


RESEARCH ARTICLE

Analyzing Internal Fragmentation of Electrosprayed Ubiquitin Ions During Beam-Type Collisional Dissociation

Kenneth R. Durbin, Owen S. Skinner, Ryan T. Fellers, Neil L. Kelleher

Departments of Chemistry and Molecular Biosciences, Northwestern University, 2170 Campus Drive, Evanston, IL 60208, USA

M Q I F V K T L T G K T
I T L E V E P S D T I E
N V K A K I Q D K E G I
P P D Q Q R L I F A G K
Q L E D G R T L S D Y N
I Q K E S T L H L V L R
L R G G

of internal fragment ions
covering residue
1  23

Abstract. Gaseous fragmentation of intact proteins is multifaceted and can be unpredictable by current theories in the field. Contributing to the complexity is the multitude of precursor ion states and fragmentation channels. Terminal fragment ions can be re-fragmented, yielding product ions containing neither terminus, termed internal fragment ions. In an effort to better understand and capitalize upon this fragmentation process, we collisionally dissociated the high (13+), middle (10+), and low (7+) charge states of electrosprayed ubiquitin ions. Both terminal and internal fragmentation processes were quantified through step-wise increases of voltage potential in the collision cell. An isotope fitting algorithm matched observed product ions to theoretical terminal and internal fragment ions. At optimal energies for internal

fragmentation of the 10+, nearly 200 internal fragments were observed; on average each of the 76 residues in ubiquitin was covered by 24.1 internal fragments. A pertinent finding was that formation of internal ions occurs at similar energy thresholds as terminal *b*- and *y*-ion types in beam-type activation. This large amount of internal fragmentation is frequently overlooked during top-down mass spectrometry. As such, we present several new approaches to visualize internal fragments through modified graphical fragment maps. With the presented advances of internal fragment ion accounting and visualization, the total percentage of matched fragment ions increased from approximately 40% to over 75% in a typical beam-type MS/MS spectrum. These sequence coverage improvements offer greater characterization potential for whole proteins with no needed experimental changes and could be of large benefit for future high-throughput intact protein analysis.

Keywords: Top-down mass spectrometry, Proteomics, Internal fragmentation

Received: 28 October 2014/Revised: 31 December 2014/Accepted: 31 December 2014/Published online: 26 February 2015

Introduction

Protein fragmentation by top-down tandem mass spectrometry produces enough data that automated data analysis is necessary [1, 2]. Potentially hundreds of product ions can result from a single ionic species of a whole protein [3]. Although different computational tools exist to aid in the analysis of protein fragmentation data, they are primarily designed to analyze protein cleavage products that contain either the amino- or carboxy-terminus or search for differences in fragment ion masses that correspond to amino acid residues [2, 4].

However, for internal fragment ions, there is a lack of available analysis software because of the computational demands of the increased search space. Owing to these limitations and the perception of ‘over-fragmentation,’ internal fragment ions have generally been avoided by practitioners of top-down mass spectrometry. Further, experimental conditions have even been sought to minimize their formation, potentially decreasing protein characterization [5]. Additionally, the value of identifying internal fragment ions, as well as how to visualize and interpret the results, is less clear than for terminal fragment ions.

Previous studies detailing internal fragmentation have been mainly limited to peptides [6–8]. For intact proteins, ubiquitin is often used as a model, and has been studied extensively to determine the influence of multiple charging on ion structure and protein dissociation [9]. Agar et al. explored the internal fragmentation of ubiquitin with declustering potential (i.e., source-induced dissociation) and at high energies found that nearly half of all product ions were internal fragments [10].

Kenneth R. Durbin and Owen S. Skinner contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s13361-015-1078-1) contains supplementary material, which is available to authorized users.

Correspondence to: Neil Kelleher; e-mail: n-kelleher@northwestern.edu

However, common high-throughput top-down proteomics methodology typically targets individual charge states for fragmentation instead of dissociating all electrosprayed ions with declustering potential in the electrospray source. Therefore, a quantitative study of the effect of fragmentation energy on individual charge states is necessary to provide insight into internal fragmentation for top-down experiments. Here, we present such a study with detailed internal and terminal fragment analysis with visualization of axial acceleration higher-energy collisional dissociation (HCD) product ions from the 13+, 10+, and 7+ precursor charge states of ubiquitin ions. The study strongly motivates the further consideration of internal ions in beam-type fragmentation, possibly important for high mass proteins where localization of variation within the middle region of sequence presents a major challenge for high-throughput tandem MS using any ion fragmentation technology.

Experimental

A solution of 2 μM bovine ubiquitin (Sigma, St. Louis, MO, USA) was analyzed by ESI-MS under denaturing solution conditions (50:50:1 MeOH/H₂O/HCOOH, pH \sim 3) for the spectra of 13+ and 10+ ions, and native solution conditions (100:1 H₂O/HCOOH, pH \sim 3) for spectra deriving from 7+ ions. Although ubiquitin is known to remain in its native form at the low pH of our native electrospray solution, many proteins are denatured under these same conditions [11]. All samples were run on a Q Exactive HF mass spectrometer (Thermo Scientific, Bremen, Germany). Precursor ions were quadrupole-isolated with a 5 m/z window. Fragmentation spectra were collected at 120,000 resolving power @ 200 m/z with 4 μ scans and a fixed injection time of 200 ms (7+), 100 ms (13+), and 50 ms (10+). Different injection times were used to avoid space charge effects as a result of different initial precursor abundances. The scan window ranged from 133 to 2000 m/z .

The HCD MS/MS spectra were analyzed with an in-house algorithm that generates the masses of all possible theoretical terminal and internal *b*- and *y*-ions from the known protein sequence. Only fragment ions that contained at least two residues were considered. Sets of unique fragments with the same chemical formula were only counted once for overall yield and length calculations and were not included in fragment maps or coverage calculations. All possible charges of each fragment ion were then searched against each spectrum at 5 ppm (13+) and 3 ppm (10+ and 7+), with these different tolerances accounting for slight differences in space charge. For positively identified fragment ions, charge-normalized intensity values were measured and tabulated.

Results and Discussion

We isolated the ubiquitin 13+, 10+, and 7+ ions individually and fragmented each at varying HCD energies (Figure S1 of

the Supplementary Material). At low energy, the isolated charge states remain nearly intact (Supplementary Figure S1, 20 V on left panels), but as the applied collision energy was increased, more fragment ions were produced (Supplementary Figure S1, 35 V in middle panels). When collisional energy was further increased, a larger number of lower m/z fragment ions were observed (Supplementary Figure S1, 50 V on right panels). Qualitatively, the number of fragment ions was observed to be dependent on the amount of energy applied to the intact precursor ion.

Fragment ions can be divided into two classes: terminal ions containing the amino or carboxy terminus, and internal ions with neither. The intensity yields from both terminal and internal ions were determined for isolated ubiquitin precursor charge states of 7+, 10+, and 13+ at stepped collision energies as shown in Figure 1a–c. The number of total matched ions for both fragmentation types is displayed for the three charge states in Figure 1d–f. Initially, raising the voltage potential for all three precursors increases the number and abundances of terminal fragment ions (Figure 1, red). However, this increase is dependent on the precursor charge: the 13+ charge state fragments at lower energies than the 10+, which in turn begins to fragment before the 7+. This effect can be attributed to the energy imparted on the parent ions, which is directly proportional to the voltage multiplied by the charge, as well as other structural and Coulombic factors [11]. As HCD energy is further increased, internal fragment ions are generated. The number and yield of internal fragment ions (Figure 1, blue) increase along with a corresponding decrease of terminal fragment ions, but with only a few volts discriminating their maxima (Figure 1). Only internal fragment ions are observed at the highest energy levels. Unexpectedly, local maxima were observed in the number and yield of both fragment types, particularly for the 7+ and 10+ parent ions. This feature may be attributable to an increase in the number of ions after cleavage of larger ions. For example, a single large ion, when fragmented, produces two product ions. The distribution of charge on the precursors likely plays a role in these patterns as evidenced by the lack of maxima in the higher charged 13+ ion fragmentation data. We attribute both the corresponding minima in product yields and the overall decrease in fragment ion intensities at higher energies to experimental limitations in detection. As fragments are first formed and then re-fragmented into smaller pieces at higher energies, more low-mass and neutral product ions are formed, including ions that may be below the minimum scan value (133 m/z) or undetectable by MS, respectively. Further differences in the fragmentation outcomes of different precursor charge states can be seen by analyzing the average length of matched fragment ions, which decreases exponentially at higher energies (Supplementary Figure S2).

Although the overall yields of fragment ions are important for studying mechanisms of protein fragmentation, the confidence in protein identification and especially characterization is dependent on the *number* of matched versus unmatched fragment ions. At 32 V (25 normalized collision energy, or NCE,

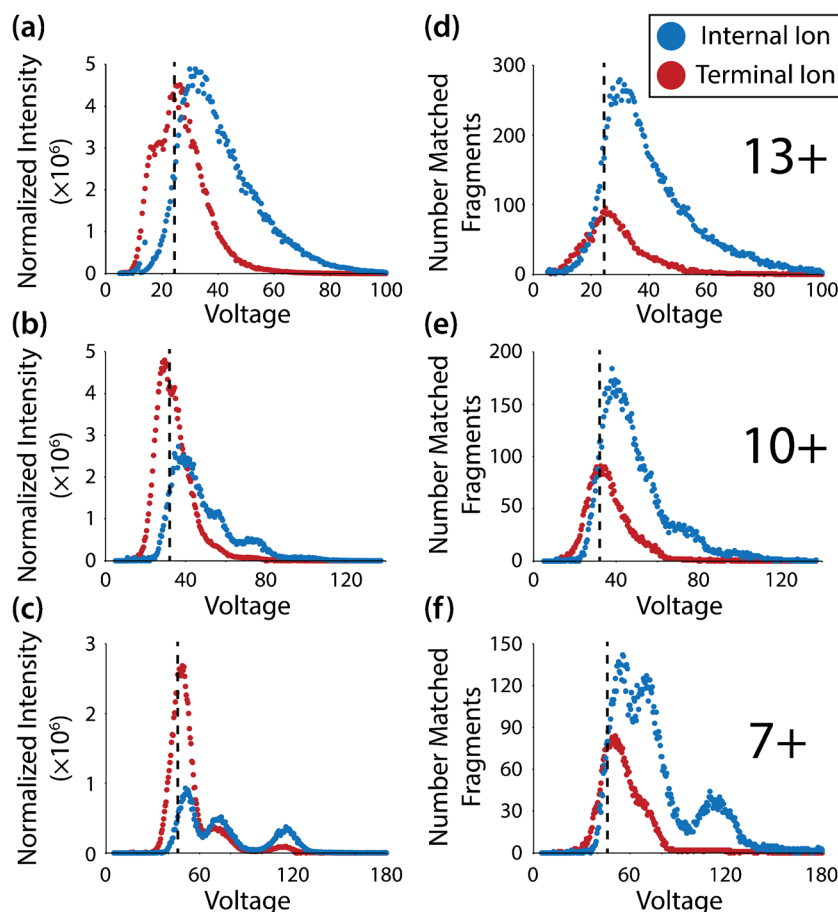


Figure 1. Differential charge state fragmentation as a function of applied voltage. Stepped HCD energies were used to fragment the 7+ (bottom), 10+ (middle), and 13+ (top) charge states of ubiquitin. Fragment ions were found by matching experimental isotope distributions to theoretical isotope distributions. **(a–c)** Intensity values for terminal and internal fragment ions were normalized by charge. Terminal fragment ions are denoted by red circles whereas internal fragment ions are signified by blue circles. **(d–f)** The number of internal and terminal matching fragment ions is shown for each of the three charge states. The dashed line present in each graph indicates the voltage that corresponds to the typical energy used in top-down proteomics analysis

the standard HCD energy setting for top-down proteomics), 100 terminal ions from the 10+ precursor ion were matched, including the 57 matched internal fragment ions results in a total of 157 matching fragments. For the 13+ precursor ion at 25 NCE (24.5 V), 95 terminal and 78 internal fragments ions were found. If the internal ions were used for proteoform characterization [12], the 173 total matching ions would represent an 82% increase in the overall number of matched fragments. For the 7+ charge state at 25 NCE (46 V), 92 terminal and 41 internal fragments were matched, a 45% improvement in total matching fragments. If energies yielding more internal fragment ions are considered, the increase in matched fragments could be even greater. At 70 V for the 7+ precursor ion, only 39 terminal ions are matched, while 105 internal ions are produced, a 269% increase in the total number of matched ions. To account for all ions in a spectrum, we manually validated one full fragmentation spectrum from the 10+ precursor at 30 V. Of the 201 matched fragment ions, matching terminal fragment ions accounted for 88 ions, or 43.7% of the total. The inclusion of internal ions led to a gain of 64 matching ions,

raising the total number of accounted ions to 75.6%. Many of the unaccounted ions are from neutral losses. The above increases in total matching fragment ions without the need for additional fragmentation events would be beneficial for improved characterization and validation during high-throughput proteomic experiments. Furthermore, internal fragment ions could help localization of proteoform features and may ultimately benefit proteoform characterization metrics, such as the C-score [13].

An example of the local effects of fragmentation can be seen by examining the intensity of several 10+ products, y_{58} and y_{40} ions (terminal fragments) and the $y_{Ib}[18-36]$ (an internal fragment depicted by nomenclature described previously [14]) across a variety of collisional energies (Figure 2a). These three fragment ions are formed from cleavages at sites 18 and 36 (Figure 2b), each directly N-terminal to two of the three prolines in ubiquitin. Cleavage sites N-terminal to proline have been shown to have high fragmentation propensity when collisionally dissociated [15]. The formation of the y_{58} fragment (a predominant product ion of ubiquitin) occurs at much lower

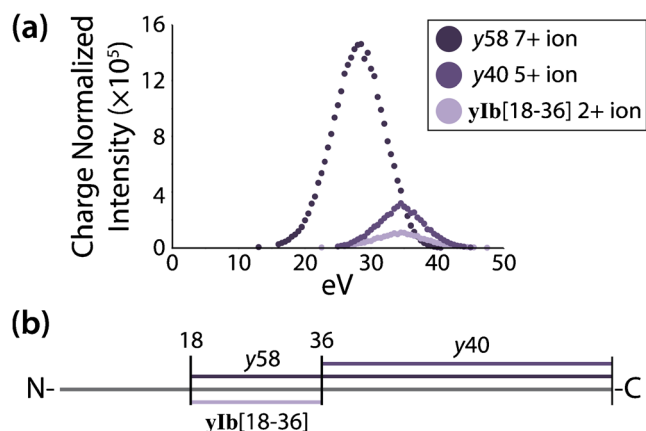


Figure 2. Individual fragment ion traces from 10+ charge state of ubiquitin. The 7+ charge state (dark) of the most abundant fragment ion from HCD fragmentation, $y58$, is plotted as a function of normalized charge intensity versus the HCD energy applied. The 5+ fragment ion (medium) from a less abundant terminal ion, $y40$, is in the medium purple color. Lastly, a 2+ ion of an internal fragment ion (light) spanning residues 19 through 36 (cleavages after residues 18 and 36) is shown

energy than the other two ions, indicating that it is among the first product ions to form upon activation. At higher energies, the abundance of $y58$ decreases while the intensities of $y40$ and $yIb[18-36]$ increase. These trends are consistent with the re-fragmentation of $y58$ into, among others, the complementary products $y40$ and $yIb[18-36]$ (Figure 2b). Thus, the new formation of both smaller terminal and internal fragment ions at higher collisional energies can be attributed to the secondary fragmentation of larger product ions.

Although the inclusion of internal fragment ions can greatly increase top-down protein characterization, the sequence space covered by all possible internals is sizable. Regarding internal ion formation from ubiquitin, 28,490 different ions between the fragment charge states 1+ through 10+ are theoretically possible for a 76-residue protein such as ubiquitin. By comparison, ubiquitin has 1480 theoretical terminal fragment ions when the same charge state range is considered. As protein size increases, the theoretical number of internal fragments increases quickly. For a protein with 500 residues, there are 1.2×10^6 possible ions even when considering only charge states at or below 10+. Only 9960 of that total can be attributed to terminal fragment ions. The number of theoretical terminal and internal fragment ions (with no charge states considered) for proteins up to 1000 residues in length are graphed on a \log_{10} scale in Supplementary Figure S3. The total number of theoretical internal fragments was determined by:

$$\frac{((n-l-1) + (n-l-1)^2)}{2} \quad (1)$$

where n is the length of the protein and l is the minimum fragment length considered. To determine the probability of spuriously matching a fragment ion in this study, we estimated our false discovery rate for the experimental observation of fragment ions. Spectral m/z values from the 10+ precursor were shifted by set amounts between -500 and $+500$ ppm. Using the altered data as a null set (where the experimental mass accuracy is far outside the expected values), we found an average of 1.4% of the number of fragments matched in the original data set (Supplementary Figure S4, blue). Further, the false positive fragment ion matches made up only 0.8% of the overall original intensity (Supplementary Figure S4, red).

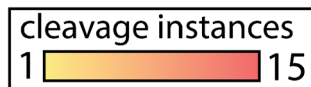
(a) Terminal Ion Graphical Fragment Map

```

N M Q I F V K T L T G K T I T L E V E P S D T I E N V 26
27 K A K I Q D K E G I P P D Q Q R L I F A G K Q L E D 52
53 G R T L S D Y N I Q K E S T L H L V L R L R G G C

```

(b) Cleavage



```

N M Q I F V K T L T G K T I T L E V E P S D T I E N V 26
27 K A K I Q D K E G I P P D Q Q R L I F A G K Q L E D 52
53 G R T L S D Y N I Q K E S T L H L V L R L R G G C

```

(c) Coverage



```

N M Q I F V K T L T G K T I T L E V E P S D T I E N V 26
27 K A K I Q D K E G I P P D Q Q R L I F A G K Q L E D 52
53 G R T L S D Y N I Q K E S T L H L V L R L R G G C

```

Figure 3. Visualization of fragmentation patterns after 30 eV HCD fragmentation of the 10+ charge state of ubiquitin. (a) Matched terminal fragment ions are mapped in the traditional graphical fragment format. (b) A heat map depicts internal fragmentation propensity based on color. Darker color (i.e., red) represents the most frequent cleavage sites, with 15 instances at the top of the range. (c) A coverage map depicts the number of times each residue is covered by internal fragment ions. As above, darker equates to higher coverage. The highest value represented is a residue covered by 23 different internal fragment ions

Although internal fragment ions provide a wealth of data, to our knowledge no straightforward method to visualize the internal fragment landscape has been devised. Terminal fragment ions from bottom-up and top-down studies are traditionally depicted by graphical fragment maps that indicate cleavage site and fragment directionality. A fragment map for the 10+ ubiquitin charge state after HCD of 30 V is shown in Figure 3a. Unlike terminal ions, uniquely defining internal fragments require two internal cleavage sites, contributing an additional dimension to the visualization process. Previous efforts to indicate internal fragments have used vertical bars [10]. A similar method, using horizontal bars to indicate both terminal and internal fragment ions, is shown for ubiquitin in Supplementary Figure S5. However, this method can be difficult to scale up with longer protein sequences and larger numbers of observed internal fragments. We therefore propose two compact, easily displayable internal fragment maps that utilize a heat map to indicate the number of internal fragmentation events occurring at each cleavage site (Figure 3b), and the number of fragment ions that cover each residue in the protein (Figure 3c). The ‘cleavage’ heat map is used to observe ‘hotspots’ that cleave more readily (such as between E18 and P19). In certain situations, cleavage propensities may also depend on higher-order gaseous protein structure. The ‘coverage’ heat map indicates the number of ions that provide information on each residue, facilitating the localization of post-translational modifications and single nucleotide polymorphisms that are crucial for characterizing proteoforms. In addition, a histogram is included to show average residue coverage over different fragmentation energies applied to the 10+ charge state (Supplementary Figure S6). Only considering internal fragments, the peak coverage is on average 24.1 fragment ions covering each residue. Overall, the highest coverage averaged a combined 64 terminal and internal fragments per residue at 34.5 V. The coverage and cleavage graphical fragment maps provide a clear approach towards visualization of all fragmentation data.

Conclusion

Internal fragment ions, often formed from terminal fragments within a narrow range of HCD acceleration, compose a large percentage of the total number of product ions produced by threshold dissociation of whole proteins. In certain instances, the number of matching fragment ions can more than double when internal ions are included in the data analysis. Therefore, the inclusion of internal fragments in top-down data analysis could substantially increase proteoform identification and characterization. More specifically, the additional characterization supplied by internal fragment ions has the potential to precisely localize modifications retained at the high collision energies that yield internal fragment ions, particularly on the middle

sequence regions of larger proteins where high sequence coverage by terminal fragment ions is difficult to achieve. In this work, we present a methodology to analyze and visualize internal fragments, and present some of the strengths and weaknesses of their inclusion. In future top-down mass spectrometry applications, including high-throughput proteomic experiments, the ability to analyze internal fragments may provide critical insight towards validating protein identifications and fully characterizing proteoforms.

Acknowledgments

The authors thank Bryan Early for helpful discussions and Phil Compton for technical assistance. O.S.S. acknowledges the NSF GRFP under fellowship #2014171659. The authors also acknowledge the generous support of the W. M. Keck Foundation, Northwestern University, and an NIH R01 (GM 067193) to N.L.K.

References

- Durbin, K.R., Fellers, R.T., Ntai, L., Kelleher, N.L., Compton, P.D.: Autopilot: an online data acquisition control system for the enhanced high-throughput characterization of intact proteins. *Anal. Chem.* **86**, 1485–92 (2014)
- Zamdborg, L., LeDuc, R.D., Glowacz, K.J., Kim, Y.-B., Viswanathan, V., Spaulding, I.T., Early, B.P., Bluhm, E.J., Babai, S., Kelleher, N.L.: ProSight PTM 2.0: improved protein identification and characterization for top down mass spectrometry. *Nucleic Acids Res.* **35**, W701–6 (2007)
- Shaw, J.B., Li, W., Holden, D.D., Zhang, Y., Griep-Raming, J., Fellers, R.T., Early, B.P., Thomas, P.M., Kelleher, N.L., Brodbelt, J.S.: Complete protein characterization using top-down mass spectrometry and ultraviolet photodissociation. *J. Am. Chem. Soc.* **135**, 12646–51 (2013)
- Karabacak, N.M., Li, L., Tiwari, A., Hayward, L.J., Hong, P., Easterling, M.L., Agar, J.N.: Sensitive and specific identification of wild type and variant proteins from 8 to 669 kDa using top-down mass spectrometry. *Mol. Cell Proteomics* **8**, 846–56 (2009)
- Breuker, K., Jin, M., Han, X., Jiang, H., McLafferty, F.W.: Top-down identification and characterization of biomolecules by mass spectrometry. *J. Am. Soc. Mass Spectrom.* **19**, 1045–53 (2008)
- Ballard, K.D., Gaskell, S.J.: Sequential mass-spectrometry applied to the study of the formation of internal fragment ions of protonated peptides. *Int. J. Mass Spectrom. Ion Process* **111**, 173–189 (1991)
- Khatun, J., Ramkissoon, K., Giddings, M.C.: Fragmentation characteristics of collision-induced dissociation in MALDI TOF/TOF mass spectrometry. *Anal. Chem.* **79**, 3032–40 (2007)
- Barran, P.E., Polfer, N.C., Campopiano, D.J., Clarke, D.J., Langridge-Smith, P.R.R., Langley, R.J., Govan, J.R.W., Maxwell, A., Dorin, J.R., Millar, R.P., Bowers, M.T.: Is it biologically relevant to measure the structures of small peptides in the gas phase? *Int. J. Mass Spectrom.* **240**, 273–84 (2005)
- Skinner, O.S., McLafferty, F.W., Breuker, K.: How ubiquitin unfolds after transfer into the gas phase. *J. Am. Soc. Mass Spectrom.* **23**, 1011–4 (2012)
- Cobb, J.S., Easterling, M.L., Agar, J.N.: Structural characterization of intact proteins is enhanced by prevalent fragmentation pathways rarely observed for peptides. *J. Am. Soc. Mass Spectrom.* **21**, 949–59 (2010)
- Price, W.D., Williams, E.R.: Activation of peptide ions by blackbody radiation: factors that lead to dissociation kinetics in the rapid energy exchange limit. *J. Phys. Chem. A* **101**, 8844–52 (1997)
- Smith, L.M., Kelleher, N.L.: The Consortium for Top Down Proteomics: Proteoform: a single term describing protein complexity. *Nat. Methods* **10**, 186–7 (2013)
- LeDuc, R.D., Fellers, R.T., Early, B.P., Greer, J.B., Thomas, P.M., Kelleher, N.L.: The C-score: a Bayesian framework to sharply improve

- proteoform scoring in high-throughput top down proteomics. *J. Proteome Res.* **13**, 3231–40 (2014)
14. Harvey, S.R., Porrini, M., Konijnenberg, A., Clarke, D.J., Tyler, R.C., Langridge-Smith, P.R.R., MacPhee, C.E., Volkman, B.F., Barran, P.E.: Dissecting the dynamic conformations of the metamorphic protein lymphotactin. *J. Phys. Chem. B* **118**, 12348–59 (2014)
15. Kruger, N.A., Zubarev, R.A., Carpenter, B.K., Kelleher, N.L., Hom, D.M., McLafferty, F.W.: Electron capture versus energetic dissociation of protein ions. *Int. J. Mass Spectrom.* **182**, 1–5 (1999)