Published in final edited form as:

Anal Chem. 2014 March 18; 86(6): 2917–2924. doi:10.1021/ac401578p.

Gas-phase Ion Isomer Analysis Reveals the Mechanism of Peptide Sequence Scrambling

Chenxi Jia[†], Zhe Wu[‡], Christopher B. Lietz^{†,‡}, Zhidan Liang[†], Qiang Cui[‡], and Lingjun Li^{†,‡,*}

[†]School of Pharmacy, University of Wisconsin–Madison, Madison, Wisconsin 53705, USA

[‡]Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53705, USA

Abstract

Peptide sequence scrambling during mass spectrometry-based gas-phase fragmentation analysis causes misidentification of peptides and proteins. Thus, there is a need to develop an efficient approach to probing the gas-phase fragment ion isomers related to sequence scrambling and the underlying fragmentation mechanism, which will facilitate the development of bioinformatics algorithm for proteomics research. Herein, we report on the first use of electron transfer dissociation (ETD)-produced diagnostic fragment ions to probe the components of gas-phase peptide fragment ion isomers. In combination with ion mobility spectrometry (IMS) and formaldehyde labeling, this novel strategy enables qualitative and quantitative analysis of b-type fragment ion isomers. ETD fragmentation produced diagnostic fragment ions indicative of the precursor ion isomer components, and subsequent IMS analysis of b ion isomers provided their quantitative and structural information. The isomer components of three representative b ions (b₉, b₁₀, and b₃₃ from three different peptides) were accurately profiled by this method. IMS analysis of the b₉ ion isomers exhibited dynamic conversion among these structures. Furthermore, molecular dynamics simulation predicted theoretical drift time values which were in good agreement with experimentally measured values. Our results strongly support the mechanism of peptide sequence scrambling via b ion cyclization, and provide the first experimental evidence to support that the conversion from molecular precursor ion to cyclic b ion (M→^cb) pathway is less energetically (or kinetically) favored.

INTRODUCTION

Mass spectrometry (MS)-based proteomics has become an attractive technology for global analysis of protein composition, modifications, and dynamics.^{1–2} Tandem mass spectrometry (MS/MS) analysis of peptides is an essential tool for proteomic studies.^{2–3} Although bioinformatics algorithms provide necessary support to process the enormous quantity of MS/MS spectra generated from large scale analysis of biological samples, false

Correspondence: Professor Lingjun Li, School of Pharmacy and Department of Chemistry, University of Wisconsin–Madison, 777 Highland Avenue, Madison, Wisconsin 53705-2222, lli@pharmacy.wisc.edu, Fax: +1-608-262-5345.

NOTES

The authors declare no competing financial interest.

Supporting Information

Additional information as noted in the text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

positive identifications of peptides and proteins may exist.² One of the reasons causing such misidentifications stems from peptide sequence scrambling that could occur during tandem MS analysis of peptides via collision-induced dissociation (CID).⁴

Typically, CID causes the peptide bonds to be sequentially cleaved along the backbone to produce so-called “direct sequence ions”.⁴ However, if a peptide chain is cyclized and rearranged in the gas phase, these resulting ion isomers will undergo further dissociation to produce “non-direct sequence ions” (Figure S1) leading to sequence scrambling.⁴⁻⁵ Interpreting these fragment ions without an efficient bioinformatics algorithm may lead to misidentification of peptides and proteins.

Recent studies⁴ reported that head-to-tail cyclization of peptide b-type fragment ions under CID fragmentation caused sequence scrambling. Two fragmentation pathways (Figure S2) were proposed and studied.⁴⁻²⁹ In $M \rightarrow ^c b$ pathway (M, molecular ion of peptides; $^c b$, cyclic b ion; see Table S1 for special ion symbols), the primary amine at the N-terminus of the peptide chain reacts with the carbonyl carbon near the C-terminus of the peptide chain via nucleophilic attack to produce $^c b$ ion. The $^c b$ ion can also be produced via $^{ol} b \rightarrow ^c b$ pathway by head-to-tail cyclization of $^{ol} b$ ion ($^{ol} b$, original linear b ions). However, for analysis of polypeptides from complicated biological samples, there are many factors influencing the propensity for various peptides to form $^c b$ ions thus leading to sequence scrambling. These factors include peptide length,^{9, 26} residue acidity,¹² proline and histidine effects,^{13, 21} and residue side chains,¹⁴ etc. Therefore, it is important to establish a method for accurate and fast identification of ion isomers, which allows for large-scale analysis of b ions to improve our understanding of the underlying fragmentation chemistry and facilitate development of bioinformatics algorithm.

Structural elucidation and profiling of ion isomers are challenging, because these gas-phase ion isomers share identical m/z values in MS measurement and they may undergo dynamic conversion which could be difficult to characterize. Previous studies have established a great variety of strategies for analysis of ion isomers, such as infrared spectroscopy,^{9, 15, 27} gas-phase H/D exchange,^{15, 24} guest-host chemistry,¹⁰ etc, which provided useful insights into underlying mechanism of peptide sequence scrambling. However, these strategies lack the capability to elucidate the precise components of various ion isomers in a complex mixture. To fill in this gap, herein we employ novel use of electron transfer dissociation (ETD), an electron-based fragmentation technique, to probe the gas-phase composition and structures of fragment ion isomers. In combination with ion mobility spectrometry (IMS) and formaldehyde (FH) labeling, this novel strategy enables simultaneous qualitative and quantitative analysis of various b ion isomers including linear and cyclic b ions such as $^{ol} b$, $^c b$, and $^{rl} b$ ions ($^{rl} b$, rearranged linear b ions).

EXPERIMENTALS

Materials and Chemicals

All chemical reagents were obtained from Sigma-Aldrich (St. Louis, MO) unless otherwise noted. Optima grade formic acid, acetonitrile (ACN) and water were purchased from Fisher Scientific (Pittsburgh, PA). Peptide standards, neurokinin and Substance P were purchased

from American Peptide Co. (Sunnyvale, CA). The crustacean hyperglycemic hormone precursor related peptide (CPRP) was isolated and purified from sinus glands of blue crabs *Callinectes sapidus* as described in our previous report.³⁰

Formaldehyde Labeling of Peptides

One microgram of peptide sample was labeled in 10 μL of water solution by adding 1 μL of borane pyridine ($\text{C}_5\text{H}_8\text{BN}$, 120 mM in 10% methanol) and then mixing with formaldehyde (15% in H_2O , 1 μL). The reaction mixture was then vortexed at room temperature for 15 min and quenched with ammonium bicarbonate solution (1 μL , 0.2 M). After drying down in Speedvac, the sample was desalted by Ziptip for direct infusion analysis on mass spectrometers.

Fragmentation of b Ions by ETD

The ETD experiments were performed on an amaZon ETD ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with a CaptiveSpray ESI source. Optimization of the CaptiveSpray source resulted in dry gas temperature, 130 $^{\circ}\text{C}$, dry gas, 6.0 L/min, capillary voltage, -1250 V , end plate offset, -500 V . The peptide sample was dissolved in ACN-water-formic acid (50:50:0.1) at a concentration of 100 ng/mL and was directly infused into mass spectrometer at 300 nL/min. The peptide molecular ions were selected with a 4 Da isolation window in MS^2 and then fragmented by collision induced dissociation (CID) with an adjusted fragmentation amplitude (0.7–1.2) that allows the parent ions to be completely fragmented. In MS^3 the CID-produced b ions were further selected with a 2 Da isolation window and fragmented with ETD. The ion charge control (ICC) target was set to 200 000, maximum accumulation time, 120 ms, 10 spectral averages, 5 rolling averaging (a filter that operates on the time series of mass spectra that are generated by the ion trap), and acquisition range of m/z 100–1800. The ETD reagent parameters were set to 400 000 ICC target for the fluoranthene radical anion and the ETD duration time was 100 ms. Data were acquired in enhanced resolution mode with ~ 3000 of resolving power for 1 hour and processed using software DataAnalysis to generate a high quality MS^3 ETD spectrum. The fragment ions were assigned within 0.1 Da of mass error.

Ion Mobility Analysis

The IMS experiments were performed using a Synapt G2 HDMS mass spectrometer equipped with a nano-ESI ion source and MassLynx data processor (Waters, Milford, MA, USA). Instrument acquisition parameters used were as follows: an inlet capillary voltage of 3.0 kV, a sampling cone setting of 35 V, and a source temperature of 120 $^{\circ}\text{C}$. The argon gas pressure in the traveling wave ion guide trap and the traveling wave ion guide transfer cell were 2.44×10^{-2} and 2.61×10^{-2} mbar, respectively. The wave height, the wave velocity, and the nitrogen pressure in the traveling wave ion mobility drift cell were 32.0 V, 800 m/s, and 2.96 mbar, respectively. Samples were directly infused into the mass spectrometer at a rate of 0.5–0.8 $\mu\text{L}/\text{min}$. The peptide molecular ions were selected and fragmented in Traveling-Wave trap cell by CID with an adjusted collision energy of 25–30 eV that allows the parent ions to be completely fragmented. All the fragment ions were submitted into drift tube with their drift time measured. The spectra were acquired for 5 min. Data processing

was conducted using Waters MassLynx 4.1 and DriftScope 2.1. The experimental procedures for collision cross section measurement are described in Supporting Information.

The dynamics study was conducted as shown in the schematic of Figure 3A. The peptides were fragmented by in-source dissociation (ISD) in nozzle-skimmer region. To result in ISD, the adjusted source parameters were used: an inlet capillary voltage of 3.5 kV, a sampling cone setting of 55 V, and a source temperature of 120 °C. The target b ions were selected by quadrupole and accumulated in trap cell, where various collision energies 0–22 eV were used to activate the b ions. Subsequently, the b ions were submitted into drift tube with the same parameters as described above.

High-resolution CID MS/MS of CPRP Neuropeptide

The high-resolution CID MS/MS of CPRP peptide was carried out on a 7T linear trap quadrupole (LTQ)/Fourier transform ion cyclotron resonance (FTICR) (LTQ-FT Ultra) hybrid mass spectrometer (Thermo Scientific Inc., Bremen, Germany) as previous report.³¹ The experimental details are described in Supporting Information.

Molecular Dynamics (MD) Simulation

MD simulations were performed using the program CHARMM (c37) with the CHARMM22 force field for proteins.³² Parameters were generated with the CHARMM General Force field (CGenFF) program³³ for the oxazolone structure. MD simulations were performed in vacuum at 450 K, which mimics the low gas density environment in the drift tube with enhanced molecular vibrations due to N₂ collision. Standard non-bonded cut-off scheme and temperature-coupling protocol in CHARMM were adopted. For each molecule, a 30 ns MD trajectory was calculated and the last 15 ns was used to generate structures for the collision cross section (Ω_{He}) calculations. For each b ion isomer in Figures S3 and S4 (protonation sites listed in Table S2), 1500 structures were collected to compute the Ω_{He} values using the Sigma program^{34–36} with a scaled Leonard-Jones projection protocol. Finally, the Ω_{He} values were converted into drift times to generate distributions of drift times for the b ion isomers. Nonlinear fitting was carried out using Gaussian functions. It should be noted that the Synapt G2 HDMS used 0.069 ms of bin size as default setting. To be consistent, the same bin size was used in theoretical simulation.

RESULTS AND DISCUSSION

Figure 1 describes the workflow of the proposed strategy for fragment ion isomer analysis. The CID-produced b ion isomers are subjected to ETD fragmentation and IMS analysis, respectively. The formaldehyde (FH)-labeled peptides are fragmented by CID to generate FH-labeled linear b ion, followed by ETD fragmentation and IMS analysis, respectively. On the left panel of Figure 1, under ETD fragmentation each b ion isomer generates unique diagnostic fragment ions. The FH-labeled b ion (dimethylation on N-terminus and lysine side chain)³⁷ is used for alignment. On the right panel, each b ion isomer is separated in IMS, followed by alignment with FH-labeled b ions. Because the dimethylation on the N-terminus of FH-labeled b ion prevents the formation of cyclized b-ion and subsequent ring

opening and rearrangement of linear b-ions, the IMS peak of the FH-labeled b ions facilitates the assignment of original linear b-ion structure in the mixture.

ETD fragmentation is based on gas-phase ion/ion reactions, where an anion transfers one electron to the protonated peptide cation to initiate radical-driven backbone cleavage for peptide fragmentation.^{38–39} Unlike CID which relies on “slow heating” threshold fragmentation, ETD occurs in a much shorter time scale than that is required for internal energy distribution, thus resulting in rapid fragmentation and reducing the propensity of b ion cyclization.³⁸ Previous studies^{40–41} have demonstrated that ETD or electron capture dissociation (ECD) of cyclic peptides can induce a free radical reaction cascade generating [z-c] and [z-a] ions. Because the ^cb ion has the same structure as a protonated cyclic peptide, ETD fragmentation of ^cb ion produces the same diagnostic ions [z-c] and [z-a] as well. In contrast, ETD fragmentation of ^lb ions generates diagnostic [z-b] and c ions. This distinction provides the key molecular signatures to differentiate ^cb from ^lb ions. Additionally, ^ol^b and ^rl^b ions can be distinguished by assigning sequential c and [z-b] ions. Figure S5 illustrates the structures of these diagnostic ions. Recently, Li et al⁴² reported observation of these diagnostic fragment ions from ^cb and ^lb ions using ECD. Herein, we extend this concept to the development of a novel diagnostic analysis of ion isomers using ETD, as ETD may minimize internal energy distribution and thus reduce secondary fragmentation.³⁸ IMS analysis offers accurate profiling of these b ions by measuring drift time differences of isobaric isomer species due to their different shapes and conformations in the gas phase. Alignment of unlabeled b ion with _{FH}.b (formaldehyde-labeled b ion) provides complementary information to ETD and IMS analyses. To demonstrate this strategy, we analyzed three representative b ions from three peptides with diverse isomer components, i.e., neurokinin b₉ ion, substance P b₁₀ ion, and CPRP b₃₃ ion.

Neurokinin b₉ ion

The ETD experiments were performed on an ion trap Bruker amaZon with electrospray ionization (ESI). The [M+2H]²⁺ ion of neurokinin (HKTDSFVGLM-NH₂) was selected in MS² and fragmented using CID to generate the b₉²⁺ ion, followed by ETD fragmentation in MS³. In the resulting ETD spectrum (Figure 2A), fragment ions from four different b ion isomers were detected, i.e., c₂~c₇ from ^ol^b₉²⁺ (HKTDSFVGL²⁺), [z-b]_{A3}~[z-b]_{A8} from ^rl^b_{A9}²⁺ (TDSFVGLHK²⁺), [z-b]_{B5}~[z-b]_{B8} from ^rl^b_{B9}²⁺ (SFVGLHKTD²⁺), as well as [z-c]₃~[z-c]₇ and [z-a]₃~[z-a]₇ from ^cb₉²⁺ (cyclo-HKTDSFVGL²⁺) (See Table S1 for special ion symbols). In contrast, the N-terminus and lysine side chain were dimethylated with formaldehyde,³⁷ which blocks cyclization reaction of b ions (mechanisms shown in Figure S6). Thus, the _{FH}.b₉²⁺ ion produced by CID should maintain linear structure. As expected, the ETD spectrum of _{FH}.b₉ in MS³ (Figure 2C) only contains c and [z-b] ions from ^lb₉, confirming that FH labeling blocks the formation of cyclic b ion. Subsequently, the IMS experiments were performed on an ESI-QTOF Waters Synapt G2 HDMS. The CID-produced b₉²⁺ ion of neurokinin was submitted into IMS drift tube with drift time measured. The resulting IMS distribution (Figure 2B) shows four peaks at 4.02, 4.29, 4.43 and 4.78 ms respectively corresponding to four ion isomers. Their collision cross sections Ω_{He} were 229, 237, 244 and 256 Å², which were converted from Ω_{N_2} according to the method reported by Bush et al.⁴³ (Table S3 lists all Ω_{He} determined in this study). This was in agreement with

the four isomers determined by ETD method (Figure 2A). It should be noted that the isomer peak at 4.29 ms was further resolved by the subsequent dynamic study. In contrast, IMS analysis of ${}_{\text{FH}}\text{b}_9^{2+}$ (Figure 2D) displayed only one peak at 4.55 ms ($248 \text{ \AA}^2$), which was also consistent with the ETD data (Figure 2C). Here, we tentatively assign the peak at 4.02 ms in Figure 2B as ${}^{\text{c}}\text{b}_9^{2+}$, since cyclic structure is more compact than the linear ones and has the highest mobility and thus shortest drift time. The peak at 4.43 ms is tentatively assigned as the ${}^{\text{ol}}\text{b}_9^{2+}$ ion, as the drift time of ${}_{\text{FH}}\text{b}_9^{2+}$ at 4.55 ms is the closest to the peak at 4.43 ms. This argument can be supported by the fact that the $[\text{M}+2\text{H}]^{2+}$ and ${}_{\text{FH}}[\text{M}+2\text{H}]^{2+}$ of neurokinin (Figure S7) also show similar drift times under the same IMS conditions. In addition, IMS can resolve conformational isomers as well as structural isomers.⁴⁴ The IMS distribution of intact peptide of neurokinin (Figure S7A) showed one major peak corresponding to the presence of a single unique IMS-resolved conformation, so most likely the four peaks in IMS distribution of b_9^{2+} (Figure 2B) were different structural isomers instead of different conformational isomers.

To study the dynamics of structural conversion among these isomers under CID activation, we designed the MS^3 -IMS experiment (Figure 3A) on Synapt G2. The b_9^{2+} ion isomers were generated by in-source dissociation (ISD) of neurokinin $[\text{M}+2\text{H}]^{2+}$ ion in nozzle-skimmer region (MS^2), isolated by quadrupole, and then was further activated by collision energy in the Triwave[®] trap cell (MS^3). After that, all the ion isomers were submitted into drift tube to monitor their conversion dynamics. The benefit of this design is that tandem MS in the trap cell can be optimized to directly activate the target fragment ions, instead of precursor ions, thus improving the capability to monitor the conversion dynamics of these ion isomers. Figure 3B shows the IMS distributions of the b_9^{2+} isomers after activation by various collision energy (CE) levels. With elevation of CE from 0 to 22 eV, the relative peak abundances at 4.02 and 4.43 ms were reduced; and those at 4.29 and 4.78 ms exhibited consistent elevation. This trend revealed the conversion dynamics of these ion isomers (depicted in Figure S3), with the peak at 4.02 ms assigned to ${}^{\text{c}}\text{b}_9^{2+}$; 4.43 ms attributed to ${}^{\text{ol}}\text{b}_9^{2+}$; and 4.29 and 4.78 ms assigned to the two ${}^{\text{r}}\text{b}_9^{2+}$ ions.

To further confirm the above assignments, we conducted the MS^3 -IMS- MS^4 experiment (Figure 3A). All the ions from the IMS drift tube were fragmented in the Triwave[®] transfer cell (MS^4) by CID with varied CE in the trap cell. This experiment allowed the mobility-separated ${}^{\text{c}}\text{b}_9^{2+}$ and ${}^{\text{ol}}\text{b}_9^{2+}$ ions to be fragmented at successive time points. Ideally, CID of ${}^{\text{c}}\text{b}_9^{2+}$ should produce higher abundance of rearranged fragment ions than that of ${}^{\text{ol}}\text{b}_9^{2+}$. As expected, we observed different intensity ratio of b_8 and ${}^{\text{r}}\text{b}_8$ (1:3.7 and 1:2) in the two resulting CID spectra (Figure 3C and D). The b_8 ion arose directly from ${}^{\text{ol}}\text{b}_9^{2+}$, and the rearranged fragment ion ${}^{\text{r}}\text{b}_8$ was produced from ${}^{\text{c}}\text{b}_9^{2+}$ via pathway ${}^{\text{c}}\text{b}_9^{2+} \rightarrow {}^{\text{r}}\text{b}_9^{2+} \rightarrow {}^{\text{r}}\text{b}_8$ (details shown in Figure S8). This result provided the third piece of evidence to confirm the assignment of ${}^{\text{ol}}\text{b}_9^{2+}$ and ${}^{\text{c}}\text{b}_9^{2+}$ in Figure 3B. Although the results from ETD and the conversion dynamics studies indicate that the peaks at 4.29 and 4.78 ms were the two rearranged ions ${}^{\text{r}}\text{b}_{\text{A}9}^{2+}$ and ${}^{\text{r}}\text{b}_{\text{B}9}^{2+}$, the order of their mobility in the IMS drift tube could not be determined.

To further assign the four ion isomers, we calculated their Ω_{He} distributions by molecular dynamics simulation.^{45–46} In practice, the ISD-produced b ion isomers were calibrated by

polyalanine under the MS³-IMS experimental condition. Figure S9 displays the measured Ω_{He} distributions. It should be noted that the individual drift times of the four isomers remained consistent under CE 0~22 eV, assuring that the experimental CE in the trap cell used here would not cause remarkable drift time shifts. For theoretical simulation,⁴⁷ the structures (Figure S3 and Table S2) of the four b ion isomers were sampled by molecular dynamics, and then the Ω_{He} values were calculated using the Sigma program³⁵ with a scaled Leonard-Jones projection approximation procedure (see Supporting Information). To make a direct comparison with Figure 3B, the theoretical Ω_{He} values were converted to drift times, followed by statistical processing to generate simulated distributions in Figure 3E. The theoretical and experimental drift time values agree with each other, supporting the assignment of the four ion isomers. Collectively, the agreement of the results from ETD fragmentation, IMS measurement, and theoretical simulation support the proposed mechanism of b ion cyclization outlined in Figure S3. In addition to the standard ^rb ions containing oxazolone structures at the C-terminus, two non-standard ^rb ions could be produced respectively via two alternative ring opening pathways due to nucleophilic attacks of the side chains of Lys and Asp (Figure S4).⁴⁸ The two non-standard ^rb ions, termed as $^r\text{b}_{\text{A9}}^{2+}$ and $^r\text{b}_{\text{B9}}^{2+}$, respectively contain the same amino acid sequences as $^r\text{b}_{\text{A9}}^{2+}$ and $^r\text{b}_{\text{B9}}^{2+}$, but different C-terminal structures (Figure S4). Figure S10 shows the theoretical drift time distributions of $^r\text{b}_{\text{A9}}^{2+}$ and $^r\text{b}_{\text{B9}}^{2+}$ centered at 4.32 and 4.64 ms, respectively, which agrees with the measured drift times of $^r\text{b}_{\text{A9}}^{2+}$ and $^r\text{b}_{\text{B9}}^{2+}$ at 4.29 and 4.78 ms (Figure 3B). Although our experimental results and theoretical calculations support the assignment of the four ion isomers in Figure 3B, further studies using other methodologies, such as infrared photodissociation spectroscopy,⁴⁹ gas-phase hydrogen/deuterium exchange,¹⁵ etc could provide further evidence to make the assignments of these ion isomer structures more confident. In addition, it was previously reported that the primary amine on the lysine side chain in a peptide may attack the C-terminus of the peptide chain via nucleophilic reaction to produce a macrocycle structure.⁵⁰ In the ETD spectrum (Figure 2A) of b_9^{2+} , we did not observe any fragment ions from this macrocycle structure, indicating that this side reaction may occur via a minor pathway (or did not happen under the experimental conditions used in this study).

Substance P b_{10} ion

Figure 2E is the ETD fragmentation of the b_{10}^{2+} ion from substance P (RPKPQQFFGLM-NH₂). Interestingly, one set of [z-b] ions was clearly detected, which arose from the $^r\text{b}_{10}^{2+}$ ion (PQQFFGLRPK²⁺). Consistently, only one peak at 4.90 ms (260 Å²) was present in the corresponding IMS distribution of this b_{10}^{2+} ion (Figure 2F). Thus, this IMS peak can be assigned as $^r\text{b}_{10}^{2+}$. Furthermore, ETD fragmentation of the $_{\text{FH}}\text{b}_{10}^{2+}$ ion (Figure 2G) produced a set of c ions with original sequence; and accordingly the IMS distribution (Figure 2H) showed one peak at 5.74 ms (278 Å²). Aligning the two IMS peaks of $^r\text{b}_{10}^{2+}$ and $_{\text{FH}}\text{b}_{10}^{2+}$ revealed a large drift time shift of 0.84 ms, in contrast to small deviation often observed for original linear b-ion and their formaldehyde labeled pair. Therefore, we conclude that the peak at 4.90 ms can be assigned as the rearranged linear b-ion $^r\text{b}_{10}^{2+}$.

CPRP b_{33} ion

Tirado and Polfer⁹ recently developed an infrared multiple photon dissociation-based strategy which allows for probing the components of b ions containing up to 12 amino acids. To the best of our knowledge, the b_{12} ion was the largest b ion whose component has been accurately analyzed so far. Thus, examination of a b-type ion with 33 amino acid residues using our strategy may provide useful insight into peptide scrambling that occurs in large b ions. We investigated the b_{33}^{4+} ion of CPRP (RSAEGLGRMGRLASLKSDTVTPLRGFEGETGHPLE), a neuropeptide isolated from the sinus gland of blue crab *Callinectes sapidus*³⁰. Figure 4A shows the CID spectrum (MS^2) of $[M+5H]^{5+}$ from CPRP on a high resolution instrument LTQ-FTICR, giving rise to abundant b_{33}^{4+} ion. CID fragmentation of b_{33}^{4+} in MS^3 (Figure 4B) produced intense $^{\circ}b$ ions (ion nomenclature proposed by Chawner et al.⁵¹) leading to sequence scrambling. The proposed pathway of b_{33} cyclization and re-arrangement is shown in Figure S12. Interestingly, two $^{\circ}b$ ions, $[b_{33}19]b_{29}^{3+}$ and $[b_{33}19]b_{14}^{+}$ with very low intensities were observed in the MS^2 spectrum. This suggests that peptide sequence scrambling may be a minor problem in top-down MS^2 , while CID fragmentation of large b ions in MS^3 can cause significant sequence scrambling, leading to misidentification of large peptides. The result obtained from the b_{33}^{4+} ion also represents the largest b-type ion that undergoes cyclization and rearrangement in the gas phase documented to date.

To investigate the isomer component of this b_{33}^{4+} ion, ETD, FH labeling and IMS experiments were performed. Surprisingly, ETD fragmentation patterns (Figures 2I and 2K) of b_{33}^{4+} and $_{FH}b_{33}^{4+}$ were remarkably consistent, where two sets of c and [z-b] ions from the original sequence were observed in each ETD spectrum. IMS distributions (Figures 2J and 2L) of b_{33}^{4+} and $_{FH}b_{33}^{4+}$ agree with each other as well, where the two peaks (6.14 and 6.00 ms) has 0.14 ms of shift due to N-terminal dimethylation. These results clearly indicate that the IMS peak at 6.14 ms is $^{ol}b_{33}^{4+}$. In other words, CID-produced b_{33}^{4+} in MS^2 maintains the original linear structure without rearrangement.

Previous studies⁴⁻⁵ proposed two pathways regarding formation of $^{\circ}b$ ions, $M \rightarrow ^{\circ}b$ and $M \rightarrow ^{ol}b \rightarrow ^{\circ}b$ (Figure S2). Harrison, Paizs and co-workers⁴ suggested that the former pathway is less energetically favored via theoretical calculation. So far, there has been no experimental evidence to support this theoretical prediction, because all of the b ions under study undergo cyclization in MS^2 . Therefore, a critical experiment is to find the b ions in MS^2 displaying original linear structures as the major molecular species. Interestingly, our results indicated that b_{33}^{4+} maintains the original linear sequence via CID fragmentation of $[M+5H]^{5+}$, suggesting that majority of b_{33}^{4+} ion did not undergo cyclization in MS^2 . This finding provided the first experimental evidence to support the calculation result⁴ that the $M \rightarrow ^{\circ}b$ pathway is indeed less energetically or kinetically favored.

Our strategy enables qualitative and quantitative analysis of b ion isomers. Table 1 summarizes the analysis results of the three representative b ions, including the ion isomer components and their corresponding ratios in relative abundances. The goal of this work is to improve our understanding of the underlying fragmentation chemistry of peptide sequence scrambling and to facilitate the development of bioinformatics algorithm for

proteomics research. Although our strategy is demonstrated in the study of multiply charged b ions, the general fragmentation mechanism and identification rules are most likely applicable to b ions at both multiply charged and singly charged states. To apply our strategy to large-scale analysis of b ion components, we may choose an alternative enzyme such as Lys-N⁵² to digest protein mixture and generate multiply charged peptide ions containing basic amino acid residues at the N-terminus, which produces high abundance of multiply charged b ions. These b ions will be subsequently analyzed by ETD and IMS, followed by computer-assisted data processing. The qualitative and quantitative information of b ion isomer composition can be extracted from the resulting ETD spectra and IMS distributions, respectively. In this study, the IMS and ETD experiments were respectively performed on two different mass spectrometers, because the Synapt G2 mass spectrometer utilized for IMS experiments in this study does not offer the capability of ETD fragmentation. If the Synapt G2 IMS mass spectrometer is upgraded by installation of an ETD source,⁵³ the peptide can be fragmented by in-source dissociation and the resulting b ions can be transferred to Triwave trap cell for ETD fragmentation. Therefore, with the upgraded instrumentation it would be possible to conduct the IMS and ETD experiments using the same instrument.

In addition, Li et al.⁴² previously applied ECD to fragment b ions and study cyclization of b ions. However, ECD-produced fragment ions were not utilized as a diagnostic tool for b ion isomer component analysis. Due to relatively higher energy level applied to these ions, ECD may cause secondary fragmentation, which may not provide accurate information to allow characterizing the components of b ions. Moreover, compared with ECD technique realized on FTICR instruments, mass spectrometers equipped with ETD technique often provide much faster scan rate, which may improve the throughput of analysis.³⁸ In addition, from the ETD spectra of the three FH .b ions (Figure 2C, G and K), only the fragment ions from linear structures were observed; and IMS distributions of the three FH .b ions (Figure 2D, H and L) only displayed as single peaks. These experimental evidence supports that *N,N*-dimethylation with formaldehyde can fully prevent the cyclization of b ions, therefore producing greatly simplified fragmentation due to elimination of sequence scrambling. Fu and Li⁵⁴ previously reported that the CID spectra of *FH*-labeled peptides were simplified due to the formation of enhanced N-terminal fragment ions and suppressed internal fragments. Our current work provided strong experimental support that such fragmentation simplification and improved *de novo* sequencing is mainly due to the blockage of cyclic b-ion formation via *N,N*-dimethylation.

CONCLUSIONS

In comparison with previous studies,^{9–10, 15} our method has several advantages: (1) all the isomers (^{ol}b , cb , and ^{rl}b) can be simultaneously probed; (2) the experiments can be conducted using commercially available mass spectrometers; and (3) the method is applicable for large-scale screening. In this study, ion isomer analysis of the three representative b ions has demonstrated the utility of a novel strategy combining experimental analysis via ETD, IMS and chemical labeling and theoretical calculation using molecular dynamic simulation. The detailed analyses of gas-phase ion isomer components and their dynamic conversions reveal the proposed fragmentation mechanism of b ion

cyclization, strategies to minimize such sequence scrambling and their potential impact on large-scale proteomic studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by the National Institutes of Health (R01DK071801 to LL) and National Science Foundation (CHE-0957784 to LL, CHE-0957285 to QC, and CHE-0840494 to phoenix cluster). LL acknowledges an HI Romnes Faculty Research Fellowship, and CBL thanks an NIH-CBIT fellowship (T32-GM008505) and a NSF graduate research fellowship. We thank Prof. Cheng Lin for his helpful discussions about fragmentation chemistry. We are also grateful to the UW School of Pharmacy Analytical Instrumentation Center and UW Human Proteomics Program for access to mass spectrometers.

References

1. Walther TC, Mann M. *J Cell Biol.* 2010; 190:491–500. [PubMed: 20733050]
2. Aebersold R, Mann M. *Nature.* 2003; 422:198–207. [PubMed: 12634793]
3. Wysocki VH, Resing KA, Zhang Q, Cheng G. *Methods.* 2005; 35:211–222. [PubMed: 15722218]
4. Harrison AG, Young AB, Bleiholder C, Suhai S, Paizs B. *J Am Chem Soc.* 2006; 128:10364–10365. [PubMed: 16895391]
5. Yague J, Paradelo A, Ramos M, Ogueta S, Marina A, Barahona F, Lopez de Castro JA, Vazquez J. *Anal Chem.* 2003; 75:1524–1535. [PubMed: 12659218]
6. Jia C, Qi W, He Z. *J Am Soc Mass Spectrom.* 2007; 18:663–678. [PubMed: 17234429]
7. Bleiholder C, Osburn S, Williams TD, Suhai S, Van Stipdonk M, Harrison AG, Paizs B. *J Am Chem Soc.* 2008; 130:17774–17789. [PubMed: 19055406]
8. Harrison AG. *Mass Spectrom Rev.* 2009; 28:640–654. [PubMed: 19338048]
9. Tirado M, Polfer NC. *Angew Chem Int Ed Engl.* 2012; 51:6436–6438. [PubMed: 22615257]
10. Somogyi A, Harrison AG, Paizs B. *J Am Soc Mass Spectrom.* 2012; 23:2055–2058. [PubMed: 22996293]
11. Polfer NC, Bohrer BC, Plasencia MD, Paizs B, Clemmer DE. *J Phys Chem A.* 2008; 112:1286–1293. [PubMed: 18215025]
12. Atik AE, Yalcin T. *J Am Soc Mass Spectrom.* 2011; 22:38–48. [PubMed: 21472542]
13. Bythell BJ, Knapp-Mohammady M, Paizs B, Harrison AG. *J Am Soc Mass Spectrom.* 2010; 21:1352–1363. [PubMed: 20541953]
14. Molesworth S, Osburn S, Van Stipdonk M. *J Am Soc Mass Spectrom.* 2010; 21:1028–1036. [PubMed: 20219391]
15. Chen X, Yu L, Steill JD, Oomens J, Polfer NC. *J Am Chem Soc.* 2009; 131:18272–18282. [PubMed: 19947633]
16. Saminathan IS, Wang XS, Guo Y, Krakovska O, Voisin S, Hopkinson AC, Siu KW. *J Am Soc Mass Spectrom.* 2010; 21:2085–2094. [PubMed: 20947370]
17. Goloborodko AA, Gorshkov MV, Good DM, Zubarev RA. *J Am Soc Mass Spectrom.* 2011; 22:1121–1124. [PubMed: 21953093]
18. Durand S, Rossa M, Hernandez O, Paizs B, Maitre P. *J Phys Chem A.* 2013; 117:2508–2516. [PubMed: 23437986]
19. Dong NP, Liang YZ, Yi LZ. *J Am Soc Mass Spectrom.* 2012; 23:1209–1220. [PubMed: 22539146]
20. Tasoglu C, Gorgulu G, Yalcin T. *Int J Mass Spectrom.* 2012; 316:108–116.
21. Miladi M, Zekavat B, Munisamy SM, Solouki T. *Int J Mass Spectrom.* 2012; 316:164–173.

22. Chen X, Tirado M, Steill JD, Oomens J, Polfer NC. *J Mass Spectrom.* 2011; 46:1011–1015. [PubMed: 22012667]
23. Molesworth SP, Van Stipdonk MJ. *J Am Soc Mass Spectrom.* 2010; 21:1322–1328. [PubMed: 20219393]
24. Fattahi A, Zekavat B, Solouki T. *J Am Soc Mass Spectrom.* 2010; 21:358–369. [PubMed: 20005740]
25. Harrison AG. *J Am Soc Mass Spectrom.* 2009; 20:2248–2253. [PubMed: 19781958]
26. Molesworth S, Osburn S, Van Stipdonk M. *J Am Soc Mass Spectrom.* 2009; 20:2174–2181. [PubMed: 19758821]
27. Erlekan U, Bythell BJ, Scuderi D, Van Stipdonk M, Paizs B, Maitre P. *J Am Chem Soc.* 2009; 131:11503–11508. [PubMed: 19637928]
28. Harrison AG. *J Am Soc Mass Spectrom.* 2008; 19:1776–1780. [PubMed: 18639467]
29. Riba Garcia I, Giles K, Bateman RH, Gaskell SJ. *J Am Soc Mass Spectrom.* 2008; 19:1781–1787. [PubMed: 18930410]
30. Hui L, Cunningham R, Zhang Z, Cao W, Jia C, Li L. *J Proteome Res.* 2011; 10:4219–4229. [PubMed: 21740068]
31. Jia C, Hui L, Cao W, Lietz CB, Jiang X, Chen R, Catherman AD, Thomas PM, Ge Y, Kelleher NL, Li L. *Mol Cell Proteomics.* 2012; 11:1951–1964. [PubMed: 23028060]
32. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M. *The journal of physical chemistry B.* 1998; 102:3586–3616.
33. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, MacKerell AD. *J Comput Chem.* 2010; 31:671–690. [PubMed: 19575467]
34. Mesleh MF, Hunter JM, Shvartsburg AA, Schatz GC, Jarrold MF. *J Phys Chem.* 1996; 100:16082–16086.
35. Wyttenbach T, Witt M, Bowers MT. *J Am Chem Soc.* 2000; 122:3458–3464.
36. Vonhelden G, Hsu MT, Kemper PR, Bowers MT. *J Chem Phys.* 1991; 95:3835–3837.
37. Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, Heck AJ. *Nature protocols.* 2009; 4:484–494.
38. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. *Proc Natl Acad Sci U S A.* 2004; 101:9528–9533. [PubMed: 15210983]
39. Cole SR, Ma X, Zhang X, Xia Y. *J Am Soc Mass Spectrom.* 2012; 23:310–320. [PubMed: 22161508]
40. Guan F, Uboh CE, Soma LR, Rudy J. *J Am Soc Mass Spectrom.* 2011; 22:718–730. [PubMed: 21472610]
41. Leymarie N, Costello CE, O'Connor PB. *J Am Chem Soc.* 2003; 125:8949–58. [PubMed: 12862492]
42. Li X, Huang Y, O'Connor PB, Lin C. *J Am Soc Mass Spectrom.* 2011; 22:245–254. [PubMed: 21472584]
43. Bush MF, Campuzano ID, Robinson CV. *Anal Chem.* 2012; 84:7124–7130. [PubMed: 22845859]
44. Pierson NA, Chen L, Valentine SJ, Russell DH, Clemmer DE. *J Am Chem Soc.* 2011; 133:13810–13813. [PubMed: 21830821]
45. McLean JR, McLean JA, Wu Z, Becker C, Perez LM, Pace CN, Scholtz JM, Russell DH. *The journal of physical chemistry B.* 2010; 114:809–816. [PubMed: 20000372]
46. Chen L, Shao Q, Gao YQ, Russell DH. *J Phys Chem A.* 2011; 115:4427–4435. [PubMed: 21476523]
47. Chirot F, Calvo F, Albrieux F, Lemoine J, Tsybin YO, Dugourd P. *J Am Soc Mass Spectrom.* 2012; 23:386–396. [PubMed: 22131226]
48. Paizs B, Suhai S. *Mass Spectrom Rev.* 2005; 24:508–548. [PubMed: 15389847]

49. Tirado M, Rutters J, Chen X, Yeung A, van Maarseveen J, Eyler JR, Berden G, Oomens J, Polfer NC. *J Am Soc Mass Spectrom.* 2012; 23:475–482. [PubMed: 22219043]
50. Tang XJ, Thibault P, Boyd RK. *Anal Chem.* 1993; 65:2824–2834. [PubMed: 7504416]
51. Chawner R, Gaskell SJ, Evers CE. *Rapid Commun Mass Spectrom.* 2012; 26:205–206. [PubMed: 22173809]
52. Taouatas N, Drugan MM, Heck AJ, Mohammed S. *Nat Methods.* 2008; 5:405–407. [PubMed: 18425140]
53. Rand KD, Pringle SD, Morris M, Engen JR, Brown JM. *J Am Soc Mass Spectrom.* 2011; 22:1784–1793. [PubMed: 21952892]
54. Fu Q, Li L. *Anal Chem.* 2005; 77:7783–7795. [PubMed: 16316189]

Figure 1.

Workflow of the developed strategy for gas-phase ion isomer analysis. On the left panel, under ETD fragmentation each b ion isomer generates certain diagnostic fragment ions and the FH-labeled linear b ion is used for alignment. On the right panel, each isomer is separated by IMS, followed by comparison with the drift time of the FH-labeled b ions. Although ETD fragments multiply charged ions, we show singly charged ions here for simplicity in illustration.

Figure 2.

ETD fragmentation spectra and IMS distributions of three b ions. (A,B) b_9^{2+} ion of neurokinin. (C,D) $FH\text{-}b_9^{2+}$ ion of neurokinin. (E,F) b_{10}^{2+} ion of substance P. (G,H) $FH\text{-}b_{10}^{2+}$ ion of substance P. (I,J) b_{33}^{4+} ion of CPRP. (K,L) $FH\text{-}b_{33}^{4+}$ ion of CPRP. The letter n denotes system noise. The details of ion assignment are listed in Tables S4–9.

Figure 3.

IMS dynamics study of neurokinin b_9 . (A) Workflow of the experimental design on ESI-QTOF-IMS. (B) Drift time distributions of b_9^{2+} under various collision energy values (CE, eV; bin, 0.069 ms). (C,D) CID spectra of b_9^{2+} from MS³-IMS-MS⁴ experiment. (E) Simulated drift time distributions of the four b_9^{2+} isomers (bin, 0.069 ms). Note that the integrated distributions from simulations are the same for the four isomers because 1500 structures are used for the simulation of each isomer. By contrast, the populations of the four isomers are different in experiment.

Figure 4.

MS² (CID) of CPRP peptide (A) and MS³ (CID) of b₃₃ ion (B) acquired on LTQ-FTICR. The 'b' ions were highlighted with red and assigned according to ion nomenclature reported in Ref. 51. The details of ion assignment are listed in Table S10. Note that the [b₃₃19]b₉~[b₃₃19]b₁₄ and [b₃₃8]b₂₅ ions could also be internal fragments produced by backbone cleavage of *ol*b₃₃.

Table 1

Summary of fragment ion isomer analysis of the three representative b ions.

The b ions under investigation ^a	Ion isomers	Ratio ^b
Neurokinin b₉²⁺ HKTDSEFVGL ²⁺	<i>c</i> b ₉ ²⁺	80%
	Cyclo-(HKTDSEFVGL) ²⁺	
	<i>ol</i> b ₉ ²⁺	16%
	HKTDSEFVGL ²⁺	
	<i>rl</i> b _{A9} ²⁺	— ^c
	TDSFVGLHK ²⁺	
Substance P b₁₀²⁺ RKPQFFGL ²⁺	<i>rl</i> b ₁₀ ²⁺	100%
	PQFFGLRKP ²⁺	
CPRP b₃₃⁴⁺ RSAEGLGRMGRLASLKS DTVTPLRGFEGETGH ⁴⁺	<i>ol</i> b ₃₃ ⁴⁺	100%
	RSAEGLGRMGRLASLKS	
	DTVTPLRGFEGETGH ⁴⁺	

^a The sequences of original linear structures are listed.^b The ratios are obtained from IMS distributions in Figure 2B, F and J, respectively.^c The *rl*b_{A9}²⁺ ion is not fully resolved between *c*b₉²⁺ and *ol*b₉²⁺ (Figure 2B). Therefore, the ratios of *c*b₉²⁺ and *ol*b₉²⁺ include small amount of *rl*b_{A9}²⁺.