Peptide sequencing by using a combination of partial acid hydrolysis and fast-atom-bombardment mass spectrometry

Francesco DE ANGELIS,* Maurizio BOTTA, Stefano CECCARELLI and Rosario NICOLETTI*

Dipartimento di Chimica, Università di Roma 'La Sapienza', P.le A. Moro 5, I-00185 Roma, Italy

To overcome the limit of the intensity of ions carrying sequence information in structural determinations of peptides by fast-atom-bombardment m.s., we have developed a method that consists in taking spectra of the peptide acid hydrolysates at different hydrolysis times. Peaks correspond to the oligomers arising from the peptide partial hydrolysis. The sequence can then be identified from the structurally overlapping fragments.

INTRODUCTION

The combination of m.s. with hydrolytic procedures in sequence studies of peptides is well known. In the past, partial acid hydrolysis has been used in combination with electron-impact m.s. (Kelley *et al.*, 1975; Nau *et al.*, 1975). By this technique, extensive chemical manipulation of fragments, in order to obtain volatile derivatives, is required; data are collected through g.l.c.-m.s. experiments.

More recently, f.a.b.-m.s. (Barber *et al.*, 1981a) has found wide application in protein chemistry, mainly as the method of choice for the assignment of M_r to unmodified peptides (Williams *et al.*, 1981).

Peptide sequencing by this method is often difficult because of the generally low abundance, or even absence, of ions carrying sequence information, and which may also be obscured by the spectrum background or by peaks from the matrix (Williams, 1983).

In order to obtain more reliable information for the determination of the primary structure of peptides by the f.a.b.-m.s. technique, we have submitted the peptide of interest to partial acid hydrolysis before f.a.b.-m.s. analysis. The number and intensity of ions carrying the sequence information in the spectrum are thus highly increased.

Our method, illustrated here by application to two model peptides, is based on the ability of the f.a.b.-m.s. technique to analyse peptide mixtures by virtue of the large intensity of molecular ions with respect to fragment ions: partial hydrolysis of a peptide in strong acid (Hill, 1965) produces a mixture of oligomers (and free amino acids) whose masses can be easily recognized on the spectrum. We have thus followed the hydrolytic process by taking the f.a.b.-m.s. spectra of the crude hydrolysis mixtures at increasing reaction times, without any prior work-up. The result is a series of f.a.b.-m.s. spectra characterized by intense molecular ions of the oligomers arising from the hydrolysis and therefore typical of the molecule under study.

A preliminary description of this work was presented as part of a plenary lecture at the Symposium 'Recent developments in l.c.-m.s., SFC-m.s. and more advanced ionization methods', Florence, 14–15 June 1985.

EXPERIMENTAL

Peptides were obtained from Fluka and were used without further purification.

The experiments were carried out as follows. Five glass capillary vials, each containing 2 μ l of a solution of 5 nmol of the peptide in aq. 5% (w/v) HCl, were sealed and incubated at 90 °C. The vials were opened at prefixed times (15 min, 30 min, 1 h, 3 h, 6 h and 8 h), and each solution was injected by a microsyringe into a thin layer of glycerol (3 μ l) previously applied to the f.a.b.-m.s. probe tip. Volatile substances were then evacuated in the mass-spectrometer vacuum lock before insertion of the probe into the ion source. Analyses of the mixtures were obtained by using a Kratos MS80 instrument equipped with post-acceleration detector and high-field magnet, a xenon beam being used to ionize the sample.

RESULTS AND DISCUSSION

Fig. 1 shows the positive-ion f.a.b.-m.s. spectrum of [2-D-Ala,5-Leu]enkephalin-NH₂ (structure 1) (Barber *et al.*, 1981b):

Tyr-D-Ala-Gly-Phe-LeuNH₂

Structure 1

with a very intense protonated molecular ion at m/z 569.

Note that ions carrying sequence information, though present, are of very low abundance, and can only be interpreted on the basis of the already known molecular structure.

The positive-ion f.a.b.-m.s. spectra of the mixtures of peptides generated by acid hydrolysis of structure 1 after 15 min and 3 h are reported in Fig. 2. The spectrum in Fig. 2(a) shows the protonated molecular ion shifted by 1 mass unit with respect to the intact molecule, to m/z 570, thus demonstrating the amidated nature of the C-terminal amino acid in the parent peptide.

As expected, peaks associated with high-mass fragments become less abundant as the hydrolysis proceeds, while peaks related to small peptides behave in the opposite way. Of course medium-size fragments (e.g. those formed by consecutive cleavage of two or more bonds) are particularly abundant at intermediate hydrolysis times.

Abbreviation used: f.a.b.-m.s., fast-atom-bombardment mass spectrometry.

* To whom correspondence should be addressed.

Fig. 1. Positive-ion f.a.b.-m.s. spectrum of [2-D-Ala,5-Leu]enkephalin-NH₂ (structure 1)

Peaks indicated by (M) correspond to glycerol clusters.

Fig. 2. Positive-ion f.a.b.-m.s. spectra of the mixtures of peptides generated by acid hydrolysis of [2-D-Ala,5-Leu]enkephalin-NH₂ (structure YAGFL-NH₂) after (a) 15 min and (b) 3 h

Peaks indicated by (M) correspond to glycerol clusters.

Ions at m/z 457 and 407 indicate that two tetrapeptides are generated from structure 1 by loss of leucine and tyrosine respectively (integral masses 113 and 163), one of the two being the *N*-terminal, and the other one the *C*-terminal residue.

Of the three possible tripeptides, two give rise to peaks at m/z 336 (Gly-Phe-Leu, very abundant) and 310 (Tyr-Ala-Gly). The third one (Ala-Gly-Phe), which should be formed by cleavage at both termini, is not present in the reaction mixture.

The most abundant peak (m/z 279) in the spectrum in Fig. 2(a) is relevant to the sequence Phe-Leu. It is worth

noting that this dipeptide, as expected, is very resistant to hydrolysis, being the only peptide present in the mixture after prolonged reaction time (8 h).

The f.a.b.-m.s. spectrum of the 3 h hydrolysis mixture (Fig. 2b) shows some enhanced peaks, namely the protonated molecular ions for the dipeptide Tyr-Ala at m/z 253 and peaks corresponding to the free amino acids. Spectra obtained from the 6 h and 8 h hydrolysis mixtures were coherent with the general trend indicated above (i.e. disappearance of the largest peptides).

The interpretation of all f.a.b.-m.s. spectra collected during the peptide hydrolysis leads to a self-consistent

Table 1. Oligomers present in the hydrolysis [5% (w/v) HCl] mixture of Substance P (structure 2) determined by f.a.b.-m.s.

Amidated peptides and those containing glutamine residues were computed together with their carboxylic analogues.

Oligomer	(M+H) ⁺	Hydrolysis time...						Abundance			
		15 min	30 min	1 h	2 h	3 h	4 h	6 h			
Arg-Pro-Lys-Pro-Glu-Glu-Phe-Phe-Gly-Leu-Met	1350	++	++	++	++	+					
Arg-Pro-Lys-Pro-Glu-Glu-Phe-Phe-Gly-Leu	1219	++	++	++	++	+					
Pro-Lys-Pro-Glu-Glu-Phe-Phe-Gly-Leu-Met	1194	++	++	++	++	+					
Arg-Pro-Lys-Pro-Glu-Glu-Phe-Phe-Gly	1106	++	++	++	++	+	+				
Lys-Pro-Glu-Glu-Phe-Phe-Gly-Leu-Met	1097	++	++	++	++	+					
Arg-Pro-Lys-Pro-Glu-Glu-Phe-Phe-Gly-Leu	1063	++	++	++	++	+	+				
Pro-Lys-Pro-Glu-Glu-Phe-Phe	1049	++	++	++	++	+					
Pro-Glu-Glu-Phe-Phe-Gly-Leu-Met	969	++	++	++	++	+					
Lys-Pro-Glu-Glu-Phe-Phe-Gly-Leu	966	++	++	++	++	+					
Pro-Lys-Pro-Glu-Glu-Phe-Phe-Gly	950	++	++	++	++	+					
Arg-Pro-Lys-Pro-Glu-Glu-Phe	902	++	++	++	++	+	+				
Glu-Glu-Phe-Phe-Gly-Leu-Met	872	++	++	++	++	+					
Lys-Pro-Glu-Glu-Phe-Phe-Gly	853	+	++	++	++	+	+				
Pro-Glu-Glu-Phe-Phe-Gly-Leu	838	++	++	++	++	+	+				
Lys-Pro-Glu-Glu-Phe-Phe	796	++	++	++	++	+	+				
Arg-Pro-Lys-Pro-Glu-Glu	755	++	++	++	++	+	+				
Glu-Phe-Phe-Gly-Leu-Met	743	++	++	++	++	+	+				
Pro-Glu-Glu-Phe-Phe-Gly	725	++	++	++	++	+	+				
Pro-Glu-Glu-Phe-Phe	668	++	++	++	++	+	+				
Lys-Pro-Glu-Glu-Phe	649	++	++	++	++	+	+				
Glu-Glu-Phe-Phe-Gly	628	++	++	++	++	+	+				
Arg-Pro-Lys-Pro-Glu	626	++	++	++	++	+	+				
Phe-Phe-Gly-Leu-Met	614	++	++	++	++	+	+				
Glu-Phe-Phe-Gly-Leu	612	++	++	++	++	+	+				
Pro-Lys-Pro-Glu-Glu	599	++	++	++	++	+	+				
Glu-Glu-Phe-Phe	571	++	++	++	++	+	+				
Pro-Glu-Glu-Phe	521	++	++	++	++	+	+				
Glu-Phe-Phe-Gly	499	++	++	++	++	+	+				
Phe-Phe-Gly-Leu	483	++	++	++	++	+	+				
Phe-Gly-Leu-Met	467	++	++	++	++	+	+				
Glu-Phe-Phe	442	++	++	++	++	+	+				
Glu-Glu-Phe	424	++	++	++	++	+	+				
Arg-Pro-Lys	400	++	++	++	++	+	+				
Lys-Pro-Glu	373	++	++	++	++	+	+				
Phe-Gly-Leu	336	++	++	++	++	+	+				
Gly-Leu-Met	320	++	++	++	++	+	+				
Phe-Phe	313	++	++	++	++	+	+				
Leu-Met	263	++	++	++	++	+	+				
Gly-Leu	189	++	++	++	++	+	+				

picture of the molecule under investigation. Peaks related to peptides generated during the hydrolysis process can be easily recognized, thus providing a complete sequence information. The only remaining ambiguity is the *N*-terminal $\rightarrow$ *C*-terminal, or the reverse, sequence of the molecule, which can be easily solved by analysing the f.a.b.-m.s. spectrum of the intact molecule.

The strategy presented here is also amenable to the study of medium-size peptides. In the positive-ion f.a.b.-m.s. spectra of the hydrolysis samples of the neuropeptide Substance P (structure 2):

Structure 2

ions carrying sequence information are numerous and intense.

Data obtained from the spectra taken at increasing hydrolysis times are analytically presented in Table 1, giving a clear and complete picture of the amino acid sequence in a more complex peptide. However, some major points deserve comment.

First, abundant $(M+H)^+$ at m/z 1350 [3 atomic mass units higher than $(M+H)^+$ of the parent peptide] was readily produced (see Table 1: 15 min-hydrolysis data), this being connected with the ready hydrolysis of the CONH_2 groups (the two glutamine and the methionine amide residues; for the possibility to distinguish lysine from glutamine by this hydrolytic procedure, see below).

It was also apparent from these series of spectra that 39 peptides, identified through their $(M+H)^+$ peaks, were obtained of the possible 64 (Table 1). $(M+H)^+$ identification of fragment peptides was also greatly aided by comparing the f.a.b.-m.s. spectra of the mixtures obtained in sequential times of hydrolysis, in a sort of kinetic analysis of the process (i.e. increase or decrease in the abundance of selected peaks while the hydrolysis proceeds). In the analysis, multiplets formed by peaks differing by 1 atomic mass unit were recognized as characteristic of peptides containing one or more amidated residues. In this way it is possible to distinguish lysine-containing from glutamine-containing peptides (i.e. glutamine-containing peptides are shifted to higher mass by 1 atomic mass unit during the hydrolysis), which, because of their equal nominal mass, are not distinguishable in the conventional f.a.b.-m.s. [The only method described to remove the mass degeneracy of lysine and glutamine is somewhat troublesome and relies on *N*-acetylation of the former (Williams *et al.*, 1982). The Leu/Ile ambiguity is obviously not solvable by this method, but requires a more sophisticated procedure (Aubagnac *et al.*, 1985).]

The data obtained by the combined use of partial hydrolysis and f.a.b.-m.s. analysis of Substance P (as well as in the example presented above) appear to be more

than sufficient to reconstruct the complete peptide sequence. In general, small and medium-size fragments in fact contain the overlapping sequences necessary for determination of the order of the amino acids in the parent peptide. On the other hand, mass differences of large fragments give supplementary amount of information on the sequence at both ends of the molecule.

We wish to point out that, with the advent of f.a.b.-m.s., peptide partial hydrolysis, which in the past suffered from the limitation due to fragment identification, may become an alternative to other sequencing techniques. In particular, the strategy presented here has the advantage, also over the g.l.c.-m.s. approach, of rapidity and simplicity of operation. Mixtures arising from partial acid hydrolysis can be directly analysed without any derivative formation and/or separation procedure.

The two examples presented in this paper show that the kinetic analysis by f.a.b.-m.s. of the partial hydrolysates of small and medium-size peptides is a sensitive technique that directly provides a complete sequence map.

F.a.b.-m.s. has also been used in combination with partial enzymic digestion, with exopeptidases (Williams, 1983; Self & Parente, 1983) or other proteolytic enzymes (Morris *et al.*, 1983), in order to obtain enhancement of selected sequence-information-carrying ions on the spectrum. Our method may be considered in some instances as alternative to the method just mentioned, but it has the definite advantage of applicability to blocked peptides or those comprising unusual or D-amino acids (Williams *et al.*, 1983).

REFERENCES

- Aubagnac, I. L., El Amrani, B., Devienne, F. M. & Combarieu, R. (1985) *Org. Mass Spectrom.* **20**, 428–429
- Barber, M., Bordoli, R. S., Sedgwick, R. D. & Tyler, A. N. (1981a) *J. Chem. Soc. Chem. Commun.* 325–327
- Barber, M., Bordoli, R. S., Garner, G. V., Gordon, D. B., Sedgwick, R. D., Tetler, L. W. & Tyler, A. N. (1981b) *Biochem. J.* **197**, 401–404
- Hill, R. L. (1965) *Adv. Protein Chem.* **20**, 37–61
- Kelley, J. A., Nau, H., Förster, H.-J. & Biemann, K. (1975) *Biomed. Mass Spectrom.* **2**, 313–325
- Morris, H. R., Panico, M. & Taylor, G. W. (1983) *Biochem. Biophys. Res. Commun.* **117**, 299–305
- Nau, H., Förster, H.-H., Kelley, J. A. & Biemann, K. (1975) *Biomed. Mass Spectrom.* **2**, 326–339
- Self, R. & Parente, A. (1983) *Biomed. Mass Spectrom.* **10**, 78–82
- William, D. H. (1983) *Gazz. Chim. Ital.* **113**, 27–35
- Williams, D. H., Bradley, C., Bojesen, G., Santikarn, S. & Taylor, L. C. E. (1981) *J. Am. Chem. Soc.* **103**, 5700–5704
- Williams, D. H., Bradley, C. V., Santikarn, S. & Bojesen, G. (1982) *Biochem. J.* **201**, 105–117
- Williams, D. H., Santikarn, S., De Angelis, F., Smith, R. J., Reid, D. G., Oelrichs, P. B. & MacLeod, J. K. (1983) *J. Chem. Soc. Perkin Trans. 1* 1869–1878