CHEMISTRY LETTERS, pp. 1369-1372, 1979. © The Chemical Society of Japan 1979

A NEW METHOD FOR SEQUENCE DETERMINATION OF PEPTIDE MIXTURES BY EDMAN-DEGRADATION AND FIELD DESORPTION MASS SPECTROMETRY

Yasutsugu SHIMONISHI*, Yeong-Man HONG,
Takekiyo MATSUO#, Itsuo KATAKUSE## and Hisashi MATSUDA#
Institute for Protein Research, Osaka University
Yamada-kami, Suita, Osaka 565
#Institute of Physics, College of General Education, Osaka University
##Department of Physics, Faculty of Science, Osaka University
Toyonaka, Osaka 560

A new method is described for sequencing of peptide mixtures by Edman-degradation and field desorption mass spectrometry. The principle is realized by measurement of the mass numbers of molecular or quasi-molecular ions of peptides and their fragments degraded by the Edman method and calculation of the mass difference.

This paper reports a novel method for sequencing of polypeptides by Edman-degradation[1] and field desorption mass spectrometry[2] (FD-MS). The method is specially referred to sequencing of peptides in mixtures obtained by enzymic or chemical cleavage of polypeptides and proteins, using the FD ionization technique that gives almost exclusively the molecular (M^+) or quasi-molecular ($[M+H]^+$) ion peaks of free peptides.[3] Edman-degradation is a very useful method for sequencing of peptides and proteins, but it has the disadvantage that the peptides obtained from polypeptides and proteins by enzymic or chemical cleavages must be separated by time-consuming procedures. Gas chromatography-mass spectrometry[4,5] (GC-MS) is also very useful for analysis of mixtures after they have been separated by GC, but this method involves considerable chemical processing to convert peptides to volatile derivatives, and hence the complexity of the test preparations increases with increase in size of peptides examined.[6] To overcome these problems, we developed a new method for sequencing of peptide mixtures, based on the following principle: (1) measurement of m/z of M^+ or $[M+H]^+$ [7] of individual peptides in the mixture by FD-MS,[3] (2) Edman-degradation of peptides present in the mixture, (3) identification of the amino acid phenylthiohydantoins (PTHs) released in (2) by thin layer chromatography[8] (TLC) or FD-MS,[9] (4) measurement of m/z of M^+ or $[M+H]^+$ of peptides lacking N-terminal amino acid residues (R_ns) obtained in (2), by FD-MS, and (5) subtraction of one of the residual weights calculated from the PTHs in (3) from given one of m/z of M^+ or $[M+H]^+$ of peptides in (1) to obtain one of m/z of peptides lacking R_ns in (4). Thus, we can determine R_n for given one of peptides in mixture and also R_ns for other peptides. Repeating the sequence of these operations, we can determine the amino acid sequences of individual peptides

in a mixture.

The possibility of sequence determination of peptides using Edman-degradation and FD-MS has not been investigated except in the case of a peptide from ribosomal protein L17, in which released PTHs were identified.[9] Instead of identification of PTH, we[10] attempted to measure m/z of M^+ or $[M+H]^+$ of peptide and its peptide fragment degraded by the Edman method and to determine R_n from the difference between the values. Recently, some of us[6,11] found that careful measurements of $[M+H]^+$ and multiply charged ions $[M+nH]^{n+}$ provided difinite information about the molecular weight (M) of a peptide. First, by this method we tried to measure m/z of $[M+H]^+$ of the synthetic single peptide Asp^1, Ile^5-angiotensin I[12] and its successively degraded peptide fragments; we confirmed that the values were 1296, 1181, 1025, 926, 763, 650, 513, 416 and 269, in order of degradation. From these values we determined the sequence Asp(115)-Arg(156)-Val(99)-Tyr(163)-Ile(113)-His(137)-Pro(97)-Phe(147)- simply by subtraction. PTH of Ile was identified at position 5 from the N-terminus by TLC, although Ile has the same residual weight as Leu. Furthermore, [2-anilino-5]-thiazolinone derivative (intermediate of PTH) of His with m/z 273 of M^+ was detected in the material with m/z 269 after degradation. Therefore, the C-terminal sequence was -His(137)-Leu(113)-OH. Thus, the method was confirmed to be useful for sequencing of peptide. The sequence of bradykinin[12] was ascertained in the same way as that of Asp^1, Ile^5-angiotensin I, as reported in detail elsewhere.[14]

Then, we studied the sequence determination of peptide mixtures derived from polypeptides by enzymic digestion. As a typical example, glucagon[15] was digested with chymotrypsin and TPCK-treated trypsin[16] at pH 7.8, as described in the literature[15] with some modifications. The FD mass spectra of the resulting

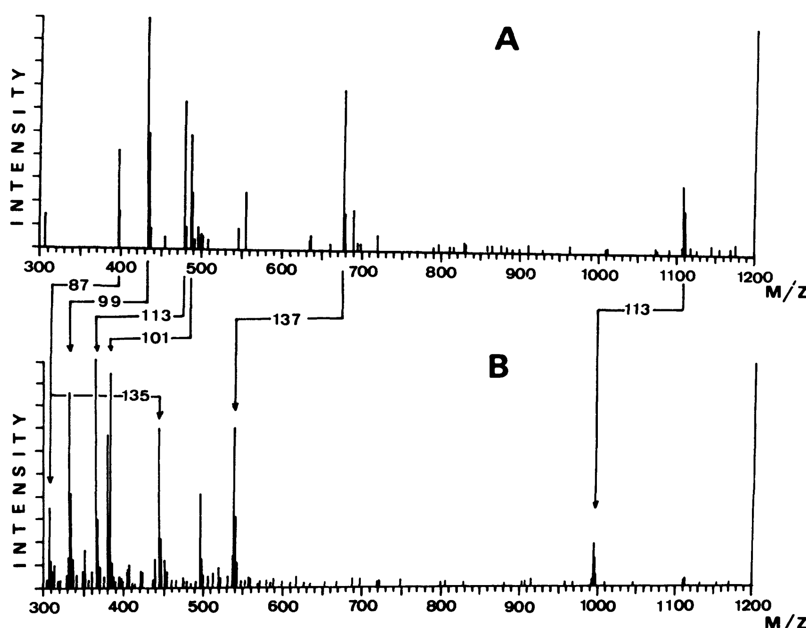

Fig. 1. FD mass spectra of chymotryptic peptides (A) and their peptide fragments after one cycle of degradation (B) of glucagon

Table 1. Mass Numbers of Observed Ions in Edman-degradation and FD-MS of Chymotryptic and Tryptic Peptides of Glucagon

0	1 PTH	1 peptide	2 PTH	2 peptide	3	4	5	Estimated sequence
Chymotryptic peptides								
1107[a]	248(L,[b,g]113)	994	250(D,115)	879	792	636	480	LDSRR(-)[g,13)]
676	272(H,137)	539	204[c](S,87)	452	324	267[d]		HSQG(TF)
485	236(T,101)	384		297	182			TSDY
478		365	266(M,131)	234	120			LMNT
432	234(V,99)	333	263(Q,128)	205				VQW
397	204[c](S,87)	310		182				SKY
		445[e]	248[f](L,113)					
Tryptic peptides								
1357		1220		1133	1005			HSQ(-)
1352		1281		1153	1038			AQD(-)
653		490		377	262			YLD(-)
175								R

a) m/z 1107 was confirmed as $[M+H]^+$ by the presence of $[M+2H]^{2+}$ with m/z 554.

b) Since glucagon does not contain Ile, the PTH with m/z 248 represents PTH of Leu.

c) PTH of Ser was observed mainly as $[M-H_2O]^+$ with m/z 204.

d) Peptide with m/z 267 as M^+ or $[M+H]^+$ contains only Thr and Phe. Since chymotrypsin generally hydrolyzes amide-bonds at the C-side of aromatic amino acid residues, the sequence Thr-Phe is assumed.

e) This m/z was considered to be derived from 310 + phenylthiocarbamyl (135). Therefore, this peptide contains Lys.

f) m/z 248 was considered to be carried over from the 1st step of degradation.

g) Abbreviations: A:Ala, D:Asp, F:Phe, G:Gly, H:His, I:Ile, K:Lys, L:Leu, M:Met, N:Asn, P:Pro, Q:Gln, R:Arg, S:Ser, T:Thr, V:Val, W:Trp, Y:Tyr.

peptide mixtures were then measured. Seven intense ion peaks were observed in the spectrum of chymotryptic peptides, as shown in Fig. 1. One of these ion peaks was confirmed to be a doubly charged ion peak. The spectrum obtained had relatively simple ion peaks, similar to those of a synthetic mixture of peptides,[6)] and hence the procedure seemed suitable for analyses of enzymic digests of polypeptides. The m/z determined are recorded in the first column of Table 1. In the first cycle of Edman-degradation of the chymotryptic peptides, five PTHs (four from organic solution, two from aqueous solution, and PTH of Thr from both solutions) and seven ion peaks of degraded peptides were observed on FD mass spectra. The sequence of these operations was repeated until the 5th step of degradation. The results are summarized in Table 1. The mass spectra of tryptic peptides and successively degraded peptide fragments were also measured like those of chymotryptic peptides and the results are shown in Table 1. The amino acid sequence of glucagon could then be confirmed by overlapping the partial sequences obtained by only 5 cycles of degradation of chymotryptic peptides and 3 cycles of

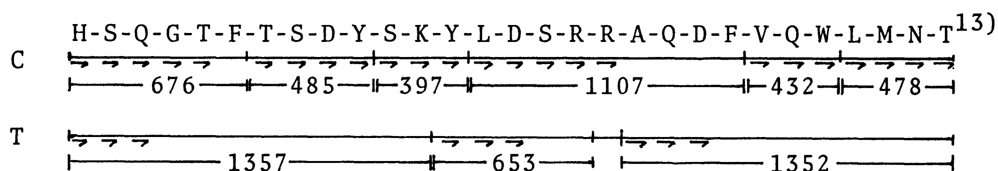

Fig. 2. Overlapping of two enzymic peptides of glucagon. C and T stand for chymotryptic and tryptic peptides, respectively. Arrows denote amino acid residues released by the Edman method. Numbers mean the mass values described in Table 1.

degradation of tryptic peptides, as shown in Fig. 2. Thus, the procedure has the following advantages as a method for sequencing of peptide mixtures; (1) substantial preseparation of individual peptides from the mixture is not necessary, (2) several R_ns can be identified at the same time, (3) R_n can be precisely determined by analyses of both PTH and peptide fragment released by degradation, and (4) complicated chemical treatments, *e.g.*, conversion of peptides to volatile derivatives for GC-MS, are not necessary.

References and Notes

1) P. Edman and A. Henschen, "Protein Sequence Determination", ed by J. M. Needleman, Springer-Verlag, Berlin-Heidelberg-New York, (1975), pp. 232.
2) H. D. Beckey, *Int. J. Mass Spectrom. Ion Phys.*, 2, 55 (1969).
3) H.-R. Schulten, "Methods of Biochemical Analysis", ed by D. Glick, Interscience Publ., J. Wiley, New York, (1977), Vol. 24, pp. 374.
4) K. Bieman, *Pure Appl. Chem.*, 50, 149 (1978).
5) D. H. Williams, *Pure Appl. Chem.*, 50, 219 (1978).
6) H. U. Winkler and H. D. Beckey, *Biochem. Biophys. Res. Commun.*, 46, 391 (1972).
7) I. Katakuse, T. Matsuo, H. Wollnik, and H. Matsuda, *Org. Mass Spectrom.*, in press.
8) See *e.g.* T. Inagami and K. Murakami, *Anal. Biochem.*, 47, 501 (1972).
9) H.-R. Schulten and B. Wittmann-Liebold, *Anal. Biochem.*, 76, 300 (1976).
10) The experimental for FD-MS are described: T. Matsuo, I. Katakuse, H. Matsuda, Y. Shimonishi, Y.-M. Hong, and Y. Izumi, Nature, in press.
11) I. Katakuse, T. Matsuo, H. Matsuda, Y. Shimonishi, and Y. Izumi, *Mass Spectrometry (Japan)*, 27, 127 (1979).
12) These peptides were synthesized by conventional methods in our laboratories.
13) One letter notation for amino acid residues by IUPAC-IUB, tentative rules: *J. Biol. Chem.*, 243, 3557 (1968). See footnote g) in Table 1.
14) Y.-M. Hong, Y. Shimonishi, T. Matsuo, I. Katakuse, and H. Matsuda, to be published elsewhere.
15) Glucagon was purified from commercial material (Lot 802521, Calbiochem) as described: W. W. Brommer, M. E. Boucher, and J. E. Koffenberger, Jr., *J. Biol. Chem.*, 246, 2822 (1971). Amino acid analysis: Trp,0.71 (in 4M CH_3SO_3H); Lys, 1.06; His,0.98; Arg,2.12; Asp,3.77; Thr,2.83; Ser,3.73; Glu,2.89; Gly,0.97; Ala,1.00; Val,0.90; Met,0.83; Leu,1.97; Tyr,1.88; Phe,1.73 (in 6M HCl at 105°C for 24 h).
16) Preparations from Worthington Biochemical Corp. TPCK stands for L-1-tosylamido-2-phenylethyl chloromethyl ketone.

(Received August 6,1979)