Vol. 12, No. 2, 2021
ISSN 2233-4203/ e-ISSN 2093-8950
ARTICLE www.msletters.org | Mass Spectrometry Letters

Statistical Characterization of the Multi-Charged Fragment Ions in the CID and HCD Spectrum

Sangeetha Ramachandran[1]* and Tessamma Thomas[2]

[1]*Department of Electronics, Cochin University of Science and Technology, Cochin-682022, Kerala, India*
[2]*Department of Electronics, Cochin University of Science and Technology, Cochin-682022, Kerala, India*

Received in March 7, 2021; Revised May 24, 2021; Accepted May 27, 2021
First published on the web June 30, 2021; DOI: 10.5478/MSL.2021.12.2.41

Abstract : Collision-induced dissociation (CID) and higher-energy collisional dissociation (HCD) are the widely used fragmentation technique in mass spectrometry-based proteomics studies. Understanding the fragmentation pattern from the tandem mass spectra using statistical methods helps to implement efficient spectrum analysis algorithms. The study characterizes the frequency of occurrence of multi-charged fragment ions and their neutral loss events of doubly and triply charged peptides in the CID and HCD spectrum. The dependency of the length of the fragment ion on the occurrence of multi-charged fragment ion is characterized here. Study shows that the singly charged fragment ions are generally dominated in the doubly charged peptide spectrum. However, as the length of the product ion increases, the frequency of occurrence of charge 2 fragment ions increases. The y- ions have more tendencies to generate charge 2 fragment ions than b- ions, both in CID and HCD spectrum. The frequency of occurrence of charge 2 fragment ion peaks is prominent upon the dissociation of the triply charged peptides. For triply charged peptides, product ion of higher length occurred in multiple charge states in CID spectrum. The neutral loss peaks mostly exist in charge 2 states in the triply charged peptide spectrum. The b-ions peaks are observed in much less frequency than y-ions in HCD spectrum as the length of the fragment increases. Isotopic peaks are occurred in charge 2 state both in doubly and triply charged peptide's HCD spectrum.

Keywords : CID, HCD, fragmentation pattern, multi-charged ions, LC-MS/MS

Introduction

Tandem mass spectrometry (MS/MS) is the most widely used technology to characterize the peptides in the complex sample.[1] The proteins are digested into peptides and analyzed in MS1. The parent peptide ion isolated from MS1 is dissociated and analyzed in MS2. Different dissociation technique creates different patterns of fragmentation spectra. HCD and CID fragment the peptide at its amide bond along its peptide backbone generates predominantly N-terminal b-ions and C-terminal y-ions. The b- & y-type fragment ions can further fragment by losing small neutral molecules, creates neutral loss b- & y-ion peaks in the spectrum. The difference in collission energy in CID and HCD dissociation causes changes in their fragmentation spectra.

The basic theory of peptide fragmentation was first explained by Wysocki et al. in their mobile proton hypothesis.[2] The parent ion proton moves along the amide bond and cleaves its backbone at the most favorable point. Kapp et al.[3] proved that the mobility of charge is hindered by the basic residues within the peptide. The different fragmentation pathways are extensively investigated in the literature. The researchers have studied the fragmentation patterns using statistical analysis of the large set of data. The data mining and machine learning approaches were used to analyze the spectral pattern, which proves the residue-specific cleavage preferences of CID fragmentation pattern.[4–8] The statistical analysis of neutral loss events also proved the influence of residues in the fragmentation pattern. Most of the works are focused on the intensity information from the mass spectrum. Shao et al.[9] characterised the HCD and CID spectrum and found that the intensity of y- ions reached a maximum in the 60–70% and 40–50% relative mass bins of HCD spectra from doubly and triply charged peptides, respectively. The multi-charged fragment ions were analyzed based on the relative mass bins. Zubarev et.al[9] compared CID cleavage selectivity to ECD (electron capture dissociation). They found that the yn-2 (n refers to peptide length) fragment had the highest intensity of all y ions.

*Reprint requests to Sangeetha Ramachandran,
orcid.org/0000-0002-8033-294X
E-mail: sangeetharamachandran.25@gmail.com

Multi charged peptide dissociates into singly charged fragment ions, or the entire charge may retain on one fragment. The fragmentation of the doubly charged peptide generates singly and doubly charged fragment ions. Similarly, the fragmentation of triply charged peptides generates singly, doubly, and triply charged fragment ions. The x-axis of the tandem mass spectra has the mass/charge (*m/z*) value of fragment ion. The *m/z* of the doubly charged fragment ion will be nearly half of that of its singly charged fragment ion. Similarly mass of triply charged fragment ion will be nearly 1/3 of its singly charged fragment ion. Hence the generation of multi-charged fragment ions creates the more complex pattern in the tandem mass spectrum. For example, the CID spectrum of the doubly charged peptide 'DKGEAENEAKPIDVK' is shown in the Figure 1. The singly and doubly charged fragment ion y10 (y10+ & y10++) are appeared at *m/z* 251.11 & 571.80 respectively in CID spectrum. Not all fragment ions appeared in both charge states. Hence it is necessary to characterize the occurrence of multi charged fragment ions in the tandem mass spectrum for the reliable interpretation of the spectrum.

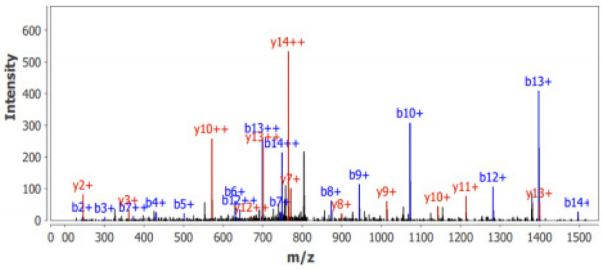

Figure 1. CID spectrum of the peptide 'DKGEAENEAKPIDVK'.

In this study, the frequency of occurrence of multi-charged fragment ions of doubly and triply charged peptides in HCD & CID spectra is analyzed. The dependency of the number of residues in the fragment is characterized here.

Method

Dataset

The CID spectrum of doubly and triply charged peptides were downloaded from the NIST human peptide spectral library (chemdata.nist.gov).[10] The spectral library provides high-quality reference consensus spectrum annotated with peptide for mass spectrometry-based proteomics studies. Dataset consists of 87661 and 29153 mass spectra of doubly & triply charged tryptic peptides, respectively, with no missed cleavage, and modifications are used for this study.

The HCD spectra of doubly and triply charged peptides were downloaded from ProteomeTool project spectral library.[11] It contains high-quality reference spectra of synthetic peptides of human proteome analyzed on LTQ Orbitrap fusion. Dataset consists of 188514 and 64087 mass spectra of doubly & triply charged tryptic peptides, respectively, with no missed cleavage and modifications are used for this study.

Methodology for analyzing multi-charged fragment ion peaks pattern

The relative frequency information was used to characterize the fragment ions and the neutral loss events in our previous studies.[12,13] Here also, the frequency of occurrence of fragment ion peaks is estimated to characterize the multi-charged fragment ion peaks in CID and HCD spectrum. The frequency statistics are calculated with respect to the length of the fragment ion or the position of the cleavage site along the length of the peptide. The maximum length of peptide considered in this study is 21. Hence the length of the fragment ion can vary from 1-20. The dataset contains large set of spectra of known sequences. From the peptide list, the possible number of times a particular ion b/y ion can be generated upon a cleavage position along the length of the peptide is estimated and is denoted as N^t_n. Here 't' denote fragment ion type b-/y- ion, 'n' denotes the number of residues in the fragment ion. From the CID & HCD spectra, the actual number of times an ion observed in the spectra is estimated and denoted as $C^t_{p,n}$. The frequency of occurrence of a fragment ion with respect to the length of the fragment is calculated using equation1.

$$F^t_p(n) = \frac{C^t_{p,n}}{N^t_n} \quad (1)$$

Where, p € { bn+, bn++, yn+, yn++ in case of doubly charged peptide spectrum
bn+, bn++, bn+++, yn+, yn++, yn+++ in case of triply charged peptide spectrum}

Here 'p' represents the singly and multi charged b- & y- ions comprising of the allied peaks of neutral losses and isotopic peaks. The study focused on the analysis of b- & y- ion peaks corresponding to the loss of nominal mass units: b/y- 17 (NH_3), 18 (H_2O), 34 (NH_3+NH_3), 35 (NH_3+ H_2O), 36 (H_2O + H_2O).

Results and Discussion

In our previous studies, we characterized the fragment ion peaks and their neutral loss peaks for the comprehensive understanding of the fragmentation pattern. The studies have proved the influence of residue-specific and position-specific cleavage preferences of CID fragmentation pattern.[12,13] The statistical features extracted were found to be consistent with the log-transformed intensity of the experimental spectrum. Hence, the features were efficiently

used to model the CID spectra and to identify the peptide from the CID spectra.[14,15] In this study, the frequency of occurrence of multi-charged fragment ions on CID and HCD spectra is characterized and compared. Shao et al. proved that the percentage of multi-charged fragment ions gradually increases with an increase in the relative mass of fragment.[9] The relative mass is the fragment ion mass divided by the precursor mass. In the triply charged peptide spectrum, the multi-charged fragment ions are prominent in the higher relative mass region. The dependency of the number of residues in the fragment ion is investigated in this study.

Characterization of multi-charged fragment ions in doubly charged peptide spectrum

The influence of the length of fragment ions for the generation of fragment ions in charge 2 states in the doubly charged peptide spectrum is characterized here. Initially, the fragment ions- b or y ions encompass all its allied peaks. The relative frequency of occurrence of charge one and two fragments of b- & y- ions plotted with respect to the length of fragment ion in CID and HCD spectrum are shown in Figure 2(a) and 2(b), respectively.

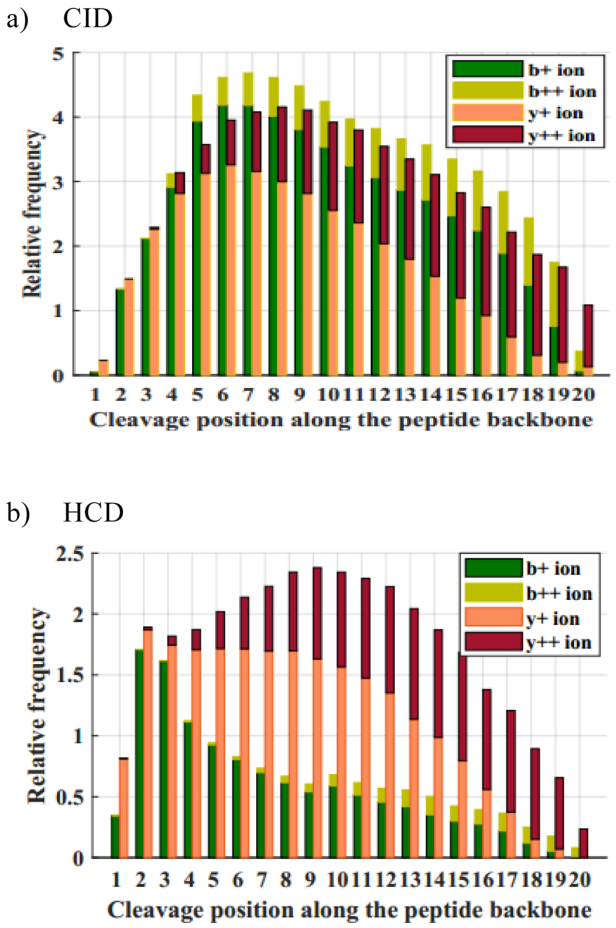

Figure 2. Charge state of fragment ions versus the length of the fragment. The relative frequency of occurrence of fragment ions b+, b++, y+, and y++ along the length of the peptide in CID and HCD spectrum is shown in Figure 2(a) and 2(b), respectively. The x-axis represents the length of the fragment ion. Y-axis represents the relative frequency of occurrence of fragment ion peaks.

Figure 2 shows that the occurrence of singly charged fragment ion peaks dominated in the doubly charged peptide's spectrum. The occurrence of charge 2 fragment ion peak increases as the length of the fragment ion increases. The y-ion has more tendencies to be in the y++ state. As the length of the fragment greater than 14, the frequency of occurrence of y++ ions are higher than the y+ ion. Compared to CID spectrum, HCD spectrum has the lesser frequency of occurrence of b-ion, and has very less b++ fragment ions. The frequency of occurrence of y-ions are higher in HCD spectrum as shown in Figure 2(b). Hence, it is inferred that, with the increase in length of the peptide, the frequency of occurrence of doubly charged fragment ion peaks in the CID & HCD spectrum increases. The b-ions peaks observed much less frequency than y-ion in HCD spectrum as the length of the fragment increases. This is because b-ions are less stable due to higher collision energy.[9]

Considering the neutral loss events and isotopic events in the CID & HCD spectrum of doubly charged peptides, the percentage of occurrence of charge 2 fragment ion of b- and y-type ion is shown in color map Figure 3(a) and 3(b), respectively. The color map emphasis how frequently the allied peaks of b-ion such as b-, b-isotopic, b-17, b-18, b-34, b-35, b-36, b-44, b-45, b-46, b+18 and y-ion such as y-, y-isotopic, y-17, y-18, y-34, y-35, y-36, y-44, y-45, y-46 occurred in charge 2 state as the length of fragment increases. The loss peak y-35 is not available in CID spectrum, hence not shown in colormap.

Figure 3 shows that as the length of the fragment increases, the neutral loss peaks are more frequently observed in charge 2 states. The loss of multiple neutral molecules, hence the generation of b-34, b-35, b-36, y-34, y-35, y-36 peaks more frequently observed in charge 2 state. In the case of y-ion, as the length of the fragment increases, the loss peaks of y-ions are frequently observed in charge 2 state. The isotopic peaks of b- & y-ions are frequently occurred in charge 1 state in CID spectrum, while in HCD spectrum, isotopic peaks mostly occur in charge 2 state in the doubly charged peptide's spectrum.

Characterization of multi-charged fragment ions in triply charged peptide spectrum

The influence of the length of the fragment for the generation of fragment ions in the charge 2 & 3 states in triply charged peptide spectrum is characterized here. The triply charged peptide undergoing fragmentation can generate singly, doubly, and triply charged fragment ions. The frequency of occurrence of b- & y- ions (encompassing its

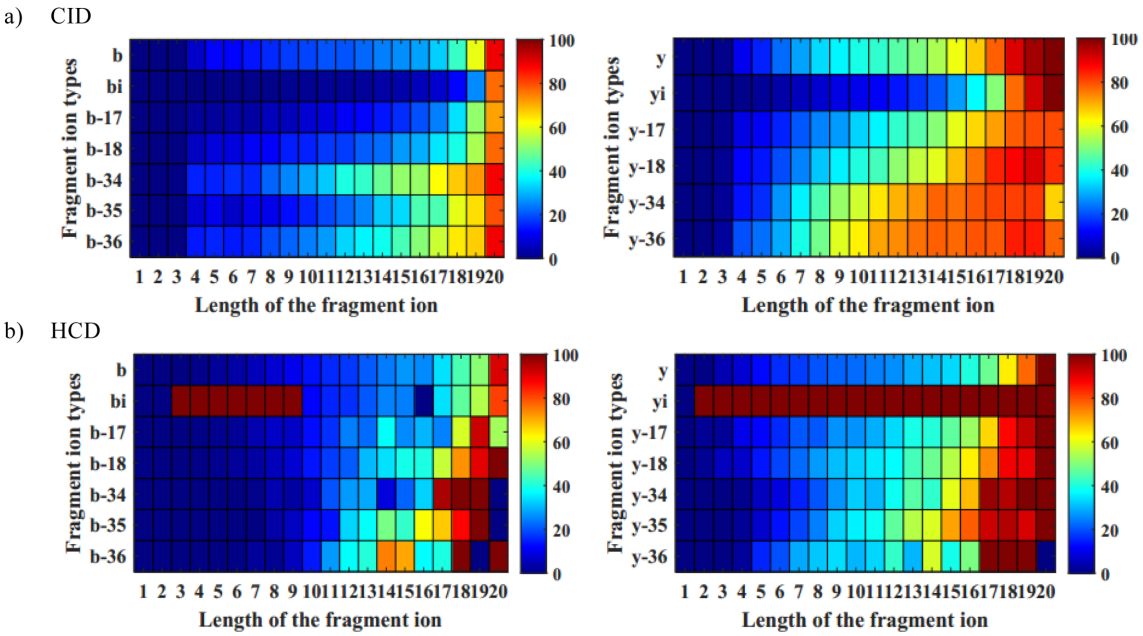

Figure 3. Percentage of occurrence of charge 2 state fragment ions among the fragment ion peaks. The x-axis represents the length of the fragment ion. Y-axis represents the allied peaks of b- & y-ions.

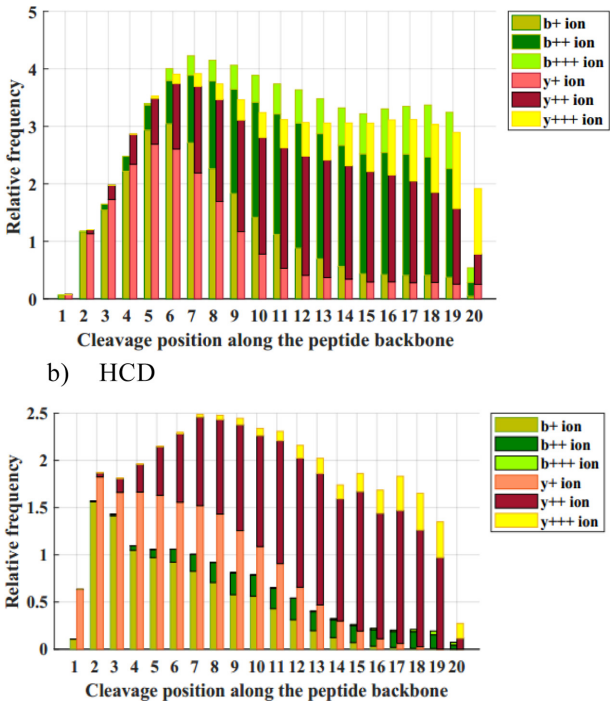

Figure 4. Charge state of fragment ions versus the length of the fragment. The figure shows the relative frequency of occurrence of fragment ions b+, b++, b+++, y+, y++, and y+++ along the length of the peptide. The x-axis represents the length of the fragment ion. Y-axis represents the relative frequency of occurrence of fragment ion peaks.

allied peaks) along the length of the peptide in charge 1, 2 & 3 states in CID and HCD spectrum are plotted in Figure 4(a) and 4(b) respectively.

From Figure 4(a) it is clear that, with the increase in the length of the fragment, the multi-charged fragment ion peaks are more frequently observed than singly charged fragment ion. Charge 2 fragment ion peaks are more prominent in the triply charged peptide spectrum. At higher length, the multiple charge state fragment ions are produced in the CID spectrum. In the HCD spectrum (Figure 4(b)), b-ions are observed in much less frequency as the length of the fragment ions increases. Singly charged b-ions are more frequently observed than the charge 2 state. Singly charged y-ions occurred in higher degree at lower length and y++ fragment ions dominated in the triply charged peptide's HCD spectrum as the length of the fragment ion increases.

Considering the neutral loss peaks and isotopic peaks in the CID & HCD spectrum of triply charged peptides, the percentage of occurrence of multi charged fragment ion of b- and y-type ion is plotted in the color map (Figure 5). Figure 5.i shows the percentage of occurrence of charge 2 b- & y- fragment ions and their allied peaks, respectively. Figure 5.ii shows the percentage of occurrence of charge 3 b- & y- fragment ions and their allied peaks, respectively.

The key results examined from the Figure 5 are discussed as follows. The frequency of occurrence of charge 2 fragment ion peaks is prominent upon the dissociation of triply charged peptides. The multi-charged neutral loss peaks increases as the length of the fragment ion increases. The neutral loss peaks of y-ions have more tendency to exist in

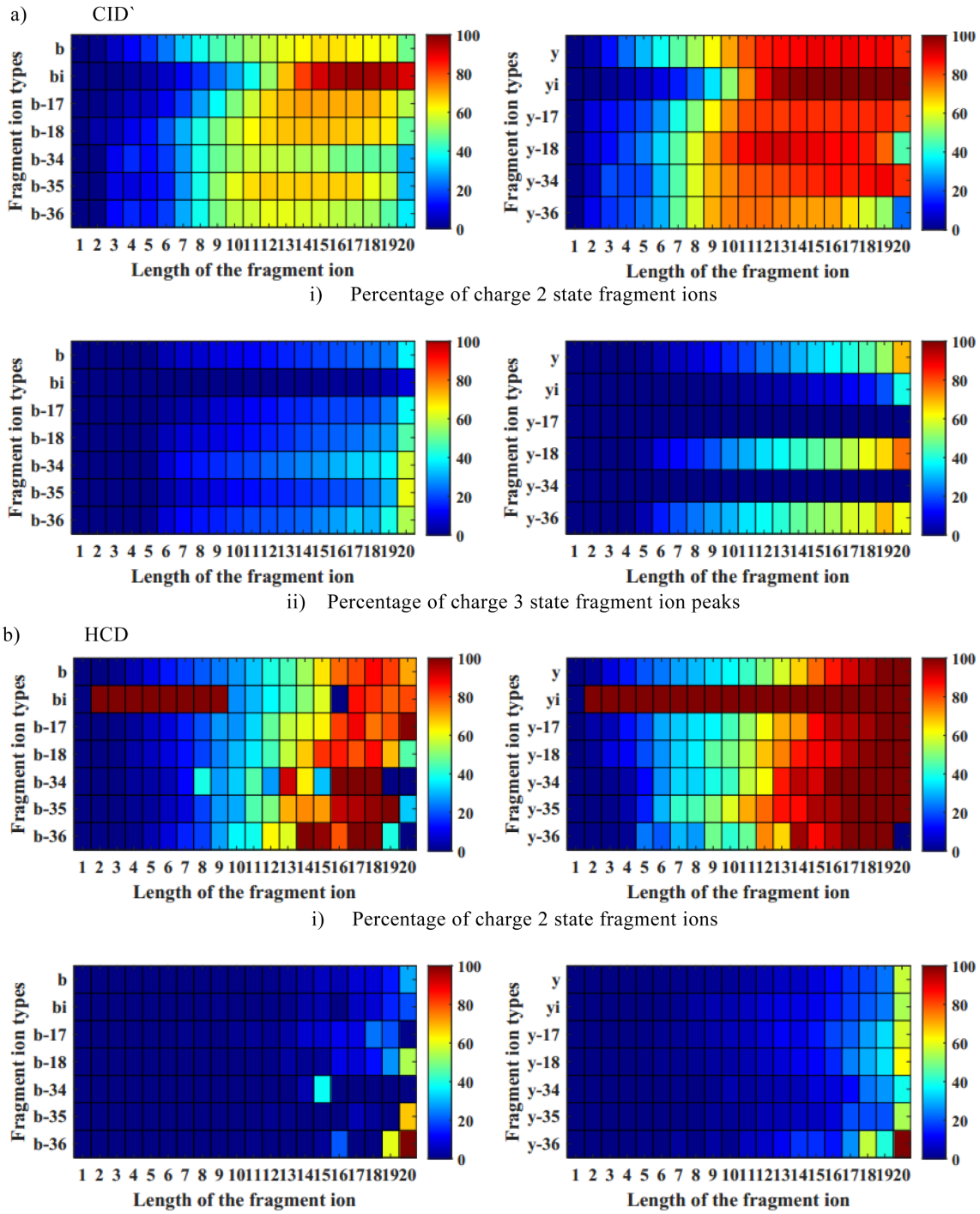

Figure 5. The percentage of occurrence of fragment ion peaks with charge 2 & 3 state. The x-axis represents the length of the fragment ion. Y-axis represents the allied peaks of b- & y-ions.

charge 2 state. The loss of water from y-ion (y-18 & y-36) has shown different patterns to exist in charge 3 state in CID spectra. Hence it can be inferred that the charge 3 fragment ion of y-ion has more tendency to lose water from its ions. Dissociation of a triply charged peptide is more likely to generate doubly charged isotopic peaks than the singly and triply charged isotopic peaks of b- & y- ions. In CID spectra, doubly charged isotopic peaks prominent in higher length fragment ion, while in HCD spectra, doubly charged isotopic peaks are prominent from lower length.

Conclusions

The occurrences of multi-charged fragment ions in the tandem mass spectrum add more complexity to the spectrum, hence the reliable interpretation of peptide sequence from the spectrum. The relative frequency of occurrence of multi-charged fragment ions in the doubly and triply charged peptide spectrum is studied here. Study shows that the singly charged fragment ions are generally dominated in the doubly charged peptide CID & HCD spectrum. However, as the length of the product ion increases, the frequency of occurrence of charge 2 fragment ions increases. The y- ions have more tendencies to generate charge 2 fragment ions than b- ions. In HCD spectrum, the relative frequency of occurrence of b-ions are much lesser than y-ion as the length of the fragment ion increases. The consecutive loss of ammonia and water from the y-ion occurs mostly from the doubly charged fragment ion y++. In the doubly charged peptide spectrum, isotopic peaks are more likely to occur in the singly charged state in CID spectrum and in the charge 2 state in HCD spectrum. For triply charged peptides, product ion of higher length occurred in multiple charge states. In both CID and HCD spectrum, isotopic peaks are more likely to occur in charge 2 state upon the dissociation of a triply charged peptide. Multi-charged neutral loss peaks gradually increases as the length of the fragment increases. Neutral loss from y-ions shows a higher tendency to exist in charge 2 state. In CID spectrum, the loss of water from y-ion (y-18 & y-36) has shown different patterns to exist in charge 3 state. Thus charge 3 fragment ion of y-ion has more tendency to lose water from its ions. The study provides the statistical trends of occurrence of fragment ion peaks in the mass spectrum, which helps in reliable interpretation of the peptide from the CID & HCD spectrum.

Acknowledgments

The authors would like to acknowledge the support from the Department of Electronics, Cochin University of Science and Technology.

References

1. Aebersold, R.; Mann, M. *Nature* **2003**, 422, 198, DOI: 10.1038/nature01511.
2. Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, 35, 1399, DOI: DOI: 10.1002/1096-9888(200012)35:12<1399::AID-JMS86>3.0.CO;2-R.
3. Kapp, E. A. Schütz, F.; Reid, G. E.; Eddes, J. S.; Moritz, R. L.; O'Hair, R. A. J.; Speed, T. P.; Simpson, R. J. *Anal. Chem.* **2000**, 75, 6251, DOI: 10.1021/ac034616t.
4. Tabb, D. L.; Smith, L. L.; Breci, L. A.; Wyscocki, V. H.; Lin, D.; Yates, J. R. *Anal. Chem.* **2003**, 75, 1155, DOI: 10.1021/ac026122m.
5. Huang, Y.; Triscari, J. M.; Tseng, G. C.; Pasa-Tolic, L,; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *Anal. Chem.* **2005**, **77**, 5800, DOI: 10.1021/ac0480949.
6. Huang, Y. Tseng, G. C.; Yuan, S.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *J. Proteome Res.* **2008**, 7, 70, DOI; 10.1021/pr070106u.
7. Loo, J. A.; Edmonds, C. G.; Smith, R. D. *Anal. Chem.* **1993**, 65, 425, DOI: 10.1021/ac00052a020.
8. Lau, K. W.; Hart, S. R.; Lynch, J. A.; Wong, S. C. C.; Hubbard, S. J.; Gaskell, S. J. *Rapid Commun. Mass Spectrom.* **2009**, 23, 1508, DOI: 10.1002/rcm.4032.
9. Shao, C.; Zhang, Y.; Sun, W. *J. Proteomics* **2014**, 109, 26, DOI: 10.1016/j.jprot.2014.06.012.
10. Stein, S. E.; Pudnick, P. A. *NIST Peptide Tandem Mass Spectral Libraries. Human Peptide Mass Spectral Reference Data*, National Institute of Standards and Technology: Gaithersburg, MD, **2008**, DOI: 10.18434/T4ZK5S.
11. Zolg, D. P.; Wilhelm, M.; Schnatbaum, K.; Zerweck, J.; Knaute, T.; Delanghe, D.; Bailey, D. J.; Gessulat, S.; Ehrlich, H.-C.; Weininger, M.; Yu, P.; Schlegl, J.; Kramer, K.; Schmidt, T.; Kusebauch, U.; Deutsch, E. W.; Aebersold, R.; Moritz, R. L.; Wenschuh, H.; Moehring, T.; Aiche, S.; Huhmer, A.; Reimer, U.; Kuster, B. *Nat. Methods* **2017**, 14, 259, DOI: 10.1038/nmeth.4153.
12. Ramachandran, S.; Thomas, T. **2019**, 10, 7, DOI: 10.5478/MSL.2019.10.2.50.
13. Ramachandran, S.; Thomas, T. *Int. J. Mass Spectrom.* **2020**, 448, 116270, DOI: 10.1016/j.ijms.2019.116270.
14. Ramachandran, S.; Thomas, T. A *ACS Omega* **2020**, 5, 12615, DOI: 10.1021/acsomega.9b03935.
15. Ramachandran, S.; Thomas, T. *J. Proteomics* **2021**, 235, 104116, DOI: 10.1016/j.jprot.2021.104116.

