 2983.

imidazole added as catalyst. A comparison [139] of the carbodi-imide method with the silver salt coupling method in the synthesis of acetylated 1-*O*-[*N*-acyl-D(and -L)-amino-acyl]-β-D-glucopyranoses shows that in the former method *N*-acetyl-amino-acids are racemised, but if benzyloxycarbonyl or phthaloyl groups are used for *N*-protection no racemisation occurs. In the silver salt method there is 80—85% retention of configuration irrespective of the protecting group. A study [140] on the comparative resistance to hydrolysis of a carbohydrate–carbohydrate *O*-glycoside bond and the bond linking a carbohydrate to serine or threonine has been carried out using the compounds (49) and (50). At pH 11 the two bonds showed





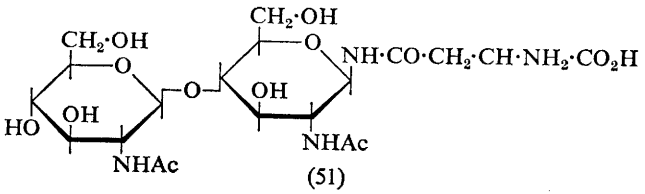
(50) R^1 = Ac, R^2 = Me

similar stability, but in acid the bond to the amino-acids was more stable. Kinetic studies,[141] using the 2,4-dinitrophenyl group for spectrophotometric monitoring, have been carried out on 3-*O*-(2,3,4,6-tetra-*O*-benzyl-α-D-glucopyranosyl)-*N*-(2,4-dinitrophenyl)-L-serine and -threonine methyl esters. Cleavage of the carbohydrate–hydroxy-amino-acid linkage took place readily in base, which is in contrast to bonds involving 3-*O*-glycosides of 4-hydroxyproline.

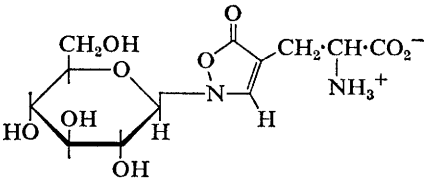
Suitably protected 3-*O*-α-D-glucopyranosyl-L-serine has been synthesised [142] from *N*-benzyloxycarbonyl-L-serine benzyl ester using the silver salt method. The compound (51) has been synthesised [143] as a model for work on the identification of (2-acetamido-2-deoxy-D-glucosyl)-(2-acetamido-2-deoxy-D-glucosyl) units in alkali-stable proteins.

A second u.v.-sensitive amino-acid from seedlings of *Pisum sativum* has been characterised [144] as (52).

[139] A. Kornhauser and D. Keglević, *Carbohydrate Res.*, 1970, **13**, 433.
[140] N. K. Kochetkov, E. M. Klimov, and V. A. Derevitskaya, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1969, 2779; N. K. Kochetkov, V. A. Derevitskaya, L. M. Likhoshertov, V. M. Kalinevich, and O. S. Novikova, *ibid.*, p. 2509.
[141] J. R. Vercellotti, N. Nienaber, and C. J. Chang, *Carbohydrate Res.*, 1970, **13**, 63.
[142] K. Kum, *Carbohydrate Res.*, 1969, **11**, 269.
[143] R. W. Jeanloz and M. Spinola, *J. Biol. Chem.*, 1970, **245**, 4158.
[144] F. Lambein and R. Van Parijs, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 557.

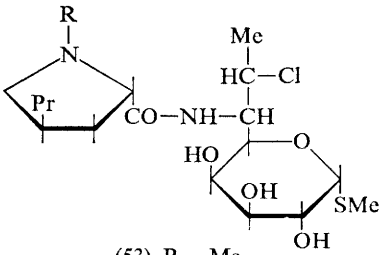


(51)



(52)

C. Lincomycin.—Oxygen and water in the presence of a platinum catalyst bring about dealkylation of the tertiary amino-group in lincomycin and its analogues. For example, (53, R = Me) gives the 1′-desmethyl-lincomycin



(53) R = Me,
(54) R = H

analogue (54, R = H) in 52% yield.[145] Other analogues give a lower yield on demethylation. Several procedures for preparing halogen analogues [146] of lincomycin have been described and many of these compounds have been found to be potent antibacterial agents. Syntheses [147] of the carbohydrate portion of the antibiotic complete the requirements for the total synthesis of lincomycin.

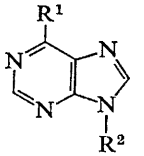
5 Peptides and Amino-acids Linked to Nucleosides and Nucleotides

A. Peptidyl- and (Amino-acyl)-tRNA.—The *o*-nitrophenylsulphenyl *N*-protecting group, in combination with the dicyclohexylcarbodi-imide–

[145] R. D. Birkenmeyer and L. A. Dolak, *Tetrahedron Letters*, 1970, 5049.
[146] R. D. Birkenmeyer and F. Kagan, *J. Medicin. Chem.*, 1970, **13**, 616.
[147] B. J. Magerlein, *Tetrahedron Letters*, 1970, 33; G. B. Haworth, W. A. Szarek, and J. K. N. Jones, *J. Chem. Soc.* (*C*), 1970, 2218.

N-hydroxysuccinimide coupling method, has been used to synthesise [148] a series of oligopeptidyl-tRNA's, *e.g.* Val-Gly-Phe-tRNA. The method also lends itself to the preparation of peptidyl-tRNA containing hydroxy-amino-acids.[149] In both of these reports the *N*-protecting group was removed using 0.2M sodium thiosulphate at 30 °C for 2 h. Various amino-acyl-tRNA's have initiated [150] the polymerisation of different *N*-carboxy-amino-acid anhydrides in alkaline dimethyl sulphoxide–water mixtures, to yield peptidylamino-acyl-tRNA's. Nearly quantitative yields of (DL-Ala)$_n$-Val-tRNA, (Sar)$_n$-Val-tRNA, [L-Asp(OBzl)]$_n$-Val-tRNA, [L-Glu(OBzl)]$_n$-Val-tRNA, and (L-Phe)$_n$-Val-tRNA were obtained by this method. The lengths of the peptide moieties were heterogeneous in all these derivatives and the ester linkage between the peptidylamino-acyl unit and the polynucleotide appears to be more stable than in the parent amino-acyl-tRNA itself. In a search [151] for new methods of preparing heavy-atom derivatives for use in *X*-ray crystallographic studies, the *N*-hydroxysuccinimide ester of chloromercuriacetic acid has been used to form the chloromercuriacetylated derivative of L-Val-tRNA.

B. Nucleopeptides.—The chemical quest for potential biologically active compounds has initiated syntheses of a number of amino-acid and peptide derivatives based on the purine and pyrimidine systems. 4-(*N*^{9}-Purinyl)-α-aminobutyric acids based on adenine, hypoxanthine, guanine, and xanthine have been synthesised [152] from the chloro-analogues of the purine bases with t-butyl α-benzyloxycarbonylamido-γ-bromobutyrate. Willardiine (55) has also been synthesised using the same principles.

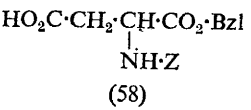


(55) $R^1 = NH_2$, $R^2 = \cdot CH_2 \cdot CH(NH_2) \cdot CO_2H$

(56) $R^1 = \cdot SCO \cdot R$, $R^2 = H$

(57) $R^1 = SCH_3$, $R^2 = COR$

where R is derived from (58)



(58)

When 6-methylthiopurine and purine-6-thiol react [153] with the glutamate derivative (58), using the dicyclohexylcarbodi-imide or the mixed anhydride

[148] Y. Lapidot, D. Elat, S. Rappoport, and N. de-Groot, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 559.

[149] M. Rubinstein, N. de-Groot, and Y. Lapidot, *Biochim. Biophys. Acta*, 1970, **209**, 183.

[150] S. A. Yankofsky, S. Yankofsky, E. Katchalski, and U. Z. Littauer, *Biochim. Biophys. Acta*, 1970, **199**, 56.

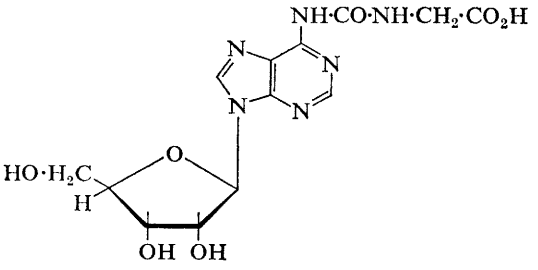
[151] G. Fölsch, *Acta Chem. Scand.*, 1970, **24**, 1115.

[152] A. J. H. Nollet, C. M. Huting, and U. K. Pandit, *Tetrahedron*, 1969, **25**, 5971, 5983; A. J. H. Nollet and U. K. Pandit, *ibid.*, p. 5989.

[153] A. Y. Veinberg, I. N. Gracheva, and G. I. Samokhalov, *Zhur. obshchei Khim.*, 1970, **40**, 484.

coupling methods, the products obtained, (56) and (57) respectively, are quite reactive and very similar to activated forms of amino-acids used in peptide coupling reactions or as enzyme inhibitors. A rearrangement of N^6-(α-amino-acyl)adenines to *N*-(purin-6-yl)amino-acids has been applied to adenosine analogues,[154] and syntheses of β-(2-aminopurin-9-yl)-α-alanine,[155] (purin-9-yl)aminobutyric acids,[156] and α-*N*-(pyrimidin-4-yl)-amino-acids [157] have been reported.

The glycine-substituted adenosine (59) has been isolated [158] from enzyme digests of unfractionated yeast tRNA. The structure was confirmed by comparison of the n.m.r. and mass spectra of the natural material



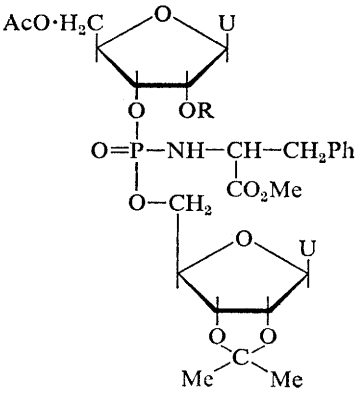
(59)

with those of a synthetic sample. In a study of the biological importance of the sulphoxides of *S*-adenosyl- and *S*-ribosyl-L-homocysteine it has been found [159] that in a synthesis of the sulphoxides using hydrogen peroxide, the sulphur atoms of the cysteine derivatives are not as readily oxidised as in L-methionine. A longer reaction time and more concentrated hydrogen peroxide had to be used for complete reaction.

Studies on the properties and stabilities of nucleotide–peptide bonds continue, using suitably synthesised model compounds. Thymidylyl-(3′ → 5′)adenosine (P_m-N) amides have been prepared [160] from thymidylyl-(3′ → 5′) adenosine and protected amino-acids, using the mixed anhydride method. The hydrolysis of phosphoramide bonds in such systems has also been investigated. Compound (60) is stable [161] at pH 10.5 for 1 h but undergoes cleavage over 20 h, which can be explained by the involvement

[154] G. B. Chheda and R. H. Hall, *J. Org. Chem.*, 1969, **34**, 3492, 3498.
[155] M. Lidaks, Y. Y. Schluke, S. Y. Poritere, and Y. P. Schvachkin, *Khim. geterotsikl. Soedinenii*, 1970, 529.
[156] Y. Y. Schluke, B. Zarina, M. Lidaks, and Y. P. Shvachkin, *Khim. geterotsikl. Soedinenii*, 1970, 534.
[157] R. Paegle, M. Plata, and M. Lidaks, *Khim. geterotsikl. Soedinenii*, 1969, 558.
[158] M. Schweizer, K. McGrath, and L. Baczynskyj, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 1046.
[159] J. A. Duerre, L. Salisbury, and C. H. Miller, *Analyt. Biochem.*, 1970, **35**, 505.
[160] N. I. Sokolova, G. I. Gurova, Z. A. Shabarova, and M. A. Prokof'ev, *Vestnik Moskov. Univ.*, 1969, **24**, 104.
[161] O. E. Vorob'ev, Z. A. Shabarova, and M. A. Prokof'ev, *Doklady Akad. Nauk S.S.S.R.*, 1970, **190**, 842.

of the deacetylated form (61). Nucleotide (P → N) peptides have also been synthesised on a polymer carrier.[162] Sequential formation of the peptides was carried out using t-butoxycarbonyl-amino-acids, and the resin-linked peptide was then linked onto the phosphonic acid–nucleotide moiety.

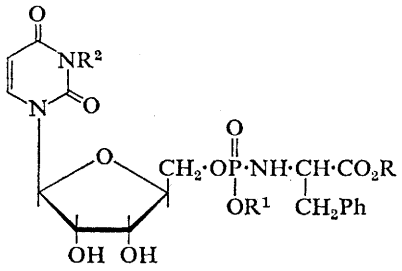


(60) R = Ac
(61) R = H
U = uridine

Model compounds relating to the terminal sequence of peptidyl-amino-acyl-tRNA have been synthesised[163] by coupling $N^{\alpha}N^{\epsilon}$-bisbenzyloxycarbonyl-L-lysine *N*-hydroxysuccinimide ester with an appropriate nucleotide intermediate to give cytidylyl(3′ → 5′)-2′(3′)-*O*-[$N^{\alpha}N^{\epsilon}$-bisZ-L-Lys-Gly]-adenosine.

The influence of substituents both in the heterocyclic ring and on the phosphate bond have been investigated in relation to their effect[164] on the reactivity of the phospho-amide bond. In acid, alkaline, and enzymic hydrolysis it has been established that a methyl group in position 3 of the uracil ring does not exert much influence on the phospho-amide bond, but methylation of the phosphate bond (62, R^1 = Me) stabilises the phospho-amide bond although it makes the 5-phospho-ester bond more labile. Substitution of bromine or hydroxy-groups at position 5 (62, R^2 = Br or OH) also changes the nature of hydrolysis of the phospho-amide bond in acid medium and makes it labile in weakly alkaline solution. The greater lability of the P—N bond in the guanylyl-5-*N*-amino-acids (63) in acid

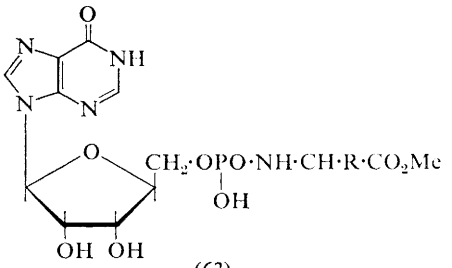
[162] V. D. Smirnov, G. A. Khangulov, Z. A. Shabarova, and M. A. Prokof'ev, *Vestnik Moskov. Univ.*, 1969, **24**, 118.
[163] S. Chládek and J. Zemlička, *Coll. Czech. Chem. Comm.*, 1970, **35**, 89.
[164] N. G. Shinskii, N. N. Preobrazhenskaya, R. K. Ledneva, Z. A. Shabarova, and M. A. Prokof'ev, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1969, 2307; N. G. Shinskii, N. N. Preobrazhenskaya, Z. A. Shabarova, and M. A. Prokof'ev, *Zhur. obshchei Khim.*, 1970, **40**, 1122.



(62)

media has been interpreted [165] as being due to hydrogen-bonding involving the 2-amino-group of guanine.

The synthesis of dipeptidyl amino-sugar nucleosides (64) related to the antibiotic gougerotin has been carried out [166] by condensing 1-(3-amino-3-deoxy-β-D-glucopyranosyl)cytosine with the azide of *N*-t-butoxycarbonyl-sarcosyl-D-serine. In contrast to the dicyclohexylcarbodi-imide and active ester methods, the azide method gave no *N*-acylation of the cytosine amino-group.

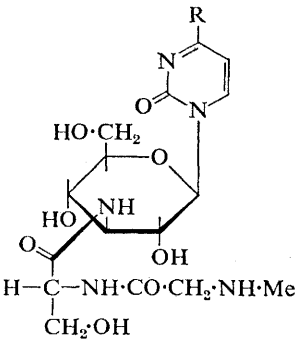


(63)

Structural analogues [167] of puromycin, *e.g.* amino-acyl derivatives of 9-(3′-amino-3′-deoxy-β-D-arabinofuranosyl)adenine, have been prepared in an effort to exploit the biological properties of this class of antibiotics. *N*-Formylmethionylpuromycin has been found [168] as a biosynthetic product in honey bees injected with a solution of puromycin and labelled methionine.

Polyoxins D, E, and F have been transformed [169] into polyoxins L, M, and K respectively, using a very mild bisulphite-catalysed decarboxylation reaction of 5-carboxyuracil derivatives.

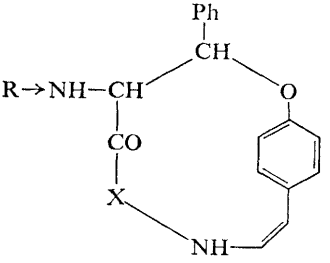
[165] N. I. Sokolova, L. G. Gatinskaya, Z. A. Shabarova, and M. A. Prokof'ev, *Zhur. Vsesoyuz. Khim. obshch. im. D.I. Mendelleeva*, 1969, **14**, 583.
[166] F. W. Lichtenthaler, G. Trummlitz, and P. Emig, *Tetrahedron Letters*, 1970, 2061.
[167] L. V. Fisher, W. W. Lee, and L. Goodman, *J. Medicin. Chem.*, 1970, **13**, 775.
[168] G. K.-Kreil and G. Kreil, *Monatsh.*, 1970, **101**, 629.
[169] K. Isono, S. Suzuki, M. Tanaka, T. Nanbata, and K. Shibuya, *Tetrahedron Letters*, 1970, 425.



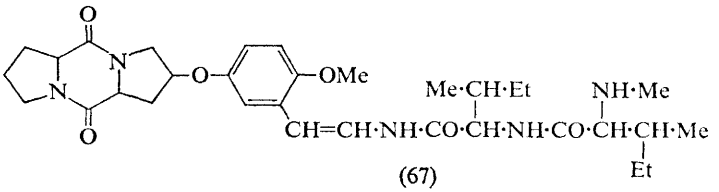
(64) R = OH or NH_2

6 Peptide Alkaloids

Mass spectrometry coupled with chemical methods have again played a major role in the structural elucidation of two peptide alkaloids containing the *p*-alkoxystyrylamino-residue. Canthuimine, a novel alkaloid from *Canthium euryoides*, has been shown [170] to be (65), while aralionine B,[171] a minor alkaloid from the leaves of *Araliorhamnus vaginata*, has the structure



(65) R = Me_2Phe, X = Pro
(66) R = MePhe, X = Ile



(67)

[170] G. Boulvin, R. Ottinger, M. Pais, and G. Chiurdoglu, *Bull. Soc. chim. belges*, 1969, **78**, 583.
[171] R. Tschesche, E. Frohberg, and H. W. Fehlhaber, *Chem. Ber.*, 1970, **103**, 2501.

(66). Physical methods have also featured in the elucidation[172] of the structure of zizyphinine (67), a peptide alkaloid from *Zizyphus oenoplia.*

Biosynthetic studies using *Claviceps purpurea* have shown[173] that DL-tryptophan is specifically incorporated into the lysergic acid moiety of the ergotoxine alkaloids, while L-proline appears to be a precursor of the proline of the ergotoxine peptide chain.

7 Peptides and Amino-acids Conjugated to Lipids

The peptidolipid $CH_3(CH_2)_nCO$-D-Phe-L-Ile-L-Phe-L-Ala-OMe and its acid analogue behave[174] differently in chloroform solutions. O.r.d. and i.r. studies show that the acid has a preferred conformation stabilised by hydrogen-bonding between two peptide chains associated as carboxylic dimers.

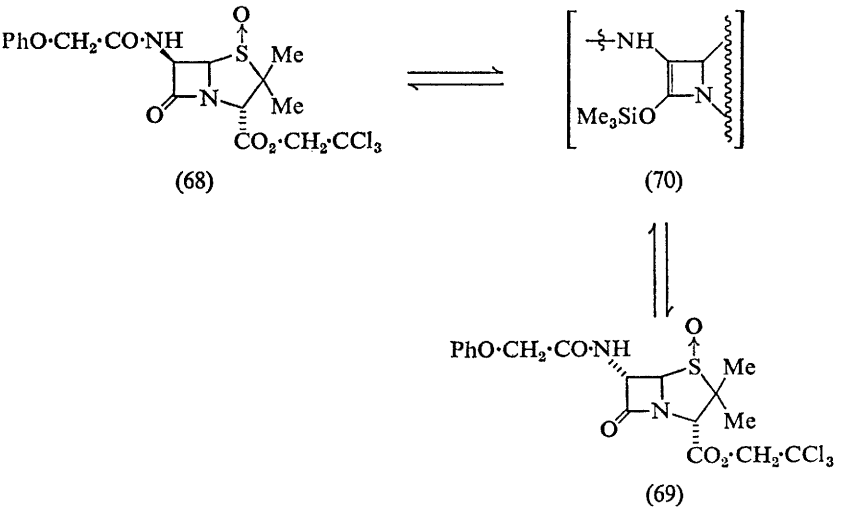
A new synthesis of α-(βγ-dipalmitoyl)glycerylaminoethylphospho-(P → N)glycine has been reported.[175]

8 Penicillins and Cephalosporins

This year again, no attempt has been made to review all the papers published in this field; the emphasis is more on the work relating to the chemistry of the β-lactam moiety.

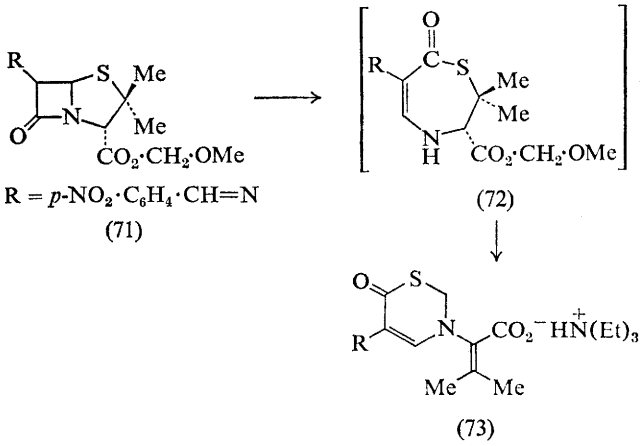
In last year's Report examples were cited of epimerisation at the C-6 position in penicillins and the C-7 position in cephalosporins. The procedures used, however, failed to epimerise penicillins containing a secondary amide side-chain. Epimerisation in such structures has been shown[176] to occur in the penicillin sulphoxide (68) when it is treated with *NO*-bis-(trimethylsilyl)acetamide. Equilibration, giving a 4 : 1 ratio of (69) to (68), requires several days at room temperature and occurs in a variety of solvents. Epimerisation using this method appears to be quite general in the penicillin series, possibly involving an intermediate such as (70). Surprisingly, the 7-position in the cephalosporins and their sulphoxides cannot be epimerised in this way. A detailed investigation of the effect of temperature, the substitution pattern at C-5 and C-6, and the nature of the base, has been made[177] to rationalise the conditions for proton abstraction at C-6 by carbanion or β-elimination pathways. In a competition between β-elimination and carbanion formation at C-6, the latter process is facilitated by the substituents, bromine, trimethylammonium, and phthaloyl at C-6, and also in the presence of strong bases such as hydroxide, sodium hydride, sodamide, and t-butoxide. The β-elimination is favoured when the

[172] H. Pailer, E. Haslinger, and E. Zbiral, *Monatsh.*, 1969, **100**, 1608.
[173] D. Groeger and D. Erge, *Z. Naturforsch.*, 1970, **25b**, 196.
[174] G. Laneelle, *F.E.B.S. Letters*, 1969, **4**, 210.
[175] M. K. Petrova, S. D. Bakalo, V. I. Shvets, and N. A. Preobrazhenskii, *Zhur. org. Khim.*, 1969, **5**, 1883.
[176] G. E. Gutowski, *Tetrahedron Letters*, 1970, 1779.
[177] S. Wolfe, W. S. Lee, and R. Misra, *Chem. Comm.*, 1970, 1067.



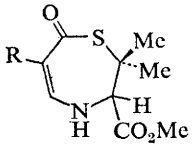
base is triethylamine. A β-elimination process has also been suggested [178] for the stereospecific conversion of penicillins G and V into their desthio-derivatives using Raney nickel. As a result of attempts [179] to bring about epimerisation at C-6, the derivative (71) has been found to rearrange in triethylamine–chloroform to give (73), probably *via* (72).

The β-lactam ring in methyl penicillinate or its 6α-chloro-analogue opens to give (74) in the presence of antimony pentachloride.[180] When



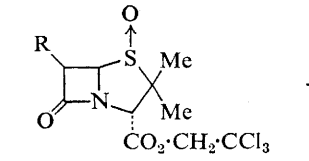
[178] S. Wolfe and S. K. Hasan, *Chem. Comm.*, 1970, 833.
[179] J. R. Jackson and R. J. Stoodley, *Chem. Comm.*, 1970, 14.
[180] J. P. Clayton, R. Southgate, B. G. Ramsay, and R. J. Stoodley, *J. Chem. Soc.* (*C*), 1970, 2089.

penicillin G is heated with trifluoroacetic acid the β-lactam ring is completely degraded[181] and the main product is 5,5-dimethyl-Δ^2-thiazoline-carboxylic acid. It is probable that the β-lactam fragment exists in the product mixture as a mixed anhydride,

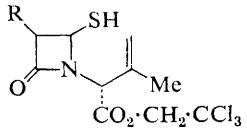


(74)

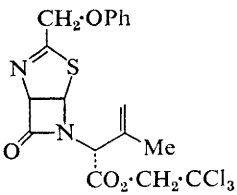
$PhCH_2 \cdot CO \cdot NH \cdot CH_2 \cdot CO_2 \cdot CO \cdot CF_3$. The penicillin sulphoxide (75) undergoes a novel rearrangement[182] to the structure (77) when it is treated with trimethyl phosphite in refluxing benzene. An intramolecular nucleophilic attack of the thiol on the amide carbonyl of intermediate (76) may



(75, R = $PhO \cdot CH_2 \cdot CO \cdot NH \cdot$)



(76)



(77)

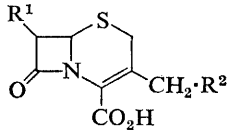
be involved in the latter stages. New studies have been reported in the search for methods to transform penicillins into cephalosporins using ring-expansion processes.[183]

The ^{13}C n.m.r. spectra of several penicillins and related sulphoxides have been studied[184] and the chemical shift assignments for γ-substituted carbons are explained in terms of a sizeable steric effect induced by the $S \rightarrow O$ bond.

[181] M. R. Bell, J. A. Carlson, and R. Oesterlin, *J. Amer. Chem. Soc.*, 1970, **92**, 2177.
[182] R. D. G. Cooper and F. L. Jose, *J. Amer. Chem. Soc.*, 1970, **92**, 2575.
[183] R. D. G. Cooper, *J. Amer. Chem. Soc.*, 1970, **92**, 5010; D. O. Spry, *ibid.*, p. 5006.
[184] P. V. Demarco, R. A. Archer, R. D. G. Cooper, and L. R. F. Johnson, *Chem. Comm.*, 1970, 1291.

Cyclohepta-amylose specifically catalyses [185] the hydrolysis of a series of penicillins, and the kinetics confirm the intermediacy of a penicilloyl-cyclo-amylose. The specificity for certain penicillin derivatives indicates that the system is an interesting model of penicillinase (β-lactamase). The synthesis of *p*-*O*-(tetra-*O*-acetyl-α-D-glucopyranosyl)penicillin G has been reported,[186] and transformations of 6-phenylacetamido- and 6-tritylamino-penicillanyl-toluene-*p*-sulphonates have been investigated.[187] Mass spectral and n.m.r. investigations on methyl derivatives of 6-aminopenicillanic acid have also been reported.[188]

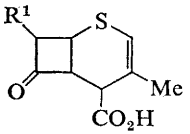
Many attempts have been made recently to clarify the mechanism by which penicillins and cephalosporins inhibit bacterial wall synthesis. In this context, a review [189] on the relationship between the chemistry and biological activity of the cephalosporins and the penicillins has appeared. Structural comparisons using *X*-ray crystallography have also been made [190] between two cephalosporins having the structural type (78), namely



(78) R^1 =  $CH_2 \cdot CO \cdot NH \cdot$, R^2 = pyridinium

$R^1 = Ph \cdot CH(NH_2) \cdot CO \cdot NH$, R^2 = OAc

cephaloridine hydrochloride and cephaloglycine, and a biologically inactive Δ^2-cephalosporin (79). The most striking difference between the active Δ^3-cephalosporins and the inactive Δ^2-form is found in the geometry of the β-lactam ring. The bridgehead nitrogen atom in the former is definitely



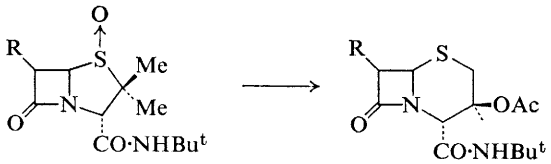
(79, $R^1 = PhOCH_2 \cdot CO \cdot NH \cdot$)

pyramidal, while in the Δ^2-form it is nearly planar. In addition, the β-lactam —N—CO— amide bond is observed to lengthen on going from the inactive to the active forms, which is consistent with an increase in the carbonyl-stretching frequency accompanying the decreased amide character.

[185] D. E. Tutt and M. A. Schwartz, *Chem. Comm.*, 1970, 113.
[186] T. L. Nagabhushan and C. G. Chin, *Canad. J. Chem.*, 1970, **48**, 3097.
[187] M. R. Bell, S. D. Clemans, and R. Oesterlin, *J. Medicin. Chem.*, 1970, **13**, 389.
[188] F. Moll and M. Hannig, *Arch. Pharm.*, 1970, **303**, 321.
[189] E. P. Abraham, *Top. Pharm. Sci.*, 1968, **1**, 1.
[190] R. M. Sweet and L. F. Dahl, *J. Amer. Chem. Soc.*, 1970, **92**, 5489.

This comparison between activity and the geometry of a reactive bond assumes, of course, that the conformation in the crystalline state remains unchanged in solution. There is recent n.m.r. evidence (last year's Report, p. 220) on penicillin sulphoxides to suggest that this might be an unjustified assumption.

A study of nuclear Overhauser effects and of the magnitude of long-range coupling indicates[191] that the major sulphoxide isomer formed on oxidation of desacetoxy-Δ^2- and -Δ^3-cephalosporin is the (*S*)-sulphoxide. The penicillin sulphoxide (80) rearranges in acetic anhydride[192] to (81).



(80, R = $PhCH_2CO\cdot NH$) (81)

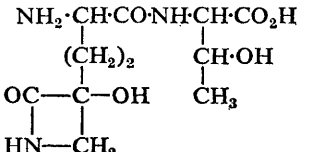
A heavy-atom derivative of the product of the latter with periodate has been subjected to *X*-ray crystallography.[193] The final link in a total synthesis of cephalosporanic acid derivatives has been made with the chemical conversion[194] of desacetylcephalothin lactone into desacetylcephalothin. Synthesis and biological properties of 3-acyloxymethyl-7(2-thienylacetamido)-3-cephem-4-carboxylic acid and derivatives have been published.[195] Cefazolin has been synthesised from 7-aminocephalosporanic acid and its *in vitro* and *in vivo* antimicrobial properties have been investigated.[196] The synthesis[197] of 7-acyl-3-methyl-2-cephem-4-carboxylic acid esters, and the transformation of a Δ^2-cephem to a Δ^3-cephem, have been reported.[198] Investigations[199] have been carried out on the influence of lipophilic character on the antibacterial activity of cephalosporins and penicillins.

9 Miscellaneous

After two decades of study, the structure of 'Wildfire Toxin', the cause of leaf-spot disease of tobacco plants, has been shown[200] to be the β-lactam-

[191] R. D. G. Cooper, P. V. Demarco, C. F. Murphy, and L. A. Spangle, *J. Chem. Soc.* (*C*), 1970, 340.
[192] D. H. R. Barton, F. Comer, D. G. T. Greig, G. Lucente, P. G. Sammes, and W. G. E. Underwood, *Chem. Comm.*, 1970, 1059.
[193] M. L. Smart and D. Rogers, *Chem. Comm.*, 1970, 1060.
[194] S. L. Neidelman, S. C. Pan, J. A. Last, and J. E. Dolfin, *J. Medicin. Chem.*, 1970, **13**, 386.
[195] S. Kukolja, *J. Medicin. Chem.*, 1970, **13**, 1114.
[196] M. Nishida, T. Matsubara, T. Murakawa, Y. Mine, Y. Yokota, S. Goto, and S. Kuwahara, *J. Antibiotics* (*Japan*), 1970, **23**, 137; K. Kariyone, H. Harada, M. Kurita, and T. Takano, *ibid.*, p. 131.
[197] C. F. Murphy and R. E. Koehler, *J. Org. Chem.*, 1970, **35**, 2429.
[198] G. V. Kaiser, R. D. G. Cooper, R. E. Koehler, C. F. Murphy, J. A. Webber, I. G. Wright,and E. M. Van Heyningen, *J. Org. Chem.*, 1970, **35**, 2430.
[199] G. L. Biagi, M. C. Guerra, A. M. Barbaro, and M. F. Gamba, *J. Medicin. Chem.*, 1970, **13**, 511.
[200] W. W. Stewart, *Nature*, 1971, **229**, 174.

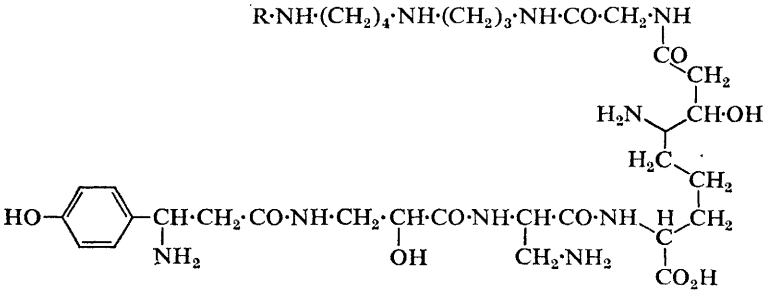
containing compound (82). This is quite different from the lactone previously proposed by Woolley for the same toxin some years ago. The



(82)

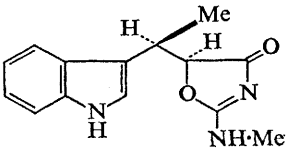
labile nature of the compound contributed greatly to the problems of structure elucidation, but mild hydrolysis gives threonine and an amino-acid, tabtoxinine. The β-lactam structure has been found only three times in Nature – in the penicillins, the cephalosporins, and the pachystermins, and was identified in wildfire toxin on the basis of an unusually low coupling constant (6—7 Hz) of the ε-proton of the tabtoxinine moiety.

Edeines A and B (see, also, last year's Report, p. 223) have each been shown [201] to be mixtures of an active peptide A_1 and B_1 respectively, and inactive isomers A_2 and B_2. The identification of peptide fragments from partial acid hydrolysates confirms the sequence (83) for edeine A_1 and (84)



(83) R = H
(84) R = C(=NH)NH$_2$

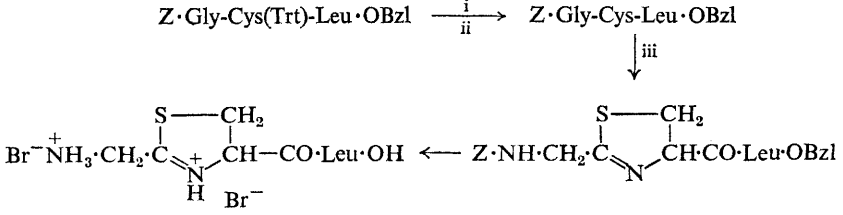
for edeine B_1. The absolute configuration of indolmycin has been designated [202] (5*S*, 6*R*)-indolmycin (85), based on a comparison of the product of hydrolysis of indolmycin (indolmycenic acid) with a synthetic sample of known configuration.



(85)

[201] T. P. Hettinger and L. C. Craig, *Biochemistry*, 1970, **9**, 1224.
[202] T. H. Chan and R. K. Hill, *J. Org. Chem.*, 1970, **35**, 3519.

Unprotected thiazoline peptides such as (86) have been synthesised [203] as outlined in Scheme 5 as part of a study on the synthesis of bacitracin. The



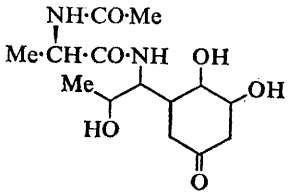
(86)

Reagents: i, $AgNO_3$–pyridine; ii, HCl; iii, hydrogen chloride–chloroform

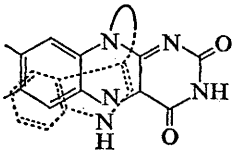
Scheme 5

peptides were synthesised either by the imino-ether coupling method or by the dehydration method. In the formation of the thiazoline ring using either of these two methods for cyclisation, racemisation [204] takes place, so it seems likely that any synthesis of bacitracin A will have to involve resolution at one stage.

The structure of the product derived from the mild basic hydrolysis of the *N*-acetyl derivative of the antibiotic actinobolin has been shown [205] by physical methods to be (87). Pepstatin has been identified [206] as 4-[4-(isovaleryl-L-valyl-L-valylamino)-3-hydroxy-6-methylheptanoyl-L-alanylamino]-3-hydroxy-6-methylheptanoic acid, using u.v. and n.m.r. spectroscopy and mass spectrometry.



(87)



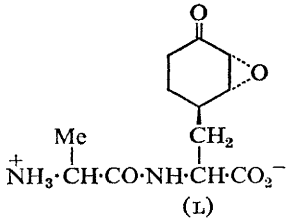
(88)

N.m.r. data [207] for flavin peptides where a 7,8-dimethyliso-alloxazine portion is linked *via* an amide linkage to alanine, tryptophan, tyrosine, or phenylalanine show that the flavin side-chain tends to fold back on the ring system. This also occurs in ω-carboxyalkylflavins. Specific upfield shift of the aromatic protons of flavinyl amino-acids making up the

[203] Y. Hirotsu, T. Shiba, and T. Kaneko, *Bull. Chem. Soc. Japan*, 1970, **43**, 1564.
[204] Y. Hirotsu, T. Shiba, and T. Kaneko, *Bull. Chem. Soc. Japan*, 1970, **43**, 1870.
[205] D. B. Nelson and M. E. Munk, *J. Org. Chem.*, 1970, **35**, 3832,
[206] H. Morishima, T. Takita, T. Aoyagi, T. Takeuchi, and H. Umezawa, *J. Antibiotics (Japan)*, 1970, **23**, 263.
[207] W. Föry, R. E. MacKenzie, F. Y. H. Wu, and D. B. McCormick, *Biochemistry*, 1970, **9**, 515.

flavinyl peptide shows that the aromatic and heteroaromatic portions of these compounds interact intramolecularly *via* vertical ring stacking, as shown in the model (88).

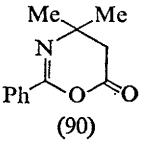
An antibiotic, bacilysin, isolated from *Bacillus subtilis*, has been shown to have [208] the structure (89). Prephenic acid appears to be a possible precursor of the *C*-terminal acid.



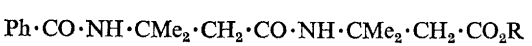
(L)

(89)

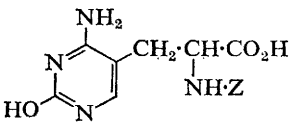
The aminolysis of the oxazinone (90) proceeds [209] in good yield to give (91). Peptides of this type, containing 3-amino-3-methylbutanoic acid, are difficult to prepare using conventional methods. As part of a study on unnatural amino-acids and their peptides, the derivative (92) can be



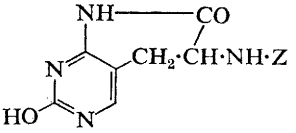
(90)



(91)



(92)



(93)

synthesised readily, but all attempts to prepare the deacylated product have failed.[210] On coupling the derivative (92) with glycine ethyl ester in the presence of dicyclohexylcarbodi-imide, only a low yield of dipeptide is obtained. The formation of lactam (93) was a major side-reaction.

There has been some evidence presented for the presence of *N*-hydroxy-peptide groups in cancer protein.[211]

[208] J. E. Walker and E. P. Abraham, *Biochem. J.*, 1970, **118**, 557, 563.
[209] C. N. C. Drey and J. Lowbridge, *Chem. Comm.*, 1970, 791.
[210] M. Shimada, K. Takeuchi, Y. Asai, S. Kutawa, and H. Watanabe, *Bull. Chem. Soc. Japan*, 1970, **43**, 2874.
[211] O. Neunhoeffer, *Z. Naturforsch.*, 1970, **25b**, 299.

5

Metal Derivatives of Amino-acids, Peptides, and Proteins

BY R. D. GILLARD AND S. H. LAURIE

1 Introduction*

This Report on the developments in the wide areas concerned with metal binding to amino-acids, peptides, and proteins covers work published in the two years, 1969 and 1970, since our last Report.[1] With the cessation of the publication of *Current Chemical Papers* at the end of 1969 we have made use of *Current Titles*[2] as our literature source, together with an independent survey of the major journals. Certain papers published towards the end of 1970 will, unfortunately, not be covered in this Report but will certainly be mentioned in our next review.

Few striking developments have occurred in the past two years, though there are pleasing signs that attention is now being given to the changes in reactivity, at least of amino-acids and peptides, on binding to metal ions. A number of valuable reviews have appeared with general titles: 'Inorganic Biochemistry,'[3] 'The Discriminating Behaviour of Metal Ions and Ligands with regard to their Biological Significance,'[4] and 'N.M.R. Studies on the Biochemistry of Biopolymers,'[5] which includes sections on metal derivatives of amino-acids, peptides, and proteins. Some more specific reviews, each devoted to a single metal, were 'Chromium Occurrence and Function in Biological Systems'[6] (a strangely neglected subject in the past), 'Biochemical Aspects of Molybdenum Co-ordination

[1] R. D. Gillard and S. H. Laurie in 'Amino-acids, Peptides, and Proteins,' ed. G. T. Young, (Specialist Periodical Reports), The Chemical Society, London, 1969, Vol. 1, p. 262.
[2] 'Current Titles,' Institute for Scientific Information, Philadelphia, U.S.A.
[3] U. Weser, *Naturwiss.*, 1969, **56**, 506.
[4] H. Sigel and D. B. McCormick, *Accounts Chem. Res.*, 1970, **3**, 201.
[5] J. J. M. Rowe, J. Hintin, and K. L. Rowe, *Chem. Rev.*, 1970, **70**, 1.
[6] W. Mertz, *Physiol. Rev.*, 1969, **49**, 163.

* The formulation of chelates of amino-acids commonly requires the amino-acid anion, *e.g.* $NH_2 \cdot CH_2 \cdot CO_2^-$, and papers in this field customarily abbreviate this as 'Gly', the amino-acid itself being 'Gly·H', *etc.* By the I.U.P.A.C.–I.U.B. Tentative Rules, 'Gly' is the abbreviation for the amino-acid, $NH_2 \cdot CH_2 \cdot CO_2H$, and the anion is formulated as 'Gly-$\bar{O}$'. In conformity with our practice throughout the volume, the I.U.P.A.C.–I.U.B. formulations are used here. In addition, the symbol α is used to denote an amino-acid anion in general, Hα then being the amino-acid itself.

Chemistry,'[7] 'Polynuclear Complexes of Iron and their Biological Significance,'[8] 'Structural Studies of Iron–Sulphur Proteins,'[9] and 'Catalase and Peroxidase Activity of Copper(II) Complexes.'[10] Absolute stereochemistries of octahedral complexes of transition metals have been surveyed.[11] Several reviews of subjects in haem protein studies include 'The Spectra of Ferric Haems and Haemoproteins,'[12] 'The Haemoglobin Molecule,'[13] and 'Haemoglobin.'[14]

Finally, we must state that the section headings we have chosen are arbitrary and inevitably some papers overlap more than one area; in such cases we have been obliged to make a choice, and we hope that authors and readers will find our choice appropriate.

2 Amino-acids

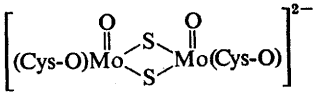
A. Binding Sites.—In addition to a plethora of sporadic studies, a good deal of work directed to more specific areas has appeared. We shall consider these directed studies first.

Molybdenum Complexes. Melby[15] has studied the molybdenum(v)–histidine system, which was already known to form a diamagnetic 1 : 1 complex at pH 4.6. Melby isolated, from solutions of pH 7—8, a complex of composition Mo_2O_4(L-His-O)$_2$, and proposed, from spectroscopic data and a preliminary *X*-ray study, possible structures with a central μ,μ'-(dioxo)dimolybdenum(v) unit, each molybdenum atom then being bonded by a further oxygen atom and a terdentate histidinate ligand. The proposed structures are very similar to that of the analogous L-cysteine complex.[16]

Complexes of molybdenum(v) and molybdenum(vi) with sulphur-containing amino-acids have been of particular interest because of their possible relevance to molybdenum-containing enzymes, such as xanthine oxidases, nitrogenases, and nitrate reductases. A long series of diamagnetic complexes have been described[17a] by Kay and Mitchell. The following are among them: $[Mo^{VI}_2O_4(OH)_3L]$, where L is L-alanine or β-alanine; $[Mo^{VI}O_2L_2]$, $[Mo^{V}_2O_3L_4]$, $[Mo^{V}_2O_2S_2L_2]$, where L is L-cysteine methyl ester; $Na_2[Mo^{V}_2O_2S_2L_2],3H_2O$, where L = L-cysteine. With hydrogen sulphide the oxo-bridges of some complexes are replaced by thio-bridges, Mo—S—Mo. Spivack and Dori, in a similar study,[17b] remind us of the comment[18] that the importance of Mo—S bonding in the enzymes, though

[7] E. Sinn and C. M. Harris, *Co-ordination Chem. Rev.*, 1969, **4**, 391.
[8] T. G. Spiro and P. Saltman, *Structure and Bonding*, 1969, **6**, 116.
[9] J. C. M. Tsibris and R. W. Woody, *Co-ordination Chem. Rev.*, 1970, **5**, 417.
[10] H. Sigel, *Angew. Chem. Internat. Edn.*, 1969, **8**, 167.
[11] R. D. Gillard and P. R. Mitchell, *Structure and Bonding*, 1970, **7**, 46.
[12] D. W. Smith and R. J. P. Williams, *Structure and Bonding*, 1970, **7**, 1.
[13] M. F. Perutz, *Proc. Roy. Soc.*, 1969, **B173**, 113.
[14] A. Antonini and M. Brunori, *Ann. Rev. Biochem.*, 1970, **39**, 977.
[15] L. R. Melby, *Inorg. Chem.*, 1969, **8**, 1539.
[16] Ref. 1, p. 264.
[17a] A. Kay and P. C. H. Mitchell, *J. Chem. Soc.* (*A*), 1970, 2421.
[17b] B. Spivack and Z. Dori, *Chem. Comm.*, 1970, 1716.
[18] J. T. Spence, *Co-ordination Chem. Rev.*, 1969, **4**, 475.

it may be due to a cysteine residue at the active site, could conceivably arise from sulphur-bridged species (like the Fe—S region in ferredoxin). They synthesized several sulphur-bridged complexes containing ethylenediaminetetra-acetic acid, cysteine, and histidine, including $Na_2[Mo_2O_2S_2(Cys\text{-}O)]$ which is formulated as containing the di-anion



They, too, made the thio-bridged complexes by reacting oxo-bridged ones with hydrogen sulphide.

Knox and Prout report [19] the full data from their *X*-ray study of the molybdenum(v)–cysteinate complex $Na_2[Mo_2O_4(Cys\text{-}O)_2],5H_2O$, briefly mentioned in our last Report.[16] The short Mo–Mo distance (2.569 Å) accounts for the diamagnetism. An interesting contribution is that [20] of Huang and Haight, who establish, largely by e.s.r. methods, the existence of an equilibrium between the diamagnetic, dimeric molybdenum(v)–cysteinate complex and its paramagnetic monomer in aqueous solutions. At concentrations of molybdenum(v) below 10^{-3} mol l^{-1} the monomer is stable. Similarities to the molybdenum of xanthine oxidase were proposed, on the bases: (*a*) that *g*-values and hyperfine splitting (studied by ^{95}Mo enrichment) are similar, (*b*) both systems show weak (perhaps *d–d*) absorption at 580 nm, (*c*) neither system shows nitrogen hyperfine splitting, (*d*) anisotropies suggest a highly distorted environment for Mo^V in each case.

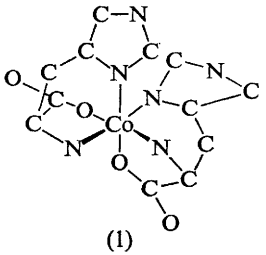
A number of cations of the formula $[Mo(\pi\text{-}C_5H_5)(NHR^1\cdot CHR^2\cdot COO)]^+$ were made [21] and separated as their chloride or hexafluorophosphate(v) salts, the amino-acids being glycine, alanine, valine, leucine, isoleucine, proline, methionine, sarcosine, and phenylalanine. All are attached to molybdenum through nitrogen and oxygen. The hydrogen atom(s) of the α-carbon atom exchange readily in mildly basic D_2O.

Complexes of Histidine. A novel form of binding of histidine was discovered [22] in the compound $Hg(C_6H_{10}N_3O_2)Cl_3$, in which each mercury ion is co-ordinated by three chlorides and one carboxylate oxygen of the histidine. No Hg—N bonds are involved. This finding may have relevance to the fact that the rhombohedral form of insulin forms only in the presence of the divalent ions of Zn, Mn, Fe, Co, Ni, Cu, or Cd, and that these can be replaced by Pb^{2+} but not by the mercury ion.

The structure of the complex $[Co^{II}(D\text{-}His\text{-}O)(L\text{-}His\text{-}O)],2H_2O$ has been established.[23] The relationship between the two terdentate histidinate

[19] J. R. Knox and C. K. Prout, *Acta Cryst.*, 1969, **B25**, 1857.
[20] T. J. Huang and G. P. Haight, *J. Amer. Chem. Soc.*, 1970, **92**, 2336.
[21] E. S. Gore and M. L. H. Green, *J. Chem. Soc.* (*A*), 1970, 2315.
[22] M. J. Adams, D. C. Hodgkin, and V. A. Raeburn, *J. Chem. Soc.* (*A*), 1970, 2632.
[23] R. Candlin and M. M. Harding, *J. Chem. Soc.* (*A*), 1970, 384.

groups, as shown in (1), differs from that in other bishistidinato-metal complexes.[24] The two histidine groups in the racemic complex have differing



(1)

conformations because of hydrogen-bonding requirements. In the context of cobalt(II)–histidine systems, an earlier observation that cobalt(II)–L-histidine solutions become deep violet at about pH 12 has now been attributed [25] to a change from octahedral to tetrahedral stereochemistry of cobalt.

The suggestion [26] that lanthanides may be useful as probes for calcium in biological systems has been taken up in a calorimetric and potentiometric study [27] of complexes of eight trivalent lanthanide ions (Ln) with histidine. Results were interpreted in terms of equilibria involving $[Ln(His\text{-}O)]^{2+}$, $[Ln(His\text{-}O)_2]^{+}$, and $[Ln(His)]^{3+}$.

Williams has also obtained [28] thermodynamic data for the complex formation between histidine and tryptophan with several first-row transition-metal divalent ions. He concludes that around pH 7.4 histidine acts as a terdentate ligand towards Mn, Fe, Co, and Ni, whereas tryptophan is bidentate. Furthermore, for most α-amino-acids the zinc complexes are more stable than those of cobalt, while for histidine the reverse is observed to be true; this is attributed to weak co-ordination of the carboxylate group, leaving a strained six-membered ring system.

Similar calorimetric studies [29] on histidine complexes of divalent ions show that ΔS for the formation of $[Cu(His\text{-}O)_2]$ is much smaller than for other divalent ions, which is interpreted in terms of a different stereochemistry of the copper(II) complex. Thornton and Skinner have obtained similar results [30] from microcalorimetric experiments on glycine, serine, threonine, and histidine with copper, nickel, and zinc ions. They also found a sharp change in enthalpy above pH $\sim$ 10 for the systems Cu–Ser and Cu–Thr, possibly due to the formation of hydroxo-bridged polynuclear complexes in the more alkaline solutions. These same systems, and

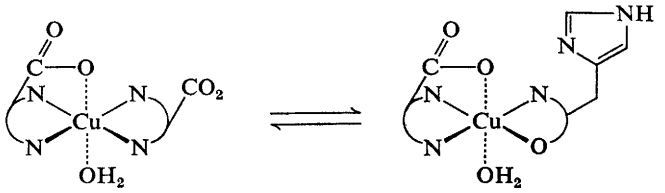
[24] Ref. 1, p. 263.
[25] P. J. Morris and R. B. Martin, *J. Amer. Chem. Soc.*, 1970, **92**, 1543.
[26] R. J. P. Williams, *Quart. Rev.*, 1970, **24**, 331.
[27] A. D. Jones and D. R. Williams, *J. Chem. Soc.* (*A*), 1970, 3138.
[28] D. R. Williams, *J. Chem. Soc.* (*A*), 1970, 1550.
[29] E. V. Raju and H. B. Mathur, *J. Inorg. Nuclear Chem.*, 1969, **31**, 425.
[30] A. C. R. Thornton and H. A. Skinner, *Trans. Faraday Soc.*, 1969, **65**, 2044.

Cu–His, but not Cu–Ala or Cu–Val, show[31] marked changes in visible absorption and c.d. spectra at pH's greater than 10; all these effects may be due to the amino-acids acting as terdentate ligands, since the pK's of the appropriate OH and NH groups are in this region.

Another study[32] of the binding of copper(II) to potentially terdentate amino-acids in aqueous solution comments on the histidine problem. The authors conclude that neutral histidine binds copper(II) through the glycine moiety with an unbound protonated imidazolium group. The first histidinate anion bound serves mainly as a terdentate ligand, with the two nitrogen atoms in the square-plane of the complex and a weaker co-ordination at the apical site by the carboxylate group; the second histidinate is thought to act as a bidentate chelating ligand. In the series of α,ω-L-diaminocarboxylate anions, lysine and ornithine bind to copper(II) as bidentate substituted glycines.

Calorimetric work[33] on complexes of copper(II) with histidine, histidine methyl ester, histamine, and imidazolepropionic acid leads to the conclusion that the two nitrogen atoms are involved in bonding from pH 2.8—10.0.

Another study,[34] using stability constant data and *d–d* absorption spectra at pH 8.3, leads to the conclusion that more than one kind of bis-ligand species is present, the major ones being (2) and (3).

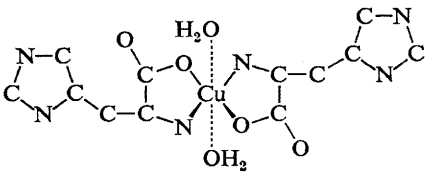


(2) (3)

From optical rotatory dispersion spectra of 1 : 1 : 1 mixtures of L-His, Cu^{2+}, and other amino-acids, compared with those of similar 1 : 1 : 1 mixtures with L-histidine methyl ester replacing L-histidine, it was concluded[35] that L-histidine co-ordinates to copper(II) through two nitrogen atoms, one being the nitrogen of the amino-group and the other the N-1 atom of the imidazole group. The same authors presented supporting evidence[36] for bonding through nitrogen atoms, using a previous observation that only moieties directly ligated to copper(II) were oxidised by hydrogen peroxide.

[31] R. D. Gillard and S. H. Laurie, unpublished observations.
[32] E. W. Wilson, M. H. Kasperian, and R. B. Martin, *J. Amer. Chem. Soc.*, 1970, **92**, 5365.
[33] J. L. Meyer and J. E. Bauman, *J. Amer. Chem. Soc.*, 1970, **92**, 4210.
[34] K. M. Wellman and B. K. Wong, *Proc. Nat. Acad. Sci. U.S.A.*, 1969, **64**, 824.
[35] H. Sigel, R. E. Mackenzie, and D. B. McCormick, *Biochim. Biophys. Acta*, 1970, **200**, 411.
[36] H. Sigel, R. Griesner, and D. B. McCormick, *Arch. Biochem. Biophys.*, 1969, **134**, 217.

Finally, the difficulties inherent in establishing structures of labile complexes in solution are emphasised by Evertsson's crystal structure [37] of the complex bis-L-histidinatocopper(II) dinitrate dihydrate, which crystallised at pH 3.7. The molecular structure has the features shown in (4). The imidazole groups are *not* co-ordinated, and turn away from the metal.



(4)

The confused situation in the copper–histidine equilibria no doubt arises from the multiplicity of possible binding sites and from the possibility of geometrical isomerism about the copper ion. There is some measure of agreement that, in neutral solution, both in the 1 : 1 and 2 : 1 complexes of histidine with copper, a nitrogen atom of imidazole is involved in bonding. Much remains to be done before a satisfactory description of the equilibria will emerge.

Complexes of Sulphur-containing Amino-acids. It has been accepted,[38] for methionine and *S*-methylcysteine, that while silver(I) bonds to the thioether group, mercury(II) does not. Natusch and Porter have, however, reported,[39] on the basis of proton resonance, that Hg^{II} bonds to the sulphur. The stability constant for the 1 : 1 complex of mercury with methionine, in 1M nitric acid, is 8.4×10^6 l mol^{-1}. Several new complexes of cobalt(III) of the type $[Co(en)_2L]^{n+}$ (en = ethylenediamine) have been reported;[40] when L = (cysteinate)$^{2-}$, or (cysteinate-*O*-ethyl ester)$^-$, bonding is through the nitrogen and sulphur atoms, but for (cysteinate-*S*-methyl ether)$^-$ bonding is through the nitrogen atom and the carboxylate oxygen atom. Methylation of the mercaptide in $[Co(en)_2(cysteinate)]^+$ results in a novel change in co-ordination from N—S chelation to N—O chelation.

Metal ions are known to catalyse the base hydrolysis of cysteine methyl (and ethyl) ester, and, in an attempt to characterise the complexes present, stability constants have been measured [41] for this ligand (HL) and several divalent metal ions. As an example of the findings, with nickel(II) the species $[NiL]^+$, $[NiL_2]$, $[NiL_3]^-$, $[Ni_4L_6]^{2-}$, $[NiHL]^{2+}$, and $[NiHL_2]^+$ occur.

The first structure of a copper complex of a sulphur-containing amino-acid to be determined is that [42] of bis(methioninato)copper(II), by Veidis

[37] B. Evertsson, *Acta Cryst.*, 1969, **B25**, 30.
[38] S. E. Livingstone and J. D. Nolan, *Inorg. Chem.*, 1968, **7**, 1447.
[39] D. F. S. Natusch and L. J. Porter, *Chem. Comm.*, 1970, 596.
[40] V. M. Kothari and D. H. Busch, *Inorg. Chem.*, 1969, **8**, 2277.
[41] L. J. Porter, D. D. Perrin, and R. W. Hay, *J. Chem. Soc.* (*A*), 1969, 118.
[42] M. V. Veidis and G. J. Palenik, *Chem. Comm.*, 1969, 1277.

and Palenik. The complex was crystallised from water; the copper ion is surrounded by a *trans* planar arrangement of the two nitrogens and two oxygens. The sulphur atoms are not attached to the copper. The authors suggest that the complex might be useful for carrying out a transmethylation on sulphur. In this paper, as in so many others, it is not specifically stated whether the amino-acid is of an optically-active form or is a racemic mixture. In view of the extra complications (discussed in the next section) which arise from possibly differing stereoselective equilibria in comparing complexes of the L-acid and the DL-acid, this information is absolutely essential. The structure of the complex of copper(II) with DL-methionine in aqueous solution has been deduced [43] from the n.m.r. spectrum of protons in both methionine and water, and again the chelation is found to be through the N—O group.

Crystal Structures. An important problem in amino-acid studies is whether structural information derived from diffraction experiments on solids can be applied to solution species in an attempt to rationalise reactions. Chiroptical properties (such as circular dichroism) which directly reflect molecular conformation would, if identical in both phases, provide some justification for the assumption that the conformations in solution and solid are the same.

This notion was applied [44] to the bis(amino-acidato) complexes of copper(II), with seemingly some promise that it might be a valid assumption. However, although the solid-state c.d. of bis(L-serinato)copper(II) is [44] closely similar to that of other bis(amino-acidato) complexes of known *trans* structure, it appears from the *X*-ray diffraction work [45] of van der Helm and Franks that the complex actually has the *cis* structure. Quite apart from reinforcing the necessity of considering *cis* ⇌ *trans* isomeric equilibria in copper(II) chemistry,[46] this finding throws doubt on the use of c.d. as a means of relating conformations in solid and solution. It appears, in fact, that the solid-state c.d. spectra of the *cis*- and *trans*-isomers of some bis(amino-acidato)copper(II) complexes may be surprisingly similar, notwithstanding their very different molecular stereochemistries.

Although the nickel(II) complex of L-serine is not isostructural with the copper(II) species (it is dihydrated, in contrast to the anhydrous copper complex), it also has [47] a *cis* arrangement of the two serinate ligands in the *xy*-plane about the nickel ion. The *z*-positions are occupied by the water molecules.

To add further to the large number of available crystal structures of copper(II) complexes of the type $[Cu(\alpha)_2]$*, we now have the five-co-

[43] M. R. Harrison and F. J. C. Rossotti, *Chem. Comm.*, 1970, 175.
[44] (*a*) R. D. Gillard and S. H. Laurie, *Chem. Comm.*, 1969, 489. (*b*) R. D. Gillard and S. H. Laurie, *J. Chem. Soc.* (*A*), 1970, 59.
[45] D. van der Helm and W. A. Franks, *Acta Cryst.*, 1969, **B25**, 451.
[46] Ref. 1, p. 263.
[47] D. van der Helm and M. B. Hossain, *Acta Cryst.*, 1969, **B25**, 457.

* See footnote on first page of this chapter.

ordinate *cis*-structure [48] of $[Cu(Ile\text{-}O)_2H_2O]$, in which the *xy*-plane is occupied by the two isoleucinate anions in a *cis* configuration. One of the L-isoleucine ligands has an unusual conformation, in that the four carbon atoms of the side-chain exist in a planar zig-zag configuration (as predicted by Bijvoet in 1954!).

In following up their observation that, in hot benzene, a reaction occurred between metallic copper and *NN*-dialkyl-α-amino-acids, Nash and Schaefer have established [49] the structure of the red complex bis(*NN*-diethylalaninato)copper(II). It has a *trans* planar configuration about the Cu^{II} ion, with a very short Cu—O bond of length 1.911 Å. The asymmetric C—O stretching frequency is very sensitive to solvent, ranging from 1684 cm^{-1} in benzene to 1657 cm^{-1} in a chloroform–t-butyl alcohol mixture. On concentrating solutions of the di-n-butylglycinate complex a colour change occurs:

$$\underset{\text{red } (\lambda_{max}\ 482\ nm)}{[Cu\{(Bu)_2Gly\text{-}O\}_2]} \longrightarrow \underset{\text{blue } (\lambda_{max}\ 590\ nm)}{[Cu\{(Bu)_2Gly\text{-}O\}_2]_n}$$

The blue form apparently contains polymeric aggregates, with the carboxylate function forming a bridge between copper ions; in this case, the asymmetric C—O stretching frequency occurs at 1645 cm^{-1}.

In a study [50] of mixed complexes of glycinate and racemic alaninate, $[Cu(Gly\text{-}O)(DL\text{-}Ala\text{-}O)],\frac{1}{2}H_2O$ was isolated as a crystalline compound, and it was shown that the unit cell contained equal quantities of the enantiomers [Cu(Gly-O)(L-Ala-O)] and [Cu(Gly-O)(D-Ala-O)]. As expected, this racemic crystal differs from the optically active crystals of $[Cu(Gly\text{-}O)(L\text{-}Ala\text{-}O)],\frac{1}{2}H_2O$.

A number of crystal structures of platinum-metal derivatives of amino-acids and peptides have appeared. These include the *trans*-isomer of bisglycinatoplatinum(II), in which [51] the dimensions of the glycine ligands are almost identical with those in the nickel(II) and copper(II) complexes. Freeman and Golomb describe [52] some valuable work on methionine complexes of platinum(II), in connection with the use of the tetrachloroplatinate(II) ion to prepare heavy-atom derivatives of proteins for structure analysis. The structures (by *X*-ray analysis) of $[Pt(DL\text{-}Met)Cl_2]$ and of $[Pt(L\text{-}Met)Cl_2]$ are reported; both have a *cis*-$PtCl_2$ moiety with methionine occupying the other two co-ordination positions by chelation through sulphur and nitrogen. The protonated carboxy-group is not involved in bonding to the metal.

The same paper summarizes the results of a combined *X*-ray and neutron diffraction study of a complex of glycyl-L-methioninate ion of the form $[NH_2\cdot CH_2\cdot CO\cdot N\cdot CH(R)\cdot CO_2H]$, which crystallised at pH *ca.* 2.5, of

[48] C. M. Weeks, A. Cooper, and D. A. Norton, *Acta Cryst.*, 1969, **B25**, 443.
[49] C. P. Nash and W. P. Schaefer, *J. Amer. Chem. Soc.*, 1969, **91**, 1319.
[50] L. F. Chapurina and A. V. Ablov, *Russ. J. Inorg. Chem.*, 1969, **14**, 796.
[51] H. C. Freeman and M. L. Golomb, *Acta Cryst.*, 1969, **B25**, 1203.
[52] H. C. Freeman and M. L. Golomb, *Chem. Comm.*, 1970, 1523.

formula [Pt(Gly-L-Met)Cl],H_2O. Co-ordination in the planar platinum complex is *via* the Cl^- ion and the N(amino), N(peptide), and S atoms of the peptide. The neutron diffraction study shows that the proton of the peptide link is dissociated and that the carboxy-group is still protonated.

It is concluded that the order of promoting the dissociation of the peptide proton is $Pt^{II} > Pd^{II} > Cu^{II} > Ni^{II}$. The structure of dichloro-DL-methioninepalladium(II) is [53] similar to that of the platinum(II) congener, in that methionine binds in the same way. The bond distances in the co-ordinated methionine ligand are similar to those in free methionine.

Other Studies. Katzin has reviewed [54] the available data on the c.d. of amino-acid complexes of nickel(II) and the lanthanides, and the results of a detailed study of the 1 : 1 and 2 : 1 complexes with nickel using c.d. are available.[55] A similar examination [56] of europium(III) complexes has also been made. Other lanthanide (Ln) complexes, of composition $LnCl_3,3(H_3N^+\cdot CH_2\cdot COO^-),3H_2O$ have been described.[57] The solvation in D_2O–D_2SO_4 media of the four isomers [58] of [Co(L-Ala-O)$_3$] has been studied [59] by n.m.r. spectroscopy.

A large number of stability constants have been reported. Those for the complexes of chromium(III) with L-aspartic acid (log K_1 = 10.1, log K_2 = 9.5) suggest that it acts as a terdentate ligand.[60] Childs and Perrin give [61] pH titration data, treated by computer, without the common assumption that only [M(α)$_n$] species exist. These data lead to equilibrium constants for the metal ions Cu^{II}, Zn^{II}, and Mn^{II} with the amino-acids (Hα) glycine, L-alanine, L-valine, and L-proline. Some novel species appear to be important, including $[Cu(H\alpha)]^{2+}$, $[Mn(H\alpha)]^{2+}$, and $[Mn(\alpha)(H\alpha)]^+$. Equally interesting is the conclusion that for [α] : [M] ratios up to at least 10 : 1, where M = Zn^{II} or Mn^{II}, the species M(α)$_3$ are unimportant. Another set of data on 1 : 1 complexes of Mn^{II} relates [62] to alanine and β-alanine, where, as commonly found, the β-alanine complex ($K_{20\,°C}$ = 3.48 × 10^2) is less stable than that of alanine ($K_{20\,°C}$ = 1.42 × 10^3).

Extensive new results [63] on the stabilities of arginine, ornithine, and citrulline with eight divalent metal ions follow the Irving–Williams order; complexes with Cu^{II}, Hg^{II}, and Zn^{II} of monoprotonated arginine are significantly more stable than those of monoprotonated ornithine.

[53] R. C. Warren, J. F. McConnell, and N. C. Stephenson, *Acta Cryst.*, 1970, **B26**, 1402.
[54] L. I. Katzin, *Co-ordination Chem. Rev.*, 1970, **5**, 279.
[55] L. I. Katzin, *J. Amer. Chem. Soc.*, 1969, **91**, 6940.
[56] L. I. Katzin, *Inorg. Chem.*, 1969, **8**, 1649.
[57] B. S. Mathur and T. S. Srivastava *J. Inorg. Nuclear Chem.*, 1970, **32**, 3277.
[58] Ref. 1, p. 268.
[59] B. M. Fung and I. H. Wang, *Inorg. Chem.*, 1969, **8**, 1867.
[60] H. Mizuochi, S. Shirakata, E. Kyuno, and R. Tsuchida, *Bull. Chem. Soc. Japan*, 1970, **43**, 397.
[61] C. W. Childs and D. D. Perrin, *J. Chem. Soc.* (*A*), 1969, 1039.
[62] L. P. Berezina, E. T. Zushinskaya, and A. I. Pozigun, *Russ. J. Inorg. Chem.*, 1969, **14**, 973.
[63] E. R. Clarke and A. E. Martell, *J. Inorg. Nuclear Chem.*, 1970, **32**, 911.

Copper(II) complexes have attracted rather less attention than in the past, although new values at 20 °C for L-proline ($\beta_2 = 3.1 \times 10^{15}$) and L-hydroxyproline ($\beta_2 = 1.0 \times 10^{18}$) are reported [64] which differ considerably from earlier data. Hydroxyproline was found to give polynuclear species. Complexes of *p*-, *m*-, and *o*-tyrosine and of β-phenylserine with copper(II) have been studied, and the suggestion has been made [65] that, for *o*-tyrosine, chelation involves the amino-nitrogen and the phenolic oxygen. The same workers have also published [66] the data shown in Table 1, which

Table 1 *Thermodynamic functions*[a] *for the formation of some bis(amino-acidato)metal*(II) *complexes at* 25 °C, I = 0.16 mol l^{-1}

Ligand	M^{2+}	Log β_2	$-\Delta G^b$	$-\Delta H^b$	ΔS^c
Serine	Cu	14.50	19.77	11.64	27.3
	Ni	9.98	13.61	9.03	18.7
Threonine	Cu	14.68	20.02	11.65	29.0
	Ni	9.97	13.60	8.17	18.2
Homoserine	Cu	14.69	20.03	11.39	28.1
	Ni	10.11	13.79	8.40	18.1
Isoserine	Cu	14.37	19.60	11.94	25.7
	Ni	7.85	10.71	6.82	13.5
4-Amino-3-hydroxybutyric acid	Cu	12.54	17.10	10.42	22.4
	Ni	7.17	9.78	5.31	15.0
β-Aminobutyric acid	Cu	12.85	17.52	10.52	23.6
	Ni	7.86	10.72	5.59	17.0

[a] Data from J. E. Letter and J. E. Bauman, *J. Amer. Chem. Soc.*, 1970, **92**, 437; [b] kcal mol^{-1}; [c] cal deg^{-1} mol^{-1}.

are in good agreement with comparable earlier results. For isoserine and Cu^{II} it is suggested that terdentate binding is significant.

On the basis of stability constants of its complexes with six divalent metal ions, *N*-(8-quinolyl)glycine is thought [67] to act as a terdentate ligand. The stabilities of mixed complexes between five divalent metal ions with nitrilotriacetate and amino-acids;[68] copper(II) with both 2,2′-bipyridyl and glycine or alanine or β-alanine;[69] and cobalt and zinc with ethylenediamine (or histamine) and serine [70] have been studied, where, in each case, the stability constant of a mixed complex is comparable with, or greater than, the mean of the constants for the homo-ligand complexes.

B. Stereochemistry.—Among the four-co-ordinate compounds of the type $[M(\alpha)_2]$, the *cis*- and *trans*-isomers of bis-serinatoplatinum(II) have been

[64] F. Karczynski and G. Kupryszewski, *Roczniki Chem.*, 1970, **44**, 967.
[65] J. E. Letter and J. E. Bauman, *J. Amer. Chem. Soc.*, 1970, **92**, 443.
[66] J. E. Letter and J. E. Bauman, *J. Amer. Chem. Soc.*, 1970, **92**, 437.
[67] T. Tanabe, K. Kimura, and S. Takamoto, *J. Chem. Soc. Japan*, 1969, **90**, 598.
[68] B. K. Afghan and J. Israeli, *Bull. Soc. chim. France*, 1969, 1393; J. Israeli and M. Cecchetti, *Canad. J. Chem.*, 1968, **46**, 3821, 3825.
[69] M. V. Chidambaram and P. K. Bhattacharya, *J. Inorg. Nuclear Chem.*, 1970, **32**, 3271.
[70] D. D. Perrin and V. S. Sharma, *J. Chem. Soc.* (*A*), 1969, 2060.

made [71] and characterised. Otherwise, most of the recent work relates to the popular problem [72] of the isomers of the complexes $[Cu(\alpha)_2]$.

E.s.r. spectra [73] of copper(II) chelates with the anions of glutamic acid, aspartic acid, leucine, threonine, L-histidine, and tyrosine gave no evidence for dimer formation at room temperature or in frozen solutions. A useful new criterion for distinguishing between geometrical isomers of $[Cu(\alpha)_2]$ appears [74] to be the multiplicity of bands in the far-i.r. spectrum in the region of the frequencies ν(Cu—N) and ν(Cu—O). In a survey of the spectra of seven complexes whose isomeric structure is known from diffraction studies, the method was borne out, and is extended [75] in a second paper to complexes of proline, where the complexes of the racemic ligand appear to have the *trans* structure, whereas for $[Cu(L\text{-Pro-O})_2]$ and its hydrate, a *cis* structure is indicated, as had been suggested [44b] on the basis of c.d. studies.

On the basis of e.s.r. spectra of solutions and powdered samples of the bis-complexes of DL- and L-alanine with copper(II), it has been suggested [76] that some stereoselective difference in electronic structure exists between the racemic and optically active forms in both phases. No comment is made in this paper on the differences in crystal lattices between the compounds.

The c.d. spectra of complexes $Cu(\alpha)_2$ have been further studied.[77, 78] Wellman and Wong [79] measured the equilibrium constant (K) for the reaction below, where L = an amino-acidate, oxalate, or ethylenediamine:

$$[Cu(Gly\text{-}O)_2] + [Cu(L)_2] \rightleftharpoons 2[Cu(Gly\text{-}O)(L)]$$

In these cases K = 4—6. However, for L = histidinol, $K > 10^3$, and this is attributed to the fact that steric hindrance in $[Cu(L)_2]$ is reduced on forming the mixed complex.

Stereoselectivity in complex formation by bidentate α-amino-acids of the type L-alanine, L-valine, *etc.* has been exhaustively sought in previous years and found to be lacking. Attention now focuses on the potentially terdentate amino-acids, such as histidine. The results in Table 2 show [80] that the complexes $[M(D\text{-His-O})(L\text{-His-O})]$ are more stable than $[M(L\text{-His-O})_2]$ {or, of course, $[M(D\text{-His-O})_2]$} for M = Co^{II} or Ni^{II}. The finding for cobalt(II) confirms an earlier report. Morris and Martin come

[71] L. M. Volshtein and T. R. Lastushkina, *Russ. J. Inorg. Chem.*, 1969, **14**, 246.
[72] Ref. 1, p. 263.
[73] J. F. Boas, J. R. Pilbrow, and T. D. Smith, *J. Chem. Soc.* (*A*), 1969, 723.
[74] A. W. Herlinger, S. L. Werhold, and T. V. Long, *J. Amer. Chem. Soc.*, 1970, **92**, 6474.
[75] A. W. Herlinger and T. V. Long, *J. Amer. Chem. Soc.*, 1970, **92**, 6481.
[76] H. Yokoi and T. Isobe, *Bull. Chem. Soc. Japan*, 1969, **42**, 2085.
[77] J. R. Gollogly, C. J. Hawkins, and C. L. Wong, *Inorg. Nuclear Chem. Letters*, 1970, **6**, 215.
[78] K. M. Wellman, S. Bogdansky, C. Piontek, C. R. Hare, and M. Mathieson, *Inorg. Chem.*, 1969, **8**, 1025.
[79] K. M. Wellman and B. K. Wong, *Chem. Comm.*, 1969, 1213.
[80] J. H. Ritsma, J. C. van de Grampel, and F. Jellinck, *Rec. Trav. chim.*, 1969, **88**, 411.

to the same conclusion [81] for the same metals and also for zinc(II). However, these latter authors find no stereoselectivity for the copper(II) system, in contrast to the work [82] of Barnes and Pettit, who have analysed their stability-constant data in terms of enthalpy and entropy contributions by

Table 2 *Stability constants*[a] *for some histidinate complexes at* 25 °C, I = 0.1 mol l^{-1}.

Ligand	M^{2+}	log K_1	log K_2
DL-His	Co	6.86_5	5.51_7
L-His	Co	6.86_4	5.39_0
DL-His	Ni	8.64_5	7.05_8
L-His	Ni	8.65_6	6.84_1

[a] Data from J. H. Ritsma, J. C. van de Grampel, and F. Jellinck, *Rec. Trav. chim.*, 1969, **88**, 411.

direct calorimetry. Their results are shown in Table 3. The ΔG values (from stability constants) and ΔH values run parallel, showing the stereoselectivity to be an enthalpy effect. Explanations of the greater stability for the racemic complex for nickel(II) and zinc(II), but for the optically

Table 3 *Enthalpy changes*[a, b] *for the reaction*

$$[M(His\text{-}O)]^+ + His\text{-}\bar{O} \rightleftharpoons [M(His\text{-}O)_2]$$

Ligand	Ni^{2+}	Cu^{2+}	Zn^{2+}
(+)-His	69.11	83.64	47.86
(−)-His	68.91	83.64	47.72
(±)-His	70.92	82.58	49.22

[a] Data from D. S. Barnes and L. D. Pettit, *Chem. Comm.*, 1970, 1000; [b] in kJ mol^{-1}.

active complex of copper(II), may reside in the varying possibilities for terdentate co-ordination which were described briefly in our last Report and in more detail elsewhere.[83]

Related to the studies mentioned above is the work [84] by Zompa, who has succeeded in preparing and separating three isomers of the bis(L-histidinato)cobalt(III) cation and has tentatively assigned structures. A similar separation is that [85] of the bis(L-2,4-diaminobutyrate)cobalt(III) ion. Nickel(II) complexes have also been studied preparatively [86] and by carbon-13 resonance.[87] Amino-acid amide complexes of nickel(II) [88] and copper(II) [89] have been described.

[81] P. J. Morris and R. B. Martin, *J. Inorg. Nuclear Chem.*, 1970, **32**, 2891.
[82] D. S. Barnes and L. D. Pettit, *Chem. Comm.*, 1970, 1000.
[83] R. D. Gillard, *Inorg. Chim. Acta, Rev.*, 1967, **1**, 69.
[84] L. J. Zompa, *Chem. Comm.*, 1969, 783.
[85] W. A. Freeman and C. F. Liu, *Inorg. Chem.*, 1970, **9**, 1191.
[86] C. A. McAuliffe and W. D. Perry, *J. Chem. Soc.* (*A*), 1969, 634.
[87] C. E. Strouse and N. A. Matwiyoff, *Chem. Comm.*, 1970, 439.
[88] T. Komorita, J. Hidaka, and Y. Shimura, *Bull. Chem. Soc. Japan*, 1969, **42**, 1782.
[89] T. Komorita, J. Hidaka, and Y. Shimura, *Bull. Chem. Soc. Japan*, 1969, **42**, 168.

Among the octahedral complexes formed by α-amino-acids, those of cobalt(III) remain pre-eminent in the number of new syntheses. Among those described recently are: six geometrical–optical isomers [90] of $[Co(C_2O_4)(L\text{-Ser-O})_2]^-$, optical isomers [91] of $[Co(C_2O_4)_2(L\text{-Ala-O})]^{2-}$ $[Co(C_2O_4)_2(\beta\text{-Ala-O})]^{2-}$, $[Co(malonate)_2(Gly\text{-O})]^{2-}$, and $[Co(malonate)_2(L\text{-Ala-O})]^{2-}$; and eleven complexes of the type [92] K[Co(NTA)(α)], where NTA is nitrilotriacetate and α = amino-acidate. Deductions of absolute stereochemistry continue; steric compression effects on the n.m.r. spectra of the amino-acid complexes have been successfully used for this purpose.[93, 94]

It has been known for some time that in a series of related complexes such as $[Co(NH_3)_4(L\text{-}\alpha)]^{2+}$, *cis*-$[Co(NH_3)_4(L\text{-}\alpha H)_2]^{3+}$, or $[Co(NH_3)_5(L\text{-}\alpha)]^{2+}$, the signs of the Cotton effects of the cobalt(III) chromophore may show a similarity throughout the series. It has been suggested that such regularities might allow convenient assignment of configuration to amino-acids. Three papers [95] take up this suggestion in detail, and in particular the authors consider the effects of added polarisable anions and changes in pH on the Cotton effects.

Several groups have worked on the stereospecificity of co-ordination of secondary amino-groups, as in sarcosine or proline. In this context, new synthetic work is described [96] on the complexes $[Co(en)_2(\alpha)]^{2+}$, where en = ethylenediamine and Hα is *N*-methyl-L-alanine, D-pipecolic acid, or L-proline, and on [97] $Na[Co(MeAla\text{-}O)_2(C_2O_4)]$.

The most impressive advances in this area come from the Australian schools, who have, in recent papers, resolved an apparent difficulty. It was known [98] that in $[Co(en)_2(Sar\text{-}O)]^{2+}$, and in $[Co(trien)(Sar\text{-}O)]^{2+}$, where trien = triethylenetetramine, co-ordination of sarcosinate is stereospecific, *i.e.* for a given configuration at the cobalt ion only one configuration at the nitrogen of sarcosine is possible. Proline on co-ordination also possesses an asymmetric nitrogen atom, and so it was considered [99] possible that with (*S*)-proline a stereospecific product should result from co-ordination to the [Co(trien)] moiety; a study of Dreiding models also apparently supported the authors' predictions. In fact, two products were found in nearly equal yield. With characteristic thoroughness, the authors separated the two products and showed them to have

[90] N. Matsuoka, J. Hidaka, and Y. Shimura, *Inorg. Chem.*, 1970, **9**, 719.
[91] K. Yamasaki, J. Hidaka, and Y. Shimura, *Bull. Chem. Soc. Japan*, 1969, **42**, 119.
[92] N. Koine, N. Sakota, J. Hidaka, and Y. Shimura, *Bull. Chem. Soc. Japan*, 1969, **42**, 1583.
[93] E. A. Berends and J. G. Brushmiller, *Inorg. Nuclear Chem. Letters*, 1970, **6**, 847.
[94] J. C. Dabrowiak and D. W. Cooke, *J. Amer. Chem. Soc.*, 1970, **92**, 1097.
[95] T. Yasui, J. Fujita, and Y. Shimura, *Bull. Chem. Soc. Japan*, 1969, **42**, 2081; C. J. Hawkins and P. J. Larsen, *Inorg. Chem.*, 1970, **9**, 6; *Austral. J. Chem.*, 1970, **23**, 1735.
[96] M. Saburi, M. Homma, and S. Yoshikawa, *Inorg. Chem.*, 1969, **8**, 367.
[97] E. A. Berends and J. G. Brushmiller, *Inorg. Nuclear Chem. Letters*, 1970, **6**, 531.
[98] Ref. 1, p. 267.
[99] D. A. Buckingham, L. G. Marzilli, I. E. Maxwell, A. M. Sargeson, and H. C. Freeman, *Chem. Comm.*, 1969, 583.

opposite configurations at the cobalt centres. Full *X*-ray analyses of the two isomers have been published [100] and show them to be L(−)$_{589}$-β_2-(*RRS*)-[Co(trien)(*S*-Pro-O)]$^{2+}$ and D(+)$_{589}$-β_2-(*SSS*)-[Co(trien)(*S*-Pro-O)]$^{2+}$. Formation of the unexpected D(+)-isomer can be attributed to angular distortions in the triethylenetetramine ligand. Isomers of the related complex [Co(trien)(*S*-MeAla-O)]$^{2+}$ have also been obtained [101] and the rates of mutarotation at the asymmetric nitrogen and carbon centres have been followed.

Little has appeared on stereoselective reactions of amino-acid complexes, though this is without doubt an area of much promise. Leach and Angelici have used [102] L-valine-*N*-monoacetato-copper(II) [Cu(L-AcOVal-O)] as a means of testing the discrimination of an optically active metal ion between the two enantiomers of a substrate amino-acid. Equilibrium constants were determined for the reactions:

$$[\text{Cu(L-AcOVal-O)}] + \text{L-}\alpha \underset{}{\overset{K_L}{\rightleftharpoons}} [\text{Cu(L-AcOVal-O)(L-}\alpha\text{)}]$$

$$[\text{Cu(L-AcOVal-O)}] + \text{D-}\alpha \underset{}{\overset{K_D}{\rightleftharpoons}} [\text{Cu(L-AcOVal-O)(D-}\alpha\text{)}]$$

For α = alanine, leucine, serine, and phenylalanine, $K_L = (3.3—6.5)K_D$. For α = valine, $K_D = 2.5K_L$. Rates of ester hydrolysis were essentially non-stereoselective; for example, the rate constants for the hydrolysis of D- and L-alanine methyl ester were equal.

Finally, in an initial study of bacterial growth on complexed amino-acids as sole nitrogen source, it was found [103] that *Enterobacter cloacae* metabolised the enantiomers of 1,2,6-trisglycinatocobalt(III) at very different rates, such that, under specific conditions, a complete resolution of the cobalt complex is possible, the isomer D(+)-[Co(Gly-O)$_3$] remaining unchanged when all L-isomer has been used.

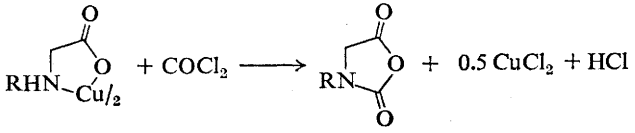
C. Reactivity.—The bulk of the recent work on reactivity of amino-acids in complexes falls under three headings: rates of complexation, studies of metal-catalysed oxidations, and metal-catalysed ester hydrolyses. We treat these in turn.

Some isolated observations, not readily categorized, are listed first. In line with Mertz's suggestion [6] that chromium(III) should be regarded as biochemically active and an essential metal, Weser and Koolman have studied [104] the effect of this and other metal ions on the rate of incorporation of ^{14}C-labelled amino-acids into whole rat-liver nuclei. Mercuric ion (known to form very stable derivatives with sulphur and nitrogen donors) reduces the incorporation to about 30% of the level in a control

[100] H. C. Freeman and I. E. Maxwell, *Inorg. Chem.*, 1970, **9**, 649; H. C. Freeman, L. G. Marzilli, and I. E. Maxwell, *ibid.*, p. 2408.
[101] D. A. Buckingham, I. E. Maxwell, and A. M. Sargeson, *Chem. Comm.*, 1969, 581.
[102] B. E. Leach and R. J. Angelici, *J. Amer. Chem. Soc.*, 1969, **91**, 6296.
[103] R. D. Gillard and C. Thorpe, *Chem. Comm.*, 1970, 997.
[104] U. Weser and J. Koolman, *Experientia*, 1970, **26**, 246.

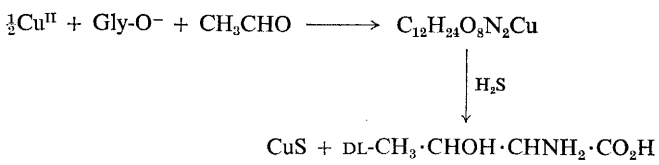
experiment. Iron(III) and manganese(II) salts have little effect, but chromium(III) salts cause at least a 400% increase in the rate. Brown and Edwards have shown [105] that the common assumption that mercurials react with thiol groups but not with disulphide links is incorrect. Mercuric salts react readily with cyclic and linear aryl and alkyl disulphides. The decomposition of cysteine–mercury compounds by aldehydes has been briefly reported,[106] as has the photochemical reaction [107] of L-histidine with ferrocyanide, yielding histidinate complexes of iron(II).

The changed reactivity of co-ordinated amino-acids has not been extensively applied recently, although Hamilton and Lyman prepared [108] *N*-carboxy-α-amino-acid anhydrides (of interest as polypeptide precursors) in good yield by reacting copper complexes of amino-acids with phosgene in tetrahydrofuran at room temperature for 1—2 h (see Scheme 1).



Scheme 1

Bisglycinatocopper(II) has been known for some time to react with aldehydes to give, after treatment with hydrogen sulphide, primarily threonine (Scheme 2). An intermediate is copper(II) bis-2,4-dimethyloxazolidine-4-

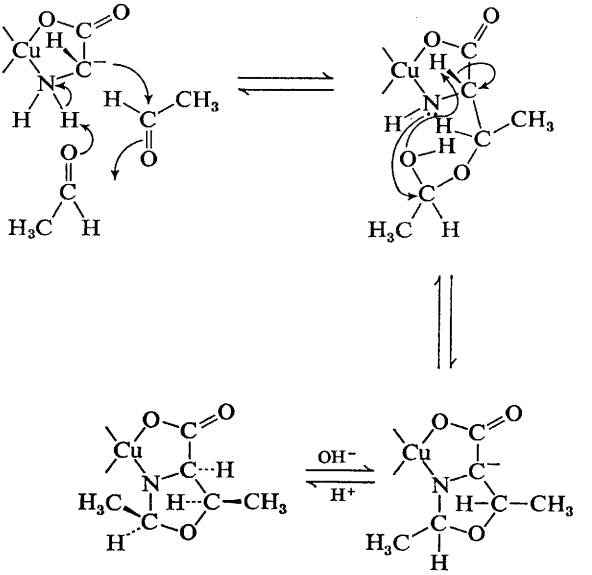


Scheme 2

carboxylate dihydrate, whose structure has now been established [109] by *X*-ray diffraction. The structure which was found fits the mechanism for the synthesis shown in Scheme 3.

In line with experience on cobalt(III) (though not with copper [110]), complexes of asymmetric α-amino-acids, it was recently found [111] that for *trans*-[Pt(L-Glu-O)$_2$]$^{2-}$ the rates of base-catalysed proton exchange and of inversion at the asymmetric carbon are comparable. The rates of the base-

[105] P. R. Brown and J. O. Edwards, *Biochemistry*, 1969, **8**, 1200.
[106] M. Wronski and W. Goworek, *Chem. analit.*, 1969, **14**, 387.
[107] W. U. Malik and M. Aslam, *Indian J. Chem.*, 1970, **8**, 736.
[108] R. D. Hamilton and D. J. Lyman, *J. Org. Chem.*, 1969, **34**, 243.
[109] J. P. Aune, P. Maldonado, G. Larcheres, and M. Pierrot, *Chem. Comm.*, 1970, 1351.
[110] R. D. Gillard and D. A. Phipps, *Chem. Comm.*, 1970, 800.
[111] L. E. Erickson, A. J. Dappen and J. C. Uhlenhopp, *J. Amer. Chem. Soc.*, 1969, **91**, 2510.



Scheme 3

catalysed N—H proton exchange in $[PtCl_2(Gly\text{-}O)]^-$ and in $[PtCl_2(Sar\text{-}O)]^-$ were also measured. In a similar paper,[112] using $[Pt(NH_3)_3(Sar\text{-}O)]^+$ and $[Pt(NH_3)_2(edda)]$, where edda = ethylenediaminediacetate, the rate of inversion at the asymmetric nitrogen was less than the N—H exchange rate by a factor of seventy.

Rates of Binding. As in our last Report, there have been several measurements of the rates of complexation of amino-acids by metal ions, using the techniques applicable to fast reactions. For example, the results shown in Table 4 were obtained [113] for serine using the *T*-jump method. The rate-determining step appears to be the loss of a water molecule in all cases except $CuL + L \rightleftharpoons CuL_2$. Similar results [114] for the complexation of sarcosine, undertaken to establish whether the *N*-methyl binding-site is important in kinetic terms, are shown also in Table 4.

Rate constants are also available [115] for the formation of copper(II) complexes with alaninate and with β-alaninate and the histidine zwitterion.

[112] L. E. Erickson, H. L. Fritz, R. J. May, and D. A. Wright, *J. Amer. Chem. Soc.*, 1969, **91**, 2513.

[113] R. L. Karpel, K. Kustin, and R. F. Pasternack, *Biochim. Biophys. Acta*, 1969, **177**, 434.

[114] R. F. Pasternack, K. Kustin, L. A. Hughes, and E. Gibbs, *J. Amer. Chem. Soc.*, 1969, **91**, 4401.

[115] W. B. Makinen, A. F. Pearlmutter, and J. E. Stuehr, *J. Amer, Chem. Soc.*, 1969, **91**, 4083.

For alanine, forward rate constants (k_1 = 1.3 × 10^9 and k_2 = 1.5 × 10^8 l mol^{-1} s^{-1}) are larger than for β-alanine (2.0 × 10^8 and 8 × 10^6 l mol^{-1} s^{-1}), and those for histidine (1.3 × 10^7 and 3.0 × 10^6 l mol^{-1} s^{-1}) are smaller yet. For both β-alanine and histidine the rate-determining step is said to be ring closure.

Table 4 *Rate-constants for* L-*serine complexes*[a] *and sarcosine complexes*[b] *for the reaction*

$$L + ML_{n-1} \underset{k_{-n}}{\overset{k_n}{\rightleftharpoons}} ML_n$$

	Co^{II}	Ni^{II}	Cu^{II}	Zn^{II}
		L-*Serine*		
k_1(l mol^{-1} s^{-1})	2.0 × 10^6	2.9 × 10^4	2.5 × 10^9	~1 × 10^8
k_{-1}(s^{-1})	93	0.11	32	~2 × 10^3
k_2(l mol^{-1} s^{-1})	2.0 × 10^6	3.4 × 10^4	5 × 10^8	—
k_{-2}(s^{-1})	930	1.6	150	—
k_3(l mol^{-1} s^{-1})	—	3.0 × 10^4	—	—
k_{-3}(s^{-1})	—	28	—	—
		Sarcosine		
k_1(l mol^{-1} s^{-1})	9.2 × 10^5	1.3 × 10^4	2.8 × 10^9	—
k_{-1}(s^{-1})	57	4.1 × 10^{-2}	32	—
k_2(l mol^{-1} s^{-1})	1.5 × 10^6	1.2 × 10^4	1.0 × 10^8	—
k_{-2}(s^{-1})	570	0.50	22	—

[a] R. L. Karpel, K. Kustin, and R. F. Pasternack, *Biochim. Biophys. Acta*, 1969, **177**, 434; [b] R. F. Pasternack, K. Kustin, L. A. Hughes, and E. Gibbs, *J. Amer. Chem. Soc.*, 1969, **91**, 4401.

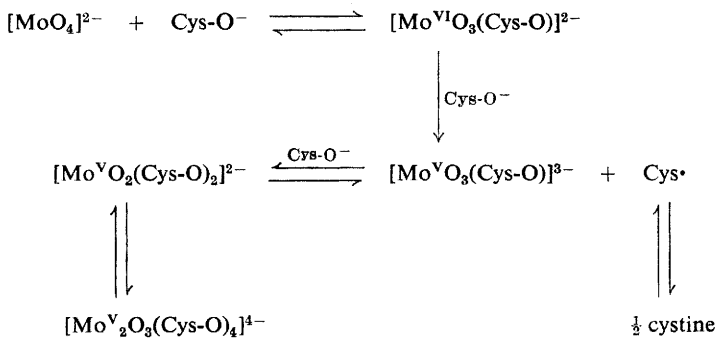
The rates of formation of the mixed complexes of copper(II)–2,2′-bipyridyl–glycine [116] and of copper(II)–histamine–serine [117] have been measured. A mechanism has been suggested [118] for the protonation of a Gly-Gly-Gly complex of nickel(II) which involves as the rate-determining step the cleavage of the Ni—N(peptide) bond after interaction with the acid.

Oxidation and Reduction. The redox reaction between the Mo^V–Mo^{VI} couple and the thiol–disulphide system provides useful background information for work on molybdenum-containing enzymes. Martin and Spence have deduced,[119] for the oxidation of L-cysteine by Mo^{VI} at pH 7.5 (phosphate buffer), the rate-law:

$$-\frac{d[Mo^{VI}]}{dt} = k_3[Mo^{VI}][Cys\text{-}O]^2$$

[116] R. F. Pasternack and H. Sigel, *J. Amer. Chem. Soc.*, 1970, **92**, 6146.
[117] V. S. Sharma and D. L. Leussing, *Chem. Comm.*, 1970, 1278.
[118] E. J. Billo and D. W. Margerum, *J. Amer. Chem. Soc.*, 1970, **92**, 6811.
[119] J. F. Martin and J. T. Spence, *J. Phys. Chem.*, 1970, **74**, 2863.

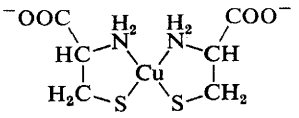
They propose the mechanism:



The oxidation of glutathione by molybdate was also studied. Other work on cysteine includes a careful polarographic study [120] of the reduction of cobalt ions in the presence of cystine and cysteine.

The rapid interaction of cysteinate complexes of iron(II) in water with oxygen has been examined,[121] a serious contribution to the perennial problem of the autoxidation of cysteine. By controlling the pH, contributions from the 1 : 1 and 2 : 1 cysteine–iron complexes could be separated, using the stopped-flow method. The rate constants (at 25 °C) for the oxygen reaction were, for the 1 : 1 species, $5 \pm 1 \times 10^3$ l mol^{-1} s^{-1}, and, for the 2 : 1 species, $2 \pm 0.5 \times 10^4$ l mol^{-1} s^{-1}, both values being a good deal greater than for other $Fe^{II} + O_2$ reactions.

An interesting paper,[122] with possible relevance to the mode of catalytic action of cysteamine oxidase, treats the copper-catalysed oxidation of cysteine to cystine. The yellow coloration [λ_{max}330 nm and 400 nm (sh)] which occurs on adding alkaline cysteine solution to a solution of copper(II) is shown to be due to the complex $[Cu(Cys\text{-}O)_2]^{2-}$, with the proposed structure (5). During the uptake of di-oxygen (O_2), this complex is present,



(5)

but it quickly disappears at the end of the reaction (when no cysteine is left). At this point, all the copper may be recovered as copper(I) by trapping with neocuproin.

[120] M. Zielinski and J. Kůta, *Coll. Czech. Chem. Comm.*, 1969, **34**, 2523.
[121] A. D. Gilmour and A. McAuley, *J. Chem. Soc.* (*A*), 1970, 1006.
[122] D. Cavallini, G. De Marco, S. Dupre, and G. Rotilo, *Arch. Biochem. Biophys.*, 1969, **130**, 354.

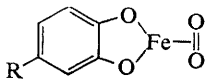
Oxygenation of solutions containing cobalt(II) and amino-acids or peptides continues to attract attention. Petru and Jursik have observed,[123] for the glycinate system, an intermediate brown species, presumably a peroxo-bridged dicobalt(III,III) complex.

At pH 9.3, the uptake of di-oxygen by a solution containing L-histidine and cobalt(II) has been shown [124] to lead to two oxygen-carrying species, each with a Co : O_2 ratio of 2 : 1. The author suggests that the orientation of the di-oxygen unit may differ in the two species. This is strongly reminiscent of a finding,[125] for peptide–cobalt(II) systems, that uptake leads rapidly to a brown peroxo-dicobalt(III,III) species, which then more slowly forms a hydroperoxo-dicobalt(III,III) complex containing the unit Co Co. The requirements and conditions for oxygenation and oxida-

tion of cobalt(II) chelates with a variety of amines, amino-acids, and peptides have been examined.[126] The conclusion is reached that a minimum of three nitrogen donors per cobalt(II) ion is necessary for an oxygenated complex to form. This was first recognised by Fallab.[127]

One of the most famous and frequently quoted observations in the field of metal catalysis of amino-acid reactions is that [128] by Shibata and Tsuchida on the rates of oxygenation of 3,4-dihydroxyphenylalanine (Dopa) in the presence of cobalt(III) complexes. They reported that, in the presence of (−)-$[Co(en)_2(NH_3)Cl]^{2+}$ ions, L-Dopa was oxidised more rapidly than was D-Dopa. This work has now been repeated [129] and the conclusions are not confirmed. The rate of di-oxygen uptake and the optical rotation are almost the same for mixtures of L-Dopa and (+)D-, (−)D-, or racemic $[Co(en)_2(NH_3)Cl]^{2+}$. The most rapid reaction occurred with *trans*-$[Co(en)_2Cl_2]^+$. The optical rotatory dispersion spectra of reaction mixtures were independent of the particular cobalt complex used, suggesting a common intermediate. A further report has also appeared [130] on the autoxidation of L-Dopa with iron salts as catalysts. The rate of oxidation is proportional to $[Fe^{3+}][O_2][H^+]^{-2}$, and the suggested intermediate is (6), the rate-determining step being electron transfer through the metal ion to the di-oxygen ligand.

[123] F. Petru and F. Jursik, *Coll. Czech. Chem. Comm.*, 1969, **34**, 3153.
[124] S. Bagger, *Acta Chem. Scand.*, 1969, **23**, 975.
[125] R. D. Gillard and A. Spencer, *J. Chem. Soc.* (*A*), 1969, 2718.
[126] M. S. Michailidis and R. B. Martin, *J. Amer. Chem. Soc.*, 1969, **91**, 4683.
[127] S. Fallab, *Angew. Chem. Internat. Edn*, 1967, **6**, 496.
[128] Y. Shibata and R. Tsuchida, *Bull. Chem. Soc. Japan*, 1929, **4**, 142.
[129] K. Yamasaki and Y. Yoshikawa, *Proc. 11th Internat. Conf. Co-ord. Chem.*, Israel, 1968, ed. M. Cais, Elsevier, New York.
[130] R. F. Jameson, *Proc. 11th Internat. Conf. Co-ord. Chem.*, Israel, 1968, ed. M. Cais, Elsevier, New York.



(6)

Catalytic decompositions of hydrogen peroxide by cobalt(II) complexes [131] of *N*-substituted amino-acids, *e.g.* *N*-benzoylglycine, and by the copper(II)–histidine system [132] have been studied kinetically.

Ester Hydrolysis. Small rate-enhancements for the base hydrolysis of cysteine methyl ester in the presence of divalent metal ions have been established,[133] *e.g.* with nickel(II), the bis-ligand complex hydrolyses seventy-five times faster than the free anion. The rates of hydrolysis of amino-acid esters have been measured [134] in the presence of a number of complexes of copper(II) with substituted iminodiacetate ligands, $[R^1N(CHR^2 \cdot COO)(CH_2 \cdot COO)]^{2-}$.

The important work of Buckingham and Sargeson on the cobalt(III)-promoted reactions of amino-acid derivatives continues. They report [135] a novel amidolysis of glycine ethyl ester, in which a co-ordinated ammonia group is the nucleophile:

$$[Co(NH_3)_5(NH_2 \cdot CH_2 \cdot CO_2C_2H_5)]^{3+} \xrightarrow{\text{pH 9—14}} \underset{\text{A}}{[Co(NH_3)_5(NH_2 \cdot CH_2 \cdot CO_2)]^{2+}} + \underset{\text{B}}{[Co(NH_3)_4(NH_2 \cdot CH_2 \cdot CO \cdot NH)]^{2+}}$$

Cation B, which contains glycinamide chelated through two nitrogen atoms, may be isolated in almost 100% yield from the reaction using 0.1M sodium hydroxide. The compound B reacts with perchloric acid to give a derivative which is protonated on the oxygen of the amide group rather than the nitrogen.

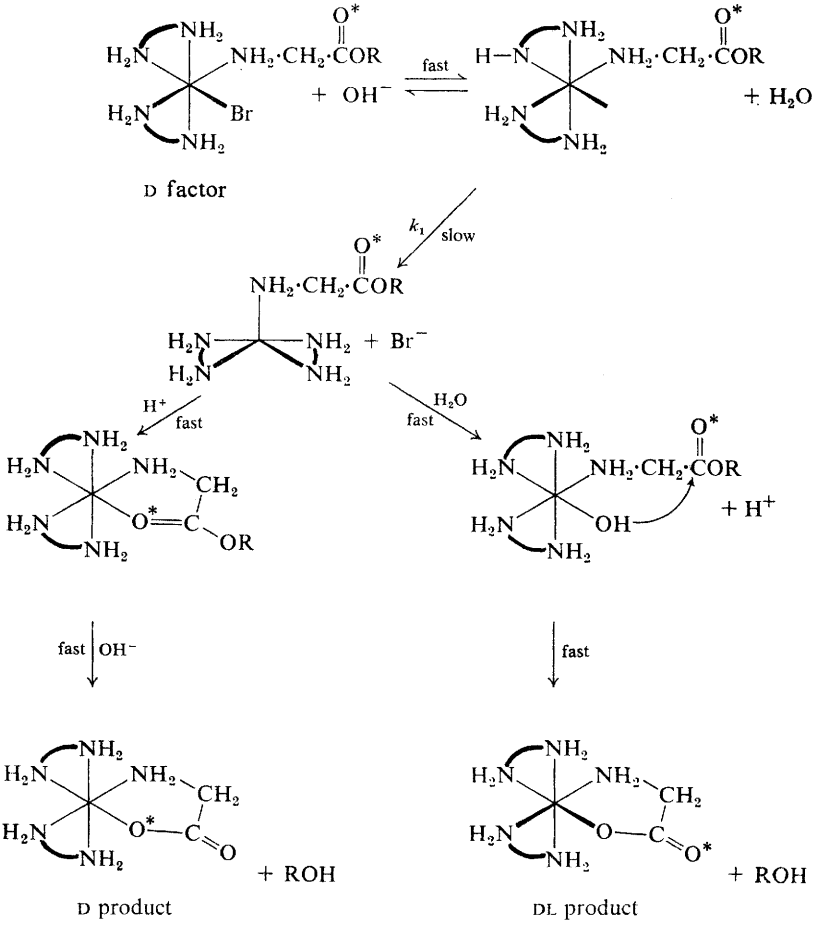
The mechanisms of the reactions:

$$cis\text{-}[Co(en)_2X(NH_2 \cdot CH_2 \cdot CO_2R)]^{2+} + OH^- \longrightarrow [Co(en)_2(Gly\text{-}O)]^{2+} + X^- + HOR$$

where X = Cl or Br, R = Me, Et, Pr^i, Bu^n, Bu^t, or $CH_2C_6H_5$, have been studied.[136] Oxygen-18 labelling experiments with *cis*-$[Co(en)_2Br(H_2NCH_2 \cdot CO_2 \cdot CHMe_2)]^{2+}$ show that around half of the

[131] A. Y. Sychev and B. N. Tkho, *Russ. J. Phys. Chem.*, 1969, **43**, 653.
[132] V. S. Sharma, J. Schubert, H. B. Brooks, and F. Sicilio, *J. Amer. Chem. Soc.*, 1970, **92**, 822.
[133] R. W. Hay and L. J. Porter, *J. Chem. Soc.* (*A*), 1969, 127.
[134] B E. Leach and R. J. Angelici, *Inorg. Chem.*, 1969, **8**, 907.
[135] D. A. Buckingham, D. M. Foster, and A. M. Sargeson, *J. Amer. Chem. Soc.*, 1969, **91**, 3451.
[136] D. A. Buckingham, D. M. Foster, and A. M. Sargeson, *J. Amer. Chem. Soc.*, 1969, **91**, 4102.

$[Co(en)_2(Gly\text{-}O)]^{2+}$ produced arises from co-ordination of the ester carbonyl oxygen, while the remainder is produced by the intervention of a solvent oxygen atom. The latter path is regarded as an internal nucleophilic displacement by co-ordinated hydroxyl ion, comparable with the amidolysis [135] described above. The suggested mechanisms (consistent with the optical activity of products) are shown in Scheme 4.



Scheme 4

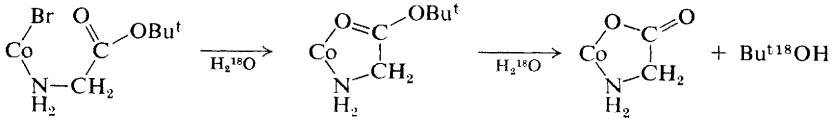
A third paper shows [137] that the large increase in the rate of hydrolysis (by *ca.* 10^6) of glycine isopropyl ester on co-ordination to cobalt(III) is attributable to a large positive entropy of activation (*ca.* 15 cal mol^{-1}

[137] D. A. Buckingham, D. M. Foster, and A. M. Sargeson, *J. Amer. Chem. Soc.*, 1969, **91**, 5701.

deg^{-1}). The acid hydrolysis of the co-ordinated glycine ethyl ester, induced by mercury(II) ion, is also described.[138] Wu and Busch have demonstrated,[139] by oxygen-18 labelling, that for the mercury-induced reaction below,

$$[Co(en)_2X(NH_2 \cdot CH_2 \cdot CO_2Bu^t)]^{2+} \longrightarrow [Co(en)_2(NH_2 \cdot CH_2 \cdot CO_2)]^{2+} + X^- + Bu^tOH$$

alkyl–oxygen bond breakage occurs, as in Scheme 5.



Scheme 5

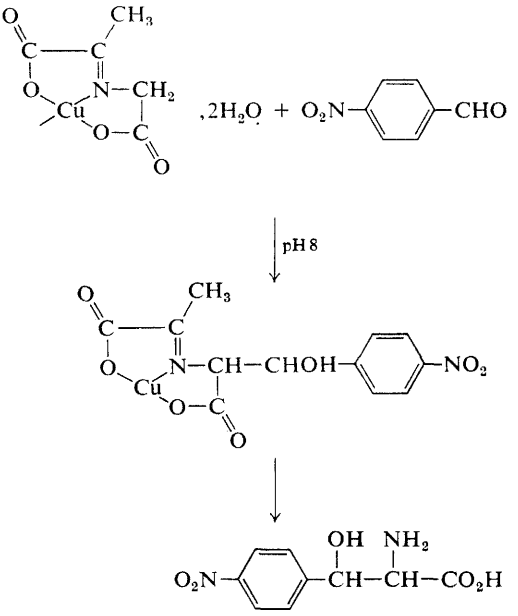
Buckingham, Foster, and Sargeson have collected together[140] many of their findings on the intramolecular and intermolecular hydrolyses of substituted glycinamides in *cis*-$[Co(en)_2Br(Gly\text{-}NR^1R^2)]^{2+}$ ions following base hydrolysis of the bromide. They also relate their findings to possible mechanisms operative for the promotion by divalent cations of ester and amide hydrolyses, and extend their comments to the possible mechanisms of hydrolytic enzyme reactions catalysed by divalent metal ions, *e.g.* zinc(II) in carboxypeptidase A.

In a related study,[141] the rates of loss of optical activity by aqueous solutions of the complexes *cis*-β_2-$[Co(trien)(L\text{-}\alpha)]I_2$ have been measured, where trien = triethylenetetramine, αH = proline or phenylalanine. Results are explained in terms of (*a*) racemisation at the cobalt(III) centre, (*b*) racemisation at the α-carbon atom of the amino-acid, and (*c*) displacement of various ligands by water or hydroxide ion.

D. Schiff Base Derivatives.—The principal interest in these systems, obtained from amino-acids or peptides and reactive carbonyl functions such as those in pyruvic acid, salicylaldehyde, or pyridoxal, is in the activation of the α-CHR position of the amino-acid. We consider recent developments in the order of increasing complexity of the carboxy-compound, with pyruvate first, then salicylaldehyde, then pyridoxal.

Japanese workers[142] have reported syntheses of β-hydroxy-DL-amino-acids by the treatment with aldehydes of the copper(II) complex of the Schiff base from glycine and pyruvic acid, as shown in Scheme 6. Good yields (*ca.* 80%) were obtained at 25 °C at low ratios of aldehydes to Schiff base complex, *e.g.* acetaldehyde gave DL-threonine in 82% yield.

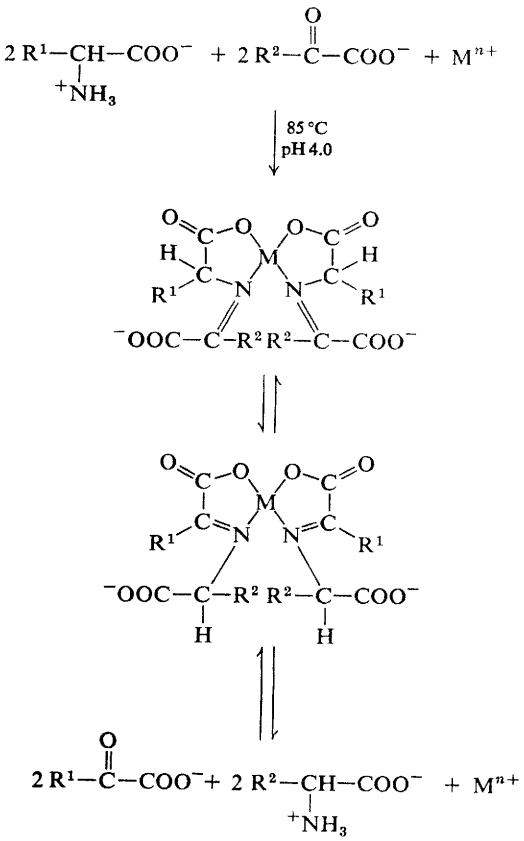
[138] D. A. Buckingham, D. M. Foster, L. G. Marzilli, and A. M. Sargeson, *Inorg. Chem.*, 1970, **9**, 11.
[139] Y. Wu and D. H. Busch, *J. Amer. Chem. Soc.*, 1970, **92**, 3326.
[140] D. A. Buckingham, D. M. Foster, and A. M. Sargeson, *J. Amer. Chem. Soc.*, 1970, **92**, 6151.
[141] M. H. Ghandihari, T. N. Anderson, and D. R. Boone, *J. Amer. Chem. Soc.*, 1970, **92**, 6466.
[142] T. Ichikawa, S. Maeda, Y. Araki, and Y. Ishido, *J. Amer. Chem. Soc.*, 1970, **92**, 5514.



Scheme 6

Non-enzymatic transaminations are the subject of a further report [143] by Doctor and Oró. An earlier report had mentioned that in non-enzymatic transaminations, histidine was more effective than other amino-acids. To elucidate this, the reaction was studied for various amino-acids and α-oxoglutaric acid, catalysed by several ions, *e.g.* Al^{3+}, Cu^{2+}, Fe^{2+}, and Fe^{3+}. The optimum pH was found to be four, and histidine was confirmed as the most active of the fourteen amino-acids, which is attributed to the involvement of the tertiary nitrogen of the imidazole in the formation of stable, sparingly soluble, metal–imidazoyl pyruvate complexes as end-products. The proposed mechanism is shown in Scheme 7. Leussing and his group have undertaken further kinetic studies on the mechanisms of transaminations, and report first [144] on the metal-dependent rate of formation of *N*-pyruvylideneglycinate in the presence of zinc(II) at pH 4.5—7.6 measured by the stopped-flow technique. In this reaction the authors state that zinc plays a 'promnastic' rôle; promnastic means 'to be a match-maker', and the distinction between this and the so-called 'template' reactions is that in these 'promnastic' cases the metal serves merely to assemble the reactants, whereas in the template case it has the further function of activating them. In a second report,[145] the starting point is the

[143] V. M. Doctor and J. Oró, *Biochem. J.*, 1969, **112**, 691.
[144] D. L. Leussing and L. Anderson, *J. Amer. Chem. Soc.*, 1969, **91**, 4698.
[145] H. Scheidegger, W. Felty, and D. L. Leussing, *J. Amer. Chem. Soc.*, 1970, **92**, 808.



Scheme 7

fact that tautomeric forms of isomeric Schiff base complexes with metal ions may have appreciably different stability constants, and in commenting on the possible disturbance of the equilibria by metal ions, a knowledge of such stability constants is vital. The results show that for the uncomplexed Schiff bases, the tautomerization constant β_t is 0.6 for $(pyr{\cdot}Glu\text{-}O)^{3-} \rightleftharpoons (\alpha Kg{\cdot}Ala\text{-}O)^{3-}$, where pyr = pyruvate, Glu-O$^-$ = glutamate, αKg = α-ketoglutarate, and Ala-O$^-$ is alaninate. However, after complexing to zinc, the isomerization constant β_{Znt} is 0.08 for $[Zn(pyr{\cdot}Glu\text{-}O)]^- \rightleftharpoons [Zn(\alpha Kg{\cdot}Ala\text{-}O)]^-$.

Using salicylaldehyde (Hsal) as the carbonyl component, Japanese workers have described [146] reactions of $Cu(sal)_2$ with glycinamide, esters of

[146] K. Hamada, H. Ueda, Y. Nakao, and A. Nakahara, *Bull. Chem. Soc. Japan*, 1969, **42**, 1297.

glycine, and the amide and esters of glycylglycine. The crystal structure of *N*-salicylideneglycinatoaquocopper(II) tetrahydrate has been reported.[147]

In complexes between copper(II) and salicylaldehyde Schiff bases with amino-acids and peptides, some specific activation is present.[148] For example, with glycylglycine in alkaline media, the α-CH_2 protons exchange with heavy water only for the *N*-terminal amino-acid. Based on the isolation of a carbinolamine complex of copper(II) from an amino-acid–salicylaldehyde system, the suggestion has been made [149] that such complexes may be the reactive intermediates of 'Schiff base' systems.

Complexes of *N*-salicylidene-amino-acids (L and DL) forms have been described [150] for nickel(II) with glycine, alanine, leucine, valine, phenylalanine, and methionine, and a similar series for vanadium in both the (IV) and (V) oxidation states is the subject of two papers.[151, 152]

Recent work on pyridoxal itself as the source of the carbonyl moiety used in forming Schiff bases with amino-acids, includes a study [153] of the reaction between pyridoxal phosphate (PLP) and glutamate in the presence of copper(II) ions. At pH 4, the rate of formation of the copper–Schiff base complex is first-order in both PLP and glutamate, but zeroth-order in copper(II) ion. The metal ions are said to act as a trap for both the Schiff base and an amino-alcohol intermediate. A detailed equilibrium study of the species in the systems Zn^{II}–PLP–αH, where αH = glycine or L-alanine, is the subject of another report.[154]

Further work has been carried out on the vexed question of the nature of the species in solutions containing an amino-acid, pyridoxal, and a metal ion. A facile method of preparing DL-[β-$^2H_3$]alanine is described.[155] In a later paper [156] summarizing the zinc(II) and aluminium(III) results, Gansow and Holm, referring to the three reactions (A), (B), and (C),

$$R^1CHNH_2\cdot CO_2H + \text{pyridoxal} \rightleftharpoons R^1CO\cdot CO_2H + \text{pyridoxamine} \quad (A)$$

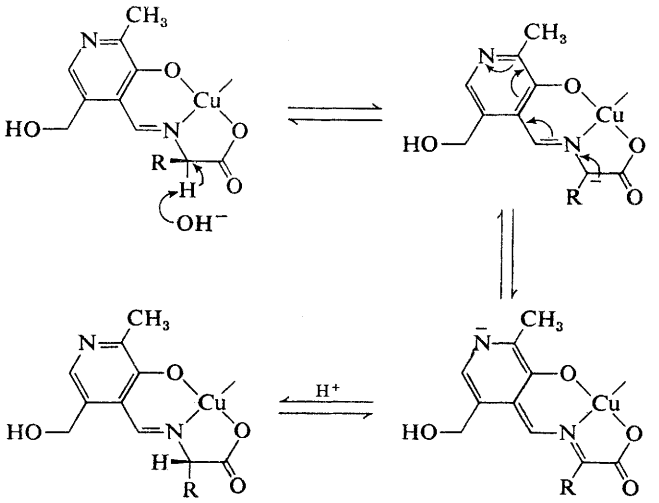
$$\text{pyridoxamine} + R^2CO\cdot CO_2H \rightleftharpoons R^2CHNH_2\cdot CO_2H + \text{pyridoxal} \quad (B)$$

$$R^1CHNH_2\cdot CO_2H + R^2CO\cdot CO_2H \rightleftharpoons R^1CO\cdot CO_2H + R^2CHNH_2\cdot CO_2H \quad (C)$$

point out that: (*a*) aldimine and ketimine complexes are formed sequentially in reactions (A) and (B), (*b*) the postulated tautomeric conversion of ketimine to aldimine in (C) unquestionably occurs, and (*c*) conversion is susceptible to direct detection by n.m.r.

[147] T. Ueki, T. Ashida, Y. Sasada, and M. Kakudo, *Acta Cryst.*, 1969, **B25**, 328.
[148] R. D. Gillard and D. A. Phipps, *Chem. Comm.*, 1970, 800.
[149] R. D. Gillard and R. Wootton, *J. Chem. Soc.* (*B*), 1970, 364.
[150] L. J. Theriot, G. O. Carlisle, and H. J. Hu, *J. Inorg. Nuclear Chem.*, 1969, **31**, 2891.
[151] L. J. Theriot, G. O. Carlisle, and H. J. Hu, *J. Inorg. Nuclear Chem.*, 1969, **31**, 2841.
[152] J. J. R. F. da Silva, R. Wootton, and R. D. Gillard, *J. Chem. Soc.* (*A*), 1970, 3369.
[153] M. E. Farago and T. Matthews, *J. Chem. Soc.* (*A*), 1970, 609.
[154] W. L. Felty, C. G. Ekstrom, and D. L. Leussing, *J. Amer. Chem. Soc.*, 1970, **92**, 3006.
[155] O. A. Gansow and R. H. Holm, *J. Amer. Chem. Soc.*, 1969, **91**, 573.
[156] O. A. Gansow and R. H. Holm, *J. Amer. Chem. Soc.*, 1969, **91**, 5984.

The racemisation of amino-acids in Schiff base complexes is thought [157] to occur by the mechanism shown in Scheme 8. Martell has also continued his work on the mechanism of the reactions promoted by zinc or aluminium ions, and using n.m.r. methods finds [158, 159] spectra closely similar to those reported by Holm's group. The interpretations placed on the results



Scheme 8

differ, however, particularly in the case of the aluminium complexes, where Holm's papers consider the spectra [155] to arise from a mixture of terdentate and bidentate modes of attachment of the Schiff bases, whereas Abbott and Martell postulate [159] terdentate co-ordination only, and the existence of isomeric complexes. The latter workers have also described in detail [160] their work on β-proton exchange in Schiff bases of α-amino-acids with vitamin B_6; in particular, the β-deuteriation of α-aminobutyric acid and of valine has been studied as a function of pD in the presence and absence of zinc or aluminium ions.

3 Peptides

A good deal of attention has focused on the usefulness of so-called chiroptical properties (c.d. and o.r.d.) in commenting on conformations and structures of peptide complexes of metal ions, particularly of copper(II).

[157] G. N. Weinstein, M. J. O'Connor, and R. H. Holm, *Inorg. Chem.*, 1970, **9**, 2104.
[158] E. H. Abbott and A. E. Martell, *J. Amer. Chem. Soc.*, 1969, **91**, 6866.
[159] E. H. Abbott and A. E. Martell, *J. Amer. Chem. Soc.*, 1970, **92**, 5845.
[160] E. H. Abbott and A. E. Martell, *J. Amer. Chem. Soc.*, 1970, **92**, 6931.

Hartzell and Gurd report [161] on the spectroscopic properties of the copper complexes of the *N*-terminal pentapeptide from sperm whale myoglobin, and on some other tetra- and penta-peptides. They show that the optical rotatory properties of the copper complex with L-Val-L-Leu-L-Ser-L-Glu-Gly can be distinguished qualitatively from those of complexes with histidine-containing peptides, *e.g.* the peptide comprising the first four residues of bovine serum albumin. Treptow has also studied [162] the c.d. of dipeptide and dipeptide amide complexes of copper. He concludes, from c.d. titrations, that the consumption of a third mole of base by the 1 : 1 complexes is probably due to the process $[LCu{-}OH_2] \rightarrow [LCu{-}OH]^-$, since the c.d. does not vary when this point is crossed. This paper and another [163] point out that the magnitude of the c.d. in copper complexes of dipeptides is an additive function of independent contributions from the *N*- and *C*-terminal residues. Similarly, the summed intensities of the c.d. of the complexes of L-Leu-Gly-Gly plus Gly-L-Leu-Gly plus Gly-Gly-L-Leu are equal to the intensity for the complex of L-Leu-L-Leu-L-Leu. The chiroptical properties of peptide complexes of copper and nickel in the u.v. have also been described.[164] Related work is that [165] on the structures of the 1 : 1 complexes of dipeptides with copper(II), in which the peptides act as terdentate ligands. From the o.r.d. spectra of the complexes of the diastereoisomers of Leu-Ala with copper(II), it was concluded [166] that the sign of the Cotton effect for the copper band depends on the asymmetric configuration in the *C*-terminal unit. It seems clear that, within the next year or two, some relationship between peptide structure and the chiroptical behaviour of copper complexes will emerge. The c.d. of palladium complexes of dipeptides has been studied.[167]

A detailed study by n.m.r. of the nickel and copper complexes of Gly-Gly, Gly-Gly-Gly, and of Gly-Gly-Gly-Gly reveals [168] that, as pH increases and each successive proton is removed from the complex, the resonance signals due to those protons nearest the site of proton dissociation shift to higher field. The method is applicable to deducing the order in which the various binding sites of the peptide become attached to the metal ion.

The changes in structure of the carnosine–copper complex, over the range pH 2—12, were followed [169] by the line-broadening effect of the paramagnetic copper(II) ion on the ligand's n.m.r. spectrum. The conclusions were: (*a*) that in acid solution, the only binding site is a carboxy-group; (*b*) in neutral solution the dominant complex involves only the

[161] C. R. Hartzell and F. R. N. Gurd, *J. Biol. Chem.*, 1969, **244**, 147.
[162] R. S. Treptow, *J. Inorg. Nuclear Chem.*, 1969, **31**, 2983.
[163] J. M. Tsangaris and R. B. Martin, *J. Amer. Chem. Soc.*, 1970, **92**, 4255.
[164] J. M. Tsangaris, J. W. Chang, and R. B. Martin, *J. Amer. Chem. Soc.*, 1969, **91**, 726.
[165] F. Karezynski and G. Kupryszewski, *Roczniki Chem.*, 1969, **43**, 1317.
[166] F. Karezynski, J. Szafranek, and G. Kupryszewski, *Zeszyty Nauk., Mat., Fiz., Chem.*, 1969, **9**, 131.
[167] E. W. Wilson and R. B. Martin, *Inorg. Chem.*, 1970, **9**, 528.
[168] M. K. Kim and A. E. Martell, *J. Amer. Chem. Soc.*, 1969, **91**, 872.
[169] M. Ihnat and R. Bersohm, *Biochemistry*, 1970, **9**, 4555.

3-nitrogen of imidazole(!), and (*c*) at alkaline pH there is a mixture of the complex with N-3 binding and another in which there is three-point binding to O(carboxy-group), N(amino-group), and N(peptide) atoms.

Bacitracin is a dodecapeptide antibiotic produced from *Bacillus licheniformis*, for which divalent metal ions seem to have a stabilizing and activating effect. The peptide binds [170] zinc(II) to a molar ratio of 1 : 1. O.r.d. establishes [171] that the thiazoline ring of bacitracin provides a binding site for zinc. A histidine unit is also implicated by potentiometric and n.m.r. spectra.

An *X*-ray structure for the disodium salt of a copper(II) complex anion of the pentapeptide tetraglycylglycine is reported.[172] Crystallised from alkaline solution, the peptide in the complex is attached to copper *via* four nitrogen atoms. The structures have been reported [173] of three new salts of the known bis(glycylglycinato)cobaltate(III) anion and, more interestingly, of the product of its protonation, in which the proton evidently is attached to the oxygen atom of the peptide, *i.e.* $H_2N \cdot CH_2 \cdot C(OH):N \cdot CH_2 \cdot COO^-$, rather than to the peptide nitrogen.

Österberg has described [174] a mixed oxidation-state complex of Gly-Gly-Gly with copper(II) and copper(I). For the electrode reaction below, $E^0 = +340$ mV at pH 7.0. This value is of the order of magnitude re-

$$[Cu^{II}{}_2H_{-4}A_2]^{2-} + e^- + H^+ \rightleftharpoons [Cu^{I}Cu^{II}H_{-3}A_2]^{2-}$$

ported for many copper proteins.

Stability constants have been measured [175] (as a function of temperature) for glycylglycine with zinc. Log K_2 fell from 7.06 at 10 °C to 6.06 at 40 °C. The stabilities of the complexes formed by the dipeptides from glycine and/or β-alanine with copper(II) have been determined potentiometrically,[176] with the finding that the tendency to form the deprotonated chelate, CuA, decreases in the order Gly-Gly > β-Ala-Gly > Gly-β-Ala > β-Ala-β-Ala.

Physico-chemical properties of the poly-L-histidine complex of copper(II) have been reported.[177] The poly-L-histidine was of molecular weight 16 000, and the complex formed at pH 4.5—6.0 had one copper(II) ion per 4—20 histidyl residues. At pH 4.6, four protons were released per copper(II) ion bound. The authors describe some interesting resemblances between this complex and haemocyanin. Previously they had reported [178] on the

[170] L. C. Craig, W. F. Phillips, and M. Burachik, *Biochemistry*, 1969, **8**, 2348.
[171] N. W. Cornell and D. G. Guiney, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 530.
[172] J. F. Blount, H. C. Freeman, R. V. Holland, and G. H. W. Milburn, *J. Biol. Chem.*, 1970, **245**, 5177.
[173] M. T. Barnet, H. C. Freeman, D. A. Buckingham, I. N. Hsu, and D. van der Helm, *Chem. Comm.*, 1970, 367.
[174] R. Österberg, *European J. Biochem.*, 1970, **13**, 493.
[175] S. Pelletier, *Compt rend.*, 1969, **268***C*, 2248.
[176] O. Yamauchi, Y. Hirano, Y. Nakao, and A. Nakahara, *Canad. J. Chem.*, 1969, **47**, 3442.
[177] A. Levitzki, I. Pecht, and A. Berger, *Proc. 11th Internat. Conf. Co-ord. Chem.*, Israel, 1968, ed. M. Cais, Elsevier, New York.
[178] I. Pecht, A. Levitzki, and M. Anbar, *J. Amer. Chem. Soc.*, 1967, **89**, 1587.

high catalytic efficiency of this complex (and its small substrate specificity) in autoxidation reactions. Following this work, and the more general statement [179] that many polymeric metal complexes have higher catalytic efficiencies than their low-molecular-weight analogues, the asymmetrically selective oxidation of Dopa was observed in the presence of a poly-L-lysine–copper(II) complex.[180] The D-isomer of Dopa was oxidised rather faster than L-Dopa.

Oxygenation of cobalt(II) solutions with glycylglycine [181] and other peptides [182] has been described. The conversion in such systems of some glycyl-L-histidine ligand to glycylkynurenic acid has been discussed in detail.

The specific hydrolyses of co-ordinated peptides, discussed in our last Report,[183] have been the subject of a good deal more work. The hydrolyses of chelated glycine amides, glycylglycine, and esters of glycylglycine have been investigated kinetically [184] and a mechanism proposed, in which solvent hydroxide attacks at the co-ordinated carboxy-groups of the chelated glycine derivatives. Hay and Morris find[185] that in β-[Co(trien)(Gly-Gly-OMe)]$^{3+}$ the base hydrolysis of the peptide bond in the complex is 6.5×10^4 times faster than in free Gly-Gly-OMe.

Collman has extended [186] his original observation with Buckingham on the *N*-terminal specific hydrolysis of dipeptides by β-[Co(trien)(H_2O)(OH)]$^{2+}$ to the use of the similar cobalt(III) complex with 2,2′,2″-triaminotriethylamine (tren). The complex *cis*-[Co(tren)(H_2O)(OH)]$^{2+}$ also specifically removes the *N*-terminal residue from dipeptides and tripeptides, although more slowly than the original reagent.

Several reports have appeared from Pasternack and his group [187] on the rates of complex formation of di-, tri, and tetra-peptides with the metal ions Co^{II}, Ni^{II}, and Cu^{II}.

4 Proteins

A novel suggestion this year lies in the use of some metal ions, which are not currently thought to be biologically significant, as probes for binding sites and mechanisms of action of metal-potentiated enzymes. The idea underlying the application is that whereas such metals as potassium and

[179] H. Sigel, *Angew. Chem.*, 1969, **81**, 161.
[180] M. Hatano, T. Nozawa, and M. Yoneyama, *Bull. Chem. Soc. Japan*, 1970, **43**, 295.
[181] E. D. McKenzie, *J. Chem. Soc.* (*A*), 1969, 1655.
[182] R. D. Gillard and A. Spencer, *J. Chem. Soc.* (*A*), 1969, 1655.
[183] Ref. 1, p. 280.
[184] D. A. Buckingham, C. E. Davis, D. M. Foster, and A. M. Sargeson, *J. Amer. Chem. Soc.*, 1970, **92**, 5571.
[185] R. W. Hay and P. J. Morris, *Chem. Comm.*, 1969, 1208.
[186] E. Kimura, S. Young, and J. P. Collman, *Inorg. Chem.*, 1970, **9**, 1183.
[187] K. Kustin and R. F. Pasternack, *J. Phys. Chem.*, 1969, **73**, 1; G. Davies, K. Kustin, and R. F. Pasternack, *Inorg. Chem.*, 1969, **8**, 1535; R. F. Pasternack, K. Kustin, L. A. Hughes, and E. Gibbs, *J. Amer. Chem. Soc.*, 1969, **91**, 4401; R. F. Pasternack, M. Angwin, and E. Gibbs, *ibid.*, 1970, **92**, 5878.

calcium are distinctly lacking in properties which can be used to comment on their environment, there are other metal ions, with closely related charge–size ratios, which have useful observable properties. Thus, thallium(I) has been studied [188] as a possible replacement for potassium ion, and has been found to be around ten times more effective than potassium in activating pyruvate kinase and the vitamin-B_{12}-dependent diol dehydratase. Similar results have emerged [189] in other biochemical systems.

In the same way, the rare-earth ions may substitute for calcium; the binding of neodymium (radius of Nd^{3+} = 0.995 Å; Ca^{2+} = 0.990 Å) to bovine serum albumin has been studied.[190]

Binding-site studies continue to dominate the literature. Fleet and Rechnitz describe [191] an elegant application of liquid membrane ion-selective electrodes as sensors in a rapid-mixing continuous flow system, which has enabled them to measure rates of complexing of ions such as calcium and magnesium with bio-ligands. The use of c.d. as a means of studying binding-sites of proteins for copper(II) and nickel(II) has been investigated.[192] The results suggest that the first equivalents of both Cu^{II} and Ni^{II} bind to the amino-terminal ends of the proteins studied; subsequent binding in the case of bovine serum albumin appears to be to sulphydryl groups.

Those methods of studying conformational changes in proteins which have been applied in examining the effects of metal ions include hydrogen–tritium exchange and n.m.r. spectroscopy. Hydrogen–tritium exchange data have been published for conalbumin (ovotransferrin) in the presence of metal ions. Emery finds [193] a retardation of the exchange by Cu^{II} or Fe^{III}; the latter finding is in agreement with those [194] of Ulmer. The proton exchange of lysozyme, serum albumin, and haemoglobin was not affected [193] by iron. The proton resonance spectrum of lysozyme has been assigned [195] on the basis of the perturbation produced by introducing (paramagnetic) cobalt(II) ions. Specific divalent metal ions, Mn^{2+} or Mg^{2+}, are known to be directly involved both in the catalytic function and in the stabilisation of the quaternary structure of glutamine synthetase. A n.m.r. study of the manganese(II) – glutamine synthetase (from *Escherichia coli*) reveals [196] the large conformational change of the protein on metal binding, and three independent binding sites are also apparent.

[188] J. P. Manners, K. G. Morallee, and R. J. P. Williams, *Chem. Comm.*, 1970, 965.
[189] C. E. Inturrisi, *Biochim. Biophys. Acta*, 1969, **173**, 567; F. J. Kayne and J. Reuben, *J. Amer. Chem. Soc.*, 1970, **92**, 214.
[190] E. R. Birnbaum, J. E. Gomez, and D. W. Darnall, *J. Amer. Chem. Soc.*, 1970, **92**, 5287.
[191] B. Fleet and G. Rechnitz, *Analyt. Chem.*, 1970, **42**, 691.
[192] J. M. Tsangaris, J. W. Chang, and R. B. Martin, *Arch. Biochem. Biophys.*, 1969, **103**, 53.
[193] T. F. Emery, *Biochemistry*, 1969, **8**, 877.
[194] D. D. Ulmer, *Biochim. Biophys. Acta*, 1969, **181**, 305.
[195] C. C. McDonald and W. D. Phillips, *Biochem. Biophys. Res. Comm.*, 1969, **35**, 43.
[196] M. D. Denton and A. Ginsburg, *Biochemistry*, 1969, **8**, 1714.

Metal requirements and potentiation of a wide range of enzymes continue to be discovered and studied. A particularly interesting observation is that [197] the nature and quantity of acid-soluble proteins separated from rat-liver nucleoprotein are greatly dependent upon the quantity of divalent metal ion present. It is thought that this arises because Mg^{2+} and Mn^{2+} play an important rôle in maintaining the structural integrity of arginine-rich nucleoproteins in the nuclei. The general rôle of bivalent cations in the phosphoglucomutase system has been studied;[198] a more detailed examination [199] reveals a 1 : 1 metal : protein binding ratio with negligible conformational change on metal binding. It is concluded [200] that the size of a crevice in the phosphoglucomutase–substrate complex can be altered through binding a metal ion. Binding of manganese to pyruvate carboxylase,[201] phosphoglucomutase,[202] and β-methylaspartase [203] has been investigated spectroscopically. Alger has distinguished [204] between inactivation of ribonuclease A and activation of substrate RNA (but not of substrate cyclic 2′,3′-cytidine monophosphate). The inactivation of ribonuclease by divalent metal ions has been attributed [205] to the binding of the metal ions to the essential histidine-12 and -119 residues of the protein.

Williamson reports [206] the interesting result that copper(II) ions accelerate the rate of formation of mixed disulphides between cystine and β-globulins; this discovery arose from the realisation that the inhibition by glutamate of the mixed disulphide formation was due to complexing of copper by the glutamate. In a related context, it was found [207] that copper functions as a catalyst in the formation of the intramolecular bonds in collagen.

5 Metalloproteins

Eichhorn has summarised [208] some of the factors which distinguish metal ion catalysis in small 'model' reactions from those of the larger biochemical macromolecules. The complexity of the ligand affects the metal catalysis in several important ways:

(i) The action of the metal ion bound at one point in the macromolecule may be aided by other groups in the macromolecule, so far as its catalytic action is concerned. For example, the action of zinc in carboxypeptidase is

[197] H. P. von Hahn, J. E. Heim, and G. L. Eichhorn, *Biochim. Biophys. Acta*, 1970, **214**, 509.
[198] W. J. Ray, *J. Biol. Chem.*, 1969, **244**, 3740.
[199] E. J. Peck and W. J. Ray, *J. Biol. Chem.*, 1969, **244**, 3748.
[200] E. J. Peck and W. J. Ray, *J. Biol. Chem.*, 1969, **244**, 3754.
[201] M. C. Scrutton and A. S. Mildvan, *Arch. Biochem. Biophys.*, 1970, **140**, 131.
[202] W. J. Ray and A. S. Mildvan, *Biochemistry*, 1970, **9**, 3886.
[203] G. A. Fields and H. J. Bright, *Biochemistry*, 1970, **9**, 3801.
[204] T. D. Alger, *Biochemistry*, 1970, **9**, 3248.
[205] G. L. Eichhorn, P. Clark, and E. Tarien, *J. Biol. Chem.*, 1969, **244**, 937.
[206] M. B. Williamson, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 379.
[207] W. S. Chou, J. E. Savage, and B. L. Odell, *J. Biol. Chem.*, 1969, **244**, 5785.
[208] G. L. Eichhorn, *Proc. 11th Internat. Conf. Co-ord. Chem.*, Israel, 1968, ed. M. Cais, Elsevier, New York.

helped by the action of a carboxy-group and of a tyrosine hydroxy-group on the peptide link of the substrate.

(ii) Properties of the co-ordination sphere may be modified by changes in the macromolecular surroundings. As an example is quoted the hydrophobic environment of the iron(II) porphyrin in haemoglobin which enables di-oxygen to be added without oxidation of the iron(II).

(iii) The metal may have a number of possible co-ordination sites in a macromolecule and can thereby induce a variety of ligand reactions in it. As an example, the effect of zinc(II) ions on the nucleic acid structure is discussed, where, if bound to phosphate in the backbone it promotes cleavage of phosphate diester bonds, but if bound to heterocyclic bases it causes helix $\leftrightarrow$ random-coil transitions.

A. Non-Haem Metalloproteins.—The bulk of work in this area is concerned with the characterization of metalloproteins and with binding-site studies. Vallee has reported a number of detailed studies on zinc proteins, including[209] the isolation of an α_2-macroglobulin from human serum which contains firmly bound zinc. The catalytic properties of liver alcohol dehydrogenase are found[210] to be retained on substituting the naturally occurring zinc(II) with cobalt(II) or cadmium(II). A more detailed study[211] of liver alcohol dehydrogenase in fact shows that two chemically distinct groups of zinc ions are present. The kinetics of the carboxypeptidase A enzyme in its zinc, cobalt, and manganese forms are described.[212] A fructose diphosphate aldolase has been obtained[213] in a stable and homogeneous form, with bound zinc. Alkaline phosphatase (from *Escherichia coli*) has been found[214] to alter conformation on removal of the bound zinc ions; the zinc can be replaced by other metal ions[215] but only the cobalt(II) derivative shows appreciable catalytic activity; the copper(II)- and manganese(II)-substituted enzymes display weak activity. The presence of zinc as well as copper in human cytocuprein, formerly classed as a copper-protein, has been demonstrated.[216] Chlorine-35 n.m.r. studies have been reported on the cobalt(II) form of carbonic anhydrase,[217] and on the native zinc enzyme.[218]

[209] A. F. Parisi and B. L. Vallee, *Biochemistry*, 1970, **9**, 2421.
[210] D. E. Drum and B. L. Vallee, *Biochem. Biophys. Res. Comm.*, 1970, **41**, 33.
[211] D. E. Drum and B. L. Vallee, *Biochemistry*, 1970, **9**, 4078.
[212] D. S. Auld and B. L. Vallee, *Biochemistry*, 1970, **9**, 4353.
[213] R. D. Kobes, R. T. Simpson, B. L. Vallee, and W. J. Rutter, *Biochemistry*, 1969, **8**, 585.
[214] J. A. Reynolds and M. J. Schlesinger, *Biochemistry*, 1969, **8**, 588; H. Csopak, *European J. Biochem.*, 1969, **7**, 186.
[215] H. Csopak and K. E. Falk, *F.E.B.S. Letters*, 1970, **7**, 147; C. Petitclerc, C. Lazdunski, D. Chappelet, A. Moulin, and M. Lazdunski, *European J. Biochem.*, 1970, **14**, 301; D. Chappelet, C. Lazdunski, C. Petitclerc, and M. Lazdunski, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 91.
[216] R. J. Carrico and H. F. Deutsch, *J. Biol. Chem.*, 1970, **245**, 723.
[217] R. L. Ward and K. J. Fritz, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 207.
[218] R. L. Ward, *Biochemistry*, 1970, **9**, 244.

The mechanism whereby cadmium and zinc alter [219] copper metabolism has been studied [220] in detail; copper and zinc effectively compete for the sulphydryl binding sites of a metallothionein-like protein. The presence of bound zinc and cobalt has been demonstrated [221] in transcarboxylase, a biotin-containing enzyme which catalyses the reversible carboxylation of pyruvate by methylmalonyl-CoA.

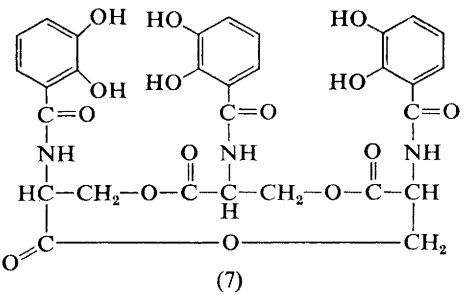
Copper proteins have been studied in many laboratories, primarily from the point of view of the nature of the binding sites. Of the many reports, we select only a few which have a more mechanistic implication. The copper states of porcine ceruloplasmin change [222] upon addition of urea; this is attributed to a change in the protein structure causing a change in the environment of the copper ions. An earlier report, that up to 50% of the copper in cytochrome *c* oxidase may be removed and then re-introduced, has been disproved; the treatment with 2,9-dimethyl-4,7-diphenyl-1,10-phenanthroline actually removes [223] very little copper from the enzyme. The loss of enzyme activity is actually connected [224] with disaggregation caused by the chelating agent.

In the area of molybdenum-containing enzymes, the crystallisation of the nitrogenase from *Azotobacter vinelandii* is described;[225] it gives white needle-like crystals, with a ratio of Mo : Fe : cysteine : labile sulphur about 1 : 20 : 20 : 15 in a molecular weight of 270 000—300 000. The crystallised protein, in conjunction with the iron protein of the nitrogenase system, is essential to nitrogenase reactions such as the reduction of N_2,C_2H_2,CN^-, acrylonitrile, and azide, and to ATP-dependent hydrogen evolution. Schrauzer and Schlesinger use [226] the reduction of acetylene to ethylene (a property of the nitrogenase systems of *Clostridium pasteurianum* and *Azotobacter vinelandii*) to test model systems. Maximum activity was observed with a 1 : 1 molybdenum : thiol (*e.g.* mercaptoethanol, cysteine) ratio. Iridium was the only other metal which showed much activity, *viz.* 15% of that for molybdenum.

Non-haem iron proteins have again attracted a great deal of attention. The ferredoxins, in particular, form the subject of many investigations. Among the papers which do not have as their primary concern the elucidation of binding-sites in intact proteins, there are several which contain novel features. Neilands, in continuing his studies on iron transport, has

[219] B. C. Starcher, *J. Nutrition*, 1969, **97**, 321; P. D. Wanger and P. H. Weswig, *ibid.*, 1970, **100**, 341.
[220] G. W. Evans, P. F. Majors, and W. E. Cornatyer, *Biochem. Biophys. Res. Comm.*, 1970, **40**, 1142.
[221] D. B. Northrop, and H. G. Wood, *J. Biol. Chem.*, 1969, **244**, 5801.
[222] H. Mukasa, Y. Nosoh, and T. Sato, *J. Biochem.*, 1969, **65**, 649.
[223] H. Beinert, C. R. Hartzell, B. F. van Gelder, K. Ganapathy, H. S. Mason, and D. C. Wharton, *J. Biol. Chem.*, 1970, **245**, 225.
[224] H. S. Mason and K. Ganapathy, *J. Biol. Chem.*, 1970, **245**, 231.
[225] R. C. Burns, R. D. Holsten, and R. W. F. Hardy, *Biochem. Biophys. Res. Comm.*, 1970, **39**, 90.
[226] G. N. Schrauzer and G. Schlesinger, *J. Amer. Chem. Soc.*, 1970, **92**, 1808.

isolated [227] from *Salmonella typhimurium* LT2 a new iron-complexing agent, which he calls enterobactin, with the structure (7). This complexes with iron(III) through the *ortho*-diphenolic residues to give an anionic six-co-ordinate complex.



(7)

Churchill and Wormald have established [228] the crystal structure of bis(imidotetramethyldithiodiphosphino-*SS*)iron(II), which they use as a structural model for the iron–sulphur core of rubredoxin, since the iron atom is tetrahedrally co-ordinated by four sulphur atoms in both molecules. In the same field, a ternary complex formed [229] by mixing ferric chloride with mercaptoethanol and sodium sulphide has spectroscopic properties resembling those of non-haem iron proteins. What is described as 'an artificial non-haem iron protein' has been made [230] from serum albumin and ribonuclease. Its spectrum is said to be like that of spinach ferredoxin.

B. Haem Proteins.—(See also Chapter 2, Part II, Section 3B.) Among the large volume of work published on the biophysics of haem proteins, there are a number of papers of direct chemical significance. Perutz has discussed [231] the stereochemistry of co-operative effects in haemoglobin, based on his exhaustive diffraction studies. He concludes that 'the essence of the mechanism may be summarized as follows. The haem group is so constructed that it amplifies a small change in atomic radius, undergone by the iron atom in the transition from the high-spin to the low-spin state, into a larger movement of the haem-linked histidine relative to the porphyrin ring. This movement, and the widening of the haem pocket required in the β-chain, triggers small changes in the tertiary structure of the reacting sub-units, which includes a movement of helix F [for the meaning of these terms, see the review [13] by Perutz] towards helix H and the centre of the molecule. The narrowing of the pocket between helices F and H results in the expulsion of the penultimate tyrosine residues. The expelled tyrosine

[227] J. R. Pollack and J. B. Neilands, *Biochem. Biophys. Res. Comm.*, 1970, **38**, 989.
[228] M. R. Churchill and J. R. Wormald, *Chem. Comm.*, 1970, 703.
[229] C. S. Yang and F. M. Huennekens, *Biochem. Biophys. Res. Comm.*, 1969, **35**, 634; *Biochemistry*, 1970, **9**, 2127.
[230] V. E. Bayer, P. Krauss, and P. Schretzmann, *Z. physiol. Chem.*, 1969, **350**, 994.
[231] M. F. Perutz, *Nature*, 1970, **228**, 726.

pulls the *C*-terminal residue with it, rupturing the salt bridges that had held the reacting sub-units to its neighbours in the deoxy tetramer. The rupture of each salt bridge removes one of the constraints holding the molecule in the deoxy conformation and tips the equilibrium between the two alternative quaternary structures some way in favour of the oxy conformation. In this conformation, the oxygen affinity is raised because the sub-units are no longer constrained by the salt-bridges to maintain the tertiary deoxy structure.' This mechanism also leads to the liberation of the Bohr protons, in agreement with Gray's recently published conclusion [232] that their release is exactly proportional to the amount of oxygen taken up, and that the two processes are synchronous. N.m.r. and e.s.r. spectroscopic studies had already [233] suggested that no changes in the haem groups themselves occur when a neighbouring haem is oxygenated, and that co-operativity between haem groups is therefore explicable in terms of changes at or near the interfaces of the subunits.

Kilmartin and Wootton have demonstrated [234] (in support of the ideas of Perutz [231, 235]) that des-His$^{146\beta}$-human haemoglobin, made by carboxypeptidase *B* digestion, lacks half the alkaline Bohr effect.

Gersonde has studied [236] the existence and nature of complexes between native haemoglobin(III) and dodecyl sulphate. The reaction of nitric oxide with methaemoglobin has been studied [237] using absorption spectra, magnetic susceptibilities, and e.s.r. spectra. A possible reaction mechanism is suggested {based on the known hydrolysis of nitroprusside to $[Fe(CN)_5(NO_2)]^{4-}$}:

$$(\text{Haem-Fe})^{3+} + \text{NO} \longrightarrow (\text{Haem-FeNO})^{3+}$$

$$(\text{Haem-FeNO})^{3+} + 2\text{OH}^- \longrightarrow (\text{Haem-FeNO}_2)^{+} + \text{H}_2\text{O}$$

$$(\text{Haem-FeNO}_2)^{+} + \text{NO} \longrightarrow (\text{Haem-FeNO})^{2+} + \text{NO}_2^-$$

pH-Dependence of conformations of haemoglobin(III) has been established; distinct e.s.r. signals for the conformers are observed.[238]

An effective reagent [239] for reducing cytochrome *c* in acid solution is made up of ferrous ions and ethylenediaminetetra-acetic acid (at pH 0.7—5.5). Ferrous ion alone does not suffice under these conditions. E.s.r. spectra show [240] that cytochrome P450 undergoes substrate-induced spin

[232] R. D. Gray, *J. Biol. Chem.*, 1970, **245**, 2914.
[233] R. G. Shulman, S. Ogawa, K. Wüthrich, T. Yamane, J. Peisach, and W. E. Blumberg, *Science*, 1969, **165**, 251.
[234] J. V. Kilmartin and J. F. Wootton, *Nature*, 1970, **228**, 766.
[235] M. F. Perutz, H. Muirhead, L. Mazzarella, R. A. Crowther, J. Greer, and J. V. Kilmartin, *Nature*, 1969, **222**, 1240.
[236] K. Gersonde, *J. Mol. Biol.*, 1969, **42**, 285.
[237] J. C. W. Chien, *J. Amer. Chem. Soc.*, 1969, **91**, 2166.
[238] K. Gersonde and A. Wollmer, *European J. Biochem.*, 1970, **15**, 226; K. Gersonde, H. Sick, and A. Wollmer, *ibid.*, p. 237.
[239] M. Kurihara and S. Sano, *J. Biol. Chem.*, 1970, **245**, 4804.
[240] J. A. Whysner, J. Ramseyer, G. M. Kazni, and B. W. Harding, *Biochem. Biophys. Res. Comm.*, 1969, **36**, 795.

changes. N.m.r. measurements at 220 MHz of nine different species of mammalian-type ferrocytochrome *c* provide evidence [241] for identical methionine–iron co-ordination in all the species studied, thus disproving that either histidine or lysine residues occupy the sixth co-ordination position of the iron atom.

Workers in Russian laboratories have studied model systems for catalase, peroxidase, and oxidase activity. The bis(2,2′-bipyridyl)copper(II) complex (a catalase model) has been shown [242] to form dimers readily. In a related study, a number of complexes of cobalt(II) have been shown [243] to be catalytically active in the oxidation of pyrogallol with hydrogen peroxide and in the oxidation of ascorbic acid with di-oxygen. The authors consider the mechanisms to involve peroxo-derivatives of the complexes, in line with a number of suggestions in the literature.

Finally, it has been thought possible for some time that peripheral reaction of the haem and related tetra-nitrogenous ligands might account for some features of enzymatic oxidations and reductions. A report germane to this question is that [244] of Castro and Davis, who present evidence for peripheral attack upon the porphyrin ring during the chemical reduction and oxidation of chloro-iron(III) octaethylporphyrin.

[241] C. C. McDonald, W. D. Phillips, and S. N. Vinogradov, *Biochem. Biophys. Res. Comm.*, 1969, **36**, 442.
[242] Y. L. Skurlatov and A. D. Purmal, *Russ. J. Phys. Chem.*, 1969, **43**, 880.
[243] G. A. Shagisultanova and N. P. Glukhova, *Russ. J. Phys. Chem.*, 1969, **43**, 890.
[244] C. E. Castro and H. F. Davis, *J. Amer. Chem. Soc.*, 1969, **91**, 5405.

Author Index