The Peptide MS/MS-Fragmentome: A Set of Predictable Fragment Ions with Highly Redundant Sequence Information

W.D. Lehmann [1] J. Wei [1#], J. Rappsilber [2], M. Salek [2]
[1] Central Spectroscopy, German Cancer Research Center (DKFZ), Heidelberg, Germany
[2] FIRC Institute of Molecular Oncology (IFOM), Milan, Italy
present address: Department of Biochemistry, University of Washington, Seattle, USA.

Extended Abstract

Low-energy collision-induced dissociation (CID) of peptides in a collision cell has been introduced as a fast method for peptide sequencing [1]. The method has been adapted to the analysis of gel-isolated proteins by the introduction of the nanoESI technology [2]. In the meantime proteins can be identified by search-engine supported interpretation of raw data of peptide MS/MS spectra in correlation to a protein sequence database [e.g. 3]. A triplet of information, consisting of the protease specificity (mostly trypsin), the molecular weight of the precursor peptide, and a series of MS/MS sequence ions usually forms the basic information for search engine-supported protein identification in sequence databases by ESI-MS/MS data. In spite of the success of this approach, in practical work the majority of automatically acquired ESI-MS/MS spectra of tryptic peptides is often left unidentified. Unknown covalent modifications, incorrect assignment of charge states of precursor and fragment ions, the absence of a sufficiently long series of sequence ions, or the selection of non-optimal instrumental conditions are considered to be the main reasons for this situation. Here, we focus on the fact that besides the highly significant y ion series, MS/MS spectra of multiply charged peptides contain a wealth of sequence information in a variety of fragment ion types. Since peptide fragmentation is severely sequence-dependent, knowledge of all sequence-specific features of a peptide MS/MS spectrum is highly useful. An individual peptide may present a certain part of its sequence only in a single type of fragmentation, for instance only by neutral loss fragmentations or by internal fragments. In this case, the sequence information would be overlooked by considering exclusively y ions. On the other hand, the recognition of redundant sequence information contained in several types of fragmentations is also highly useful, since this phenomenon increases the reliability of the sequence assignment. Moreover, redundancy of sequence information may enable an otherwise impossible clear decision in case more than one precursor ion is (automatically) selected in the first stage of MS/MS. Using the phosphopeptide SA-pT-PEALAFVR as an example, a description of the redundant structural information is provided that can be extracted from the peptide MS/MS-fragmentome, which is the universe of MS/MS fragments of multiply charged peptide ions.

In low energy collision-cell CID, different portions of the MS/MS-fragmentome show up in dependence of the applied collision offset. In proportion to this offset, the precursor ions gain additional kinetic energy, which is partially converted into internal energy after collision with the reagent atoms in the collisions cell (mostly nobel gas atoms, such as He or Ar). A higher collision offset corresponds to an increased probability of fragmentation, which includes the probability for formation of secondary (or even higher order) fragments, since fragment ions can undergo additional collisions with the reagent gas as they travel through the collision cell. Figure 1 shows the MS/MS spectra of the $[M+2H]^{2+}$ ion of the model peptide SA-pT-PEALAFVR as recorded using low, moderate and high collision offset conditions.

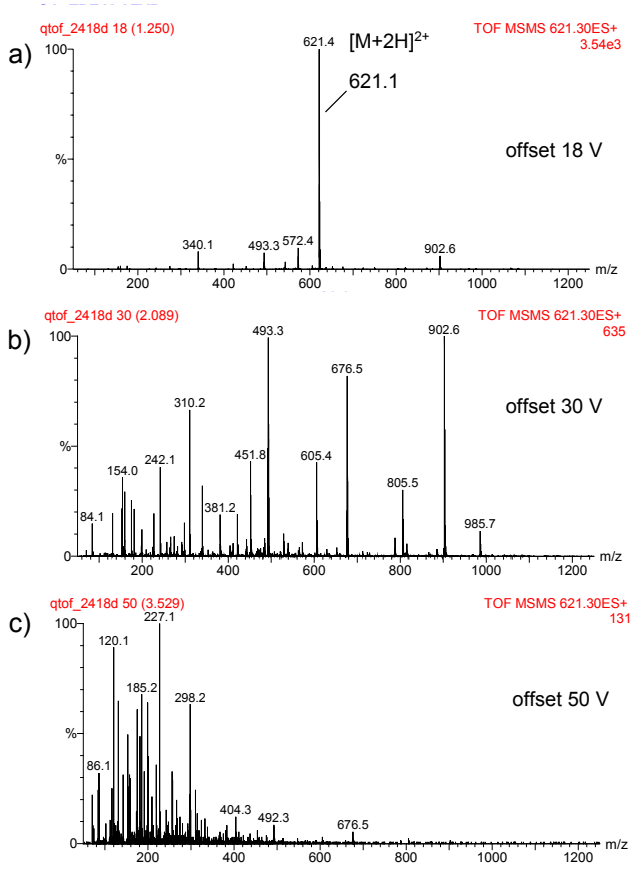

Figure 1:

Variability in shape and intensity of the collision cell MS/MS spectrum of the peptide SA-pT-PEALAFVR as a function of the collision offset:

a) low offset (18 V), mainly first generation backbone cleavages and neutral loss reactions are observed;
100 % intensity = 3.5 x 10^3 counts.
b) intermediate offset (30 V), a mixture of primary and higher order backbone cleavages and some internal fragments are observed;
100 % intensity = 635 counts.
c) high offset (50 V), the spectrum shows immonium ions and di- and tripeptide internal fragments;
100 % intensity = 131 counts.

MS/MS spectra of peptides recorded at low collision offset (Fig. 1a) show mainly first generation peptide backbone cleavages, for instance proline-induced cleavages, and neutral loss reactions, such as loss of H_3PO_4 for phosphopeptides and neutral loss reactions from the peptide N-terminus [4]. Collision cell MS/MS spectra recorded at low offset resemble ion trap MS/MS spectra [5], which are also governed by first generation fragment ions. A particular feature of peptide backbone cleavages is that they result in complementary b/y ion pairs, which by principle must be formed in exactly equimolar amounts (see also Figure 2). Cleavages at the N-terminal site of Pro or at the C-terminal site of Asp and Glu are the prototype forms of such preferred backbone cleavages resulting in complementary fragment

ions. Deviations from equal intensity of the two complementary fragment ions may occur due to secondary fragmentation of one of the fragments, by mass-dependent ion transmission of the analyser or by a mass-dependent detection efficiency. Under intermediate collision offset conditions (Fig. 1b), the largest variety of sequence-specific fragment ions are observed, since both first and higher order fragments are observed. In addition, some internal fragment ions occur. Figure 2 shows the MS/MS spectrum in Figure 1a (low offset) with enhanced intensity.

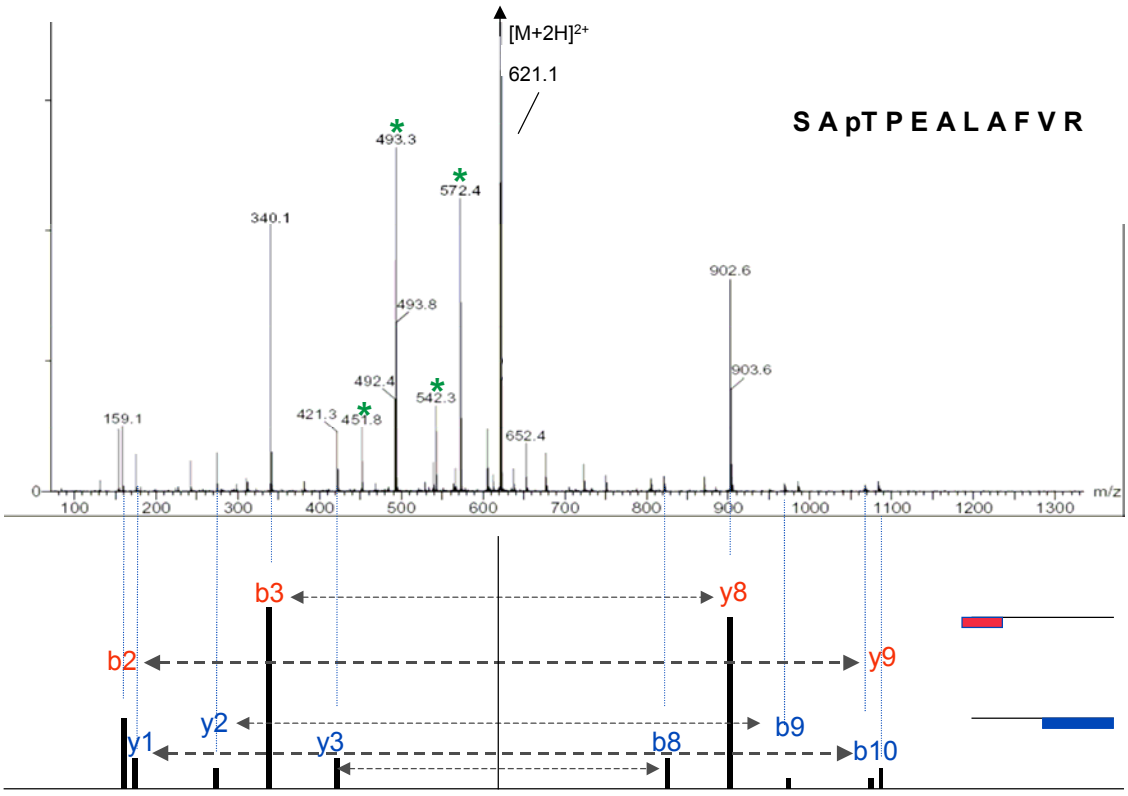

Figure 2:
MS/MS spectrum of the $[M+2H]^{2+}$ ion of the peptide SA-pT-PEALAFVR (18 V, Figure 1a) with expanded intensity scale. The neutral loss fragments (marked with an asterisk) indicate the following losses: 572.4 = $-H_3PO_4$; 542.3 = - SA; 493.3 = $-H_3PO_4$ – SA; 451.8 = -SApT. Due to the absence of secondary fragmentations, backbone cleavages show up mainly through complementary b/y ion pairs. These cleavages are concentrated in the N- and C-terminal region. Annotations are given in the lower panel. Most of the other fragment ions are derived from the b/y ion pairs by neutral loss of small units, such as H_2O, NH_3 or CO.

The MS/MS spectra recorded at low or moderate collision offset are most favourable in terms of absolute ion intensity and background level. As shown in Figure 2, most fragment ions at low offset are first order fragment ions, which support each other by their complementary nature. They can be unambiguously correlated to the precursor ion, since their mass values add up to the precursor ion mass. The sequence information derived from the neutral loss ions is complementary to the b ion series indicating the N-terminal sequence.

A dynamic fragmentation of $[peptide]^{n+}$ ions in a collision cell includes recording of MS/MS spectra at different collision offset values (from low to high) so that accumulated MS/MS spectra with an optimal content of sequence information are obtained. The partially redundant sequence information is highly valuable in the practice of peptide sequencing, since a variety of factors, such as sequence-specific fragmentation behaviour or selection of multiple precursor ions may generate complex MS/MS spectra, which may be difficult to interpret with only partial utilization of the sequence information. This situation may be improved by a more complete use of the structural information present in the peptide MS/MS-fragmentome and by recognition of its highly abundant and self-supporting sequence information.

Further information may be obtained by considering the low-mass fragment ions, comprising a ions, b ions, y ions, and internal b ions [6]. In the low-mass region, the unique set of 19 y_1 ions and of the 190 b_2 ions carries a particular message, since these ions define the N-or C-terminal amino acid(s). Further, the b_1 ions of the basic residues K, H, W, and R carry a specific N-terminal information, which is redundant to that contained in the corresponding b_2 ions and in the N-terminal neutral loss peaks. Redundant information is also found in b and y ion series and in complementary b/y ion pairs. From the latter, the molecular weight and the charge state of the precursor ion can be reconstructed. This is helpful in case a mixture of precursor ions or a precursor ion of very low abundance is isolated. Finally, reporter ions or reporter neutral losses for modifications, such as the loss of H_3PO_4 from pSer- or pThr-peptides, are useful for recognition of covalently modified peptides. Search tools, which fully incorporate the current knowledge about the peptide MS/MS-fragmentome will increase the scores of peptide/protein identifications by MS/MS and thus will increase the fraction of automatically assigned MS/MS spectra in proteomic studies.

References

1. Hunt DF, Buko AM, Ballard JM, Shabanowitz J, Giordani AB. *Biomed. Mass Spectrom.* 1981, 8, 397-408. Sequence analysis of polypeptides by collision activated dissiociation on a triple quadrupole mass spectrometer.
2. Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T, Mann M. *Nature* 1996, 379, 466-469. Femtomole sequencing of propteins from pollyacrylamide gels by nano-electrospray mass spectrometry.
3. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Electrophoresis 1999, 20, 3551-3567. Probability-based protein identification by searching sequence databases using mass spectrometry data.
4. Salek M, Lehmann WD. *J. Mass Spectrom.* 2003, 38, 1143-1149. Neutral loss of amino acid residues from protonated peptides in collision-induced dissociation generates N- or C-terminal sequence ladders.
5. Schwartz JC, Jardine I. *Methods Enzymol.* 1996, 270, 552-586. Quadrupole ion trap mass spectrometry.
6. Schlosser A, Lehmann WD. *Proteomics* 2002, 2, 524-533. Patchwork peptide sequencing – extraction of sequence information from accurate mass data of peptide tandem mass spectra recorded at high resolution.