Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
International Journal of Mass Spectrometry 252 (2006) 204–212
International Journal of Mass Spectrometry
www.elsevier.com/locate/ijms

Determination of the location of positive charges in gas-phase polypeptide polycations by tandem mass spectrometry

Frank Kjeldsen *, Mikhail M. Savitski, Christopher M. Adams, Roman A. Zubarev

BioMedical Center, Laboratory for Biological and Medical Mass Spectrometry, Box 583, Uppsala University, S-75123 Uppsala, Sweden

Received 18 August 2005; received in revised form 18 October 2005; accepted 23 October 2005
Available online 19 April 2006

Abstract

Location of protonated sites in electrospray-ionized gas-phase peptides and proteins was performed with tandem mass spectrometry using ion activation by both electron capture dissociation (ECD) and collisional activation dissociation (CAD). Charge-carrying sites were assigned based on the increment in the charge state of fragment ions compared to that of the previous fragment in the same series. The property of ECD to neutralize preferentially the least basic site was confirmed by the analysis of three thousand ECD mass spectra of doubly charged tryptic peptides. Multiply charged cations of bradykinin, neurotensin and melittin were studied in detail. For n+ precursors, ECD revealed the positions of $(n-1)$ most basic sites, while CAD could in principle locate all n charges. However, ECD introduced minimal proton mobilization and produced more conclusive data than CAD, for which N- and C-terminal data often disagreed. Consistent with the dominance of one charge conformer and its preservation in ECD, the average charge states of complementary fragments of n+ ions almost always added up to $(n-1)$+, while the similar figure in CAD often deviated from n+, indicating extensive charge isomerization under collisional excitation. For bradykinin and neurotensin, the charge assignments were largely in agreement with the intrinsic gas-phase basicity of the respective amino acid residues. For melittin ions in higher charge states, ECD revealed the charging at both intrinsically basic as well as at less basic residues, which was attributed to charge sharing with other groups due to the presence of secondary and higher order structures in this larger polypeptide.

Keywords: Electron capture dissociation; Proton localization; Polypeptide

1. Introduction

In mass spectrometry (MS), the emergence of the soft ionization techniques electrospray ionization (ESI) [1] and matrix assisted laser desorption ionization (MALDI) [2,3] have greatly assisted the study of thermodynamic and physical properties of polypeptide ions in the gas phase. The solvent-free environment stands in sharp contrast to the far more complex conditions found in solution. Understanding the anhydrous behaviour of peptide and protein ions may provide fundamental insight into the role of water in protein folding and the biological activity of polypeptides. Solving protein gas-phase structures will expand our understanding of laws governing formation of secondary and tertiary structures of polypeptides. Gas-phase analysis can be performed by MS with at least 10^3 times greater sensitivity compared to conventional structure-determination techniques such as NMR and X-ray crystallography.

There are several MS-based strategies of investigating gas-phase polypeptide structures. One efficient approach is to use hydrogen-deuterium exchange (H/D-X) in solution [4] or in the gas phase [5]. H/D-X allows for determination of surface exposed amide hydrogens, which indirectly provides information about structural features of polypeptide ions (folded/unfolded states). Another useful technique is ion mobility spectrometry (IMS) [6], where ion-neutral collisional cross-sections are derived from measurements of ion drift times in a pressure-elevated tube with an electrostatic field. The experimentally measured collisional cross-sections are then compared with theoretical cross-sections derived from molecular modeling of lowest-energy structures [7]. The result of molecular modeling is strongly influenced by the assumed positions of charges in the gas-phase polypeptide. As an example, dissimilar structures with distinctly different collisional cross-sections were calculated for a number of common peptides in

* Corresponding author. Tel.: +46 18 471 5729; fax: +46 18 471 7209.
E-mail address: Frank.Kjeldsen@bmms.uu.se (F. Kjeldsen).

1387-3806/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2005.10.009

the same charge state but with different charge configurations [8].

Generally speaking, the protons in polypeptide cations produced by electrospray ionization should preferentially occupy the most basic sites. The gas-phase basicity (GB) of individual sites is determined by a variety of factors. The intrinsic gas-phase basicity (GB_{int}) of amino acids is an important, but not necessarily decisive parameter, as other factors also play prominent roles. As an example, proton sharing with other chemical groups (charge solvation) increases the energy required to deprotonate the charged site, which is equivalent to an increase in the proton affinity of that site in the neutral state [9–11]. Thus, the availability of charge solvation groups is a factor influencing the protonation pattern. Simultaneously, Coulomb repulsion with already existing neighbouring proton reduces the proton affinity of a potential protonation site by the amount equivalent to the potential Coulombic energy. This reduction in proton affinity may make an intrinsically less basic but more distant residue more favourable for protonation. For example, Fenselau and co-workers concluded from experiments and molecular modeling that in doubly charged des-Arg^9-bradykinin (RPPGFSPF) the second protonation occurs at the C-terminus (Phe^8) and not at the more intrinsically basic N-terminus. This was explained by the shorter distance in the latter configuration to the first protonation site at Arg^1 [12]. Similarly, Cassady and co-workers [13] measured the GB_{app} of $[M+4H]^{4+}$ ions of the peptides K_4G_8, $(K_2G_4)_2$ and $(KGG)_4$, and determined K_4G_8 to have the lowest GB_{app} due to the higher potential Coulomb energy between the charged sites. Finally, interplay between charges of opposite sign, such as in gas-phase salt bridges, also influences the gas-phase basicity of various sites. These effects should be especially strong in larger polypeptide ions that provide more opportunity for charge solvation and salt bridging. Thus, for these ions, intrinsic basicity can be a poor guide for predicting charge location, except for special cases, such as the protonation of two arginine residues in the bradykinin analogue RRPPGFSPF that has been confirmed in SID experiments [14].

The absence of a simple algorithm for prediction of charge-distribution presents a significant problem for gas-phase protein structure determination. For instance, the small protein ubiquitin (76 residues) has 10^{14} different charge configurations if 13 protons are distributed [15]. For solving this problem and finding the lowest free energy charge configuration, Williams and co-workers suggested the use of a "pseudo-random walk" algorithm. This method successfully predicted the maximum charge state and suggested the charge positions in many common polypeptide ions [15]. The predicted maximum charge state was most often equal to the number of basic residues including the N-terminus. However, the protons were not always assigned to residues with the highest GB_{int}, but other basic sites, such as Pro, Trp and Gln, were also assumed protonated when charging exceeded 60% of the maximum theoretical value. In order to achieve realistic computational times, this model made simplified assumptions on the polypeptide gas-phase structures and intra-molecular solvation patterns. For instance, higher charge states of proteins were all assumed to be completely elongated linear strings and the model did not account for the basicity of the electrospray solvent [15]. Recently, Kebarle and co-workers developed another, quantitative model for prediction of the basicities of the charged sites and the average charge states of folded proteins [16]. The model starts with the "native" protein structure determined by X-ray or NMR, assigns to every basic site the basicity equal to its GB_{int}, increasing its value by the amount depending upon the local sequence context to account to charge solvation. Then the model simulates the interaction of each protonated site with an external base to account for the charge transfer to solvent. Although this model was successful in accounting for the decrease of the average charge states of proteins sprayed from basic solutions, the authors admitted a number of shortcomings. One of them was that the model determined the basicities of each site only for the total charge state 1+ of the molecule due to the huge number of charge isomers of a large molecule. The authors saw the strong side of their model in assigning the basicity order to every particular charged site.

In contrast to these theoretical advances, experimental determination of the charge location has proved more difficult. In principle, this can be done in a rather straightforward manner by fragmenting the molecule and observing the charge state of N- and C-terminal fragments. However, one needs to avoid proton mobilization and rearrangement before backbone cleavage is induced. Unfortunately, traditional fragmentation techniques utilizing vibrational excitation (VE) seem to mobilize at least one proton [14,17], whose movement along the peptide backbone can greatly complicate charge site localization [18]. In contrast to that, electron capture dissociation (ECD) is believed to produce $N–C_\alpha$ bond fragmentation without inducing notable VE [19,20]. This should limit proton scrambling before the bond cleavage, and thus preserve the information on the original proton position.

Neutralization of a protonated site erases information of its location. McLafferty and co-workers could probe the positions of 12 protons in 13+ ubiquitin ions [21]. The charge locations suggested by the experimental data and the positions of intrinsically basic sites not always agreed, which has been attributed to $N–C_\alpha$ bond cleavage three to four residues away from the protonation site. No probability assignment could be made for neutralization of different protonation sites.

Recently, Williams and co-workers showed experimentally that, in mixed cationized polypeptides $[M+n\mathrm{Cat}^1+m\mathrm{Cat}^2]^{(n+m)+}$, the cation with the largest recombination energy is neutralized preferentially [22]. The recombination energy RE(n+) of proton neutralization in n+ precursor depends upon the proton affinity PA([$n-1$]+) of that site in the neutral state [23]:

$$\mathrm{RE}(n+) = 13.6\,\mathrm{eV} - \mathrm{PA}([n-1]+) + \mathrm{HA}([n-1]+) \qquad (1)$$

where HA([$n-1$]+) is the hydrogen atom affinity of the neutral site. According to (1), the highest recombination energy in multiply protonated polypeptides is released at the site with the lowest proton affinity, which is the least basic site. This hypothesis has been put forward in reference [24] and applied to charge localization in the 20-residue Trp-cage protein. In this work we confirmed this hypothesis with a large set of data

(ca. 3000 molecules) acquired in a LC/MS/MS proteomics experiment.

Because of the preferential neutralization of one charged site, the more basic $n-1$ sites remain largely unaffected in ECD of a n+ polypeptide cation until the nth charge is neutralized. Then the charges in the $(n-1)^{\bullet}$+ reduced molecular species may rearrange under the influence of changed basicity due to disappearance of the coulombic interaction with the neutralized nth charge. Such a rearrangement will take time, while in most ECD models N–C_α bond cleavage occurs shortly after (in some models, simultaneously with) the electron landing. Thus, it is reasonable to assume that the charge rearrangement will happen in c, z fragments after their separation. Therefore, the charge state of ECD fragments should reflect the charge location in the precursor ion.

In ECD, the position of the neutralized nth charge remains obscure. Consequent ECD of $(n+1)$+ ions should locate n+ charged sites, thus identifying the nth (least basic) charge. By comparing ECD of consecutive charge states of the same molecule, one could determine both the proton locations as well as the gas-phase basicity order of various sites in polypeptide ions.

The above view [24] was a starting point for this work. Here we investigated the charge locations in a range of common polyprotonated polypeptide ions. For comparison, both ECD and CAD data were collected for the same charge states. The work revealed surprising patterns of protonation and charge solvation, as well as the presence of charge reorganization in higher charge states, complicating the above simple picture.

2. Experimental

2.1. Materials and sample preparation

All polypeptides were synthesized in-house by solid-phase peptide synthesis using N-α-Fmoc protocol [25]. A research micro-scale ResPep peptide synthesizer (Intavis AG, Gladbach, Germany) was used to synthesis five micromole of each polypeptide. Capping (acetylation of the N-terminus) between each amino acid coupling reduced the production of undesired truncated polypeptides. The produced peptides were then purified by RP-HPLC (HP1100, Agilent, CA) using a C18 column. For MS analysis polypeptides were diluted to a final concentration of ca. 10^{-6} M in water/methanol/acetic acid (49:49:2 v/v/v) solutions.

2.2. Mass spectrometry

A 7 T ESI LTQ-FT mass spectrometer (Thermo Electron, Bremen, Germany) equipped with an indirectly heated dispenser cathode as an electron source was used for ECD experiments. All spectra were obtained in broadband detection mode with a resolution of 100,000 at mass 400 Da. External calibration provided better than 2 ppm mass accuracy for precursor as well as product ions. Nanoflow (10–100 nL/min) ESI infusion was performed with metal-coated pulled-glass needles (Proxeon, Odense, Denmark) using an electric field potential of 700 V between the spraying needle and the mass spectrometer inlet. Molecular ions in a desired charge state were isolated in the linear ion trap and fragmented by collisions with He gas or transferred to the Penning (ICR) trap for ECD experiments. An electron current of 30 μA was used in all experiments; the irradiation time was 70 ms, the electron energy—below 1 eV. In each experiment, a total of ten acquisitions each containing ten individual spectra were averaged.

3. Results and discussion

3.1. Testing the preferential charge neutralization hypothesis

The assumption of preferential neutralization of one charge was tested against ECD data on 2927 doubly charged tryptic peptides that were reliably identified by Mascot database search (score ≥34). These were high-quality ECD spectra that contained sequence tags five residues long on average [26], while the average length of the peptides was close to nine residues. If low-energy electrons neutralized both charges with similar probabilities, one would expect a significant number of complementary fragments. However, 86% of all ECD data contained no complementary fragment ions, consistent with preferential neutralization of one charge. As a typical tryptic peptide contains one basic amino acid (Lys or Arg) at the C-terminus, it is the other charge that was expected to be neutralized, giving rise to preferentially C-terminal fragments. Indeed, three quarters of all fragment ions were $z^{\bullet}$ ions, despite the fact that, being radicals, they are less stable than even-electron, N-terminal c'-ions. This result can be compared with the typical 3:2 ratio between c' and $z^{\bullet}$ ions for non-tryptic polypeptides. Fig. 1a shows a typical ECD mass spectrum of a doubly charged tryptic peptide, dominated by a series of $z^{\bullet}$ ions.

In 60% of the cases concerning the ECD data that did contain complementary fragment ions, the ratio between the abundances of the two complementary fragments was 3:1 or larger. This decisive dominance of one of the fragments is also supportive of the preferential-neutralization hypothesis. Investigation of the remaining data revealed that in most cases, the sequences involved were either due to missed trypsin cleavages or contained His (see Fig. 1b as an example). In either case, there was more than one basic residue in the sequence. Thus, by a conservative estimate, ECD was selectively neutralizing one proton in 90% of doubly charged precursors, which strongly supports the tested preferential-neutralization hypothesis. Preferential charge neutralization in ECD opens a way for charge localization, as non-neutralized protons do not participate in bond cleavage and remain at their original location.

This also highlights a problem with the current understanding of the ECD fragmentation mechanism assuming cleavages in near proximity to the charge solvated site. According to that mechanism, preferential neutralization of just one charge should not lead to long series of N–C_α cleavages found in ECD. We are currently working on an alternative explanation of abundant backbone ECD cleavages consistent with the preferential-neutralization hypothesis.

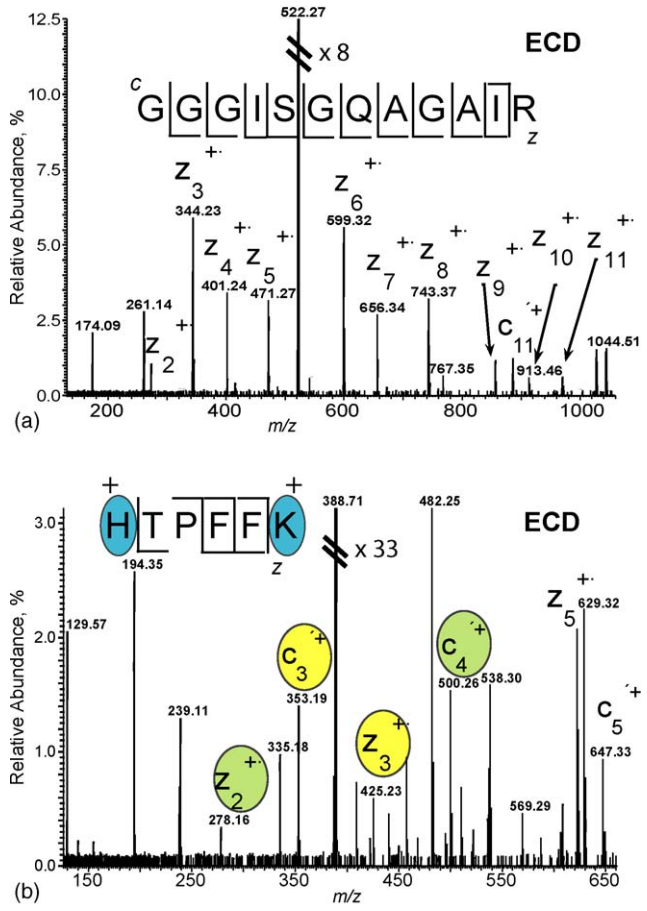


Fig. 1. ECD mass spectra of doubly charged tryptic peptides: (a) a typical spectrum, dominated by a series of z• ions and no complementary fragments; (b) a spectrum of a peptide with one more basic residue, His, showing complementary fragments c_3', $z_3^{\bullet}$ and c_4', $z_2^{\bullet}$.

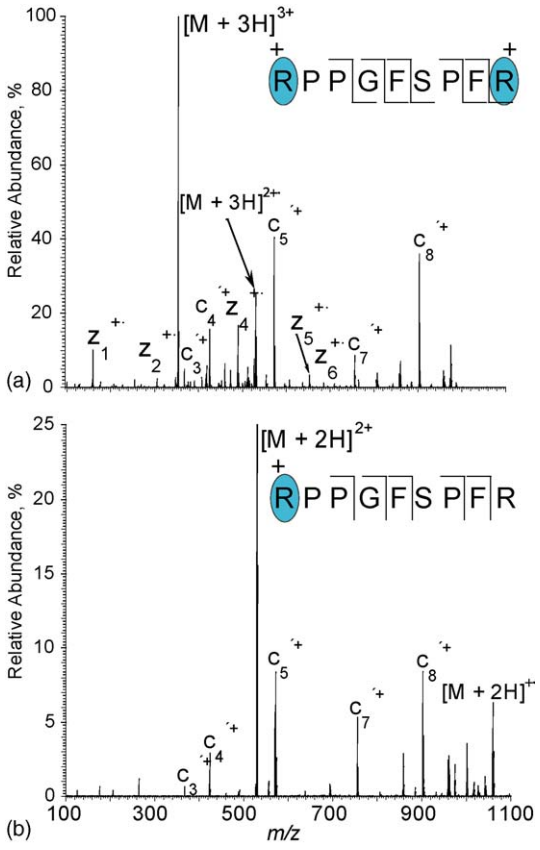
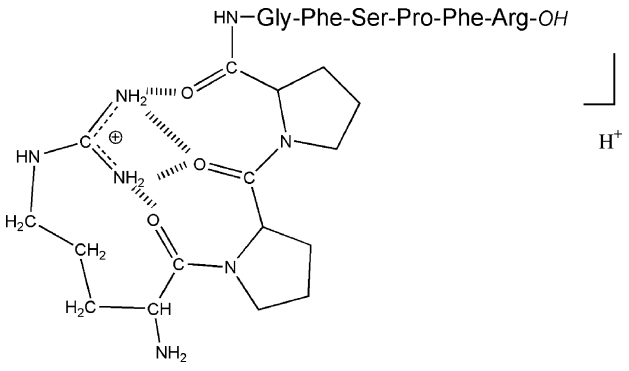


Fig. 2. ECD mass spectra of (a) $[M+3H]^{3+}$ and (b) $[M+2H]^{2+}$ bradykinin ions.

3.2. Bradykinin

Bradykinin (RPPGFSPFR) has arginine residues at both N- and C-termini. Its gas-phase structure has been subject of a large number of studies using ion mobility spectrometry [7,27], hydrogen-deuterium exchange [28,29], kinetic method [11,30], bracketing method [18,31], and combinations of molecular mechanics with conformational search [7,32]. Because of the wealth of structural information available for this peptide, bradykinin served a natural starting point in this study. Fig. 2 displays the ECD mass spectra of 3+ and 2+ ions, respectively. For triply charged bradykinin (Fig. 2a), only singly charged product ions were observed: five c'^{+} ($c_{3-5}'^{+}$ and $c_{7-8}'^{+}$) and five $z^{+\bullet}$ ($z_{1-2}^{+\bullet}$ and $z_{4-6}^{+\bullet}$) ions. ECD missed only N–C_α bonds N-terminal to proline residues (Pro^{2-3} and Pro^{7}), as was expected due to the immunity of tertiary amides to ECD [19]. Thus, the location of the N-terminal charge was restricted to the Arg^{1}-Pro^{2}-Pro^{3} residues. However, based on the superior intrinsic basicity, the charge was assigned to the arginine. The second charge was unambiguously assigned to the C-terminal arginine residue due to the presence of $z_1^{+\bullet}$ fragment ion. This charge configuration also satisfies the maximum charge separation and minimum Coulomb repulsion between the two assigned charges. The location of the third, neutralized charge, remained undetermined.

ECD of 2+ bradykinin (Fig. 1b) gave only a series of N-terminal c type fragment ions ($c_{3-5}'^{+}$ and $c_{7-8}'^{+}$), thus locating the most basic charge at the N-terminal arginine residue. Protonation of Arg^{1} is in agreement with findings for 3+ ions, as well as with earlier studies that pointed towards protonation in 2+ ions of both arginine residues [18,31,33,34]. The preference for Arg^{1} protonation is consistent with extensive charge stabilization by the backbone carbonyl oxygen of the second and third proline [33] (Scheme 1), which should be more favourable than competing charge solvation opportunities for the C-terminal arginine.



Scheme 1. Solvation of protonated Arg^{1} on backbone carbonyls in 2+ bradykinin ions. Adopted after Ref. [33].

In 3+ ions, such preference may disappear because of the protonation of the one of proline residues [31].

Literature suggests that, in 1+ of bradykinin that forms a salt bridge, both arginines are protonated [7,32,33,35]. Arg^9 should be more basic due to the proximity to the deprotonated C-terminus. In zwitterionic neutral bradykinin, Arg^1 should again be more basic than Arg^9 due to charge solvation on the proline carbonyls as in Scheme 1 and deprotonation of the C-terminus. This switching of the basicity order between the two arginines in n+ bradykinin as a function on n highlights an important issue: the basicity order in polypeptides can only be assigned correctly if the charge state of the molecule is specified. If this is true for bradykinin, it can also be true for proteins, with implications for theoretical models.

Because of the possible competition in ECD between proton rearrangement after electron capture and product ion separation (backbone bond cleavage is believed to be fast (non-ergodic), whereas product ion separation can be slow), there is a possibility that the fragment charge states actually reflect an intermediate charge distribution in $(n-1)$+ precursor ions rather than in n+ ions. We investigated the possibility of proton rearrangement before fragment separation for bradykinin 2+ precursors. If such a possibility was realized, the charge position in the fragments should reflect more that in 1+ bradykinin. In the latter species, the C-terminal arginine is more basic, and upon collision-induced fragmentation the charge is retained on C-terminal product ions as witnessed by the dominance of y′ type fragment ions (data not shown). However, only N-terminal fragment ions are observed in ECD of 2+ bradykinin, indicating that the proton rearrangement prior to fragment separation does not play an important role. Such a process however cannot be totally excluded for other polypeptide ions.

In comparison, CAD of 3+ bradykinin ions (Fig. 3a) gave product ions in multiple charge states (e.g., b_5 and b_6 fragments were found in charge states 1+ and 2+), making exact charge localization impossible. To highlight changes in the product ion charge states, shaded areas are overlaid on bars (Fig. 3). These areas start at charge state 0 and end with the charge state n+ of the precursor polypeptide ion. In principle, the sum of the charge states indicated by the shaded areas at a given cleavage should equal to n. Thus, a decrease in the charge state of one product ion should be followed by the same increase in the charge of the complementary product ion.

CAD of 2+ (Fig. 3b) also gave multiple charging of fragments, but to a lesser degree. The y′ ions gave lesser spread of charge states and thus more consistent information compared to b ions. By taking into account the most abundant charge states, the positions of two protons was possible to assign to Arg^1 and Arg^9 residues, in agreement with ECD findings.

3.3. Neurotensin

Neurotensin (pELYENKPRRPYIL) exhibits in electrospray ionization abundant charge states 2+ and 3+. Neurotensin contains three intrinsically basic residues, K^6, R^8 and R^9. The N-terminus is pyroglutamated and thus its basicity is reduced [36]. The shifts in the N-terminal fragment charge states (appear-

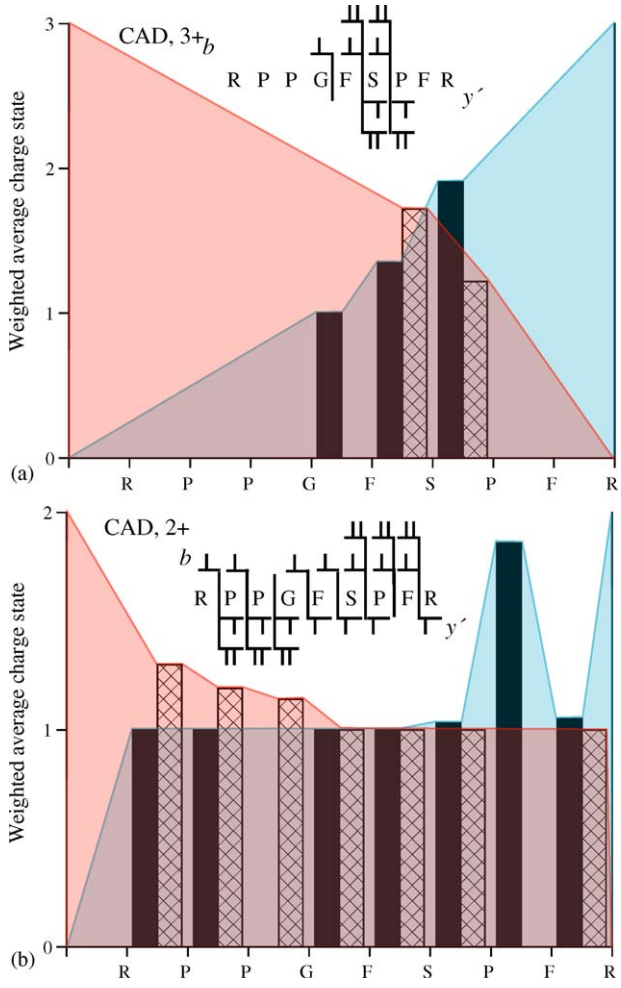


Fig. 3. Average charge states of CAD fragments of (a) $[M+3H]^{3+}$ and (b) $[M+2H]^{2+}$ bradykinin ions. (Light columns) C-terminal y′ fragments; (black columns) N-terminal b fragments. Regarding insets, the number of ticks on the bond cleavage sites represents the observed charge states of the product ions.

ance of the first c′ ion after P^7 and shift from 1+ to 2+ from $c_8{}'^+$ to $c_{10}{}'^{2+}$) in ECD of 3+ ion (Fig. 4a) located two charges at K^6P^7 and R^9P^{10}. C-terminal fragments told a similar story. Consistent with protonation at K^6 and R^9, the first z• type fragment ion was $z_5{}^{+\bullet}$, and the charge increase from 1+ to 2+ was observed between $z_6{}^{+\bullet}$ and $z_8{}^{2+\bullet}$.

ECD of 2+ ions (Fig. 4b) identified one protonated site, R^8. The shift in proton location from R^8 to R^9 in 3+ ions is likely due to the decreased basicity of R^8 because of the Coulomb repulsion with the protonated K^6. As in bradykinin, the basicity order in neurotensin is specific to the charge state. While in 2+ ions the most basic site is R^8, in 3+ its basicity is below that of both R^9 and K^6.

Note that the presence of only one charge state for each of the fragments in ECD of 2+ and 3+ ions indicates the absence of proton mobilization in ECD, as well as the presence of just one charge conformer. This does not necessarily mean that neurotensin ions have in each charge state a unique three-dimensional structure. The latter is highly unlikely, as such a structure would produce very limited fragmentation in ECD, similar to Trp-cage dications [24]. It is more likely that there is

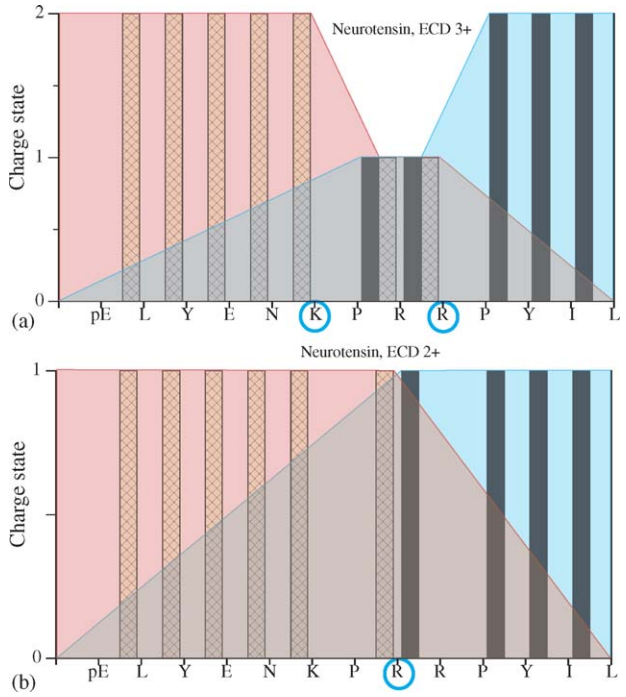


Fig. 4. Average charge states of ECD fragments of (a) $[M+3H]^{3+}$ and (b) $[M+2H]^{2+}$ neurotensin ions. (Light columns) C-terminal $z^{\bullet}$ fragments; (black columns) N-terminal c' fragments. Circles denote the assigned charge positions.

a multitude of structures, all corresponding to the same unique charge conformer.

CAD of 3+ and 2+ neurotensin ions (Fig. 5a and b) produced multiple charges for the same ions and conflicting data for proton

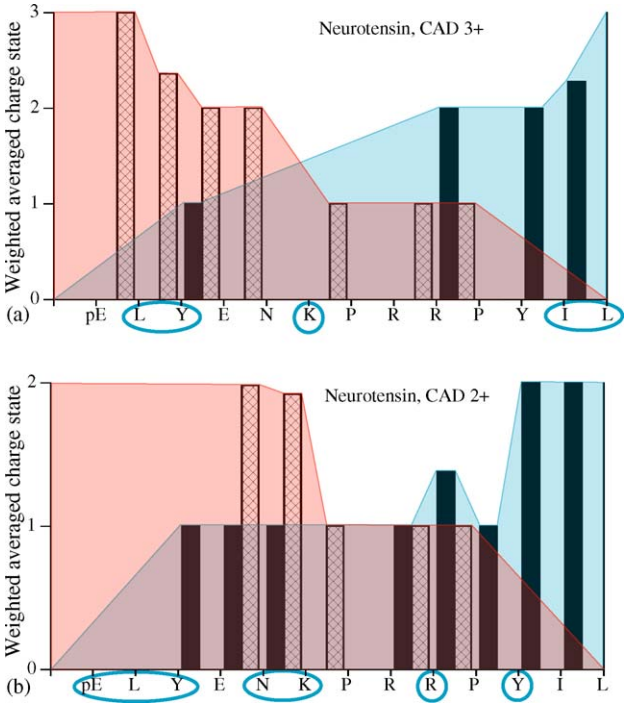


Fig. 5. Average charge states of CAD fragments of (a) $[M+3H]^{3+}$ and (b) $[M+2H]^{2+}$ neurotensin ions. (Light columns) C-terminal y' fragments; (black columns) N-terminal b fragments. Circles denote the assigned charge positions.

locations. According to the increments in fragment ion charge states, the most likely protonated sites in 3+ ions are L^2Y^3, K^6 and $I^{12}L^{13}$, which agrees with ECD only for K^6. CAD of 2+ ions showed even higher degree of proton scrambling. Charge increments found for b and y' fragments were inconsistent with each other and the two protonated sites could be assigned with similar probabilities to $pELY^{1–3}$, $NK^{5–6}$, R^9 and Y^{11}, of which one site (R^9) is the same as in ECD. These observations highlight pitfalls of using CAD for location of protonation sites [18].

3.4. Melittin

Melittin (GIGAVLKVLTTGLPALISWIKRKRQQ-NH_2) has a predominantly hydrophobic N-terminal sequence besides four basic residues near the C-terminus. The MS spectrum exhibits molecular ions in charge states 3+, 4+ and 5+. Although there are totally five intrinsically basic amino acids plus the N-terminal amino group, only a small peak of 6+ ions could be observed in the electrospray mass spectrum. Indeed, protonation of all residues in the $KRKR^{21–24}$ region would be thermodynamically unfavourable due to strong Coulomb repulsion between proximal charges. Therefore, protonation at other, less intrinsically basic residues was expected even for 5+ ions, as has been earlier suggested by Williams and co-workers [15]. ECD of 5+ ions (Fig. 6a) allowed to assign charges to the $IG^{2–3}$, $T(GLPALI)^{11–(17)}$, R^{22} and Q^{26} residues. The whole central region comprising seven amino acids was selected as a protonation site because of a small inconsistency between the charge states of the complementary c' and $z^{\bullet}$ ions originating from cleavages in that region: while c' ions show charge state shift after $LP^{12–13}$, the $z^{\bullet}$ ions shift their charge the most between T^{10} and T^{11}. This inconsistency may be due to the presence of several charge isomers of melittin 5+, with heterogeneity of protonation in the central region. The basicity order and the position of the most basic site may be different in different charge isomers, which would result in different fragmentation patterns and charge states of the fragment. Consistent with this suggestion, several of other c' and $z^{\bullet}$ fragments appeared in more than one charge state. However, the sum of the average charge states of complementary c' and $z^{\bullet}$ fragments is always ≤ 4, with the possible exception of $KR^{21–22}$, which means that charge isomerization is not very extensive.

In 4+ ions (Fig. 6b), the central region is not protonated and fragments appear preferentially in one charge state. The protonation could be assigned to $AV^{4–5}$, $IK^{20–21}$ and $RQ^{24–25}$. The shift in the charge locations compared to 5+, noticeably from Q^{26} to $RQ^{24–25}$ and from R^{22} to $IK^{20–21}$, was probably due to the reduced Coulomb repulsion. In ECD of 3+ ions (Fig. 6c), all fragments appear in a single charge state and the two apparent charges were consistent with protonation at $AV^{4–5}$ and $KR^{23–24}$.

The preferential protonation in the aliphatic region in presence of intrinsically much more basic Lys^7 and a number of polar residues was puzzling. It can only be explained in terms of the secondary and higher order structure of this polypeptide. Melittin has no defined structure in water solutions, and it acquires its biologically active conformation as it binds to

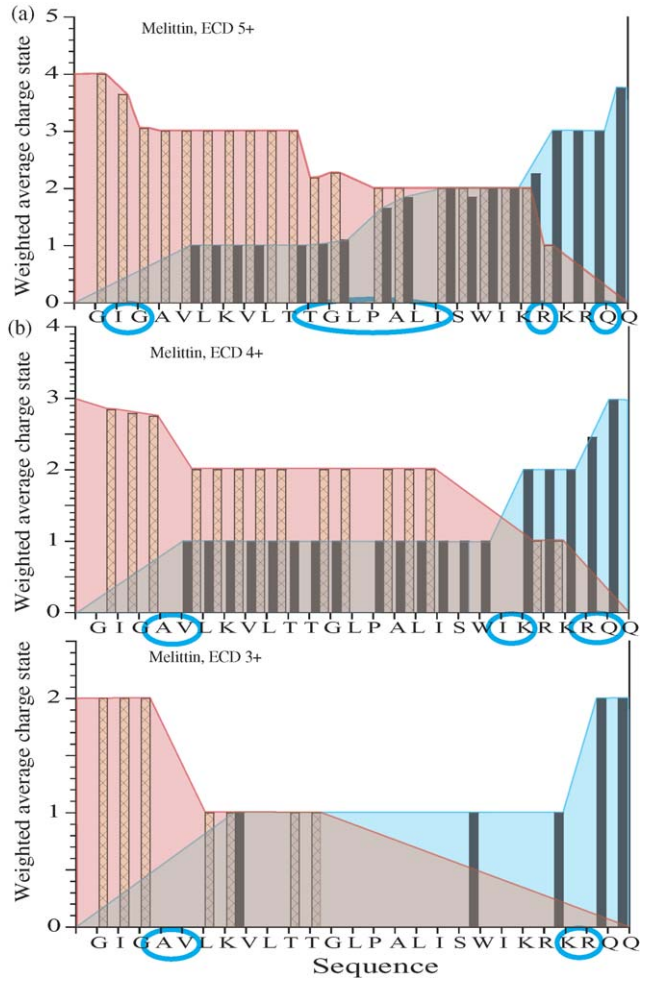


Fig. 6. Average charge states of ECD fragments of (a) $[M+5H]^{5+}$, (b) $[M+4H]^{4+}$ and (c) $[M+3H]^{3+}$ melittin ions. (Light columns) C-terminal $z^{\bullet}$ fragments; (black columns) N-terminal c′ fragments. Circles denote the assigned charge positions.

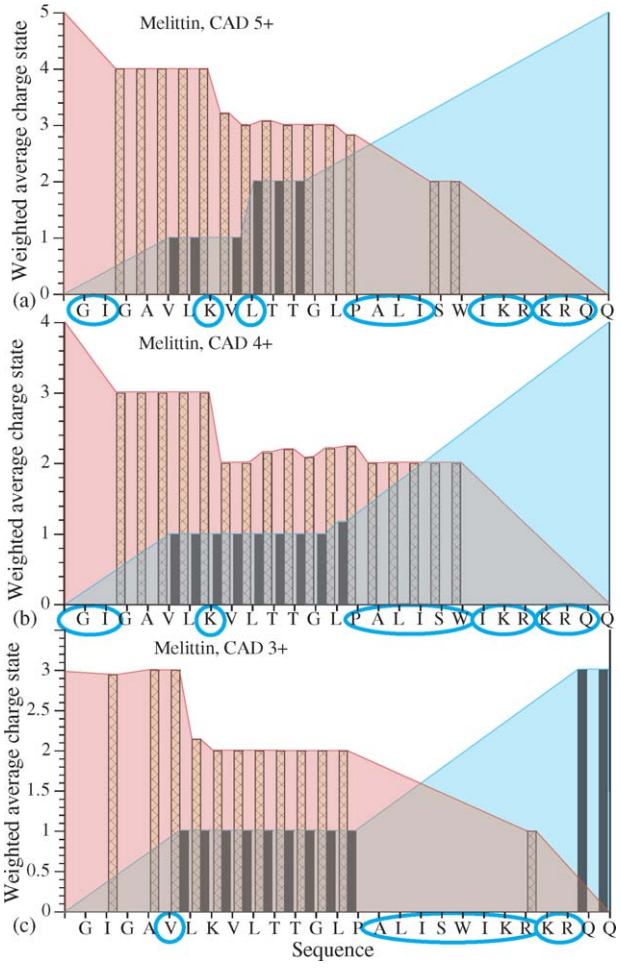


Fig. 7. Average charge states of CAD fragments of (a) $[M+5H]^{5+}$, (b) $[M+4H]^{4+}$ and (c) $[M+3H]^{3+}$ melittin ions. (Light columns) C-terminal y′ fragments; (black columns) N-terminal b fragments. Circles denote the assigned charge positions.

the cell membrane. Melittin structure in hydrophobic environment, including vacuum, may resemble the native conformation [11,30], in which the N-terminal forms an α-helix, probably imperfect. A local rupture of neutral hydrogen bonding in the α-helix would present an opportunity for charge solvation for basic backbone functionalities, first of all carbonyls. Multiple charge solvation, e.g., on backbone carbonyls can successfully compete with charge solvation on a single but intrinsically more basic group—note the preference for proline carbonyls in Scheme 1 over more basic N-terminus. If a charged side chain is solvated on some backbone carbonyl and after N–C_α bond cleavage at the same inter-residue link the proton remains with the N-terminal fragment (which may occur, e.g., due to simultaneous charge solvation on more than one carbonyl of that fragment, as in Scheme 1), then this proton will be attributed in our analysis to the residue of that carbonyl and not to the residue of the side chain. This may explain the formal assignment of the charge location to aliphatic residues.

CAD data were even more puzzling than ECD data (Fig. 7a–c), with product ions appearing in different charge states even for 3+ precursors. CAD of 5+ revealed six possible protonation sites: N-terminus with the first two residues, K^7, L^9 (according to the clear shift in b-ion charge state), $PALI^{14–17}$ and the double protonation in the $KRKR^{21–24}$ region. CAD of 4+ ions showed a similar pattern without the protonation at L^9. Note that the protonation at K^7 is supported by the increase in charge state from $y_{19}{}'^{2+}$ to $y_{20}{}'^{3+}$ ions, while the complementary $b_6{}^+$ and $b_7{}^+$ ions, both being singly charged, do not confirm this assignment. For both 5+ and 4+ precursor ions the localization of more charges than present were observed (six and five, respectively).

In 3+ ions, protonation at the N-terminus and K^7 was absent, but a new site at V^5 appeared. Note that the protonation pattern for 3+ ions exposed by CAD agrees well with that revealed by ECD of 3+ and 4+ precursors (Fig. 6b and c), including the absence of protonation at K^7. Such an agreement is consistent with the reduced scrambling of protons in CAD of lower charge states due to the larger difficulty in proton mobilization, and the absence of scrambling in ECD of even high charge states. Regarding the protonation at K^7, it is likely to occur when secondary and higher order structures are destroyed by vibrational excitation. Such destruction could also occur due to Coulomb

repulsion in highly charged molecules. Therefore, K^7 could be the fifth (most basic) protonation site in 5+ melittin. This arrangement would also correspond to maximum separation of the charges and minimum Coulomb energy.

4. Conclusions

The location of protonated sites in polypeptide cations up to 26 residues long was studied with ECD and CAD for comparison. Even with a large data set (ca. 3000 peptides), no inconsistency was found with the assumption that charge neutralization in ECD is favoured at the highest-recombination-energy site, which is the one with the lowest basicity. Thus, it was generally possible to locate with ECD $(n-1)$ charges for n+ precursor ions, although not always to a single residue. The obtained results for smaller molecules bradykinin and neurotensin were largely consistent with theoretical expectations and literature data, while melittin results were surprising as they provided evidence of protonation at unexpected residues, likely arising due to the secondary and higher order structures involving multiple charge sharing with backbone carbonyls.

In terms of the determination of the gas-phase basicity order of protonated sites, the situation was found to be more complicated than was originally thought. Additional protonation of an n+ polypeptide can result in one of the two outcomes. Either the n protons remain in the $(n+1)$+ molecule at the same positions, or they will rearrange to accommodate the growing coulombic repulsion between the charges. In the latter case, the basicity order in $(n+1)$+ ions will most likely change compared to n+ species. But even in the first case, the basicity order of protonated sites can change due to the creation/disruption of salt bridging and other weak bonding, as was found for bradykinin. Thus, the basicity order can only be established for a given charge state. That means that theoretical models cannot rely on intrinsic basicities of individual amino acids, nor on their apparent basicities determined for other charge states. As for the direct determination from ECD mass spectra, one can only establish which $(n-1)$ sites are the most basic in n+ ions, but not the basicity order among these sites.

The charge state distributions of backbone fragments revealed the presence or dominance of one charge conformer for all molecules studied, with a degree of heterogeneity for the highest charge state (5+) of the largest molecule studied, melittin. The heterogeneity was especially noticeable in the central region of the molecule. This region was the least basic, as ECD of lower charge states revealed. However, for all $N–C_\alpha$ cleavages, the sum of the average charge states of complementary fragments was very close to $(n-1)$+, which testifies to relatively low degree of charge isomerization. This is good news for the prospect of determination of three-dimensional gas-phase structures of proteins in low charge states, where extensive charge isomerization would greatly complicate the simulations and their comparison with the experiment.

CAD of the same peptide ions could potentially locate all n+ charges, but CAD data lacked self-consistency and showed evidence of proton rearrangement before fragmentation, especially for N-terminal b-ions and for higher charge states. The sum of the average charge states of complementary b and y′ fragments frequently deviated from n+, meaning significant charge isomerization during CAD excitation. This confirms the previous reports that localization of charged sites by CAD should be approached with caution. On the other hand, low-charge state CAD data were largely consistent with ECD results.

Acknowledgements

This work was supported by the Knut and Alice Wallenberg Foundation as well as the Swedish research council (grant nos. 621-2004-4897 and 621-2003-4877 to RZ).

References

[1] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Science 246 (1989) 64.
[2] M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, Int. J. Mass Spectrom. Ion Process. 78 (1987) 53.
[3] M. Karas, F. Hillenkamp, Anal. Chem. 60 (1988) 2299.
[4] V. Katta, B.T. Chait, Rapid Commun. Mass Spectrom. 5 (1991) 214.
[5] S. Campbell, M.T. Rodgers, E.M. Marzluff, J.L. Beauchamp, J. Am. Chem. Soc. 117 (1995) 12840.
[6] D.E. Clemmer, R.R. Hudgins, M.F. Jarrold, J. Am. Chem. Soc. 117 (1995) 10141.
[7] T. Wyttenbach, G. vonHelden, M.T. Bowers, J. Am. Chem. Soc. 118 (1996) 8355.
[8] H.H. Hill, C.H. Hill, G.R. Asbury, C. Wu, L.M. Matz, T. Ichiye, Int. J. Mass Spectrom. 219 (2002) 23.
[9] Z.C. Wu, C. Fenselau, J. Am. Chem. Soc 3 (1992) 863.
[10] J.Y. Wu, C.B. Lebrilla, J. Am. Chem. Soc. 115 (1993) 3270.
[11] I.A. Kaltashov, C.C. Fenselau, J. Am. Chem. Soc. 117 (1995) 9906.
[12] I.A. Kaltashov, D. Fabris, C.C. Fenselau, J. Phys. Chem. 99 (1995) 10046.
[13] X. Zhang, N.P. Ewing, C.J. Cassady, Int. J. Mass Spectrom. 175 (1998) 159.
[14] A.R. Dongre, J.L. Jones, A. Somoygi, V.H. Wysocki, J. Am. Chem. Soc. 118 (1996) 8365.
[15] P.D. Schnier, D.S. Gross, E.R. Williams, J. Am. Soc. Mass Spectrom. 6 (1995) 1086.
[16] M. Peschke, U.H. Verkerk, P. Kebarle, Eur. J. Mass Spectrom. 10 (2004) 993.
[17] X.-J. Tang, P. Thibault, R.K. Boyd, Anal. Chem. 65 (1993) 2824.
[18] G.A. Pallante, C.J. Cassady, Int. J. Mass Spectrom. 219 (2002) 115.
[19] R.A. Zubarev, N.L. Kelleher, F.W. McLafferty, J. Am. Chem. Soc. 120 (1998) 3265.
[20] K. Breuker, H.B. Oh, C. Lin, B.K. Carpenter, F.W. McLafferty, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 14011.
[21] K. Breuker, H.B. Oh, D.M. Horn, B.A. Cerda, F.W. McLafferty, J. Am. Chem. Soc. 124 (2002) 6407.
[22] A.T. Iavarone, K. Paech, E.R. Williams, Anal. Chem. 76 (2004) 2231.
[23] R.A. Zubarev, Mass Spectrom. Rev. 22 (2003) 57.
[24] C.M. Adams, F. Kjeldsen, R.A. Zubarev, B.A. Budnik, K.F. Haselmann, J. Am. Soc. Mass Spectrom. 15 (2004) 1087.
[25] L.A. Carpino, G.Y. Hann, J. Org. Chem. 37 (1972) 3404.
[26] M.M. Savitski, M.L. Nielsen, R.A. Zubarev, Mol. Cell. Proteomics 4 (8) (2005) 1180.
[27] A.E. Counterman, S.J. Valentine, C.A. Srebalus, S.C. Henderson, C.S. Hoaglund, D.E. Clemmer, J. Am. Chem. Soc. 9 (1998) 743.

[28] M.A. Freitas, C.L. Hendrickson, M.R. Emmett, A.G. Marshall, J. Am. Soc. Mass Spectrom. 9 (1998) 1012.
[29] M.A. Freitas, R.A.J. O'Hair, Int. J. Mass Spectrom. 175 (1998) 107.
[30] I.A. Kaltashov, D. Fabris, C.C. Fenselau, J. Phys. Chem. 99 (1995) 10046.
[31] N.P. Ewing, G.A. Pallante, X. Zhang, C.J. Cassady, J. Mass Spectrom. 36 (2001) 875–881.
[32] E.F. Strittmatter, E.R. Williams, J. Phys. Chem. A 104 (2000) 6069.
[33] P.D. Schnier, W.D. Price, R.A. Jockusch, E.R. Williams, J. Am. Chem. Soc. 118 (1996) 7178.
[34] T.G. Schaaff, J.L. Stephenson, S.A. McLuckey, J. Am. Chem. Soc. 121 (1999) 8907.
[35] J. Gidden, E.S. Baker, A. Ferzoco, M.T. Bowers, Int. J. Mass Spectrom. 240 (2005) 183.
[36] J.L. Stephenson, S.A. McLuckey, Anal. Chem. 69 (1997) 281.