The Mass Spectra of Amino-acid and Peptide Derivatives

By J. H. Jones
DYSON PERRINS LABORATORY, OXFORD UNIVERSITY

1 Introduction

The various types of absorption spectroscopy which are such powerful tools in most areas of organic chemistry find limited application in peptide chemistry; these methods reveal the main structural features but give no information about the amino-acid sequences of proteins, although they are used in a semi-empirical manner to procure data about secondary structure. In the elucidation of the primary structure of a protein,[1] it is necessary to determine the amino-acid sequences of a large number of small peptides which are obtained by partial degradation. Although chemical methods for the determination of oligopeptide amino-acid sequences have attained a very high level of sophistication (*e.g.*, the recent introduction of repeated Edman degradation of resin-bound peptides[2]), they are time-consuming and tedious. There has therefore been considerable interest during the last few years in the behaviour of amino-acid and peptide derivatives on electron impact, since mass spectrometry provides a rapid means, requiring very small amounts of material (0·2 μg. is sufficient[3]), of discovering the amino-acid sequences of oligopeptides.

For discussion of the principles of instrumentation and of the fragmentation of ionised organic molecules, the reader is directed to a previous Quarterly Review[4] and standard texts.[5]

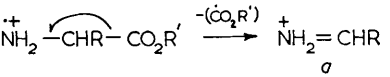
In the compilation of this Review, papers published up to the end of 1967 have been considered.

2 Amino-acids and their Derivatives

A. Free Amino-acids.—The zwitterionic nature of free amino-acids renders them very involatile, and thermal reactions can occur at the high temperatures required for vaporisation. However, this difficulty has been partly overcome by

[1] E. Y. Spencer, ch. 15 in 'Techniques of Organic Chemistry', vol. XI, part II, ed. K. W. Bentley, Interscience, New York, 1963.
[2] R. A. Laursen, *J. Amer. Chem. Soc.*, 1966, **88**, 5344.
[3] M. Barber, P. Powers, M. J. Wallington, and W. A. Wolstenholme, *Nature*, 1966, **212**, 784.
[4] R. I. Reed, *Quart. Rev.*, 1966, **20**, 527.
[5] (*a*) K. Biemann, 'Mass Spectrometry', McGraw-Hill, New York, 1962; (*b*) H. C. Hill, 'Introduction to Mass Spectrometry', Heyden and Son Ltd., London, 1966; (*c*) F. W. McLafferty, 'Interpretation of Mass Spectra', W. A. Benjamin, Inc., New York, 1966; (*d*) H. Budzikiewicz, C. Djerassi, and D. H. Williams, 'Mass Spectrometry of Organic Compounds', Holden-Day, Inc., San Francisco, 1967; (*e*) J. H. Beynon, 'Mass Spectrometry and its Applications to Organic Chemistry', Elsevier, Amsterdam, 1960.

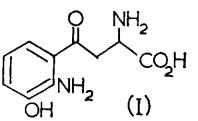
improved techniques for the direct introduction of samples into the ion source.[6] The most important fragmentation is the formation of 'amine fragments' (*a*) by loss of the carboxyl group,[6, 7] after primary ionisation at the nitrogen atom[6] (Scheme 1, R′ = H). Side-chain cleavages are discussed in a separate section.



Scheme 1

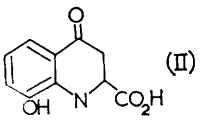
Mass spectrometry can be used for quantitative analysis of amino-acid mixtures:[8] the method depends on comparison of the relative intensities of peaks peculiar to each component. However, if any particularly involatile amino-acids (*e.g.*, tryptophan) are present, the contribution of these constituents to the spectrum is too small for accurate relative intensity measurement. Qualitative analysis of amino-acid mixtures can also be performed by using mass spectrometry to identify characteristic decomposition products after separation by gas–liquid chromatography (g.l.c.) of the products of controlled pyrolysis.[9] Time-of-flight mass spectrometers[10] have been favoured in work with free amino-acids[11, 12] and other biologically important compounds such as nucleosides,[13] and the fact that these instruments can be made in a compact form suitable for rocket transport gives a possible method for detecting life-characteristic molecules in extraterrestrial environments.

Comparison of the mass spectra of ordinary leucine and leucine recovered after equilibration with pyridoxal–alum–D_2O[14] revealed a deuteriation pattern which was consistent with the generally accepted[15] Schiff-base mechanism for pyridoxal-catalysed transamination. Another example, illustrating one of the pitfalls of mass spectrometry, is the case of the yellow pigment (I) isolated from South American butterflies. The ion of highest mass-to-charge ratio (*m/e*) in the spectrum was erroneously identified as the molecular ion, and the structure (II) was assigned.[16] It was later found that (I) was converted into (II) by cyclisation with loss of ammonia in the heated inlet system.[17]



(I)

[6] G. Junk and H. Svec, *J. Amer. Chem. Soc.*, 1963, **85**, 839.
[7] K. Heyns and H. F. Grützmacher, *Annalen*, 1963, **667**, 194.
[8] G. A. Junk and H. J. Svec, *Analyt. Chim. Acta*, 1963, **28**, 164.
[9] C. Merritt and D. H. Robertson, *J. Gas Chromatog.*, 1966, **5**, 96.
[10] D. B. Harrington, 'Advances in Mass Spectrometry', ed. J. D. Waldron, Pergamon Press, Oxford, 1959, p. 249.
[11] K. Biemann and J. A. McCloskey, *J. Amer. Chem. Soc.*, 1962, **84**, 3192.
[12] N. Martin, NASA report No. CR-68768 (*Chem. Abs.*, 1967, **66**, 1587).
[13] K. Biemann and J. A. McCloskey, *J. Amer. Chem. Soc.*, 1962, **84**, 2005.
[14] G. A. Junk and H. J. Svec, *J. Org. Chem.*, 1964, **29**, 944.
[15] A. E. Braunstein, 'The Enzymes', vol 2, ed. P. D. Boyer, H. Lardy, and K. Mrybäck, Academic Press, New York, 1960, p. 137.
[16] K. S. Brown, *J. Amer. Chem. Soc.*, 1965, **87**, 4202.
[17] K. S. Brown and D. Becker, *Tetrahedron Letters*, 1967, 1721.

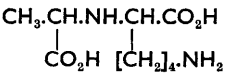


(II)

B. Amino-acid Alkyl Esters.—Because of their greater volatility, α-amino-esters[18, 19] are more suitable for mass spectrometric investigation than the free acids. The fragmentation mechanisms of these compounds have been discussed at length in a well-known treatise[20] and it suffices here to say that the most intense peaks are usually the amine fragments (*a*).

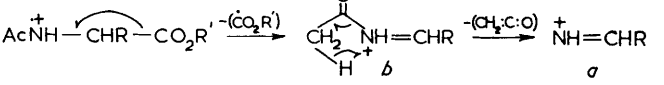
Mixtures of amino-acid esters can be analysed quantitatively[21] by use of the principle outlined above for mixtures of the free acids. Such procedures do not compare favourably with orthodox methods based on ion-exchange chromatography and colorimetric estimation,[22] but may find use when only very limited amounts of material are available.

Mass spectrometry has been used to distinguish between possible structures for the γ-hydroxy-α-amino-acids obtained by degradation of toxins from *Amanita phalloides*. The acids were converted into the corresponding lactone hydrochlorides, which underwent sufficient thermal dissociation at the source temperature for the spectra of the free bases to be observed.[23] Another example is the determination of the structure (III) of lysopine (an amino-acid isolated in very small amount from crown gall tissue) from the mass spectrum of its ethyl ester.[24]



(III)

C. *N*-Protected Amino-acids and their Derivatives.—The mass spectra of *N*-acetylamino-acids[7] and their alkyl esters[25] all show an intense acetyl ion (*m/e* 43) and the chief common feature is loss of the carboxyl (or alkoxycarbonyl) group to give an acyliminium ion (*b*) which then ejects keten with formation of an amine fragment (*a*) (Scheme 2). In addition, side-chain cleavages characteristic of the amino-acid concerned are observed (see p. 309).



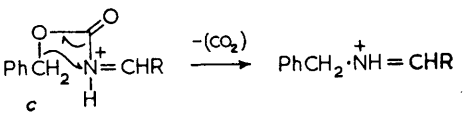
Scheme 2

[18] K. Biemann, J. Siebl, and F. Gapp, *J. Amer. Chem. Soc.*, 1961, **83**, 3795.
[19] C.-O. Andersson, R. Ryhage, S. Stalberg-Stenhagen, and E. Stenhagen, *Arkiv Kemi*, 1962, **19**, 405.
[20] Ref. 5 (*a*), p. 260.
[21] K. Biemann and W. Vetter, *Biochem. Biophys. Res. Comm.*, 1960, **2**, 93.
[22] D. H. Spackman, W. H. Stein, and S. Moore, *Analyt. Chem.*, 1958, **30**, 1190.
[23] P. Pfaender and T. Wieland, *Annalen*, 1966, **700**, 126.
[24] K. Biemann, G. G. J. Deffner, and F. C. Steward, *Nature*, 1961, **191**, 380.
[25] C.-O. Andersson, R. Ryhage, and E. Stenhagen, *Arkiv Kemi*, 1962, **19**, 417.

N-Trifluoroacetylamino-acid methyl esters are suitable derivatives for the separation of amino-acid mixtures by g.l.c.,[26] and the constituents of the mixture can be identified by passing the eluant from the g.l.c. column directly into the ion source of a fast-scanning mass spectrometer. This technique ('Combination g.l.c.–mass spectrometry'[27]) has been used for analysis of the mixture produced by trifluoroacetylation and esterification of the amino-acids obtained by acid hydrolysis after treatment of butyl-lithium with atomic nitrogen.[28]

The mass spectra of the methyl esters of a few *N*-formylamino-acids[29] and one *N*-isovalerylamino-acid[30] have been recorded.

N-Benzyloxycarbonyl derivatives[31] give spectra consisting essentially of the superimposed spectra of benzyl alcohol and the corresponding isocyanate (formed by pyrolysis) if the source temperature exceeds *ca.* 200°, but if the sample is sufficiently volatile for a spectrum to be obtained at a lower temperature, the protecting group can either break down to give a tropylium ion[32] or it can lose a benzyloxy-radical. Cleavage of the C—CO bond gives ions (*c*) which then expel carbon dioxide with concomitant migration of the benzyl group (Scheme 3). Rearrangements involving the ejection of neutral



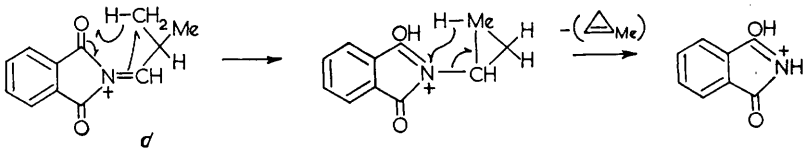
Scheme 3

fragments from non-terminal positions are of some importance because of their mechanistic interest and relevance to the 'element mapping' technique:[33] a detailed review of such processes has appeared.[34]

Phthaloylamino-acids[35] have a tendency to sublime, and good spectra are obtained at moderate source temperatures. A number of pathways can be discerned, but the most interesting is the loss of novel neutral fragments from the intense ions (*d*) formed by loss of the carboxyl group from phthaloylamino-acids with an alkyl side chain [*e.g.*, the mechanism which has been suggested for the case of ion (*d*) from phthaloylvaline (Scheme 4)].

Trimethylsilylation is a general procedure for the preparation of volatile derivatives from compounds with polar functionalities such as hydroxyl,

[26] F. Weygand, B. Koeb, A. Prox, M. A. Tilak, and J. Tomida, *Z. physiol. Chem.*, 1960, **322**, 38.
[27] F. A. J. M. Leemans and J. A. McCloskey, *J. Amer. Oil Chemists' Soc.*, 1967, **44**, 11.
[28] J. Winkler, quoted in ref. 61.
[29] K. Heyns and H. F. Grützmacher, *Z. Naturforsch.*, 1961, **16b**, 293.
[30] K. Tanaka and K. J. Isselbacher, *J. Biol. Chem.*, 1967, **242**, 2766.
[31] R. T. Aplin, J. H. Jones, and B. Liberek, *J. Chem. Soc.* (*C*), 1968, 1011.
[32] S. Meyerson, P. N. Rylander, E. L. Eliel, and J. D. McCollum, *J. Amer. Chem. Soc.*, 1959, **81**, 2606.
[33] K. Biemann, P. Bommer, and D. M. Desidero, *Tetrahedron Letters*, 1964, 1725.
[34] P. Brown and C. Djerassi, *Angew. Chem. Internat. Edn.*, 1967, **6**, 477.
[35] R. T. Aplin and J. H. Jones, *Chem. Comm.*, 1967, 261.



Scheme 4

amino-, and carboxyl groups (*e.g.*, g.l.c. work with carbohydrates,[36] amino-acids,[37] etc.), but only incomplete details of the mass spectra of a few trimethylsilylamino-acid trimethylsilyl esters are available.[38]

Mass spectrometry has been recommended as a means of analysing mixtures of 2,4-dinitrophenylamino-acids,[39] which are produced in Sanger's *N*-terminal amino-acid determination.[40] A second application of mass spectrometry in conjunction with 'wet' methods of sequence analysis is in the identification of phenylthiohydantoins,[41] which are produced by Edman degradation.[42, 43]

3 Peptides which give only Amino-acids on Hydrolysis

A. Free Peptides.—Free peptides are very involatile, but the mass spectra of a few free oligopeptides (*e.g.*, glycylleucyltyrosine[44]) have been reported. Dipeptides[45, 46] undergo thermal cyclisation to dioxopiperazines, and the sequential individuality of the amino-acids is thereby lost. Because of their very low vapour pressures and susceptibility to thermal decomposition, peptides are normally subjected to chemical modification before mass spectrometric examination.

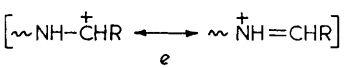
B. Chemical Modification of Peptides for Mass Spectrometry.—Reduction of peptides or acetylpeptides with lithium aluminium hydride (Scheme 5) gives

Scheme 5

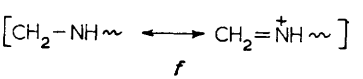
polyamino-alcohols, which are fairly volatile, separable by g.l.c., and give simple mass spectra.[47] The principal ions are of type (*e*), formed by cleavage α to the

[36] C. C. Sweeley, R. Bentley, M. Makita, and W. W. Wells, *J. Amer. Chem. Soc.*, 1963, **85**, 2497.
[37] E. D. Smith and H. Sheppard, *Nature*, 1965, **208**, 878.
[38] R. M. Teeter, paper presented at the ASTM Committee E-14 Mass Spectrometry Conference, New Orleans, 1962.
[39] T. J. Prenders, H. Copier, W. Heerma, G. Dijkstra, and J. F. Arens, *Rec. Trav. chim.*, 1965, **85**, 216.
[40] F. Sanger, *Biochem. J.*, 1945, **39**, 507; 1949, **45**, 563.
[41] N. S. Wulfson, V. M. Stepanov, V. A. Puchkov, and A. M. Zyakoon, *Izvest. Akad. Nauk S.S.S.R., Ser. khim.*, 1963, 1524.
[42] P. Edman, *Acta Chem. Scand.*, 1950, **4**, 283.
[43] P. Edman, *Acta Chem. Scand.*, 1956, **10**, 761.
[44] Ref. 5 (*a*), p. 285.
[45] G. A. Junk and H. J. Svec., *J. Amer. Chem. Soc.*, 1964, **86**, 2278.
[46] K. Heyns and H. F. Grützmacher, *Annalen*, 1963, **669**, 189.
[47] Ref. 5 (*a*), p. 284.

nitrogen atoms. As a series of ions (*e*) corresponding to stepwise loss of amino-acid units from the carboxyl end are observed, the sequence of the original peptide can be deduced. The alternative α cleavage generates a positive charge on a primary carbon atom, and the resulting ions (*f*) are of lower abundance than the ions (*e*). When the polyamino-alcohol contains side-chain hydroxyl groups, further chemical modification (thionyl chloride treatment followed by lithium aluminium deuteride reduction) becomes necessary, and the complexity of the chemical treatment is probably responsible for the fact that this method has not gained popularity.



e



f

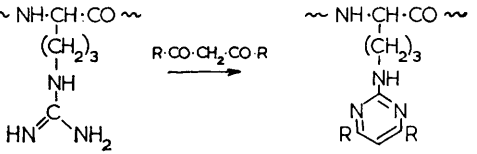
Techniques for the direct insertion of samples into the ion source make it possible to examine peptide derivatives with intact amide bonds, and all recent work has been on *N*-protected peptides. *N*-Protection eliminates the zwitterionic character, and it is in any case necessary to mark the *N*-terminal amino-acid so that ions with the same sequence as the original peptide up to the point of cleavage can be identified. In addition, esterification of the carboxyl terminal (methyl,[48] ethyl,[49] or t-butyl[50] esters) gives increased volatility, though it has been stated[51] that this is advantageous in marginal cases only.

Andersson's preliminary studies on trifluoroacetyl peptide esters[52] were extended by Stenhagen,[48] who was the first to point out the implications of the mode of fragmentation of the peptide bond for sequence determination, and the cleavage patterns were further clarified by extensive investigations on acetylpeptides,[46] trifluoroacetylpeptide esters,[53] and peptide esters with fatty acid ($C_{>10}$) acyl groups.[54] Techniques for the acylation and esterification of peptides on a very small scale have been described.[55, 56] Other *N*-protected

[48] E. Stenhagen, *Z. analyt. Chem.*, 1961, **181**, 462.
[49] V. G. Manusadzhyan, A. M. Zyakoon, A. V. Chuvilin, and Ya. M. Varshavskii, *Izvest. Akad. Nauk Arm. S.S.R., Ser khim.*, 1964, **17**, 143.
[50] M. M. Shemyakin, Yu. A. Ovchinnikov, A. A. Kiryushkin, E. I. Vinogradova, A. I. Miroshnikov, Yu. B. Alakhov, V. M. Lipkin, Yu. B. Shetsev, N. S. Wulfson, B. V. Rosimov, V. N. Bocharev, and V. M. Burikov, *Nature*, 1966, **211**, 361.
[51] K. Biemann, C. Cone, B. R. Webster, and G. P. Arsenault, *J. Amer. Chem. Soc.*, 1966, **88**, 5598.
[52] C.-O. Andersson, *Acta Chem. Scand.*, 1958, **12**, 1353.
[53] F. Weygand, A. Prox, H. H. Fessel, and K. K. Sun, *Z. Naturforsch.*, 1965, **20b**, 1169.
[54] E. Bricas, J. van Heijenoort, M. Barber, W. A. Wolstenholme, B. C. Das, and E. Lederer, *Biochemistry*, 1965, **4**, 2254.
[55] M. Senn, R. Venkataraghavan, and F. W. McLafferty, *J. Amer. Chem. Soc.*, 1966, **88**, 5593.
[56] A. A. Kiryushkin, Yu. A. Ovchinnikov, M. M. Shemyakin, V. N. Bocharev, B. V. Rosinov, and N. S. Wulfson, *Tetrahedron Letters*, 1966, 33.

peptides which have been examined include benzyloxycarbonyl,[31] phthaloyl,[51] ethoxycarbonyl,[56] hexanoyl,[57] and 2,4-dinitrophenyl[58] derivatives.

When functional side chains are present, there are difficulties owing to decreased volatility and additional routes for thermal degradation. Arginine-containing peptides were for these reasons outside the realm of mass spectrometric investigation until recently, when it was shown that conversion of arginine residues into δ-(2-pyriminidyl)ornithine residues[59] (Scheme 6) gave derivatives



*Scheme 6**

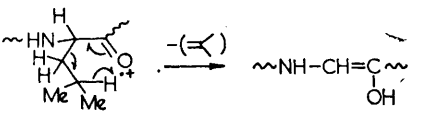
* In this and other schemes and formulae, a wavy line indicates the remainder of the peptide chain.

with sufficient volatility and simple mass spectra.[60] Side-chain amino-groups (lysine, ornithine) and carboxyl groups (aspartic and glutamic acids) are modified simultaneously with the terminal groups during the acylation–esterification procedure.[50] Alcoholic side-chains (serine, threonine) can be left free or converted into their *O*-acetyl derivatives,[61] but with some tyrosine-containing peptides, methylation of the phenolic group is essential in order to obtain satisfactory spectra.[62] Cysteine peptides have only been examined with the thiol function protected by means of a methoxycarbonylmethyl[50] or benzyl[63] group. Although peptides containing unmodified asparagine, glutamine, histidine, or tryptophan residues have been investigated sucessfully, the presence of more than one of these amino-acids prevents the observation of spectra with peptides greater than tetrapeptides.[51] Even in the absence of polar side-chains, there is a limit (at present eight[62] or nine[50, 64] amino-acids) to the length of peptide chains which can be investigated. It is general experience that acetylpeptides containing *N*-substituted amino-acids are more volatile than ordinary acylpeptides.[62] This is because the principal reason for the low volatility of acylpeptides is inter-chain hydrogen bonding, and it was therefore

[57] J. P. Flikwert, W. Heerma, H. Copier, G. Dijkstra, and J. F. Arens, *Rec. Trav. chim.*, 1967, **86**, 293.
[58] T. J. Penders, W. Heerma, H. Copier, G. Dijkstra, and J. F. Arens, *Rec. Trav. chim.*, 1966, **85**, 879.
[59] T. P. King, *Biochemistry*, 1966, **5**, 3454.
[60] A. A. Kiryushkin, M. Yu. Feigina, E. I. Vinogradova, Yu. A. Ovchinnikov, M. M. Shemyakin, Yu. B. Alakhov, N. A. Aldanova, B. V. Rosinov, and V. M. Lipkin, *Experientia*, 1967, **23**, 428.
[61] K. Heyns and H. F. Grützmacher, *Fortschr. Chem. Forsch.*, 1966, **6**, 536.
[62] J. van Heijenoort, E. Bricas, B. C. Das, E. Lederer, and W. A. Wolstenholme, *Tetrahedron*, 1967, 23, 3403.
[63] E. Bayer, C. Jung, and W. König, *Z. Naturforsch.*, 1967, **22b**, 924.
[64] M. Barber, P. Jolles, E. Vilkas, and E. Lederer, *Biochem. Biophys. Res. Comm.*, 1965, **18**, 469.

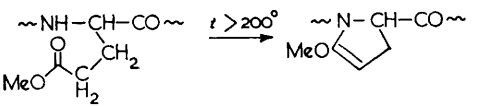
suggested[65] that methods for the replacement of the peptide hydrogen atoms would ameliorate the volatility problems. Treatment of acylpeptides with methyl iodide in dimethylformamide in the presence of silver oxide[66a] results in quantitative methylation of the peptide nitrogen atoms and side-chain functionalities, yielding permethyl derivatives which are more volatile than the original acylpeptides and give excellent mass spectra. The fact that the reaction is heterogeneous is a disadvantage, but providing the sample requirement can be drastically reduced from the present level (*ca.* 2 mg.[66b]) this procedure will probably prove to be of great value.

C. Features of the Spectra Due to the Side Chains.—All amino-acid and peptide derivatives show fragmentations which are characteristic of the side chains. Loss of the side chain *via* a McLafferty rearrangement[67] occurs with leucine, isoleucine, and valine (*e.g.*, Scheme 7), and the observation of peaks



Scheme 7

corresponding to loss of olefin fragments is diagnostic of these amino-acids.[54] Alkyl side chains can also be partly or wholly lost as neutral radicals, and leucine differs sufficiently from isoleucine in this respect to permit distinction.[53] Functional groups in the β position are eliminated (probably by a thermal reaction): *e.g.*, serine derivatives usually show $[M-H_2O]$ peaks. Sometimes thermal degradation is useful: thus in the spectra of α-(γ-methylglutamyl) peptides, all fragments (*m*) retaining the glutamic acid residues are accompanied by $(m-18)$ ions corresponding to pyrolysis of the peptide (Scheme 8), but such

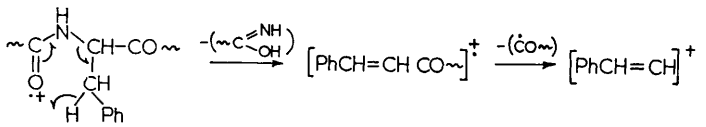


Scheme 8

ions are not observed in the spectra of the isomeric γ-(α-methylglutamyl) peptides.[68] When the side chain is of the type aryl-CH_2– (phenylalanine, tyrosine, histidine, tryptophan), side-chain cleavage occurs with preferential

[65] E. Lederer and B. C. Das, 'Peptides', ed. H. C. Beyerman, A. van de Linde, and W. Massen van den Brink, North Holland Publishing Co., Amsterdam, 1967, p. 131.
[66] (*a*) B. C. Das, S. D. Gero, and E. Lederer, *Biochem. Biophys. Res. Comm.*, 1967, **29**, 211; (*b*) D. S. Millington, personal communication.
[67] Ref. 5*b*, p. 66.
[68] A. A. Kiryushkin, A. I. Miroshnikov, Yu. A. Ovchinnikov, B. V. Rosinov, and M. M. Shemyakin, *Biochem. Biophys. Res. Comm.*, 1966, **24**, 943.

retention of charge by the side-chain moiety.[69] Simple fission of the C_α–C_β bond gives ions of type [aryl–CH_2]$^+$, which, in the case of the $[C_7H_7]^+$ ion[32] at least, have the aromatic tropylium ion structure. Alternatively, transfer of one of the benzylic hydrogen atoms gives rise to characteristic conjugated ions as shown in Scheme 9. Cleavages involving the side chain are always prominent



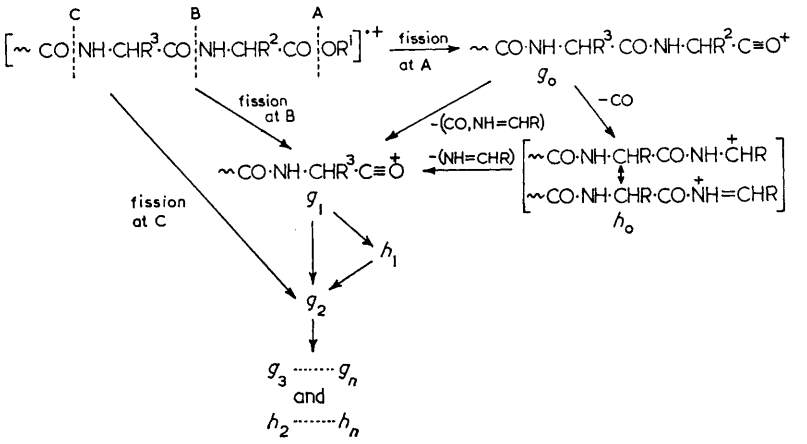
Scheme 9

features in the mass spectra of peptides containing aromatic amino-acids,[69] and usually dominate the spectra of tryptophan derivatives.[70] If an acylpeptide methyl ester has any basic heterocyclic side-chains, the ions (*m*) containing the heterocyclic ring often have satellite ($m + 14$) ions.[50, 55, 60] These peaks probably arise by a thermal transfer of a methyl group from the *C*-terminal ester function to the basic nitrogen atom. Similar complications have been observed in the mass spectrum of the indole alkaloid voacamine.[71]

Other types of side-chain also have their characteristic modes of cleavage (*e.g.*, ring contraction of proline derivatives[50]) but these are not yet as fully documented as those mentioned above. The generalisation has been made[72] that 'the specific behaviour of the individual amino-acids is more strongly expressed the closer the particular amino-acid is to the *N*-terminus', but the basis for this statement is not clear as no systematic comparison of sequential isomers has been made. Indeed, in the case of some simple acetyldipeptides containing leucine,[46] the reverse appears to be true, as expulsion of isobutene from the molecular ions of these compounds is more marked when leucine is in the *C*-terminal position.

D. Fragmentation of the Peptide Bond.—Ionised peptide chains rupture at the amide bonds in two main modes:[46, 50, 53, 54] (i) Cleavage of the CO–N bond (or, at the *C*-terminus, the CO–O bond) gives acylium ions (*g*) which then lose the next amino-acid residue either by concerted loss of carbon monoxide and a neutral imine fragment or *via* the acyliminium ion (*h*) (Scheme 10). When $R^1 = H$ (*i.e.*, glycine residues), the ions (*h*) are of lower abundance than in other cases, as the electron-donating effect of alkyl groups is a stabilising influence. The acylium ions (*g'*) formed by β ($n = 1$) and γ ($n = 2$) peptides do not lose carbon monoxide, as this would give an ion (*h'*) in which the positive charge could not be delocalised as in (*h*): the absence of ions (*h'*) has therefore

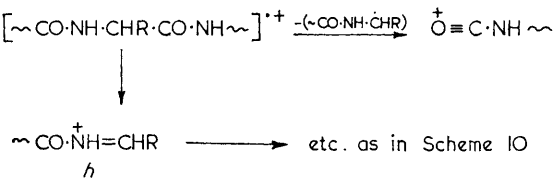
[69] K. Heyns and H. F. Grützmacher, *Annalen*, 1966, **698**, 24.
[70] P. Pfaender, *Annalen*, 1967, **707**, 209.
[71] D. W. Thomas and K. Biemann, *J. Amer. Chem. Soc.*, 1965, **87**, 5447.
[72] M. M. Shemyakin, Yu. A. Ovchinnikov, and A. A. Kiryushkin, 'Peptides', ed. H. C. Beyerman, A. van de Linde, and W. Massen van den Brink, North Holland Publishing Co., Amsterdam, 1967, p. 155.



Scheme 10

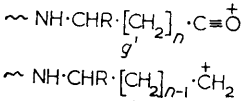
been recommended as a criterion for distinguishing β-alanyl, β-aspartyl, and γ-glutamyl peptides from their α-linked isomers.[62] The ions (*h*) also decompose by hydrogen transfer giving an amine fragment and a neutral substituted keten (cf. Scheme 2).

(ii) Cleavage of the C–CO bond is generally less favoured than the alternative fission between carbon and nitrogen, but occurs with retention of charge by either moiety (Scheme 11).



Scheme 11

E. Determination of Amino-acid Sequence.—It follows from the above discussion that the mass spectrum of an acylpeptide ester will contain a series of ions (g_0), (g_1), (g_2) etc., and a series (h_0), (h_1), (h_2) etc. corresponding to loss of 0, 1, 2 etc. amino-acid residues. Clearly, if these ions (the 'sequence ions') can be identified, the amino-acid sequence of the peptide can be deduced. The



sequence ions frequently comprise only a small proportion of the total ion current, but their recognition can be facilitated by judicious choice of the protecting groups. Thus acylation of the *N*-terminus with an equimolecular mixture of acetic and trideuterioacetic acids[62] causes the ions with an intact *N*-terminus (*i.e.*, the sequence ions) to appear as pairs of peaks of equal intensity separated by 3 mass units. An equimolecular mixture of hepta- and octa-decanoic acids[54] can be used in a similar manner, with the additional advantage that the sequence ions appear in the high mass range, clear of the majority of the other fragments. Interpretation of the spectra can also be simplified by comparing the fragmentation of the methyl and trideuteriomethyl esters or the acylpeptide, which identifies ions retaining the *C*-terminus.[57]

Because the mass spectra of peptide derivatives are so complex, data-processing techniques are of great importance. Weygand's 'Differenzschema'[53] consists of a systematic examination of the difference in mass between the peaks in the spectrum. Starting from the molecular ion, a search is made for differences which correspond to loss of amino-acid residues, and the relationships between the sequence ions are checked by use of metastable[73] ions, *i.e.*, broad low-intensity peaks due to ions which decompose during acceleration. The mass (m^*) of such ions is *ca.* m_2^2/m_1, where m_1 is the mass of the parent ion and m_2 that of the fragment ion, so that observation of a metastable ion can be used to relate the peaks at m_1 and m_2.[73]

Recent technical advances[74] mave made it possible, with the aid of computers to determine complete high-resolution mass spectra (*i.e.*, spectra in which the *m/e* of every peak is determined to a degree of accuracy which permits unambiguous assignment of elemental composition). Because the primary results come from a computer, and because these spectra comprise such an enormous amount of precise information, it is logical to extend the application of the computer to analysis of the spectra.

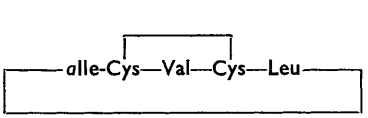
In the computer program devised by McLafferty *et al.*,[55] the *N*-terminal amino-acid is identified by checking combinations of the exact mass of the *N*-protecting group (43·01539 for $CH_3.CO$–, etc.) and the exact masses of each of the possible amino-acid residues against the observed exact masses. The computer then identifies the next amino-acid by the same procedure, and continues the process until addition of the exact mass of the *C*-terminal protecting group locates the molecular ion. The program can also accommodate cases where no molecular ion is observed by searching for typical fragment ions such as $[M-H_2O]^{\cdot+}$ $[M-CO_2]^{\cdot+}$ etc. The logic of the method of Biemann *et al.*[51] is similar, but in this case the *N*-terminal amino-acid is identified by checking combinations of the exact mass of the protecting group plus –NHCH– with the exact masses of each of the possible side chains against the observed exact masses. The subsequent sequence is determined analogously. The program of Barber *et al.*,[3] on the other hand, elaborates the sequence starting from the

[73] Ref. 5*e*, p. 251.
[74] F. W. McLafferty, *Science*, 1966, **151**, 641; also relevant literature cited in refs. 3, 51, and 55.

C-terminus: the need for a molecular ion is a limitation, but this program is more versatile than the two mentioned above, as it applies to cyclopeptides and cyclodepsipeptides as well as linear peptides. The programs of Biemann[51] and McLafferty[55] aim to eliminate all chemical examination by determining not only the sequence but also the identity of the amino-acids, which seems unnecessarily restrictive as amino-acid analyses are now easily obtained from very small samples: Barber's[3] program can use any available chemical information.

The use of high-resolution spectra makes it unnecessary to use special protecting groups for the identification of the sequence ions: even a simple *N*-acetyl group gives the sequence ions unique masses which are determined with a degree of accuracy which eliminates any ambiguity. Further, the subjective element of interpretation is removed, because the computer considers all the peaks in the spectrum, irrespective of their intensity.

F. Cyclopeptides.—Apart from cyclodipeptides (2,5-dioxopiperazines),[45, 46] only a small number of cyclopeptides[46, 75] has been examined. These compounds exhibit comparatively abundant molecular ions: the cross-linked cyclopentapeptide malformin B_1[76] (IV) is an outstanding example in this respect.

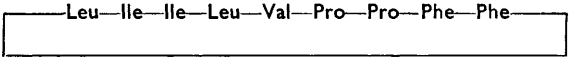


(IV)*

* In this and other formulae, abbreviated designations for amino-acids and their mode of use in representing peptide derivatives is as recommended in I.U.P.A.C. Information Bulletin No. 25, 1966.

Cyclopeptide rings open on electron impact giving acyclic ions which lose amino-acid residues in a stepwise manner in the same way as linear peptides. The amino-acid sequence can be deduced from the mass spectrum, but the occurrence of ring opening at more than one site complicates the interpretation.

The case of the cyclononapeptide (V)[77] isolated from linseed is instructive. Partial methanolysis of the peptide followed by trifluoroacetylation gave a mixture of trifluoroacetyl-di, -tri-, and -tetra-peptide methyl esters which were separated by g.l.c. The amino-acid sequences of these degradation products were determined by mass spectrometry, by use of the 'Differenzschema' (see p. 312), and consideration of the ways in which the sequences overlapped indicated structure (V), which was consistent with the mass spectrum of the intact cyclopeptide.

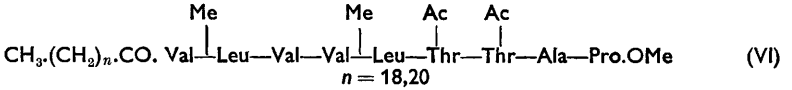


(V)

[75] B. J. Millard, *Tetrahedron Letters*, 1965, 3041.
[76] S. Takeuchi, M. Senn, R. W. Curtis, and F. W. McLafferty, *Phytochemistry*, 1967, **6**, 287.
[77] A. Prox and F. Weygand, 'Peptides', ed. H. C. Beyerman, A. van de Linde, and W. Massen van den Brink, North Holland Publishing Co., Amsterdam, 1967, p. 158.

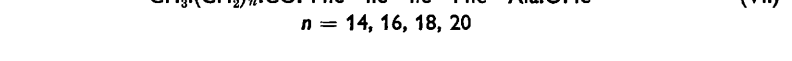
4 Peptides which are not Constructed Exclusively from Amino-acids.
The mass spectra of peptides of this type ('conjugated' peptides such as peptidolipids, glyco-peptides, peptide-alkaloids etc. and depsipeptides) have been of particular interest for two main reasons: (i) certain special aspects of the spectra (especially of peptidolipids) often make them relatively simple to interpret, so that this area has been a convenient testing ground for ideas about mass spectrometric sequential analysis; and (ii) mass spectrometry is well suited for solving structures with unusual features whereas conventional methods of sequence determination are only well developed for normal peptides containing the amino-acids commonly found in proteins.

The first amino-acid sequence to be determined by mass spectrometry was that of fortuitine[64] (VI), an acylnonapeptide ester isolated from *Mycobacterium*



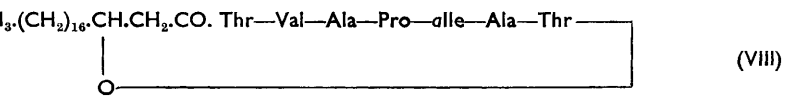
(VI)

fortuitum. The mass spectrum contained a prominent and complete set of sequence ions which were easily recognised because (i) they all occurred at $m/e > 400$, clear of most of the other ions; and (ii) all ions with an intact *N*-terminus appeared as pairs of equally intense peaks separated by 28 mass units, owing to the fact that the lipid group was an equimolecular mixture of homologues. A further example is afforded by the peptidolipid isolated from *Mycobacterium johnei*, for which a tetrapeptidolipid structure had been considered[78] on the basis of hydrolytic degradation. Mass spectrometry,[79] however, indicated the pentapeptide structure (VII): the amino-acid analysis was incorrect because the hindered isoleucylisoleucine peptide bond is only hydrolysed under extreme conditions.



(VII)

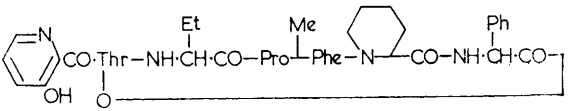
Peptidolipin NA (VIII),[80] Val[6]-peptidolipin NA,[81] and α-aminobutyryl[1]-peptidolipin NA[82] (a series of closely related peptidolipid lactones from *Nocardia asteroides*) all underwent ring opening on electron impact by loss of carbon



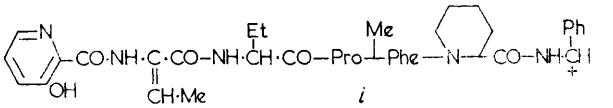
(VIII)

[78] G. Laneele and J. Asselineau, *Biochim. Biophys. Acta*, 1962, **59**, 731.
[79] G. Laneele, J. Asselineau, W. A. Wolstenholme, and E. Lederer, *Bull. Soc. chim. France*, 1965, 2133.
[80] M. Barber, W. A. Wolstenholme, M. Guinand, G. Michel, B. C. Das, and E. Lederer, *Tetrahedron Letters*, 1965, 1331.
[81] M. Guinand, M. J. Vacheron, G. Michel, B. C. Das, and E. Lederer, *Tetrahedron*, 1966, Suppl. No. 7, 271.
[82] M. Guinand, G. Michel, B. C. Das, and E. Lederer, *Vietnamica Chim. Acta*, 1966, 37.

dioxide from the ester function giving open-chain ions, the subsequent fragmentation of which revealed the amino-acid sequences. The antibiotic Staphlomycin S (IX),[83] which is particularly interesting because of the predominance of unusual amino-acids, behaved in a similar manner: the molecular ion lost carbon dioxide and a hydrogen atom to give the acylpeptide acylium ion (*i*). The stepwise splitting of (*i*) confirmed the amino-acid sequence

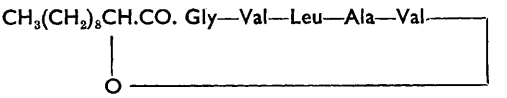


(IX)



i

which had been suggested from other evidence. A group of peptidolipid lactones isolated from the genus *Isaria* has also been investigated,[84] but the spectra of these compounds were more complex than in the examples given above. Thus the confirmation of the structure of isariin (X)[85] required an examination of



(X)

the spectra of partial degradation products in addition to that of the intact natural product.

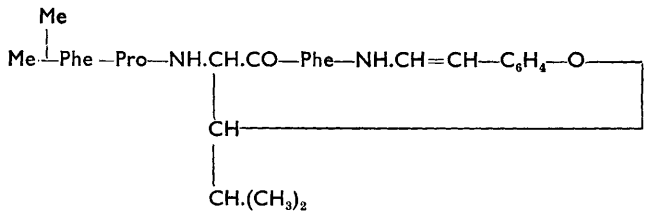
There have been a number of detailed studies[86, 87] of the mass spectra of cyclodepsipeptides containing more than one ester link, but as the fragmentation pathways have been summarised in a previous Quarterly Review,[88] no further discussion will be given here.

Examination of mycoside C*b*[89] has shown that mass spectrometry is also useful in structural work with glycopeptides, and investigations of some synthetic *N*-acylaminoacyl-2-deoxy-2-acetamido-3,4,6-tri-*O*-acetyl-β-D-glucosylamines indicate that the presence of an amino-acid–hexosamine link in a glycopeptide can be inferred from the mass spectrum of a suitable derivative.[90]

A combination of mass spectrometry and chemical degradation has proved

[83] A. A. Kiryushkin, V. M. Burikov, and B. V. Rosinov, *Tetrahedron Letters*, 1967, 2675.
[84] L. H. Briggs, B. J. Fergus, and J. S. Shannon, *Tetrahedron*, Suppl. No. 8 (I), 1966, 269.
[85] W. A. Wolstenholme and L. C. Vining, *Tetrahedron Letters*, 1966, 2785.
[86] C. H. Hassal and J. O. Thomas, *Tetrahedron Letters*, 1966, 4485.
[87] N. S. Wulfson, V. A. Puchkov, V. N. Bocharev, B. V. Rozinov, A. M. Zyakoon, M. M. Shemyakin, Yu. A. Ovchinnikov, V. T. Ivanov, A. A. Kiryushkin, E. I. Vinogradova, M. Yu. Feigina, and N. A. Aldanova, *Tetrahedron Letters*, 1964, 951.
[88] D. W. Russell, *Quart. Rev.*, 1966, **20**, 559.
[89] E. Vilkas, A. Rojas, B. C. Das, W. A. Wolstenholme, and E. Lederer, *Tetrahedron*, 1966, **22**, 2809.
[90] L. Mester, A. Schimpl, and M. Senn, *Tetrahedron Letters*, 1967, 1697.

very useful in the assignment of the related cyclic peptide alkaloids scutianin (XI),[91] pandamine,[92] zizyphin,[93] and ceanthonine B.[94]



(XI)

5 Conclusion

Mass spectrometry is now firmly established as a powerful tool for the solution of structural problems involving conjugated peptides. With peptides of the type obtained by partial hydrolysis of proteins, however, improvements in procedures for the preparation and separation of suitable derivatives are required before the full potential of the method can be harnessed. Many workers are applying themselves assiduously to the remaining problems, and we can expect an increasing swing towards mass spectrometric sequential analysis in the near future.

Since this review was completed, an important new method (treatment with mythyl iodide in dimethylsulphoxide containing sodium methylsulphinylmethide) for the permethylation of peptides for mass spectrometric purposes has been described.[95]

I thank Balliol College, Oxford, for a Junior Research Fellowship and Drs. R. T. Aplin, I. Eland, and G. T. Young for discussions.

[91] R. Tschesche, R. Welters, and H.-W. Fehlhaber, *Chem. Ber.*, 1967, **100**, 323.
[92] M. Pais, X. Monseur, X. Lusinchi, and R. Goutarel, *Bull. Soc. chim. France*, 1964, 817.
[93] E. Zbiral, E. L. Menard, and J. M. Müller, *Helv. Chim. Acta*, 1965, **48**, 404.
[94] E. W. Warnhoff, J. C. N. Ma, and P. Reynolds-Warnhoff, *J. Amer. Chem. Soc.*, 1965, **87**, 4198.
[95] E. Vilkas and E. Lederer, *Tetrahedron Letters*, 1968, 3089.