THE USE OF DEUTERATED DERIVATIVES IN PEPTIDE MIXTURE ANALYSIS BY MASS SPECTROMETRY

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

Received 29 December 1971

1. Introduction

Mass spectrometrical sequence determination of the single peptides in mixtures of synthetic peptides [1, 2] and of protein-derived peptides [3, 4] has been reported. Exact mass determination and metastable ion data have been used for the interpretation [1]. Also low resolution mass spectrometry of permethylated peptide mixtures combined with partial vaporization [2, 4] and knowledge of the amino acid composition of the mixture have been applied [2]. The techniques described above may, however, not always provide sufficient information in order to obtain a correct interpretation of the mass spectra of peptide mixtures. Techniques for obtaining additional information are therefore required.

The interpretation of mass spectra of peptides and peptide mixtures is mostly carried out by matching numerical mass differences between peaks with the integral masses of different derivatized amino acid residues. For a list of the permethylated amino acid residues see [4]. However, considering only the sequence peaks, it can be foreseen that several possibilities for ambiguities exist when dealing with peptide mixtures. Some examples are presented in table 1.

The alternative sequences in group A and B in table 1 are identical in elementary composition with the correct sequences and thus cannot be eliminated by high resolution mass spectrometry. On the other hand high resolution measurements would eliminate the alternative sequences in group C and D. The ambiguities in group A, C and D, however, depend on the derivatives used. The present experiments demonstrate that alternative sequences, as shown in group A, C and D, can be eliminated by comparing the sequences deduced from the spectra of the acetylated and permethylated mixture with those deduced from the spectra of the deuteroacetylated and deuteropermethylated mixture. Simultaneously, most of the alternative sequences due to non-sequence peaks are eliminated.

Table 1
Examples of ambiguities in mass spectrometry of permethylated peptide mixtures.

	Actual sequence		Alternative sequence
A)	–MeGly–MeLeu–	→	–ε-Ac-Me_2Lys–
	–MeAla–MeVal–	→	–ε-Ac-Me_2Lys–
	–MeGly–MeVal–	→	–δ-Ac-Me_2Orn–
	–MeAla–MeAla–	→	–Me_3Gln–
	–MeAla–MeGly–	→	–Me_3Asn–
B)	Ac-MeLeu–MeAla– Ac-MeAla–	→	Ac–MeAla–MeLeu–
	Ac-MeGly–Me_2Glu– Ac-MeAla–	→	Ac-MeAla–Me_2Asp–
C)	–MeAla–Me_2Thr–	→	–Me_2Try–
D)	Ac-MeLeu–MeMet– Ac-Me_2Asp–	→	Ac-Me_2Asp–Me_2Thr–

When the sequences in the first column are present, the sequences in the second column are also deduced.

North-Holland Publishing Company – Amsterdam

2. Experimental

Two mixtures of synthetic peptide derivatives were used in the experiments:

*I	Boc–Ala–Pro–Leu–Phe–Val–Gly–OMe	1.40 mg ∿ 1.9 μmole
	H–Asp–Glu–Ala–Asp–Pro-OH	1.05 mg ∿ 1.9 μmole
II	Z–Leu–Val–Glu(OBut)–Ala–OMe	0.77 mg ∿ 1.2 μmole
	Z–Gly–Pro–Ala–Thr–OME	0.56 mg ∿ 1.2 μmole

The peptides in I are synthesized by the solid phase technique and the peptides in II in solution.

Removal of the Boc group was carried out by treatment with 0.5 ml trifluoroacetic acid for 30 min at room temp. The trifluoroacetic acid was removed by evaporation. Removal of the benzyloxycarbonyl group and the t-butylester group was carried out as previously described [2]. After removal of the protecting groups the mixtures were dissolved in 1 ml of methanol. Half of each mixture, a, was acetylated by addition of 0.2 ml acetic anhydride; the other half, b, was deuteroacetylated by the addition of 0.2 ml hexadeuteroacetic anhydride (Merck, Darmstadt). After reaction for 1 hr at room temp the samples were evaporated to dryness. The acetylated mixtures, a, were permethylated with methyliodide, using sodiumhydride/dimethylsulfoxide as base [5, 6], and the deuteroacetylated mixtures, b, deuteropermethylated with trideuteromethyliodide (Merck, Darmstadt).

In order to investigate whether the C-terminal methylester was quantitatively transesterified during deuteropermethylation, CD_3CO–Ser–Glu–Pro–Ala–OMe was deuteropermethylated. A molecular peak at m/e 460 corresponding to the deuteromethylester was found and no peak at m/e 457. This peptide was chosen for the experiment because the C-terminal ions (H_2-Pro–MeAla–OMe) and (H_2-MeGly–Pro–MeAla–Ome) give rather intense peaks, allowing a confirmation of the quantitative transesterification.

A Perkin-Elmer 270 mass spectrometer operating at 70 eV was used. The resolution was approx. 1000. The samples were introduced directly into the ion source (temp 150°) and the solids inlet probe was slowly heated from 70° to 200°. During this period spectra were recorded at regular intervals.

* *Abbreviations:* Boc: t-butyloxycarbonyl; Z: Benzyloxycarbonyl; OBut: γ-t-butylester.

3. Results

3.1. *Peptide mixture I*

Analysis of the intensities of the sequence peaks in the spectra obtained at different temperatures showed that no noticeable partial vaporization had taken place. The intensities of the sequence peaks normally decrease with a factor of 2–10 from one sequence peak to the next at higher mass [7]. In order to compensate for the decrease, a stepwise increasing amplification was used. In all amplifications applied, peaks higher than 5% of the base peak were considered during the interpretation. It is thus not a relative intensity understood in the usual way, but a false relative intensity, where a correction for the decrease was aimed at. The interpretation was carried out by locating peaks possibly corresponding to N-terminal acyl amino acid residues. From these peaks, sequences were deduced by trial and error. The results are shown in table 2.

The sequences common for the 2 derivatives are in italics. It is observed that the comparison, in spite of the large number of possibilities, in this case leads to the determination of the correct sequences.

3.2. *Peptide mixture II*

In this experiment the partial vaporization was also negligible. Analysis of the spectra resulted in a list of possible sequences very similar to table 2. Comparison of the results obtained with the 2 derivatives reduced the number of possible sequences to those listed in table 3. It is seen that more sequences are still possible on the tetrapeptide level and additional information must be used in order to obtain an unambiguous interpretation.

In the present example, knowledge of the amino acid composition confirms Leu–Val–Glu–Ala and Gly–Pro–Ala–Thr as the correct sequences. The 2 sequences are also the only ones in table 3 where a termination of the sequence was found. For the first sequence the molecular ion was present and for the second a peak corresponding to loss of methanol from the molecular ion.

The sequence Leu–Val–Glu–Ser can also be eliminated due to the lack of peaks corresponding to a loss of methanol from either the sequence ion or the molecular ion. The peaks on each side of the intense sequence peak of Leu–Val–Glu give rise to both Gly–Pro–Ala–

Table 2

Possible sequences from peptide mixture Ia		
N-terminal	*Ala*(128)	*Asp*(186)
Di-peptides	*Ala-Pro*(225)	*Asp–Glu*(343)
Tri-peptides	*Ala–Pro–Leu*(352) Ala–Pro–Phe(386)	*Asp–Glu–Ala*(428) Asp–Glu–Asn(499) *Asp–Glu–Gln(513)
Tetrapeptides	*Ala–Pro–Leu–Phe*(513) Ala–Pro–Phe–Val(499) Ala–Pro–Phe–Leu(513) Ala–Pro–Phe–Orn(570)	Asp–Glu–Ala–Gly(499) *Asp–Glu–Ala–Ala(513) Asp–Glu–Ala–Pro(525) *Asp–Glu–Ala–Asp*(571) *Asp–Glu–Ala–Lys(626) Asp–Glu–Asn–Gly(570) Asp–Glu–Asn–Leu(626) Asp–Glu–Asn–Thr(628) Asp–Glu–Asn–Lys(697) Asp–Glu–Asn–Try(713) *Asp–Glu–Gln–Val(626) Asp–Glu–Gln–Ser(628) *Asp–Glu–Gln–Orn(697)
Pentapeptides	*Ala–Pro–Leu–Phe–Val*(626) Ala–Pro–Leu–Phe–Ser(628) *Ala–Pro–Leu–Phe–Orn(697) Ala–Pro–Phe–Val–Gly(570) Ala–Pro–Phe–Val–Leu(626) Ala–Pro–Phe–Val–Thr(628) Ala–Pro–Phe–Val–Lys(697) Ala–Pro–Phe–Val–Try(713) Ala–Pro–Phe–Leu–Val(626) Ala–Pro–Phe–Leu–Ser(628) Ala–Pro–Phe–Leu–Orn(697) Ala–Pro–Phe–Orn–Leu(697) Ala–Pro–Phe–Orn–Thr(699) Ala–Pro–Phe–Orn–Asp(713)	Asp–Glu–Ala–Gly–Gly(570) Asp–Glu–Ala–Gly–Leu(626) Asp–Glu–Ala–Gly–Thr(628) Asp–Glu–Ala–Gly–Lys(697) Asp–Glu–Ala–Gly–Try(713) *Asp–Glu–Ala–Ala–Val(626) Asp–Glu–Ala–Ala–Ser(628) Asp–Glu–Ala–Ala–Asn(669) *Asp–Glu–Ala–Ala–Orn(697) Asp–Glu–Ala–Pro–Asp(668) *Asp–Glu–Ala–Asp–Pro*(668) Asp–Glu–Ala–Asp–Leu(698) Asp–Glu–Ala–Asp–Thr(700) Asp–Glu–Ala–Asp–Asp(714) Asp–Glu–Ala–Asp–Glu(728) *Asp–Glu–Ala–Lys–Gly(697) Asp–Glu–Asn–Gly–Leu(697) Asp–Glu–Asn–Gly–Thr(699) Asp–Glu–Asn–Gly–Asp(713) Asp–Glu–Asn–Leu–Gly(697) Asp–Glu–Asn–Thr–Gly(699) Asp–Glu–Asn–Thr–Ala(713) *Asp–Glu–Gln–Val–Gly(697) Asp–Glu–Gln–Ser–Gly(699) Asp–Glu–Gln–Ser–Ala(713)

Table 2 continued

Hexapeptides

Ala–Pro–Leu–Phe–Val–Gly(697)	Asp–Glu–Ala–Gly–Gly–Leu(697)
Ala–Pro–Leu–Phe–Ser–Gly(699)	Asp–Glu–Ala–Gly–Gly–Thr(699)
Ala–Pro–Leu–Phe–Ser–Gly(713)	Asp–Glu–Ala–Gly–Gly–Asp(713)
Ala–Pro–Phe–Val–Gly–Leu(697)	Asp–Glu–Ala–Gly–Leu–Gly(697)
Ala–Pro–Phe–Val–Gly–Thr(699)	Asp–Glu–Ala–Gly–Thr–Gly(699)
Ala–Pro–Phe–Val–Gly–Asp(713)	Asp–Glu–Ala–Gly–Thr–Ala(713)
Ala–Pro–Phe–Val–Leu–Gly(697)	*Asp–Glu–Ala–Ala–Val–Gly(697)
Ala–Pro–Phe–Val–Thr–Gly(699)	Asp–Glu–Ala–Ala–Ser–Gly(699)
Ala–Pro–Phe–Val–Thr–Ala(713)	Asp–Glu–Ala–Ala–Ser–Ala(713)
Ala–Pro–Phe–Leu–Val–Gly(697)	
Ala–Pro–Phe–Leu–Ser–Gly(699)	
Ala–Pro–Phe–Leu–Ser–Ala(713)	

Possible sequences from peptide mixture Ib

Ala(134)	*Asp*(195)	Met(194)
Ala–Pro(231)	Asp–Asn(360)	Met–Asn(359)
	Asp–Glu(358)	Met–Phe(358)
	Asp–Phe(359)	
Ala–Pro–Leu(361)	Asp–Asn–Asn(525)	Met–Asn–Asn(524)
	Asp–Asn–Phe(524)	Met–Phe–Ala(446)
	Asp–Glu–Ala(446)	
	Asp–Phe–Asn(524)	
Ala–Pro–Leu–Asn(526)	Asp–Asn–Asn–Val(641)	Met–Asn–Asn–Val(640)
Ala–Pro–Leu–Glu(524)	Asp–Asn–Phe–Val(640)	Met–Asn–Asn–Tyr(721)
Ala–Pro–Leu–Phe(525)	Asp–Asn–Phe–Tyr(721)	Met–Phe–Ala–Orn(639)
	Asp–Glu–Ala–Asp(595)	Met–Phe–Ala–Tyr(643)
	Asp–Glu–Ala–Orn(639)	
	Asp–Glu–Ala–Tyr(643)	
	Asp–Phe–Asn–Val(640)	
	Asp–Phe–Asn–Tyr(721)	
Ala–Pro–Leu–Asn–Val(642)	Asp–Asn–Asn–Val–Gly(713)	
Ala–Pro–Leu–Asn–Asp(675)		
Ala–Pro–Leu–Glu–Vεl(640)	*Asp–Glu–Ala–Asp–Pro*(692)	
Ala–Pro–Leu–Glu–Tyr(721)	Asp–Glu–Ala–Asp–Ser(716)	
Ala–Pro–Leu–Phe–Val(641)		
Ala–Pro–Leu–Asn–Val–Gly(716)		
Ala–Pro–Leu–Asn–Asp–Gly(749)		
Ala–Pro–Leu–Phe–Val –Gly(715)		

The sequences common for the two derivatives are in italic. * See text. For the simplicity the sequences are indicated non-derivatized. The m/e of the peaks corresponding to the shown sequences are indicated in parenthesis.

Table 3
Possible sequences after comparison of the results from the two derivatives of peptide mixture II.

Gly	Leu
Gly–Pro	Leu–Val
Gly–Pro–Ala	Leu–Val–Glu
Gly–Pro–Ala–Thr	Leu–Val–Glu–Ala
Gly–Pro–Ala–Asp	Leu–Val–Glu–Ser
Gly–Pro–Ala–Met	

For the simplicity the sequences are indicated non-derivatized.

Asp and Gly–Pro–Ala–Met. Compensation for the ^{13}C contribution from Leu–Val–Glu to the peak at m/e 441 in the spectrum of peptide mixture IIa eliminates Gly–Pro–Ala–Met.

Gly–Pro–Ala–Asp cannot be eliminated by pure mass spectrometrical considerations. The probability for the existence of this sequence must, however, be considered very slight as the peaks on which it is based only slightly exceed the threshold level and as no termination or continuation was found.

4. Discussion

The possible sequences of peptide mixture Ia (table 2) marked with an asterix or in italics are all deduced from the real sequence peaks thus demonstrating the insufficient information obtainable by mass spectrometry of a single derivative. In several cases the listed hexapeptide sequences are terminated with Val–Gly. Due to alternatives listed in table 1, group A, the corresponding pentapeptide sequences with C-terminal ornithine are also deduced. The sequence peaks of Asp–Glu–Ala and Ala–Pro–Leu–Phe imply an alternative sequence, Asp–Glu–Ala–Ala, which is of the type listed in D. This latter sequence results in Asp–Glu–Gln (type A). The sequence peaks of Ala–Pro–Leu–Phe–Val and Asp–Glu–Ala imply the alternative sequence Asp–Glu–Ala–Lys (type D). This sequence could also be deduced from Asp–Glu–Ala–Ala–Val (type A).

In peptide mixture IIa the presence of Gly–Pro–Ala–Thr leads to an alternative sequence of type C, Gly–Pro–Try. In peptide mixture IIb Leu–Val–Met is an alternative sequence of type D obtained by combination of the sequence peaks of Gly–Pro–Ala–Thr and Leu–Val.

It is observed that all the alternative sequences of type A, C and D can be eliminated by a comparison of the results obtained with the 2 derivatives. Furthermore, all the alternative sequences due to non-sequence peaks in the mass spectra of peptide mixture I are eliminated and most of them deduced from the mass spectra of peptide mixture II. From table 2 it is seen that most of the alternative sequences due to non-sequence peaks are based upon the peaks on each side of the sequence peaks. Although these peaks normally are not considered by the interpretation of mass spectra of single peptides, they must be taken into account when dealing with peptide mixtures.

The experiments here described thus demonstrate that the interpretation of mass spectra of peptide mixtures may be improved by carrying out the measurements on two different derivatives.

References

[1] F.W. McLafferty, R. Venkataraghavan and P. Irving, Biochem. Biophys. Res. Commun. 39 (1970) 274.
[2] P. Roepstorff, R.K. Spear and K. Brunfeldt, FEBS Letters 15 (1971) 237.
[3] F. Franek, B. Keil, D.W. Thomas and E. Lederer, FEBS Letters 2 (1969) 309.
[4] H.R. Morris, D.H. Williams and R.P. Ambler, Biochem. J. 125 (1971) 189.
[5] E. Vilkas and E. Lederer, Tetrahedron Letters (1968) 3089.
[6] D.W. Thomas, Biochem. Biophys. Res. Commun. 33 (1968) 483.
[7] E. Lederer, Pure Appl. Chem. 17 (1968) 489.