© American Society for Mass Spectrometry (outside the USA), 2017

J. Am. Soc. Mass Spectrom. (2018) 29:463–469
DOI: 10.1007/s13361-017-1842-5

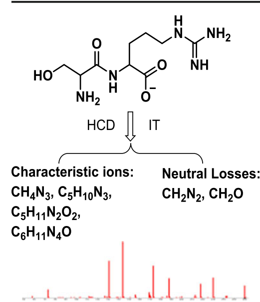
CrossMark

RESEARCH ARTICLE

# Collision-Induced Dissociation of Deprotonated Peptides. Relative Abundance of Side-Chain Neutral Losses, Residue-Specific Product Ions, and Comparison with Protonated Peptides

Yuxue Liang, Pedatsur Neta, Xiaoyu Yang, Stephen E. Stein

Mass Spectrometry Data Center, Biomolecular Measurement Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA




**Abstract.** High-accuracy MS/MS spectra of deprotonated ions of 390 dipeptides and 137 peptides with three to six residues are studied. Many amino acid residues undergo neutral losses from their side chains. The most abundant is the loss of acetaldehyde from threonine. The abundance of losses from the side chains of other amino acids is estimated relative to that of threonine. While some amino acids lose the whole side chain, others lose only part of it, and some exhibit two or more different losses. Side-chain neutral losses are less abundant in the spectra of protonated peptides, being significant mainly for methionine and arginine. In addition to the neutral losses, many amino acid residues in deprotonated peptides produce specific negative ions after peptide bond cleavage. An expanded list of fragment ions from protonated peptides is also presented and compared with those of deprotonated peptides. Fragment ions are mostly different for these two cases. These lists of fragments are used to annotate peptide mass spectral libraries and to aid in the confirmation of specific amino acids in peptides.

**Keywords:** CID, Deprotonated peptides, Protonated peptides, Side-chain neutral losses, Residue-specific product ions

Received: 1 September 2017/Revised: 27 October 2017/Accepted: 31 October 2017/Published Online: 15 November 2017

## Introduction

The application of tandem mass spectrometry of protonated peptides as a powerful analytical technique for determining the amino acid identities and the sequence of peptides and proteins has been essentially developed in last 30 y. Numerous studies regarding the interpretation of protonated peptide spectra and the ion fragmentation mechanisms have been published [1–5]. In comparison with the well documented mass spectrometry of protonated peptides, mass spectrometry of deprotonated peptides has received less attention in proteomic research, although spectra of negative ions of peptides provide an important complement to positive ion spectra in sequencing peptides in many cases. Bowie and coworkers studied the mass spectrometry of deprotonated peptides [6–20] and described the fundamental backbone cleavages and the characteristic side-chain fragmentations of amino acids residues. These studies also demonstrated that often there are fragmentation pathways that depend on the specific side-chain structure of the amino acid. Harrison and coworkers [21–25] reported the sequence-specific fragmentation of deprotonated peptides containing alkyl side chains and indicated that there were more extensive sequence-specific fragmentations using low energy collision-induced dissociation (CID) mode instead of high energy CID. Subsequent studies were carried out on deprotonated peptides containing phenylalanine, glutamic acid, and proline. Cassady and coworkers [26–31] and other groups [32, 33] also have studied the fragmentations of a variety of deprotonated peptides.

The studies mentioned above show that the side chains of some amino acids affect the fragmentation reactions of deprotonated peptides, both in peptide backbone cleavage and in side-chain fragmentation. Peptide bond dissociation of the negative ions can provide sequence information similar to that obtained from positive peptide ions dissociation, although the negative ions exhibit additional dissociation pathways (such as c-ions). Fragmentation of the amino acid side chain is especially important in the identification of certain amino

*Correspondence to:* Yuxue Liang; *e-mail:* yuxue.liang@nist.gov

acid residues by their characteristic fragments. Bowie and coworkers [15, 17, 20] published a table that includes side-chain fragmentations for 15 amino acid residues in small peptides; many of them cannot be observed in the spectra of peptides that contain more than four amino acid residues. In this work, we studied the MS/MS spectra of deprotonated peptides using 390 dipeptides and many longer peptides, mainly commercial bioactive peptides. The spectra were acquired with a Thermo Orbitrap Elite mass spectrometer to collect both ion-trap and HCD (Higher Energy Collision Dissociation) spectra for inclusion in the NIST MS/MS library [34]. We extend the previous results on the side chain neutral losses and also present a comprehensive fragment ion list from 17 amino acid residues, confirmed with high mass accuracy. These fragment ions include those produced from the side chains as well as the fragments produced from whole amino acid residues. Statistical methods were applied to estimate the relative abundance of the characteristic side chain fragmentations.

## Experimental[1]

The dipeptides were purchased from GenScript (Atlanta, GA, USA), LifeTein (Somerset, NJ, USA, or Sigma-Aldrich (St. Louis, MO, USA) and the other peptides were purchased mainly from American Peptide Company (Sunnyvale, CA, USA). The peptides were dissolved in acetonitrile/water/formic acid (50:50:0.1) (v:v:v) at a concentration of about 0.1 mg/mL and infused into an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) via a nano-electrospray source. Ion trap (IT and FT-CID) and HCD spectra were acquired in the positive and negative mode. The gases $N_2$ (99.999 %) and He (99.999 %) were utilized as collision gases for the HCD and IT spectra, respectively. Ion trap spectra were acquired at a relative collision energy of 35%. The HCD spectra were collected by using up to 18 different collision voltages ranging from 2 V to 180 V, which is beyond the voltage where no precursor ions remained. The resolution for $MS^2$ was set at 30,000. Spectra were acquired in 'profile' mode for both FT-CID and HCD.

## Results and Discussion

Tandem mass spectra of 1800 peptides, as protonated and deprotonated ions, were acquired at different collision energies, with high mass accuracy, and are included in the 2017 version of the NIST tandem mass spectral library. The results for the 527 deprotonated peptides containing two to six amino acid residues were analyzed and are discussed below.

### *Neutral Losses from the Side Chains of Amino Acid Residues*

Spectra of negative ions of peptides are as informative as those of positive ions since sequence information is provided from backbone cleavage. Both positive and negative peptide ions produce the corresponding y-, b-, and a-ions, but the negative peptides also produce abundant c-ions and rarely observed x- and z-ions. Additionally, there are certain side-chain cleavage reactions of negative peptide ions that readily identify particular amino acid residues. Bowie and coworkers [6–20] reported some characteristic negative-ion fragmentations of side chains of amino acid residues, which are useful to identify these residues and to interpret mass spectra. However, a limited number of peptides were studied to address the neutral losses of side chains. It appears that many of the neutral losses that were proposed are not present in the spectra of peptides that contain more than four amino acid residues. The results presented here from measurement of 527 peptides, comprised of 390 dipeptides and 137 peptides with three to six amino acids, using high mass accuracy, provide a more detailed analysis and permit us to estimate the relative abundances of the neutral losses from the side chains of amino acid residues. The spectra of several hundred peptides containing seven or more amino acids were deposited in the NIST library, but not used in the statistical analysis because correlations with specific residues became less certain.

The side chain neutral losses from deprotonated peptides are summarized in Table 1. For each deprotonated peptide, we searched for all the ions that were produced by side-chain neutral loss from each amino acid residue. Using the HCD and FT-IT spectra, we identified the most abundant peak for neutral loss from each residue, including loss from the precursor ion as well as losses from y-, b-, a-, or c-ions. Then for each amino acid we averaged the abundances of all the peaks observed with all the peptides containing that residue, including the results for peptides which contain that residue but lack the corresponding peak in their spectra. The most frequently observed ion with the highest peak intensity is that for the loss of $C_2H_4O$ (acetaldehyde) from threonine. This loss was observed in all 52 spectra of peptides containing threonine, with an average intensity of 78.4% of base peak. The abundance of this neutral loss peak was set as 100 and the losses from the other amino acids were estimated relative to threonine. It should be noted that a similar calculation for the spectra of 475 peptides which do not contain threonine gave a value of 0.2%, probably due to incidental small peaks. This small percentage was not deducted from the positive result except as discussed below for the loss of water and $CO_2$.

The second most abundant neutral loss is that of $CH_2O$ (formaldehyde) from serine. It was observed in 61 out of 65 peptides containing serine, with an average intensity of 66%. Its relative abundance compared with threonine is estimated as 80

[1]Certain commercial equipment, instruments, or materials are identified in this document. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products identified are necessarily the best available for the purpose.

**Table 1.** Neutral Losses from Side Chains of Amino Acid Residues in the Tandem Mass Spectra of Deprotonated Peptides and Their Estimated Relative Abundances

| Amino acid residue | Neutral loss formula | Neutral loss mass | Number of peptides containing aa | Relative abundance |
|---|---|---|---|---|
| T | $C_2H_4O$ | 44.02621 | 52 | 100 |
| S | $CH_2O$ | 30.01056 | 65 | 80 |
| E | $H_2O$ | 18.01056 | 54 | 70 |
| D | $H_2O$ | 18.01056 | 69 | 60 |
| W | $C_9H_7N$ | 129.05785 | 51 | 60 |
| R | $CH_2N_2$ | 42.0218 | 94 | 60 |
| C | $H_2S$ | 33.98772 | 41 | 54 |
| M | $CH_4S$ | 48.00337 | 50 | 27 |
| Y | $C_7H_6O$ | 106.04186 | 70 | 20 |
| E | $CO_2$ | 43.98983 | 54 | 16 |
| C | $CH_2S$ | 45.98772 | 41 | 7 |
| H | $C_4H_4N_2$ | 80.03745 | 51 | 6 |
| F | $C_7H_8$ | 92.0626 | 74 | 3 |
| D | $CO_2$ | 43.98983 | 69 | 2 |
| M | $C_3H_6S$ | 74.01902 | 50 | 0.4 |
| M | $C_2H_6S$ | 62.01902 | 50 | 0.1 |
| P, V, L, I | $C_2H_4$ | 28.0313 | | < 0.5 |
| P, V, L, I | $C_3H_6$ | 42.04695 | | < 0.5 |
| L, I | $C_4H_8$ | 56.0626 | | < 0.5 |
| M(O) | $CH_3$ | 15.02348 | 10 | 70 |
| M(O) | $CH_3SOH$ | 63.99829 | 10 | 70 |
| M(O) | $C_3H_7OS$ | 91.02176 | 10 | 55 |
| M(O) | $C_2H_6OS$ | 78.01394 | 10 | 2 |

(Table 1). Other neutral losses listed in Table 1 occur with specific amino acids and their abundances were estimated from peptides containing the specific residue. Tryptophan loses the whole indolylmethyl side chain but arginine loses only the tail end of the side chain, $CH_2N_2$, with the same relative abundance, 60. Tyrosine, histidine, and phenylalanine lose the whole side chain with decreasing relative abundances, 20, 6, and 3, respectively. It is noted, however, that phenylalanine loses $C_6H_5CH_3$ whereas tyrosine loses an oxidized form, $CH_2{=}C_6H_4{=}O$. Histidine also loses an oxidized form of its side chain.

Loss of $H_2O$ and $CO_2$, which take place from the side chains of aspartic and glutamic acids, also occur from the terminal carboxyl group of many peptides not containing D or E. Therefore, to estimate the abundance of these losses from D and E, we compared their abundances in peptides containing D or E with those in peptides containing neither D nor E. Loss of water was found to be five to seven times more abundant if the peptide contained D or E, respectively. The relative abundances given for the loss of water in Table 1 were estimated by subtracting the abundance in peptides containing neither D nor E from the values found in peptides containing D or E. A similar comparison for the loss of $CO_2$ did not show significant differences between peptides containing D or E, or neither of these residues. Therefore, we examined the loss of a second $CO_2$ from the various peptides and found significant abundances only in peptides containing D or E. Table 1 shows that loss of $CO_2$ from the side chain of E is eight times more abundant than loss from D. Cysteine also undergoes two neutral losses, an abundant loss of $H_2S$ and a loss of $CH_2S$, which is eight times less abundant. Three losses were detected from methionine: an abundant loss of $CH_4S$ and losses of $C_3H_6S$ and $C_2H_6S$, which are 70 and 270 times less abundant, respectively.

Losses of alkenes were detected from P, V, I, and L residues but their relative abundances are very low (<0.5). They were observed only at high collision energies after the more favorable dissociations had taken place. Proline appears to lose $C_2H_4$ from its ring much more abundantly than $C_3H_6$. However, V, I, and L appear to lose the whole side chain more abundantly than a smaller alkene.

Oxidized methionine undergoes abundant losses of $CH_3$, $CH_3SOH$, or the whole side chain $C_3H_7OS$, and a very minor loss of $C_2H_6OS$. These losses are listed separately in the bottom of Table 1 because they were estimated from the spectra of only 10 dipeptides containing oxidized methionine.

The relative abundances in Table 1 were estimated from the average intensity of the corresponding peaks in all the peptides containing the amino acid residue irrespective of its position within the peptide chain. It was noted, however, that while the neutral losses from the side chains of threonine and serine were independent of the location of the amino acid within the peptide, other losses may be dependent on location. The most pronounced difference was found in dipeptides containing arginine, where the loss of $CH_2N_2$ is about five times more abundant from C-terminal than from N-terminal arginine, and with cysteine, where the loss of $H_2S$ is at least 10 times more abundant from the N-terminus than from the C-terminus.

By comparison with the neutral losses from the side chains of peptide-negative ions, neutral losses from positive ion side chains are abundant mainly from methionine and arginine. From the spectra of 525 singly protonated peptides containing two to six amino acid residues we found the following neutral losses (relative abundance): from methionine $CH_3SH$ (53), from arginine $CH_2N_2$ (19), from threonine $C_2H_4O$ (8), from serine $CH_2O$ (8), from cysteine $H_2S$ (7), from tryptophan $C_9H_7N$ (2), and from tyrosine $C_7H_6O$ (1). There is also significant loss of $H_2O$ from glutamic acid (40) but much less from aspartic acid (6); the latter abundance values were estimated by deducting the abundance of water loss from peptides containing neither E nor D. The losses from serine and threonine are site-specific, as discussed before [35].

## *Fragment Ions from Amino Acid Residues*

Although earlier studies [7–11, 15–18, 20] on side-chain fragmentation discussed mostly neutral losses, the authors also mentioned two negative ions formed from the side chains: $C_6H_5CH_2^-$ from F and $HOC_6H_4CH_2^-$ from Y. In the present study we confirm these findings and extend the results to other amino acid residues, such as the formation of $C_9H_8N^-$ from W or $C_4H_5N_2^-$ from H. Such characteristic ions can be used to identify amino acids within the peptide sequence. Side chain-negative ions may be produced while the amino acid residue is retained within the peptide backbone or after that residue is released following backbone cleavage. After the amino acids are released from the peptide chain as negative ions, they may dissociate further to produce smaller negative ions. In the current measurements, HCD spectra were acquired at various collision energies, including high energies in which the

fragmentation of the single residues is observed. It should be pointed out, however, that since deprotonated peptides often form c-ions, and sometimes x-ions, what we consider here to be a single residue often includes an additional NH group from the adjacent residue.

In Table 2 we list the significant negative ions observed from single residues, including those obtained from the side chain alone and those from the whole residue. Fragment ions that are specific to one amino acid and are observed with significant abundance in the spectra of at least 20% of the peptides containing that amino acid are denoted with a bold mass. When the mass value is not bold, that fragment ion is produced either with low abundance or from more than one amino acid and thus is not useful for amino acid identification. No fragment ions are listed for glycine and alanine because all the ions that can be formed from these residues can be formed also from larger amino acids following side-chain losses. Also, no fragment ions are listed for threonine because this residue loses acetaldehyde from the side chain very rapidly, even before peptide bond cleavage, leaving behind a glycine residue. Serine loses its side chain (formaldehyde) slightly more slowly than threonine and thus the c-ion product from serine can be observed, but with very low abundance. It should be noted, however, that an earlier paper [36] reported the observation of abundant c-ions from deprotonated peptides containing serine and threonine. The difference between those results and our current work is likely due to the difference in collision energy; they used low collision energy CID (SORI-CID) whereas we used higher energy collision dissociation (HCD).

The last column in Table 2 points out the relation between the observed fragment ion and the precursor amino acid (aa). In most cases this is indicated by showing the aa losing a proton and then losing identifiable neutral species such as $H_2O$, $NH_3$, or $CO_2$. When the fragment ion is produced from a c-ion product, the relation is indicated by "aa – OH + $NH_2$ – H". An example of an ion derived from a c-ion, i.e., from a residue with an added $NH_2$ group, is the fragment ion at *m/z* 119.0285 from cysteine, which corresponds to the formula $NH_2CH(CH_2S^-)CONH_2$. This ion is not observed when cysteine is in the C-terminal position in the peptide, except when the peptide is amidated, but it can be formed from cysteine at any other position.

Table 2 shows that several amino acids have multiple characteristic negative ions. For example, the following ions are derived from histidine (Scheme 1). These nine ions are observed in significant abundance and are characteristic of histidine only. Their abundances vary with peptide composition and with collision energy. For example, the peak of the fragment ion $C_4H_5N_2^-$ at *m/z* 81.0458 is observed in the spectra of all dipeptides containing histidine with intensity >5%, 86% of the spectra exhibit the peak with intensity >25%, and 68% of the spectra show intensity >50%. The abundance of this peak in the spectra of peptides containing histidine decreases as the peptide length increases, clearly due to competing pathways. For example, the peak is exhibited with intensity >5% by 100%

**Table 2.** Amino Acid Fragment Ions from Deprotonated Peptides

| Amino acid residue | Fragment ion mass[a] | Formula | Relation to amino acid (aa)[b] |
|---|---|---|---|
| C | **119.02846** | $C_3H_7N_2OS$ | [aa-OH+NH2-H]-* |
| D | 71.01385 | $C_3H_3O_2$ | [aa-H-NH3-CO2]- |
| D | 96.0091 | $C_4H_2NO_2$ | [aa-H-2H2O]- |
| D | 113.03565 | $C_4H_5N_2O_2$ | [aa-OH+NH2-H-H2O]-* |
| D | **115.00368** | $C_4H_3O_4$ | [aa-H-NH3]- |
| E | 82.02984 | $C_4H_4NO$ | [aa-H-H2O-CO2-H2]- |
| E | 84.04549 | $C_4H_6NO$ | [aa-H-H2O-CO2]- |
| E | 101.07204 | $C_4H_9N_2O$ | [aa-OH+NH2-H-CO2]-* |
| E | **102.05605** | $C_4H_8NO_2$ | [aa-H-CO2]- |
| E | 109.04074 | $C_5H_5N_2O$ | [aa-OH+NH2-H-2H2O]-* |
| E | 128.03532 | $C_5H_6NO_3$ | [aa-H-H2O]- |
| F | **91.05532** | $C_7H_7$ | [aa-H-C2H3NO2]- |
| F | **103.05532** | $C_8H_7$ | [aa-H-NH3-CO2]- |
| F | **118.06622** | $C_8H_8N$ | [aa-H-CH2O2]- |
| F | **146.06114** | $C_9H_8NO$ | [aa-H-H2O]- |
| F | **163.08769** | $C_9H_{11}N_2O$ | [aa-OH+NH2-H]-* |
| H | **67.03017** | $C_3H_3N_2$ | [aa-H-C3H5NO2]- |
| H | **80.038** | $C_4H_4N_2$ | [aa-H-C2H4NO2]- |
| H | **81.04582** | $C_4H_5N_2$ | [aa-H-C2H3NO2]- |
| H | **93.04582** | $C_5H_5N_2$ | [aa-H-NH3-CO2]- |
| H | **108.05672** | $C_5H_6N_3$ | [aa-H-CH2O2]- |
| H | **110.07237** | $C_5H_8N_3$ | [aa-H-CO2]- |
| H | **118.04107** | $C_6H_4N_3$ | [aa-H-2H2O]- |
| H | **136.05164** | $C_6H_6N_3O$ | [aa-H-H2O]- |
| H | **153.07818** | $C_6H_9N_4O$ | [aa-OH+NH2-H]-* |
| I | 82.06622 | $C_5H_8N$ | [aa-H-CH2O2-H2]- |
| I | 112.07679 | $C_6H_{10}NO$ | [aa-H-H2O]- |
| I | 129.10334 | $C_6H_{13}N_2O$ | [aa-OH+NH2-H]-* |
| K | **97.07712** | $C_5H_9N_2$ | [aa-H-CH2O2-H2]- |
| K | 99.09277 | $C_5H_{11}N_2$ | [aa-H-CH2O2]- |
| K | 144.11424 | $C_6H_{14}N_3O$ | [aa-OH+NH2-H]-* |
| L | 82.06622 | $C_5H_8N$ | [aa-H-CH2O2-H2]- |
| L | 112.07679 | $C_6H_{10}NO$ | [aa-H-H2O]- |
| L | 129.10334 | $C_6H_{13}N_2O$ | [aa-OH+NH2-H]-* |
| M | 100.0404 | $C_4H_6NO_2$ | [aa-H-CH3SH]- |
| M | **147.05976** | $C_5H_{11}N_2OS$ | [aa-OH+NH2-H]-* |
| N | **95.02509** | $C_4H_3N_2O$ | [aa-H-2H2O]- |
| N | 96.0091 | $C_4H_2NO_2$ | [aa-H-NH3-H2O]- |
| N | 98.02475 | $C_4H_4NO_2$ | [aa-H-NH3-O]- |
| N | 113.03565 | $C_4H_5N_2O_2$ | [aa-H-H2O]- |
| N | **114.01967** | $C_4H_4NO_3$ | [aa-H-NH3]- |
| N | 130.0622 | $C_4H_8N_3O_2$ | [aa-OH+NH2-H]-* |
| P | **96.04549** | $C_5H_6NO$ | [aa-H-H2O]- |
| P | 113.07204 | $C_5H_9N_2O$ | [aa-OH+NH2-H]-* |
| Q | 82.02984 | $C_4H_4NO$ | [aa-H-NH3-CO2-H2]- |
| Q | 84.04549 | $C_4H_6NO$ | [aa-H-NH3-CO2]- |
| Q | 101.07204 | $C_4H_9N_2O$ | [aa-H-CO2]- |
| Q | 109.04074 | $C_5H_5N_2O$ | [aa-H-2H2O]- |
| Q | 127.0513 | $C_5H_7N_2O_2$ | [aa-H-H2O]- |
| Q | 128.03532 | $C_5H_6NO_3$ | [aa-H-NH3]- |
| Q | 145.06187 | $C_5H_9N_2O_3$ | [aa-H]- |
| R | **58.04107** | $CH_4N_3$ | [aa-H-C5H9NO2]- |
| R | **112.08802** | $C_5H_{10}N_3$ | [aa-H-NH3-CO2]- |
| R | **130.09859** | $C_5H_{12}N_3O$ | [aa-OH+NH2-H-CH2N2]-* |
| R | **131.0826** | $C_5H_{11}N_2O_2$ | [aa-H-CH2N2]- |
| R | **155.09383** | $C_6H_{11}N_4O$ | [aa-H-H2O]- |
| R | 156.07785 | $C_6H_{10}N_3O_2$ | [aa-H-NH3]- |
| R | 172.12038 | $C_6H_{14}N_5O$ | [aa-OH+NH2-H]-* |
| S | 103.0513 | $C_3H_7N_2O_2$ | [aa-OH+NH2-H]-* |
| V | 115.08769 | $C_5H_{11}N_2O$ | [aa-OH+NH2-H]-* |
| W | **116.05057** | $C_8H_6N$ | [aa-H-C3H5NO2]- |
| W | **130.06622** | $C_9H_8N$ | [aa-H-C2H3NO2]- |
| W | **142.06622** | $C_{10}H_8N$ | [aa-H-NH3-CO2]- |
| W | **159.09277** | $C_{10}H_{11}N_2$ | [aa-H-CO2]- |
| W | 186.05605 | $C_{11}H_8NO_2$ | [aa-H-NH3]- |
| W | 202.09859 | $C_{11}H_{12}N_3O$ | [aa-OH+NH2-H]-* |
| Y | **93.03459** | $C_6H_5O$ | [aa-H-NH3-CO2-C2H2]- |
| Y | **106.04241** | $C_7H_6O$ | [aa-H-NH3-CO2-CH]- |
| Y | **107.05024** | $C_7H_7O$ | [aa-H-NH3-C2O2]- |

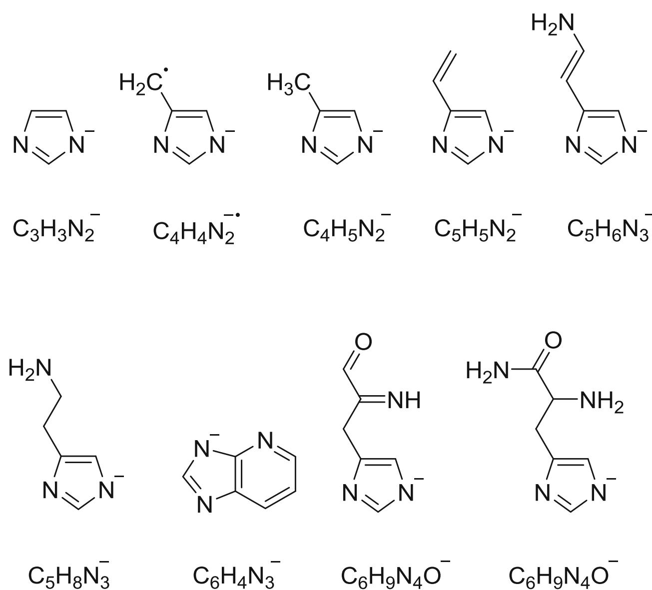
**Table 2.** (continued)

| Amino acid residue | Fragment ion mass[a] | Formula | Relation to amino acid (aa)[b] |
|---|---|---|---|
| Y | **119.05024** | $C_8H_7O$ | [aa-H-$NH_3$-$CO_2$]- |
| Y | 133.05331 | $C_8H_7NO$ | [aa-H-$CH_3O_2$]- |
| Y | **134.06114** | $C_8H_8NO$ | [aa-H-$CH_2O_2$]- |
| Y | **136.07679** | $C_8H_{10}NO$ | [aa-H-$CO_2$]- |
| Y | **162.05605** | $C_9H_8NO_2$ | [aa-H-$H_2O$]- |
| Y | **163.04007** | $C_9H_7O_3$ | [aa-H-$NH_3$]- |
| Y | 179.0826 | $C_9H_{11}N_2O_2$ | [aa-OH+$NH_2$-H]-* |

[a]The more abundant ions are indicated in bold characters
[b]Fragment ions formed by neutral loss from a c-ion are indicated with a *

of dipeptides, but only by 70% of peptides with three to six residues, by 54% of peptides with seven to egjt residues, and by <1% of the peptides with eight to 15 residues. Thus, the fragment ion peak at *m/z* 81.0458 can be used to identify histidine in peptides with less than eight amino acids. All nine ions shown above are good candidates for identifying histidine in peptides. Identification with multiple characteristic negative ions formed from a single amino acid provides greater confidence.

In parallel with histidine, phenylalanine produces five characteristic fragment ions with significant abundance, which are not observed with peptides not containing F. Similarly, tyrosine produces 10 characteristic fragment ions, not observed in the absence of Y, and eight of them appear with significant abundance. Tryptophan produces six characteristic fragment ions, four of which with significant abundance. Arginine may produce eight fragment ions, but three of them appear with low abundance or are formed from other residues. Proline produces two fragment ions but only one of them is characteristic and can be used to identify proline. The characteristic ion from proline is $C_5H_6NO^-$ with *m/z* 96.0455. However, this peak must be measured with high mass accuracy to distinguish it from a




**Scheme 1.** Proposed structures for negative fragment ions from histidine

peak at *m/z* 96.0091 of the ion $C_4H_2NO_2^-$ formed from D or N. Other amino acids have only one characteristic fragment ion but some amino acids have none, such as glutamine. Seven fragment ions are listed in Table 2 for glutamine, but none of them are useful for identification because they can be produced also from glutamic acid, although sometimes with lower abundance (and thus not listed under glutamic acid). On the other hand, glutamic acid produces one characteristic fragment ion that is not observed from glutamine.

## *Comparison to Fragments of Positive Ions*

The negatively charged fragment ions produced from deprotonated peptides (Table 2) are mostly different from the positively charged fragment ions produced from protonated peptides. Several authors have discussed "immonium ions" and other single amino acid fragments from protonated peptides [37, 38] and have incorporated them into software for interpretation of tandem mass spectra of peptides, such as Mascot [39]. During our analysis of peptide mass spectra for inclusion in the NIST peptide library, we began by using the list of immonium and other single residue fragment ions from Mascot for annotating the spectra, and then, by manual analysis of specific peptide spectra, we found additional fragment ion candidates for specific amino acid residues. The candidate ions with high mass accuracy were then confirmed by statistical analysis of all the peptides containing the specific residue. Our modified list is presented in Table 3. Comparison of Tables 2 and 3 shows that 18 fragments appear in positive and negative forms that differ by two protons. All these ions are formally derived from the protonated or deprotonated single residue by loss of the same neutrals ($NH_3$, $H_2O$, $2H_2O$). In addition, 17 of the fragment ions have the same formula with the negative and positive charge, differing in mass by two electrons only. All the other fragments have unrelated formulas due to differences in fragmentation pathways between positive and negative peptide ions. Even for those cases with the same formulas, the structures of the positive and negative ions may be different because the positive and negative charges are localized at different sites. For example, the immonium ion derived from glutamic acid has the formula $C_4H_8NO_2$ and the structure $^+NH_2$=$CHCH_2CH_2CO_2H$, but the observed negative ion fragment with the same formula is more likely to have the structure $NH_2CH_2CH_2CH_2CO_2^-$. Similarly, the fragment from tyrosine with the formula $C_8H_{10}NO$ has the structure $^+NH_2$=$CHCH_2C_6H_4OH$ in positive mode but $NH_2CH_2CH_2C_6H_4O$- in negative mode.

The fragment ion with the formula $C_6H_7O^+$, i.e., protonated phenol, is formed from tyrosine by dissociating the phenol group from the side chain along with a proton. The $C_6H_7O^+$ ion, however, is also found to be formed from protonated peptides containing phenylalanine or tryptophan but no tyrosine. This ion is clearly formed from a phenyl cation, produced from these amino acids, by reaction with water in the collision cell, as reported before for many other aromatic compounds [40]. In addition, the phenyl cation also reacts with $N_2$, present in the collision cell of the Orbitrap, to produce the stable

**Table 3.** Amino Acid Fragment Ions from Protonated Peptides[a]

| Amino acid residue | Fragment ion mass[b] | Formula | Proposed annotation[c] |
|---|---|---|---|
| A | 44.04948 | $C_2H_6N$ | ImA |
| A | *72.04439* | $C_3H_6NO$ | A-$H_2O$ |
| C | **76.02155** | $C_2H_6NS$ | ImC |
| C | ***86.0059*** | $C_3H_4NS$ | C-2$H_2O$ |
| C | ***104.0165*** | $C_3H_6NOS$ | C-$H_2O$ |
| D | 70.02874 | $C_3H_4NO$ | ImD-$H_2O$ |
| D | 88.0393 | $C_3H_6NO_2$ | ImD |
| D | *98.02365* | $C_4H_4NO_2$ | D-2$H_2O$ |
| D | *116.0342* | $C_4H_6NO_3$ | D-$H_2O$ |
| E | *84.04439* | $C_4H_6NO$ | ImE-$H_2O$ |
| E | 102.055 | $C_4H_8NO_2$ | ImE |
| E | *112.0393* | $C_5H_6NO_2$ | E-2$H_2O$ |
| E | *130.0499* | $C_5H_8NO_3$ | E-$H_2O$ |
| F | *77.03858* | $C_6H_5$ | fF |
| F | 91.05423 | $C_7H_7$ | fF |
| F | *95.04914* | $C_6H_7O$ | fF |
| F | *103.0542* | $C_8H_7$ | ImF-$NH_3$ |
| F | *105.0447* | $C_6H_5N_2$ | fF |
| F | **120.0808** | $C_8H_{10}N$ | ImF |
| F | *130.0651* | $C_9H_8N$ | F-2$H_2O$ |
| F | *148.0757* | $C_9H_{10}NO$ | F-$H_2O$ |
| G | 30.03383 | $CH_4N$ | ImG |
| H | ***81.04472*** | $C_4H_5N_2$ | fH |
| H | **82.05255** | $C_4H_6N_2$ | ImH-$H_2O$ |
| H | ***83.06037*** | $C_4H_7N_2$ | fH |
| H | ***93.04472*** | $C_5H_5N_2$ | ImH-$NH_3$ |
| H | ***95.06037*** | $C_5H_7N_2$ | fH |
| H | **110.0713** | $C_5H_8N_3$ | ImH |
| H | ***120.0556*** | $C_6H_6N_3$ | H-2$H_2O$ |
| H | **121.0396** | $C_6H_5N_2O$ | H-$H_2O$ |
| H | 123.0553 | $C_6H_7N_2O$ | fH |
| H | **138.0662** | $C_6H_8N_3O$ | H-$H_2O$ |
| H | **166.0611** | $C_7H_8N_3O_2$ | fH |
| I | 44.04948 | $C_2H_6N$ | fI |
| I | 86.09643 | $C_5H_{12}N$ | ImI |
| I | *96.08078* | $C_6H_{10}N$ | I-2$H_2O$ |
| I | *114.0913* | $C_6H_{12}NO$ | I-$H_2O$ |
| K | *56.04948* | $C_3H_6N$ | fK |
| K | 70.06513 | $C_4H_8N$ | fK |
| K | *82.06513* | $C_5H_8N$ | fK |
| K | **84.08078** | $C_5H_{10}N$ | ImK-$NH_3$ |
| K | **101.1073** | $C_5H_{13}N_2$ | ImK |
| K | ***111.0917*** | $C_6H_{11}N_2$ | K-2$H_2O$ |
| K | **112.0757** | $C_6H_{10}NO$ | K-$H_2O$-$NH_3$ |
| K | **129.1022** | $C_6H_{13}N_2O$ | K-$H_2O$ |
| L | 44.04948 | $C_2H_6N$ | fL |
| L | 86.09643 | $C_5H_{12}N$ | ImL |
| L | *96.08078* | $C_6H_{10}N$ | L-2$H_2O$ |
| L | *114.0913* | $C_6H_{12}NO$ | L-$H_2O$ |
| M | **61.01065** | $C_2H_5S$ | fM |
| M | **104.0529** | $C_4H_{10}NS$ | ImM |
| M | *132.0478* | $C_5H_{10}NOS$ | M-$H_2O$ |
| N | 70.02874 | $C_3H_4NO$ | ImN-$NH_3$ |
| N | **87.05529** | $C_3H_7N_2O$ | ImN |
| N | ***97.03964*** | $C_4H_5N_2O$ | N-2$H_2O$ |
| N | *98.02365* | $C_4H_4NO_2$ | N-$H_2O$-$NH_3$ |
| N | *115.0502* | $C_4H_7N_2O_2$ | N-$H_2O$ |
| P | 70.06513 | $C_4H_8N$ | ImP |
| P | *80.04948* | $C_5H_6N$ | P-2$H_2O$ |
| P | *98.06004* | $C_5H_8NO$ | P-$H_2O$ |
| Q | 56.04948 | $C_3H_6N$ | fQ |
| Q | 84.04439 | $C_4H_6NO$ | ImQ-$NH_3$ |
| Q | **101.0709** | $C_4H_9N_2O$ | ImQ |
| Q | *111.0553* | $C_5H_7N_2O$ | Q-2$H_2O$ |
| Q | *112.0393* | $C_5H_6NO_2$ | Q-$H_2O$-$NH_3$ |
| Q | **129.0659** | $C_5H_9N_2O_2$ | Q-$H_2O$ |
| R | ***60.05562*** | $CH_6N_3$ | fR |
| R | 70.06513 | $C_4H_8N$ | fR |
| R | ***85.07602*** | $C_4H_9N_2$ | fR |
| R | 87.09167 | $C_4H_{11}N_2$ | fR |
| R | 100.0869 | $C_4H_{10}N_3$ | fR |
| R | **112.0869** | $C_5H_{10}N_3$ | ImR-$NH_3$ |
| R | ***113.0709*** | $C_5H_9N_2O$ | fR |
| R | ***115.0866*** | $C_5H_{11}N_2O$ | fR |
| R | *116.0706* | $C_5H_{10}NO_2$ | fR |
| R | 129.1135 | $C_5H_{13}N_4$ | ImR |
| R | *139.0978* | $C_6H_{11}N_4$ | R-2$H_2O$ |
| R | ***140.0818*** | $C_6H_{10}N_3O$ | R-$H_2O$-$NH_3$ |
| R | ***157.1084*** | $C_6H_{13}N_4O$ | R-$H_2O$ |
| S | **60.04439** | $C_2H_6NO$ | ImS |
| S | *70.02874* | $C_3H_4NO$ | S-2$H_2O$ |
| S | *88.0393* | $C_3H_6NO_2$ | S-$H_2O$ |
| T | **74.06004** | $C_3H_8NO$ | ImT |
| T | *84.04439* | $C_4H_6NO$ | T-2$H_2O$ |
| T | *102.055* | $C_4H_8NO_2$ | T-$H_2O$ |
| V | 41.03858 | $C_3H_5$ | fV |
| V | **55.05423** | $C_4H_7$ | ImV-$NH_3$ |
| V | 72.08078 | $C_4H_{10}N$ | ImV |
| W | 77.03858 | $C_6H_5$ | fW |
| W | ***89.03858*** | $C_7H_5$ | fW |
| W | *105.0447* | $C_6H_5N_2$ | fW |
| W | ***115.0542*** | $C_9H_7$ | fW |
| W | **117.0573** | $C_8H_7N$ | fW |
| W | 130.0651 | $C_9H_8N$ | fW |
| W | **132.0808** | $C_9H_{10}N$ | fW |
| W | ***142.0651*** | $C_{10}H_8N$ | ImW-$NH_3$ |
| W | ***144.0808*** | $C_{10}H_{10}N$ | fW |
| W | *146.06* | $C_9H_8NO$ | fW |
| W | **159.0917** | $C_{10}H_{11}N_2$ | ImW |
| W | ***169.076*** | $C_{11}H_9N_2$ | W-2$H_2O$ |
| W | **170.06** | $C_{11}H_8NO$ | W-$H_2O$-$NH_3$ |
| W | ***187.0866*** | $C_{11}H_{11}N_2O$ | W-$H_2O$ |
| W | ***188.0706*** | $C_{11}H_{10}NO_2$ | W-$NH_3$ |
| Y | *77.03858* | $C_6H_5$ | fY |
| Y | 91.05423 | $C_7H_7$ | fY |
| Y | *95.04914* | $C_6H_7O$ | fY |
| Y | *105.0447* | $C_6H_5N_2$ | fY |
| Y | **107.0491** | $C_7H_7O$ | fY |
| Y | ***119.0491*** | $C_8H_7O$ | ImY-$NH_3$ |
| Y | **136.0757** | $C_8H_{10}NO$ | ImY |
| Y | *146.06* | $C_9H_8NO$ | Y-2$H_2O$ |
| Y | *164.0706* | $C_9H_{10}NO_2$ | Y-$H_2O$ |

[a]High-accuracy tandem mass spectra of about 1800 peptides were analyzed to derive a modified list of immonium ions and other fragment ions from single amino acid residues

[b]The more abundant ions are indicated in bold characters and the ions that were absent in the Mascot list are indicated in italics

[c]The following abbreviations are proposed for annotating the peaks of fragment ions produced from single amino acid residues: ImX = immonium ion of amino acid X; X-$H_2O$ = amino acid X + H – $H_2O$ (similarly for X-$NH_3$, X-2$H_2O$, etc.); fX = fragment ion from amino acid X (not easily annotated with simple neutral loss)

phenyldiazonium cation, $C_6H_5N_2^+$. This ion is also observed to be formed from phenylalanine, tyrosine, and tryptophan, and was reported before for many other compounds [41]. This reaction with $N_2$ is specific to positive ions and is not observed with negative ions. The reaction with water is observed mainly with positive ions and rarely with negative ions.

## Conclusions

Highly specific and abundant neutral losses and negative ion products from many amino acid residues were determined with

high mass accuracy from the MS/MS spectra of 527 deprotonated peptides, and are useful in the identification of specific residues within peptides. Side chain neutral losses are less abundant with protonated peptides. Single residue fragment ions are also different in positive and negative mode, with less than a third of the ions having related formulas. An updated list of immonium and other fragments from single residues in protonated peptides is proposed for future annotation of fragments in peptide mass spectra.

## References


1. Biemann, K., Martin, S.A.: Mass spectrometric determination of the amino acid sequence of peptides and proteins. Mass Spectrom. Rev. **6**, 1–76 (1987)
2. Polce, M.J., Ren, D., Wesdemiotis, C.: Dissociation of peptide bonds in protonated peptides. J. Mass Spectrom. **35**, 1391–1398 (2000)
3. Wysocki, V.H., Tsaprailis, G., Smith, L.L., Breci, L.A.: Mobile and localized protons: A framework for understanding peptide dissociation. J Mass Spectrom. **35**, 1399–1408 (2000)
4. Godovac-Zimmermann, J., Brown, L.R.: Perspectives for mass spectrometry and functional proteomics. Mass Spectrom. Rev. **20**, 1–57 (2001)
5. Paizs, B., Suhai, S.: Fragmentation pathways of protonated peptides. Mass Spectrom. Rev. **24**, 508–548 (2005)
6. Eckersley, M., Bowie, J.H., Hayes, R.N.: Collision-induced dissociation of deprotonated peptides: dipeptides and tripeptides with hydrogen and alkyl alpha-groups – an aid to structural determination. Org. Mass Spectrom. **24**, 597–602 (1989)
7. Waugh, R.J., Eckersley, M., Bowie, J.H., Hayes, R.N.: Collision-induced dissociation of deprotonated peptides – dipeptides containing serine and threonine. Int. J. Mass Spectrom. Ion Processes. **98**, 135–145 (1990)
8. Waugh, R.J., Bowie, J.H., Hayes, R.N.: Collision-induced dissociation of deprotonated peptides – dipeptides containing aspartic or glutamic acids. Org. Mass Spectrom. **26**, 250–256 (1991)
9. Waugh, R.J., Bowie, J.H., Hayes, R.N.: Collision-induced dissociation of deprotonated peptides – dipeptides containing phenylalanine, tyrosine, histidine, and tryptophan. Int. J. Mass Spectrom. Ion Processes. **107**, 333–347 (1991)
10. Waugh, R.J., Bowie, J.H., Gross, M.L.: Collision-induced dissociation of deprotonated peptides – dipeptides containing Asn, Arg, and Lys. Austral. J. Chem. **46**, 693–702 (1993)
11. Waugh, R.J., Bowie, J.H., Gross, M.L.: Collision-induced dissociation of deprotonated peptides – dipeptides containing methionine and cysteine. Rapid Commun. Mass Spectrom. **7**, 623–625 (1993)
12. Waugh, R.J., Bowie, J.H., Gross, M.L., Vollmer, D.: Collision-induced dissociation of deprotonated peptides. Dipeptides and tripeptides containing proline. Int. J. Mass Spectrom. Ion Processes. **133**, 165–174 (1994)
13. Waugh, R.J., Bowie, J.H.: A review of the collision-induced dissociation of deprotonated dipeptides and tripeptides. An aid to structural identification. Rapid Commun. Mass Spectrom. **8**, 169–173 (1994)
14. Bradford, A.M., Waugh, R.J., Bowie, J.H.: Characterization of underivatized tetrapeptides by negative ion fast atom bombardment. Rapid Commun. Mass Spectrom. **9**, 677–685 (1995)
15. Steinborner, S.T., Bowie, J.H.: A comparison of the positive- and negative-ion mass spectra of bio-active peptides from the dorsal secretion of the Australian red tree frog, *Litoria rubella*. Rapid Commun. Mass Spectrom. **10**, 1243–1247 (1996)
16. Brinkworth, C.S., Dua, S., McAnoy, A.M., Bowie, J.H.: Negative ion fragmentations of deprotonated peptides: backbone cleavages directed through both Asp and Glu. Rapid Commun. Mass Spectrom. **15**, 1965–1973 (2001)
17. Bowie, J.H., Brinkworth, C.S., Dua, S.: Collision-induced fragmentations of the $(M–H)^-$ parent anions of underivatized peptides. An aid to structure determination and some unusual negative ion cleavages. Mass Spectrom. Rev. **21**, 87–107 (2002)
18. Brinkworth, C.S., Bowie, J.H.: Negative ion electrospray mass spectra of the maculatin peptides from the tree frogs *Litoria genimaculata* and *Litoria eucnemis*. Rapid Commun. Mass Spectrom. **17**, 2215–2225 (2003)
19. Bilusich, D., Brinkworth, C.S., Bowie, J.H.: Negative ion mass spectra of Cys-containing peptides. The characteristic Cys γ backbone cleavage: a joint experimental and theoretical study. Rapid Commun. Mass Spectrom. **18**, 544–552 (2004)
20. Bilusich, D., Bowie, J.H.: Fragmentations of $(M–H)^-$ anions of underivatized peptides. Part 2: Characteristic cleavages of Ser and Cys and of disulfides and other post-translational modifications, together with some unusual internal processes. Mass Spectrom. Rev. **28**, 20–34 (2009)
21. Harrison, A.G.: Sequence-specific fragmentation of deprotonated peptides containing H or alkyl side chains. J. Am. Soc. Mass Spectrom. **12**, 1–13 (2001)
22. Harrison, A.G.: Effect of phenylalanine on the fragmentation of deprotonated peptides. J. Am. Soc. Mass Spectrom. **13**, 1242–1249 (2002)
23. Harrison, A.G.: Characterization of α- and γ-glutamyl dipeptides by negative ion collision-induced dissociation. J. Mass Spectrom. **39**, 136–144 (2004)
24. Harrison, A.G., Young, A.B.: Fragmentation reactions of deprotonated peptides containing proline. The proline effect. J, Mass Spectrom. **40**, 1173–1186 (2005)
25. Harrison, A.G., Young, A.B.: Fragmentation reactions of deprotonated peptides containing aspartic acid. Int. J. Mass Spectrom. **255**, 111–122 (2006)
26. Jai-nkuknan, J., Cassady, C.J.: Anion and cation post-source decay time-of-flight mass spectrometry of small peptides: Substance P, angiotensin II, and renin substrate. Rapid Commun. Mass Spectrom. **10**, 1678–1682 (1996)
27. Jai-nkuknan, J., Cassady, C.J.: Negative ion post-source decay time of-flight mass spectrometry of peptides containing acidic amino acid residues. Anal. Chem. **70**, 5122–5128 (1998)
28. Ewing, N.P., Cassady, C.J.: Dissociation of multiply charged negative ions from hirudin (54–65), fibrinopeptide B, and insulin A (oxidized). J. Am. Soc. Mass Spectrom. **12**, 105–116 (2001)
29. Clipston, N.L., Jai-nkuknan, J., Cassady, C.J.: A comparison of negative and positive ion time-of-flight post-source decay mass spectrometry for peptides containing basic residues. Int. J. Mass Spectrom. **222**, 363–381 (2003)
30. Li, Z., Yalcin, T., Cassady, C.J.: C-terminal amino acid residue loss for deprotonated peptide ions containing glutamic acid, aspartic acid, or serine residues at the C-terminus. J. Mass Spectrom. **41**, 939–949 (2006)
31. Pu, D., Clipston, N.L., Cassady, C.J.: A comparison of positive and negative ion collision-induced dissociation for model heptapeptides with one basic residue. J. Mass Spectrom. **45**, 297–305 (2010)
32. Marzluff, E.M., Campbell, S., Rodgers, M.T., Beauchamp, J.L.: Low energy dissociation pathways of small deprotonated peptides in the gas phase. J. Am. Chem. Soc. **116**, 7787–7796 (1994)
33. Grzetic, J., Oomens, J.: Spectroscopic evidence for an oxazolone structure in anionic b-type peptide fragments. J. Am. Soc. Mass Spectrom. **23**, 290–300 (2012)
34. NIST MS/MS library (http://chemdata.nist.gov/mass-spc/msms-search/) (2017)
35. Neta, P., Pu, Q., Yang, X., Stein, S.: Consecutive neutral losses of $H_2O$ and $C_2H_4O$ from N-terminal Thr–Thr and Thr–Ser in collision-induced dissociation of protonated peptides: Position-dependent water loss from single Thr or Ser. Int. J. Mass Spectrom. **267**, 295–301 (2007)
36. Pu, D., Cassady, C.J.: Negative ion dissociation of peptides containing hydroxyl side chains. Rapid Commun. Mass Spectrom. **22**, 91–100 (2008)
37. Falick, A.M., Hines, W.M., Medzihradszky, K.F., Baldwin, M.A., Gibson, B.W.: Low-mass ions produced from peptides by high-energy collision-induced dissociation in tandem mass spectrometry. J. Am. Soc. Mass Spectrom. **4**, 882–893 (1993)
38. Papayannopoulos, I.A.: The interpretation of collision-induced dissociation tandem mass spectra of peptides. Mass Spectrom. Rev. **14**, 49–73 (1995)
39. Mascot program for analysis of peptide mass spectra (http://www.matrixscience.com), accessed 02/08/2017
40. Neta, P., Farahani, M., Simon-Manso, Y., Liang, Y., Yang, X., Stein, S.: Unexpected peaks in tandem mass spectra due to reaction of product ions with residual water in mass spectrometer collision cells. Rapid Commun. Mass Spectrom. **28**, 2645–2660 (2014)
41. Liang, Y., Neta, P., Simon-Manso, Y., Stein, S.: Reaction of arylium ions with the collision gas $N_2$ in electrospray ionization mass spectrometry. Rapid Commun Mass Spectrom. **29**, 629–636 (2015)