FEBS Letters 357 (1995) 65–69

MALDI-MS for C-terminal sequence determination of peptides and proteins degraded by carboxypeptidase Y and P

Bernd Thiede[a], Brigitte Wittmann-Liebold[a], Michael Bienert[b], Eberhard Krause[b,*]

[a]*Max-Delbrück-Zentrum für Molekulare Medizin, Abt. Proteinchemie, D-13125 Berlin-Buch, Germany*
[b]*Forschungsinstitut für Molekulare Pharmakologie, Alfred-Kowalke-Str. 4, D-10315 Berlin, Germany*

Received 7 October 1994; revised version received 21 November 1994

Abstract Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) has been used for C-terminal amino acid sequence determination of peptides and proteins. The usefulness of MALDI-MS was demonstrated by analyzing peptide mixtures (C-terminal peptide ladder) which were generated by enzymatic digestion of substance P, glucagon, angiotensinogen, insulin B chain and myoglobin with the exopeptidases carboxypeptidase Y and P. The results clearly show that up to 11 amino acid residues can be determined in the pmol range by analyzing the molecular masses of the truncated peptides. For proteins it is possible to investigate enzymatic or chemical digests in the same manner.

Key words: MALDI-MS; C-Terminal sequencing; Carboxypeptidase

1. Introduction

In recent years, matrix-assisted UV laser desorption/ionization mass spectrometry (MALDI-MS), introduced by Karas and Hillenkamp [1], has become a useful tool to determine the molecular mass of biomolecules. Applications of MALDI-MS for analysis of proteins [2], glycoproteins [3], oligosaccharides [4], and oligonucleotides [5,6] have been reported. Because of the sensitivity in the pmol range and the tolerance toward many buffer systems, MALDI-MS is a suitable technique for analysis of peptide mixtures resulting from enzymatic protein digestion [7]. A new approach to N-terminal protein sequencing that combines multiple steps of Edman degradation chemistry with a MALDI-MS determination of the concatenated set of peptide fragments was described as 'protein ladder sequencing' [8]. However, the method is limited to the N-terminus of peptides or proteins and requires the formation of stable phenylcarbamyl peptides.

Sequence information from the C-terminus of proteins is especially of interest in the investigation of N-terminal blocked peptides and proteins and for confirmation of DNA sequence data and cloning experiments. Chemical methods for C-terminal sequencing have been developed, but no method for the C-terminal sequence analysis has yet been described which works satisfactorily in cases of all amino acids [9–11], comparable with Edman degradation for the N-terminal region [12]. Besides chemical modification, carboxypeptidases have frequently been used for C-terminal sequencing [13], either by determining the release of free amino acids or by identifying the truncated peptides by mass spectrometry. To analyse C-terminally truncated peptides, fast atom bombardment (FAB-MS) [14–16], plasma desorption mass spectrometry (PDMS) [17], thermospray LC-MS [18,19] and electrospray ionization [20] have been used for monitoring carboxypeptidase digestions.

The kinetics for the enzymatic release of different amino acids varies considerably. Therefore, the amino acid analysis is very complex and misinterpretations due to internal cleavages of the polypeptide chain and/or the release of dipeptides may occur. Although carboxypeptidase Y (CPY) releases all C-terminal amino acids from peptides or proteins, including Pro [13], in some cases the reaction conditions are not sufficient to obtain an interpretable sequential release of amino acids [21]. To overcome these problems and to obtain additional information we used carboxypeptidase Y (slow release of Gly and Asp) and P (slow release of Ser and Gly) to generate C-terminal truncated peptides ('C-terminal peptide ladder') since CPP can supplement CPY in releasing all amino acids. In the present work we study the combination of carboxypeptidase digestion with monitoring of the C-terminal peptide ladder by the fast and sensitive MALDI-MS technique.

2. Materials and methods

2.1. Materials

Myoglobin (horse heart), insulin B chain (porcine, oxidized), angiotensinogen (1–14, porcine), cyanogen bromide, and α-cyano-4-hydroxycinnamic acid were obtained from Sigma, Deisenhofen, Germany. Glucagon was purchased from Bachem, Heidelberg, Germany. Carboxypeptidase Y and carboxypeptidase P (EC 3.4.16.1), both sequencing grade, were obtained from Boehringer-Mannheim, Germany. Formic acid, acetic acid, and trifluoroacetic acid, all analytical grade, were supplied by Merck, Darmstadt, Germany. Substance P and ovine corticotropin releasing factor (CRF) were synthesized automatically (MilliGen 9050 Peptide Synthesizer) by solid-phase methods using standard Fmoc chemistry .

2.2. Carboxypeptidase digestion of substance P and CRF

The peptide (500 pmol) was dissolved in 10 μl sodium citrate buffer (50 mmol/l, pH 6.0). CPY was added to give an enzyme:peptide ratio of 1:50 by weight. Aliquots of 1 μl were taken after 1, 2, 5, 10, 15, 30 and 60 min incubation at ambient temperature (24°C) and mixed with 1.5 μl of the MS matrix solution (α-cyano-4-hydroxycinnamic acid) on the target for mass spectral analysis.

2.3. Carboxypeptidase digestion of glucagon, myoglobin, angiotensinogen, and insulin B chain

The peptide (500 pmol) was dissolved in 10 μl of 0.1% aqueous TFA, evaporated under reduced pressure and then dissolved in 10 μl of aqueous acetic acid at pH 5.0. CPY was added to give an enzyme:peptide ratio of 1:50 by weight. After 10 min incubation the same

*Corresponding author. Fax: (49) (30) 512 8014.

Abbreviations: MALDI-MS, matrix assisted laser desorption/ionization mass spectrometry; LC-MS, liquid chromatography mass spectrometry; CPY, carboxypeptidase Y; CPP, carboxypeptidase P; TFA, trifluoroacetic acid; Da, Dalton.

SSDI 0014-5793(94)01323-3

amount of CPP was added. Aliquots were taken and mixed with matix solution as described above (section 2.2).

2.4. Cyanogen bromide digestion of myoglobin

Myoglobin was dissolved in aqueous 70% formic acid to give a final concentration of 30 μM. A 200 molar excess of cyanogen bromide was added to 50 μl of the solution and the reaction proceeded in the dark for 3 h. The reaction mixture was then evaporated under reduced pressure. The myoglobin fragments were further digested with carboxypeptidase Y and P as described in section 2.3.

2.5. Mass spectrometry

The mass spectra were recorded in the linear mode on a time of flight VG Tof Spec (Fisons, Manchester, England) or a Maldi II (Kratos, Manchester, England). The mass spectrometers were equipped with a nitrogen laser (337 nm, pulse duration 4 and 3 ns). The acceleration voltage was 20 kV. The spectra were obtained by summing over 50 laser pulses. A 50 mmol solution of α-cyano-4-hydroxycinnamic acid in aqueous 50% acetonitrile and 0.1% TFA was used as the matrix. 1.5 μl of the matrix and 1.0 μl of the sample were mixed on the target and air dried.

3. Results

The molecular masses of the fragments obtained when the peptides were incubated with CPY and CPP are listed in Table 1. The MALDI-MS spectra of the CPY digestion of substance P show the $[M+H]^+$ and $[M+Na]^+$ peaks and demonstrate that 5 amino acids were released within 30 min. After 5 min no

Table 1
MALDI-MS determined masses of C-terminal truncated peptide fragments after carboxypeptidase digestion

Sequence	M_{cal}	$[M+H]^+_{found}$	ΔM (amino acid)
Substance P			
RPKPQQFFGLM	1347	1349	-----
RPKPQQFFGL	1216	1218	131 (M)
RPKPQQFFG	1103	1105	113 (L)
RPKPQQFF	1046	1049	56 (G)
RPKPQQF	899	901	148 (F)
RPKPQQ	752	754	147 (F)
CRF			
[1-35]KLLDIA	4669	4669	-----
[1-35]KLLDI	4599	4600	69(A)
[1-35]KLLD	4486	4488	112(I)
[1-35]KLL	4371	-----	-----
[1-35]KL	4258	4260	228 (L+D)
Angiotensinogen [1-14]			
DRVYIHPFHLLVYS	1759	1760	-----
DRVYIHPFHLLVY	1672	1672	88 (S)
DRVYIHPFHLLV	1509	1509	163 (Y)
DRVYIHPFHLL	1410	1410	99 (V)
DRVYIHPFHL	1297	1297	113 (L)
DRVYIHPFH	1184	1184	113 (L)
DRVYIHPF	1047	1046	138 (H)
DRVYIHP	900	898	148 (F)
DRVYIH	803	802	96 (P)
DRVYI	667	665	137 (H)
DRVY	554	552	113 (I)
Glucagon			
[1-16]RRAQDFVQWLMNT	3483	3483	-----
[1-16]RRAQDFVQWLMN	3382	3381	102 (T)
[1-16]RRAQDFVQWLM	3268	3267	114 (N)
[1-16]RRAQDFVQWL	3137	3136	131 (M)
[1-16]RRAQDFVQW	3024	3023	113 (L)
[1-16]RRAQDFVQ	2838	2836	187 (W)
[1-16]RRAQDFV	2710	2709	127 (Q)
[1-16]RRAQDF	2607	2609	100 (V)
[1-16]RRAQD	2460	2461	148 (F)
[1-16]RRAQ	2345	2345	116 (D)
[1-16]RRA	2217	2217	128 (Q)
[1-16]RR	2146	2148	69 (A)

Table 1 (*continued*)

Sequence	M_{cal}	$[M+H]^+_{found}$	ΔM (amino acid)
Insulin B chain			
[1-19]GERGFFYTPKA	3496	3498	-----
[1-19]GERGFFYTPK	3425	3427	71 (A)
[1-19]GERGFFYTP	3297	3300	127 (K)
[1-19]GERGFFYT	3200	3203	97 (P)
[1-19]GERGFFY	3099	3099	104 (T)
[1-19]GERGFF	2936	2938	161 (Y)
[1-19]GERGF	2789	2790	148 (F)
[1-19]GERG	2642	2643	147 (F)
[1-19]GER	2585	2586	57 (G)
[1-19]GE	2429	2430	156 (R)
Myoglobin/ CNBr-digest			
[1-9]DIAAKYKELGFQG	2513	2516	-----
[1-9]DIAAKYKELGFQ	2456	2459	57 (G)
[1-9]DIAAKYKELGF	2328	2332	127 (Q)
[1-9]DIAAKYKELG	2181	2184	148 (F)
[1-9]DIAAKYKEL	2124	2126	58 (G)
[1-9]DIAAKYKE	2011	2014	112 (L)
[1-9]DIAAKYK	1882	1884	130 (E)
[1-9]DIAAKY	1756	1757	127 (K)
[1-9]DIAAK	1593	1592	165 (Y)
[1-9]DIAA	1465	1464	128 (K)
[1-9]DIA	1394	1393	71 (A)
[1-9]DI	1323	1321	72 (A)

substance P mass peak (1347 Da) was detected. The mass differences, indicating the released amino acids, deviate from the expected masses by only 1 Da. The digestion of CRF (4669 Da) with CPY leads to the subsequent release of 4 amino acids within 10 min. However, the peptide CRF (1–38) with leucine in the C-terminal position was not detected. The mass accuracy of the detected truncated peptides was ± 2 Da.

Because of the different specificities of the serine carboxypeptidases [21], a combination of CPY and CPP was used for sequencing of angiotensinogen, glucagon, insulin B chain and myoglobin. The cleavage of the first and second amino acid of angiotensinogen (1–14) proceeded very rapidly. Therefore, the first incubation was carried out for 10 min at 0°C followed by an incubation at ambient temperature. The masses of 10 truncated peptides were found, with the peak intensities for angiotensinogen (1–5) and (1–6) being relatively weak.

The mass spectra of the digestion of glucagon, a peptide containing 29 amino acids, show that after 60 min 9 amino acids were cleaved off (Fig. 1). The signal intensities of glucagon fragments 1–27 and 1–28 were weak. However, the complete C-terminal peptide ladder was detected with a mass accuracy of ± 2 Da. The enzymatic digestion of the insulin B chain at ambient temperature (24°C) releases the C-terminal alanine alone after 15 min. After 60 min at 37°C it was possible to detect the 9 C-terminal amino acids with a mass accuracy of ±2 Da.

The digestion of myoglobin with CNBr generates three fragments with molecular weights of 8179 Da, 6233 Da and 2513 Da. The treatment of this mixture with carboxypeptidases resulted in the cleavage of up to 11 amino acid residues from the C-terminal fragment (2513 Da) in 60 min. The mass accuracy of the myoglobin fragments was determined to be ± 3 Da.

4. Discussion

The results indicate that the combination of carboxypeptidase digestion and MALDI-MS is an effective and convenient method for the analysis of the C-terminal sequence of peptides. MALDI-MS has some advantages over other methods of detection because of its speed and low sample consumption.

Carboxypeptidases have different preferences for each amino acid and individual peptide bonds [21,22]. Usually, the detection of truncated peptide of a slowly released amino acid followed by a rapidly released amino acid is difficult. However, the results show that strong peak signals with a mass accuracy of ± 1 Da were obtained in most cases. On the other hand weak mass peaks with a mass accuracy of ± 3 Da were also observed. Because of the same molecular weight of the amino acids the MALDI-MS is not able to differentiate between leucine/isoleucine and between lysine/glutamine. In addition, at a mass precision of 3 Da it is impossible to distinguish between asparagine/aspartic acid, glutamine/glutamic acid/lysine, proline/valine, valine/threonine, and threonine/cysteine. Another limitation of this method is the upper limit of molecular weight which can be determined with necessary precision. The practical mass accuracy for MALDI-MS should be 0.01% [2] but for weak peaks the mass accuracy was only <0.1%. For this reason at

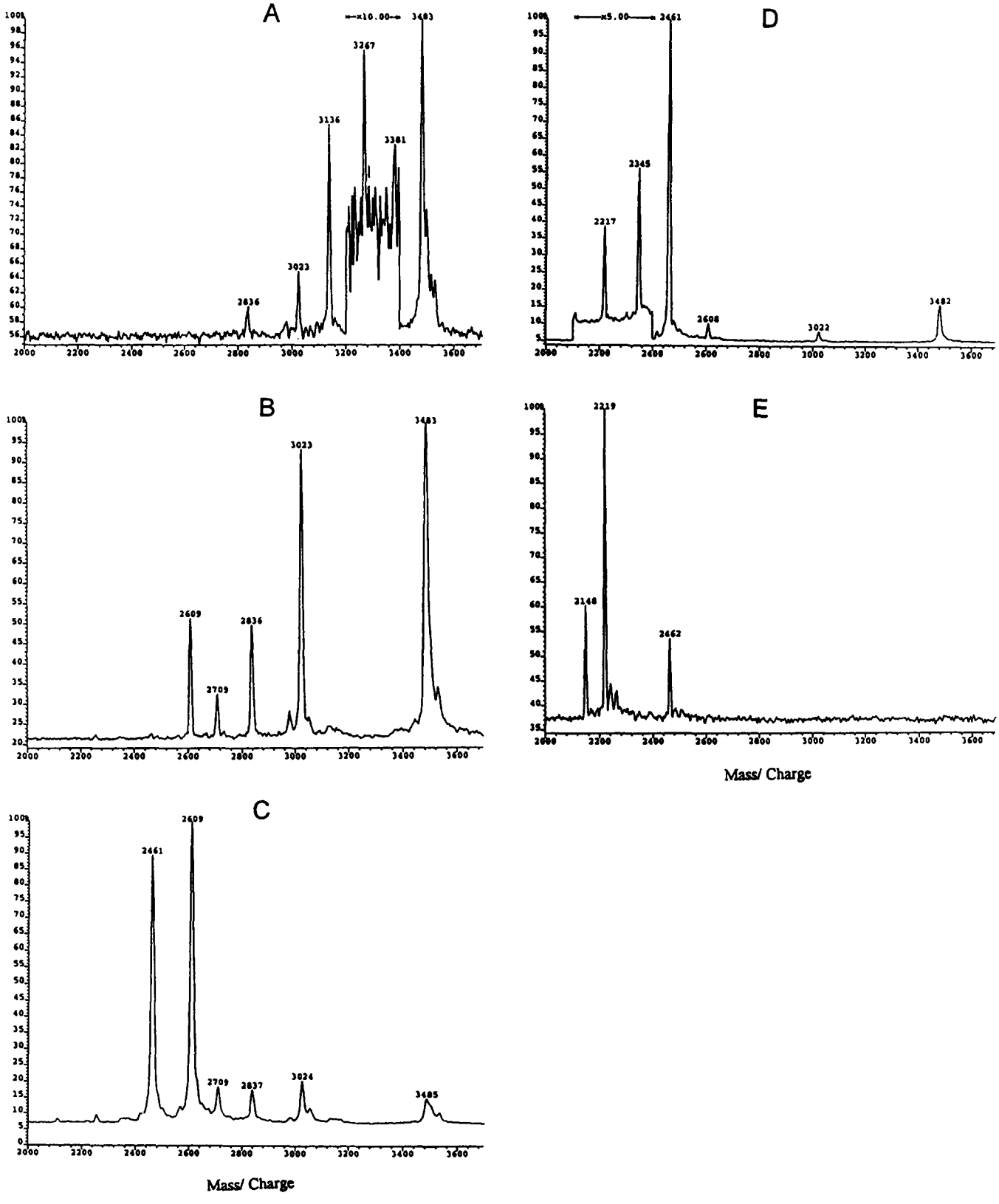

Fig. 1. MALDI mass spectra of glucagon after (A) 2 min, (B) 10 min, (C) 15 min, (D) 30 min, and (E) 60 min digestion with CPY and CPP.

masses higher than 5000 Da the mass resolution is not sufficient for sequence analysis, and in addition, the cleavage by carboxypeptidase is often hindered. For peptides with higher masses or proteins it is better to sequence chemical or enzymatic digests, as we have shown for a cyanogen bromide digestion of myoglobin. The use of a reflectron increases the accu-

racy in time-of-flight mass detection, but the intensity of weak peaks is very often lower than in the linear mode. The combined use of two carboxypeptidases, CPY and CPP, enabled us to determine serine and aspartic acid in the same time–course as other amino acids. Furthermore there was no slow release of glycine.

The combination of carboxypeptidase hydrolysis with MALDI-MS is an effective tool for C-terminal sequence analysis, because MALDI-MS is a rapid and sensitive detection method. The easy sample preparation, the matrix which terminates the enzymatic reaction, the ability to tolerate contaminations and the easily interpretable data especially in a mixture, indicate that MALDI-MS has advantages in comparison with other mass spectrometrical methods for determination or confirmation of the C-terminal sequence of polypeptides.

References

[1] Karas, M. and Hillenkamp, F. (1988) Anal. Chem. 60, 2299–2301.
[2] Beavis, R.C. and Chait, B.T. (1990) Anal. Chem. 62, 1836–1840.
[3] Allmaier, G., Schmid, E.R., Hagspiel, K., Kubicek, C.P., Karas, M. and Hillenkamp, F. (1990) Anal. Chim. Acta 241, 321–327.
[4] Stahl, B., Steup, M., Karas, M. and Hillenkamp, F. (1991) Anal. Chem. 63, 1463–1466.
[5] Wang, B.H. and Biemann, K. (1994) Anal. Chem. 66, 1918–1924.
[6] Wu, K.J., Shaler, T.A. and Becker, C.H. (1994) Anal. Chem. 66, 1637–1645.
[7] Billeci, T.M. and Stults, J.T. (1993) Anal. Chem. 65, 1709–1716.
[8] Chait, B.T., Wang, R., Beavis, R.C. and Kent, S.B.H. (1993) Science 262, 89–92.
[9] Boyd, V.L., Bozzini, M., Zon, G., Noble, R.L. and Mattaliano, R.J. (1992) Anal. Biochem. 206, 344–352.
[10] Inglis, A.S. (1991) Anal. Biochem. 195, 183–196.
[11] Tsugita, A., Takamoto, K., Kamo, M. and Iwadate, H. (1992) Eur. J. Biochem. 206, 691–696.
[12] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[13] Hayashi, R. (1977) Methods Enzymol. 47, 84–93.
[14] Bradley, C.V. and Williams, D.H. (1982) Biochem. Biophys. Res. Commun. 104(4) 1223–1230.
[15] Caprioli, R.M. and Fan, T. (1986) Anal. Biochem. 154, 596–603.
[16] Wagner, R.M. and Fraser, B.A. (1987) Biomed. Environ. Mass Spectrom. 14, 235–239.
[17] Klarskov, K., Breddam, K. and Roepstorff, P. (1989) Anal. Biochem. 180, 28–37.
[18] Kim, H.K., Pilosof, D., Dyckes, D.F. and Vestal, M.L. (1984) J. Am. Chem. Soc. 106, 7304–7309.
[19] Self, R. and Parente, A. (1983) Biomed. Mass Spectrom. 10(2) 78–82.
[20] Smith, C.E and Duffin, K.L. (1993) in: Techniques in Protein Chemistry IV, pp. 463–470, Academic Press, San Diego.
[21] Breddam, K. (1986) Carlsberg Res. Commun. 51, 83–126.
[22] Hayashi, R., Bai, Y and Hata, T. (1975) J. Biochem. 77, 69–79.