RESEARCH ARTICLE

Structural Heterogeneity of Doubly-Charged Peptide b-Ions

Xiaojuan Li,[1] Yiqun Huang,[1] Peter B. O'Connor,[1,2] Cheng Lin[1]

[1]Mass Spectrometry Resource, Department of Biochemistry, School of Medicine, Boston University, Boston, MA, USA
[2]Department of Chemistry, the University of Warwick, Coventry, UK

Abstract

Performing collisionally activated dissociation (CAD) and electron capture dissociation (ECD) in tandem has shown great promise in providing comprehensive sequence information that was otherwise unobtainable by using either fragmentation method alone or in duet. However, the general applicability of this MS^3 approach in peptide sequencing may be undermined by the formation of non-direct sequence ions, as sometimes observed under CAD, particularly when multiple stages of CAD are involved. In this study, varied-sized doubly-charged b-ions from three tachykinin peptides were investigated by ECD. Sequence scrambling was observed in ECD of all b-ions from neurokinin A (HKTDSFVGLM-NH$_2$), suggesting the presence of N- and C-termini linked macro-cyclic conformers. On the contrary, none of the b-ions from eledoisin (pEPSKDAFIGLM-NH$_2$) produced non-direct sequence ions under ECD, as it does not contain a free N-terminal amino group. ECD of several b-ions from Substance P (RPKPQQFFGLM-NH$_2$) showed series of c$_m$-Lys fragment ions which suggested that the macro-cyclic structure may also be formed by connecting the C-terminal carbonyl group and the ε-amino group of the lysine side chain. Theoretical investigation of selected Substance P b-ions revealed several low energy conformers, including both linear oxazolones and macro-ring structures, in corroboration with the experimental observation. This study showed that a b-ion may exist as a mixture of several forms, with their propensities influenced by its N-terminus, length, and certain side-chain groups. Further, the presence of several macro-cyclic structures may result in erroneous sequence assignment when the combined CAD and ECD methods are used in peptide sequencing.

Key words: Electron capture dissociation, Peptide sequencing, Sequence scrambling, Peptide b-ion structure, Tachykinin neuropeptides, Tandem mass spectrometry

Introduction

Tandem mass spectrometry (MS/MS) is a powerful technique in peptide sequencing and protein identification, with collisionally activated dissociation (CAD) [1, 2] being the most routinely used method. In CAD, energy transfer through collisions with the inert gas molecules can lead to internal excitation and eventual dissociation of the precursor ions, usually via backbone amide bond cleavages, producing b- and y-type ions. The mass difference between fragment ions of the same type resulting from cleavages of adjacent amide bonds can be used to deduce the identity of the amino acid residue in between. When such a series of fragment ion masses are present in the tandem MS spectrum, the peptide sequence may be reconstructed. CAD has been extended to large biological

ions [3, 4], and in-source CAD of intact protein ions up to 200 kDa in mass has been reported [5]. Because of CAD's tendency to preferentially break the weakest bond in peptides [6–8], high sequence coverage is often difficult to achieve, particularly in the presence of labile amide bonds, such as the Asp/Glu-Pro sequence [9, 10]. Multiple stages of CAD may be performed to improve the sequence coverage. However, in such MS^n experiments, sequence rearrangements of b-ions may occur as evident by the observation of non-direct sequence fragment ions with loss of internal amino acid residues, which can lead to erroneous peptide sequence assignments [11, 12]. On the other hand, the recently developed electron capture dissociation (ECD) method tends to cleave, nonselectively, many more inter-residue $N–C_a$ bonds along the peptide backbone, and has proved to be an extremely valuable tool in post-translational modification (PTM) characterization because of its ability to preserve labile modifications while breaking backbone bonds. In general, though, as the size of the protein increases, ECD efficiency decreases, often due to the existence of extensive noncovalent interactions which both restrict the conformational space and prevent fragment ion separation and detection. The ECD efficiency and sequence coverage can be improved by precursor ion activation, via either infrared irradiation or collisions with inert gases, as employed in activated ion (AI)-ECD [13]. It is often best to perform CAD and ECD in duet, where the complementary information provided by each method individually can be combined to allow higher confidence in sequence assignment and protein identification, either *de novo* or through database searching [14]. Conceivably, sequence coverage may also be increased by performing CAD and ECD in tandem, where a large protein ion is first fragmented by CAD, and a fragment ion that is both smaller in size and somewhat activated is then subjected to ECD [15]. Because non-direct sequence ions have been observed in b-ion CAD experiments, it is important to investigate whether they are also produced when ECD is used in the second tandem MS stage. Since the fragmentation behavior of an ion is influenced by its structure, a better understanding of the fragment ion structure is beneficial for spectral interpretation and protein sequencing using the MS^3 approach.

Of the major product ions produced by CAD, the y ions generally have the same structure as the protonated truncated peptides [16], but the b-ions may exist as several structural variants. The structure of peptide b-ions has been the subject of a long-standing debate. Originally, the structure of b-ion was proposed as an acylium [17], although the oxazolone is now generally regarded as the kinetically and energetically favored product based on its fragmentation behavior, a conclusion that is also supported by both the theoretical modeling [18] and the infrared (IR) spectroscopic study [19]. There has been growing evidence, recently, showing that a b-ion may also assume a macro-cyclic form, which can lead to extensive sequence scramblings upon CAD [11, 12, 20–27]. Most of the macro-cyclic structures are formed via the nucleophilic attack on the C-terminus by the N-terminal

amino group, while in some lysyl and ornithyl peptides, the ε-amino group on the basic amino acid side chain may also be involved [28, 29]. Upon further collisional activation, these cyclic b-ions can reopen at various positions to form a mixture of linear oxazolones giving rise to non-direct sequence ions. A recent review presented an extensive discussion on the structures of small b_{1-4} ions and medium-sized b_n ions generated in CAD [30].

Factors influencing the b-ion cyclization have been investigated, although no obvious correlation was established. Medium-sized b-ions are generally considered to have a greater potential to form macro-cyclic structure than large b-ions or small b_2 to b_4 ions, which follows stepwise degradation via "the oxazolone rule" [20, 25, 31, 32]. The effect of amino acid composition was also investigated. The aliphatic amino acids, such as Ile, Leu, Val, and Ala were found to be prone to undergo internal elimination [12]. Selective ring opening of protonated cyclic peptides [24] may be influenced by the Pro [33] and the Asn/Gln effects [34], as well as certain acidic, basic, and amide side chains [35]. The doubly charged b-ions have a higher tendency to form cyclic structures than their singly charged counterparts [12]. Further, activation methods and collisional cooling rates may also play important roles in CAD chemistry [32], and the experimental conditions, such as the collisional activation energy and activation time, can affect the formation of macro-cyclic structure as well [27].

Very few ECD studies on b-ions have been performed. In one ECD study, abundant CO loss was observed and used as evidence to support the acylium structure [36], although it was shown and later demonstrated that CO may also be lost from the oxazolone form [37]. In the current study, ECD was performed on various doubly charged medium-sized b_n ions (n=5–10) from several tachykinin peptides with multiple basic residues near their N-termini to investigate whether sequence scrambling presents a problem in the tandem CAD/ECD experiments, and to probe the b-ion structures based on their fragmentation behaviors. Theoretical calculations were also performed to identify the low energy conformers of selected b-ions.

Experimental

All experiments were performed on a solariX 12 T hybrid Qh-Fourier transform ion cyclotron resonance (FTICR) MS instrument (Bruker Daltonics, Billerica, MA, USA). Substance P (RPKPQQFFGLM-NH$_2$) and neurokinin A (HKTDSFVGLM-NH$_2$) were purchased from Sigma-Aldrich (St. Louis, MO, USA), and eledoisin (pEPSKDA-FIGLM-NH$_2$) was purchased from AnaSpec (Fremont, CA, USA), and used without further purification. Peptides were dissolved in standard electrospray solution containing 49.5:49.5:1 methanol:water:formic acid (vol/vol/vol) to a concentration of ~2 μM. Micro-electrospray was used to introduce the sample into the mass spectrometer with a flow rate of ~100 μL/h. In-source dissociation in the funnel-

skimmer region was used to fragment the precursor ions, and the collision energy was adjusted to maximize the abundance of the specific fragment ion of interest for further ECD analysis. All fragment ions were focused by and pre-stored in the source octopole, and the doubly charged b-ions of interest were isolated by the mass filtering quadrupole with an isolation window of $\sim 2 m/z$ unit, accumulated in the collision cell hexapole for up to 8 ms before being transferred to the ICR cell. Low energy electrons (~ 1 eV) were generated by a hollow cathode dispenser mounted on the rear side of the ICR cell, with an extraction lens located between the cathode and the cell to help guide the electrons into the ICR cell. The spectra were internally calibrated by the precursor ion and charge reduced species, analyzed using the DataAnalysis software (Bruker Daltonics, Billerica, MA, USA), and manually interpreted, with a typical mass accuracy around 0.5 ppm.

Low energy conformers of selected b-ions were investigated theoretically. Molecular dynamics (MD) simulations were performed to generate the energetically favored conformer candidates. The CHARMm force field was applied in MD simulations, using the NVT canonical ensemble technique. In order to explore the conformation space as comprehensively as possible, the system temperature was chosen to be at 1000 K. Each MD step was integrated to 1 fs, and each trajectory lasted 200 ps. For each trajectory, geometries were recorded every 5 ps, followed by temperature-independent energy minimization using the molecular mechanics methodology. The MD simulations were performed using the Discovery Studio 5.0 software (Accelrys, San Diego, CA, USA).

The energies of the favorable conformer candidates generated by MD simulations were further calculated using the quantum chemistry methodology. Due to the relatively large sizes of the peptide ions studied, geometry optimizations were performed using the restricted self-consistent field based ab initio- (RHF) approach with the small 3-21 G basis set. To minimize the energy inaccuracy resulting from the omission of the correlation energy, density functional theory (DFT)-based single point energy calculations were performed on optimized HF/3-21 G geometries using the hybrid of Becke's exchange and Lee-Yang-Parr's correlation functionals (B3LYP) with the 6-31 G(d) basis set. The B3LYP/6-31 G(d) electronic energies were used as zero Kelvin enthalpy without zero-point vibration energy correction. All calculations were carried out using Gaussian 03 program suite [38] at the supercomputing facility at Boston University.

Results and Discussion

Three tachykinin neuropeptides [39–41] (neurokinin A, eledoisin, and Substance P) were used as model peptides. The tachykinin peptides are neurotransmitters and can rapidly induce the contraction of gut tissue [42]. They have a common C-terminal sequence: Phe-Xxx-Gly-Leu-Met-

NH_2, which is considered as the message domain, with Phe important for receptor binding and Xxx important for receptor selectivity [43]. The ECD spectra of varied sized b-ions from neurokinin A, eledoisin, and Substance P are shown in Figures 1, 2, 3, and 4. The calculated low energy conformers of the b_8^{2+} and b_6^{2+} ions of Substance P are shown in Figures 5 and 6. The major peaks are labeled in the spectra, and the detailed peak lists can be found in the Supplemental Tables 1–11.

ECD of b-Ions from Neurokinin A and Eledoisin

Neurokinin A (HKTDSFVGLM-NH_2), also known as Substance K, is a homologue of Substance P (RPKPQQFFGLM-NH_2) [41]. ECD of the b_9^{2+} and b_8^{2+} ions from neurokinin A generated a near complete series of c-ions (Figure 1a and b), which can be produced from the linear oxazolone structure. However, these c-ions comprise only a small portion of the product ions. In the ECD spectrum of the b_9^{2+} (HKTDSFVGL^{2+}) ion of neurokinin A, the most intense peaks correspond to a series of non-direct sequence ions of the z-b-type with HK as their C-terminal sequence: LHK, GLHK, VGLHK, FVGLHK, SFVGLHK, and DSFVGLHK. A z-b-ion is a fragment ion that resembles a normal z•-ion at its N-terminus and a normal b-ion at its C-terminus. Other types of non-direct sequence ions are also present, labeled as the z-a- and z-c-ions in the spectrum. These ions do not necessarily terminate with the HK at their C-termini, and are in general of lower abundances. The ECD spectrum of the b_8^{2+} (HKTDSFVG^{2+}) ion of neurokinin A showed a similar cleavage pattern to the b_9^{2+} ion, with the major peaks corresponding to the normal c-ions and sequence scrambled GHK, VGHK, FVGHK, SFVGHK, and DSFVGHK (z-b-ion) fragments. The z-a- and z-c-ions are also observed at lower abundances, again not necessarily with an HK C-terminal sequence. ECD of the smaller b_7^{2+} (HKTDSFV^{2+}) ions from neurokinin A produced predominantly non-direct sequence ions, VHK, FVHK, and SFVHK (z-b-ion), with no c-ions, (Figure 1c).

The abundant z-b-ions observed may be explained by the formation of a macro-ring structure of the original b-ion, connecting its C-terminal carbonyl group to its N-terminal amino group, followed by ring reopening to re-form a new sequence rearranged linear oxazolone, preferentially with lysine as its C-terminus, (Scheme 1). The preference is likely driven by the favorable formation of an oxazolone with histidine as the amino acid residue next to the C-terminus, characteristic of an ergodic process [44]. ECD of the sequence rearranged linear b-ions can produce a series of z-b fragment ions as observed in Figure 1a and b. As the charge carrying histidine and lysine residues were located near the C-terminus of the newly formed linear oxazolone, no non-direct sequence c-ions were observed. Meanwhile, ring reopening by cleaving the N–C_α or C_α–C(=O) bond is energetically unfavorable, so that the z-a- and z-c-ions observed were likely formed via a different mechanism.
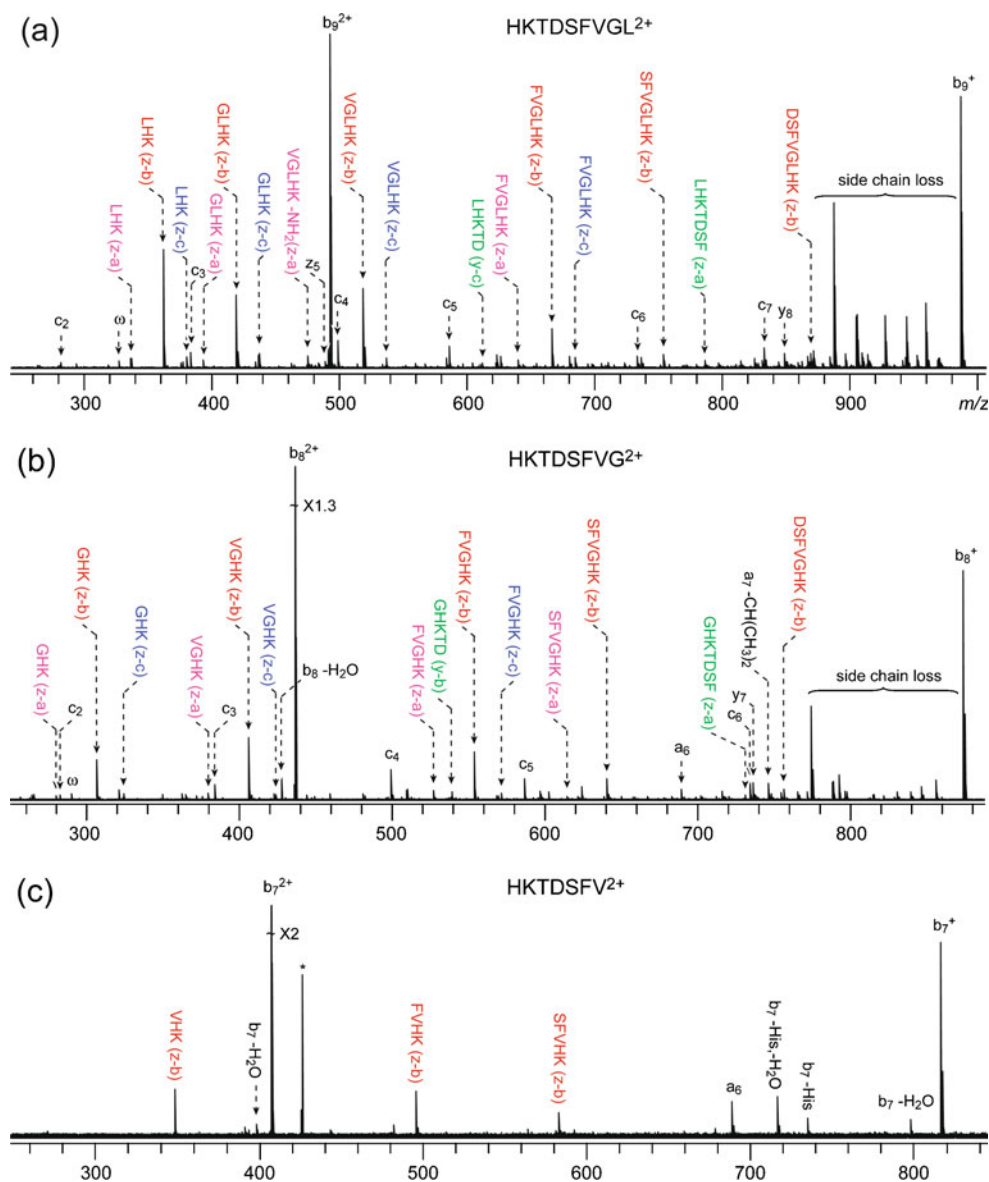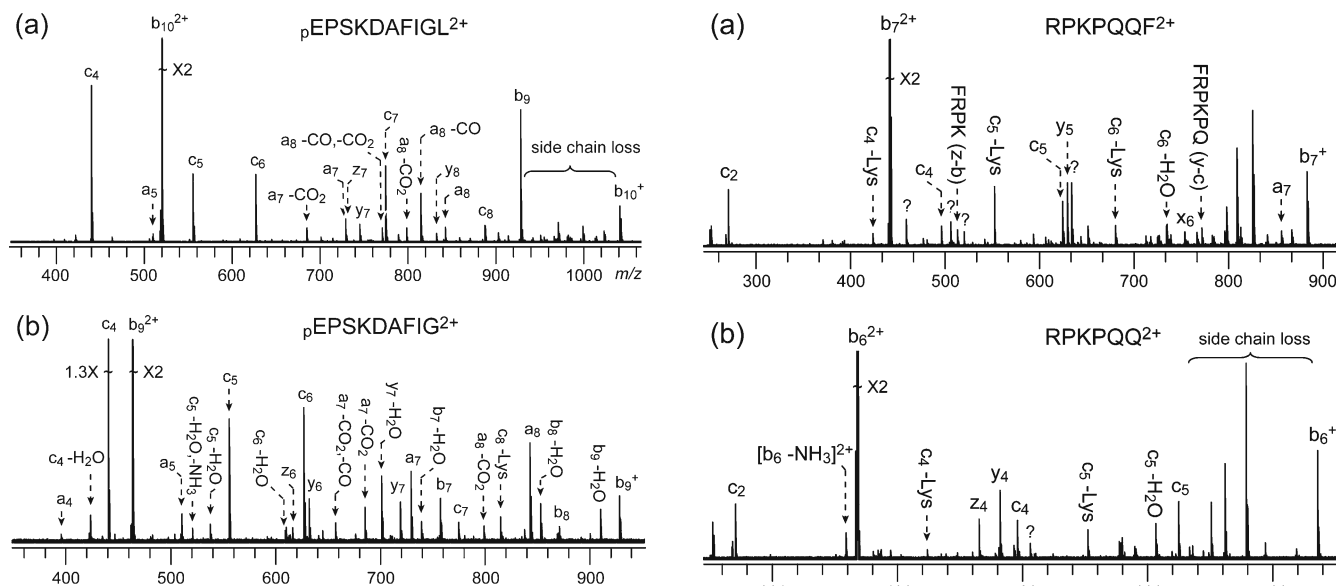
Figure 1. ECD spectra of the doubly charged b_9 (**a**), b_8 (**b**), and b_7 (**c**) ions from neurokinin A (HKTDSFVGLM-NH$_2$). The z-a, z-b, and z-c-ions are fragment ions with their N-termini resembling that of a z-ion, and C-termini resembling that of an a-, b-, and c-ion, respectively. Partial side chain loss and small molecule loss are labeled with the molecular composition of the leaving group(s). Peaks marked with "-amino acid residue" represent the entire side chain loss. *Marks noise peaks, and ω marks harmonic peaks

Note that the b-ions, if assuming a macro-ring structure, are essentially the same as the cyclic peptide ions. It has been reported that ECD of doubly charged cyclic peptides can generate numerous internal fragment ions, which were proposed to be formed via the free radical cascade (FRC) mechanism [45]. Likewise, the z• radical produced by the N-C$_\alpha$ bond cleavage within the macro-ring of the b-ions studied here can also initiate free radical cascade, leading to the formation of the z-a- and z-c-ions, via apparent loss of one or multiple internal amino acid residues (Scheme 1). The major difference between the formation of the z-b-ions and the z-a- and z-c-ions is that the former proceeds via ring-opening prior to ECD and the latter via ECD before ring-opening

followed by FRC. Thus, z-b-ions were selectively formed with HK as their C-terminal sequence, as expected from an ergodic ring-opening process; while z-a- and z-c-ions tend to contain more random sequences, as expected from the nonstatistical behavior of ECD. These results indicated that b-ions can exist as a mixture of linear and cyclic structures, and their interconversion in the gas phase occurred prior to ECD. Since sequence scrambling occurred before ECD, peptide sequencing using the MS^3 data could be erroneous, even when ECD was used in the second tandem MS stage.

Although eledoisin (pEPSKDAFIGLM-NH$_2$) has a similar C-terminal sequence as the neurokinin A, non-direct sequence ions were not observed in the ECD spectra of its b_{10}^{2+}-

Figure 2. ECD spectra of the doubly charged b_{10} (**a**) and b_9 (**b**) ions from eledoisin (pEPSKDAFIGLM-NH$_2$). Peak labeling follows the same convention as that in Figure 1
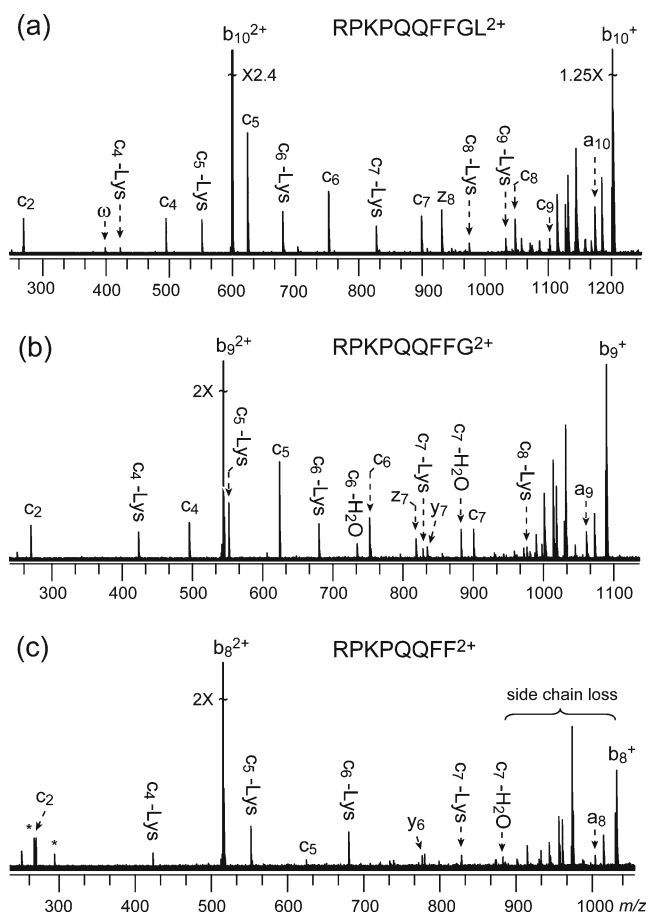
Figure 3. ECD spectra of the doubly charged b_{10} (**a**), b_9 (**b**), and b_8 (**c**) ions from Substance P (RPKPQQFFGLM-NH$_2$). Peak labeling follows the same convention as that in Figure 1
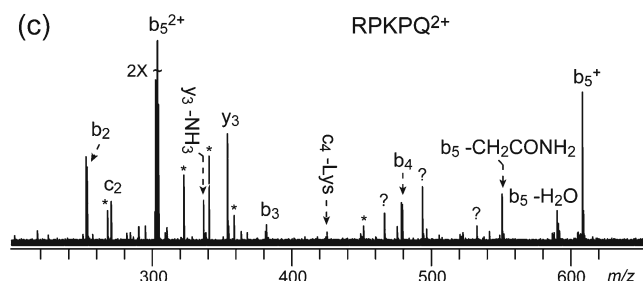
Figure 4. ECD spectra of the doubly charged b_7 (**a**), b_6 (**b**), and b_5 (**c**) ions from Substance P (RPKPQQFFGLM-NH$_2$). Peak labeling follows the same convention as that in Figure 1

and $b_9{}^{2+}$-ions, as shown in Figure 2a and b, respectively. This is perhaps not surprising in light of the presence of the N-terminal pyroglutamic acid residue, as its N-terminal amino group is blocked by the succinimide formation, thus preventing its conversion to the cyclic form via the N- and C-termini connection. This phenomenon is similar to the previous observation that the N- and C-termini cyclization of b-ions can be effectively blocked by acetylation of the N-terminal amino group [12, 21, 22, 26].

ECD of b-Ions from Substance P

In addition to the normal c-ions, ECD of doubly charged $b_n{}^{2+}$-ions ($n=10$–5) from Substance P (RPKPQQFFGLM-NH$_2$) produced a series of fragment ions that correspond to c-ions with additional lysine side-chain loss, labeled as c_m-Lys, m=4 to (n–1), (Figures 3 and 4). This is at first surprising, because c-ions do not usually contain a radical that is capable of initiating further side-chain losses as in the case of z• ions [46–49]. The formation of these unusual fragments may be explained by the presence of another type of cyclic b-ion structure, where the C-terminal carbonyl forms an amide bond with the lysine side-chain amino group. In ECD of
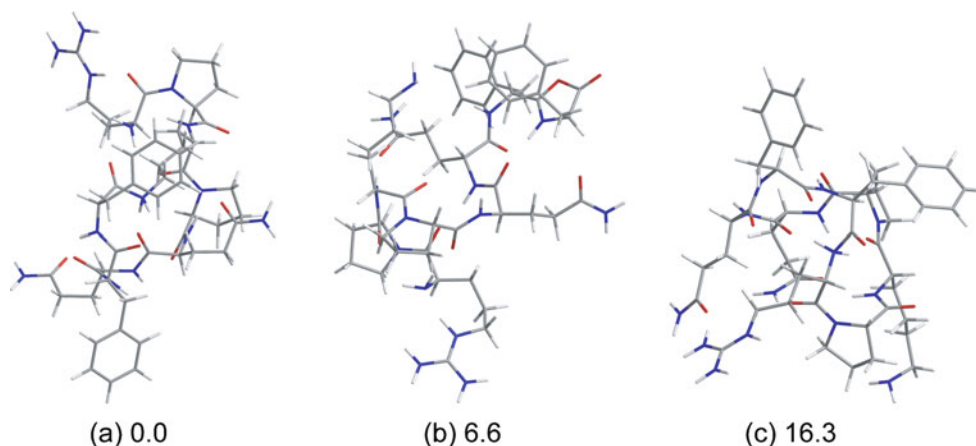
Figure 5. Lowest energy conformers of the Substance P b_8^{2+} ion and their relative energies obtained at the B3LYP/6-31 G(d)// RHF/3-21 G level of theory. (**a**) The FK-linked cyclic structure, (**b**) the oxazolone structure, and (**c**) the FR-linked cyclic structure. Numbers are in the unit of kcal/mol

this C-terminus-Lys linked cyclic b-ion, the initial $N–C_a$ bond cleavage will not result in the formation of individual c and $z\bullet$ fragments, as they are still held together by covalent bonds. The radical on the a-carbon of the $z\bullet$ moiety can migrate through space to induce the lysine side-chain loss and produce the c-Lys fragment (Scheme 2). Because the initial $N-C_a$ bond cleavage within the ring can occur at various positions, a series of c_m-Lys fragments can be generated. Similar cyclization involving the ε-NH_2 group of the Lys side chain to form a caprolactam structure has been reported previously [50].

The relative abundance of the c_m-Lys fragments to that of the c_m-ions depends on the size of the precursor b_n ions: going up from $n=10$ to 8 (Figure 3) and then down from $n=7$ to 5 (Figure 4). The c_4-Lys fragment ions in the ECD spectrum of the b_5^{2+}-ion was assigned with a big error (~2 ppm) compared to the ~0.5 ppm mass accuracy in all other data generated in this study. This may be due to the very low signal-to-noise ratio (S/N) of the monoisotopic peak of the c_4-Lys fragment in the b_5^{2+}-ion ECD spectrum, as it was difficult to isolate and accumulate the very low abundance b_5^{2+}-precursor ions generated by funnel-skimmer dissociation. The ECD spectrum of the b_8^{2+}-ion was dominated by the c_m-Lys

ions, with almost no c-ions. Thus, the majority of b_8^{2+}-ions were likely present in the C-terminus-Lys cyclized form, which is consistent with the previous study of oligoglycine b_8-ions by another group [32]. For most other b_n^{2+}-ions, both the c_m- and c_m-Lys-ions were abundantly present in the ECD spectra, indicating that these b-ions likely existed as a mixture of the two forms, both in appreciable quantities. Non-direct sequence ions were seldom observed in any Substance P b-ion ECD spectrum. Thus, the N- and C-termini connected macro-ring structure was either not formed, or formed but preferentially reopened at its original position.

In addition, the electron transfer dissociation (ETD) study [51] of the doubly charged b_{10}-ion of Substance P was performed on an ion-trap instrument as described previously [46]. Although a series of peaks corresponding to the c_m-Lys fragment ions were also observed, they were in general present in much lower intensities (data not shown). The lower abundance of the c_m-Lys fragment ions in ETD might be due to the lower amount of energy available comparing to ECD, reducing the likelihood of secondary fragmentations.

Cyclic structure formation involving the Lys side chain has been proposed previously. In a CAD study of the doubly
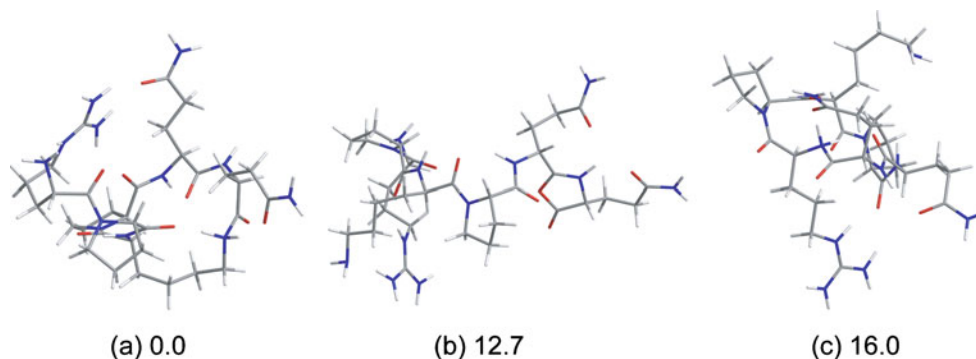
Figure 6. Lowest energy conformers of the Substance P b_6^{2+} ion and their relative energies obtained at the B3LYP/6-31 G(d)// RHF/3-21 G level of theory. (**a**) The QK-linked cyclic structure, (**b**) the oxazolone structure, and (**c**) the QR-linked cyclic structure. Numbers are in the unit of kcal/mol
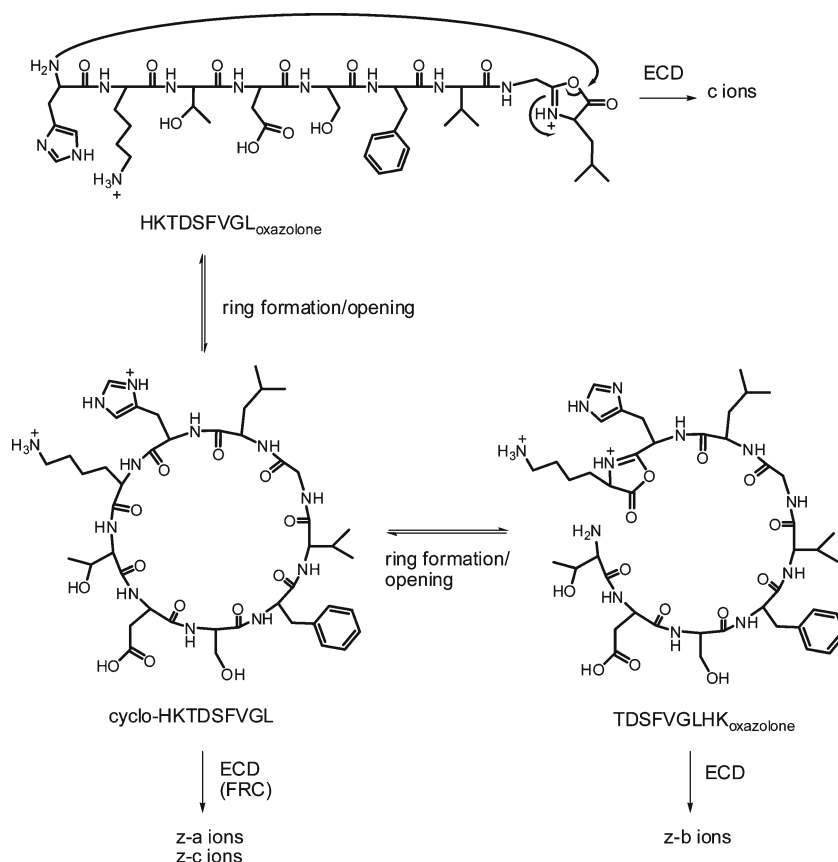
Scheme 1. Proposed ECD fragmentation pathways of the neurokinin A b_9^{2+} ion. FRC stands for the free radical cascade

charged b_{10}, b_9, and b_7 ions of Substance P, transfer of one (L) or two (GL) residues from the C-terminus of b-ions to lysine side chain were observed, suggesting the formation of a cyclic structure connecting the lysine side chain to the C-terminus [28, 29]. For other peptides, fragmentation after ring opening at one of the several ring positions was also observed. The presence of a proline residue close to the lysine or ornithine residue on the C-terminal side was considered as a feature for the ring structure formation.

To better understand the structure and ECD behavior of these doubly charged b-ions, geometric optimization and ab initio calculation of the relative energy of low energy conformers of selected b-ions were carried out at the B3LYP/6-31 G(d)//RHF/3-21 G level. Since no transition states were considered in these calculations, prediction of the fragmentation behavior can be difficult, as some of the calculated thermodynamically favorable structures may not be readily accessible kinetically.

The optimized structures and relative stabilities of the lowest energy linear (oxazolone), C-terminus-Lys linked and C-N termini linked cyclic structures of the b_8^{2+} and b_6^{2+} ions generated from Substance P are shown in Figures 5 and 6, respectively. For the b_8^{2+} (RPKPQQFF^{2+})-ion, the lowest energy oxazolone structure, which has the protonated guanidine and oxazolone groups solvated by the N-terminal amino nitrogen and backbone carbonyl oxygen, respectively, is 6.6 kcal/mol higher in energy compared to its FK-linked cyclic

counterpart that has both positive charge sites solvated by backbone carbonyl oxygen atoms. Consequently, c_m-Lys-ions are expected to be the dominant fragments in the ECD spectrum of the b_8^{2+}-ion, in agreement with the experimental observation.

For the b_6^{2+} (RPKPQQ^{2+})-ion, the QK-linked cyclic structure (Figure 6a) is 12.7 kcal/mol more preferred energetically over the oxazolone structure (Figure 6b). If energetics is the only factor taken into consideration, one would expect the predominance of the QK-linked cyclic structure over the linear oxazolone one, which would lead to the predominant formation of c_m-Lys fragment ions as well. The coexistence of both c_m- and c_m-Lys-ions with comparable abundance in the b_6^{2+}-ion ECD spectrum indicates that other factors may also play important roles in the ECD process. In particular, this low energy conformer of the QK-linked b_6^{2+}-ion has a compact and most likely inflexible structure, with the charge on the lysine side chain involved in hydrogen bonding to the carbonyl groups of the Lys_3 and Gln_5 residues. Electron capture at the protonated Lys_3 carbonyl will not lead to backbone cleavage due to the presence of its C-terminal proline residue, while electron capture at the protonated Gln_5 residue can produce the c_5-Lys side chain ion, which is the only abundantly formed c_m-Lys-ion observed in the ECD spectrum of the b_6^{2+}-ion. On the contrary, the linear oxazolone structure is much more flexible, which affords the conformational heterogeneity to produce various sized c_m-ions, as observed experimentally.
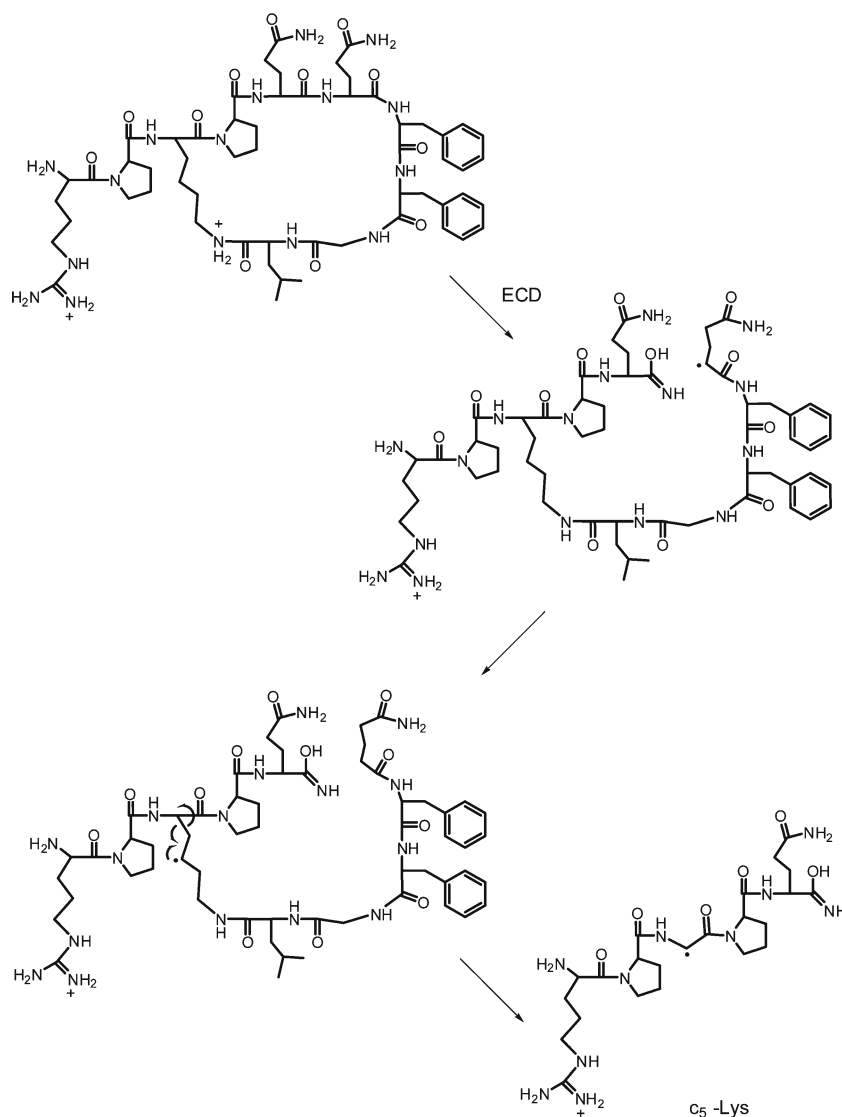
Scheme 2. Proposed mechanism for the formation of c_m-Lys side chain ions in ECD of the Substance P b_{10}^{2+} ion

As can be seen from Figures 5 and 6, the lowest energy N- and C-termini linked cyclic structures, FR- for the b_8^{2+}-ion (Figure 5c) and QR- for the b_6^{2+} (Figure 6c), are considerably higher in energy compared with the lysine linked cyclic structures. This is perhaps not all that surprising, as the localized charge on the N-terminal arginine is often stabilized by bridging to the N-terminus (Figure 5b), providing no free N-terminal amino group for macrocycle formation. Thus, the N- and C-termini linked cyclic structures are not expected to comprise an appreciable fraction of the total ion population and, consequently, no non-direct sequence ions were observed in the ECD spectra of the b_6^{2+}- and b_8^{2+}-ions.

With the sole exception of b-ions from eledoisin, which does not contain a free N-terminal amino group, macro-ring formation was observed in all b-ions studied here, either via the N- and C-termini linkage or via the Lys side chain and C-terminus connection. The identity of the C-terminal amino acid residues of b-ions, on the other hand, did not seem to play an important role in the cyclic structure formation of the

limited samples investigated. The macro-cyclic structure appeared to be abundantly formed in medium-sized b-ions, with their sizes ranging from b_6^{2+} to b_{10}^{2+}. The abundance of the macro-cyclic structure in smaller b-ions (e.g., the b_5^{2+}-ion from Substance P) is significantly lower, possibly due to the ring strain present in smaller macrocyclic b-ions containing only *trans*-amide bonds [18]. Meanwhile, the trend observed in ECD of Substance P b_n ions ($n=8–10$) indicated that the cyclic structure formation may also be disfavored as the size of the b-ions increases. This could be due to the entropic factor, i.e., it is more difficult for larger b-ions to find the low-energy conformation that brings the N-terminal or side-chain amino group close to the C-terminal oxazolone in the increased conformational space. ECD of even larger b_n ions ($n>10$) is currently under investigation.

Finally, it is important to note that while sequence scrambling seems to be a common feature among small b_n ions ($n≤10$), ECD of small b-ions are rarely needed in peptide sequencing and protein identifications, as the N-

terminal sequence of a protein is usually adequately covered in MS^2 experiments. In a recent study done here [15], it has been reported that the CAD/ECD approach can be a valuable tool to provide sequence information in regions that CAD or ECD alone or CAD based MS^n cannot access. In that study, no sequence scrambled fragment ions have been observed in ECD of large b_n ions ($n>30$) from medium-sized proteins, hemoglobin, and transthyretin (TTR). In addition to the entropic factor mentioned above, lack of cyclization in these very large b-ions may also be due to the extensive charge solvations that stabilize the oxazolone structure.

Conclusion

Both the ECD experimental and theoretical works clearly demonstrated that a peptide b-ion may exist as a mixture of several different forms in the gas phase, with their propensities influenced by its size, N-terminus, and possibly the side chains of basic amino acid residues. Blockage of the N-terminal amino group eliminated the formation the N- and C-termini linked cyclic structure. When a lysine residue is present in the sequence, cyclization via the C-terminus-Lys linkage may also occur. The possibility of cyclization is the highest in medium-sized b-ions, and becomes lower in very small b-ions due to the ring strain, as well as in large b-ions due to the decreased likelihood of finding the geometry necessary for ring formation. Non-direct sequence ions can be formed from the cyclized structure after ring reopening at different positions from the original linkage site, with ring-opening following the common preference observed in CAD experiments. When the C-terminus-Lys linkage was formed, unusual fragments corresponding to the lysine side-chain loss from the c-ions were observed. The presence of these non-direct sequence ions and unusual secondary fragment ions can complicate the spectral interpretation and may also lead to errors in peptide sequencing. The current results underscored a potential problem in protein identification and peptide sequencing based on the data generated from multiple stages of tandem mass spectrometry. However, such problem may not be as serious when using the CAD/ECD approach to sequence large proteins, where MS^3 is needed the most.

Acknowledgments

References

1. Hayes, R.N., Gross, M.L.: Collision-induced dissociation. *Method Enzymol.* **193**, 237–263 (1990)

2. Wells, J.M., McLuckey, S.A.: Collision-induced dissociation (CID) of peptides and proteins. In: Burlingame, A.L. (ed.) Biol Mass Spectrom, pp. 148–185. Elsevier Academic Press Inc, San Diego, CA (2005)

3. Biemann, K., Papayannopoulos, I.A.: Amino-acid sequencing of proteins. *Acc. Chem. Res.* **27**, 370–378 (1994)

4. Laskin, J., Futrell, J.H.: Activation of large ions in FT-ICR mass spectrometry. *Mass Spectrom. Rev.* **24**, 135–167 (2005)

5. Han, X.M., Jin, M., Breuker, K., McLafferty, F.W.: Extending top-down mass spectrometry to proteins with masses greater than 200 kilodaltons. *Science* **314**, 109–112 (2006)

6. Syrstad, E.A., Turecek, F.: Toward a general mechanism of electron capture dissociation. *J. Am. Soc. Mass Spectrom.* **16**, 208–224 (2005)

7. Zubarev, R.A., Haselmann, K.F., Budnik, B., Kjeldsen, F., Jensen, F.: Towards an understanding of the mechanism of electron-capture dissociation: a historical perspective and modern ideas. *Eur. J. Mass Spectrom.* **8**, 337–349 (2002)

8. Zubarev, R.A., Kelleher, N.L., McLafferty, F.W.: Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* **120**, 3265–3266 (1998)

9. Loo, J.A., Edmonds, C.G., Smith, R.D.: Tandem mass-spectrometry of very large molecules. 2. Dissociation of multiply charged proline-containing proteins from electrospray ionization. *Anal. Chem.* **65**, 425–438 (1993)

10. Yu, W., Vath, J.E., Huberty, M.C., Martin, S.A.: Identification of the facile gas-phase cleavage of the asp pro and asp Xxx peptide-bonds in matrix-assisted laser-desorption time-of-flight mass-spectrometry. *Anal. Chem.* **65**, 3015–3023 (1993)

11. Bleiholder, C., Osburn, S., Williams, T.D., Suhai, S., Van Stipdonk, M., Harrison, A.G., Paizs, B.: Sequence-scrambling fragmentation pathways of protonated peptides. *J. Am. Chem. Soc.* **130**, 17774–17789 (2008)

12. Yague, J., Paradela, A., Ramos, M., Ogueta, S., Marina, A., Barahona, F., de Castro, J.A.L., Vazquez, J.: peptide rearrangement during quadrupole ion trap fragmentation: added complexity to MS/MS spectra. *Anal. Chem.* **75**, 1524–1535 (2003)

13. Horn, D.M., Ge, Y., McLafferty, F.W.: Activated ion electron capture dissociation for mass spectral sequencing of larger (42 KDa) proteins. *Anal. Chem.* **72**, 4778–4784 (2000)

14. Zubarev, R.A., Zubarev, A.R., Savitski, M.M.: Electron capture/transfer versus collisionally activated/induced dissociations: solo or duet? *J. Am. Soc. Mass Spectrom.* **19**, 753–761 (2008)

15. Theberge, R.; Mccomb, M. E.; Lin, C.; Costello, C. E. Top-Down Sequencing of Hemoglobin Variants with Multiple Activation Techniques. *Proceedings of the 58th ASMS Conference on Mass Spectrometry and Allied Topics*; Salt Lake City, UT, May, 2010.

16. Papayannopoulos, I.A.: The interpretation of collision-induced dissociation tandem mass-spectra of peptides. *Mass Spectrom. Rev.* **14**, 49–73 (1995)

17. Biemann, K.: Contributions of mass-spectrometry to peptide and protein-structure. *Biomed. Environ. Mass Spectrom.* **16**, 99–111 (1988)

18. Paizs, B., Suhai, S.: Towards understanding the tandem mass spectra of protonated oligopeptides. 1: mechanism of amide bond cleavage. *J. Am. Soc. Mass Spectrom.* **15**, 103–113 (2004)

19. Polfer, N.C., Oomens, J., Suhai, S., Paizs, B.: Spectroscopic and theoretical evidence for oxazolone ring formation in collision-induced dissociation of peptides. *J. Am. Chem. Soc.* **127**, 17154–17155 (2005)

20. Erlekam, U., Bythell, B.J., Scuderi, D., Van Stipdonk, M., Paizs, B., Maitre, P.: Infrared spectroscopy of fragments of protonated peptides: direct evidence for macrocyclic structures of B_5 ions. *J. Am. Chem. Soc.* **131**, 11503–11508 (2009)

21. Harrison, A.G.: Peptide sequence scrambling through cyclization of B_5 ions. *J. Am. Soc. Mass Spectrom.* **19**, 1776–1780 (2008)

22. Harrison, A.G.: Cyclization of peptide B_9 ions. *J. Am. Soc. Mass Spectrom.* **20**, 2248–2253 (2009)

23. Harrison, A.G., Young, A.B., Bleiholder, C., Suhai, S., Paizs, B.: Scrambling of sequence information in collision-induced dissociation of peptides. *J. Am. Chem. Soc.* **128**, 10364–10365 (2006)

24. Jia, C.X., Qi, W., He, Z.M.: Cyclization reaction of peptide fragment ions during multistage collisionally activated decomposition: an inducement to lose internal amino-acid residues. *J. Am. Soc. Mass Spectrom.* **18**, 663–678 (2007)

25. Molesworth, S., Osburn, S., Van Stipdonk, M.: Influence of size on apparent scrambling of sequence during Cid of B-type ions. *J. Am. Soc. Mass Spectrom.* **20**, 2174–2181 (2009)

26. Riba-Garcia, I., Giles, K., Bateman, R.H., Gaskell, S.J.: Evidence for structural variants of a- and B-type peptide fragment ions using combined ion mobility/mass spectrometry. *J. Am. Soc. Mass Spectrom.* **19**, 609–613 (2008)

27. Vachet, R.W., Bishop, B.M., Erickson, B.W., Glish, G.L.: Novel peptide dissociation: gas-phase intramolecular rearrangement of internal amino acid residues. *J. Am. Chem. Soc.* **119**, 5481–5488 (1997)

28. Tang, X.J., Thibault, P., Boyd, R.K.: Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass-spectrometry with low-energy collision-induced dissociation. *Anal. Chem.* **65**, 2824–2834 (1993)

29. Tang, X.J., Boyd, R.K.: Rearrangements of doubly-charged acylium ions from lysyl and ornithyl peptides. *Rapid Commun. Mass Spectrom.* **8**, 678–686 (1994)

30. Harrison, A.G.: To B or Not to B: the ongoing saga of peptide B ions. *Mass Spectrom. Rev.* **28**, 640–654 (2009)

31. Allen, J.M., Racine, A.H., Berman, A.M., Johnson, J.S., Bythell, B.J., Paizs, B., Glish, G.L.: Why are A_3 ions rarely observed? *J. Am. Soc. Mass Spectrom.* **19**, 1764–1770 (2008)

32. Chen, X., Yu, L., Steill, J.D., Oomens, J., Polfer, N.C.: Effect of peptide fragment size on the propensity of cyclization in collision-induced dissociation: oligoglycine B_2-B_8. *J. Am. Chem. Soc.* **131**, 18272–18282 (2009)

33. Grewal, R.N., El Aribi, H., Harrison, A.G., Siu, K.W.M., Hopkinson, A.C.: Fragmentation of protonated tripeptides: the proline effect revisited. *J. Phys. Chem. B* **108**, 4899–4908 (2004)

34. Jonsson, A.P., Bergman, T., Jornvall, H., Griffiths, W.J., Bratt, P., Stromberg, N.: Gln-Gly cleavage: correlation between collision-induced dissociation and biological degradation. *J. Am. Soc. Mass Spectrom.* **12**, 337–342 (2001)

35. Molesworth, S., Osburn, S., Van Stipdonk, M.: Influence of amino acid side chains on apparent selective opening of cyclic B(5) ions. *J. Am. Soc. Mass Spectrom.* **21**, 1028–1036 (2010)

36. Haselmann, K.F., Budnik, B.A., Zubarev, R.A.: Electron capture dissociation of B^{2+} peptide fragments reveals the presence of the acylium ion structure. *Rapid Commun. Mass Spectrom.* **14**, 2242–2246 (2000)

37. Chen, X.H., Turecek, F.: Simple B ions have cyclic oxazolone structures. A neutralization-reionization mass spectrometric and computational study of oxazolone radicals. *J. Am. Soc. Mass Spectrom.* **16**, 1941–1956 (2005)

38. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Montgomery Jr., J.A., Vreven, T., Kudin, K.N., Burant, J.C., Millam, J.M., Iyengar, S.S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G.A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J.E., Hratchian, H.P., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Ayala, P.Y., Morokuma, K., Voth, G.A., Salvador, P., Dannenberg, J.J., Zakrzewski, V.G., Dapprich, S., Daniels, A.D., Strain, M.C., Farkas, O., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Ortiz, J.V., Cui, Q., Baboul, A.G., Clifford, S., Cioslowski, J., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R.L., Fox, D.J., Keith, T., Al-Laham, M.A., Peng, C.Y., Nanayakkara, A., Challacombe, M., Gill, P.M.W., Johnson, B., Chen, W., Wong, M.W., Gonzalez, C., Pople, J.A.: Gaussian 03 C.02. Gaussian, Inc, Wallingford CT (2004)

39. Conlon, J.M., Adrian, T.E., Secor, S.M.: Tachykinins (Substance P, Neurokinin A, and Neuropeptide γ) and neurotensin from the intestine of the burmese python. Python molurus. *Peptides* **18**, 1505–1510 (1997)

40. Helke, C.J., Krause, J.E., Mantyh, P.W., Couture, R., Bannon, M.J.: Diversity in mammalian tachykinin peptidergic neurons: multiple peptides, receptors, and regulatory mechanisms. *FASEB J.* **4**, 1606–1615 (1990)

41. Regoli, D., Boudon, A., Fauchere, J.L.: Receptors and antagonists for substance P and related peptides. *Pharmacol. Rev.* **46**, 551–599 (1994)

42. Carter, M.S., Krause, J.E.: Structure, expression, and some regulatory mechanisms of the rat preprotachykinin gene encoding substance P, neurokinin A, neuropeptide K, and neuropeptide γ. *J. Neurosci.* **10**, 2203–2214 (1990)

43. Grace, R.C.R., Chandrashekar, I.R., Cowsik, S.M.: Solution structure of the tachykinin peptide eledoisin. *Biophys. J.* **84**, 655–664 (2003)

44. Wysocki, V.H., Tsaprailis, G., Smith, L.L., Breci, L.A.: Special feature commentary—mobile and localized protons: a framework for understanding peptide dissociation. *J. Mass Spectrom.* **35**, 1399–1406 (2000)

45. Leymarie, N., Costello, C.E., O'Connor, P.B.: Electron capture dissociation initiates a free radical reaction cascade. *J. Am. Chem. Soc.* **125**, 8949–8958 (2003)

46. Li, X., Lin, C., Han, L., Costello, C.E., O'Connor, P.B.: Charge remote fragmentation in electron capture and electron transfer dissociation. *J. Am. Soc. Mass Spectrom.* **21**, 646–656 (2010)

47. Sun, Q.Y., Nelson, H., Ly, T., Stoltz, B.M., Julian, R.R.: Side chain chemistry mediates backbone fragmentation in hydrogen deficient peptide radicals. *J. Proteome Res.* **8**, 958–966 (2009)

48. Zhang, L., Reilly, J.P.: Radical-driven dissociation of odd-electron peptide radical ions produced in 157 nm photodissociation. *J. Am. Soc. Mass Spectrom.* **20**, 1378–1390 (2009)

49. Fung, Y.M.E., Chan, T.W.D.: Experimental and theoretical investigations of the loss of amino acid side chains in electron capture dissociation of model peptides. *J. Am. Soc. Mass Spectrom.* **16**, 1523–1535 (2005)

50. Yalcin, T., Harrison, A.G.: Ion chemistry of protonated lysine derivatives. *J. Mass Spectrom.* **31**, 1237–1243 (1996)

51. Syka, J.E.P., Coon, J.J., Schroeder, M.J., Shabanowitz, J., Hunt, D.F.: Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U.S A* **101**, 9528–9533 (2004)