Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER
UNCLASSIFIED	498478

TITLE

Identification of Bioactive Peptides by High Resolution Liquid Chromotography-ElectroSpray Mass Spectrometry

System Number:

Patron Number:

Requester:

Notes:

DSIS Use only:

Deliver to: FF

DRES

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD

SR 637

UNCLASSIFIED

Identification of Bioactive Peptides by High Resolution Liquid Chromatography-ElectroSpray Mass Spectrometry

BY

P. A. D'Agostino, J. R. Hancock and L. R. Provost

APRIL 1996

CRAD

National Defence Défense nationale

Canada

UNCLASSIFIED

Suffield Report No. 637

Identification of Bioactive Peptides by High Resolution Liquid Chromatography - Electrospray Mass Spectrometry

by

P. A. D'Agostino, J. R. Hancock and L. R. Provost

Executive Summary

Title: P.A. D'Agostino, J.R. Hancock and L.R. Provost, "Identification of Bioactive Peptides by High Resolution Liquid Chromatography - Electrospray Mass Spectrometry", Suffield Report No. 637, 1995, UNCLASSIFIED.

Introduction: The Canadian Forces (CF) may be called on to perform peacekeeping or peacemaking operations in regions of the world where there is a significant threat of chemica/biological warfare agent use. To operate effectively in these theaters the CF must be able to identify the exact nature of the chemical/biological agent(s). Mass spectrometry (MS), is a powerful analytical technique for the identification of both known and unknown compounds and DRE Suffield is currently investigating this instrumental technique in fulfilment of CF agent identification requirements.

Results: A database of electospray mass spectra (ESI-MS) for proteins which have biological activity was established during the development of MS techniques for identifying toxins. Mass spectra for a number of proteins which are simulants for mid-spectrum agents, including substance P (and related peptides), bradykinins, bombesins (and related peptides) and a Conus snail toxin, were acquired over a wide range of mass spectral conditions. Very accurate molecular weights were obtained for all of the materials analyzed, and an excellent database of their mass spectral fingerprints was created. The peptide mass spectra contained in this database may now be used for identification of these materials in any samples provided.

Significance of Results: The CF may be deployed in regions of the world where there is a significant threat of chemicalibiological warfare agent use. Identification of the agent is of importance since the results of such analyses would contribute to the development of strategic and political positions regarding future Canadian military operations and would facilitate the dissemination of technical advice to in-theater field commanders and medical personnel.

Future Goals: The CB threat spectrum includes chemical and biological warfare agents, and toxins of biological origin in the "mid-spectrum" between these agents. The identification research effort has been focused on the detection and identification of these toxins of biological origin. Use of these warfare agents could easily go unconfirmed, as analytical methods have not been fully developed for their identification. DRE Suffield is now actively addressing this deficiency through the application and development of MS methods for identification these agents.

ABSTRACT

A bioactive peptide database of electospray mass spectra (ESI-MS) was established during the development of high resolution liquid chromatographic ESI-MS analytical methods with a magnetic sector instrument. High resolution ESI-MS data for a variety of bioactive peptides, including substance P (and related peptides), bradykinins, bombesins (and related peptides) and a Conus snail toxin, were acquired over a wide mass range by scanning the magnetic sector and calibrating externally with polyethylene glycol standards. Multiply charged ions were observed and errors between observed and theoretical monoisotopic molecular weights were typically in the 5 to 30 ppm range for the bioactive peptides with magnetic sector resolutions between 2500 and 4000 (10% valley definition). Isotopic clusters for charge states of up to +5 were fully resolved, facilitating the rapid and unambiguous assignment of charge states and calculation of monoisotopic molecular weights. Under CAD/MS conditions both b_n- and y_n-series sequence ions were generally observed, enabling either full or partial amino acid sequencing of bioactive peptides, their degradation products and fragments generated by tryptic digestion.

TABLE OF CONTENTS

INTRODUCTION

The Canadian Forces (CF) may be called on to perform peacekeeping or battlefield operations in regions of the world where there is a significant threat of chemical/biological warfare (CBW) agent use. To operate effectively in these theaters the CF must be able to identify the exact nature of the chemical/biological agent(s). This requirement is being actively addressed at DRES by conducting research and development into new methods for the identification of chemical/biological warfare agents. DRES has adopted a multi-disciplinary approach that includes instrumental analytical techniques, immunological methods and other technologies in order to identify as wide a variety of CB agents as possible.

This research encompasses the detection, identification and confirmation of CW agents, toxins and BW agents of concern. In order to meet these objectives a research effort was initiated with the following general aim:

> **To provide the CF with an independent chemical, toxin and biological agent identification and confirmation capability.**

To meet this aim, a strong capability must exist in core identification technologies. This will allow the laboratory to respond quickly and provide the CF with the ability to unambiguously detect and confirm the presence of known and novel chemical, toxin and biological agents in samples suspected to contain these materials.

The current CB identification research effort has focused on the development of new techniques for the identification and confirmation of toxins and BW agents. Recent advances in biotechnology, including, the isolation and production of peptides using solid-phase synthesis and recombinant DNA-modified microorganisms, has opened up new avenues for the preparation of

militarily significant quantities of toxic agents in the "mid-spectrum" between classical chemical and classical biological warfare agents. Mid-spectrum agents, including bioactive peptides and proteins have emerged as a real threat. The current lack of appropriate identification methods makes confirmation of allegations of use of mid-spectrum agents extremely difficult and would likely result in inconclusive evidence. Development of suitable instrumental analytical methods for the identification and confirmation of these threat compounds must be addressed. A number of instrumental analytical technologies including mass spectrometry, liquid chromatography and capillary electrophoresis, have been targeted as candidate analytical technologies for the identification and confirmation of mid-spectrum agents.

The application of mass spectrometry to the analysis of bioactive peptides has undergone tremendous growth due to the recent development of the electrospray interface, a technique that allows rapid identification and characterization of high molecular weight peptides and proteins. When employed with collisionally activated dissociation mass spectrometry (CAD/MS) or tandem mass spectrometry (MS/MS) confirmed identification is possible for complete unknowns.

Electrospray mass spectrometry has been used at DRES for the identification of both target toxins and novel unknown toxins. The large number of potential bioactive peptides precludes the development of individual methods for all the candidate toxins. Therefore, a general approach that will provide maximal structural information and enable identification of toxins in general has been adopted. Briefly, the electrospray mass spectrometry identification strategy follows:

i) Determination of the monoisotopic or average molecular weight for toxic peptides or proteins (provisional or tentative identification) and,

ii) Sequence determination, based on amino acid residue mass losses, of toxic peptides or proteins (confirmed identification). Unambiguous confirmation would require the comparison of acquired spectrometric data with that obtained for authentic reference material.

In the early 1980's, Barber and co-workers revolutionized mass spectrometry by demonstrating the use of fast atom bombardment (FAB) mass spectrometry for the analysis of peptides (1). In the ensuing years, the useful mass range of mass spectrometry for biomolecule applications increased, particularly with the demonstration by Fenn and co-workers that electrospray ionization could be used to form multiply charged gaseous ions from large biomolecules (2,3). Biemann (4) has reviewed the mass spectrometry of peptides and proteins and the current status of biological mass spectrometry was reviewed in 1994 by Burlingame, Boyd and Gaskell (5). Numerous references, reviews and books are cited in these reviews and they serve as a good starting point for researchers interested in assessing the potential of biological mass spectrometry.

Electrospray ionization was initially interfaced to a quadrupole mass spectrometer (2,3) and most of the applications in the literature deal with this type of instrumentation. Many users in the mass spectrometry community conduct research with higher resolution instruments, in large part because of the accurate mass measurement capabilities of these instruments. High resolution data has been collected following electrospray introduction of biomolecules into Fourier transform mass spectrometers and this topic was recently reviewed by Buchanan and Hettich (6). Di- or tri- sector geometry mass spectrometers, although not capable of the resolution of Fourier transform mass spectrometers, are more common and have been used extensively for the acquisition of high resolution data.

Development of suitable electrospray interfaces for high resolution magnetic sector use (7) was in large part driven by the potential to increase mass measurement accuracy. Use of high resolution enables the assignment of charge state to multiply charged isotope clusters and aids in the interpretation of peptide primary sequence data during CAD in the region between the capillary exit and skimmer in the electrospray interface (8-15). Particularly valuable is the ability to determine monoisotopic molecular weight, as these values are independent of $^{12}C/^{13}C$ variations.

Relatively low resolution CAD/MS spectra, yielding characteristic a_n- b_n- and y_n-series ions, have been acquired in the region between the capillary exit and the skimmer region in an electrospray interface at a resolution of 1000 to 1500 (10% valley definition) for three model peptides (8). Starrett and DiDonato, working at a resolution of 5000 (5% valley definition) under voltage scanning conditions over a narrow mass range concentrated on the accurate mass measurement of product ions generated during CAD/MS (14). Use of an internal calibrant minimized differences between the theoretical and observed product ion masses to about 5 ppm for several peptides including, angiotensin II and substance P (fragment 1-9). In one case the full primary sequence of a peptide, human renin substrate, was determined with the exception that the Leucine and Isoleucine isomers could not be differentiated (8). The value of high resolution for the assignment of charge state for CAD/MS product ions was demonstrated by Loo's group in a paper focusing on the determination of protein structural information following electrospray introduction (13).

Larsen and McEwen (9) employed resolutions of 5000 and 10000 (10% valley definition) for accurate molecular weight determination and found that errors seldom exceeded 25 ppm for several pure peptides. Calibration was done internally and the isotopic cluster for the +5 charge state of insulin was resolved. Similar ppm errors were reported in a second paper for a number of standard peptides and proteins under lower resolution conditions during average molecular weight determinations. Average molecular weight accuracy was sufficient to allow for the differentiation of a single point modification differing by only 1 Da for biomolecules up to 20000 Da (11).

High resolution separation of a $(M+9H)^{9+}$ isotopic cluster was demonstrated by Cody, Tamura and Musselman for lysozyme at resolution of 10000 (10% valley). Errors associated with these measurements were in the 5 to 20 ppm range when internal calibration was employed (12). Higher errors, typically in the 5 to 90 ppm range, were observed during average molecular weight determinations when external calibration was used over a 10 hour period.

The molecular weights for a series of thirty-seven unknown synthetic peptides, used in research studies involving synthetic vaccines, antibacterial peptides or the de novo design of helical peptides and proteins, were determined with a magnetic sector instrument (15). All data were obtained with external calibration over a wide mass range during magnetic scanning. Errors between observed and theoretical monoisotopic molecular weights were typically in the 5 to 60 ppm range for the unknowns at sector resolutions between 2500 and 9000 (10% valley). Isotopic clusters for charge states up to +10 were resolved through the use of high resolution. CAD in the electrospray interface resulted in product ions that enabled either full or partial sequencing of most unknown peptides below 2000 Da. The complete primary sequence for one peptide was determined and the importance of high resolution was demonstrated by the differentiation of lysine from glutamine, two amino acids differing in residue mass by only 0.0364 Da. Two other peptides, with identical monoisotopic masses, but different primary sequences, were differentiated by CAD/MS.

Bioactive peptides, including bradykinins, substance P and bombesin have been analysed frequently by ESI-MS (16). Considerable effort has gone into analyses of bradykinin and substance P and these two bioactive peptides have been used extensively to demonstrate instrument performance. Most of this effort has involved the use of quadrupole instrumentation with no reports of the use of liquid chromatography prior to high resolution ESI-MS analysis with a magnetic sector instrument (5).

DRES, during development of analytical methods and a database of bioactive peptides, acquired ESI-MS data under high resolution LC-ESI-MS conditions, for a number of bioactive peptides including substance P (and related peptides), bradykinins, bombesins (and related peptides) and a Conus snail toxin. LC-ESI-MS data were acquired over a wide mass range by scanning the magnetic sector and calibrating externally with polyethylene glycol standards. Error between observed and theoretical monoisotopic molecular weights were typically in the 5 to 30 ppm range for the bioactive peptides with magnetic sector resolutions between 2500 and 4000 (10% valley

definition) and isotopic clusters for charge states of up to +5 were fully resolved. Under CAD/MS conditions both b_n- and y_n -series sequence ions were generally observed, enabling amino acid sequencing of intact peptides, degradation products and tryptic fragments.

EXPERIMENTAL

Samples

All the bioactive peptides with the exception of ω-conotoxin MVII-C were purchased from Sigma Chemical Company (St. Louis, MO, USA). ω-conotoxin MVII-C was purchased from Bachem (Torrance, CA, USA). Each peptide was dissolved to an initial concentration of 0.3 to 1 mg/mL in distilled water or distilled water containing 0.05% HPLC grade trifluoroacetic acid (TFA) (Aldrich, Milwaukee, WI, USA). Standard solutions were stored at −20°C and diluted by a factor of 10 with distilled water prior to analysis. Further dilutions were made to estimate detection limits.

Distilled-in-glass water was filtered through a 0.45 μm Millipore filter prior to use in the mobile phase or for diluting the peptide samples. Acetonitrile was Burdick and Jackson UV grade (Muskegon, MI, USA).

Tryptic digests were prepared by dissolving modified sequencing grade trypsin (Boehringer Mannheim, Mannheim, Germany) in 0.1M ammonium bicarbonate and adding this trypsin solution (50 μL of 0.1 mg/mL) to the peptide (200 μL of 0.5 mg/mL), giving an enzyme:substrate ratio of about 1:20. Digestions were carried out for about 18 hours at 37° C and were quenched with TFA (to pH 2-3) prior to LC-ESI-MS analyses.

Instrumental

All electrospray mass spectra were acquired using a VG (Fisons) Autospec-Q mass spectrometer (Manchester, UK) equipped with the VG (Fisons) Mark II electrospray interface. The electrospray needle was operated at 7.6 kV and ions were accelerated into the mass spectrometer at 4 kV. Sampling cone voltages up to 175 volts were initially investigated, with all subsequent data being acquired with sampling cone voltages in the 75 to 150 volt range. Nitrogen (Very Dry, Liquid Carbonic Inc., Scarborough, Ont., Canada) bath gas was introduced into the interface (80°C) at a flow rate of 500 L/hr. Nitrogen nebulizer gas was introduced at a flow rate of 14 L/hr. The electrospray interface was pumped with both a rotary and a turbomolecular pump, which enabled maintenance of $3x10^{-6}$ and $5x10^{-8}$ Torr within the source and analyser regions of the instrument, respectively.

Electrospray data were acquired in the continuum mode by scanning the magnet over a variety of mass ranges such that each scan took about six to eight seconds. This resulted in the acquisition of five to ten scans for each sample component during LC-ESI-MS analysis. These scans were averaged to enhance the signal-to-noise ratio and the data were smoothed using VG (Fisons) OPUS software. Resolutions of 2500 to 4000 (10% valley definition) were employed during magnetic sector scanning to facilitate accurate mass measurement of the ions formed during ESI-MS analyses. External calibrations were performed with solutions of polyethylene glycol in distilled water. Monoisotopic molecular weights for the bioactive peptides were calculated in triplicate from the observed $(M+nH)^{n+}$ ions, where the charge states were determined by the isotopic cluster spacings.

HPLC (LC) separations of the bioactive peptides were performed with an Applied Biosystems Model 140B dual syringe pump (Foster City, CA) equipped with a 15 cm × 0.53 mm i.d. C_{18} (5 μm) packed J&W DB-1 fused-silica capillary column (Courtesy of Mr. L. Hogge and Mr.

D. Olson, NRC, Saskatoon, Canada). The following solvent compositions were prepared for ESI-MS sample introduction: Solvent A (0.05% TFA in water) and Solvent B (0.05% TFA in acetonitrile/water (80:20)). Chromatographic separations were performed using a 5% to 65%B linear gradient over 30 or 60 minutes. In order to minimize dead volume and ensure reproducible mixing, the mobile phase was delivered at 200 μL/min and split prior to the injector such that the flow through the column was 20 μL/min.

RESULTS AND DISCUSSION

The ESI-MS data for most peptides and proteins have been acquired under low resolution conditions with quadrupole instruments, with relatively few examples of high resolution mass measurement with magnetic sector instruments (8-15). Of those reported, none appear to have been obtained following on-line separation by LC or other separatory techniques (5). A prior DRES study indicated the value of high resolution mass measurement for monoisotopic molecular weight determination, charge assignment of $(M+nH)^{n+}$ cluster and, perhaps most importantly, for the acquisition of amino acid sequence information during CAD/MS analyses (15). Demonstration of high resolution ESI-MS data acquisition for bioactive peptides following LC separations would be of considerable value, since samples arriving at DRES for analysis will likely consist of mixtures of peptides. The current study focussed on the demonstration of on-line high resolution LC-ESI-MS analysis for the identification of bioactive peptides with several goals:

a) the development of a ESI-MS database,

b) the acquisition of monoisotopic molecular weight information for bioactive peptides,

c) the identification of observed peptide degradation products and,

d) the amino acid sequencing of both known and unknown peptides using enzymatic digestions and CAD/MS procedures.

Bioactive Peptide Database

DRES maintains a comprehensive chemical warfare agent mass spectrometric database of electron impact and chemical ionization data for CW agent identification. A similar database containing ESI-MS data was envisioned for the toxins of biological origin, beginning with bioactive peptides. A number of commercially available candidate bioactive peptides (mid-spectrum agents) were analysed by high resolution LC-ESI-MS under sampling cone conditions that generally favor observation of isotopic cluster ions with little or no product ion formation. Table I lists the peptides analysed and the relative ion intensities for the ^{12}C containing ion in each $(M+nH)^{n+}$ isotopic cluster. In general $(M+H)^{+}$ isotopic clusters were most significant for lower mass peptides. Higher mass peptides typically stabilize more charge and tend to form isotopic clusters of higher charge state, provided sufficient basic sites (e.g., lysine) are available for protonation. This database, listed in Table I, will be continually updated with new peptide and protein ESI-MS data.

The ESI-MS data contained in Table I was acquired with a magnetic sector resolution of 3000 or 4000 (10% valley definition) over a wide mass range (e.g., 1300 to 300 Da). In each case the peptide or peptides to be characterized were introduced by LC. The chromatogram illustrated in Figure 1 was typical of a gradient separation (1%B/min) of a mixture of bioactive peptides. The acquired ESI-MS data for the three bradykinins contained $(M+nH)^{n+}$ isotopic clusters from which the monoisotopic molecular weight may be calculated. The monoisotopic molecular weight for each peptide was determined in triplicate using the ^{12}C ion in each of the +1, +2 and +3 isotopic clusters. Errors were in the 2 to 13 ppm range, demonstrating the mass accuracy of high resolution during LC-ESI-MS analysis. Monoisotopic molecular weight data were also acquired for the highest molecular peptide analysed, ω-conotoxin MVII-C, using a magnetic sector resolution of 4000 (10% valley definition). The +4 isotopic cluster was completely resolved with spacings of 0.25 Da and the calculated monoisotopic molecular weight was within 4 ppm of the theoretical value.

Detection limits during high resolution LC-ESI-MS were estimated for two bioactive peptides, Ile-Ser-bradykinin and substance P, at the 20 pmole level (Figure 3). In each case the signal-to-noise ratio (S/N) was in excess of 50:1, with calculated monoisotopic molecular weights being within 15 ppm of theoretical values. Based on this data a detection limit of 1 pmole (S/N = 5:1) was estimated .

Tryptic Digestion of Met-Lys-Bradykinin

The ESI-MS acquired for Met-Lys-Bradykinin with a sampling cone voltage of 125 volts provided monoisotopic molecular weight data in good agreement with theoretical data (Table II), but the product ion relative intensities were often low (Figure 4). A partial amino acid sequence was accessed based on the observation of two doubly charged product ions, y_{10}^{+2} (m/z 594.82) and y_9^{+2} (m/z 530.78) and a relatively weak series of y_n (n = 2 to 8) product ions at m/z 322.18, 419.23, 506.27, 653.32, 710.35, 807.41 and 904.43, respectively. Tryptic digestion was performed to provide peptide fragments that often give rise to more intense product ions during CAD/MS. Figure 5 illustrates the total-ion-current chromatogram obtained for the tryptic digest of this bioactive peptide. Three tryptic fragments, two of which were predicted by cleavage at the C-terminus of the lysine residue, were observed. The third component was identified following amino acid sequencing under CAD/MS conditions.

Figure 6 illustrates the ESI-MS data obtained for all three components with a sampling cone voltage of 75 volts. The monoisotopic molecular weights were within 25 ppm of predicted values (Table II) and the monoisotopic molecular weight of tryptic fragment 3 suggested the presence of an unexpected cleavage of an arginine amino acid residue from either the N- or C- terminus.

Under CAD/MS conditions, with a sampling cone voltage of 150 volts, it was possible to completely sequence tryptic fragment 2 (Figure 7). Overlap of the y_n- and b_n-series was observed

and amino acid residue losses typically within 0.008 ± 0.006 Da (n=11) of those predicted were observed for this tryptic fragment (Table II). The CAD/MS data for tryptic fragment 3 were less informative but the presence of a_6 (m/z 614.34) and a_5 (m/z 527.32) ions were indicative of loss of proline and phenylalanine and loss of serine, proline and phenylalanine, respectively, from the C-terminus. This was sufficient to confirm that the difference between tryptic fragments 2 and 3 was due to the loss of an arginine from the C-terminus of tryptic fragment 2.

Deamidation of Substance P (fragment 4-11)

During analysis of substance P (fragment 4-11) two additional components with slightly longer retention times were observed (Figure 8). The two additional components differed in monoisotopic molecular weight from substance P (fragment 4-11) by +0.985 Da and + 1.962 Da, respectively (Table III), with the most likely cause for the mass shifts being deamidation of the glutamine(s) in the peptide (17). Monoisotopic molecular weights determined during LC-ESI-MS analysis at a sector resolution of 3000 (10% valley definition) were consistent with this prediction, giving rise to errors of less than 10 ppm (Table III). Figure 9 illustrates the CAD/MS data obtained for substance P (fragment 4-11) and the two deamidation products. Partial amino acid sequence data were accessed for all three compounds in the form of singly and doubly charged b_n-series ions. The most complete series were observed for substance P (fragment 4-11), enabling sequencing of the first four amino acids (from the C-terminal) of this eight amino acid peptide. Unfortunately y_n-series were not observed so it was not possible to determine which of the two sites underwent deamidation in the first product (observed monoisotopic molecular weight: 966.468 Da). Both glutamine sites were converted for the second deamidation product (observed monoisotopic molecular weight: 967.445 Da), giving rise to a sequence of PEEFFGLM-NH_2. Deamidation of a glutamine amino acid residue to glutamic acid could be determined by more time consuming classical approaches to amino acid sequence determination (Edman degradation) since the observed modification leads to the formation of another amino acid residue. However many other modifications don't give rise to

another amino acid residue and as such would not be readily detected by classical means (17). Clearly mass spectrometry offers a considerable advantage for amino acid sequencing in cases involving modification .

Dimer Formation at Higher Concentration

At higher analyte concentrations (approximately 1000 pmoles) during preliminary ESI-MS analyses the data acquired in the $(M+H)^+$ region for a number of bioactive peptides exhibited the presence of additional ions separated by 0.5 Da. The ESI-MS data illustrated in Figure 10 for bradykinin at these levels was typical. At a resolution of 2500 (10% valley definition) the expected cluster of ions at m/z 1060.54, 1061.56 and 1062.56 were observed (Figure 10c), but additional ions at m/z 1061.05 and 1062.03 were also observed. Dimerization, resulting in the formation of $(M+H)_2^{2+}$ ions was suspected. The theoretical distributions for both the $(M+H)$ and $(M+H)_2$ species were generated (Figure 10a and 10b) and supported the assumption that a percentage of bradykinin $(M+H)^+$ ions dimerize during ESI-MS analysis at higher concentrations. As a result, the acquired of ESI-MS data contains the sum of contributions from both the $(M+H)^+$ and $(M+H)_2^{2+}$ isotopic clusters.

CONCLUSIONS

Considerable effort has gone into the ESI-MS analysis of bioactive peptides, but relatively little effort has been directed towards the identification of these compounds under high resolution conditions with a magnetic sector mass spectrometer. Past efforts have focussed on the use of loop injection and reports of chromatographic separation prior to ESI-MS analysis under high resolution conditions with magnetic sector instruments have not been reported.

A bioactive peptide database of high resolution ESI-MS data was established during the development of LC-ESI-MS analytical methods for a number of bioactive peptides including substance P (and related peptides), bradykinins, bombesins (and related peptides) and a Conus snail toxin. LC-ESI-MS data were acquired over a wide mass range by scanning the magnetic sector and calibrating externally with polyethylene glycol standards. Multiply charged ions were observed and errors between observed and theoretical monoisotopic molecular weights were typically in the 5 to 30 ppm range for the bioactive peptides with magnetic sector resolutions between 2500 and 4000 (10% valley definition). Isotopic clusters for charge states of up to +5 were fully resolved, facilitating the rapid and unambiguous assignment of charge states and calculation of monoisotopic molecular weights.

Under CAD/MS conditions both b_n- and y_n-series sequence ions were generally observed, enabling amino acid sequencing of both intact bioactive peptides and fragments generated following tryptic digestion. Over time bioactive peptides may degrade, leading to the formation of new products. During analysis of substance P (fragment 4-11) a degradation pathway involving the transformation of glutamine amino acid residue(s) to glutamic acid was observed and the products were identified on the basis of the monoisotopic molecular weight and partial amino acid sequence data acquired following high resolution LC-ESI-MS analysis. Finally dimerization of $(M+H)^+$, leading to the formation of $(M+H)_2^{2+}$ ions, was observed for higher peptide concentrations.

REFERENCES

1. Barber, M., Bordoli, R.S., Sedgwick, R.D. and Tyler, A.N., J. Chem. Soc. Chem. Commun., 325, 1981.

2. Meng, C.K., Mann, M. and Fenn, J.B., Z. Phys. D, **10**, 361, 1988.

3. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. and Whitehouse, C.M., Science, **246**, 64, 1989.

4. Biemann, K., Annu. Rev. Biochem., **61**, 977, 1992.

5. Burlingame, A.L., Boyd, R.K. and Gaskell, S.J., Anal. Chem., **66**, 634R, 1994.

6. Buchanan, M.V. and Hettich, R.L., Anal. Chem., **65**, 245A, 1993.

7. Meng, C.K., McEwen, C.N. and Larsen, B.S., Rapid Commun. Mass Spectrom., **4**, 147, 1990.

8. Meng, C.K., McEwen, C.N. and Larsen, B.S., Rapid Commun. Mass Spectrom., **4**, 151, 1990.

9. Larsen, B.S. and McEwen, C.N., J. Am. Soc. Mass Spectrom., **2**, 205, 1991.

10. Chapman, J.R., Gallagher, R.T., Barton, E.C., Curtis, J.M. and Derrick, P.J., Org. Mass Spectrom., **27**, 195, 1992.

11. McEwen, C.N. and Larsen, B.S., Rapid Commun. Mass Spectrom., **6**, 173, 1992.

12. Cody, R.B., Tamura, J. and Musselman, B.D., Anal. Chem., **64**, 1561, 1992.

13. Loo, J.A., Ogorzalek Loo, R.R. and Andrews, P.C., Org. Mass Spectrom., **28**, 1640, 1993.

14. Starrett, A.M. and DiDonato, G.C., Rapid Commun. Mass Spectrom., **7**, 12, 1993.

15. D'Agostino, P.A., Hancock, J.R., Provost, L.R., Semchuk, P.D. and Hodges, R.S., Rapid Commun. Mass Spectrom., **9**, 597, 1995.

16. Smith, R.D., Loo, J.A., Edmonds, C.G., Barinaga, C.J. and Udseth, H.R., Anal. Chem., **62**, 882, 1990.

17. Post-Translational Modifications of Peptides and Proteins (Data Interpretation Aid), Fisons Instruments, Manchester, UK, April 1993.

Table I: Database of Bioactive Peptide Electrospray Mass Spectral Data at Low Sampling Cone Voltage (75 volts).

Bioactive Peptide (Primary Sequence)	Monoisotopic Mol. Wt.	Electrospray MS Data (Relative Intensity) $(M+H)^+$	$(M+2H)^{2+}$	$(M+3H)^{3+}$
Substance P (7-11) (FFGLM-NH_2)	612.3094	100		
Substance P (4-11) (PQQFFGLM-NH_2)	965.4793	100	44	
Substance P (2-11) (PKPQQFFGLM-NH_2)	1190.6270	9	100	
Substance P (RPKPQQFFGLM-NH_2)	1346.7281	5	100	
Bradykinin (RPPGFSPFR)	1059.5614	5	100	6
Lys-Bradykinin (KRPPGFSPFR)	1187.6563	3	100	52
Ile-Ser-Bradykinin (ISRPPGFSPFR)	1259.6775	4	100	31
Met-Lys-Bradykinin (MKRPPGFSPFR)	1318.6968	2	100	79
Bombesin (8-14) (WAVGHLM-NH_2)	811.4163	41	100	
Bombesin (pEQRLGNQWAVGHLM-NH_2)	1618.8150	3	100	12
[Tyr^4]-Bombesin (pEQRYGNQWAVGHLM-NH_2)	1668.7942	5	100	13
		$(M+3H)^{3+}$	$(M+4H)^{4+}$	$(M+5H)^{5+}$
ω-Conotoxin MVII-C	2747.1576	100	75	7

Table II: Electrospray MS Data Obtained for Met-Lys-Bradykinin and the Tryptic Fragments of Met-Lys-Bradykinin.

a) Molecular Weight Determinations

Peptide Structure (Primary Sequence)	Observed Monoisotopic Molecular Weight[a]	Theoretical Monoisotopic Molecular Weight	Error (ppm)
Met-Lys-Bradykinin (MKRPPGFSPFR)	1318.701 ± 0.007 (n=3)	1318.6968	3
Tryptic Fragment 1 (MK)	278.1470	278.1538	24
Tryptic Fragment 2 (RPPGFSPFR)	1059.561 ± 0.013 (n=3)	1059.5614	0
Tryptic Fragment 3 (RPPGFSPF)	903.472 ± 0.008 (n=2)	903.4603	13

[a] Average of $(M+H)^+$, $(M+2H)^{2+}$ and/or $(M+3H)^{3+}$ data. Data were obtained at a magnetic sector resolution of 3000 and a sampling cone voltage of 75 volts.

b) Primary Sequence Data for Tryptic Fragment 2[b]

Series	m/z	Amino Acid Residue	Theoretical Residue Mass	Observed Residue Mass	Difference (Da)
y_8	904.4618	Arg (R)	156.10111	156.1045	0.003
y_7	807.4071	Pro (P)	97.05276	97.0547	0.002
y_6	710.3468	Pro (P)	97.05276	97.0603	0.008
y_5	653.3364	Gly (G)	57.02146	57.0104	0.011
y_4	506.2741	Phe (F)	147.06841	147.0623	0.006
y_3	419.2448	Ser (S)	87.03203	87.0293	0.003
y_2	322.1749	Pro (P)	97.05276	97.0699	0.017

Table II (con't): Electrospray MS Data Obtained for Met-Lys-Bradykinin and the Tryptic Fragments of Met-Lys-Bradykinin.

Series	m/z	Amino Acid Residue	Theoretical Residue Mass	Observed Residue Mass	Difference (Da)
b_8	886.4427	Arg (R) + H_2O	174.11168	174.1236	0.012
b_6	642.3392	Phe (F) + Pro (P)	244.12117	244.1035	0.018
b_5	555.3049	Ser (S)	87.03203	87.0343	0.002
b_4	408.2328	Phe (F)	147.06841	147.0721	0.004

[b] Data were obtained by scanning from 1200 to 250 Da (8 sec/decade) at a magnetic sector resolution of 3000 and a sampling cone voltage of 150 volts.

Table III: Electrospray MS Data Obtained for Substance P (fragment 4-11) and Two Decomposition Products Resulting from Deamidation of Glutamine (Gln, Q).

a) Molecular Weight Determinations

Peptide Structure Observed (Primary Sequence)	Monoisotopic Molecular Weight[a]	Theoretical Monoisotopic Molecular Weight	Error (ppm)
Substance P (4-11) (PQQFFGLM-NH_2)	965.483 ± 0.001	965.4793	4
Deamidation of Substance P (4-11) (P[QE or EQ]FFGLM-NH_2)	966.468 ± 0.010	966.4633	5
Deamidation of Substance P (4-11) (PEEFFGLM-NH_2)	967.445 ± 0.006	967.4473	2

[a] Average of $(M+H)^+$ and $(M+2H)^{2+}$ data. Data were obtained at a magnetic sector resolution of 3000 and a sampling cone voltage of 75 volts.

b) Primary Sequence Data for Substance P (fragment 4-11)[b]

Series	m/z	Amino Acid Residue	Theoretical Residue Mass	Observed Residue Mass	Difference (Da)
b_7	818.4270	Met (M) + NH_3	148.06704	148.0645	0.003
b_6	705.3451	Leu (L)	113.08406	113.0819	0.002
b_5	648.3013	Gly (G)	57.02146	57.0438	0.022
b_4	501.2289	Phe (F)	147.06841	147.0724	0.004
b_8^{2+}	475.2345	NH_3	17.02655	17.0284	0.002
b_7^{2+}	409.7069	Met (M)	131.04049	131.0552	0.015

[b] Data were obtained by scanning from 1025 to 400 Da (15 sec/decade) at a magnetic sector resolution of 3000 and a sampling cone voltage of 75 volts.

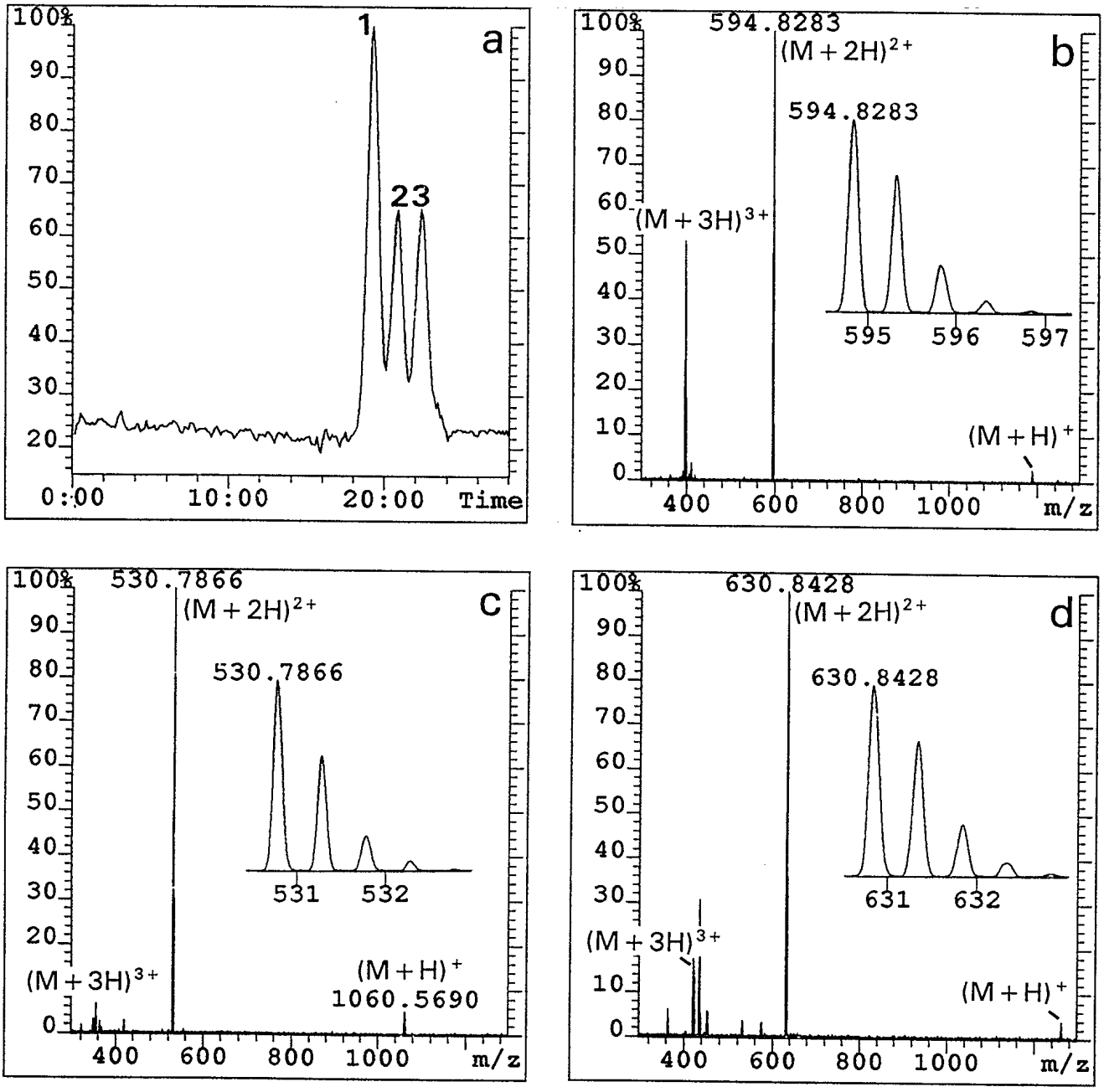

Figure 1: a) Capillary column LC-ESI-MS total-ion-current (1300 to 300 Da) chromatogram for Lys-Bradykinin, [1], Bradykinin, [2], and Ile-Ser-Bradykinin, [3] with a magnetic sector resolution of 3000 and a sampling cone voltage of 75 volts. ESI-MS data for: b) Lys-Bradykinin, Calculated Mol. Wt. 1187.64 ± 0.02 (n=3), 13 ppm error; c) Bradykinin, Calculated Mol. Wt. 1059.559 ± 0.003 (n=3), 2 ppm error; and d) Ile-Ser-Bradykinin, Calculated Mol. Wt. 1259.67 ± 0.03 (n=3), 5 ppm error.

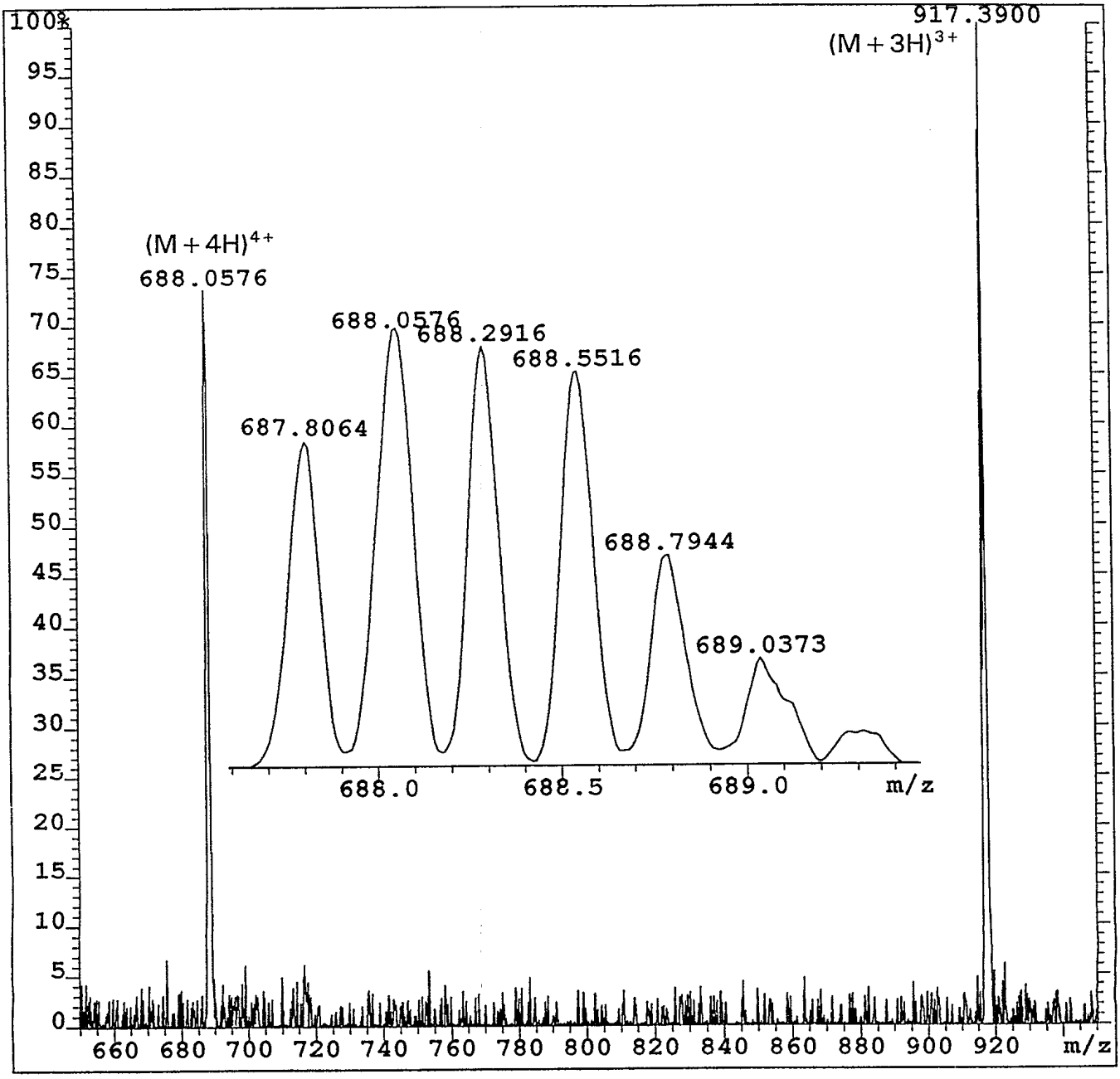

Figure 2: ESI-MS data for ω-Conotoxin MVII-C following capillary column LC-ESI-MS analysis (1025 to 400 Da) with a magnetic sector resolution of 4000 and a sampling cone voltage of 75 volts (Inset: Resolved $(M+4H)^{4+}$ isotopic cluster). The calculated Mol. Wt. for ω-Conotoxin MVII-C was 2747.17 ± 0.03 (n=2), (4 ppm error).

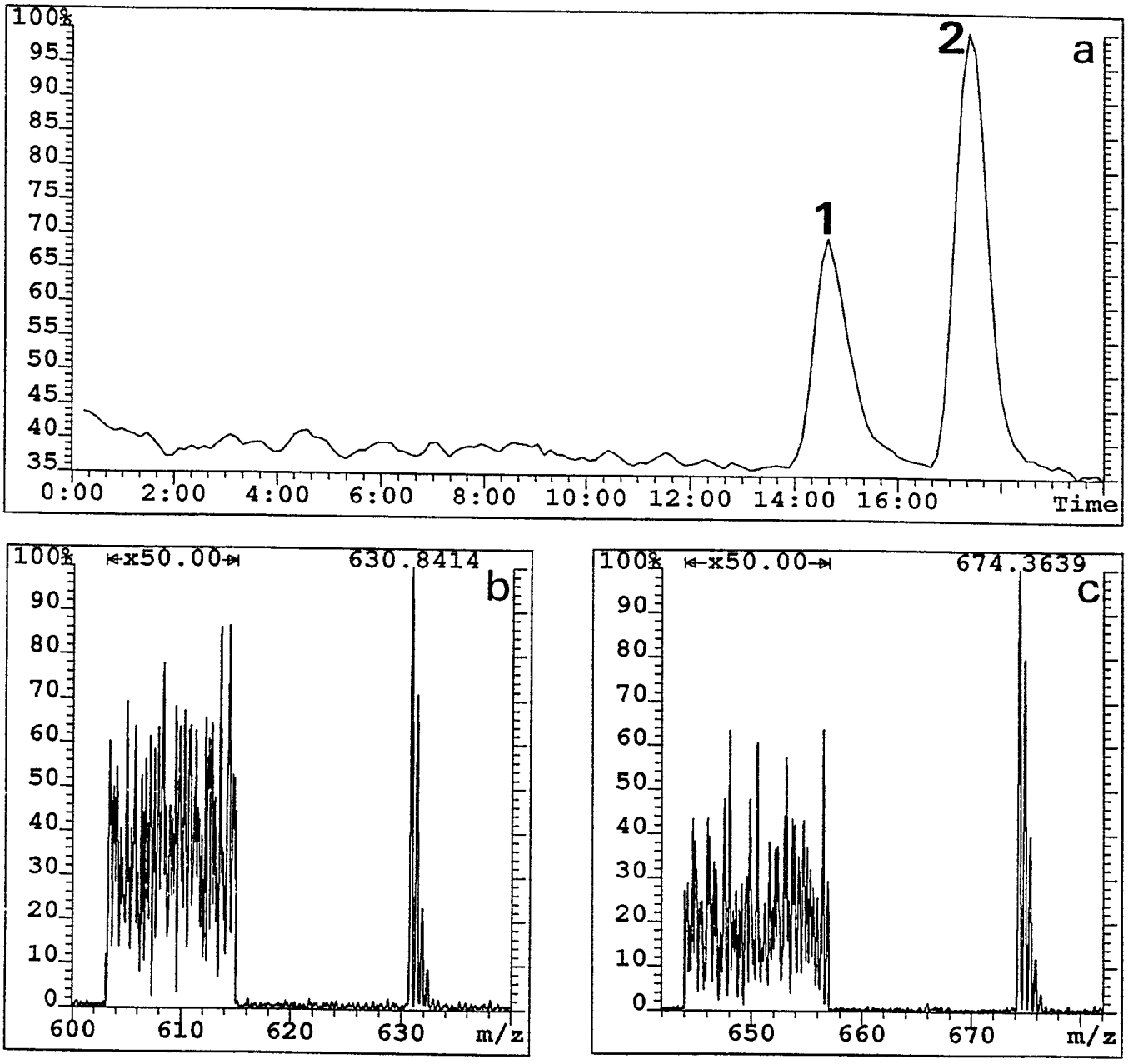

Figure 3: a) Capillary column LC-ESI-MS total-ion-current (720 to 600 Da) chromatogram for 20 pmoles of Ile-Ser-Bradykinin [1] and Substance P [2] with a magnetic sector resolution of 2500 and a sampling cone voltage of 75 volts. S/N was greater than 50:1 for: b) Ile-Ser-Bradykinin (calculated Mol. Wt. 1259. 667, 8 ppm error); and c) Substance P, (calculated Mol. Wt. 1346.712, 12 ppm error).

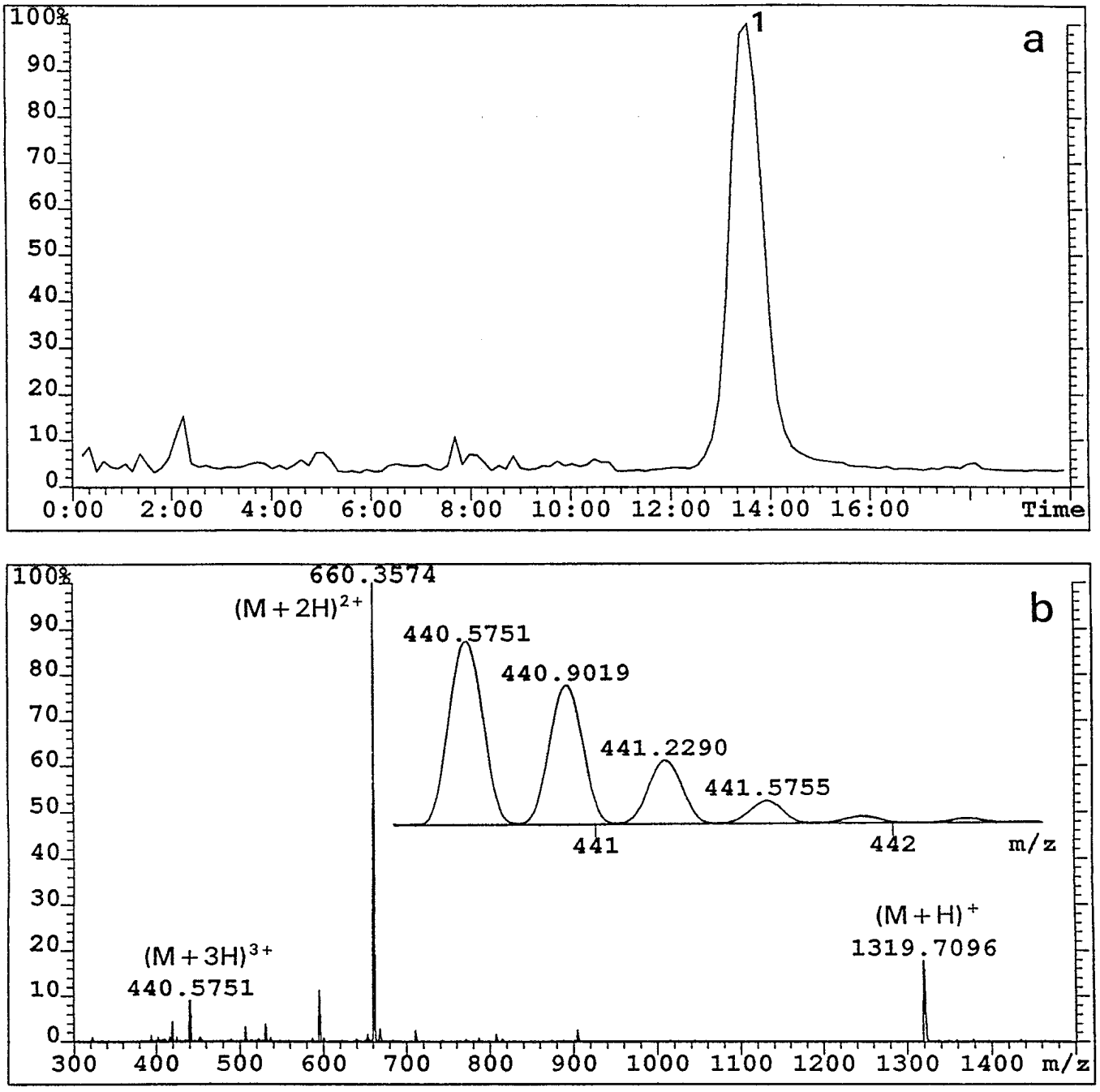

Figure 4: a) Capillary column LC-ESI-MS total-ion-current (1500 to 250 Da) chromatogram for Met-Lys-Bradykinin [1] with a magnetic sector resolution of 3000 and a sampling cone voltage of 125 volts. b) ESI-MS data for Met-Lys-Bradykinin (Inset: Resolved $(M+3H)^{3+}$ isotopic cluster). The calculated Mol. Wt. for Met-Lys-Bradykinin was 1318.701 ± 0.007 (n=3), (3 ppm error).

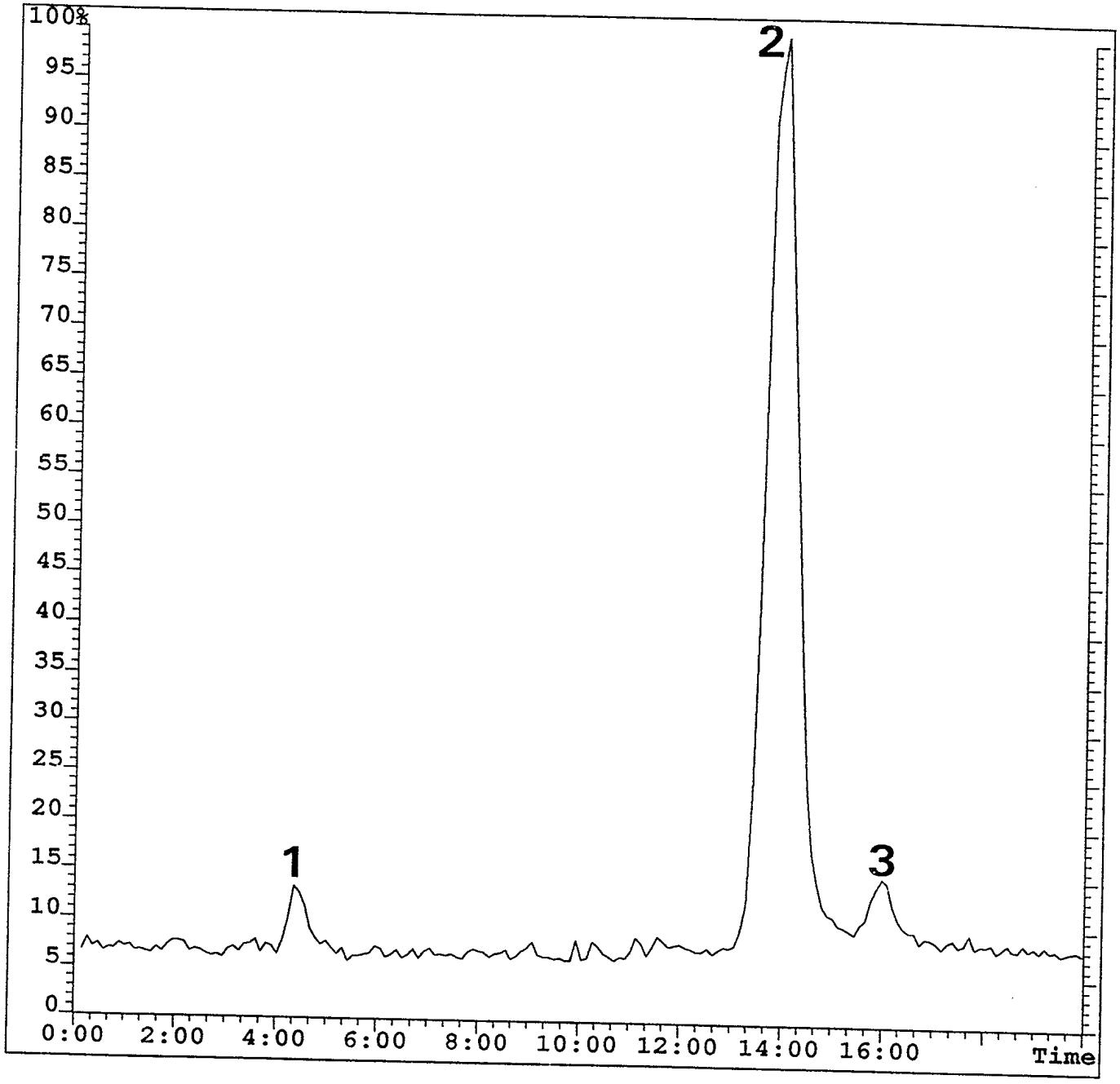

Figure 5: Capillary column LC-ESI-MS total-ion-current (1200 to 250 Da) chromatogram for tryptic digest of Met-Lys-Bradykinin with a magnetic sector resolution of 3000 and a sampling cone voltage of 75 volts. ESI-MS data for tryptic fragments 1, 2 and 3 are illustrated in Figure 6.

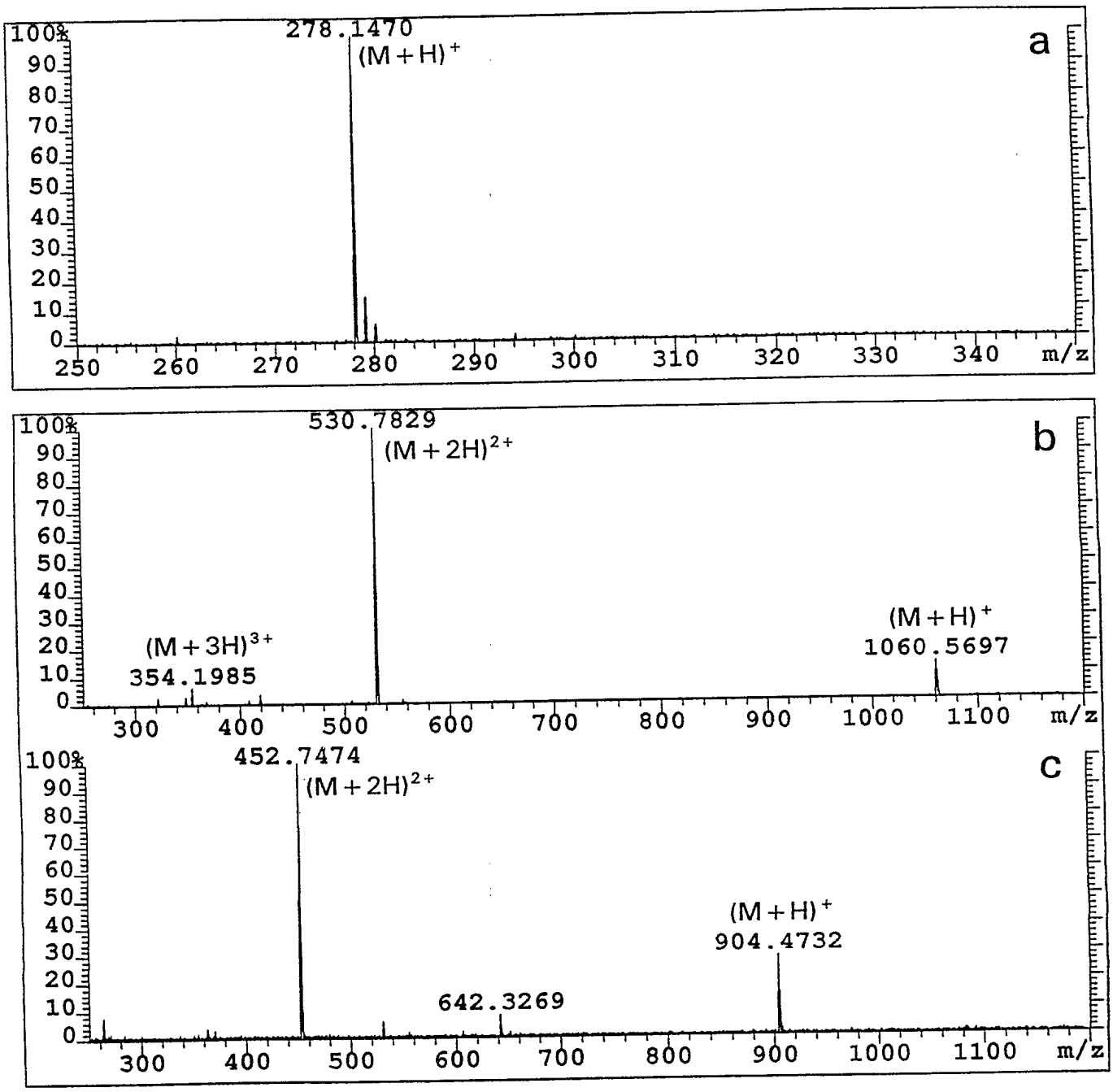

Figure 6: ESI-MS data for Met-Lys-Bradykinin (MKRPPGFSPFR) tryptic fragments a) 1 (MK), b) 2 (RPPGFSPFR) and c) 3 (RPPGFSPF) obtained following capillary column LC-ESI-MS analysis with a magnetic sector resolution of 3000 and a sampling cone voltage of 75 volts. Calculated molecular weight data are summarized in Table II.

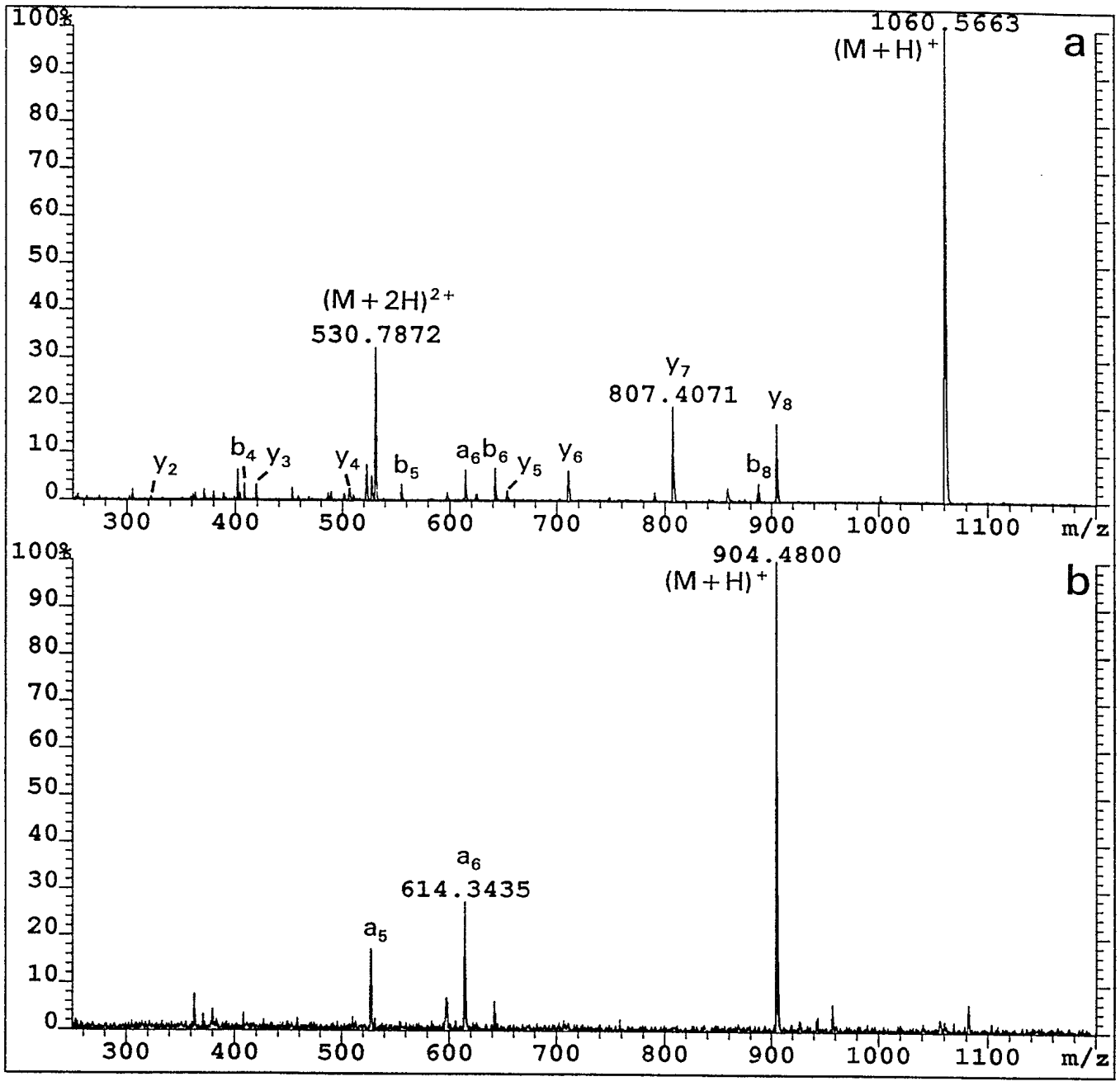

Figure 7: CAD/MS data for tryptic fragments a) 2 (RPPGFSPFR) and b) 3 (RPPGFSPF) obtained following capillary column LC-ESI-MS analysis with a magnetic sector resolution of 3000 and a sampling cone voltage of 150 volts (Refer to Table II).

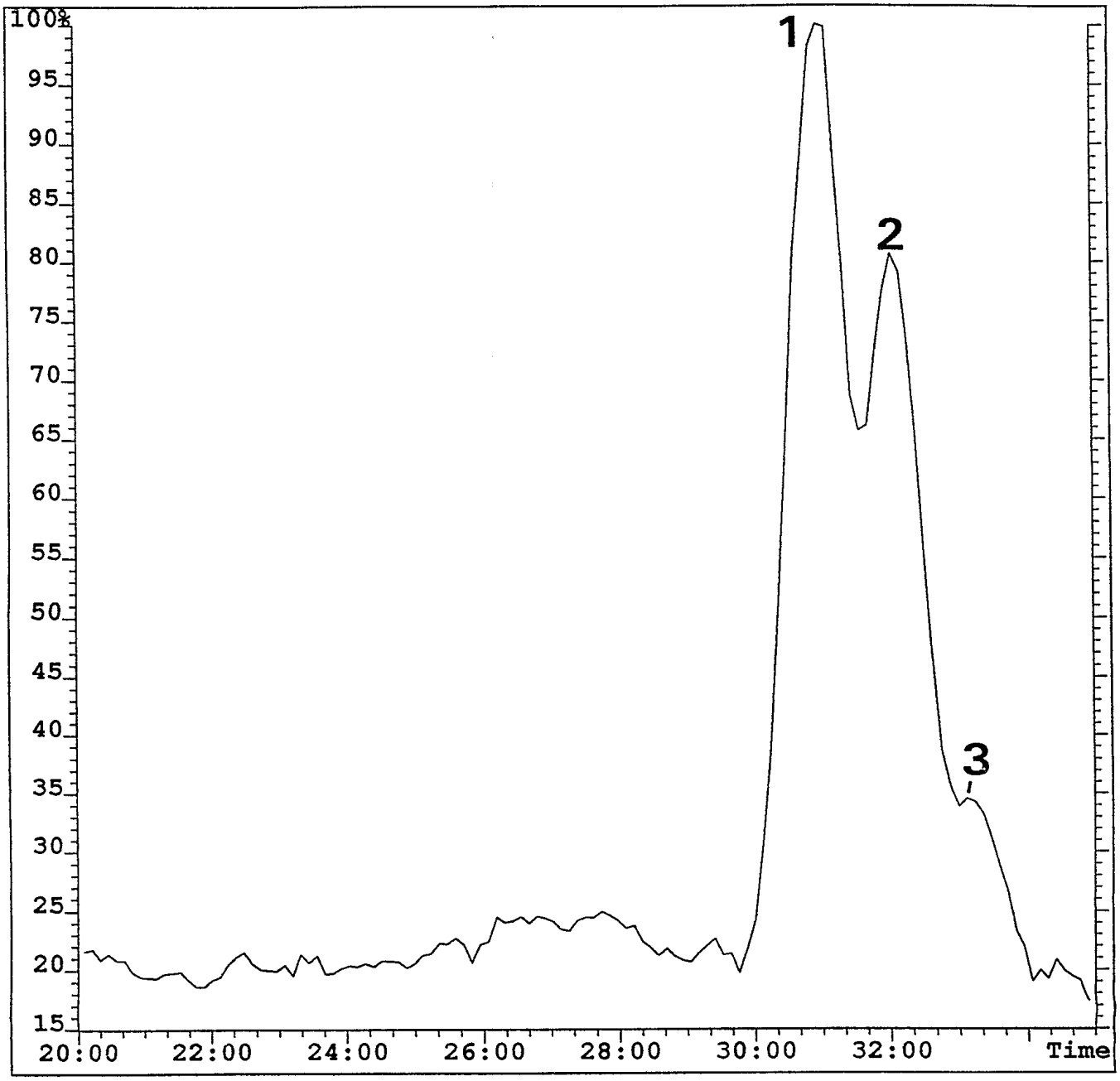

Figure 8: Capillary column LC-ESI-MS total-ion-current (1025 to 400 Da) chromatogram for Substance P (fragment 4-11), [1], and two deamidation products, [2] and [3], with a magnetic sector resolution of 3000 and a sampling cone voltage of 75 volts. ESI-MS data for the three sample components are illustrated in Figure 9.

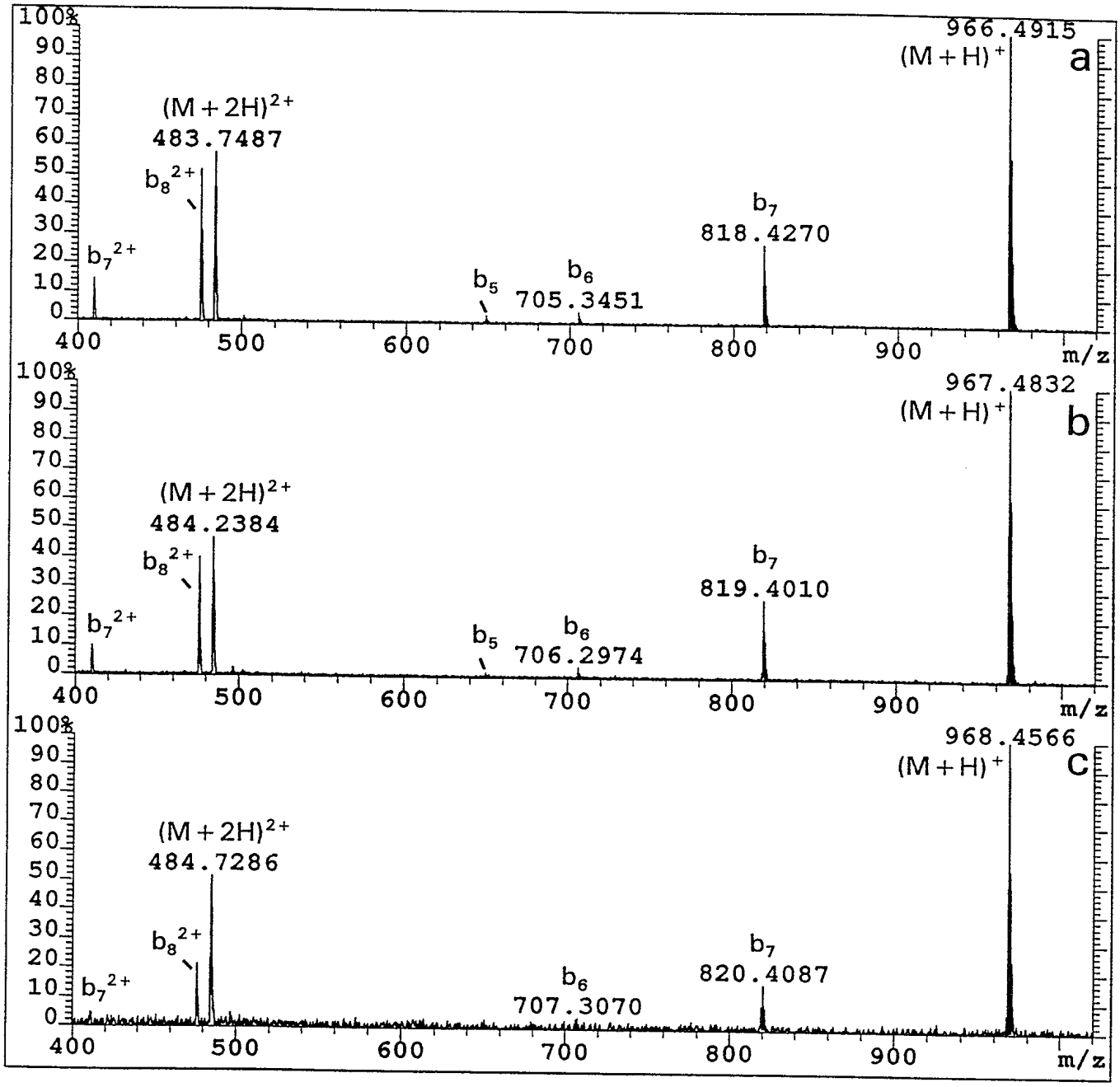

Figure 9: CAD-MS data for a) Substance P (fragment 4-11), ([1], PQQFFGLM-NH_2), b) the first deamidation product of Substance P (fragment 4-11), ([2], P[QE or EQ]FFGLM-NH_2) and c) the second deamidation product of Substance P (fragment 4-11), ([3], PEEFFGLM-NH_2) obtained following capillary column LC-ESI-MS analysis with a magnetic sector resolution of 3000 and a sampling cone voltage of 75 volts (Refer to Table III).

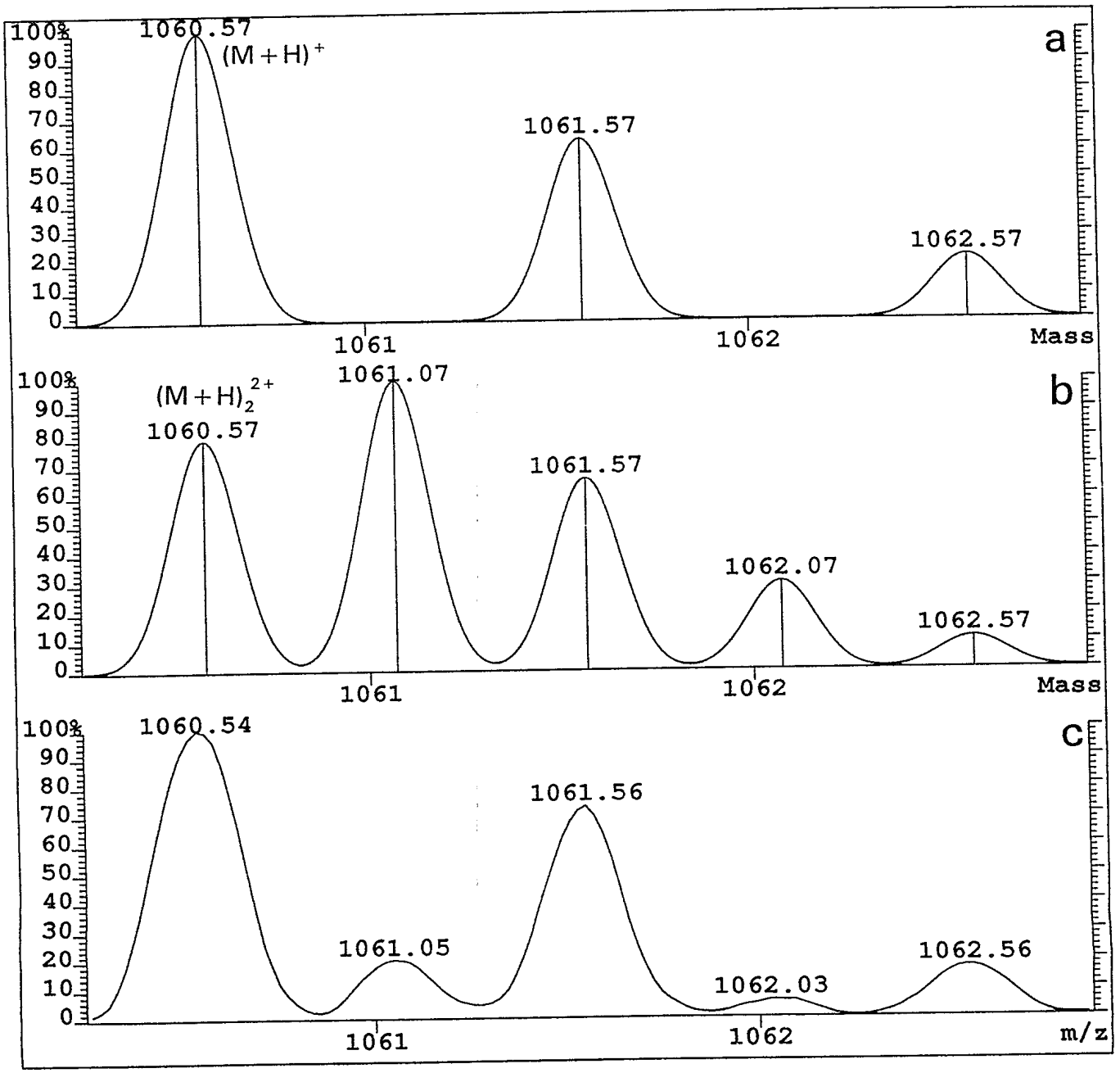

Figure 10: Theoretical isotopic distribution for a) the $(M+H)^+$ cluster and b) the $(M+H)_2^{2+}$ cluster for Bradykinin. c) Data acquired during ESI-MS analysis of approximately 1000 pmoles of Bradykinin that indicates some non-covalent interaction between two protonated Bradykinin molecules (magnetic sector resolution of 2500 and a sampling cone voltage of 75 volts).

UNCLASSIFIED

SECURITY CLASSIFICATION OF FORM
(highest classification of Title, Abstract, Keywords)

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

Defence Research Establishment Suffield

2. SECURITY CLASSIFICATION (overall security classification of the document. including special warning terms if applicable)

Unclassified

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title.)

Identification of Bioactive Peptides by High Resolution Liquid Chromatography - Electrospray Mass Spectrometry

4. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.)

D'Agostino, P.A., Hancock, J.R. and Provost, L.R.

5. DATE OF PUBLICATION (month and year of publication of document)

March 1996

6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)

31

6b. NO. OF REFS (total cited in document)

17

6. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Suffield Report

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

SR 637

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availabilty (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

UNCLASSIFIED

SECURITY CLASSIFICATION OF FORM

UNCLASSIFIED

SECURITY CLASSIFICATION OF FORM

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both offical languages unless the text is bilingual).

A bioactive peptide database of electospray mass spectra (ESI-MS) was established during the development of high resolution liquid chromatographic ESI-MS analytical methods with a magnetic sector instrument. High resolution ESI-MS data for a variety of bioactive peptides, including substance P (and related peptides), bradykinins, bombesins (and related peptides) and a Conus snail toxin, were acquired over a wide mass range by scanning the magnetic sector and calibrating externally with polyethylene glycol standards. Multiply charged ions were observed and errors between observed and theoretical monoisotopic molecular weights were typically in the 5 to 30 ppm range for the bioactive peptides with magnetic sector resolutions between 2500 and 4000 (10% valley definition). Isotopic clusters for charge states of up to +5 were fully resolved, facilitating the rapid and unambiguous assignment of charge states and calculation of monoisotopic molecular weights. Under CAD/MS conditions both b_n- and y_n-series sequence ions were generally observed, enabling either full or partial amino acid sequencing of bioactive peptides, their degradation products and fragments generated following tryptic digestion.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Toxin

Peptide

Protein

Mass Spectrometry

High Resolution

Liquid Chromotography

UNCLASSIFIED

* 498478

NO. OF COPIES NOMBRE DE COPIES	COPY NO. COPIE N°	INFORMATION SCIENTIST'S INITIALS INITIALES DE L'AGENT D'INFORMATION SCIENTIFIQUE
1	1	JC

AQUISITION ROUTE FOURNI PAR ▶	DRES
DATE ▶	27 May 1996
DSIS ACCESSION NO. NUMÉRO DSIS ▶	

DND 1158 (6-87)

National Defence / Défense nationale

PLEASE RETURN THIS DOCUMENT TO THE FOLLOWING ADDRESS:

DIRECTOR
SCIENTIFIC INFORMATION SERVICES
NATIONAL DEFENCE
HEADQUARTERS
OTTAWA, ONT. - CANADA K1A 0K2

PRIÈRE DE RETOURNER CE DOCUMENT À L'ADRESSE SUIVANTE:

DIRECTEUR
SERVICES D'INFORMATION SCIENTIFIQUES
QUARTIER GÉNÉRAL
DE LA DÉFENSE NATIONALE
OTTAWA, ONT. - CANADA K1A 0K2

UNCLASSIFIED