



RESEARCH ARTICLE

Sequencing Lys-N Proteolytic Peptides by ESI and MALDI Tandem Mass Spectrometry

Mathieu Dupré, Sonia Cantel, Pascal Verdié, Jean Martinez, Christine Enjalbal

Institut des Biomolécules Max Mousseron (IBMM), UMR 5247 CNRS—Universités Montpellier 1 et 2, Bâtiment Chimie (17), Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France

Abstract

In this study, we explored the MS/MS behavior of various synthetic peptides that possess a lysine residue at the N-terminal position. These peptides were designed to mimic peptides produced upon proteolysis by the Lys-N enzyme, a metalloendopeptidase issued from a Japanese fungus *Grifola frondosa* that was recently investigated in proteomic studies as an alternative to trypsin digestion, as a specific cleavage at the amide X-Lys chain is obtained that provides N-terminal lysine peptide fragments. In contrast to tryptic peptides exhibiting a lysine or arginine residue solely at the C-terminal position, and are thus devoid of such basic amino acids within the sequence, these Lys-N proteolytic peptides can contain the highly basic arginine residue anywhere within the peptide chain. The fragmentation patterns of such sequences with the ESI-QqTOF and MALDI-TOF/TOF mass spectrometers commonly used in proteomic bottom-up experiments were investigated.

Key words: Lys-N proteolytic peptides, MS/MS, MALDI-TOF/TOF, ESI-QqTOF

Introduction

In “bottom-up” proteomics [1], mass spectrometry analyses are undertaken at the peptide level after sample digestion, which is usually performed with trypsin [2]. The digest obtained either from an isolated protein, or from a complex protein mixture, is analyzed by various complementary mass spectrometry techniques such as MALDI-TOF for peptide mass fingerprinting (PMF), MALDI-TOF/TOF, and/or LC/ESI-MS/MS for peptide sequence information [3]. The recorded mass spectra are correlated to genomic data using a computer algorithm that virtually cleaves known archived proteins into constitutive peptides according to the specific enzymatic treatment [4–6]. Whatever the method used, provided that trypsin is used for the proteolysis

stage, the recovered peptides contain either arginine or lysine at the C-terminal position [2]. The presence of such a basic residue at the peptide C-terminus favors the formation of multiply charged ions where at least one of the ionizing protons is sufficiently mobile in MS/MS experiments to trigger multiple amide bond ruptures along the peptide backbone, as depicted by the mobile proton model [7–9]. Thus, in most proteomic studies, tryptic peptides are usually efficiently sequenced by low-energy collision-induced dissociation (CID) of doubly charged precursor ions. Therefore, fragmentation rules governing low-energy dissociations of protonated peptides are of the utmost importance, as bioinformatic tools utilize these established MS/MS behaviors to perform the automated interpretation of mass spectra and to subsequently achieve efficient protein retrieval from database searches [10, 11].

The potential of the metalloendopeptidase Lys-N [12] in proteomics has recently been reported in the literature [13, 14]. This protease was found to be a good alternative to trypsin, providing efficient digestion at every predicted cleavage site [13], even under harsh sample handling

Electronic supplementary material The online version of this article (doi:10.1007/s13361-010-0022-7) contains supplementary material, which is available to authorized users.

Correspondence to: Christine Enjalbal; e-mail: enjalbal@univ-montp2.fr

Received: 13 July 2010
Revised: 15 September 2010
Accepted: 22 October 2010
Published online: 22 January 2011

conditions [14]. Combining Lys-N and trypsin proteolysis substantially enhanced protein coverage. Peptides issued from biological protein mixtures that were subjected to Lys-N treatments were sequenced by MALDI-TOF/TOF instruments operated under CID conditions [15, 16]. The MS/MS data recorded from singly charged ions exhibited a more extensive b-ion series compared to related tryptic peptides. Although the presented results were very convincing and promising, several reasons prompted us to prepare model peptides that mimic Lys-N digests to broaden the scope of the MS/MS experiments. First, only Lys-N proteolytic peptides that did not contain basic residues within their chains were studied, with the others being discarded prior to MS/MS analyses upon strong cation exchange chromatography [15]. Second, guanidination of the N-terminal lysine residue was also undertaken to modify peptide basicity and thereafter improve its fragmentation by promoting the formation of a single ion series [17]. In our opinion, such additional sample handling (a chromatographic step or chemical derivatization) limits the interest of Lys-N proteolysis compared to trypsin digestion. Moreover, only singly charged ions activated by CID (with MALDI-MS/MS and to a less extent with ESI-MS/MS) were investigated, whereas MS/MS experiments on ions exhibiting higher charge states (+2, +3) were commonly conducted with ESI-based mass spectrometers, especially in shotgun strategies [18].

The present study was undertaken using 15 model Lys-N peptides with masses ranging from 500 to 1700 g/mol, with the C-terminal position being occupied by any of the 20 proteinogenic amino acids. Some of the prepared sequences bear a basic residue within their polyamide chain (arginine, histidine), lysine being excluded. Besides, to cover all aspects of MS/MS experiments performed in proteomic bottom-up applications, these synthetic model peptides were studied with mass spectrometers displaying different ionization and mass analyzer configurations (ESI-QqTOF and MALDI-TOF/TOF) allowing both CID fragmentation and laser-induced dissociation (LID) [19, 20]. In the chosen mass range, the synthetic peptides exhibited both singly and doubly charged ions in ESI-QqTOF analyses, allowing comparison of the recorded peptide sequence information obtained (i) from the fragmentation of the singly charged precursors in all investigated MALDI and ESI MS/MS equipment and (ii) from the fragmentation of the singly charged compared to the doubly charged parent ion in ESI-QqTOF.

Experimental

Chemicals and Reagents

All solvents were of analytical grade. Methanol was purchased from Fisher and acetonitrile from Carlo Erba. The deionized water used in all the experiments was obtained using a Milli-Q system (Millipore, Milford, MA, USA). Hydrochloric acid (37%) was obtained from VWR.

The MALDI matrix (2,5 dihydroxybenzoic acid and α -cyano-4-hydroxycinnamic acid) was purchased from Fluka. A peptide calibration standard kit was purchased from Bruker Daltonics.

Syntheses

Peptides were synthesized on the IBMM peptide synthesis platform using the conventional methodology of solid-phase peptide synthesis using Fmoc chemistry [21], and purified by preparative LC/MS (autopurification system, Waters, Milford, MA, USA). Due to the chosen linker, all peptides were prepared as C-terminal carboxylic acids. More details are available in the Electronic Supplementary Material (ESM).

Mass Spectrometry

Mass spectrometry analyses were conducted on instruments located in the Laboratoire de Mesures Physiques of University Montpellier 2 (the IBMM instrument platform).

ESI mass spectra were recorded on a Q-TOF I mass spectrometer (Waters) fitted with an electrospray ion source. Data were acquired and processed with the Masslynx software. The mass spectrometer was calibrated in the positive-ion mode using 1% phosphoric acid in water/acetonitrile solution ($\text{H}_2\text{O}/\text{CH}_3\text{CN}$, 50/50, v/v). Data were acquired by the TOF analyzer at 1 acquisition/s from m/z 50 to m/z 2000 with a resolution of 5000. Depending on the sample, 100 acquisitions were summed to produce the final spectrum. Samples were dissolved in an $\text{H}_2\text{O}/\text{CH}_3\text{CN}$ (50/50, v/v) mixture and infused into the ESI source at a flow rate of 10 $\mu\text{l}/\text{min}$. Voltages were set at +2.0 kV for the capillary and adjusted for the sampling cone. The source was heated at 80 °C. Nitrogen was used as both nebulizing and desolvation gas. The latter was heated to 150 °C. In the MS/MS experiments, the second quadrupole analyzer was set to the rf mode, with argon used as collision gas. The collision energy was varied from 25 to 90 eV depending on the sample in order to get optimal fragmentation, as displayed in Table 1.

MALDI mass spectra were recorded on an Ultraflex III TOF/TOF instrument (Bruker Daltonics, Wissembourg, France) equipped with a LIFT capability. A pulsed Nd:YAG laser at a wavelength of 355 nm was operated at a frequency of 100 Hz (MS data) or 200 Hz (MS/MS data) with a delayed extraction time of 30 ns. The source was operated in the positive mode. Data were acquired with the Flex Control software and processed with the Flex Analysis software. A solution of the HCCA matrix in water/acetonitrile (70/30, v/v) at a concentration of 10 mg/ml was mixed with an equal amount of the peptide sample, and 0.5 μl of this solution was deposited onto the MALDI target according to the dried droplet procedure. After evaporating the solvent, the MALDI target was introduced into the mass spectrometer ion source. External calibration was performed

Table 1. Studied synthetic Lys-N peptides

| Peptide | Peptide sequence | (M+H) ⁺ (Da) | MALDI-MS/MS | | ESI-MS/MS | |
|-----------|------------------|-------------------------|--|---------|--|-----------------|
| | | | Laser fluence (%) to fragment MH ⁺ by | | Collision energy (eV) to fragment by CID | |
| | | | LID | LID/CID | Charge state +1 | Charge state +2 |
| 1 | KGFLC | 567.29 | 20 | 30 | 45 | nd |
| 2 | KAFLHG | 672.37 | 20 | 30 | 50 | 35 |
| 3 | KAFHLG | 672.37 | 20 | 30 | 50 | 30 |
| 4 | KAFLRG | 691.42 | 20 | 30 | 60 | 30 |
| 5 | KAFRLG | 691.42 | 20 | 30 | 60 | 30 |
| 6 | KYPFEAL | 867.45 | 20 | 30 | 55 | 25 |
| 7 | KMVNLHIQ | 982.54 | 20 | 30 | 65 | 40 |
| 8 | KADFPQLMV | 1048.54 | 20 | 30 | 70 | 35 |
| 9 | KEDFPQLMV | 1106.54 | 20 | 30 | 65 | 40 |
| 10 | KYIWLRAV | 1135.65 | 20 | 30 | 70 | 45 |
| 11 | KEAHMTIVFCSA | 1336.63 | 30 | 40 | 75 | 45 |
| 12 | KSWDTLCMAGRF | 1414.65 | 20 | 30 | 70 | 35 |
| 13 | KWFGMLADQATYN | 1544.71 | 30 | 40 | 65 | 25 |
| 14 | KATCGRVNSASMYDAI | 1686.79 | 30 | 40 | 90 | 35 |
| 15 | KTRYNGMGEQWDPD | 1696.73 | 25 | 35 | 90 | 35 |

nd not detected

with the commercial peptide mixture (Peptide Calibration Standard 2, Bruker Daltonics, Wissembourg, France).

MS data were acquired under the following MS conditions. An acceleration voltage of 25.0 kV (IS1) was applied for a final acceleration of 21.95 kV (IS2). The reflectron mode was used for the TOF analyzer (voltages of 26.3 kV and 13.8 kV). Mass spectra were acquired from 250 laser shots, the laser fluence being adjusted for each studied sample (laser fluence 1). Ions were detected over the mass range m/z 200–2000.

MS/MS data were acquired under the following conditions. An acceleration voltage of 8.0 kV (IS1) was applied for a final acceleration of 7.25 kV (IS2). The reflectron mode was used for the TOF analyzer (voltages of 29.5 kV and 13.9 kV). Mass spectra were acquired from 600 laser shots, with the laser fluence (given in Table 1) being adjusted for each studied peptide above the threshold for generating molecular ions (laser fluence 2 > laser fluence 1). MS/MS experiments were performed under LID conditions with the LIFT cell voltage parameters set at 19.0 kV (LIFT 1) and 3.2 kV (LIFT 2) for a final acceleration of 29.5 kV (reflector voltage), and the pressure in the LIFT cell was around 4×10^{-7} mbar. The precursor ion selector was set manually to the first monoisotopic peak of the molecular ion pattern for all analyses. For LID experiments, no collision gas was added (gas-off spectra). For LID/CID experiments (gas-on spectra), the same tuning was used but argon was allowed to fill the collision cell (to a pressure of around 2×10^{-7} mbar).

Results and Discussion

Design of Lys-N Proteolytic Peptides

We synthesized 15 peptides to cover a mass range from 500 Da (5 residues) to 1700 Da (16 residues), as shown in

Table 1. Peptides were designed to present an overall amino acid occurrence close to their natural relative abundance in proteins [22] (Figure S1 of the ESM). Except for lysine, which was only present at the N-terminal position, the five classes of residues—aliphatic (L, A, G, V, I), basic (R, H), acidic (E, D), aromatic (W, F, Y) and uncharged (N, Q, S, P, T, M, C)—were inserted anywhere within the polyamide chain. It has been demonstrated that some amino acids greatly influence the peptide fragmentation output, their side chains being involved in specific low-energy CID fragmentation reactions that are well documented [23–27]. Acid–base [23] and nucleophile–electrophile [28, 29] interactions provide various gas-phase chemistries involving both backbone and side-chain functional groups. For instance, basic sites participate in mobile proton sequestration [24–26] and salt bridge formation [30, 31], whereas nucleophiles trigger intramolecular attacks (backbone/backbone, backbone/side chain, side chain/side chain, cyclization, ...) [28, 29, 32]. A great deal of attention has thus been devoted to the influence on the fragmentation patterns of amino acids bearing a basic, acidic, nucleophilic, or electrophilic site, such as arginine [23–25, 31, 33], lysine [34, 35], histidine [24, 26, 36–38], proline [39–41], aspartic acid, and glutamic acid [23–27, 29]. These residues are all known to contribute to particular dissociation behaviors irrespective of the peptide sequence (synthetic or tryptic peptides). Thus, it would be interesting to probe the effects of these amino acids in N-terminated lysine containing peptides. In contrast to previously reported data concerning Lys-N peptides [15], two basic residues, arginine (pK_a of the guanidinium side chain 12.48) and histidine (pK_a of the imidazole side chain 6.04), were positioned within some of the studied peptides (4, 5, 10, 12, 14, and 15 for arginine-containing peptides and 2, 3, 7, and 11 for histidine-containing peptides). Some positional scanning was envisaged by moving the arginine or histidine

residue along the sequence to probe any difference in the MS/MS data. Proline, which triggers specific fragmentation pathways known as the “proline effect,” involving the peptide backbone cleavage N-terminal to this amino acid [40], was also included in the prepared sequences (6, 8, 9, 12, 15). Also, one acidic amino acid, aspartic acid (D) or glutamic acid (E), was positioned next to lysine in order to investigate the potential formation of a salt bridge between the two basic/acidic side chains on polyamide dissociation (8 and 9, which differ only by the substitution of the N-terminal dipeptide KA for KE, respectively). Finally, although peptides were more easily prepared with a C-terminal amide function in solid-phase strategies [17], all peptide chains were terminated by a C-terminal carboxylic acid to mimic Lys-N proteolysis and to avoid the loss of ammonia from the peptide backbone [42], which could compete with the loss of such molecule from the lysine side chain.

MALDI vs ESI MS/MS Experiments on Lys-N Proteolytic Peptides

As displayed in Table 1, four sets of experiments were performed: CID fragmentation of singly charged and doubly charged ions in ESI-MS/MS, and dissociation of singly charged species in MALDI-MS/MS but under two precursor ion activations (LID or LID combined with CID). Except for the smallest peptide, which exhibited only a singly charged ion in the corresponding ESI-QTOF mass spectrum, all other sequences provided the expected singly and doubly charged species suitable for all dissociation experiments. It should be noted that the longest chains were difficult to sequence when the singly charged precursor ion was selected, limiting the frame of the studies to peptides with masses below 2000 Da. The presence of a basic residue that could sequester the ionizing proton of singly charged ions is known to affect the extent of peptide backbone cleavage [24, 25, 43]. A comparison of the fragmentation behaviors of singly charged and doubly charged arginine- and (to a lesser extent) histidine-containing Lys-N proteolytic peptides would provide an insight into the effect of a localized versus a mobile proton without being biased by the presence of an arginine residue only at the C-terminal position (tryptic peptides), as encountered in large-scale proteomics experiments.

To ascertain the clarity of Lys-N proteolytic peptide MS/MS spectra that were reported to exhibit distinguishable b ion series [15], detected fragment ions with S/N > 5 were classified into three categories whatever their charge state: N-terminal fragment ions (b and a ion series), C-terminal fragment ions (y ion series), and all other recorded ions. In this third group, non-direct sequence ions [32] that correspond to ions which were not expected from the backbone fragmentation model and/or were not directly related to the primary sequence, such as internal ions, immonium ions, scrambled ions, and

neutrol losses, were not distinguished. This category also included some ions that were not identified. To deduce general trends from the recorded data, the overall intensity of each type of signal—a and b ions, y ions, and others—was normalized by adding the relative abundances of all ions belonging to that category and then dividing by the sum of all relative ion abundances. The relative abundance of the parent ion was discarded, as ESI and MALDI data were not acquired under the same MS/MS process. The relative abundances of C-terminal fragment ions (y ion series), N-terminal fragment ions (a and b ion series) and other non-direct sequence ions were plotted for each peptide sequence for all of the investigated tandem mass spectrometry experiments. Since the overall ion abundance could be mainly generated by a few intense signals or derive from a large number of weak signals, the number of detected ions in each category was also indicated in the schemes to give an account of the ion population.

MALDI-TOF/TOF

First of all, LID and LID combined with CID in MALDI experiments gave the same MS/MS spectra as previously observed for other synthetic peptides [20] (Fig. S2 of the ESM). The results recorded for MALDI tandem mass spectrometry upon LID/CID activation are displayed in Fig. 1. Three representative MS/MS spectra are reported in Fig. 2 (peptides 5, 8, and 11). Overall, b ions were more abundant than y ions, as expected from literature data. Indeed, 13/15 peptides showed this behavior. The best example is displayed in Fig. 2c for peptide 11, which exhibited a highly abundant N-terminal ion series from b11 to b4 that allowed the great majority of the sequence to be deduced. However, the clear detection of the N-terminal b ion series was not always so obvious. Although the presence of non-direct sequence fragment ions was usually observed, interferences with the expected backbone ions became significant when intense high-mass signals were produced. Moreover, some very abundant high-mass y ions were also recorded. Thus, the ions that dominated the MS/MS spectra were not necessarily b ions. Deducing the peptide sequence from the recorded data was not a straightforward task, as illustrated in Fig. 2a and b for peptides 5 and 8, respectively.

ESI-Q/TOF

The same peptides were subjected to low-energy CID experiments, selecting as a precursor ion the singly protonated species as well as the doubly charged ions. The collision-energy values that were applied for each MS/MS analysis are indicated in Table 1. The recorded fragment ions were classified into the three categories previously discussed

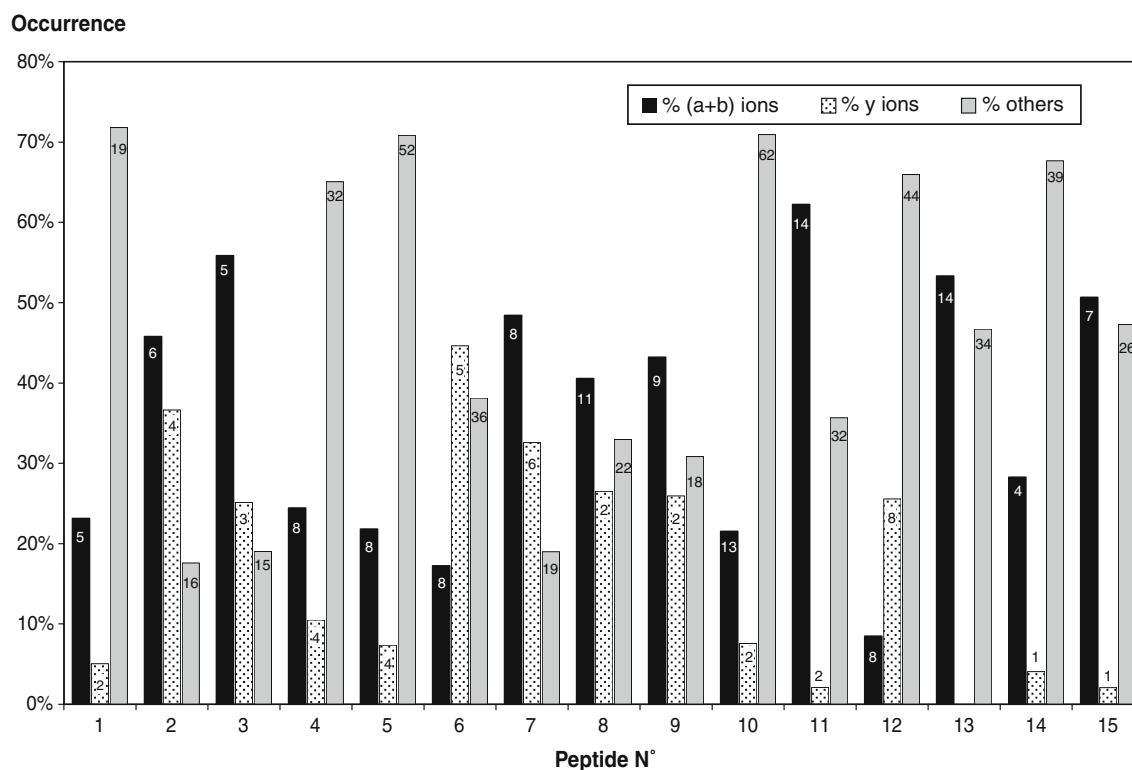


Figure 1. Occurrence of peptide fragment ions in LID/CID MALDI-TOF/TOF experiments (the number of fragment ions detected is indicated for each population)

(a and b ions, y ions, and others), as represented in Figs. 3 and 4, which relate to the CID of the $(M + H)^+$ and $(M + 2H)^{2+}$ precursor ions, respectively. From a general point of view, two major facts can be deduced from these MS/MS data. Firstly, provided that singly charged ions were dissociated, the fragmentation response was very similar in terms of ion type whatever the mass spectrometry technique used. Indeed, the studied peptides followed the same trends in Fig. 1 (MALDI-TOF/TOF) and Fig. 3 [ESI-Q/TOF on the $(M+H)^+$ precursor ion]. Secondly, the striking difference between the two sets of ESI experiments derives from the enhanced production of immonium ions in the case of doubly charged ion dissociation. Such fragmentation behavior increased the population of non-direct sequence ions, fragment ions related to truncated sequences (internal ions), and fragment ions produced from neutral losses, which were globally detected at the same level. As observed previously in MALDI-TOF/TOF, the b ions did not necessarily dominate the ESI-MS/MS spectra. The MS/MS mass spectra of peptides 5, 8 and 11, shown in Fig. 2 for MALDI-TOF/TOF experiments, are displayed in Figs. 5 and 6 for the CID of the $(M + H)^+$ and $(M + 2H)^{2+}$ precursor ions, respectively.

The choice of the precursor-ion charge state that undergoes the dissociation events greatly influences the output of the analysis, much more so than the mass spectrometry technique used (i.e., MALDI-TOF/TOF vs ESI-Q/TOF). Thus, the fragmentation features from singly

charged precursor ions and from doubly charged ions are discussed separately below.

Effect of Precursor Ion Charge State on Lys-N peptide MS/MS Behavior

LID vs CID of the $(M+H)^+$ Precursor Ion

Deviation from the reported ideal MS/MS behavior of singly charged ions—the production of abundant b ions—was attributed to the presence of arginine, proline and histidine in the sequence, as well as to the presence of lysine at the N-terminal position. In order to discuss more detailed assignments, the influence of each of these residues in the Lys-N peptides is described and compared with the specific MS/MS behaviors that they trigger in model synthetic peptides [23–27]. The effect of acidic residues, namely aspartic acid and glutamic acid, is also discussed.

Arginine (Peptides 4, 5, 10, 12, 14, 15)

Due to the strong binding of the charge to the guanidinium side chain [23, 24], the contribution of arginine to multiple fragmentation pathways was by far the most striking structural character, as it provided highly populated MS/MS spectra devoid of a clear b ion series (Fig. 2a vs 2b). Increased formation of y ions was observed for all arginine-containing peptides. Some of these y ions lost ammonia

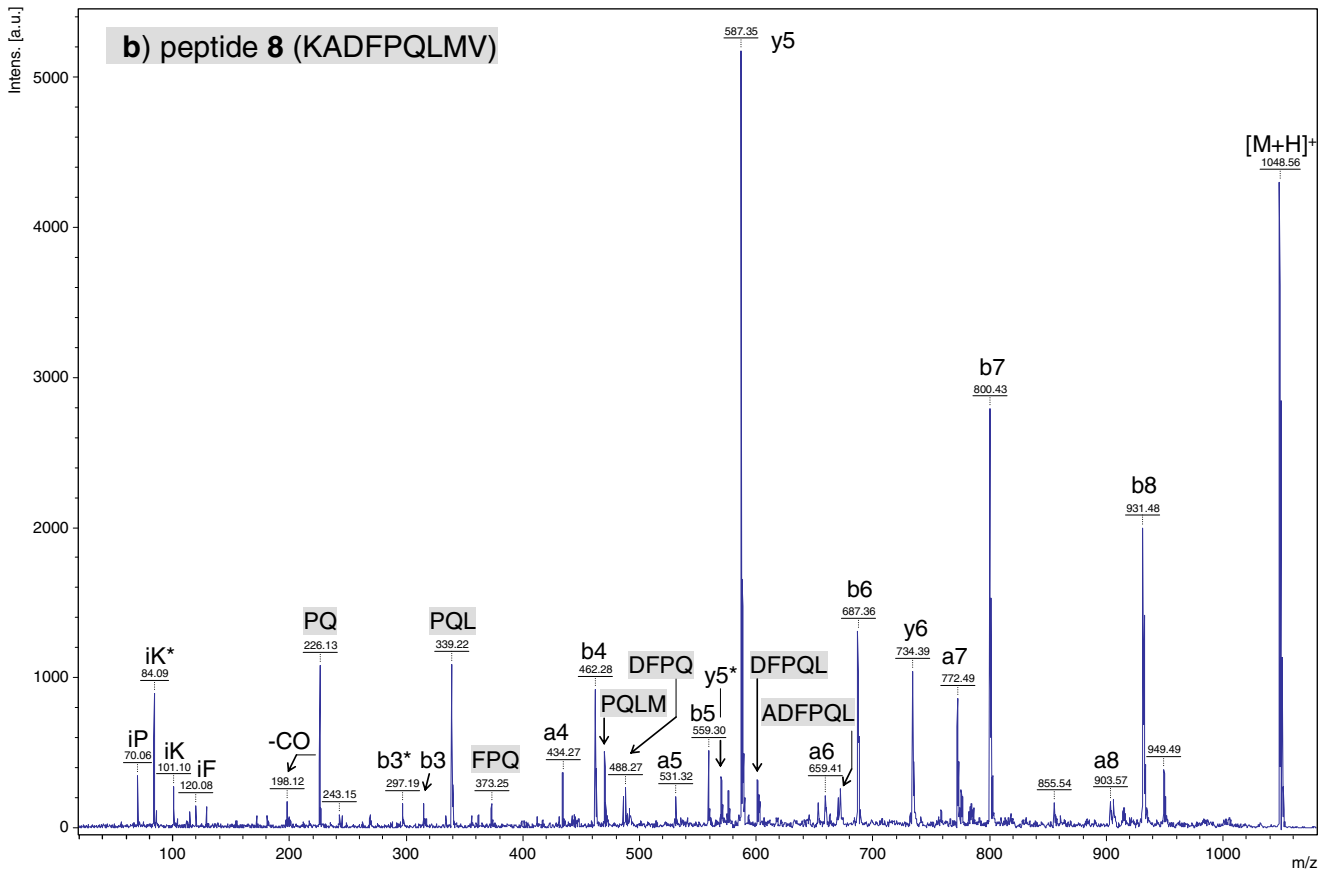
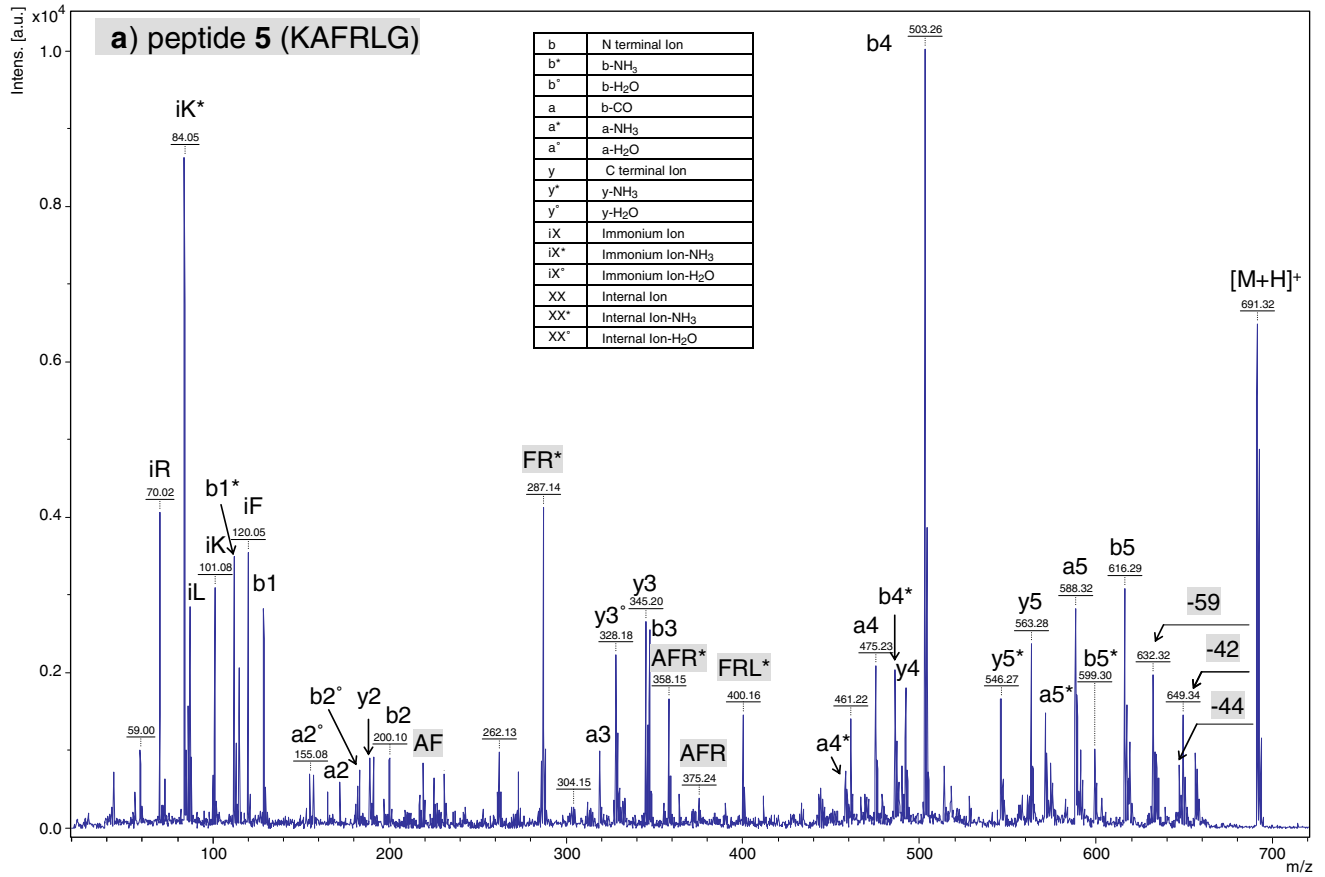


Figure 2. Continued on next page

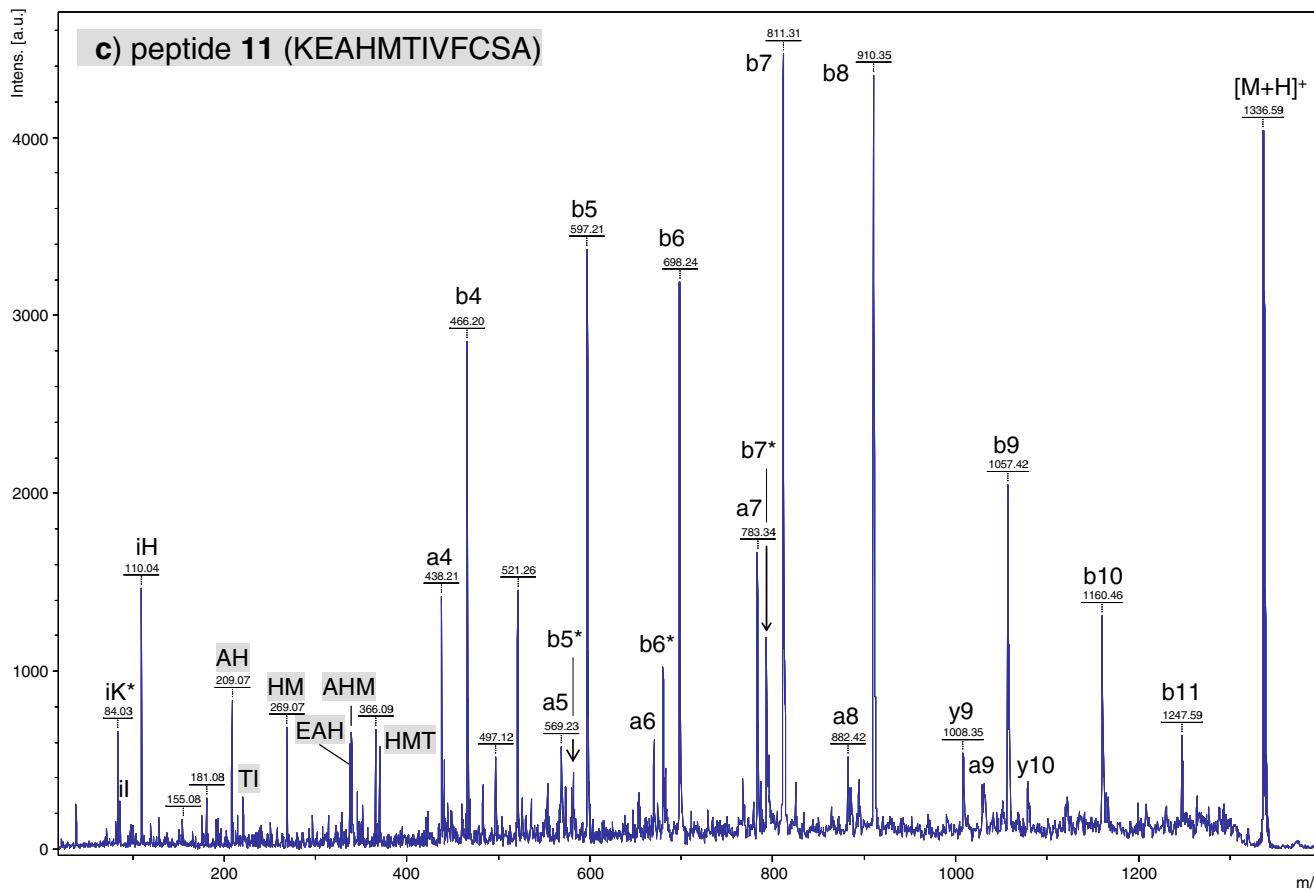


Figure 2. LID/CID MALDI MS/MS spectra of (a) peptide 5 (KAFRLG), (b) peptide 8 (KADFPQLMV), and (c) peptide 11 (KEAHMTIVFCSA)

Occurrence

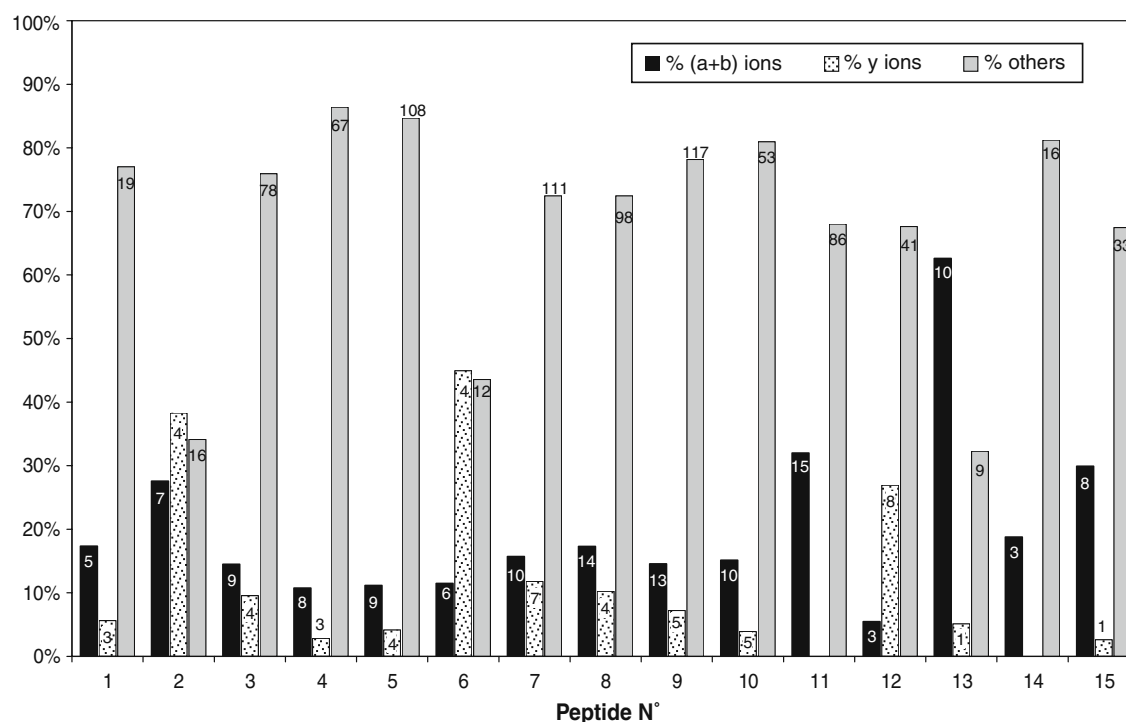


Figure 3. Occurrence of peptide fragment ions in ESI-QqTOF experiments from CID of the $(M + H)^+$ precursor ion (the number of fragment ions detected are indicated for each population)

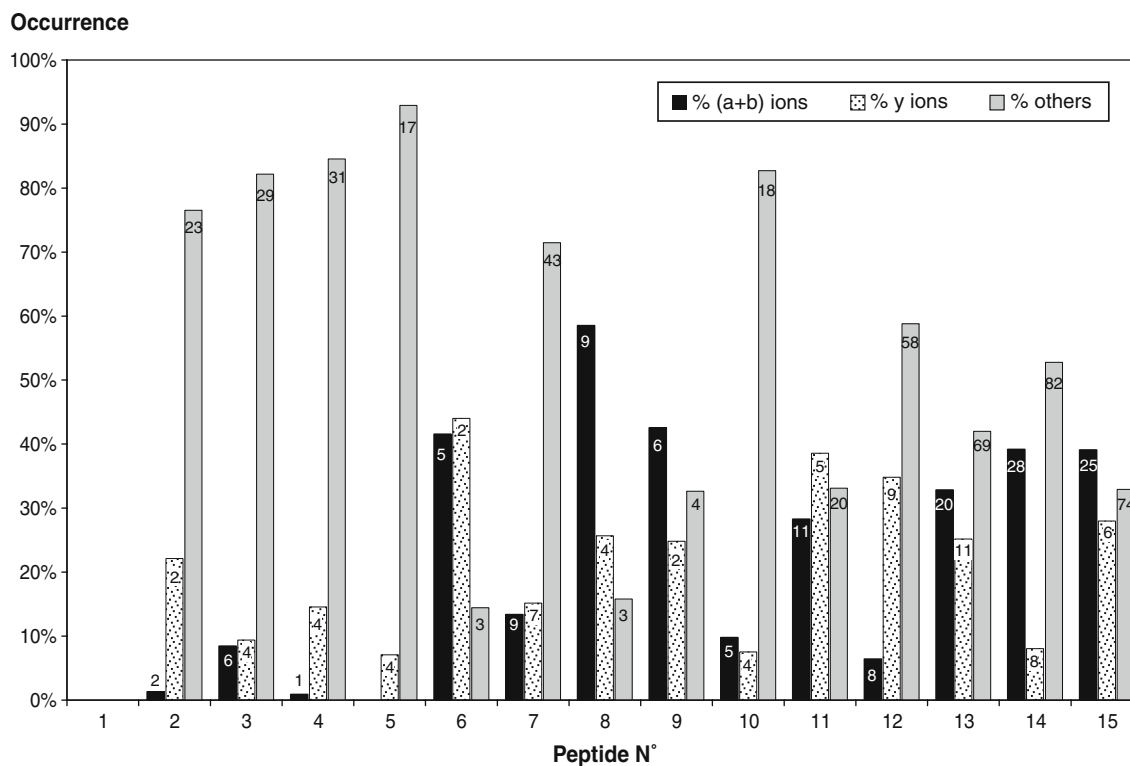


Figure 4. Occurrence of peptide fragment ions in ESI-QqTOF experiments from CID of the $(M + 2H)^{2+}$ precursor ion (the number of fragment ions detected are indicated for each population)

from the arginine side chain, duplicating the C-terminal signals and thus spreading the sequence information across multiple ions. Surprisingly, only in LID/CID MALDI-MS/MS spectra, two other neutral losses of molecules weighing 42 and 59 Da from the $(M + H)^+$ ion were seen solely with arginine-containing peptides (Fig. 2a, peptide 5). The protonated guanidinium may be involved in such fragmentations, as depicted in Scheme 1. The same results were observed whatever the position of arginine within the Lys-N peptide chain (Fig. S3 of the ESM, peptide 10). Such dissociation behavior has not been observed previously in the literature, irrespective of the nature of the peptide (i.e., for tryptic and nontryptic peptides [20, 23, 24]), indicating a specific fragmentation behavior of arginine-containing Lys-N peptides in MALDI-TOF/TOF experiments. The loss of 42 and/or 59 Da from the protonated precursor ion constituted a mass tag indicating the presence of arginine in the Lys-N peptide chain. Taking into account that this residue may cause difficulties during sequence assignment, the detection of the 42/59 Da loss could trigger the corresponding MALDI-MS/MS spectrum to be discarded, instead of relying on a chromatographic step as reported in the literature [15].

Proline (Peptides 6, 8, 9, 15)

As expected from literature data [39–41], as long as a proline residue was present in the peptide, whatever its

location within the sequence, the cleavage of the amide bond N-terminal to this residue was systematically observed. As already described, the presence of an acidic residue next to the proline (like the NP dipeptide in peptide 15) favored the rupture of the proline amide bond [44, 45]. In contrast to literature data on Lys-N peptides [15], the preferential formation of the corresponding y ion was not systematic in that situation; the production of internal sequences predominated (Fig. S4 of the ESM, peptides 6, 9). Both peptides showed several proline-containing internal ions, with this residue being positioned at the N-terminus, but only one also exhibited the corresponding y ion. For peptides containing both arginine and proline (peptide 15 in Fig. S5 of the ESM), the sequence identification from the recorded MS/MS spectra was even more difficult to interpret, as the effects of arginine and proline were combined (more y ions than b ions, loss of neutrals, a very abundant y ion due to the facile rupture of the proline amide bond, internal sequence ions).

Histidine (Peptides 2, 3, 7, 11)

Histidine, through its protonated side chain, directs a specific backbone cleavage that enhances the production of fragment ions terminating at this residue, especially when histidine is close to the C-terminus [24, 26, 36–38]. In the studied Lys-N peptides, the contribution of

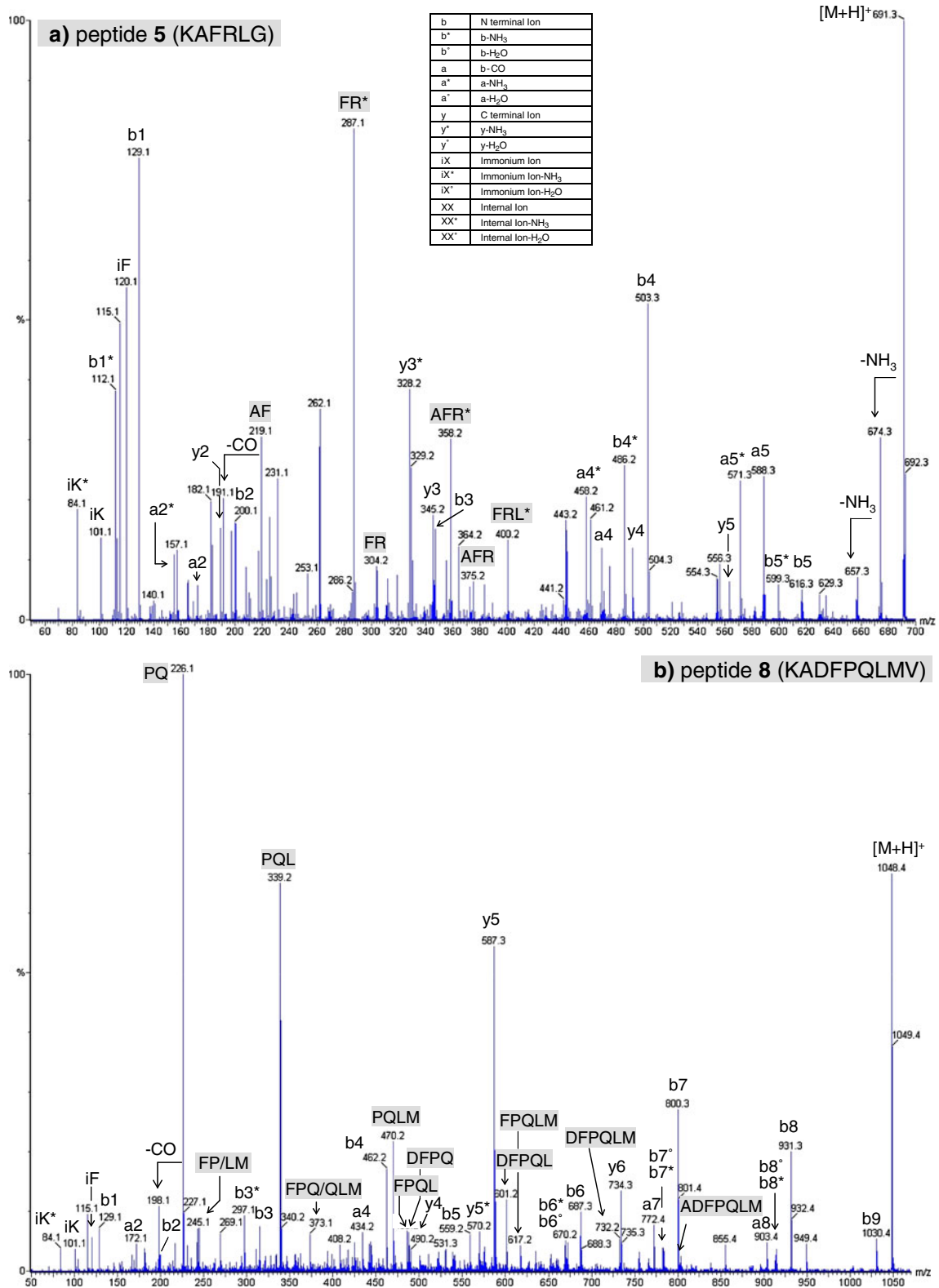


Figure 5. Continued on next page

histidine was found to be very similar to that of proline. A lot of non-direct sequence ions related to internal sequences were obtained regardless of the ionization method and precursor ion activation (MS/MS spectra of

peptide **11** shown in Figs. 2c, 5c and 6c) that was used. Just as in the literature data [24, 26, 36–38], histidine was found to trigger the rupture of the amide bonds next to it, or even peptide bonds in the close vicinity. The

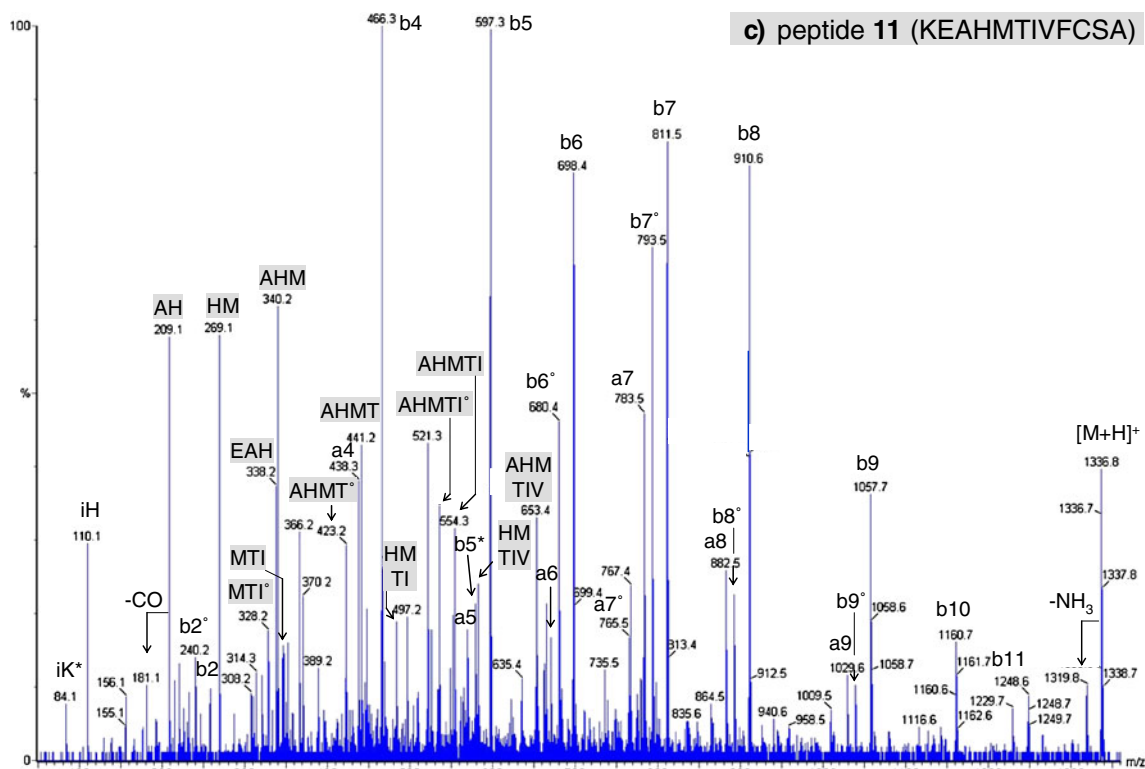


Figure 5. ESI MS/MS spectra upon CID of the $(M + H)^+$ precursor ion of (a) peptide 5 (KAFRLG), (b) peptide 8 (KADFPQLMV), and (c) peptide 11 (KEAHMTIVFCSA)

production of the corresponding y ion by cleavage of the amide bond N-terminal to histidine was found to be quite abundant in most of the peptides, especially the small ones (peptides 2, 3, 7, see Figs. S6a and S6b of the ESM for peptide 7). Also, a highly abundant immonium ion at m/z 110 was observed for all histidine-containing peptides, with this low mass signal being very intense in MALDI-TOF/TOF mass spectra.

Lysine at the N-Terminal Position (All Peptides)

The lysine immonium ion at m/z 101 and its related ion at m/z 84 obtained through the loss of ammonia were also very frequently produced. The ion detected at m/z 129 was produced by an intramolecular nucleophilic attack of the ϵ -NH₂ moiety on the first C-terminal amide bond [28]. This ion is sometimes referred to as a “b1 ion.” Such feature ions generated from the lysine residue (m/z 84/101/129) were quite abundant in nearly all recorded LID/CID MALDI MS/MS spectra. These ions were also detected in ESI-MS/MS experiments, but their signals were not very intense, as they were produced together with other immonium ions (from H, F, L, I, ...). The competitive production of the b2 ion through the oxazolone pathway was very scarce. Of the 15 MALDI-MS/MS recorded spectra, only eight produced the b2 ion, which was not abundant. Similarly, the same eight peptides produced the b2 ion in both sets of CID ESI-

MS/MS experiments, while four other sequences also produced the b2 ion, although only from singly charged or doubly charged precursor ion dissociation. In order to assess if this behavior is specific to Lys-N peptides, we performed a comparison between tryptic and nontryptic sequences that had already been studied in the laboratory [16]. Upon LID/CID dissociation, peptides that contain lysine inside their chains were found to generate the same feature ions (m/z 84/110/129, with the latter generally being the most abundant) as Lys-N peptides. However, tryptic peptides with lysine at the C-terminal position did not produce these ions. Such behavior is illustrated in the ESM (Fig. S7). Overall, the lysine tag ions that were clearly detected in MALDI-TOF/TOF spectra (m/z 84/110/129) provided unambiguous information on lysine located anywhere within the sequence (nontryptic peptides), except at the C-terminal position (tryptic peptides).

Acidic Amino Acids (Peptides 6, 8, 9, 11, 12, 13, 14, 15)

Selective cleavage of the amide bond adjacent to the acidic residues (C-side) was reported in the literature for protonated peptides with ionizing protons sequestered at basic sites [25, 26]. In the studied Lys-N peptides, the contributions of aspartic acid and glutamic acid were minor compared to those of basic residues. Except when

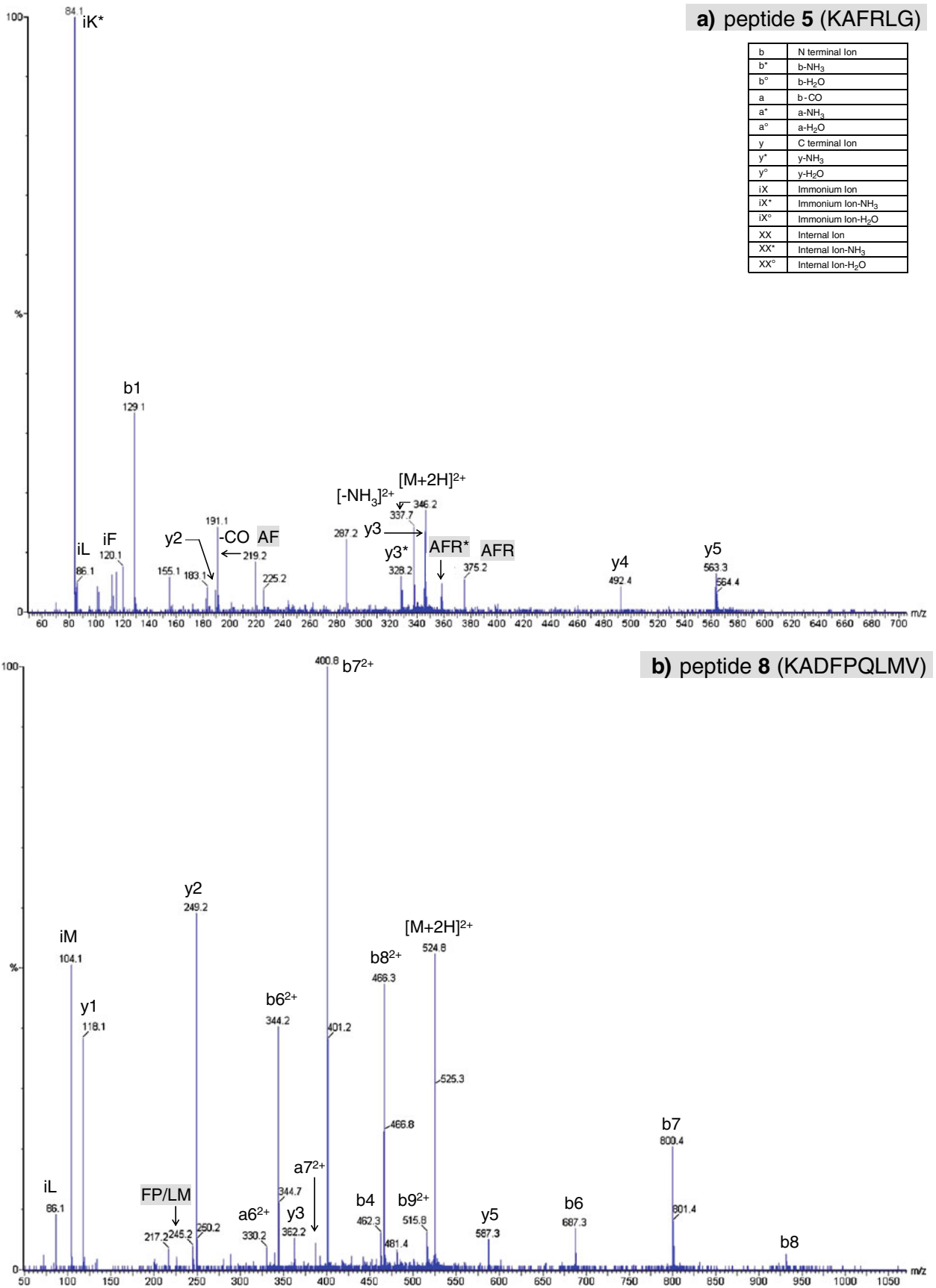


Figure 6. Continued on next page

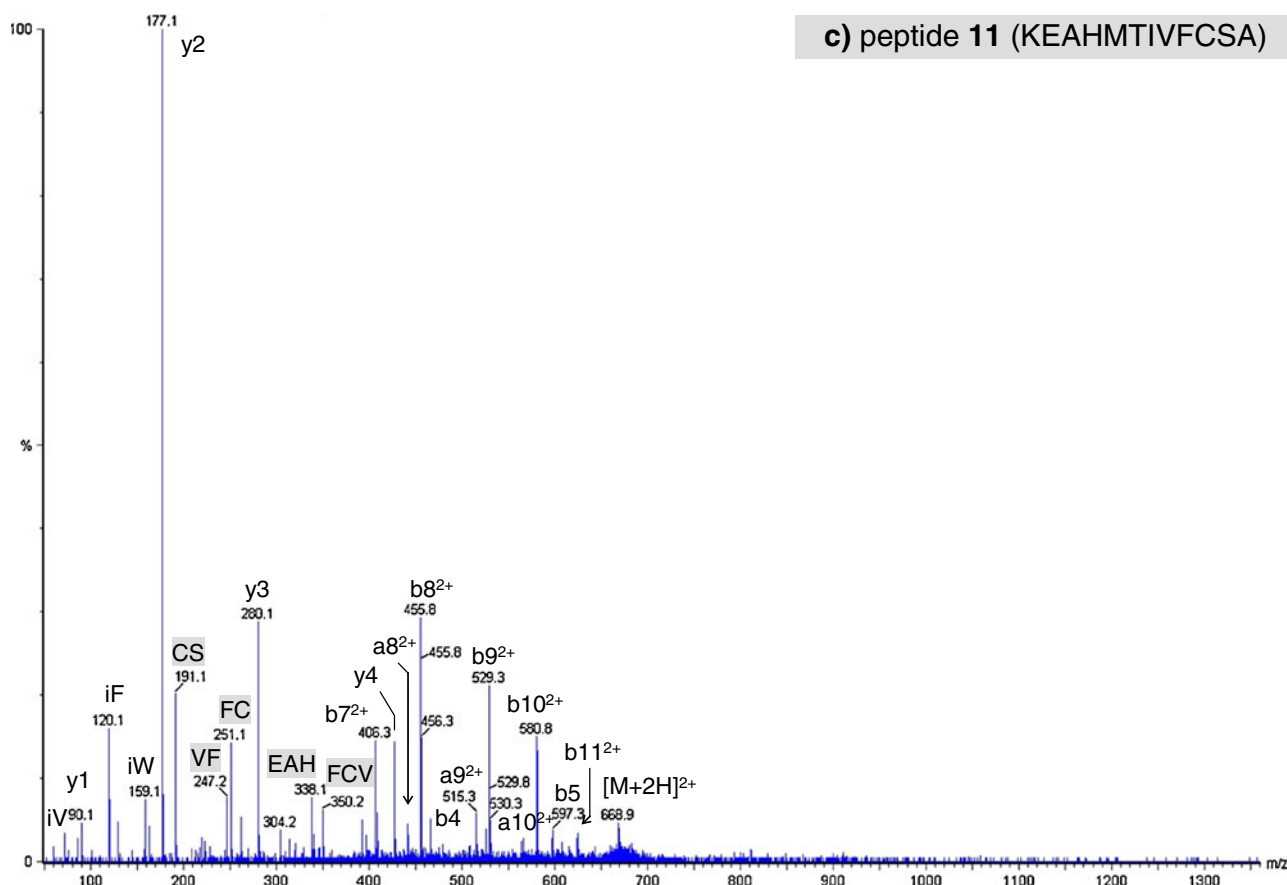


Figure 6. ESI MS/MS spectra upon CID of the $(M + 2H)^{2+}$ precursor ion of (a) peptide 5 (KAFRLG), (b) peptide 8 (KADFPQLMV), and (c) peptide 11 (KEAHMTIVFCSA)

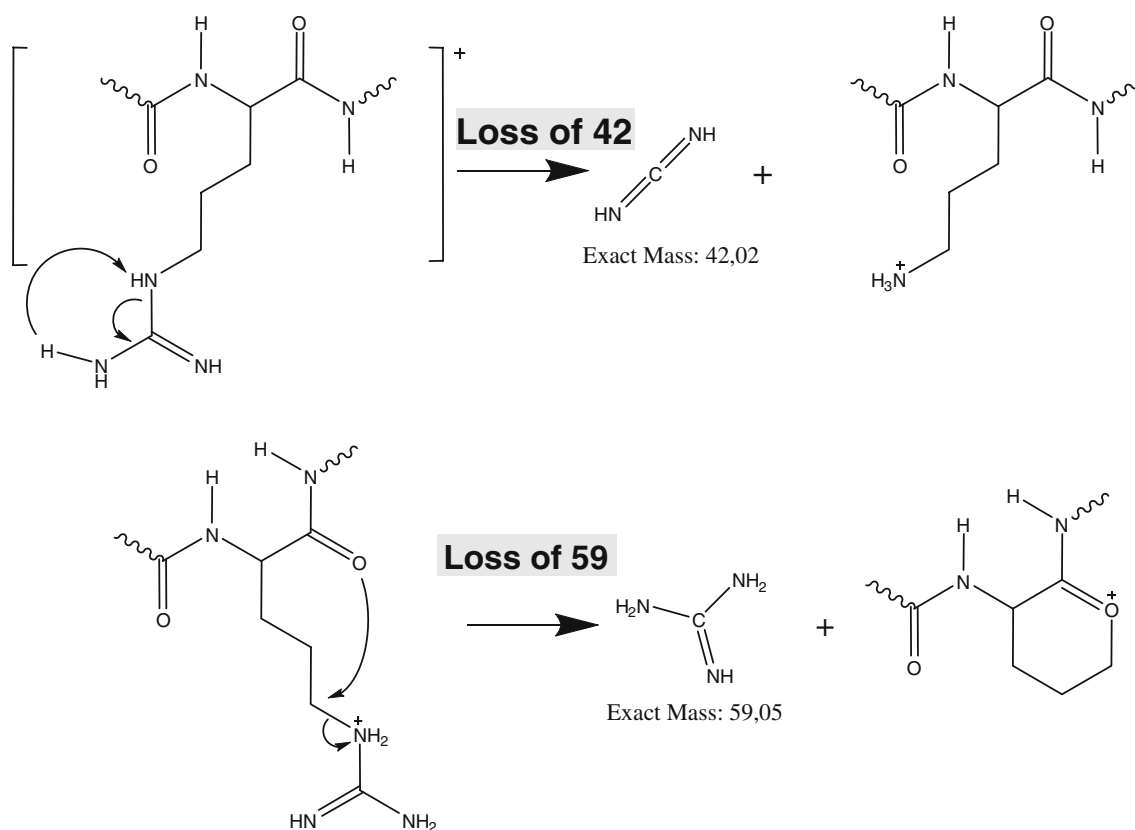
located next to proline in peptide 15, which increased the lability of the DP peptide bond [44, 45], no specific fragmentation behavior was observed. Peptides 8 and 9 (which differ only by the substitution of the N-terminal dipeptide KA for KE) behaved similarly, indicating that there was no influence on the fragmentation pattern of a potential salt bridge between the adjacent basic and acidic side chains.

In summary, nearly half of the studied Lys-N peptides (peptides 1, 4, 5, 10, 12, 14, 15) exhibited a lot of non-direct sequence fragment ions upon MALDI-TOF/TOF and ESI-QqTof, which accounted for the most intense group of signals that hamper facile sequence reading (Figs. 1, 3 and 4). Among all of the ions that were clustered in that category, low-mass ions issued from the N-terminal lysine residue (m/Z 84/101/129) largely accounted for the non-direct sequence ion intensity. Proline was the second most important contributor to this class of non-direct sequence ions, and produces a lot of internal sequences. Finally, other non-direct sequence ions, which were largely related to the truncated part of the peptide chain that was not triggered by proline but mainly by histidine, together with the loss of ammonia and guanidine from the arginine side chain, also contributed to this populated category that may compli-

cate MS/MS data deciphering. It should be noted that some sequence scrambling was also observed in a few cases, but this only led to high-mass b type ions.

CID of $(M+H)^+$ vs $(M+2H)^{2+}$ Precursor Ions

According to the mobile proton model [7, 8], ionizing protons must be transferred intramolecularly from basic sites to amide bonds along the peptide backbone in order to initiate the production of b and y ions. Thus, nonselective peptide backbone fragmentations occur when the number of ionizing charges at least equals the number of arginines (which are known to strongly sequester the proton) [23, 24]. In general, doubly charged tryptic peptides are dissociated more easily than the singly charged species since arginine is only located at the C-terminus, so one of the two ionizing protons is therefore available to trigger amide bond fragmentations. In the studied Lys-N peptides, arginine is not restricted at the C-terminus position, and the impact of this residue on ionizing proton mobility is thus crucial. As discussed previously, the loss of ammonia and water from a/b/y fragment ions increased in the case of doubly charged precursor ion dissociation. This behavior was reinforced



Scheme 1. Fragmentation of protonated guanidinium for arginine-containing peptides observed in LID/CID MALDI-TOF/TOF experiments

when an arginine residue was present. Another difference came from the detection of small y ions (such as y_1 , y_2 , and y_3) that were never (or were very scarcely) seen from $(M + H)^+$ dissociation but were very often observed from the doubly protonated parent ion. Moreover, the b ion series that was not systematically detected from the singly charged precursor ion was more thoroughly produced from the $(M + 2H)^{2+}$ dissociation; these N-terminal ions were mainly produced as doubly charged ions. The presence of an arginine in the vicinity of the C-terminal part of the studied peptide prevented such massive production of doubly charged b ions, thus increasing the detection of singly charged b and y ions (Fig. S3 of the ESM). Feature ions of lysine residues (at m/z 84/101/129) were not systematically produced from the doubly protonated parent ions. At the same time, the detection of b_2 ion was observed more frequently. Finally, the effect of proline was very limited, the dissociation of the amide bond N-terminal to this amino acid never dominated the MS/MS spectra of proline-containing peptide $(M + 2H)^{2+}$ ions. In general, a high number of internal sequences triggered by the presence of proline and histidine residues—and, more surprisingly, by some other residues—were detected, increasing the population of non-direct sequence ions. These b -type ions were much more abundant than those detected from

singly charged parent ions in both MALDI and ESI MS/MS dissociation experiments of $(M + H)^+$ ions.

Comparison of Lys-N Peptide Sequence Coverage

For each peptide, the detected b/y ions were combined in order to assess the peptide sequence tags that were unambiguously identified with each of the investigated techniques, as depicted in Scheme 2. The sequencing efficiency was much more peptide dependent than influenced by the MS/MS conditions. As expected from the mobile proton model, peptides bearing an arginine and/or a proline were better fragmented from the doubly charged precursor ion in ESI-QqTOF analyses. For other samples, choosing singly charged parent ion dissociation in MALDI or ESI-QqTOF experiments was found to be more judicious, as fewer non-direct sequence ions were produced, leading to less-crowded MS/MS spectra and thus more easily detected b - or y -ion series. Also, knowledge of the N-terminal amino acid (i.e., lysine) allowed the nature of the N-terminal dipeptide to be determined through the detection of the b_2 ion, even when the complementary y_{n-1} was not produced. The production of rather abundant low-mass ions derived from the lysine immonium ion (at m/z 84/101/129) confirmed the presence of this residue at the N-terminal position, thus providing a tool for identifying any peptides



Scheme 2. Lys-N peptide sequence coverage in all MS/MS experiments

from the N-terminal part of the studied protein in a proteolytic Lys-N digest (lysine is just one of over 20 naturally occurring amino acids that can occur at this location).

Conclusion

The MS/MS behavior of Lys-N mimicking peptides under various conditions showed that the influence of basic residues (arginine and to a lesser extent histidine) as well as proline on the sequencing efficiency must be considered. The massive production of b ions was not systematically retrieved, whichever dissociation technique (LID, CID) was used. Indeed, the peptide sequence was found to greatly influence the fragmentation output, especially through the selection of the singly or doubly protonated precursor.

Whereas arginine-free Lys-N peptides could be easily sequenced via MALDI LID/CID experiments, ESI-QqTOF analyses allowing the dissociation of the doubly charged precursor ion were preferred for arginine-containing chains. For the other peptides, this method more frequently yielded competitive fragmentation pathways leading to abundant non-direct sequence ions such as immonium ions, internal sequences, scrambled ions, and the loss of neutral molecules from side chains. Less-crowded MS/MS spectra with more intense sequence signals (a/b and y ions) were thus obtained with MALDI-TOF/TOF compared to ESI-QqTOF for Lys-N peptides that did not bear arginine. Moreover, the detection of specific low-mass ions with MALDI-TOF/TOF indicated the presence of a lysine residue at the N-terminal position, allowing Lys-N peptides produced during protein digestion from the native N-terminal protein part to be identified.

Acknowledgments

We thank G. Fromenty for participating in the ESI-QqTOF experiments.

References

- Mann, M., Aebersold, R.: Mass spectrometry-based proteomics. *Nature*. **422**, 198–207 (2003)
- Olsen, J.V., Ong, S.-E., Mann, M.: Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics*. **3**, 608–614 (2004)
- Baldwin, M.A.: Protein identification by mass spectrometry. *Mol. Cell Proteomics*. **3**, 1–9 (2004)
- Perkins, D.N., Pappin, D.J., Creasy, D.M., Cottrell, J.S.: Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. **20**, 3551–3567 (1999)
- Frank, A.M., Pevzner, P.A.: PepNovo: de novo peptide sequencing via probabilistic network modeling. *Anal. Chem.* **77**, 964–973 (2005)
- Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., Lajoie, G.: PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass. Spectrom.* **17**, 2337–2342 (2003)
- Dongré, A.R., Jones, J.L., Somogyi, A., Wysocki, V.H.: Influence of peptide composition, gas-phase basicity and chemical modification on fragmentation efficiency: evidence for the mobile proton model. *J. Am. Soc. Mass. Spectrom.* **118**, 8365–8374 (1996)
- Paizs, B., Suhai, S.: Fragmentation pathways of protonated peptides. *Mass. Spectrom. Rev.* **24**, 508–548 (2005)
- Boyd, R., Somogyi, A.: The mobile proton hypothesis in fragmentation of protonated peptides: a perspective. *J. Am. Soc. Mass. Spectrom.* **21**, 1275–1278 (2010)
- Kapp, E.A., Schutz, F., Reid, G.E., Eddes, J.S., Moritz, R.L., O'Hair, R. A., Speed, T.P., Simpson, R.J.: Mining a tandem mass spectrometry database to determine the trends and global factors influencing peptide fragmentation. *Anal. Chem.* **75**, 6251–6264 (2003)
- Baudet, M., Ortet, P., Gaillard, J.-C., Fernandez, B., Guérin, P., Enjalbal, C., Subra, G., de Groot, A., Barakat, M., Dedieu, A., Armengaud, J.: Proteomic-based refinement of *Deinococcus deserti* genome annotation reveals an unwonted use of non-canonical translation initiation codons. *Mol. Cell Proteomics*. **9**, 415–426 (2010)
- Nonaka, T., Hashimoto, Y., Takio, K.: Kinetic characterization of lysine-specific metalloendopeptidases from *Grifola frondosa* and *Pleurotus ostreatus* fruiting bodies. *J. Biochem.* **124**, 157–162 (1998)
- Hohmann, L., Sherwood, C., Eastham, A., Peterson, A., Eng, J.K., Eddes, J. S., Shteynberg, D., Martin, D.B.: Proteomic analyses using *Grifola frondosa* metalloendopeptidase Lys-N. *J. Proteome. Res.* **8**, 1415–1422 (2009)
- Taouatas, N., Heck, A.J.R., Mohammed, S.: Evaluation of metalloendopeptidase Lys-N protease performance under different sample handling conditions. *J. Proteome. Res.* **9**, 4282–4288 (2010)
- Taouatas, N., Drugan, M.M., Heck, A.J.R., Mohammed, S.: Straightforward ladder sequencing of peptides using Lys-N metalloendopeptidase. *Nat. Methods* **5**, 405–407 (2008)
- Boersema, P.J., Taouatas, N., Altelaar, A.F., Gouw, J.W., Ross, P.L., Pappin, D.J., Heck, A.J., Mohammed, S.: Straightforward and de novo peptide sequencing by MALDI-MS/MS using a Lys-N metalloendopeptidase. *Mol. Cell Proteomics*. **8**, 650–660 (2009)
- Carabetta, V.J., Li, T., Shakya, A., Todd, M.G., Cristea, I.M.: Integrating Lys-N proteolysis and N-terminal guanidination for improved fragmentation and relative quantification of singly-charged ions. *J. Am. Soc. Mass. Spectrom.* **21**, 1050–1060 (2010)
- Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D. R., Garvik, B.M., Yates III, J.R.: Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **17**, 676–682 (1999)
- Mouls, L., Aubagnac, J.-L., Martinez, J., Enjalbal, C.: Low energy peptide fragmentations in ESI-Q-ToF type mass spectrometers. *J. Proteome. Res.* **6**, 1378–1391 (2007)
- Shenar, N., Sommerer, N., Martinez, J., Enjalbal, C.: Comparison of LID versus CID activation modes in tandem mass spectrometry of peptides. *J. Mass. Spectrom.* **44**, 621–632 (2009)
- Stewart, J.M., Young, J.D.: Solid phase peptide synthesis, 2nd edn. Pierce Chemical Company, Rockford, IL (1984)
- Swiss Institute of Bioinformatics: Swiss-Prot archive. <http://www.expasy.ch/sprot/>
- Tsapraillis, G., Nair, H., Somogyi, A., Wysocki, V.H., Zhong, W., Futrell, J.H., Summerfield, S.G., Gaskell, S.J.: Influence of secondary structure on the fragmentation of protonated peptides. *J. Am. Chem. Soc.* **121**, 5142–5154 (1999)
- Wysocki, V.H., Tsapraillis, G., Smith, L.L., Breci, L.A.: Mobile and localized protons: a framework for understanding peptide dissociation. *J. Mass. Spectrom.* **35**, 1399–1406 (2000)
- Tsapraillis, G., Somogyi, A., Nikolaev, E.N., Wysocki, V.H.: Refining the model for selective cleavage at acidic residues in arginine-containing protonated peptides. *Int. J. Mass. Spectrom.* **195/196**, 467–479 (2000)
- Huang, Y., Wysocki, V.H., Tabb, D.L., Yates III, J.R.: The influence of histidine on cleavage C-terminal to acidic residues in doubly protonated tryptic peptides. *Int. J. Mass. Spectrom.* **219**, 233–244 (2002)
- Molesworth, S., Osburn, S., Van Stipdonk, M.J.: Influence of amino acid side chains on apparent opening of cyclic b₅ ions. *J. Am. Soc. Mass. Spectrom.* **21**, 1028–1036 (2010)
- O'Hair, R.A.J.: The role of nucleophile–electrophile interactions in the unimolecular and bimolecular gas-phase ion chemistry of peptides and related systems. *J. Mass. Spectrom.* **35**, 1377–1381 (2000)
- Schlosser, A., Lehmann, W.D.: Five-membered ring formation in unimolecular reactions of peptides: a key structural element controlling low-energy collision-induced dissociation of peptides. *J. Mass. Spectrom.* **35**, 1382–1390 (2000)
- Paizs, B., Suhai, S., Hargittai, B., Hruby, V.J., Somogyi, A.: Ab initio and MS/MS studies on protonated peptides containing basic and acidic amino acid residues: I. Solvated proton vs. salt-bridged structures and the cleavage of the terminal amide bond of protonated RD-NH₂. *Int. J. Mass. Spectrom.* **219**, 203–232 (2002)
- Farrugia, J.M., O'Hair, R.A.J.: Involvement of salt bridges in a novel gas phase rearrangement of protonated arginine-containing dipeptides, which precedes fragmentation. *Int. J. Mass. Spectrom.* **222**, 229–242 (2003)
- Bleilholder, C., Osburn, S., Williams, T.D., Suhai, S., Van Stipdonk, M., Harrison, A.G., Paizs, B.: Sequence-scrambling fragmentation pathways of protonated peptides. *J. Am. Chem. Soc.* **130**, 17774–17789 (2008)
- Molesworth, S., Van Stipdonk, M.J.: Apparent inhibition by arginine of macrocyclic b ion formation from singly charged protonated peptides. *J. Am. Soc. Mass. Spectrom.* **21**, 1322–1328 (2010)
- Yalcin, T., Harrison, A.G.: Ion chemistry of protonated lysine derivatives. *J. Mass. Spectrom.* **31**, 1237–1243 (1996)
- Kish, M.M., Wesdemiotis, C.: Selective cleavage at internal lysine residues in protonated vs. metalated peptides. *Int. J. Mass. Spectrom.* **227**, 191–203 (2003)
- Farrugia, J.M., Taverner, T., O'Hair, R.A.J.: Side-chain involvement in the fragmentation reactions of protonated methyl esters of histidine and its peptides, which precedes fragmentation. *Int. J. Mass. Spectrom.* **209**, 99–112 (2001)
- Tsapraillis, G., Nair, H., Zhong, W., Kuppanan, K., Futrell, J.H., Wysocki, V.H.: A mechanistic investigation of the enhanced cleavage at histidine in the gas-phase dissociation of protonated peptides. *Anal. Chem.* **76**, 2083–2094 (2004)
- Bythell, B.J., Knapp-Mohammady, M., Paizs, B., Harrison, A.G.: Effect of the His residue on the cyclization of b ions. *J. Am. Soc. Mass. Spectrom.* **21**, 1352–1363 (2010)
- Vaisar, T., Urban, J.: Probing the proline effect in CID of protonated peptides. *J. Mass. Spectrom.* **31**, 1185–1187 (1996)
- Breci, L.A., Tabb, D.L., Yates, J.R., Wysocki, V.H.: Cleavage N-terminal to proline: analysis of a database of peptide tandem mass spectra. *Anal. Chem.* **75**, 1963–1971 (2003)
- Unnithan, A.G., Myer, M.J., Veale, C.J., Danell, A.S.: MS/MS of protonated polyproline peptides: the influence of N-terminal protonation on dissociation. *J. Am. Soc. Mass. Spectrom.* **18**, 2198–2203 (2007)
- Mouls, L., Subra, G., Aubagnac, J.-L., Martinez, J., Enjalbal, C.: Tandem mass spectrometry of amidated peptides. *J. Mass. Spectrom.* **41**, 1470–1483 (2006)
- Bythell, B.J., Suhai, S., Somogyi, A., Paizs, B.: Proton-driven amide bond-cleavage pathways of gas-phase peptide ions lacking mobile protons. *J. Am. Chem. Soc.* **131**, 14057–14065 (2009)
- Maux, D., Enjalbal, C., Martinez, J., Aubagnac, J.-L.: New example of a proline-induced fragmentation in ESI mass spectrometry of peptides. *Rapid Commun. Mass. Spectrom.* **16**, 1470–1475 (2002)
- Sachon, E., Clodic, G., Blasco, T., Jacquot, Y., Bolbach, G.: In-source fragmentation of very labile peptides in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal. Chem.* **81**, 8986–8992 (2009)