Mass Spectrometry of Proteins

DOI: 10.1002/anie.201304046

Photodissociation and Dissociative Photoionization Mass Spectrometry of Proteins and Noncovalent Protein–Ligand Complexes**

*Francis Canon, Aleksandar R. Milosavljević, Guillaume van der Rest, Matthieu Réfrégiers, Laurent Nahon, Pascale Sarni-Manchado, Véronique Cheynier, and Alexandre Giuliani**

Tandem mass spectrometry (MS2) is a widely used method in structural analysis and biopolymer sequencing.[1] In MS2, an ion of interest is isolated, activated, and brought to dissociation. After analysis, the generated fragments provide structural information on the precursor ion.[2] Among the different activation methods, low-energy collision induced dissociation (CID), which relies on slow heating of ions through multiple inelastic low-energy collisions with a neutral gas, is the most widely used.[3] Another set of activation techniques, referred to as the EXD methods (electron capture dissociation (ECD) or electron-transfer dissociation (ETD)), involves the dissociative recombination reaction of multiply protonated ions with electrons.[4] The EXD set of techniques preserve labile bonds, which makes them particularly suited to localize post-translational modifications (PTM) and noncovalent binding sites on protein backbones.[3,5] Another way to increase the internal energy of an ion is by absorption of energetic photons. Methods involving ultraviolet (UV) laser-based photodissociation (PD) of electrosprayed ions are attracting a growing interest since the pioneering work of Williams and McLafferty[6] and have recently found promising applications in proteomics analysis.[7] Indeed, UV activation produces fragments complementary to those generated by other methods and provides high coverage in peptide sequence, that is almost the complete sequence can be determined from the experiments. Among the different wavelengths used, vacuum ultraviolet (VUV) photons of 193 nm (6.2 eV) and 157 nm (7.8 eV) have demonstrated the highest potential to be of use.[7,8]

However, very little is known about fragmentations induced above 8 eV. This problem pertains to the limitations in terms of photon energy delivered by lasers. Surprisingly, synchrotron radiation (SR), which is a widely tunable photon source in the VUVs, has been used only very recently for ion activation. Preliminary work has indicated some potential to bring sequence information on small peptides.[9]

Herein, we report on the potential of SR as a credible activation method in MS2 for structural analysis. The performances of photon activation over a wide energy range are compared to those of CID and ECD in terms of protein sequence coverage. More interestingly, the ability of this new method to preserve noncovalent interactions and to permit identification of the binding sites of a ligand on a protein is investigated and its outcomes are compared to those obtained by ECD.

This study has been performed on a human intrinsically disordered protein (IDP), named IB5. The only known function of IB5 and other basic proline-rich proteins (PRPs) is to bind and scavenge tannins, and thus constitute a first-line of defense against tannin anti-nutritional effects. These interactions are also thought to play a role in the sensation of astringency,[10] which is one of the most important organoleptic properties of food regarding consumer acceptability. Hitherto, the tannin binding sites on IB5 could not be precisely determined by classical structural approaches, such as X-rays diffraction or NMR spectroscopy, because of the lack of crystals for diffraction and of difficulties in spectra interpretation for spectroscopy because of the multiple repeated sequences of this protein. In contrast, the SR-based tandem mass spectrometry activation method presented herein has achieved, for the first time, unambiguous determination of the binding site of the tannin procyanidin B2 3′OG ((−)-epicatechin (4β-8)-(−)-epicatechin 3-O-gallate) on IB5.

[*] Dr. F. Canon, Dr. M. Réfrégiers, Dr. L. Nahon, Dr. A. Giuliani
Synchrotron SOLEIL
L'Orme des Merisiers, Saint Aubin
91192 Gif-sur-Yvette (France)
E-mail: giuliani@synchrotron-soleil.fr
Dr. F. Canon
INRA, UMR1324 Centre des Sciences du Goût et de l'Alimentation
21065 Dijon Cedex (France)
Dr. A. Giuliani
UAR1008 CEPIA, INRA
44316 Nantes (France)
Dr. A. R. Milosavljević
Institute of Physics Belgrade, University of Belgrade
Pregrevica 118, 11080 Belgrade (Serbia)
Dr. P. Sarni-Manchado, Dr. V. Cheynier
INRA, UMR1083 Sciences pour l'Œnologie
2 place Viala, 34060 Montpellier Cedex (France)
Dr. G. van der Rest
Laboratoire de Chimie Physique, Université Paris Sud
91405 Orsay Cedex (France)

[**] This work was supported by the Agence Nationale de la Recherche, France, under the projects ANR-08-BLAN-0065 and ANR-BLAN-0279. A.R.M. acknowledges support by the Ministry of Education, Science and technical development of Republic of Serbia (Projects No. 171020). Financial support from the TGE FT-ICR for conducting ECD dissociation experiments is gratefully acknowledged. The SOLEIL synchrotron radiation facility is acknowledged for providing beamtime under project 20090862.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201304046.

The experimental set up, described in detail elsewhere,[11] is based upon the coupling of a linear ion-trap mass spectrometer (LTQ XL, Thermo Finnigan) with a VUV beamline[12] at the SOLEIL synchrotron radiation facility (Figure 1). The mass spectrometer is equipped with a nano-

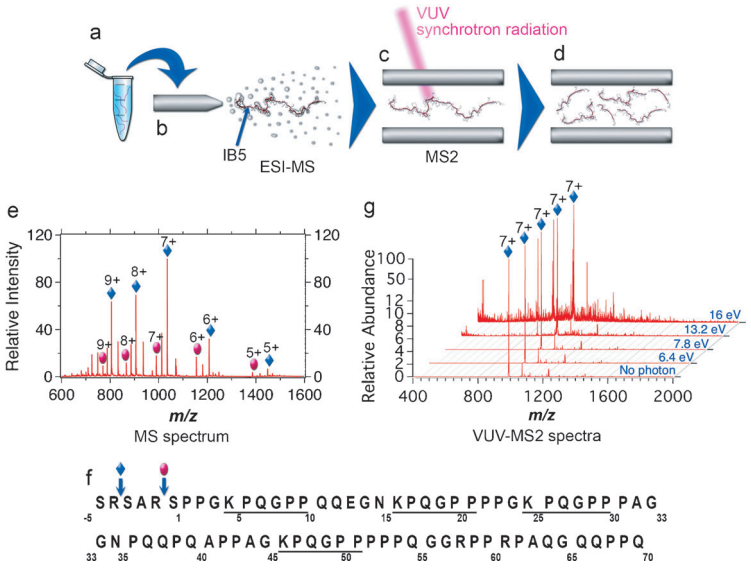

Figure 1. Synchrotron-radiation-based photon-induced activation. Solutions containing the protein, IB5, or its noncovalent complexes (a) are ionized (b), transferred into the ion trap (c), where the ions of interest are isolated and irradiated to generate fragment ions (d). The mass spectrum (e) shows a charge state distribution ranging from +5 to +9 for the two protein isoforms. The sequences of the isoforms are indicated in (f) by the red (IB5b) and blue (IB5a) symbols. Tandem mass spectra recorded at several photon energies (6.4, 7.8, 13.2, and 16 eV) are shown in (g) for $[IB5a]^{7+}$, spectra offset for clarity.

spray ion source allowing soft ionization of native proteins and noncovalent supramolecular complexes. The targeted protein or noncovalent complex can then be selected and activated with energetic VUV photons (Figure 1c) delivered by the beamline, thereby generating MS2 spectra. The fragments may eventually be activated in a subsequent tandem MS step (MS3) upon further VUV irradiation or CID.

The mass spectrum of the protein (Figure 1e), shows a charge-state distribution characteristic of a disordered protein. It also reveals that several protein isoforms have been co-purified.[13] The two isoforms, whose sequences are indicated in the Figure 1f have been activated with photons during 500 ms. The resulting MS2 spectra for IB5a (Figure 1g) reveal that the fragmentation patterns are strongly dependent on the photon energy. Figure 2 shows both the nature and the location of the bond cleavages on the protein backbone as a result of VUV irradiation at 6.4, 7.8, 13.2, and 16 eV, and ECD and CID, respectively. As expected, CID generates mainly y- and b-type ions, whereas ECD sequence ions are mainly c and z-type in nature. In contrast, both MS2 spectra of $[IB5a]^{7+}$ at 6.4 and 7.8 eV are dominated by a-type ions, which result from C_α–C bond cleavages as observed from laser-based PD experiments at the same photon energies on protonated peptides.[7,8b] At 13.2 and 16 eV, the mass spectra exhibit significantly richer fragmentation patterns (Figure 2a) into a, b, c, x, y and z-type ions, which resembles the outcomes of electron ionization dissociation (EID).[14] The abundances of the precursor (Figure 3a) and the product ions (Figure 3b,c) reveal that the fragmentation efficiency of the protein increases with the photon energy, which is consistent with the energy dependence of the photoabsorption cross-section of ions.[9] It leads us to distinguish between two different photon energy regimes. The fundamental difference between the low-photon-energy range (6.4 and 7.8 eV) and the high-energy one (13.2 and 16 eV) lies in the appearance of the $[M+7H]^{\bullet 8+}$ and $[M+7H]^{\bullet 9+}$ radical cations produced upon photoionization of the precursor. These characteristics are similar to those reported[14] under EID conditions of protein ions. From the relative abundances of $[M+7H]^{\bullet 8+}$ and $[M+7H]^{\bullet 9+}$ displayed for different photon energies in Figure 3b, it clearly appears that the ionization threshold of IB5 lies in between 7.8 and 13.2 eV. Above this threshold, photon activation leads to the ionization of the precursor ion and generates fragments through dissociative photoionization (DPI), with abundant formation of sequence ions of various natures. It has been shown theoretically that the removal of an electron from the peptidic backbone significantly weakens all the bonds, making them prone to fragmentation,[15] which is consistent with the present observations. Below the ionization threshold, photon absorption populates electronic excited state, which relax partly through dissociation. This regime, usually referred to as photodissociation (PD), has been shown to lead mainly to C_α–C bond cleavage.[8b] Tuning the photon energy below or above the ionization onset provides control over the extent and the nature of fragmentation pattern.

The following sequence coverage of the IB5 protein has been obtained: 46% for PD at 6.4 eV, 56% for PD at 7.8 eV, 58% for CID, 68% for ECD, and 91% for DPI at 16 eV. Although the CID spectrum is dominated by fragments resulting from cleavages close to proline residues, known as the proline effect[16] and previously described for this protein,[13] DPI leads to cleavage evenly spread along the backbone. The improved sequence coverage of ECD over CID remains however inferior to that obtained from DPI, because the abundant proline residues preclude the formation of c- and z-type fragments[16] under ECD conditions.

The photoabsorption cross-section is proportional to the number of electrons involved in the process.[17] Above the ionization threshold, photoionization is the main relaxation process following photoabsorption. This feature has the important practical consequence that the probability of ionizing a target increases with its mass or its size, which is totally opposite to CID which loses efficiency as the size of the target increases.[3]

The mass spectra of a solution containing IB5 and B2 3′OG typically show IB5·B2 3′OG supramolecular edifices with several stoichiometries in agreement with the tannin-

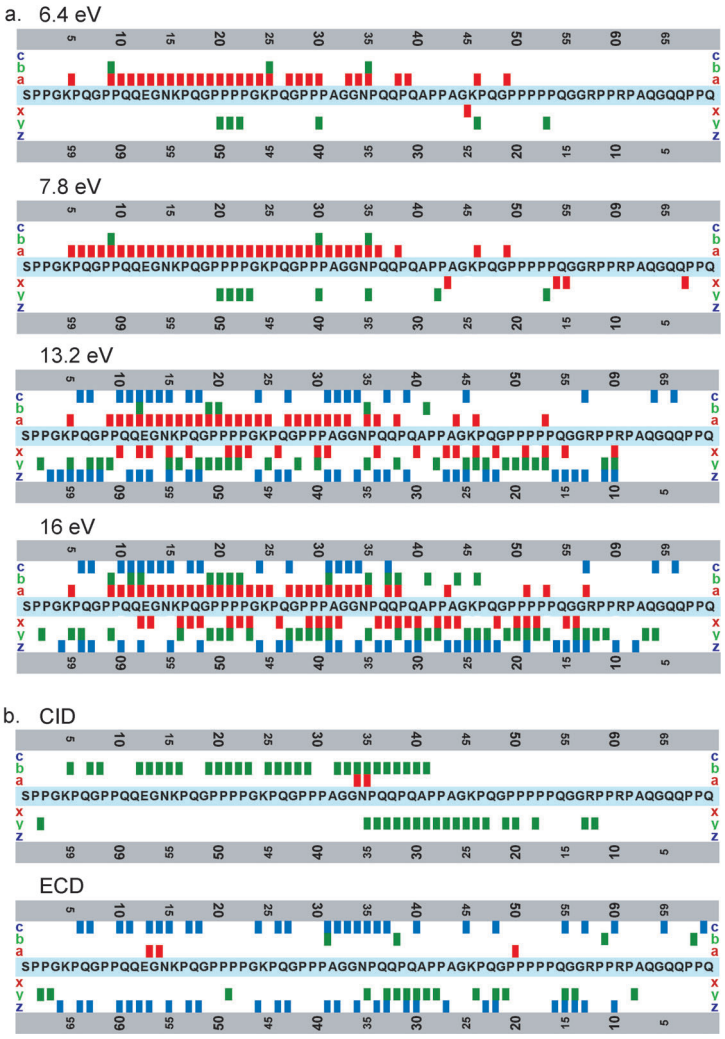

Figure 2. a) pattern of $[IB5]^{7+}$ fragmentation as a function of photon energy; b) pattern of $[IB5]^{7+}$ fragmentation resulting from CID and ECD experiments.

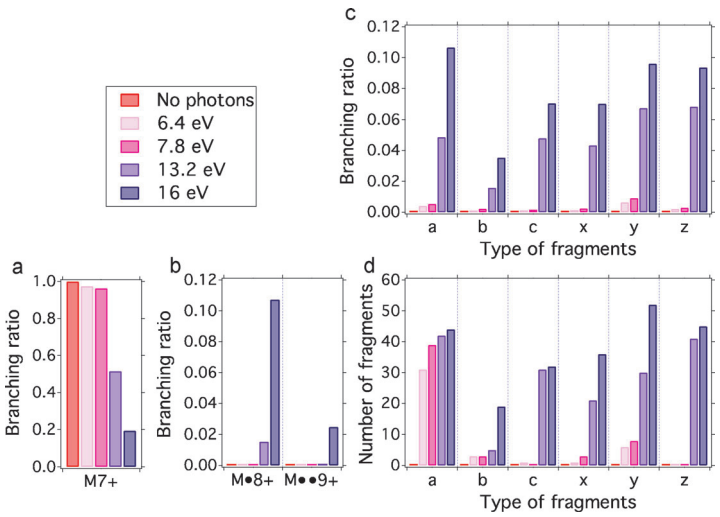

Figure 3. Branching ratio of the parent ion (a), the photo-ionized $[IB5]^{\bullet 8+}$ and $[IB5]^{\bullet\bullet 9+}$ (b), and the a, b, c, x, y, and z sequence ions (c) as a function of the photon energy. (d) Histogram of the number of fragment ions produced as a function of photon energy.

scavenging function of IB5.[18] The clusters of proline present in the IB5 sequence (Figure 1 f), as in other salivary PRPs, are believed to provide multiple tannin binding sites.[19] NMR experiments have confirmed the involvement of proline clusters in these interactions, but were not able to identify precisely which part of the sequence was involved.[20] To assess the potential of photon activation to localize the binding sites of the ligand (B2 3′OG) to the protein, IB5 1 B2 3′OG has been submitted to 16 eV DPI (Figure 4). This regime was chosen because of the large sequence coverage and abundance of fragments it had provided on the bare protein. The comparison with the MS2 spectrum of the bare protein (Figure 4 a) revealed peaks that are shared by both spectra and others that are not. Analysis of the specific peaks revealed that the masses of these fragments correspond to that of the specific fragments identified in the MS2 spectrum of $[IB5+7H]^{7+}$ shifted by the mass of the ligand. These fragments have been selected (Figure 4 b) and a subsequent CID analysis has been conducted. This MS3 step unambiguously confirmed the presence of the ligand through its detachment from the protein fragment (Figure 4 c). The list of fragments exhibiting such behavior is presented in Figure 4 d. This procedure has been repeated for the second isoform of the protein, IB5b (Figure 1 f), and the results are given in Figure 4 e. As expected, the fragmentation patterns of $[IB5a{\cdot}B2\ 3'OG+7H]^{7+}$ and $[IB5b{\cdot}B2\ 3'OG+7H]^{7+}$ are very similar, showing the reliability of our method (Figure 4 d,e). In both experiments more than fifty peaks have been identified and interpreted as fragments of IB5 noncovalently bound to B2 3′OG, contrasting with the four fragments identified using ECD (Figure 4 f). It appears that all these fragments from both N- and C-terminal series contain the KPQGPPPPPQGG segment of the sequence, indicating therefore a strong interaction between B2 3′OG and this part of the protein. Note that comparison of the DPI fragmentation patterns obtained for the bare and the complexed protein, as shown in the Supporting Information, reveals that the presence of the tannin does not affect significantly the fragmentation pattern of the protein, the ligand appearing as a spectator of the fragmentation process. This observation is consistent with the DPI mechanism suggested herein, where most of the photon absorption is ascribed to the protein. The sequence identified with a cluster of five prolines very likely adopts a PPI or a PPII helix conformation in solution.[20b,21] Such structural elements are thought to be crucial for IDPs in the binding with their partner,[22] as this stable segment might provide an initial contact point.[19a,20b] Moreover, the phenolic rings of the tannins can stack on the Pro-S face of proline.[20a] Therefore, the sequence identified displays all the features required for it to be the preferential binding site of B2 3′OG on IB5 and the role of the proline clusters in PRPs sequence is thus unequivocally confirmed.

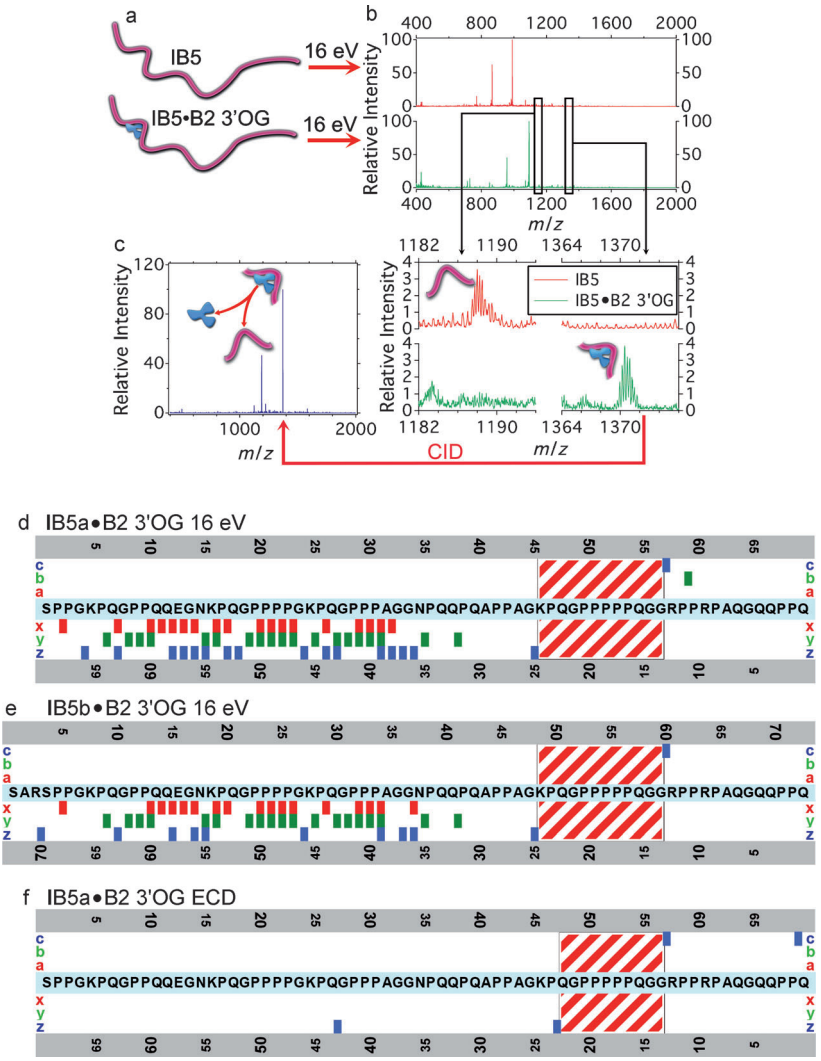

Figure 4. Localization of the tannin (B2 3′OG) binding site on IB5. a) MS2 spectra of $[IB5+7H]^{7+}$ and $[IB5{\cdot}B2\ 3'OG+7H]^{7+}$ after dissociative photoionization (DPI) at 16 eV. b) Enlargement on two parts of the DPI MS2 spectra for the bare and the complexed protein showing the shift in mass of one fragment as a result of the conservation of the ligand on it. c) DPI/CID MS3 spectrum of the *m/z* 1369.95 ion confirming the presence of the ligand on the DPI fragment ion. Patterns of the fragments linked to B2 3′OG for d) $[IB5a{\cdot}B2\ 3'OG+7H]^{7+}$ and e) $[IB5b{\cdot}B2\ 3'OG+7H]^{7+}$ after 16 eV DPI. f) Patterns of the fragments linked to B2 3′OG for $[IB5a{\cdot}B2\ 3'OG+7H]^{7+}$ after ECD activation. The binding site of B2 3′OG is highlighted by the dashed rectangle.

In summary, the present study shows that the nature of the fragments generated upon VUV photon activation is controlled by the photon energy. Therefore, synchrotron radiation, by providing access to a wide variety of photon activation regimes, ranging from PD to DPI, clearly appears to be very promising sequencing method, complementary to laboratory activation techniques. The sequence coverage, obtained both in PD and DPI regimes, makes synchrotron-radiation activation an appealing method for sequencing. DPI allowed the determination of the ligand binding site on an IDP. Therefore the new method, based upon DPI, should aid research in the growing and challenging field of protein noncovalent interactions.

Experimental Section

IB5 and B2 3'-O-gallate were produced and purified as described elsewhere.[18] Stock solutions of B2 3'-O-gallate were prepared in water/ethanol, 88:12 (v/v) acidified to pH 3.3 with acetic acid, mimicking the consumption of red wine. Final solutions of 10 μM of IB5 were electrosprayed, giving an overall consumption of 300 fmol per tandem mass spectrum. For IB5-B2 3′OG interactions, tannin and protein solutions were mixed extemporaneously at room temperature (regulated at 24 °C) to obtain a protein/polyphenol molar ratio of 1:10 (10 μM:100 μM).

The experimental setup is based upon a linear ion-trap mass spectrometer (LTQ XL, Thermo Electron, San Jose, CA, USA)[11] coupled to the DESIRS beamline[12] of the SOLEIL synchrotron radiation facility (France). The synchrotron radiation beam was injected along the ion-trap axis. An MgF_2 window (below 10 eV) and gas filter filled with argon (8–16 eV) were used to suppress the high harmonic content delivered by the undulator that would be transmitted by the monochromator's grating. An electromechanical shutter was placed along the SR beams, allowing synchronization of the SR beam with the LTQ time sequence. For MS2 experiments, the precursor ions were irradiated during 500 ms. For MS3 experiments, a supplementary step of CID was applied on the targeted VUV product ion during which the shutter was kept close.

Mass spectra were analyzed using mMass software[23] and Igor Pro (Wavemetrics, Portland, USA). For each spectrum, the relative abundance (R) of the different ions has been established by the measurement of their peak area (A). The contribution of fragments formed without photon activation has been subtracted. Each of the peak areas was normalized by the sum of the area of all ions, and to the photon flux.

The ECD experiments were carried out with a 7-T APEX III FT-ICR mass spectrometer (Bruker Daltonik GmbH, Germany) equipped with an electrospray ion source. The isolated precursor ions were irradiated with electrons having less than 1 eV kinetic energy during 150 ms.

Received: May 11, 2013
Published online: July 19, 2013

Keywords: dissociative photoionization · mass spectrometry · proteins · tannins · VUV photon activation

[1] R. Aebersold, M. Mann, *Nature* **2003**, *422*, 198–207.
[2] F. W. McLafferty, *Science* **1981**, *214*, 280–287.
[3] N. L. Kelleher, *Anal. Chem.* **2004**, *76*, 196A–203A.
[4] a) R. A. Zubarev, N. L. Kelleher, F. W. McLafferty, *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266; b) J. E. P. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz, D. F. Hunt, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9528–9533.
[5] Y. Xie, J. Zhang, S. Yin, J. A. Loo, *J. Am. Chem. Soc.* **2006**, *128*, 14432–14433.
[6] E. R. Williams, F. W. McLafferty, *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 361–365.

[7] J. A. Madsen, D. R. Boutz, J. S. Brodbelt, *J. Proteome Res.* **2010**, *9*, 4205 – 4214.
[8] a) H. Zhou, Z. Ning, A. E. Starr, M. Abu-Farha, D. Figeys, *Anal. Chem.* **2012**, *84*, 720 – 734; b) L. Zhang, J. P. Reilly, *J. Proteome Res.* **2010**, *9*, 3025 – 3034.
[9] S. Bari, O. Gonzalez-Magana, G. Reitsma, J. Werner, S. Schippers, R. Hoekstra, T. Schlatholter, *J. Chem. Phys.* **2011**, *134*, 024314 – 024319.
[10] E. C. Bate-Smith, *Phytochemistry* **1973**, *12*, 907 – 912.
[11] A. R. Milosavljević, C. Nicolas, J.-F. Gil, F. Canon, M. Refregiers, L. Nahon, A. Giuliani, *J. Synchrotron Radiat.* **2012**, *19*, 174 – 178.
[12] L. Nahon, N. de Oliveira, G. A. Garcia, J.-F. Gil, B. Pilette, O. Marcouille, B. Lagarde, F. Polack, *J. Synchrotron Radiat.* **2012**, *19*, 508 – 520.
[13] F. Canon, F. Paté, E. Meudec, T. Marlin, V. Cheynier, A. Giuliani, P. Sarni-Manchado, *Anal. Bioanal. Chem.* **2009**, *395*, 2535 – 2545.
[14] a) Y. M. E. Fung, C. M. Adams, R. A. Zubarev, *J. Am. Chem. Soc.* **2009**, *131*, 9977 – 9985; b) R. A. Zubarev, H. Yang, *Angew. Chem.* **2010**, *122*, 1481 – 1483; *Angew. Chem. Int. Ed.* **2010**, *49*, 1439 – 1441.
[15] R. Parthasarathi, Y. He, J. P. Reilly, K. Raghavachari, *J. Am. Chem. Soc.* **2010**, *132*, 1606 – 1610.
[16] N. Leymarie, E. A. Berg, M. E. McComb, P. B. O'Connor, J. Grogan, F. G. Oppenheim, C. E. Costello, *Anal. Chem.* **2002**, *74*, 4124 – 4132.
[17] T. N. Olney, G. Cooper, W. F. Chan, G. R. Burton, C. E. Brion, K. H. Tan, *Chem. Phys.* **1997**, *218*, 127 – 149.
[18] a) F. Canon, A. Giuliani, F. Paté, P. Sarni-Manchado, *Anal. Bioanal. Chem.* **2010**, *398*, 815 – 822;b) F. Canon, R. Ballivian, F. Chirot, R. Antoine, P. Sarni-Manchado, J. Lemoine, P. Dugourd, *J. Am. Chem. Soc.* **2011**, *133*, 7847-7852.
[19] a) A. J. Charlton, N. J. Baxter, T. H. Lilley, E. Haslam, C. J. McDonald, M. P. Williamson, *FEBS Lett.* **1996**, *382*, 289 – 292; b) F. Canon, F. Paté, V. Cheynier, P. Sarni-Manchado, A. Giuliani, J. Pérez, D. Durand, J. Li, B. Cabane, *Langmuir* **2013**, *29*, 1926 – 1937.
[20] a) N. J. Baxter, T. H. Lilley, E. Haslam, M. P. Williamson, *Biochemistry* **1997**, *36*, 5566 – 5577; b) C. Pascal, F. Paté, V. Cheynier, M.-A. Delsuc, *Biopolymers* **2009**, *91*, 745 – 756.
[21] H. Boze, T. Marlin, D. Durand, J. Pérez, A. Vernhet, F. Canon, P. Sarni-Manchado, V. Cheynier, B. Cabane, *Biophys. J.* **2010**, *99*, 656 – 665.
[22] M. Fuxreiter, I. Simon, P. Friedrich, P. Tompa, *J. Mol. Biol.* **2004**, *338*, 1015 – 1026.
[23] M. Strohalm, D. Kavan, P. Novák, M. Volný, V. r. Havlíček, *Anal. Chem.* **2010**, *82*, 4648 – 4651.

