Journal of Chromatography, 646 (1993) 169–174
Elsevier Science Publishers B.V., Amsterdam

CHROMSYMP. 2802

Rapid high-sensitivity peptide mapping by liquid chromatography–mass spectrometry

Kuldip Mock*, Mark Hail, Iain Mylchreest, Joe Zhou, Ken Johnson and Ian Jardine

Finnigan MAT, 355 River Oaks Parkway, San Jose, CA 95134 (USA)

ABSTRACT

Toward a complete LC–MS mapping system for peptides and proteins, we have coupled a precision-flow microbore HPLC system to an electrospray single quadrupole mass spectrometer. The HPLC system allows fast separation of protein digests with UV detection at the low pmol level. A 2 μl/min portion (1:25) of the effluent is passed into a high-sensitivity electrospray MS system. The electrospray source allows for molecular mass associated ions (MH^+, MH_2^{2+}, MH_3^{3+}, etc.) to be generated as well as collision induced dissociation of these ions before MS analysis. After LC–MS runs, with or without partial fragmentation, the data generated are largely interpreted by identification of predicted peptides, incompletely digested peptides, unusual peptide cleavages, and so on, using appropriate integrated software (PEPMAP, PEPMATCH). Examples of peptide mapping at the low pmol level using this integrated system will be shown (*e.g.*, of the protein human growth hormone and of the glycoprotein, tissue plasminogen activator).

INTRODUCTION

The ability to determine the primary structure of proteins is essential to understanding the biological function of important macromolecules. Although the amino acid sequence is fundamental to their activity, what is also very important is the way in which the molecules have been post-translationally modified, *e.g.* glycosylated, disulfide bridged, and so on. The coupling of liquid chromatography (LC) with electrospray ionization (ESI) on a single-stage quadrupole mass spectrometer offers a unique investigative tool for the molecular biologist. We report here, the systematic use of ESI-MS in conjunction with precise-flow, gradient microbore LC for verifying the primary structure of proteins at the micro-scale level.

Microbore column chromatography (less than 2.1 mm I.D.), first proposed by Ishii *et al.* [1] and Scott and Kucera [2,3], has become a fundamental technique for isolating low microgram quantities of proteins and peptides for automated sequence analysis using Edman degradation [4].

An important aspect of the successful utilization of microbore columns is the ability to employ reproducible gradient elution chromatography. This was accomplished using micro-LC designed specifically for ultrafast protein analysis. The LC system consisted of a dual microbore binary gradient pumping system which allows rapid gradient formation (1–3 min) with precise flows 10–1000 μl/min range. Band broadening is minimized (less than 0.1 μl) by using special columns which directly couple to the micro flow cell, eliminating the use of tubing or fittings. In the microflow mode the pumps deliver 0.1–100 μl/min allowing precise, reproducible gradients from 10–100 μl/min. At these flow-rates the 3-μl mixing tee and the 17-μl static mixer ensure proper mixing and minimize gradient delay vol-

* Corresponding author.

0021-9673/93/$06.00

ume to less than 50 μl [5]. The data presented here shows confirmation of the structure of human growth hormone, a recombinant protein (rhGH) and a recombinant glycoprotein, tissue plasminogen activator (rt-PA), using less than 100 pmol of material. This analysis, coupled with automated recording and interpretation of data [6] enabled rapid achievement of the following processes: (i) isolation of amino and carboxy terminii; (ii) location of disulfide bridged peptides; (iii) location of glycosylated peptides; and (iv) internal sequence information.

EXPERIMENTAL

Chromatography

Tryptic digests of native rhGH and reduced carboxymethylated rt-PA were analyzed. The micro-column LC system used in this work consisted of a Michrom BioResources (Pleasanton, CA, USA) capable of delivering linear reproducible gradients, equipped with an automated injector (Valco Instruments, Houston, TX, USA) and a thermally controlled mobile phase and coupled to a variable-wavelength UV detector equipped with a microflow cell (0.8 μl). All chromatographic separations were accomplished using a Reliasil, 15 × 0.1 cm, 5 μm, microbore column packed with a C_{18} stationary phase. The elution profile was monitored at 214 nm.

The mobile phase A consisted of water containing 0.1% trifluoroacetic acid. Mobile phase B was acetonitrile containing 0.085% trifluoroacetic acid. The gradient used to separate the rhGH tryptic peptides was linear from 0 to 60% B in 30 min. The separation of tryptic peptides of rt-PA utilized a linear gradient 0 to 25% B over 50 min followed by 25 to 60% B in 35 min. The column effluent (50 μl/min at 40°C) was split at a ratio of 25:1 to give a flow-rate of 2 μl/min into the mass spectrometer.

Electrospray mass spectrometry

Experiments were conducted on a Finnigan MAT SSQ 710 single-stage quadrupole mass spectrometer fitted with a 20-kV post-acceleration detector for high sensitivity and an ESI interface. The electrospray used a multilayer flow system delivering a sheath liquid flow and a concentric gas flow at the terminus of the ESI needle in order to allow for efficient droplet dispersion (and ion production) over a wide range of mobile phase conditions. Typically a 1:1 eluate:sheath liquid flow is utilized in order to maintain flow surface tension over gradient separation. 2-Methoxyethanol (2 μl/min) was used as the sheath liquid.

Collisionally activated dissociation (CAD) spectra of peptides were generated between the capillary and the skimmer. The voltages were increased from 36 to 100 V on the capillary lens and to 195 from 136 for the tube lens. The quadrupole was scanned from 330–2000 u in 4 s.

Materials

Tryptic digests of native rhGH and reduced carboxymethylated rt-PA were analyzed. These were proprietary samples and were a gift from a biotechnology company. HPLC-grade water and acetonitrile were obtained from Scientific Products (Hayward, CA, USA), sequanal grade trifluoroacetic acid was from Pierce (Rockford, IL, USA) and 2-methoxyethanol from Aldrich (Milwaukee, WI, USA).

RESULTS AND DISCUSSION

An important link between data generation and problem solving, is data interpretation. The

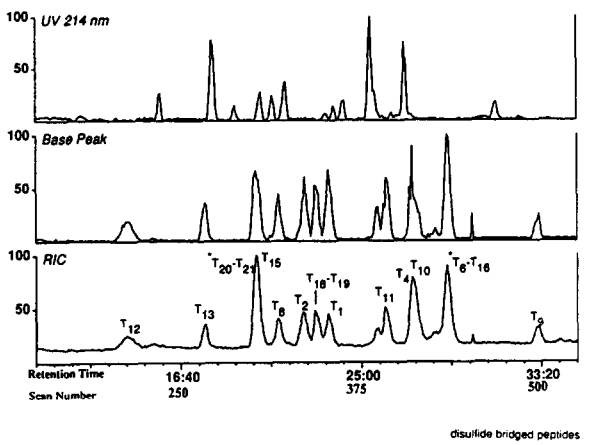

Fig. 1. LC–ESI-MS of tryptic digest of rhGH. A 78-pmol amount was injected on the column and 3 pmol were analyzed by ESI-MS. Retention time in min:s.

mass spectrometer data system has software tools which facilitate interpretation of electrospray data. For example the PEPMAP program automatically searches the LC–ESI-MS chromatogram for all predicted peptides including incompletely cleaved and disulfide-linked peptides of a protein digest. The LC–ESI-MS chromatograms of a tryptic digest of native rhGH are presented in Fig. 1. Peak retention times are different for reconstructed ion current (RIC) chromatogram and UV absorbance due to the transfer line between the UV detector and the mass spectrometer. In addition, peak intensities are different for RIC and UV absorbance because they are based on different physical measurements. After LC–MS, the data generated were interpreted using integrated software (PEPMAP, PEPMATCH). The results of the PEPMAP search are listed in Table I. The tryptic peptides found are designated by "T" and are consecutively ordered from amino terminus (T_1) to the carboxy terminus peptide. The program identified 15 out of the 21 fragments and in addition, located two disulfide-bridged peptides ($T_6 \cdot T_{16}$, $T_{20} \cdot T_{21}$), and an incompletely digested fragment ($T_{18} \cdot T_{19}$).

Fig. 2A and B are mass spectra of tryptic fragment T_{11} of rhGH obtained before and after fragmentation from an on-line LC–ESI-MS experiment. Rapid confirmation of known sequences is attained with an automated PEPMATCH program. PEPMATCH is a peptide sequence matching program which automatically compares MS–MS data for putative known peptides with their predicted MS/MS ions as shown in Table II. In addition, a semi-automatic sequencing program (PEPSEQ) allows rapid first pass analysis of unknown peptide sequence data.

The LC–ESI-MS tryptic map of reduced carboxymethylated rt-PA is shown in Fig. 3. Each peak is labelled according to the corresponding tryptic peptide. The RIC chromatogram compares well with UV absorbance profile monitored at 214 nm. After LC–MS, an automated two- or three-dimensional search is carried out on the data. Fig. 4 is a two-dimensional display of multiply-charged ions of the rt-PA digest. The results indicated that there were ions of different molecular mass but similar retention time, as highlighted by intense regions. On closer examination , it was found that these regions contained different glycoforms. This ability is most useful

TABLE I

RESULTS OF PEPMAP CHROMATOGRAM SEARCH

Position	M + H	Scan	Sequence[a]
T_1	930.54	346	FPTIPLSR
T_2	979.50	330	LFDNAMLR
T_4	2342.13	404	LHQLAFDTYQEFEEAYIPK
T_8	844.49	312	SNLELLR
T_{10}	2262.13	407	SVFANSLVYGASDSNVYDLLK
T_{11}	1361.67	386	DLEEGIQTLMGR
T_{12}	773.38	209	LEDGSPR
T_{13}	693.39	261	TGQIFK
T_{15}	1489.69	296	FDTNSHNDDALLK
$T_6S–ST_{16}$	3761.77	428	YSFLQNPQTSLCFSESIPTPSNR[b] | NYGLLYCFR
$T_{18}–T_{19}$	1253.62	337	DMDKVETFLR[c]
$T_{20}S–ST_{21}$	1400.63	298	IVQCR[b] | SVEGSCGF

[a] Peptide amino acid sequence.
[b] Disulfide bridged peptides.
[c] Incompletely cleaved fragment.

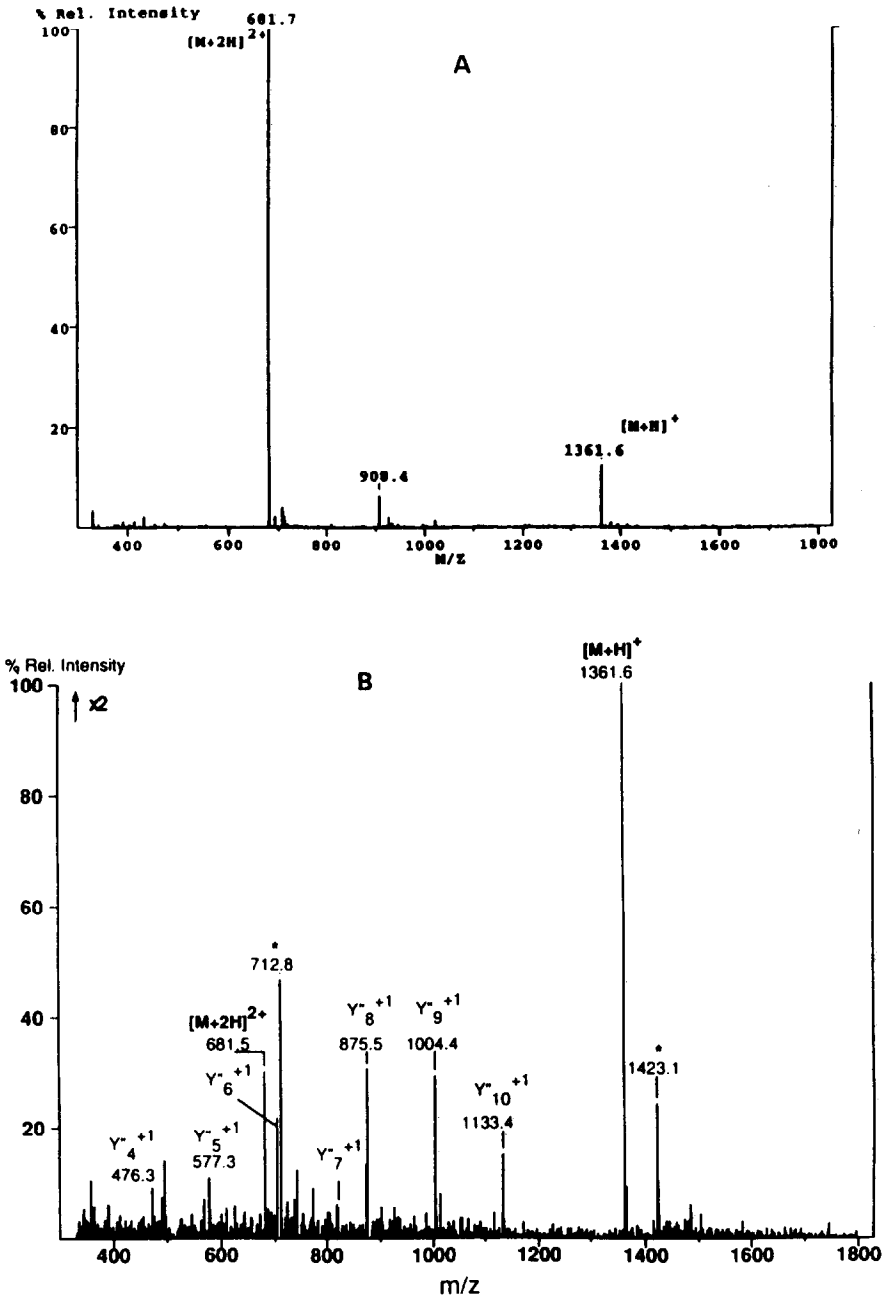

Fig. 2. (A) Mass spectrum of tryptic fragment T_{11} of rhGH. Peptide sequence: DLEEGIQTLMGR. (B) Source-CAD spectrum of fragment T_{11} of rhGH obtained from an on-line LC–ESI-MS experiment. A 78-pmol amount was injected on the column and 3 pmol were analyzed by ESI-MS. The asterisk indicates a co-eluted component (M_r 1423).

in identifying regions of the molecule that are heterogeneous, *e.g.* glycoforms, isomers, etc. It would not be possible to distinguish these properties, by UV detection alone. Fig. 5 is an example of a mass spectrum of glycopeptide T_{45}. The spectrum showed the glycoforms of T_{45} differed is sialic acid content.

LC–ESI-MS analysis of peptides coupled with

TABLE II

SEQUENCE-SPECIFIC FRAGMENT IONS IDENTIFIED FROM SOURCE-CAD SPECTRUM OF TRYPTIC PEPTIDE T_{11}

No.	Sequence	Ion type			No.
		A	B	Y″	
1	Asp	88.0	116.0	1361.7[a]	12
2	Leu	201.1	229.1	1246.6	11
3	Glu	330.2	358.2	1133.6[a]	10
4	Glu	459.2	487.2	1004.5[a]	9
5	Gly	516.2	544.2	875.5[a]	8
6	Ile	629.3	657.3	818.5[a]	7
7	Gln	757.4	785.4	705.4[a]	6
8	Thr	858.4	886.4	577.3[a]	5
9	Leu	971.5	999.5	476.3[a]	4
10	Met	1102.5	1130.5	363.2	3
11	Gly	1159.6	1187.6	232.1	2
12	Arg	1315.7	1343.7	175.1	1

[a] Sequence ions present in CAD spectrum.

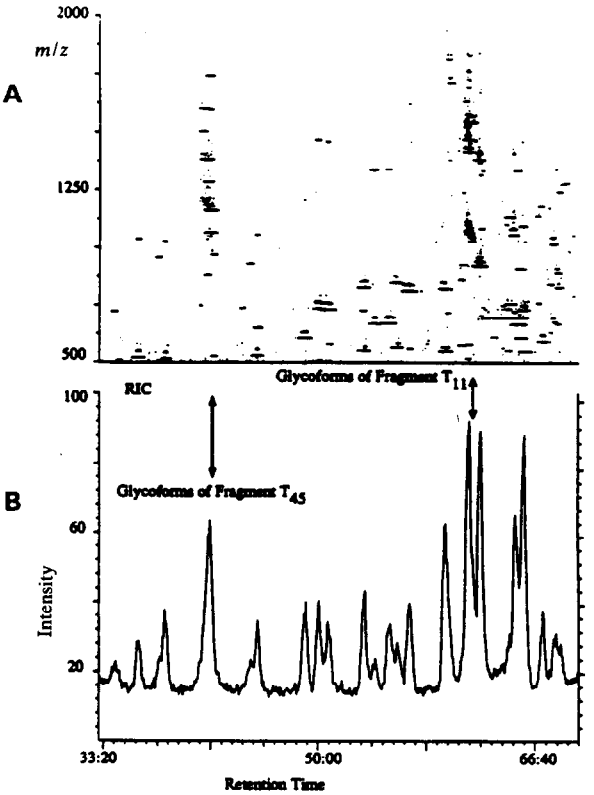

Fig. 4. Two dimensional map of multiply charged ions of tryptic digest of rt-PA. (A) Plot of mass-to-charge ratio *versus* intensity and retention time. A 50-pmol amount was loaded on the LC column and 2 pmol were analyzed by ESI-MS. The intense regions contained different glycoforms. (B) Plot of RIC *versus* retention time (in min:s).

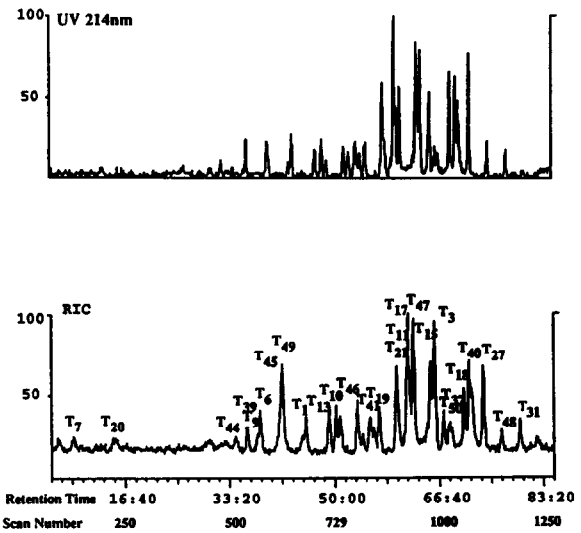

Fig. 3. LC–ESI-MS map of tryptic digest of reduced carboxymethylated rt-PA. A 50-pmol amount was loaded on the LC column and 2 pmol were analyzed by ESI-MS. Retention time in min:s.

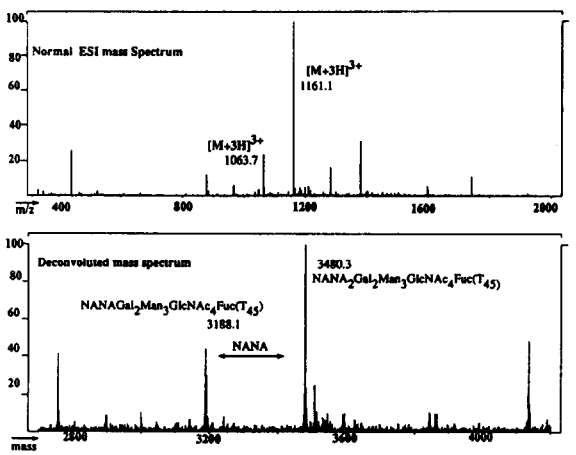

Fig. 5. Mass spectra of glycopeptide T_{45} before and after deconvolution. The spectrum showed the glycoforms of T_{45} differed in sialic acid content. NANA = N-acetylneuraminic acid.

automated recording and interpretation of data on a single-stage quadrupole mass spectrometer is routine, accommodating full LC gradients, including trifluoroacetic acid with excellent sensitivity. Source-CAD works well, providing internal sequence information, with good sensitivity. In addition, the UV trace is acquired directly on the data system of the mass spectrometer for comparison.

In our laboratory, recent improvements to the ESI source have enhanced the ability to cope with high flow rates up to 1000 μl/min, without any compromise in sensitivity (data not shown). The system (available for commercial release from Finnigan MAT, San Jose, CA, USA) accepts effluent from 1–4.6 mm I.D. HPLC columns, as well as packed column LC or CE, with direct flows of 1–1000 μl/min. This combination could allow ESI to emerge as the ultimate ionization technique for the on-line analysis of peptides and proteins.

REFERENCES

1 D. Ishii, K. Asai, K. Hibi, T. Jonokuchi and M. Nagaya, *J. Chromatogr.*, 144 (1977) 157–168.
2 R.P. Scott and P. Kucera, *J. Chromatogr.*, 185 (1979) 27–41.
3 R.P.W. Scott and P. Kucera, *J. Chromatogr.*, 169 (1979) 51–72.
4 R.J. Simpson and E.C. Nice, in A.K. Walsh (Editor), *Methods in Protein Sequence Analysis*, Humana Press, Clifton, NJ, 1986, pp. 213–228.
5 K.D. Nugent and P.W. Nugent, *BioChromatography*, 5 (1990) 142–148.
6 J.R. Yates, P.R. Griffin, L.E. Hood and J.X. Zhou, in J. Villifranca (Editor), *Techniques in Protein Chemistry, II*, Academic Press, San Diego, CA, 1991, 477–485.