Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER
UNCLASSIFIED	152705

TITLE

HIGH RESOLUTION ELECTROSPRAY MAS SPECTROMETRY WITH A MAGNETIC SECTOR INSTRUMENT: ACCURATE MASS MEASUREMENT AND PEPTIDE SEQUENCING

System Number:

Patron Number:

Requester:

Notes:

DSIS Use only:

Deliver to:

RAPID COMMUNICATIONS IN MASS SPECTROMETRY, VOL. 9, 597–603 (1995)

High Resolution Electrospray Mass Spectrometry with a Magnetic Sector Instrument: Accurate Mass Measurement and Peptide Sequencing

P. A. D'Agostino,* J. R. Hancock and L. R. Provost
Defence Research Establishment Suffield, Box 4000, Medicine Hat, Alberta, Canada, T1A 8K6.

P. D. Semchuk and R. S. Hodges
Department of Biochemistry and Protein Engineering Network of Centres of Excellence, University of Alberta, Edmonton, Alberta, Canada, T6G 2H7

The accurate molecular weights for a series of 37 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. 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 5–60 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. Collisionally activated dissociation (CAD) in the electrospray interface resulted in product ions that enabled either full or partial sequencing of most unknown peptides of molecular weights 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 based on CAD-MS data.

In the early 1980s, 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 biomolecular 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.

The electrospray interface 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 sector use[7] was in large part driven by the potential to increase mass measurement accuracy. Particularly valuable is the ability to determine monoisotopic molecular weights, as these values are independent of $^{12}C/^{13}C$ variations. Use of high resolution enables the assignment of charge state to multiply charged molecular and product-ion isotopic clusters which aids in the interpretation of peptide primary sequence data acquired during collisionally activated dissociation (CAD) in the region between the capillary exit and skimmer in the electrospray interface.[8–14]

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) for three model peptides.[8] Starrett and DiDonato, working at a resolution of 5000 (5% valley) 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 a fragment of substance P. In one case, the full primary sequence of a peptide, human renin substrate, was determined with the exception that Leu (L) and Ile (I) could not be differentiated.[8] The value of high resolution, with a magnetic sector instrument, for the assignment of charge state for CAD-MS product ions was also demonstrated by Loo's group in a paper focusing on the determination of protein structural information following electrospray ionization.[13]

Larsen and McEwen employed resolutions of 5000 and 10 000 (10% valley) 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

* Author for correspondence.

CCC 0951–4198/95/070597–07

Received 14 February 1995
Accepted (revised) 19 April 1995

state of insulin was resolved.[9] 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 a resolution of 10000 (10% valley). Errors associated with these measurements were in the 5–20 ppm range when internal calibration was employed.[12] Higher errors, typically in the 5–90 ppm range, were observed during average molecular weight determinations when external calibration was used over a 10 h period.

Reports of accurate mass measurement by high resolution electrospray mass spectrometry with a magnetic sector instrument are limited, and all examples have dealt with the analysis of pure peptide standards, generally obtained from a commercial supplier. In these publications the accuracies reported represent an improvement over quadrupole data, where 100 ppm errors were typical.[15] We report the accurate molecular weight data for a series of 37 unknown synthetic peptides used in research studies involving synthetic vaccines, antibacterial peptides or the *de novo* design of helical peptides and proteins.[16–20] Differences between observed and theoretical monoisotopic molecular weights were typically in the 5–60 ppm range and sector resolutions between 2500 and 9000 (10% valley) were employed to resolve the isotopic clusters. One peptide was fully sequenced by CAD-MS and partial sequencing, typically three or more amino acid residues, was possible for most of the unknown peptides with molecular weights below 2000 Da.

EXPERIMENTAL

Samples

Thirty-seven peptides (Table 1), synthesized by Dr R. Hodge's group (member of Protein Engineering Network of Centres of Excellence, University of Alberta, Edmonton, Canada), were received as unknowns for analysis by electrospray mass spectrometry. The peptides, synthesized in support of research studies, were shipped and received at Defence Research Establishment Suffield (DRES) as frozen aqueous solutions (0.5 mL) in the 0.1 to 1.5 mg/mL range. All samples were kept frozen at −18 °C until just prior to electrospray mass spectrometric analysis, at which time individual samples were either analysed directly or diluted with distilled water to give a concentration of approximately 0.1 mg/mL.

Instrumental

All electrospray mass spectra were acquired using a VG (Fisons) (Manchester, UK) Autospec-Q mass spectrometer (with EBEQQ geometry) equipped with the recently released VG (Fisons) Mark II electrospray interface. The electrospray needle was operated at 7.8 kV and ions were accelerated into the mass spectrometer at 4 kV. The effect of varying sampling cone voltage (0 to 250 V) was investigated during earlier analyses of bioregulators and toxic peptides in the 500–2000 molecular weight range, including bombesin, bradykinins, substance P and its fragments. A good compromise between molecular-weight and fragmentation ion information was found at about 100 V and this voltage was used for all electrospray analyses of the peptide unknowns. Nitrogen (Very Dry, Liquid Carbonic Inc.) (Scarborough, Ontario, Canada) bath gas was introduced into the interface (80 °C) at a flow rate of 500 L/h. Nitrogen nebulizer gas was introduced at a flow rate of 14 L/h. The electrospray interface was pumped with both a rotary and a turbomolecular pump, which enabled maintenance of 2×10^{-6} and 3×10^{-8} Torr within the source and analyser regions of the instrument, respectively.

Electrospray data were acquired in the continuum mode by scanning the magnet from 1550 to 400 Da or 1000 to 400 Da exponentially at a scan rate of 25 and 30 s/decade, respectively. Five to ten scans were averaged to enhance the signal-to-noise ratio and the data were smoothed using VG (Fisons) OPUS software, with a smooth number of 2 and a window of 7. Monoisotopic molecular weights for all but two samples (9 and 30) were calculated in triplicate from $[M+nH]^{n+}$ and/or $[M+nH+m\mathrm{Na}]^{(n+m)+}$ ions, where the charge states were determined by the isotopic cluster spacings. Average molecular weights were determined in a similar manner in triplicate for the two higher molecular weight peptide samples, 9 and 30, due to incomplete resolution of the isotope clusters. The following resolutions, based on the 10% valley definition, were employed:

Peptide mol. wt.	*Resolution employed*
1000–2000	2500
2000–3000	3000
3000–4000	4500
4000–6000	6000
7000–8000	8000
>8000	9000

External calibrations were performed with a polyethylene glycol (PEG) standard containing both PEG 600 and PEG 1540 (0.5 mg/mL in distilled water).

Loop injections of the samples were made under conditions that simulated those employed during liquid chromatographic separations. An Applied Biosystems (Foster City, CA, USA) dual syringe pump was used to mix and deliver 20 μL/min of a 1:1 mixture of A and B: A, 100% water with 0.05% trifluoroacetic acid and B, 20% water + 80% acetonitrile with 0.05% trifluoroacetic acid. Two Rheodyne 8125 sample injectors (Cotati, CA, USA) were placed in series with a 7 μm, 10 cm × 1 mm i.d. Aquapore RP300 column (Applied Biosystems) located after the first injector and before the second. The principal advantage of this liquid chromatographic configuration was that it allowed the analyst the flexibility either of sample introduction through the first injector and column or loop injection. Data were acquired following a 50 μL injection of the peptide or calibration sample into the second injector.

Water was distilled in glass and filtered through a 0.45 μm Millipore (Bedford, MA, USA) filter prior to use in the mobile phase or for diluting the provided peptide samples. Acetonitrile was Burdick & Jackson (Muskegon, MI, USA) UV grade and trifluoroacetic

Table 1. Analysis of peptide samples received from University of Alberta

Peptide number	Primary sequence	Theoretical monoisotopic molecular weight	Observed monoisotopic molecular weight Mean ± standard deviation ($n = 3$)	Difference
1	Ac-KCEALEGKLGAVEEKLGAV EEKLGAVEEKLEALEG-NH_2	3708.9494	3708.99 ± 0.02	+ 0.04 (11)[a]
3	Ac-KCEALEGKLEEVGAKLEEV GAKLEEVGAKLEALEG-NH_2	3708.9494	3708.97 ± 0.07	+ 0.02 (5)
4	Ac-EAEALKKEIEALKK-NH_2	1639.9508	1639.99 ± 0.02	+ 0.04 (24)
5	Ac-EIEALKKEIEALKK-NH_2	1681.9978	1682.06 ± 0.01	+ 0.06 (36)
6	Ac-EIAELKKEIAELKK-NH_2	1681.9978	1682.04 ± 0.01	+ 0.04 (24)
7	Ac-EALKKEIEALKK-NH_2	1439.8711	1439.92 ± 0.01	+ 0.05 (35)
8	Ac-KCGALEEKLGALEEKAGAL EEKLGALEEKLGALEE-NH_2	3708.9494	3709.01 ± 0.03	+ 0.06 (16)
9	(Ac-EIETLRAEYEALKHEVESLR DEIEFLKAEVEACRS-NH_2)$_2$ Dimer with C-C bridge	8353.2904[b]	8352.78 ± 0.15[b]	− 0.51 (6)
11	Ac-ELEKLLKELEKLLKEKEK-NH_2	2280.3668	2280.34 ± 0.04	− 0.03 (13)
12	Ac-ELEKLLKEQEKLLKELEK-NH_2	2266.3511	2266.34 ± 0.03	− 0.01 (4)
14	Ac-SCATTVDAAFRPNGCTD Internal C-C bridge	1767.7193	1767.86 ± 0.02	+ 0.14 (79)
15	Ac-ELEKLLKECEKLLKELEK-NH_2	2255.2810	2255.26 ± 0.04	− 0.02 (9)
16	Ac-ELEKLLKEYEKLLKELEK-NH_2	2315.3351	2315.32 ± 0.04	− 0.02 (9)
17	Ac-EEAQAQAEAEAKAKAEEK-NH_2	1970.9545	1971.14 ± 0.01	+ 0.19 (96)
18	Ac-ELEKLLKEWEKLLKELEK-NH_2	2338.3511	2338.34 ± 0.03	− 0.01 (4)
19	Ac-ACKSTQDPMFTPKGCDN Internal C-C bridge	1881.7696	1881.98 ± 0.03	+ 0.21 (112)
20	FPVKLFPVKL (ring)	1168.7372	1168.90 ± 0.01	+ 0.16 (137)
21	FPVKLFPVKL	1186.7478	1186.90 ± 0.01	+ 0.15 (126)
22	Ac-ACASTQDPMFTPKGCDN Internal C-C bridge	1824.7117	1824.79 ± 0.02	+ 0.08 (44)
23	Ac-ACKSTQDPMFAPKGCDN Internal C-C bridge	1851.7590	1851.85 ± 0.02	+ 0.09 (49)
24	Ac-ACKSTQDPMFTAKGCDN Internal C-C bridge	1855.7539	1855.87 ± 0.02	+ 0.12 (65)
25	Ac-ACKSTQDPMFTPKGCDA Internal C-C bridge	1838.7638	1838.87 ± 0.02	+ 0.11 (60)
26	ATKKEVPLGVAADANKLG	1781.0046	1781.00 ± 0.02	0.00 (0.3)
27	BB-GGDEQFIPK-NH_2	1196.5502	1196.55 ± 0.02	0.00 (0.3)
28	Ac-EIEALKAEIEALKAGGDEQF IPKGGEIEALKAEIEALKA-NH_2	4162.2412	4162.30 ± 0.02	+ 0.06 (15)
29	ATKKEVPLGVAADANKLGEIEAL KAEIEALKAGGDEQFIPKGGEIEA LKAEIEALKA-NH_2	5883.2247	5883.44 ± 0.12	+ 0.22 (37)
30	(Ac-EIEALKAEIEALKAEIEALK AGGDEQFIPKGGEIEALKAEIEAL KAEIEACKA-NH_2)$_2$ Dimer with C-C bridge	11326.9950[b]	11326.33 ± 0.15[b]	− 0.67 (59)
31	Ac-KCGALEKKLGALEKKAGA LEKKLGALEKKLGALEK-NH_2	3704.2113	3704.24 ± 0.03	+ 0.03 (8)
32	(Ac-QCGALQKQVGALQKQVGA LQKQVGALQKQVGALQK-NH_2)$_2$ Dimer with C-C bridge	7368.1718	7368.36 ± 0.08	+ 0.19 (26)
33	(Ac-ECGALQKQVGALEKEVGA LQKQVGALEKEVGALEK-NH_2)$_2$ Dimer with C-C bridge	7379.9800	7380.27 ± 0.07	+ 0.29 (39)
34	(Ac-ECGALEKEVGALEKEVGA LEKQVGALQKQVGALQK-NH_2)$_2$ Dimer with C-C bridge	7379.9800	7380.16 ± 0.22	+ 0.18 (24)
35	Ac-ECGALEKEVGALEKEVGAL EKEVGALEKQVGALQK-NH_2	3692.9658	3693.04 ± 0.03	+ 0.07 (19)
36	(Ac-ECGALEKQVGALQKQVGA LQKQVGALQKQVGALQK-NH_2)$_2$ Dimer with C-C bridge	7372.1078	7372.45 ± 0.05	+ 0.34 (46)
37	(Ac-ECGALEKEVGALEKQVGA LQKQVGALQKQVGALQK-NH_2)$_2$ Dimer with C-C bridge	7376.0438	7376.36 ± 0.13	+ 0.32 (43)
38	(Ac-QCGALQKQVGALQKEVGA LEKQVGALQKQVGALQK-NH_2)$_2$ Dimer with C-C bridge	7372.1078	7372.26 ± 0.08	+ 0.15 (20)
39	Ac-QVGALKEKVGALQEQVGA LKEKVGALQEQVGACKE-NH_2	3675.9981	3676.04 ± 0.01	+ 0.04 (11)
40	(Ac-QVGALEKEVGALQKQVGA LEKEVGALQKQVGACEK-NH_2)$_2$ Dimer with C-C bridge	7349.9806	7350.17 ± 0.06	+ 0.19 (26)

[a] Da (ppm).
[b] Average molecular weight.

acid was high performance liquid chomatographic (HPLC) grade (Pierce, Rockford, IL, USA).

RESULTS AND DISCUSSION

Thirty-seven synthetic peptides, used in research studies involving synthetic vaccines, antibacterial peptides or the *de novo* design of helical peptides and proteins,[16–20] were provided as unknowns to DRES for analysis by electrospray mass spectrometry. The data obtained using this technique provide primarily molecular weight information but when used with a higher cone voltage setting or under tandem mass spectrometric (MS/MS) conditions, full or partial primary sequence data may be determined. The acquisition of both accurate molecular weight and primary sequence data for a given peptide is particularly important during analyses of unknowns, as the data allow the analyst to identify the peptide with a greater level of certainty. The principal objective of the present study was the accurate determination of the monoisotopic molecular weights of the unknown peptides, through the use of high resolution electrospray mass spectrometry. Secondary to this goal, was the determination of full or partial primary sequence information for the unknown peptides.

Each peptide was analysed under electrospray with a cone voltage setting of 100 V. This condition resulted in the formation of the $[M+nH]^{n+}$ series of ions for each peptide, in some instances sodium adduct ions, (e.g., $[M+nH+mNa]^{(n+m)+}$) and structurally significant product ions. The resolution was set at a level such that the isotopic cluster ions were well separated. The spacing between ions in each cluster relates directly to the charge state of the cluster; the spacing on the m/z scale in a mass spectrum for $[M+nH]^{n+}$ ions would be $1/n$ Da.

Table 1 contains the observed monoisotopic molecular weight data for all but two of the unknown peptides analysed and compares the acquired data with the theoretical values. Average molecular weight data were calculated for these two peptides, as the higher molecular weights of these compounds resulted in incomplete resolution of the isotopic clusters. Monoisotopic molecular weights were calculated from three different ion clusters in the electrospray mass spectrum, to illustrate the degree of precision that could be expected during typical analyses with external calibration. With the exception of peptide samples 14, 17, 19, 20 and 21, the accuracies, indicated in the difference column, were typically in the 5–60 ppm range.

Following submission of the observed monoisotopic molecular weight data, the University of Alberta provided us with the peptide primary sequences and theoretical monoisotopic molecular weight data. Subsequent inspection of the raw electrospray mass spectrometric data indicated that the five peptide samples 14, 17, 19, 20 and 21 were all acquired in the same data set, using the same external calibration file. No other peptide samples were acquired in this data set. Review of the raw data indicated more variability in the acquired data for the PEG standard, probably due to minor magnetic scanning fluctuations during the course of acquisition. This resulted in a poorer than expected calibration for these five peptides and all monoisotopic molecular weight differences were about 0.2 Da (80 to 140 ppm) above theoretical values. Following the study, one of the these peptides, 21, was selected for reanalysis with a resolution of 2500. External calibration was performed and the monoisotopic molecular weight differences for duplicate analyses of this peptide were consistent with the 5–60 ppm range stated above. Increased accuracy would be possible if internal calibration under voltage scanning conditions were employed.[14] However, this means of calibration increases analysis times, making it less suitable for routine screening of unknowns over a large mass scale.

One peptide was fully sequenced and partial sequencing, typically of three or more amino acid residues, was possible for most of the unknown peptides with molecular weights below 2000 Da. The structurally significant product ions observed were the result of collisionally activated dissociation in the electrospray interface. Selection of a relatively high voltage difference of 100 V increased the residence time for the multiply charged ions formed during the electrospray process. This increased residence time resulted in increased fragmentation, due to more opportunities for collision with neutrals in this region. Characteristic ions of the b_n- or y_n-series type were generally detected.

For example, unknown peptide 27 (Fig. 1, Table 2) could be completely sequenced based on the acquired CAD-MS data, with the exception that Leu (L) and Ile (I) could not be differentiated. Prior to receiving the sequence data it was possible to assign all the amino acids in the primary sequence, but the identity of the N-terminal blocking group was uncertain. The benzoyl benzoyl (BB) could only be confidently assigned following receipt of the actual primary sequence data.

Figure 1 illustrates the electrospray mass spectrum acquired under cone voltage conditions that facilitated production of diagnostic product ions. Overlap of the b_n- and y_n-series singly charged fragmentation ions enabled identification of this unknown as a peptide with

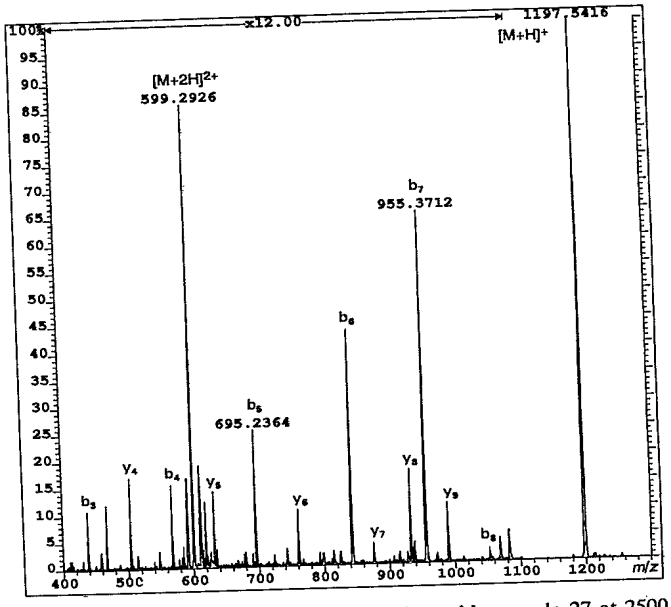

Figure 1. Electrospray mass spectrum of peptide sample 27 at 2500 resolution. Interpretation of the b_n- and y_n-series ions enabled complete sequencing of this unknown (Refer to Table 2). Observed monoisotopic molecular weight: 1196.55 ± 0.02; sequence: BB-GGDEQFIPK-NH_2.

Table 2. Sequence determination for peptide sample 27

Series	m/z	Amino acid residue	Theoretical residue mass	Observed residue mass	Difference
b_8	1052.4226	Lys + NH_3 (Gln + NH_3)	145.1215 (145.0851)	145.1190 (145.1190)	−0.0025 (+0.0339)
b_7	955.3712	Pro	97.0528	97.0514	−0.0014
b_6	842.2960	Ile (Leu)	113.0805	113.0752	−0.0089
b_5	695.2364	Phe	147.0684	147.0596	−0.0088
b_4	567.1774	Gln (Lys)	128.0586 (128.0950)	128.0590 (128.0590)	+0.0004 (−0.0360)
b_3	438.1435	Glu	129.0426	129.0339	−0.0087
y_9	989.5035				
y_8	932.4641	Gly	57.0215	57.0394	+0.0179
y_7	875.4632	Gly	57.0215	57.0009	−0.0206
y_6	760.4436	Asp	115.0269	115.0196	−0.0073
y_5	631.4058	Glu	129.0426	129.0378	−0.0048
y_4	503.3389	Gln (Lys)	128.0586 (128.0950)	128.0669 (128.0669)	+0.0083 (−0.0281)

Peptide sequence: BB-GGDEQFIPK-NH_2
BB (benzoyl benzoyl crosslinking group)

the following primary sequence, BB-GGDEQFIPK-NH_2. High resolution enabled confident assignment of the charge state of all the ions in the electrospray mass spectrum, based on the isotopic cluster spacings. The residue masses for the amino acids were sufficiently accurate to distinguish between Lys (K) and Gln (Q), two amino acids that differ by only 0.0364 Da. This finding was confirmed in both the b_n- and y_n-series ions, and both Gln (Q) and Lys (K) were correctly assigned. The alternative possibility in each case, given in parentheses in Table 2, exceeds by a factor of three to four the typical amino acid residue mass measurement errors (0.008 ± 0.006 Da, $n = 11$) obtained during analysis of this peptide sample.

Two of the unknown peptides, 5 and 6, were partially sequenced and exhibited within experimental uncertainty the same monoisotopic molecular weight (1682.06 and 1682.04 Da, respectively). The two peptides differ only in the relative positions of the amino acids, Glu (E) and Ala (A) at the third and fourth, and tenth and eleventh amino acid positions, where the first position refers to the N-terminal amino acid. Both peptides could be differentiated, based on the interpretation of the multiply charged y_n^{2+} and y_n^{3+} series ions observed (Fig. 2, Table 3). The charge state of the characteristic series ions could be confidently assigned, based on the charge spacing (1/2 Da or 1/3 Da) between the isotopic cluster ions for each. In both cases the first four amino acids in the sequence, beginning from the N-terminal, were confirmed at a resolution of 2500. Errors were consistent with the previous example, where the difference between the observed and theoretical amino acid residue losses were 0.009 ± 0.004 Da ($n = 6$) and 0.005 ± 0.004 Da ($n = 6$) for peptide samples 5 and 6, respectively.

The data obtained for peptide sample 7 (Ac-EALKKEIEALKK-NH_2) at a resolution of 2500 were sufficient for the assignment of the charge state of all ion clusters in this mass spectrum. Ions with 1/2 Da spacing were observed for the $[M + 2H]^{2+}$ isotopic cluster at m/z 720.965 and an impurity was postulated based on the observation of a second $[M + 2H]^{2+}$ cluster at m/z 656.916. The mass difference between the principal component, 7, and the impurity was 128.098 Da. Lys (K) has a monoisotopic mass of 128.09496 and this residue occupies several positions in the sequence in a repeating manner. It seems probable, based on the high resolution evidence, that during peptide synthesis one Lys (K) was not incorporated in the sequence. Further LC/MS experiments at the University of Alberta confirmed this conclusion and indicated that the impurity co-eluted with the analyte.

Figures 3 and 4 illustrate the utility of high resolution electrospray mass spectrometry for assigning the charge state of molecular ion isotopic clusters. Increased resolution is required with increasing mass to resolve the isotopic clusters formed during the electrospray process. Resolution of $[M + nH]^{n+}$ cluster ions differing in mass by 1/8 and 1/10 Da were possible and have been expanded and illustrated below the full electrospray mass spectra obtained for each of the examples. Determination of charge state may be made based solely on the spacing, or by simply counting the number of ions over a 1 Da range in the expanded portions of the electrospray mass spectra.

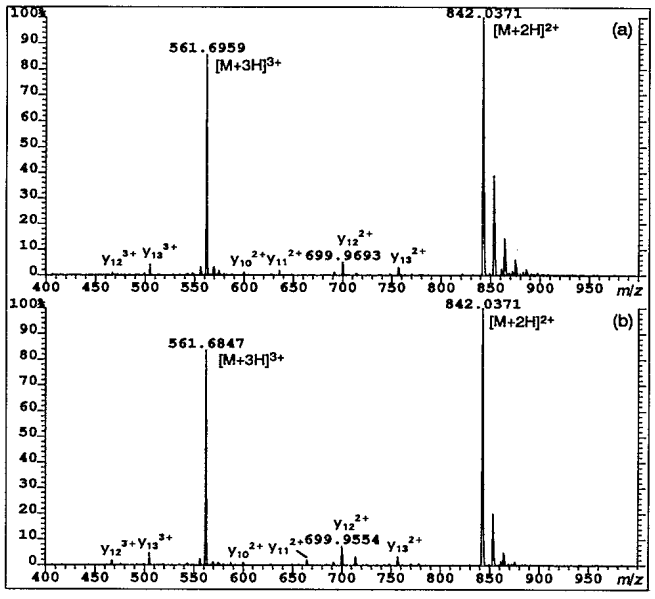

Figure 2. Electrospray mass spectra of peptide samples (a) 5 and (b) 6 at 2500 resolution. Interpretation of the y_n^{n+}-series ions enabled differentiation of the two peptides with identical monoisotopic molecular weights (Refer to Table 3). Sample 5 observed monoisotopic molecular weight: 1682.06 ± 0.01; sequence: Ac-EIEALKKEIEALKK-NH_2. Sample 6 observed monoisotopic molecular weight: 1682.04 ± 0.01; sequence: Ac-EIAELKKEIAELKK-NH_2.

Table 3. Series ions for peptides: (a) sample 5 and (b) sample 6

Series	*m/z*	Amino acid residue	Theoretical residue mass	Observed residue mass	Difference
(a) Sample 5					
y_{13}^{2+}	756.5055	Glu + C_2H_2O	85.5266	85.5316	+0.0050
y_{12}^{2+}	699.9693	Ile (Leu)	56.5420	56.5362	−0.0058
y_{11}^{2+}	635.4377	Glu	64.5213	64.5316	+0.0103
y_{10}^{2+}	599.9374	Ala	35.5186	35.5003	−0.0183
y_{13}^{3+}	504.6703	Glu + C_2H_2O	57.0177	57.0256	+0.0079
y_{12}^{3+}	466.9676	Ile (Leu)	37.6947	37.7027	+0.0080
(b) Sample 6					
y_{13}^{2+}	756.5055	Glu + C_2H_2O	85.5266	85.5316	+0.0050
y_{12}^{2+}	699.9554	Ile (Leu)	56.5420	56.5501	+0.0081
y_{11}^{2+}	664.4480	Ala	35.5186	35.5074	−0.0112
y_{10}^{2+}	599.9254	Glu	64.5213	64.5226	+0.0013
y_{13}^{3+}	504.6703	Glu + C_2H_2O	57.0177	57.0144	−0.0033
y_{12}^{3+}	466.9790	Ile (Leu)	37.6947	37.6913	−0.0034

Peptide sequence: sample 5 Ac-EIEALKKEIEALKK-NH_2
sample 6 Ac-EIAELKKEIAELKK-NH_2

The resolution of the mass spectrometer was always set, using a PEG ion at *m/z* 701.4, to the value theoretically required to produce a 10% valley. With increased resolution, the transmission decreases such that a transmission of 10% (relative to 100% transmission at 1000 resolution) is typically observed at 9000 to 10000 resolution. The observed isotopic ions, and their relative intensities, for the peptides, were in general agreement with theoretically produced isotopic patterns. However, at higher resolution and, as is clearly the case for the $[M+8H]^{8+}$ (Fig. 3(b)) and $[M+10H]^{10+}$ ions (Fig. 4(b)), the observed valley approached 50%. The reduced signal, relative to the background, could be a contributing factor to the observation of a higher than expected valley between ions. Alternately, slight energy spreading during the electrospray process may be leading to some signal broadening. Finally, the fact that several weaker electrospray spectra were summed could lead to some unanticipated broadening. From a purely empirical standpoint, it appears that about 25% more resolution, than calculated by the software for a PEG ion, would be required to resolve isotopic clusters to the 10% valley definition.

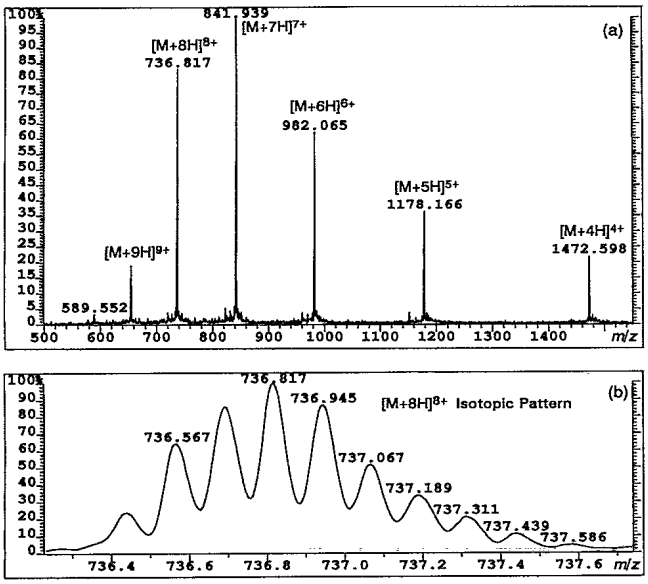

Figure 3. (a) Electrospray mass spectrum of peptide sample 29 at 6000 resolution. (b) Isotopic pattern of $[M+8H]^{8+}$ ions, illustrating resolution of ions differing by 1/8 Da. Observed monoisotopic molecular weight: 5883.44 ± 0.12; sequence: ATKKEVPLGVAADANKLGEIEALKAEIEALKAGGDEQFIPKGGEIEALKAEIEALKA-NH_2.

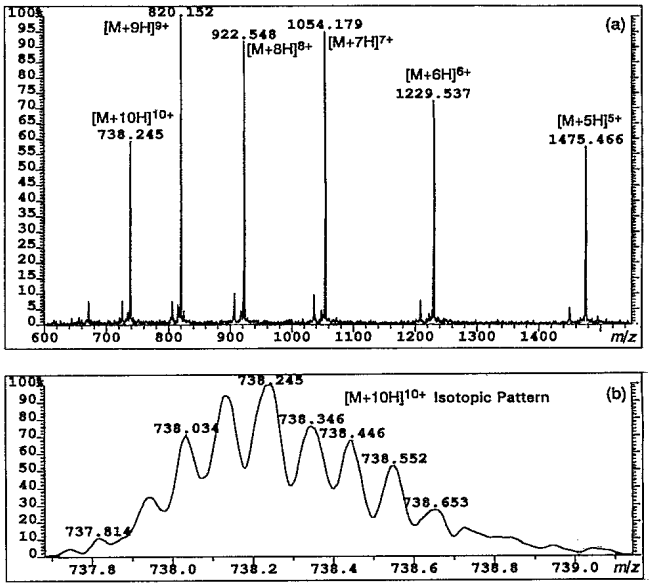

Figure 4. (a) Electrospray mass spectrum of peptide sample 32 at 8000 resolution. (b) Isotopic pattern of $[M+10H]^{10+}$ ions, illustrating resolution of ions differing by 1/10 Da. Observed monoisotopic molecular weight: 7368.36 ± 0.08; sequence: (Ac-QCGALQKQVGALQKQVGALQKQVGALQK-NH_2)$_2$, dimer with C-C bridge.

The monoisotopic molecular weights of two peptides, above 8000 Da, could not be confidently assigned and for this reason average molecular weight data under high resolution conditions were reported for both peptide samples 9 and 30. This was in large part due to the fact that with increasing molecular weight the relative intensity of the ion used for the monoisotopic molecular ion calculations (containing only ^{12}C) decreases until it is not easily resolved from the background. For example, the relative intensity of the ion

used for this purpose for peptide sample 32 (monoisotopic molecular weight, 7368.36 Da), at m/z 737.814 exhibited a relative intensity of about 10%. This compares to 100% relative intensity (for the ^{12}C-only ion) at lower molecular weights, (e.g., monoisotopic molecular weight, 1439.92 Da). This fact, along with reduced transmission with increasing resolution, made confident assignment of these ions for peptides above 8000 Da impossible under the experimental conditions employed. The average molecular weights were calculated following centroiding of the cluster, with errors between the theoretical and observed average molecular weights being within the 5–60 ppm window (Table 1).

CONCLUSIONS

Accurate molecular weights of unknown synthetic peptides, used in research studies involving synthetic vaccines, antibacterial peptides, or the *de novo* design of helical peptides and proteins, may be determined with a magnetic sector instrument. Electrospray data were acquired over a wide mass range during magnetic scanning and calibrated externally. Good agreement between observed and theoretical monoisotopic molecular weights was demonstrated with differences typically in the 5–60 ppm range. Isotopic clusters for charge states up to 10+ were resolved through the use of magnetic sector resolutions of up to 9000. Peptide sequencing was possible for unknown peptides below 2000 Da. One peptide was completely sequenced by CAD-MS and the value of high resolution was demonstrated in this example by the differentiation of Lys (K) from Gln (Q). Two other peptides, with identical monoisotopic masses, but different primary sequences, were partially sequenced and differentiated based on their CAD-MS data.

Acknowledgements

We thank Dr James Fethiere, Dr Les Kondejewski, Dr Wah Wong, Dr David Wade, Dr Nian Zhou and Wayne Kohn of the Protein Engineering Network of Centres of Excellence, whose research projects stimulated the synthesis of the peptides described in this manuscript. In addition, we thank Iain Wilson, Len Daniels and Cindy Gannon for technical assistance in the synthesis, HPLC purification and characterization of the peptides described. We thank Terri Keown for performing the amino acid analyses. The research of Dr Robert S. Hodges was supported by the Government of Canada through the Protein Engineering Network Centres of Excellence (PENCE).

REFERENCES

1. M. Barber, R. S. Bordoli, R. D. Sedgwick and A.N. Tyler, *J. Chem. Soc. Chem. Commun.* 325 (1981).
2. C. K. Meng, M. Mann and J. B. Fenn, *Z. Phys. D* **10**, 361 (1988).
3. J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong and C.M. Whitehouse, *Science* **246**, 64 (1989).
4. K. Biemann, *Ann. Rev. Biochem.* **61**, 977 (1992).
5. A. L. Burlingame, R. K. Boyd and S. J. Gaskell, *Anal. Chem.* **66**, 634R (1994).
6. M. V. Buchanan and R. L. Hettich, *Anal. Chem.* **65**, 245A (1993).
7. C. K. Meng, C. N. McEwen and B. S. Larsen, *Rapid Commun. Mass Spectrom.* **4**, 147 (1990).
8. C. K. Meng, C. N. McEwen and B. S. Larsen, *Rapid Commun. Mass Spectrom.* **4**, 151 (1990).
9. B. S. Larsen and C. N. McEwen, *J. Am. Soc. Mass Spectrom.* **2**, 205 (1991).
10. J. R. Chapman, R. T. Gallagher, E. C. Barton, J. M. Curtis and P. J. Derrick, *Org. Mass Spectrom.* **27**, 195 (1992).
11. C. N. McEwen and B. S. Larsen, *Rapid Commun. Mass Spectrom.* **6**, 173 (1992).
12. R. B. Cody, J. Tamura and B. D. Musselman, *Anal. Chem.* **64**, 1561 (1992).
13. J. A. Loo, R. R. Ogorzalek Loo and P. C. Andrews, *Org. Mass Spectrom.* **28**, 1640 (1993).
14. A. M. Starrett and G. C. DiDonato, *Rapid Commun. Mass Spectrom.* **7**, 12 (1993).
15. T. R. Covey, R. F. Bonner, B. I. Shushan and J. Henion, *Rapid Commun. Mass Spectrom.* **2**, 249 (1988).
16. W. Y. Wong, R. T. Irvin, W. Paranchych and R. S. Hodges, *Protein Science* **1**, 1308 (1992).
17. T. J. Sereda, C. T. Mant, F. D. Sonnichsen and R. S. Hodges, *J. Chromatogr. A* **676**, 139 (1994).
18. N. E. Zhou, C. M. Kay and R. S. Hodges, *J. Molecular Biology* **237**, 500 (1994).
19. D. S. Wishart, L. H. Kondejewski, P. D. Semchuk, C. M. Kay, R. S. Hodges and B. D. Sykes, in: *Techniques in Protein Chemistry V*, J. W. Crabb (Ed.), Academic Press Inc., San Diego (in press).
20. W. D. Kohn, C. M. Kay and R. S. Hodges, *Protein Science* (in press).

#152705

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 ▶	12 Jul 91	
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