Methods of Peptide Sequencing. Part I. Conversion of Oligopeptides into Cyclic Dipeptides: a Gas Chromatographic–Mass Spectrometric Study

By **Robert A. W. Johnstone*** and **T. Jeffery Povall**, The Robert Robinson Laboratories, The University, Liverpool L69 3BX

The chemical degradation of oligopeptides to cyclic dipeptides and identification of the latter by gas chromatography–mass spectrometry has been investigated as a method of amino-acid sequencing. Advantages and disadvantages of the method are described.

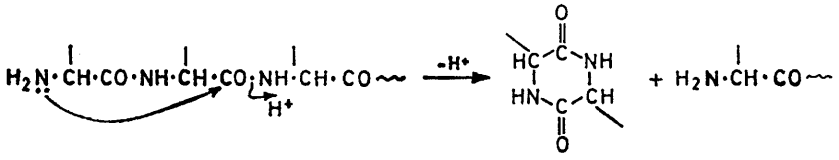
THE transformation of peptides into cyclic dipeptides (dioxopiperazines) by pyrolysis, followed by identification of the latter by g.l.c. has been suggested as a method of amino-acid sequencing.[1] This method yields all possible 'sequential' cyclic dipeptides from an oligopeptide but it is not yet clear whether misinformation can be generated by application of the vigorous pyrolytic conditions to peptides larger than the actinomycins. We have examined an alternative acid-catalysed conversion of oligopeptides into cyclic dipeptides [2] based on earlier limited examples in which peptides were heated with 2-naphthol [3] or acetic acid.[4] As examples the tetrapeptide Leu-Gly-Gly-Leu gave only the cyclic dipeptide *cyclo*(-Leu-Gly-),[3] and the tripeptide Phe-Gly-Gly gave *cyclo*(-Phe-Gly-).[4] There are other reports [5] of the formation of cyclic dipeptides from tripeptides and of their formation in pyridine from 2,4,5-trichlorophenyl esters of tetrapeptides.[6]

As an illustration of the procedure envisaged in this work, a tetrapeptide, ABCD, would be converted sequentially into two cyclic dipeptides, AB and CD. Removal of the *N*-terminal amino-acid by the Edman method would enable this unit to be identified and give a tripeptide, BCD, which could be converted into one cyclic dipeptide, BC. This method was suggested in essence many years ago [4] but was not pursued because identification of the components of the product mixture was impossible at the time on the small scale required. However, the identification of cyclic peptides by g.l.c.–mass spectrometry provides a sensitive means of investigating such mixtures on a small scale and was used in this work. The use of g.l.c. alone to identify cyclic dipeptides is not attractive because of the large number of possible cyclic dipeptides (210 from the 21 common amino-acids). Use of relative retention times requires synthesis of all possible cyclic dipeptides, not to mention

[1] A. B. Mauger, *Chem. Comm.*, 1971, 39.
[2] J. D. Baty, R. A. W. Johnstone, and T. J. Povall, *J.C.S. Chem. Comm.*, 1973, 392.
[3] N. Lichtenstein, *J. Amer. Chem. Soc.*, 1938, **60**, 560.
[4] N. F. Albertson and F. C. McKay, *J. Amer. Chem. Soc.*, 1953, **75**, 5323; F. C. McKay and N. F. Albertson, *ibid.*, 1957, **79**, 4686.
[5] J. Meienhofer, Y. Sano, and R. P. Patel in 'Peptides: Chemistry and Biochemistry,' ed. B. Weinstein, Marcel Dekker, New York, 1970, pp. 419—434; J. C. Sheehan and D. N. McGregor, *J. Amer. Chem. Soc.*, 1962, **84**, 3000; G. Lucente and P. Frattesi, *Tetrahedron Letters*, 1972, **42**, 4285.
[6] K. Titlestad, *Chem. Comm.*, 1971, 1527.

diastereoisomers, and still does not guarantee a positive identification. Use of mass spectrometry with g.l.c. removes these difficulties.

Several methods of converting peptides into cyclic dipeptides were examined by using the tetrapeptide Val-Gly-Leu-Phe as test material. The peptide was heated in (i) dimethylformamide (DMF), (ii) DMF containing phenol, (iii) DMF containing acetic acid, (iv) acetic acid alone, and (v) ethyl phosphorodichloridite in pyridine.[7] Only glacial acetic acid alone or in DMF proved successful. Thus, acidic catalysis is necessary for the sequential production of cyclic dipeptides, perhaps by the mechanism shown in Scheme 1. Although acidic conditions are necessary, more strongly acidic conditions (HCl) degrade a peptide into amino-acids and small peptides but prevent cyclic dipeptide formation, probably by protonation of the terminal amino-group. Additional evidence for the sequential formation of cyclic dipeptides is provided by the pentapeptide [Glu-Gly-Pro-Trp-Leu, which has no free terminal amino-group and did not form any cyclic dipeptides.



SCHEME 1

Our study of the conversion of peptides into cyclic dipeptides followed by mass spectrometry of the product mixture was encouraging, since ions due only to the expected cyclic dipeptides were observed (Table 1). However, unequivocal identification of the cyclic dipeptides is only possible through separation of the products by g.l.c. followed by mass spectrometry of the separated components, and so g.l.c.–mass spectrometry of the *O*-trimethylsilyl derivatives of the cyclic peptides was carried out (Table 2). Thus, the pentapeptide H-Gly-Leu-Leu-Gly-Gly-OH gave only one cyclic dipeptide, *cyclo*(-Gly-Leu-), on treatment with glacial acetic acid. Edman degradation with phenyl isothiocyanate enabled identification of the terminal glycine as the corresponding phenylthiohydantoin (identified by mass spectrometry), and gave a tetrapeptide, Leu-Leu-Gly-Gly. With acetic acid, this tetrapeptide gave two cyclic dipeptides, *cyclo*(-Leu-Leu-) and *cyclo*(-Gly-Gly-). These, and similar experiments with other peptides, were carried out on about 20 μmol of material and yields of cyclic dipeptides were about 50%. Since these were readily identifiable on this scale, the method appeared suitable for use on a much smaller scale. All except one of the simple peptides shown in Tables 1 and 2 behaved similarly to the above pentapeptide and no difficulties were encountered.

TABLE 1

Degradation of peptides to cyclic dipeptides [a]

Peptide [b]	Cyclic dipeptides detected	Identifying ions in mass spectrum [c]
Ala-Phe-Leu	*cyclo*(-Ala-Phe-)	44, 91, 114, 120, 128, 175, *218*
Gly-Pro-Trp-Leu	*cyclo*(-Gly-Pro-)	70, 83, 98, 111, *154*
	cyclo(-Trp-Leu-)	130, *299*
Trp-Gly-Leu-Phe	*cyclo*(-Trp-Gly-)	30, 77, 103, 130, *243*
	cyclo(-Leu-Phe-)	91, 113, 120, 141, 169, 204, *260*
Trp-Met-Asp(OMe)-Phe	*cyclo*(-Trp-Met-)	77, 130, *317*
	cyclo[-Asp(OMe)-Phe-]	91, 114, 120, 125, 157, 185, 232, 245, *276*
Gly-Phe-Gln-Gly-Gly	*cyclo*(-Gly-Phe-)	91, 113, *204*
	cyclo(-Gln-Gly-)	84, 126, 168, *185*

[a] Identified by mass spectrometry of the mixture without g.l.c. separation. [b] For each peptide, the same experiment was repeated on the peptide remaining after removal of the *N*-terminal amino-acid by the Edman method. [c] *m/e* values; those italicised are molecular ions.

TABLE 2

Degradation of peptides to cyclic dipeptides [a]

Peptide [b]	Cyclic dipeptide detected	Identifying ions in mass spectrum [c]
Gly-Lys-Gly	*cyclo*(-Gly-Lys-)	174, 257, 458, *473*
Ala-Tyr-Leu-Phe	*cyclo*(-Ala-Tyr-)	179, 271, 435, *450*
	cyclo(-Leu-Phe-)	257, 285, 313, 347, 389, *404*
Ser-Gly-Leu-Phe	*cyclo*(-Ser-Gly-)	103, 147, 257, 345, *360*
	cyclo(-Leu-Phe-)	See above
Ala-Glu(OMe)-Leu-Gly	*cyclo*[-Ala-Glu(OMe)]	241, 257, 326, 327, 343, *358*
	cyclo(-Leu-Gly-)	156, 184, 257, 258, 271, 299, *341*
Lys-Val-Phe-Gly	*cyclo*(-Phe-Gly)	91, 229, 257, 333, *348*
Gly-Leu-Leu-Gly-Gly	*cyclo*(-Leu-Gly-)	See above

[a] Identified by mass spectrometry of g.l.c. separated components (trimethylsilyl derivatives). [b,c] As Table 1.

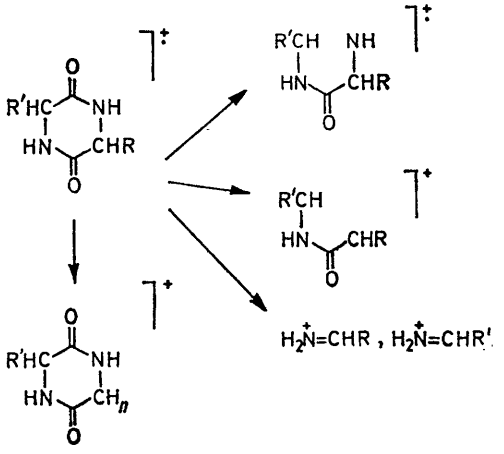
[7] Personal communication from G. W. Kenner; original work by J. M. Turner, Ph.D. Thesis, Cambridge, 1955.

The first indication that the method might not be general, at least under the conditions used here, was provided by the tetrapeptide Lys-Val-Phe-Gly, from which only the cyclic dipeptide, *cyclo*(-Phe-Gly-) was identified although presumably *cyclo*(-Lys-Val-) must have been formed initially (Table 2). However, in the cases of two hexapeptides, H-Thr-Ala-Ile-Gly-Val-Gly-OH and H-Ala-Phe-Ile-Gly-Leu-Val-OH, none of the expected cyclic dipeptides could be detected by g.l.c.–mass spectrometry. Similarly, the B-chain of insulin yielded no identifiable cyclic dipeptides with acetic acid. Thus, the yields of cyclic dipeptides varied from *ca.* 50% down to 0%. The reaction of a peptide with acetic acid afforded, in addition to cyclic dipeptides, brown materials which gave no peaks on g.l.c.–mass spectrometry and are probably products of extensive degradation and/or polymerisation. Mass spectrometry of the crude products at high temperature indicated the presence of considerable amounts of high molecular weight material.

For reference purposes, a variety of cyclic dipeptides was synthesised, mostly by heating the methyl ester of the corresponding dipeptide. For g.l.c., the cyclic dipeptides were made less polar by conversion into trimethylsilyl derivatives,[8] thereby allowing the more complex ones containing amino-acid residues such as serine, tyrosine, and lysine to be chromatographed as well as those containing simpler amino-acids. The polarity of the cyclic peptides was also reduced by permethylation,[9] but the mass spectra of these permethylated compounds were inferior to those of the trimethylsilyl derivatives for identification purposes.

The mass spectra of several cyclic dipeptides have been reported;[10,11] linear dipeptides gave mass spectra due to a combination of the spectra of the linear and cyclic dipeptides,[10] the latter being formed in the mass spectrometer. The mass spectra of the cyclic peptides investigated here generally gave abundant molecular ions; fragment ions at $M - 28$, $M - 29$, and $M - 43$ were sometimes but not invariably present and were often of low abundance, in contrast with earlier suggestions.[10,11] Similarly, fragment ions of the type $H_2N{=}CHR$ (R = amino-acid side-chain), were often but not always present (Scheme 2). With the exception of those containing proline, the cyclic dipeptides showed abundant fragment ions due to loss of the side-chain of one of the amino-acid residues with or without hydrogen transfer. The loss of the side-chain of leucine, valine, or serine occurred so readily that no molecular ion, or only one of very low abundance, was present (Scheme 2). For cyclic dipeptides containing aromatic side-chains, the major ions were due to the side-chains themselves, *e.g.* *m/e* 81 and 82 for histidine. In cyclic dipeptides containing tryptophan, the mass spectra showed little else than an ion at *m/e* 130 together with a molecular ion. Lysine-containing cyclic dipeptides gave an ion at *m/e* 84, characteristic of lysine, as well as ions due to loss of NH_3; for glutamine-containing cyclic dipeptides, by far the most abundant fragment ion was produced by loss of NH_3 from the molecular ion. The mass spectra of permethylated cyclic dipeptides showed little or no improvement over those of unmethylated ones.

The mass spectra of trimethylsilylated cyclic dipeptides were much more satisfactory. Apart from ions at *m/e* 73 and 75 from the trimethylsilyl group itself, the derivatives gave abundant molecular and $M - 15$ ions. Trimethylsilylated cyclic dipeptides containing aromatic side-chains gave fragment ions resulting from loss of the side-chain; for phenylalanine the charge was retained on the cyclic dipeptide ring but for tyrosine and tryptophan the charge was retained by the side-chain (*m/e* 179 and 202, respectively). Lysine-containing cyclic dipeptides



(n = 1 or 2)

SCHEME 2

gave a tristrimethylsilyl derivative under mild conditions but longer reaction times gave the tetrakistrimethylsilyl derivative.

Conclusion.—A general theoretical approach to obtaining the sequence of amino-acids in an oligopeptide by converting it into cyclic dipeptides has been shown to work practically in several instances. Analysis and identification of cyclic dipeptides in mixtures by g.l.c.–mass spectrometry of their trimethylsilyl derivatives has been demonstrated. The actual sequential conversion of oligopeptides into cyclic dipeptides was found to give variable yields of products under the conditions used making the method at present unacceptable as a general procedure where only limited amounts of peptide are available.

EXPERIMENTAL

The following analytical conditions were satisfactory for cyclic dipeptides. T.l.c. was carried out on silica gel with $CHCl_3$–MeOH (9 : 1) or $CHCl_3$–MeOH–AcOH (70 : 10 : 5).

[8] A. B. Mauger, *J. Chromatog.*, 1968, **37**, 315.

[9] G. Marino, L. Valente, R. A. W. Johnstone, F. Mohammadi-Tabrizi, and L. Sodini, *J.C.S. Chem. Comm.*, 1972, 357.

[10] H. J. Svec and G. A. Junk, *J. Amer. Chem. Soc.*, 1964, **86**, 2278.

[11] K. Jankowski and L. Varfalvy, *Bull. Acad. polon. Sci., Sér. Sci. chim.*, 1971, **19**, 651; 1972, **20**, 423; R. Nagarajan, J. L. Occolowitz, N. Neuss, and S. M. Nash, *Chem. Comm.* 1969, 359.

G.l.c. was performed on 5% OV1 columns (2·1 m × 6 mm) at 190—320 °C with a nitrogen flow rate of 45 ml min^{-1}. Mass spectra were obtained with an A.E.I. MS12 spectrometer at 70 eV. G.l.c.–mass spectrometry was carried out on a column (2·1 m × 6 mm) of either 5% OV1 on GasChrom Q or 3% OV 17 on GasChrom Q initially at 166—170 °C and then temperature-programmed at 4° min^{-1} to 250 °C with the carrier gas flowing at 30 ml min^{-1}.

Synthesis of Cyclic Dipeptides.—In a typical experiment, the dipeptide H-Leu-Phe-OMe (20 mg) was heated in refluxing dimethylformamide (3 ml) for 12 h. The solution (ninhydrin-negative) was evaporated *in vacuo* and the residue sublimed at 140—170 °C and 0·5 mmHg to give the cyclic dipeptide. The purity was checked by analysis, mass spectrometry, t.l.c., and in some instances, ^1H n.m.r. spectroscopy. Similarly, the following cyclic dipeptides were prepared [*cyclo*(-A-B-)]; A = Gly, B = Gly, Ala, Val, Leu, Pro, Lys, Ser, Phe, or Glu; A = Pro, B = Val, Pro, Leu, Ile, or Phe; A = Phe, B = Leu or Asp(OMe); A = Tyr, B = Ala; A = Trp, B = Leu or Met; A = Lys, B = Val; A = Leu, B = Leu; A = Ile, B = His.

Trimethylsilylation. The cyclic dipeptide (<1 mg) was left in a stoppered vial with *NO*-bistrimethylsilylacetamide (0·1—0·2 ml) at 80 °C for 1½ h. A sample of this solution was injected straight onto the g.l.c. column.

Edman degradation. This was carried out in the usual way with phenyl isothiocyanate to give a phenylthiohydantoin (identified by mass spectrometry) and a residual peptide which was then degraded to cyclic dipeptides.

Degradation of Peptides to Cyclic Dipeptides.—(*a*) The peptide (1·5 mg) was dissolved in dry pyridine (1 ml) and ethyl phosphorodichloridite (0·01 ml) was added at 0 °C. The solution was allowed to reach room temperature and was then heated on a steam-bath for 48 h. The solution was evaporated *in vacuo* to leave a brown gum in which no cyclic dipeptides were detected.

(*b*) The peptide (4 mg) was heated at 120 °C in dimethylformamide (6 ml) containing phenol (50 mg) for 5 h. Evaporation left a brown gum in which no cyclic dipeptides were detected.

(*c*) The peptide (10 mg) was heated in glacial acetic acid (4 ml) at 100—110 °C for 10 h. Evaporation *in vacuo* left a brown residue containing cyclic dipeptides.

[5/100 *Received*, 16*th January*, 1975]