EXOPEPTIDASE DIGESTION IN COMBINATION WITH FIELD DESORPTION MASS SPECTROMETRY FOR AMINO ACID SEQUENCE DETERMINATION

A. TSUGITA, R. VAN DEN BROEK and M. PRZYBYLSKI

European Molecular Biology Laboratory, Meyerhofstr. 1, 6900 Heidelberg and **Institut für Organische Chemie, Universität Mainz, J. J. Becher-Weg 18-20, 6500 Mainz, FRG*

Received 29 October 1981

1. Introduction

Numerous studies have been devoted in the last years to the development of mass spectrometric methods for the sequence determination of peptides [1,2]. Most advanced among this work has been so far the rigorous chemical derivatization of oligopeptides to achieve sufficient volatility for the application of conventional (electron impact, EI) mass spectrometry [2–4]. For example, the analysis by gas chromatography–mass spectrometry (GC–MS) of mixtures of oligopeptide fragments derivatized after chemical or enzymatic hydrolysis of polypeptides has been successfully used for sequence determinations [2,5]. Major limitations of this approach are that only small peptide derivatives are amenable to GC–MS, and the relatively large amount of original peptide needed. More recently, mass spectra of underivatized peptides have been obtained by field desorption (FD) mass spectrometry [6], a technique that does not require any volatility of the sample to be ionized. It was shown that FD mass spectra of oligopeptides usually exhibit molecular ions (e.g., protonated, MH^+ ions) of high intensity, with relatively little fragmentation [7,8]. While studies of thermally or collision-induced fragmentation in FD spectra of peptides [9–11] so far indicate only limited utility for obtaining structural information, the possibility of analysing peptide mixtures by means of the predominant molecular ions has been suggested [12]. FD mass spectrometry in combination with Edman degradation–subtractive analysis has been reported [13].

Digestion with exopeptidases is one of the classical methods which has been widely used for peptide sequencing [14,15]. The complexity resulting from the non-linear kinetics, the relatively low purity of the available enzymes (e.g., contamination with endopeptidase or different types of exopeptidase activity), and the problem of determining repetitive amino acid residues have limited their application so far. Attempts to overcome these difficulties have been made such as the addition of endopeptidase inhibitors [14,15] and solid-phase sequencing with exopeptidases [16]. However, the characteristic feature of the inhomogeneous kinetics of exopeptidase digestion which, in principle, should match well the capability of FD mass spectrometry to analyse peptide mixtures prompted our interest to test the combination of both methods for peptide sequencing. Here, fast and unequivocal sequence determinations were obtained for 2 oligopeptides by direct FD mass spectral analysis of aliquots taken from partial amino- and carboxypeptidase digests. The finding that FD mass spectra of exopeptidase digests provided reliable sequence data without any separation step required, and with amounts of material at the sub-nanomole level indicate that this combination may prove to be a valuable approach among the instrumentarium for the sequence determination of peptides.

2. Materials and methods

Leu–Trp–Met–Arg–Phe–Ala, Glu–Thr–Tyr–Ser–Lys and peptides with corresponding partial sequences were obtained from Serva (Heidelberg). The enzymes were from the following sources: Carboxypeptidase A and B (Cpase A,B, both diisopropylfluorophosphate-treated), Worthington Biochemicals (Freehold NJ); aminopeptidase M (Apase M), PL-Biochemicals (Milwaukee WI). Cpases were dialysed against 0.1 M pyridinium acetate as in [16]. Apase was dia-

Published by Elsevier/North-Holland Biomedical Press
00145793/82/0000–0000/$02.75

lysed against water at 4°C. Three-times distilled deionized water and acetic acid of highest available purity (Merck, Darmstadt) were used as solvents for FD mass spectrometry.

A Durrum D500 amino acid analyzer was used for amino acid determinations directly from partly digested mixtures [16,17]. FD mass spectrometry was performed with a Finnigan MAT 312 spectrometer/SS 188 data system. The combined FD/EI ion source of the instrument was equipped with a precisely adjustable direct insertion probe [18]. Activated carbon field ion emitters with average needle lengths of 30–40 μm were prepared by the high-temperature activation procedure [19]. A three-dimensional micromanipulator [7,18] was used for sample loading onto the emitter with a microliter syringe. An emitter heating current (EHC) programmer (Linden, Bremen) was used for heating the emitter at linear rates or in the emission-controlled mode [20]. Accumulation of spectra acquired during the temperature-programmed desorption was performed by the data system [18]. Mass calibration was carried out by EI spectra of fomblin (perfluoro-polypropylether, PCR, Gainesville FL). Other instrumental conditions were as in [18].

Enzymatic digestions were carried out with 5 nmol substrate and 5 μg enzyme in 0.1 M pyridinium acetate buffer (pH 8.2) for Cpases A and B, and pH 7.0 for Apase M. Other reaction conditions are given in the figure legends and in table 2. Reactions were terminated at various time intervals by heating to 100°C for 1 min, and the reaction mixture was lyophilized. An aliquot of 1 nmol was subjected to amino acid analysis. The remaining aliquot was dissolved in 50 μl water, and 1 μl acetic acid was added. This solution (2 μl) was loaded on the FD emitter. After evaporation of the solvent, the emitter was introduced in the ion source and heated at a rate of 4 mA/min. Acquisition of spectra by the data system was started at ~5 mA EHC with a scan rate of 2 s/decade, and continued up to a maximum EHC of 30 mA. The desorption profile consisting typically of 60 single scans over 17–25 mA was then accumulated into a resultant spectrum which was evaluated for sequence determinations.

3. Results

As typically found in previous FD mass spectral studies of oligopeptides [7,8], the FD spectra of the hexapeptide, Leu–Trp–Met–Arg–Phe–Ala and the corresponding partial sequences readily yielded molecular ions of predominant abundance (table 1). All peptides, when prepared at conditions as those used for the enzymatic digestions showed a homogeneous series of MH^+ ions. No formation of molecular cluster ions such as $[MA]^+$ to be expected in the presence of

Table 1
FD mass spectra of Leu–Trp–Met–Arg–Phe–Ala, Glu–Thr–Tyr–Ser–Lys and related peptides with partial sequences

Peptide	M_r	m/e, MH^+ (%)[b]	EHC (mA)	m/e, Fragment ions[a] (%)[b]
Leu–Trp–Met–Arg–Phe–Ala	823[c]	824(100)	18–27	751(3) 673(8) 544(12) 509(4) 447(12) 299(8) 225(23)
Leu–Trp–Met–Arg–Phe	752[c]	753(100)	17–25	709(6) 605(3) 604(8) 545(7) 544(10) 447(7) 412(2)
Leu–Trp–Met–Arg	604	605(100)	18–26	561(8) 545(6) 492(6) 412(3) 306(10) 289(8)
Leu–Trp–Met	448	449(100)	17–25	309(5) 289(3)
Leu–Trp	317	318(100)	15–24	299(41) 205(3)
Met–Arg–Phe–Ala	523	524(100)	18–24	491(9) 451(6) 375(12) 305(16) 304(18) 219(4) 218(3)
Arg–Phe–Ala	392	393(100)	18–24	320(5) 319(7) 253(7) 211(8)
Phe–Ala	236	237(100)	16–22	–
Glu–Thr–Tyr–Ser–Lys	626	–	20–25	609, $[MH-H_2O]^+$ (100) 591(98) 480(6) 375(7)
Glu–Thr–Tyr–Ser	498	–	19–24	481, $[MH-H_2O]^+$ (100) 463(45) 375(5)
Glu–Thr–Tyr	411	–	18–23	394, $[MH-H_2O]^+$ (15) 393(25) 376(55) 375(100) 247(8)
Tyr–Ser–Lys	400	401(100)	17–25	400(20) 277(12) 276(6)

[a] Fragment ions >m/e 200 are listed
[b] Relative intensity
[c] Nominal M_r with consideration of mass defect according to the mass calibration used (section 2)

The peak intensities given refer to spectra accumulated from single scans at the EHC range indicated (section 2)

alkali metal ions A^+ [21] was observed. The formation of fragment ions was notable at ≳20 mA heating currents. However, the intensity of fragments was quite low even in spectra accumulated at ⩽28–30 mA EHC. A preferred fragmentation pathway is the thermal cleavage of amino-alkyl linkages leading to ions with carboxamide end-groups [7,10], such as at Ala (*m/e* 751), Phe (*m/e* 604) and Arg (*m/e* 447). While providing some sequence information from the N-terminus, this fragmentation does not interfere with the identification of MH^+ ions of the C-terminal partial sequences to be expected by Cpase digestion (table 1). Other fragment ions in the spectra of these peptides are difficult to rationalize and do not correspond to simple cleavages.

FD mass spectra of the peptide mixtures produced by partial digestion of Leu–Trp–Met–Arg–Phe–Ala with Apase M are shown in fig.1a–d. With the M_r for the complete peptide and the amino acid composition at hand, the partial sequence, Leu–Trp–Met– is immediately apparent by subtractive analysis from the MH^+ ions at *m/e* 824, 711, 524 and 393 (fig.1a–c). Note that the nominal mass difference for the Trp residue between *m/e* 711 and 524 is to be incremented by 1 mass unit according to the hydrogen mass defect above about *m/e* 700 due to the mass calibration employed. The identification of the MH^+ sequence ions is very clear, and their relative abundance distribution is, qualitatively, well consistent with time course and temperature of the enzymatic digestion (see below). Fig.1d shows that all the possible sequence information, Leu–Trp–Met–Arg– is revealed by a single FD spectrum obtained from a mixture of Apase digest aliquots (*m/e* 237 for Phe–Ala; *m/e* 165 and 89 for Phe and Ala). Digestion of the present peptide with Cpases A and B and FD mass spectral analysis of aliquots at various times or mixtures of them yielded the same definite sequence data (fig.1e,f; table 2). MH^+ ions at *m/e* 753, 605, 449 and 318 correspond to the partial sequence, –Met–Arg–Phe–Ala–OH and are thus complementary to the Apase data. In the Cpase digests, the ions at *m/e* 753 (Leu–Trp–Met–Arg–Phe) were found with very low abundances in all but the initial digest aliquots, which reflects the high rate of release of the second residue Phe after release of the C-terminal Ala by this enzyme (fig.1f). Considering the mass of this peak as a sequence ion, the mass difference of 218 to the next sequence ion, *m/e* 605 does not correspond to any amino acid residue but to combinations of two amino acids such

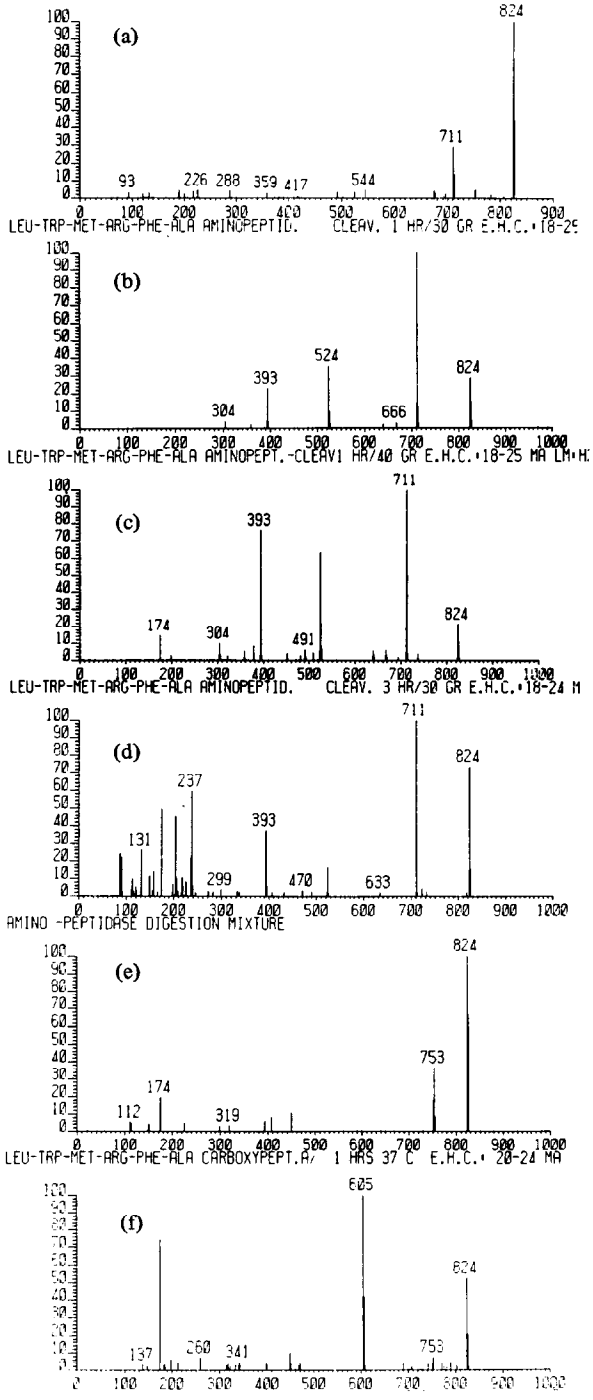

Fig.1. FD mass spectra of aliquots obtained from partial digestion of Leu–Trp–Met–Arg–Phe–Ala with Apase M and Cpase A: (a) Apase M, 1 h, 30°C. (b) Apase M, 1 h, 40°C; (c) Apase M, 3 h, 30°C; (d) mixture of equal volumes of digests (a)–(c) plus Apase M digest aliquot, 6 h, 30°C; (e) Cpase A, 1 h, 37°C; (f) Cpase A, 2 h, 37°C. Amounts containing ~0.2 nmol substrate mixtures were loaded on the FD emitter (section 2).

Table 2
Percentage of released amino acids *vs* MH^+ sequence ions observed by FD mass spectrometry upon digestion of Leu–Trp–Met–Arg–Phe–Ala with Cpase A,B, and Apase M

	Apase:823[a] Leu		710[a]		Trp 317		Met 448		Arg 604		Phe 752[a]		823:Cpase Ala–OH
					523		392		236		165,89		
Enzyme	Time h (°C)	% digest[b]	MH^+ ion[c]	% digest[b]	MH^+ ion[c]	% digest[b]	MH^+ ion[c]	% digest[b]	MH^+ ion[c]	% digest[b]	MH^+ ion[c]	% digest[b]	MH^+ ion[c]
Cpase A	1 (37)	0	–	0	–	0	–	0	–	18	753	28	824
Cpase A + B	1 (37)	0	–	0	318	42	449	34	605	66	753	88	824
Cpase A + B	6 (37)	0	–	0	318	70	449	80	605	84	–	88	–
Apase M[d]	1/6 (30)	85	824	80	711	38	524	4	–	0	–	0	–
Apase M[d]	1 (30)	95	–	90	711	90	524	85	393	18	237	20	165
Apase M[d]	3 (30)	100	–	100	–	100	–	98	–	83	237	84	165,89

[a] Nominal M_r-value with consideration of mass defect due to the mass calibration used (section 2)
[b] Expressed as percentage of free amino acid recovered
[c] MH^+ ion observed for residual peptide from which amino acid is released
[d] Enzyme preparation different from that used for experiments shown in fig.1

as Ser + Met or Cys + Asp which, however, are ruled out by the amino acid composition. With regard to the possible complication of missing sequence ions of low abundance, some experiments were made to correlate the release of amino acids as determined by amino acid analysis with the FD mass spectral detection of MH^+ ions of residual fragments. These data (table 2) show that in all cases the presence of MH^+ ions of residual peptide fragments was indeed qualitatively consistent with the release of a significant portion of the corresponding amino acid. No satisfactory quantitative correlation with the relative intensity distribution of MH^+ sequence ions was obtained so far, although the overall correspondence of peak intensities with the time course of the enzymatic digestion is evident (cf. fig.1). Also, molecular ions of the free amino acids were not detected in all cases where the corresponding MH^+ ion of the residual peptide fragment was present. This effect may be due to several factors, such as the inhomogeneous kinetics of release of individual amino acids, and differences in the relative FD ionization efficiencies among the substrate mixtures.

A most interesting result from this study was the absence of any qualitative and quantitatively significant interference by the exopeptidases used. Almost all additional ions of notable abundance were recognized as fragments of the oligopeptide substrates. FD spectra of Cpase and Apase blank samples did not show any significant ions of higher masses in the temperature range of interest (~15–28 mA EHC). Ions indicative of thermal decomposition were observed at high temperatures (>30 mA EHC) but appeared predominantly in the lower mass range ($m/e < 200$). Also, no significant reduction of the FD sensitivity for the MH^+ ions ($1–4 \times 10^{-11}$ C/μg for MH^+ of Leu–Trp–Met–Arg–Phe–Ala) was observed in spectra of enzyme samples spiked with the present oligopeptides.

Since any successful sequence analysis by the present approach depends critically on the certainty of M_r determination, another model peptide, Glu–Thr–Tyr–Ser–Lys, was studied which did not yield a molecular ion. As reported previously [22] (cf. table 1), FD spectra of peptides with α-Glu at the N-terminus do not show any significant molecular ions because of a predominant thermal elimination of water and formation of pyroglutamyl residues. This effect leads to $[MH–H_2O]^+$ ions of high abundance. In the FD spectrum of Glu–Thr–Tyr–Ser–Lys

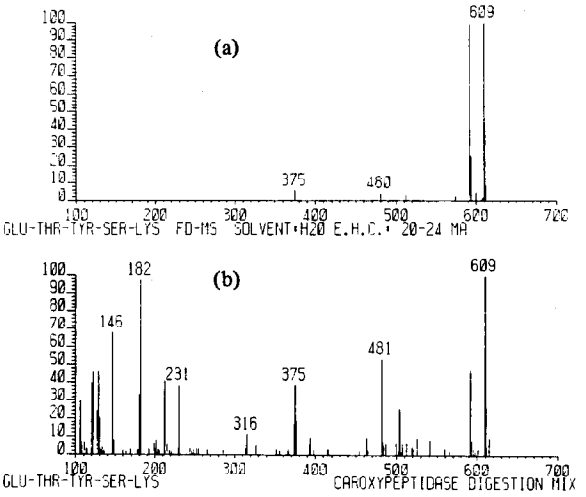

Fig.2. FD mass spectra of (a) Glu–Thr–Tyr–Ser–Lys and (b) a mixture of equal volumes of digest aliquots after Cpase digestion at 1 h, 3 h and 6 h, 37°C. The sample amount loaded on the emitter contained ~0.4 nmol substrate mixture.

(fig.2a), the $[MH–H_2O]^+$ ion (m/e 609) appears as base peak and is accompanied by an abundant ion at m/e 591 according to the elimination of a second molecule of water (from the Thr residue). This finding may be explained by the lability of the particular arrangement, α-Glu–Thr–, since the spectrum of Tyr–Ser–Lys (table 1) showed an intensive MH^+ ion without any appreciable fragmentation. Digestion of the pentapeptide with Cpase A combined with FD mass spectral analysis demonstrated that the complete sequence information is obtained even in the case of this complication, given the additional information about the presence of N-terminal α-Glu. The spectrum of a mixture of partial digest aliquots (fig.2b) contains all $[MH–H_2O]^+$ ions at m/e 609, 481, 394 and 231 associated with the ions due to the additional elimination of H_2O (m/e 591, 463, 375/376 and 213). Furthermore, analysis of the lower mass region shows the presence of free Tyr (m/e 181/182), Ser (m/e 105) and Lys (m/e 146).

4. Discussion

The ease and certainty with which complete sequence data were obtained with the peptides used here indicates that the specific combination of exopeptidase digestion and FD mass spectrometry may develop as a powerful approach for amino acid sequence determinations. The requirement that the

data obtained from MH^+ sequence ions plus molecular ions of released amino acids, and that of total and released amino acid analyses be consistent with each other should provide a high degree of confidence. The sequence determination may be obtained by a single proper digest aliquot (or appropriate mixtures), but the identification of the sequencing molecular ions will be facilitated, in any problem case, by comparison of spectra of aliquots through the time course of enzymatic digestion, and by enzymatic digestion from the other terminus. This method provides reliable sequence data in cases where exopeptidase digestion and conventional analysis typically fail, such as repetitive amino acid residues, or very inhomogeneous kinetics. The FD mass spectral analysis is rather tolerant to intrinsic mis-information from exopeptidase digestion data that may confuse the analysis of released amino acids, such as autodigestion of the enzyme (producing mainly amino acids) and endopeptidase activity present (decreasing immediately the substrate size). The two crucial prerequisites for the successful FD mass spectral analysis are:

(i) The detection of all peptide fragments in a particular digest mixture;
(ii) The production of a homogeneous set of molecular ion species.

Precise control of the emitter temperature [12] and spectra accumulation techniques as applied in this study are essential to meet the first requirement, while the use of pyridinium acetate as a buffer solvent was found quite advantageous to obtain exclusively MH^+ ions. Principal limitations for the FD mass spectral sequencing are identical amino acid M_r-values (Leu/Ile and Gln/Lys), and the M_r-limit of the peptide amenable to FD mass spectrometry. FD spectra of ~100 model peptides containing proteinogenic amino acids obtained in our laboratory so far suggest an upper limit of ~2000 for a relatively safe and easy M_r determination, with our present instrumentation. Major advantages of the combination, exopeptidase digestion and FD mass spectrometry are high sensitivity, simplicity and speed of analysis, and the absence of interfering ions from the enzyme. The latter finding may prove advantageous in comparison with other methods of stepwise cleavage producing by-products by chemical reactions, such as Edman degradation combined with FD mass spectrometry [12,13]. Work on the sequence determination of several cleavage products from T4 bacteriophage ligase using this method is in progress.

Acknowledgements

This work was supported by a grant from the Deutsche Forschungsgemeinschaft to M. P. The expert technical assistance of Mrs I. Dietrich is gratefully acknowledged.

References

[1] Biemann, K. (1972) in: Biochemical Applications of Mass Spectrometry (Waller, G. R. ed) pp. 405–428, Wiley, New York.
[2] Bieman, K. (1978) Pure Appl. Chem. 50, 149–158.
[3] Kelley, J. A., Nau, H., Förster, H. J. and Biemann, K. (1975) Biomed. Mass Spectrom. 2, 313–325.
[4] Morris, H. R., Dickinson, R. J. and Williams, D. H. (1973) Biochem. Biophys. Res. Commun. 51, 247–255.
[5] Nau, H. (1976) Angew. Chem. Int. Ed. 15, 75–87.
[6] Beckey, H. D. and Schulten, H. R. (1975) Angew. Chem. Int. Ed. 14, 403–415.
[7] Schulten, H. R. (1977) Methods Biochem. Anal. 24, 368–413.
[8] Przybylski, M. and Voelter, W. (1981) in: Structure and Activity of Natural Peptides (Weitzel, G. and Voelter, W. eds) pp. 141–165, De Gruyter, Berlin.
[9] Asante-Poku, S., Wood, G. W. and Schmidt, D. E. jr (1975) Biomed. Mass Spectrom. 2, 121–125.
[10] Przybylski, M., Lüderwald, I., Kraas, E., Voelter, W. and Nelson, S. D. (1979) Z. Naturforsch. 34b, 736–743.
[11] Matsuo, T., Matsuda, H., Katakuse, I., Shimonishi, Y., Higuchi, T. and Kubota, E. (1981) Anal. Chem. 53, 416–421.
[12] Schulten, H. R. (1979) Int. J. Mass Spectrom. Ion Phys. 32, 97–283.
[13] Shimonshi, Y., Hong, Y., Kitagishi, T., Matsuo, T., Matsuda, T. and Katakuse, T. (1980) Eur. J. Biochem. 112, 251–264.
[14] Ambler, R. P. (1972) Methods Enzymol. 25, 262–272.
[15] Light, A. (1972) Methods Enzymol. 25, 253–262.
[16] Tsugita, A. and Van den Broek, R. (1980) in: Methods in Peptide and Protein Sequence Analysis (Birr, C. ed) pp. 359–369, Elsevier/North-Holland, Amsterdam, New York.
[17] Isobe, T., Yangida, M., Boosman, A. and Tsugita, A. (1978) J. Mol. Biol. 125, 339–356.
[18] Przybylski, M., Preiss, J., Dennebaum, R. and Fischer, J. (1981) Biomed. Mass Spectrom. in press.
[19] Schulten, H. R. and Beckey, H. D. (1972) Org. Mass Spectrom. 6, 885–895.
[20] Schulten, H. R. and Nibbering, N. M. M. (1977) Biomed. Mass Spectrom. 4, 55–61.
[21] Rinehart, K. L. jr, Cook, J. C. jr, Meng, H., Olson, K. L. and Pandey, R. C. (1977) Nature 269, 832–833.
[22] Lüderwald, I., Przybylski, M., Ringsdorf, H., Silberhorn, D., Zech, K. and Voelter, W. (1978) Adv. Mass Spectrom. 7B, 1513–1517.