MASS SPECTROMETRIC SEQUENCE DETERMINATION OF PERMETHYLATED PEPTIDE MIXTURES

P. ROEPSTORFF, R.K. SPEAR and K. BRUNFELDT
The Danish Institute of Protein Chemistry, affiliated to the Danish Academy of Technical Sciences, 33, Finsensvej, DK-2000, Copenhagen, F, Denmark

Received 5 April 1971

1. Introduction

Mass spectrometry of a peptide mixture obtained by cleavage of a protein has been proposed as a means to determine the amino acid sequence of the intact protein [1]. It has been shown that sequence determination of the single peptides in a mixture of synthetic *N*-acetyl-tripeptide-methylesters has been possible utilizing exact mass determination, partial vaporization and metastable ion data [2]. It is demonstrated below that sequence determination may be performed by low resolution mass spectrometry of permethylated mixtures of two or three synthetic peptide derivatives, provided knowledge of the amino acid composition of the mixture.

2. Experimental

The following mixtures of synthetic peptide derivatives were used for the experiments:

(I)*	a)	Z–Gly–Pro–Ala–Thr–OMe	(1.4 mg, 2.9 μmoles)
	b)	Z–Leu–Val–Glu(OBut)–Ala–OMe	(2.0 mg, 3.1 μmoles)
(II)	a)	Ac–Ala–Gly–Ala–Gly–OMe	(1.6 mg, 4.8 μmoles)
	b)	Ac–Asp–Glu–Ala–Asp–Pro–OMe	(2.2 mg, 3.7 μmoles)
(III)	a)	Ac–Ala–Leu–Phe–Gly–OMe	(1.5 mg, 3.2 μmoles)
	b)	Ac–Phe–Gly–Leu–Ala–OMe	(1.5 mg, 3.2 μmoles)
(IV)	a)	Ac–Ala–Leu–Phe–Gly–OMe	(0.9 mg, 1.9 μmoles)
	b)	Ac–Phe–Gly–Leu–Ala–OMe	(0.9 mg, 1.9 μmoles)
	c)	Ac–Ala–Pro–Leu–Phe–Val–Gly–OMe	(1.2 mg, 1.9 μmoles)

Removal of the benzyloxycarbonyl- and the tert-butylestergroup and acetylation of I was carried out as follows: I was dissolved in glacial acetic acid (0.2 ml) and 40% HBr in glacial acetic acid (0.2 ml) was added. After reaction for 30 min at room temperature the peptide methylester hydrobromides were precipitated by the addition of dry ether (4 ml). The precipitate was washed with dry ether and dissolved in methanol (0.5 ml). The solution was neutralized (pH 8) with triethylamine and acetic anhydride (0.2 ml) was added. After reaction for 1 hr at room temperature the reaction mixture was evaporated to dryness.

All the mixtures were permethylated with methyliodide using sodium hydride-dimethylsulfoxyde as base [3, 4].

The mass spectra were obtained on a Perkin Elmer 270 mass spectrometer operating at 70 eV. The resolution of the mass spectrometer was approximately 1000. The samples were introduced directly into the ion source and the temperature

* Abbreviations: Z: Benzyloxycarbonyl, OBut: γ-tert-butylester.

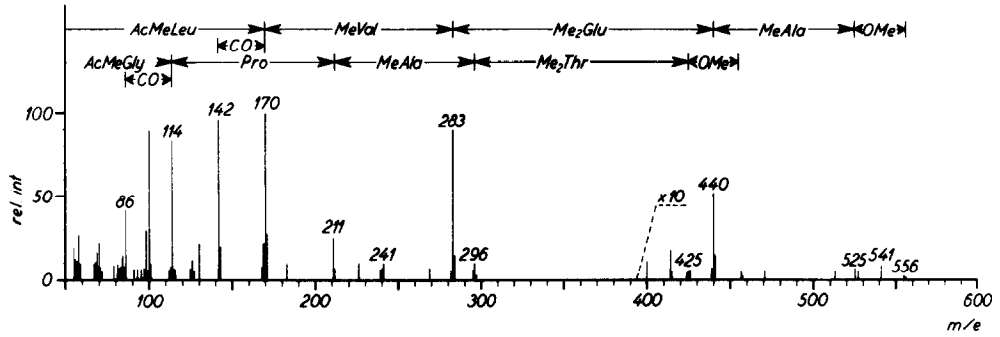

Fig. 1. Peptide mixture I (110°).

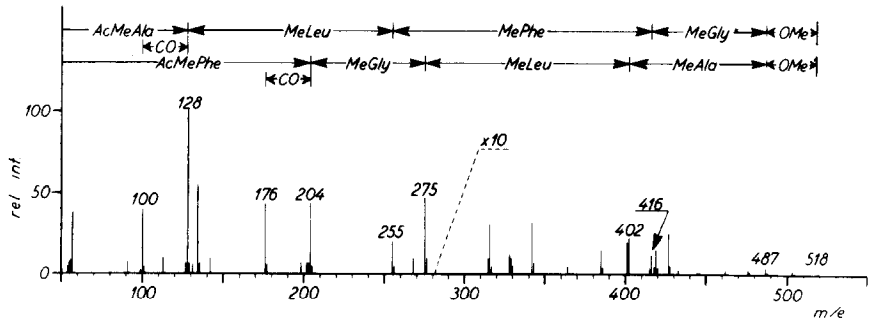

Fig. 2. Peptide mixture III (110°).

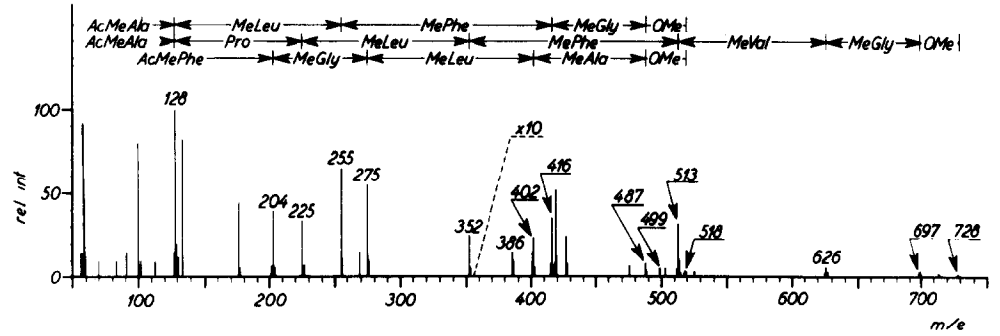

Fig. 3. Peptide mixture IV (150°).

of the solids inlet probe was slowly increased from 70° to 200°. During this period the spectra were recorded at regular intervals in order to observe if partial vaporization occurred. The temperature of the ion source was 150°.

3. Results

3.1. *Peptide mixture I*

The spectrum of peptide mixture I (fig. 1) shows peaks corresponding to three possible *N*-terminal acetylamino acid residues: Ac–MeGly (m/e 114). Ac–MeLeu (m/e 170) and Ac_2Me_2 Lys (m/e 241). The first two of the previous three are further confirmed by losses of CO at m/e 86 respectively m/e

Table 1
Sequences deduced from the mass spectrum of peptide mixture IV.

Ala–Pro (225)	Ala–Leu (255)	Phe–Gly (275)
Ala–Pro–Leu (352) Ala–Pro–Phe (386)	Ala–Leu–Phe (416) Ala–Leu–Pro (352)	Phe–Gly–Leu (402)
Ala–Pro–Leu–Phe (513) Ala–Pro–Phe–Val (499) Ala–Pro–Phe–Leu (513)	Ala–Leu–Phe–Gly (487) Ala–Leu–Phe–Pro (513) Ala–Leu–Pro–Phe (513)	Phe–Gly–Leu–Ala (487) Phe–Gly–Leu–Pro (499)
Ala–Pro–Leu–Phe–Val (626) Ala–Pro–Phe–Val–Leu (626) Ala–Pro–Phe–Leu–Val (626)	Ala–Leu–Phe–Pro–Val (626) Ala–Leu–Pro–Phe–Val (626)	Phe–Gly–Leu–Pro–Leu (626)
Ala–Pro–Leu–Phe–Val–Gly (697) Ala–Pro–Phe–Val–Leu–Gly (697) Ala–Pro–Phe–Leu–Val–Gly (697)	Ala–Leu–Phe–Pro–Val–Gly (697) Ala–Leu–Pro–Phe–Val–Gly (697)	Phe–Gly–Leu–Pro–Leu–Gly (697)

For the simplicity the possible sequences are indicated non-derivatized. The m/e of the peaks corresponding to the shown sequences is indicated in parenthesis.

142. Knowing the amino acid composition of the mixture, Ac_2Me_2 Lys can be rejected. Further analysis of the spectrum based on the possible *N*-terminals and the amino acid composition gives the following possible dipeptides: Ac–MeGly–Pro (m/e 211), Ac–MeGly–MeLeu (m/e 241), Ac–MeLeu–MeGly (m/e 241) and Ac–MeLeu–MeVal (m/e 283). The following tripeptides: Ac–MeGly–Pro–MeAla (m/e 296) and Ac–MeLeu–MeVal–Me_2Glu (m/e 440) and the following tetrapeptides: Ac–MeGly–Pro–MeAla–Me_2Thr (m/e 425) and Ac–MeLeu–MeVal–Me_2Glu–Me–Ala (m/e 525). In the latter case the sequence terminates with peaks at m/e 541 and 556, which indicate the presence of a methylester group. Peptide Ia shows no peak corresponding to the molecular ion, but a peak at m/e 424 corresponds to the loss of methanol from the molecular ion, a loss which is often observed for permethylated peptide derivatives containing serine or threonine.

The possible dipeptide sequences Ac–MeGly–MeLeu and Ac–MeLeu–MeGly do not show corresponding tri- and tetrapeptides or a termination with OMe. These sequences are probably artefacts which may be attributed to the loss of a valine side chain by a McLafferty rearrangement in peptide Ib. Peptide mixture I showed no noticeable partial vaporization effect. The spectra remained unchanged until 170°, after which the intensity of the high molecular weight peaks decreased rapidly.

3.2. *Peptide mixture II*

In peptide mixture II the effect of partial vaporization was very pronounced. At 90–110° the peaks corresponding to IIa dominated the spectra. IIa had nearly disappeared at 150° and the pentapeptide IIb appeared clearly at 160–210°

3.3. *Peptide mixture III*

These two peptides have an identical amino acid content. With knowledge of the amino acids composition of the mixture only the correct sequences could be deduced from the spectrum (fig. 2). Without this information the interpretation of the spectrum results in 6 dipeptide sequences, 19 tripeptide sequences and 11 tetrapeptide sequences.

3.4. *Peptide mixture IV*

This peptide mixture contains the two peptides in mixture III plus a hexapeptide with the same *N*-terminal as IIIa. When considering the amino acids present, the *N*-terminals Ac–MeAla (m/e 128) and Ac–MePhe (m/e 204) and the dipeptide sequences: Ac–MeAla–Pro (m/e 225) Ac–MeAla–MeLeu (m/e 255) and Ac–MePhe–MeGly (m/e 275) are the only ones possible (fig. 3). The possible sequences are listed in table 1. Only two peaks, m/e 487 and 697, involved in the assignment of the sequences in table 1 are followed by peaks indicating a sequence termination by a methyl ester. The first corresponds to the two tetrapeptides IVa and IVb, the second to

all the possible hexapeptides. This is supported by the fact that the amino acid composition, assuming an equimolar mixture can be obtained from the possible sequences only with a mixture of IVa and b and one of the five hexapeptides with *N*-terminal alanine.

Among these hexapeptides only three contain the *N*-terminal Ala–Pro, which was found on the dipeptide level, and which cannot be accounted for by the two accepted tetrapeptides. The relative intensities of the peaks which indicate the presence of these three hexapeptides are as follows:

100 : 35 : 25 : 3.2 : 0.6 : 0.2
100 : 35 : 1.0 : 0.5 : 0.6 : 0.2
100 : 35 : 1.0 : 3.2 : 0.6 : 0.2

A decrease in intensity with a factor of 2 to 10 is normally observed between two successive sequence determining peaks [5]. The hexapeptide Ala–Pro–Leu–Phe–Val–Gly is the only one of the three hexapeptides in question, which is in accordance with this observation. The use of relative intensities must, however, be considered questionable.

By increasing the temperature from 110° to 150° it was observed that the intensity of the peaks which could be assigned to the two tetrapeptides increased with a factor of 2–3 relative to the peaks which could be assigned to the hexapeptide.

4. Discussion

The results show that sequence information may be obtained by low resolution mass spectrometry of permethylated peptide mixtures combined with amino acid analysis of the mixture. Without knowledge of the amino acid composition the sequence assignment would be very ambiguous or impossible from the spectra reported. The advantages of a partial vaporization has already been described [3] and was evident in peptide mixture II. In peptide mixture IV the partial vaporization was not quite so obvious. Based upon the relative intensity data obtained at the two temperatures it was, however, possible to distinguish between two groups of peaks. In peptide mixture IV several possibilities existed for the final assignment of the sequence of the hexapeptide. In this case the peak intensities were useful for indication of the most probable sequence. As mentioned above, however, the use of relative intensities is often incumbered with great uncertainty.

References

[1] G.E. VanLear and F.W. McLafferty, Ann. Rev. Biochem. 38 (1969) 289.
[2] F.W. McLafferty, R. Venkataraghavan and P. Irving, Biochem. Biophys. Res. Commun. 39 (1970) 274.
[3] E. Vilkas and E. Lederer, Tetrahedron Letters (1968) 3089.
[4] D.W. Thomas, Biochem. Biophys. Res. Commun. 33 (1968) 483.
[5] E. Lederer, Pure Appl. Chem. 17 (1968) 489.