

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/43180381>

# MALDI-In Source Decay Applied to Mass Spectrometry Imaging: A New Tool for Protein Identification

ARTICLE *in* ANALYTICAL CHEMISTRY · MAY 2010

Impact Factor: 5.64 · DOI: 10.1021/ac902875q · Source: PubMed

---

CITATIONS

55

READS

123

## 5 AUTHORS, INCLUDING:



[Delphine Debois](#)

ZenTech S.A.

26 PUBLICATIONS 546 CITATIONS

[SEE PROFILE](#)



[Loïc Quinton](#)

University of Liège

38 PUBLICATIONS 618 CITATIONS

[SEE PROFILE](#)



[Edwin De Pauw](#)

University of Liège

383 PUBLICATIONS 7,232 CITATIONS

[SEE PROFILE](#)

# MALDI-In Source Decay Applied to Mass Spectrometry Imaging: A New Tool for Protein Identification

Delphine Debois,<sup>\*,†</sup> Virginie Bertrand,<sup>†,‡</sup> Loïc Quinton,<sup>†</sup> Marie-Claire De Pauw-Gillet,<sup>‡</sup> and Edwin De Pauw<sup>†</sup>

LSM-CART-GIGA-R and Laboratory of Histology-Cytology (GIGA-R), University of Liège,  
B-4000 Liège (Sart-Tilman), Belgium

Matrix-assisted laser desorption ionization-mass spectrometry (MALDI-MS) imaging is a powerful technique giving access to the distribution of a large range of biomolecules directly from a tissue section, allowing, for example, the discovery of new pathological biomarkers. Nevertheless, one main difficulty lies in the identification of the detected species, especially proteins. MALDI-in source decay (ISD) is used to fragment ions directly in the mass spectrometer ion source. This technique does not require any special sample treatment but only the use of a specific MALDI matrix such as 2,5-dihydroxybenzoic acid or 1,5-diaminonaphthalene. MALDI-ISD is generally employed on classical, purified samples, but here we demonstrate that ISD can also be performed directly on mixtures and on a tissue slice leading to fragment ions, allowing the identification of major proteins without any further treatment. On a porcine eye lens slice, *de novo* sequencing was even performed. Crystallins not yet referenced in databases were identified by sequence homology with other mammalian species. On a mouse brain slice, we demonstrate that results obtained with ISD are comparable and even better than those obtained with a classical *in situ* digestion.

Matrix-assisted laser desorption ionization (MALDI)-mass spectrometry imaging (MSI) is a powerful technique allowing the mapping of any compound present at the surface of a biological sample (the more often a tissue slice). The tissue section is divided into pixels, and a mass spectrum is recorded for each one. A two-dimensional map is then generated, in which the intensity of the compound is depicted by a color scale, associated with the spatial coordinates of each pixel (*x,y*). The proof of concept of this technique has been demonstrated in 1997 by Caprioli et al.<sup>1</sup> The main advantage of the technique is its lack of *a priori* knowledge since any compound present at the surface of the sample can potentially be detected. Thus, for 12 years, different kinds of

compounds such as proteins and peptides,<sup>2–5</sup> lipids,<sup>6,7</sup> and drugs and their metabolites<sup>8</sup> have been successfully studied. Nowadays, MSI is used more and more in clinical proteomics approaches as a tool for biomarkers hunting and/or diagnosis.<sup>9,10</sup> Nevertheless, classical protein identification through standard proteomics studies are still at work. We can summarize them in two main approaches: “bottom-up” and “top-down”. The bottom-up methodology is based on an enzymatic digestion, most of the time using trypsin. The resulting mix of peptides is then analyzed by MALDI-time-of-flight (TOF)/MS. The corresponding masses are submitted to a protein databases query to identify the unknown protein. This method is commonly named peptide mass fingerprinting (PMF) and may be combined with tandem MS (MS/MS) to sequence peptides and confirm the protein identity.<sup>11</sup> The top-down methodology or sequencing (TDS) relies on the fragmentation of protein(s) ion(s) without any enzymatic digestion prior to analysis.<sup>12</sup> TDS is generally performed on multiply charged ions, generated by electrospray ionization. Different modes of fragmentations are then implemented: collision-induced dissociation, electron capture dissociation, and electron transfer dissociation.<sup>13–15</sup>

- (2) Burnum, K. E.; Frappier, S. L.; Caprioli, R. M. *Annu. Rev. Anal. Chem.* **2008**, *1*, 689–705.
- (3) Chaurand, P.; Schwartz, S. A.; Caprioli, R. M. *Anal. Chem.* **2004**, *76*, 87A–93A.
- (4) Taban, I. M.; Altelaar, A. F. M.; van der Burgt, Y. E. M.; McDonnell, L. A.; Heeren, R. M. A.; Fuchser, J.; Baykut, G. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 145–151.
- (5) Grey, A. C.; Chaurand, P.; Caprioli, R. M.; Schey, K. L. *J. Proteome Res.* **2009**, *8*, 3278–3283.
- (6) Chan, K.; Lanthier, P.; Liu, X.; Sandhu, J. K.; Stanimirovic, D.; Li, J. *Anal. Chim. Acta* **2009**, *639*, 57–61.
- (7) Murphy, R. C.; Hankin, J. A.; Barkley, R. M. *J. Lipid Res.* **2009**, *50* (S3), 17–22.
- (8) Fuchser, J.; Cornett, S.; Becker, M. *Bruker Daltonics Application Note No. FTMS-37*, **2008**.
- (9) Fournier, I.; Wisztorski, M.; Salzet, M. *Expert Rev. Proteomics* **2008**, *5*, 413–424.
- (10) Meistermann, H.; Norris, J. L.; Aerni, H. R.; Cornett, D. S.; Friedlein, A.; Erskine, A. R.; Augustin, A.; De Vera Mudry, M. C.; Ruepp, S.; Suter, L.; Langen, H.; Caprioli, R. M.; Ducret, A. *Mol. Cell. Proteomics* **2006**, *5*, 1876–1886.
- (11) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198–207.
- (12) Kelleher, N. L. *Anal. Chem.* **2004**, *76*, 197A–203A.
- (13) Kruger, N. A.; Zubarev, R. A.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass Spectrom.* **1999**, *185–187*, 787–793.
- (14) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266.
- (15) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528–9533.

\* To whom correspondence should be addressed. Dr. Delphine Debois, University of Liège, LSM-CART-GIGA-R, Allée de la Chimie, 3, Bâtiment B6c, B-4000 Liège, Belgium. Phone: 00324 366 3382. Fax: 00324 366 3413. E-mail: delphine.debois@ulg.ac.be.

† LSM-CART-GIGA-R.

‡ Laboratory of Histology-Cytology (GIGA-R).

(1) Caprioli, R. M.; Farmer, T. B.; Gile, J. *Anal. Chem.* **1997**, *69*, 4751–4760.

Concerning MALDI-MSI, a lot of work has been done to improve the experimental workflow, such as for sample preparation and matrix deposition on the tissue.<sup>16</sup> Another field of research concerns the identification of the detected molecules. Indeed, mass spectrometry imaging provides information on molecular masses and spatial distributions of the compounds but does not allow direct identification. For this purpose, two methods are currently used. The first is tandem mass spectrometry, directly on tissue, for small molecules such as drugs<sup>17</sup> or for peptides up to 3.5 kDa.<sup>18</sup> The second method corresponds to an *in situ* enzymatic digestion prior to MS analysis. For example, this method, based on a “bottom-up” proteomic approach, has been successfully used for the characterization of tumor marker proteins.<sup>19,20</sup> Nevertheless, this method presents some drawbacks: the incubation of the enzyme takes time (usually 2 h) and some identification difficulties can arise because the complexity of the sample (more than one protein can be digested within the same spot of enzyme) influences the MS/MS spectra quality and therefore the data interpretation. To improve this method, some authors recently proposed to add a derivatization reaction after the digestion step to enhance the ion fragmentation yield.<sup>21</sup>

We chose to work with in-source decay (ISD) which corresponds to a fragmentation process occurring in the source region, right after the laser shot.<sup>22,23</sup> As ISD fragments are formed before the extraction, they are detected at their real *m/z* ratios in full scan MS mode (either in linear or reflector mode). ISD of proteins and peptides generally leads to c- and z-fragment ions, according to Roepstorff and Biemann’s nomenclature.<sup>24,25</sup> These fragments correspond to the cleavage of the N–C<sub>a</sub> bonds on the peptide backbone. The major drawback of ISD is the lack of precursor ion selection, which could lead to a complicated mass spectrum if more than one protein is present at the laser shot position. On the other hand, the main advantage of ISD is that there is no mass limitation since fragmentations occur in a very short time scale before ion acceleration (prompt fragmentations) and thus allow top-down sequencing approaches to be implemented.<sup>26–28</sup> However, first c- or z-ions are often absent from the mass spectrum because of matrix adduct peaks (until around *m/z* 1000). To circumvent this issue, it is possible to perform a “pseudo-MS<sup>3</sup>” experiment, called T<sup>3</sup>-sequencing.<sup>29,30</sup> In this technique, an ISD fragment is isolated and then fragmented as in a classical MS/MS experiment, allowing sequencing of the N-terminus or the C-terminus part of the protein if the selected ion is a c- or a z-ion, respectively. At last, ISD is very easy to perform since it only

- (16) Kaletas, B. K.; van der Wiel, I. M.; Stauber, J.; Guzel, C.; Kros, J. M.; Luider, T. M.; Heeren, R. M. *Proteomics* **2009**, *9*, 2622–2633.
- (17) Khatib-Shahidi, S.; Andersson, M.; Herman, J. L.; Gillespie, T. A.; Caprioli, R. M. *Anal. Chem.* **2006**, *78*, 6448–6456.
- (18) Minerva, L.; Clerens, S.; Baggerman, G.; Arckens, L. *Proteomics* **2008**, *8*, 3763–3774.
- (19) Groseclose, M. R.; Massion, P. P.; Chaurand, P.; Caprioli, R. M. *Proteomics* **2008**, *8*, 3715–3724.
- (20) Djidja, M. C.; Claude, E.; Snel, M. F.; Scriven, P.; Francese, S.; Carolan, V.; Clench, M. R. J. *Proteome Res.* **2009**, *8*, 4876–4884.
- (21) Franck, J.; El Ayed, M.; Wisztorski, M.; Salzet, M.; Fournier, I. *Anal. Chem.* **2009**, *81*, 8305–8317.
- (22) Brown, R. S.; Lennon, J. J. *Anal. Chem.* **1995**, *67*, 3990–3999.
- (23) Brown, R. S.; Feng, J.; Reiber, D. C. *Int. J. Mass Spectrom. Ion Processes* **1997**, *169–170*, 1–18.
- (24) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601.
- (25) Biemann, K. Mass Spectrometry. In *Methods Enzymology*, Vol. 193; McCloskey, J. A., Ed. Academic Press/Elsevier Inc.: London, U.K., 1990; pp 455–479.

requires the use of an “ISD favorable MALDI matrix” such as 2,5-dihydroxybenzoic acid (2,5-DHB) which is the most commonly used matrix for ISD.<sup>31,32</sup> However, since recent years, 1,5-diaminonaphthalene (1,5-DAN) has been shown to be very efficient to produce ISD fragments by increasing the fragmentation yield.<sup>33</sup> Our objective was then to develop a new original method allowing the identification of proteins directly on a tissue slice, without any further treatment.

## MATERIALS AND METHODS

**Materials.** The MALDI matrix 1,5-diaminonaphthalene (1,5-DAN) was purchased from Acros Organics (Geel, Belgium). All solvents used were HPLC grade quality (purity >99.9%).

**Harvesting of Tissues.** Eyes coming from pigs aged from 18 to 24 weeks were purchased from the slaughterhouse of Aubel (Detry S.A, Belgium). Immediately after harvesting, the eyes were dissected to extract the lenses which were frozen on powdered dry ice for 5 min. Eye lenses were kept at –80 °C and then placed at –20 °C 24 h before use.

A 6-months-old Balb C mouse was provided by the Central Animal Housing of Liège University. During intraperitoneal anesthesia with 60 mg/kg pentobarbital, the brain was harvested after sectioning the medulla and immediately frozen on powdered dry ice for 1 min. The brain was stored at –80 °C until use. All animal use was performed in strict accordance with the Animal Ethic Committee guidelines in force at Liège University.

**Tissues Preparation.** For the porcine eye lens, 18 μm thick equatorial sections have been realized on a cryostat (Microm, HM 5000, Prosan, Merelbeke, Belgium). To collect frozen sections, the “soft landing” technique using a thin layer of cold methanol on a ITO-coated glass slide (Bruker Daltonics, Bremen, Germany) was applied.<sup>34</sup> After the tissue sections were dried, they were dipped in an acetonitrile (ACN)/H<sub>2</sub>O solution (50:50, vol/vol) for 2 s. Sections were then allowed to dry in a desiccator for 10 min at room temperature before matrix deposition.

For the mouse brain, 12 μm-thick sections have been realized on a cryostat and thaw-mounted on a cold ITO-coated glass slide. Tissue sections were then washed with graded ethanol baths (70%, twice for 1 min and 100%, for 1 min) and allowed to dry in a desiccator at ambient temperature for at least 30 min before use. For more clarity, the experimental workflow (from this point) is summarized in a flowchart (see the Supporting Information, Figure S-1).

**ISD on Tissue.** To demonstrate the feasibility of ISD on tissue, we first applied the 1,5-DAN matrix solution with a micropipet. For both kinds of tissue, 0.3 μL of a saturated 1,5-DAN solution (ACN/H<sub>2</sub>O 50:50 vol/vol) was deposited and allowed to air-dry prior to MS analysis.

- (26) Quinton, L.; Demeure, K.; Dobson, R.; Gilles, N.; Gabelica, V.; De Pauw, E. *J. Proteome Res.* **2007**, *6*, 3216–3223.
- (27) Hardouin, J. *Mass Spectrom. Rev.* **2007**, *26*, 672–682.
- (28) Demeure, K.; Quinton, L.; Gabelica, V.; De Pauw, E. *Anal. Chem.* **2007**, *79*, 8678–8685.
- (29) Raska, C. S.; Parker, C. E.; Huang, C.; Han, J.; Glish, G. L.; Pope, M.; Borchers, C. H. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1034–1041.
- (30) Suckau, D.; Resemann, A. *Anal. Chem.* **2003**, *75*, 5817–5824.
- (31) Reiber, D. C.; Grover, T. A.; Brown, R. S. *Anal. Chem.* **1998**, *70*, 673–683.
- (32) Reiber, D. C.; Brown, R. S.; Weinberger, S.; Kenny, J.; Bailey, J. *Anal. Chem.* **1998**, *70*, 1214–1222.
- (33) Takayama, M.; Tsugita, A. *Electrophoresis* **2000**, *21*, 1670–1677.
- (34) Grey, A. C.; Schey, K. L. *Mol. Vis.* **2008**, *14*, 171–179.

**MALDI Imaging.** MALDI images of ISD fragments were recorded on a porcine eye lens section. 1,5-DAN solution was prepared at a concentration of 5 mg/mL in ACN/0.2% trifluoroacetic acid (TFA) 50:50 vol/vol. Application of the matrix solution was performed with an ImagePrep automated sprayer device equipped with the new spray head (Bruker Daltonics). The method was adapted from the Bruker standard method for 2,5-DHB, but the deposition was finally applied with two distinct stages. The first consisted of a phase of initialization with a spray power of  $18 \pm 23\%$  and a drying phase of 30 s. Once this first stage was finished, the sample was removed and the ImagePrep cleaned with methanol, because the 1,5-DAN matrix (even at 5 mg/mL) dirties the spray head and the aluminum membrane leading to a deleterious clogging of these parts. The second stage performed consisted of three phases with increasing spray powers (from  $18 \pm 23\%$  to  $25 \pm 30\%$ ) and number of cycles between a complete drying (every four cycles for the first phase and every six cycles for the last two). The complete preparation for 1,5-DAN deposition took approximately 90 min.

**Data Acquisition.** All data were acquired with an UltraFlex II TOF/TOF mass spectrometer (Bruker Daltonics) equipped with a Smartbeam laser (MLN 202, LTB) controlled by the FlexControl 3.0 software (Bruker Daltonics). Mass spectra were acquired in positive reflector mode, using a 25 kV acceleration voltage, 21.85 kV pulse voltage, 30 ns ion extraction delay, 1.8 kV detector gain, and 100 mV digitizer sensitivity. A low-mass ion deflector cutoff was set to 740 Da. The laser intensity was set to 60% (laser attenuator 64%), and 500 shots per spectrum were averaged. For T<sup>3</sup>-sequencing, the ISD fragment was selected and further fragmented using the “PSD-like” LIFT method. The acceleration voltage was 8 kV in the MALDI ion source and 19 kV for fragment postacceleration in the LIFT cell. The laser intensity was set to 98% (same attenuation as the MS mode), and a total of 12 000 mass spectra were accumulated.

For imaging experiments, FlexImaging 2.0 software (Bruker Daltonics) was used for acquisition and construction of two-dimensional density maps. To reduce the quantity of data, only half of the eye lens section was analyzed. The surface of the sample was divided into 100  $\mu\text{m}$ -wide pixels (leading to approximately 2000 points), and 500 laser shots were averaged on each. The laser frequency was set to 200 Hz, and the intensity was adjusted to 65%. The laser intensity was thus slightly higher than in previous settings because of the difference of crystallization between the manually and automated depositions.

**Data Analysis.** All data were treated with FlexAnalysis 3.0 and BioTools 3.1 (Bruker Daltonics). For ISD, T<sup>3</sup>-sequencing, and ISD fragments imaging, the mass spectra were smoothed (Gauss algorithm, 0.1  $m/z$ , 1 cycle) and the baseline was subtracted (TopHat algorithm).

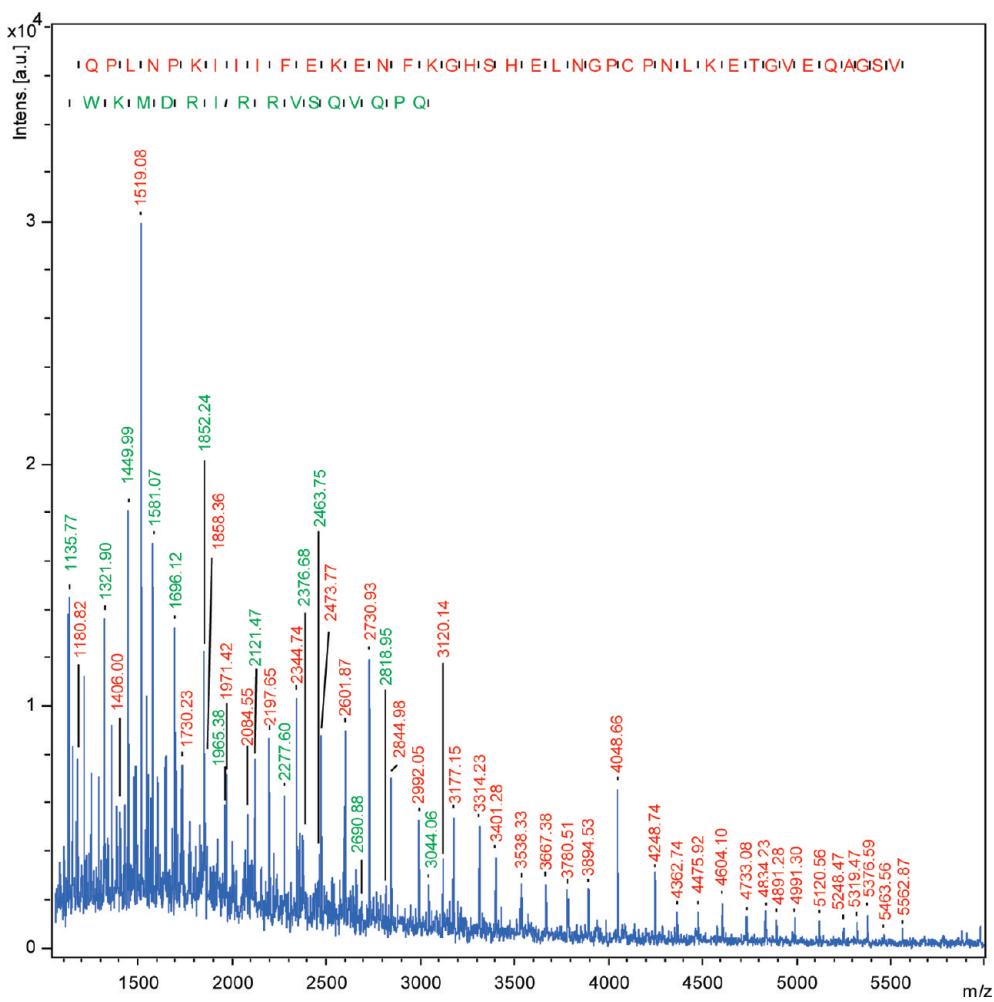
For ISD and T<sup>3</sup>-sequencing experiments, peaks on the mass spectrum were manually picked and the mass difference between two consecutive peaks was then automatically calculated by the software. For each mass difference, the software suggested an amino acid residue, leading to a sequence tag which was submitted to a query in protein databases using a BLAST program. Since the ISD process leads to a break of the N–C<sub>α</sub> bond, the proline residue cannot produce ISD fragments due to its cyclic nature. As this observation is very

general, a gap in c-ions of an ISD-sequencing experiment can be interpreted as a proline residue on the N-side followed by another amino acid, in which mass and position in the sequence can easily be determined. The BLAST (Basic Local Alignment Search Tool, <http://www.expasy.ch/tools/blast/>) research consists in the comparison between the submitted tag and every protein sequence existing in the databases (Swiss-Prot and TrEMBL for the program used). Some parameters can be modified, such as the taxonomy and the *E*-threshold, which represents the precision of the comparison (number of amino acids which must be identical in the tag and in the protein sequence). The *E*-threshold can be set from 0.001 to 10 000. For our interrogations, taxonomy was set either as “Mammalia” (for porcine eye lens analysis) or “Mus musculus” (for mouse brain analysis) and the *E*-threshold was set to 0.01 (to ensure both confidence and biological relevance). For each proposed protein, a score and an *E*-value are given. The *E*-value corresponds to the theoretical probability for an alignment with a higher score to occur randomly (lower is better). The score is a function of the length of the tag and of the similarities between consecutive amino acids of the tag with proteins from databases (higher is better).

For imaging, an ion density map was created for each signal present on the whole sample average mass spectrum. For signals exhibiting close localizations, a mass spectrum associated with a high-intensity pixel was extracted. The same data analysis procedure was then applied (manual peak picking, mass difference between two peaks, tag creation, and databases interrogation with BLAST).

## RESULTS AND DISCUSSION

**ISD on Porcine Eye Lens Slice.** Figure 1 depicts the ISD mass spectrum recorded on a porcine eye lens section with 1,5-DAN as a matrix, deposited by hand, for the mass range 1000–6000. Several peaks are visible, with a good signal-to-noise ratio. The mass spectrum seems to be complicated, and this is due to the lack of precursor ion selection. In our technique, any protein present at the surface of the sample may be fragmented, leading to one or even two ion series peaks, which can exhibit close *m/z* ratios or even be overlapped. It is true that an increasing number of proteins makes the ISD spectrum difficult to interpret and, for more than 3 proteins in a mix, the complexity of the spectrum cannot be overcome in the case of a manual interpretation. That is the reason why a good bioinformatics tool is necessary for interpretation of ISD mass spectra of complex mixtures (as tissue sections). Nevertheless, on the ISD mass spectrum of Figure 1, we were able to obtain a 40 amino acids-long tag: QPLNPKIIIF EQENFQGHSH ELNGPCPNLK ETGVEKAGSV. Because of the mass accuracy of mass measurements, some residues cannot be distinguished as Gln/Lys. For Leu/Ile, the distinction is even impossible in our conditions but, for more clarity, the tag corresponding to the sequence having the best score and *E*-value will always be given. To identify the protein corresponding to these ISD fragments, we performed a search in databases using BLAST. Eye lenses are rich in soluble proteins named crystallins which are divided into different families ( $\alpha$ -,  $\beta$ -, and  $\gamma$ -crystallins). Although their sequences are known for many animal species, only the  $\alpha$  family (chains A and B) and the  $\beta$ -crystallin B1 are sequenced for the pig (*Sus scrofa* taxonomy).



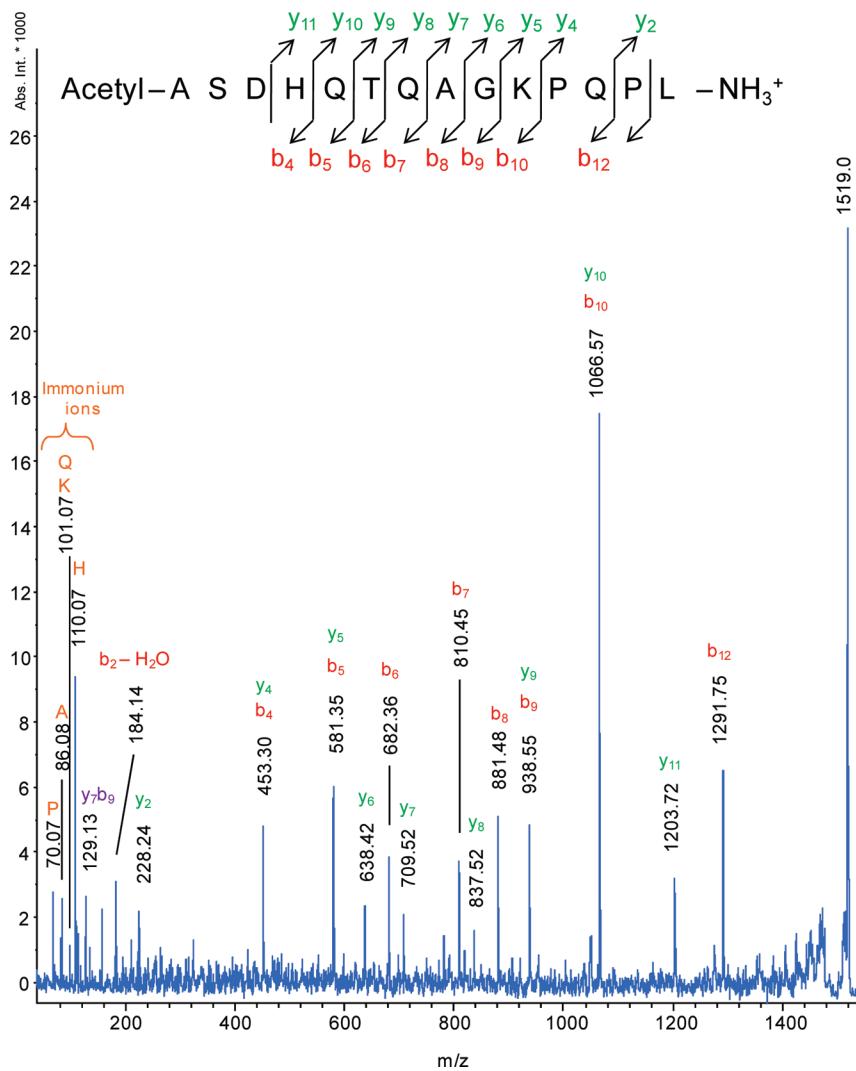
**Figure 1.** ISD mass spectrum recorded on a porcine eye lens slice with 1,5-DAN as the matrix. The tag and the masses indicated in red correspond to the N-terminus part of the  $\beta$ -crystallin B2 (c-ions series) and the green ones to the C-terminus part (z-ion series).

That is the reason why, for the interrogations in BLAST, we selected “Mammalia” as taxonomy instead of “*Sus scrofa*”. The protein with the highest score (102) and the best *E*-value ( $2 \times 10^{-22}$ ) was found to be the  $\beta$ -crystallin B2, for different species such as rabbit, beef, and guinea pig. From the primary sequence of this protein, we deduced that the observed peaks characterize ions for which the charge has been kept by the N-terminus part. Therefore, c-type ions were used to create the sequence tag. After examination of the primary sequences of these three proteins, some differences appeared. In Table S-1 in the Supporting Information are presented the sequence of the  $\beta$ -crystallin B2 for each species. In red is indicated the tag we obtained thanks to ISD fragments. For rabbit and beef, the N-terminus parts of the proteins are the same but for guinea pig, one amino acid is different (indicated in blue). T<sup>3</sup>-sequencing on the c<sub>14</sub>-ion (*m/z* 1519.08) was set to assess which sequence is the good one. The corresponding MS/MS spectrum is shown in Figure 2. We observed the quasi complete y- and b-ion series, confirming the sequence ASDHQTQAGK PQQPL, meaning that the N-terminal part of the porcine  $\beta$ -crystallin B2 is not identical to the one of the guinea pig. To determine if the porcine protein sequence is closer to the bovine or the rabbit, we used the z-ions

series (green annotations in Figure 1). We obtained a tag of 15 amino acids QPQVQSVRRIR DMQW which is found in the bovine sequence but not in the rabbit one, the first residue (Gln) being replaced by a His residue (indicated in orange in Table S-1 in the Supporting Information). We can thus state that the porcine sequence of the  $\beta$ -crystallin B2 is closer to the bovine one. Nevertheless, even if the primary sequences between positions +181 and +195 are the same for these proteins, the calculated masses of the z-ion series for each one exhibit a mass difference of 28 Da, showing an amino acid mutation at the very end of the protein.

To sequence the C-terminus part of the porcine  $\beta$ -crystallin B2, T<sup>3</sup>-sequencing was performed on z-ions (*m/z* 1449.99, 1321.9, and 1135.77) but not one led to a MS/MS spectrum with enough peaks to be interpretable. Consequently, we are not able to give the right sequence of the C-terminus part of the porcine  $\beta$ -crystallin B2. We can nevertheless conclude that sequences of porcine and bovine  $\beta$ -crystallins B2 are very close, which is probably true for the whole  $\beta$ -crystallin family as it was already established for  $\beta$ -crystallin B1.<sup>35</sup>

On the *m/z* 1100–2700 region of the same ISD mass spectrum (Figure S-2 in the Supporting Information), in spite of an important background noise, some other c- and z-ion series can be annotated. Thus, we obtained two tags of 12 and 11 amino acids long,



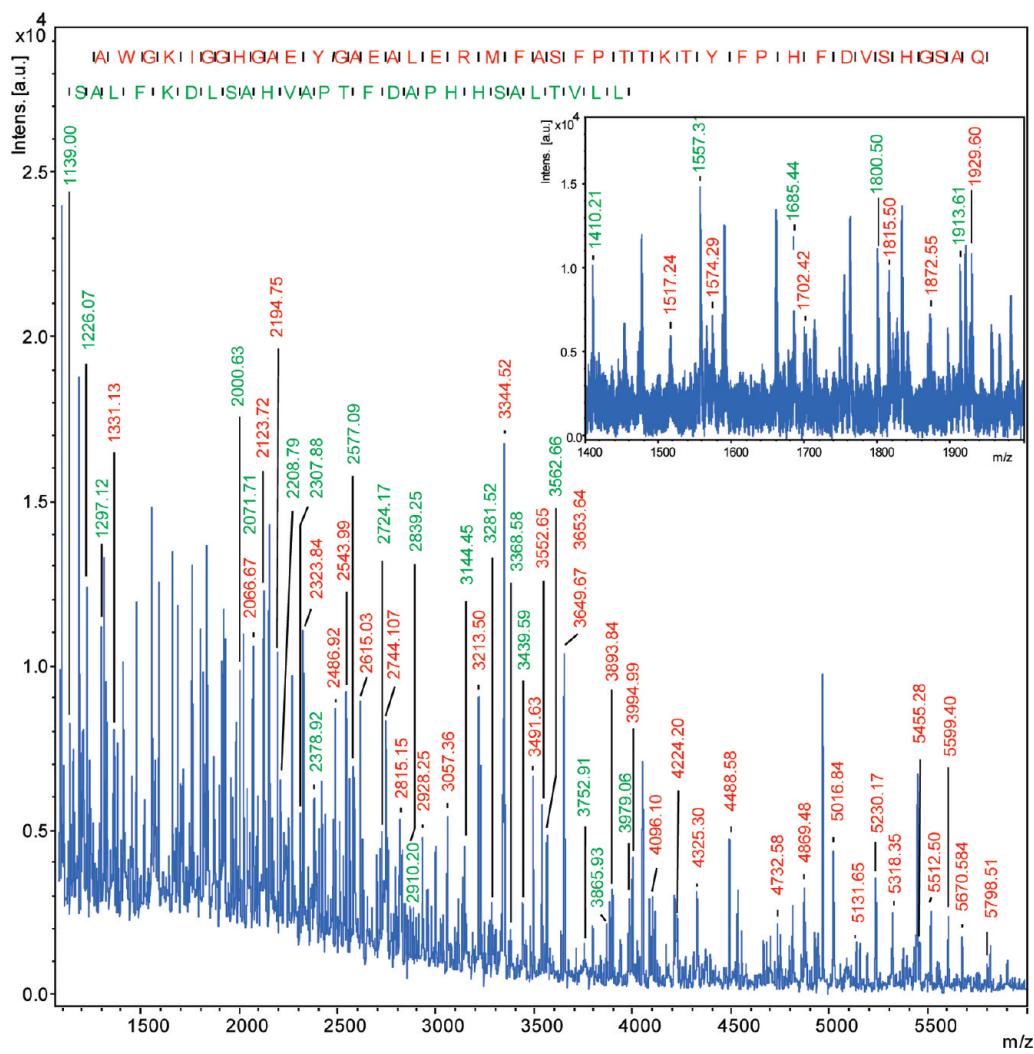
**Figure 2.** T<sup>3</sup>-sequencing mass spectrum recorded on a porcine eye lens slice with 1,5-DAN as the matrix. The ion at *m/z* 1519.0 was selected and fragmented.

respectively, SFQGRCYECS SD and LDWGAMNAKV G. The interrogation of databases for the first tag gave 13 results for 5 different species (mouse, rat, dog, Western gray kangaroo, and beef), all being  $\gamma$ -family crystallins. For the second tag (*z*-ion series), we obtained a more restricted list of 4  $\gamma$ -crystallins for 3 different species, all with a score of 39 and an *E*-value of  $2 \times 10^{-3}$ . These proteins are the  $\gamma$ -crystallin D of the Western gray kangaroo, the  $\gamma$ -crystallin B for the dog, and the  $\gamma$ -crystallins B and A for the beef (the last one will not be further considered since 4 amino acids out of 12 are different for the N-terminus part, see Table S-2 in the Supporting Information). These results show how it is important to have access to N-terminus and C-terminus ion series because a more precise identification can be made if both types of ions are used to search in the databases. As for the  $\beta$ -crystallin B2, there is a difference between the experimental and calculated (based on the sequences) masses for both ions series. For the N-terminus part, a mass difference of 30 Da is observable for the *c*<sub>10</sub> ion (which is identical for all proteins). This difference is explained by the modification of the Gly residue by a Ser residue in the porcine protein. The mass difference between these two residues is 30.0106 Da

which matches perfectly with the experimental mass difference. On the other hand, for the C-terminus part, the mass differences vary from +1 to +67 Da approximately, according to the considered species. These differences could be explained by the mutation of one amino acid or more. Unfortunately, we were not able to determine the nature of this modification because the intensity of all *z*-ions was too weak to allow an efficient fragmentation in T<sup>3</sup>-sequencing.

From these experiments, we can conclude that ISD (together with T<sup>3</sup>-sequencing) is a valuable tool to sequence and identify proteins directly on a tissue section, even if the genome of the studied species has not been fully sequenced. One critical aspect to perform T<sup>3</sup>-sequencing is the need for an intense ion peak to be selected and fragmented. For less abundant species, a solution could be to perform ISD and T<sup>3</sup>-sequencing with a mass spectrometer allowing the accumulation of the signal of interest, as in a Fourier transform mass spectrometer (FTMS) instrument.

**ISD on Mouse Brain Slice.** Figure 3 depicts the mass spectrum recorded on a mouse brain section with 1,5-DAN as a matrix, deposited by hand, for the mass range 1000–6000. Several peaks are visible and, in spite of a deformed baseline in the low mass



**Figure 3.** ISD mass spectrum recorded on a mouse brain slice with 1,5-DAN as the matrix. The tag and the masses indicated in red correspond to the N-terminus part of the hemoglobin subunit  $\alpha$  (c-ion series) and the green ones to the C-terminus part (z-ion series). The inset is a zoom of the 1400–2000  $m/z$  region of the spectrum.

range of the spectrum, the signal-to-noise ratio is high enough to allow some interpretation. By calculating the mass difference between two consecutive peaks, we were able to obtain the following 42 amino acid-long tag AWGKIGGHGA EYGAEALERMR FASFPTTKTY FPHFDVSHGS AQ. The protein with the highest score (104) and the best  $E$ -value ( $1 \times 10^{-23}$ ) was found to be the hemoglobin subunit  $\alpha$  (Swiss Prot entry P01942). From the primary sequence of this protein, we deduced that the observed peaks correspond to c-type ions.

Consequently, we looked for their counterparts (z-type ions) and we were able to create a 27 amino acid-long tag (LLVTLASHHP ADFTPAVHAS LDKFLAS). When submitting this tag to a search using BLAST, we also identified the hemoglobin subunit  $\alpha$ , with a score of 86 and an  $E$ -value of  $4 \times 10^{-18}$ . The detection of hemoglobin on a mouse brain section is thus unambiguous and relevant but not surprising since blood naturally irrigates the brain. Nevertheless, these two ion-series are not the only ones we can detect on the mass spectrum. Figure S-3 in the Supporting Information shows the same mass spectrum with another peak picking, allowing the creation of another tag SKYLATASTM DHARHGFLPR HR. The submission of this tag in databases led to the identification of the myelin basic protein (Swiss Prot entry P04370) with a score of 78

and an  $E$ -value of  $1 \times 10^{-15}$ . This identification is relevant since MBP is one of the most abundant proteins in the central nervous system.<sup>36</sup> As for the hemoglobin subunit  $\alpha$ , we deduced from the examination of the primary sequence that the ions detected were c-type ions, corresponding to the N-terminus part of the protein. However, not less than 13 isoforms are known for the murine MBP. Other information can be found in the mass spectrum, allowing us to go further in the identification of the protein. The first c-type ion we observed on the mass spectrum is at  $m/z$  1098.86 (smaller ions cannot be detected since they are lost in the signals of intense matrix aggregates). The mass of this ion means that the serine (first amino acid of the tag) is localized at the beginning of the primary sequence of the detected isoform (around the amino acid +10). Sequences and modifications of each isoform are summarized in Table 1. For the two first isoforms, which are composed of 250 and 195 amino acids, respectively, our tag should begin at position +143, which is incompatible with a c-ion at  $m/z$  1098. We can also exclude the third isoform as its primary sequence does not contain the tag found.

(36) Boggs, J. *Cell. Mol. Life Sci.* 2006, 63, 1945–1961.

**Table 1. Primary Sequences and Modifications of the 13 Isoforms of the MBP<sup>a</sup>**

Isoform	Primary sequence	Other modifications
#1	MGNHSGKRELSAEKASKDGEIHRGEAGKKRSVGKLSQTASESDVFGEADAIQNNG TSAEDTAVTDSKHTADPKNNWQGAHPADPGNRPHLIRLFSRDAKGREDNTFKDRPS ESDELOTIQEDPTAASGGLDVMASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGI LDSIGRFFSGDRGAPKRGSQKDSHTRTTHYGLSPQKSQHGRQDENPVVHFFKNIVT PRTPPSQQGKGGDRSRSGSPMARR	-
#2	MGNHSGKRELSAEKASKDGEIHRGEAGKKRSVGKLSQTASESDVFGEADAIQNNG TSAEDTAVTDSKHTADPKNNWQGAHPADPGNRPHLIRLFSRDAKGREDNTFKDRPS ESDELOTIQEDPTAASGGLDVMASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGI LDSIGRFFSGDRGAPKRGSQKVSSEPE	-
#3	MGNHSGKRELSAEKASKDGEIHRGEAGKKRSVGKLSQTASESDVFGLTHENYPLW LPAPEVAARPDPW	-
#4	MASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGILDSIGRFFSGDRGAPKRGSQK VPWLQSRSPLPShARSRPGLCHMYKDSHTRTTHYGLSPQKSQHGRQDENPVVHF FKNIVTPRTPPSQGKGRGLSLSRFSWGAEGQKPGFYGGGRASDYKSAHKGFKGAY DAQGTLSKIFKLGGDRSRSGSPMARR	Met-1 is removed. Contains N-acetylalanine at position 2.
#5	MASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGILDSIGRFFSGDRGAPKRGSQK DSHTRTTHYGLSPQKSQHGRQDENPVVHFKNIVTPRTPPSQGKGRGLSLSRFSW GAEGQKPGFYGGGRASDYKSAHKGFKGAYDAQGTLSKIFKLGGDRSRSGSPMARR	Met-1 is removed. Contains N-acetylalanine at position 2.
#6	MASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGILDSIGRFFSGDRGAPKRGSQK VPWLQSRSPLPShARSRPGLCHMYKDSHTRTTHYGLSPQKSQHGRQDENPVVHF FKNIVTPRTPPSQGKGRGLSLSRFSWGGDRSRSGSPMARR	Met-1 is removed. Contains N-acetylalanine at position 2.
#7	MASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGILDSIGRFFSGDRGAPKRGSQK DSHTRTTHYGLSPQKSQHGRQDENPVVHFKNIVTPRTPPSQGKGAEGQKPGFGY GGGRASDYKSAHKGFKGAYDAQGTLSKIFKLGGDRSRSGSPMARR	Met-1 is removed. Contains N-acetylalanine at position 2.
#8	MASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGILDSIGRFFSGDRGAPKRGSQK DSHTRTTHYGLSPQKSQHGRQDENPVVHFKNIVTPRTPPSQGKGRGLSLSRFSW GGDRSRSGSPMARR	Met-1 is removed. Contains N-acetylalanine at position 2.
#9	MASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGILDSIGRFFSGDRGAPKRGSQK VPWLQSRSPLPShARSRPGLCHMYKDSHTRTTHYGLSPQKSQHGRQDENPVVHF FKNIVTPRTPPSQGKGAEGQKPGFYGGGRASDYKSAHKGFKGAYDAQGTLSKIFKL GGDRSRSGSPMARR	Met-1 is removed. Contains N-acetylalanine at position 2.
#10	MASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGILDSIGRFFSGDRGAPKRGSQK DSHTRTTHYGLSPQKSQHGRQDENPVVHFKNIVTPRTPPSQGKDFVPGDHVVNV SVTVTSFSSSQGRGLSLSRFSWGAEGQKPGFYGGGRASDYKSAHKGFKGAYDAQGT LSKIFKLGGDRSRSGSPMARR	-
#11	MASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGILDSIGRFFSGDRGAPKRGSQK VPWLQSRSPLPShARSRPGLCHMYKDSHTRTTHYGLSPQKSQHGRQDENPVVHF FKNIVTPRTPPSQGKDFVPGDHVVNVSVTVTSFSSSQGRGLSLSRFSWGGDRSRSG PMARR	-
#12	MASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGILDSIGRFFSGDRGAPKRGSQK VPWLQSRSPLPShARSRPGLCHMYKDSHTRTTHYGLSPQKSQHGRQDENPVVHF FKNIVTPRTPPSQGKGRDRSRSGSPMARR	-
#13	MASQKRPSQRSKYLATASTMDHARHGFLPRHRDTGILDSIGRFFSGDRGAPKRGSQK DSHTRTTHYGLSPQKSQHGRQDENPVVHFKNIVTPRTPPSQGKGGDRSRSGSPM ARR	-

<sup>a</sup> The tag corresponding to the N-terminus part is indicated in red. In blue are indicated the amino acids differing from a sequence to another.

For isoforms 4–13, many differences can be observed if compared with the primary sequence of the isoform 1 (considered as the “canonical” sequence). One of these modifications is common to all and consists in the missing of the 133 first amino acids. By removing them, the serine residue is now at position +11 (MASQKRPSQRS), which should fit perfectly with the mass of the first detected c-ion. Therefore we calculated the theoretical mass of this first c-ion, corresponding to the complete tag MASQKRPSQRS, and we observed a mass difference of +88.78 Da if compared to the experimental mass. This result led us to eliminate isoforms 10–13. For the remaining isoforms (no. 4–9), two more modifications are described. The first amino acid (Met) is removed, and the Ala residue at position +2 is acetylated. We

calculated the mass of the c-ion corresponding to the tag ASQKRPSQRS, with an acetyl group on the alanine residue. We compared to the experimental mass and we found an error of –0.360 Da (328 ppm), probably due to a poor calibration of the mass spectrum. This confirms that the detected protein is among these six isoforms. At last, as shown in Table 1, other modifications can differentiate the isoforms no. 4–9. At the end of the sequences, some differences could help to reduce the number of putative isoforms. As previously said, the detection of z-type ions is of great importance to determine which sequence is the good one. On the mass spectrum, a series of z-ions allowed us to obtain the tag LSRFSWGGRD. In comparison with the sequence of isoforms no. 4–9, we eliminated isoforms no. 4, no. 5, no. 7, and no. 9, as

**Table 2. Comparison of Protein Identification Results Obtained by *in Situ* Digestion on a Rat Brain Section and ISD on a Mouse Brain Section<sup>a</sup>**

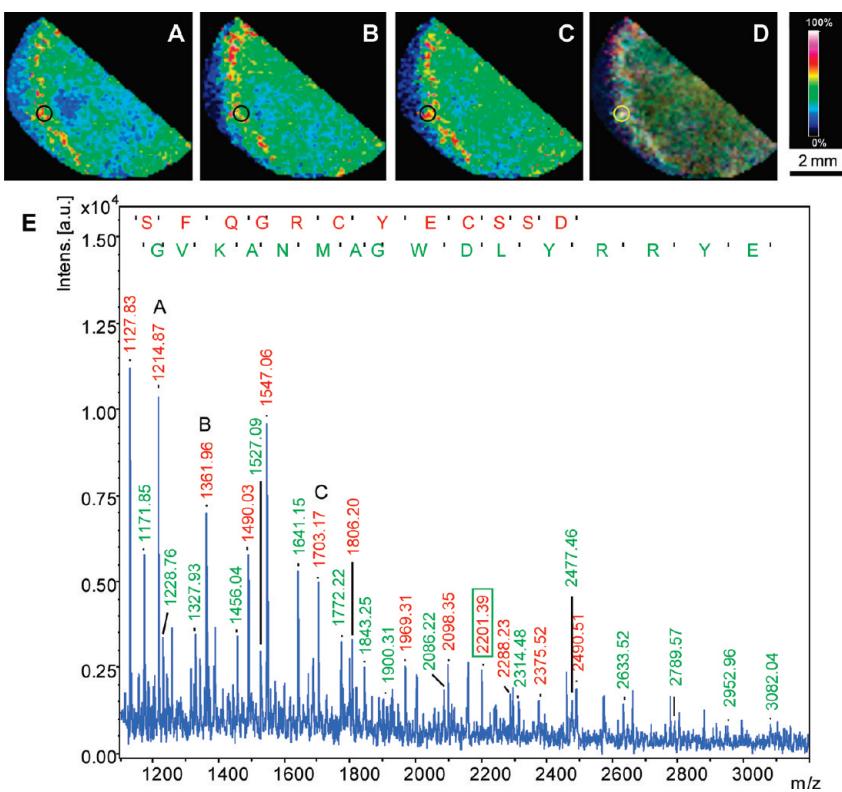
method	<i>in situ</i> tryptic digestion followed by the derivatization of peptides on a rat brain section	ISD on a mouse brain section
results	identification of hemoglobin subunit $\beta$ with a score of 103 and an <i>E</i> -value of $1.8 \times 10^{-9}$	identification of hemoglobin subunit $\alpha$ with a score of 104 and an <i>E</i> -value of $1 \times 10^{-23}$
sample preparation time	identification of myelin basic protein with a score of 94 and an <i>E</i> -value of $9.6 \times 10^{-9}$	identification of myelin basic protein with a score of 78 and an <i>E</i> -value of $1 \times 10^{-15}$
	3 h 0 min	1 h 30 min

<sup>a</sup> *In situ* digestion results come from ref 21. The sample preparation time for *in situ* digestion + derivatization protocol was estimated according to information given by the authors in ref 21 (1 h for enzyme incubation, 1 h for derivatization reaction). We added 1 h more for matrix deposition.

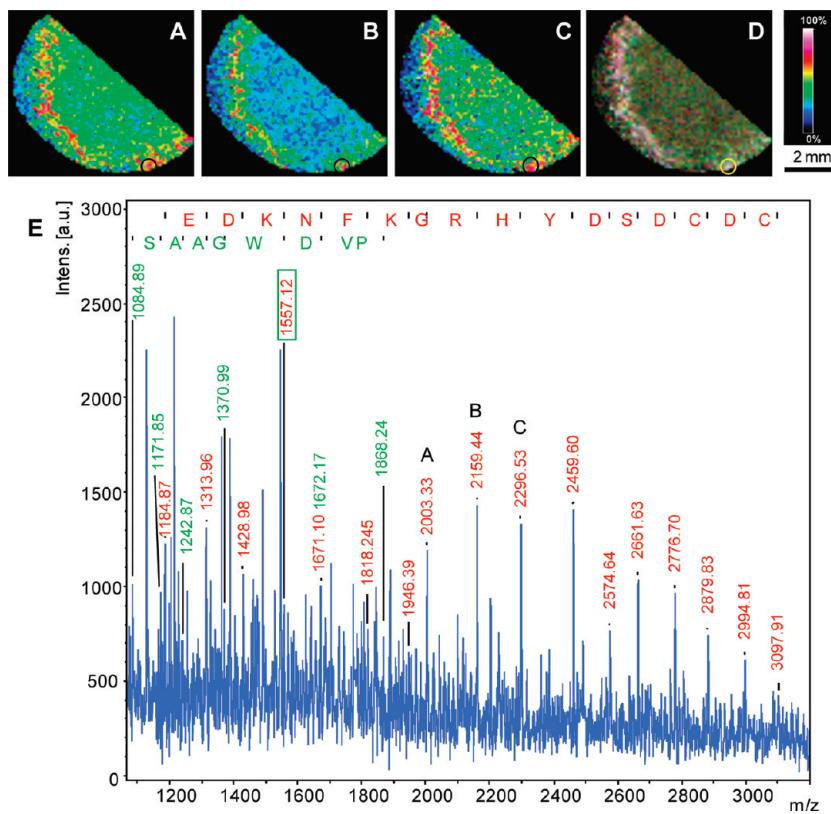
their sequences are, for the same amino acids positions, LSKIFKLG-GRD. The remaining isoforms are the no. 6 and no. 8, but we are not able to distinguish the one we really detect, since the difference between their sequences consists in an addition of 26 residues between positions +58 and +83 (for the isoform no. 6) which are inaccessible with our technique. From this example, we can conclude how it is important to obtain the N-terminus and C-terminus ion series to improve the quality of the identification and that ISD can help to distinguish different isoforms of a unique protein.

**Comparison of ISD Results with *in Situ* Digestion Results.** Very recently, Franck et al. proposed to add a derivatization reaction using TMPP (*N*-succinimidyl carbonylmethyltris(2,4,6-trimethoxyphenyl)phosphonium bromide) to the *in situ* digestion protocol used to identify proteins directly from a tissue section.<sup>21</sup> They applied this new method for the analysis of a rat brain section. Table 2 shows a comparison of the results obtained by

Franck et al. with an *in situ* tryptic digestion followed by the derivatization reaction and our results, obtained with ISD. In both cases, a hemoglobin subunit was identified. The score given to the identification is the same but the *E*-value, which represents the probability of a wrong identification, is far better in our case. Also, the myelin basic protein is detected and identified with a better score for *in situ* digestion but the *E*-value is one million-times better for the ISD method. Moreover, when we compare the time needed, from the beginning of the sample preparation to the MS analysis, two times less time is needed for ISD than for *in situ* digestion associated with derivatization. We should also factor MS/MS data acquisition times ( $T^3$  and after *in situ* digestion) as well as total data analysis times, but this can only be done with difficulty because conditions in both workflows vary in such a degree. Finally, results are comparable (Franck et al. only identified one more protein as a tubulin), leading,



**Figure 4.** MALDI images of ISD fragments at (A)  $m/z$  1214.9, (B)  $m/z$  1361.9, and (C)  $m/z$  1703.2, recorded on a porcine eye lens slice. (D) Overlay of the three previous ion images. The black and the yellow circles indicate the region where the mass spectrum has been extracted (high intensity pixel). (E) ISD mass spectrum extracted from a high intensity pixel. The tag and the masses indicated in red correspond to the N-terminus part of the  $\gamma$ -crystallin B (c-ions series), the green ones to the C-terminus part (z-ions series). The red green-framed annotation indicates a peak for which the  $m/z$  ratio corresponds to both the c- and z-ions. The ion peaks in which images are shown are labeled with letters A, B, and C.



**Figure 5.** MALDI images of ISD fragments at (A)  $m/z$  2003.4, (B)  $m/z$  2159.4, and (C)  $m/z$  2296.6 recorded on a porcine eye lens slice. (D) Overlay of the three previous ion images. The black and the yellow circles indicate the region where the mass spectrum has been extracted (high-intensity pixel). (E) ISD mass spectrum extracted from a high-intensity pixel. The tag and the masses indicated in red correspond to the N-terminus part of  $\beta$ -crystallin S (c-ion series), the green ones to the C-terminus part (z-ion series). The red green-framed annotation indicates a peak for which the  $m/z$  ratio corresponds to both the c- and z-ions. The ion peaks in which images are shown are labeled with letters A, B, and C.

in both cases, to unambiguous identifications, but our  $E$ -values are better because longer tags are obtained with ISD, enhancing the sequence coverage.

**ISD Imaging on Porcine Eye Lens Slice.** It is also possible to map ISD fragments during a MALDI imaging experiment. The only requirement is then to use an “ISD-promoting matrix” and to record data over the right mass range. Figure 4 shows ion images for signals at  $m/z$  1214.9 (Figure 4A), 1361.9 (Figure 4B), and 1703.2 (Figure 4C), respectively. The very close localizations of these fragments led us to think that they came from the same protein. The extracted mass spectrum from a high-intensity pixel is shown in Figure 4E. The background noise is a little bit high, but the signal-to-noise ratio of detected peaks is high enough to allow some sequencing. Thus we were able to create two tags: SFQGRCYECSD and EYRRYLDWGA MNAKVG. These tags are almost the same than those described previously. We even improved the length of the C-terminus tag (from 12 to 16 amino acids). The interrogation in databases using BLAST gave the same result, that is to say the sequence of the protein which led to these fragments is very close to the ones of the bovine, canine, and gray kangaroo  $\gamma$ -crystallin B.

Another example is given in Figure 5. Parts A, B, and C of Figure 5 show the ion images of ISD fragments at  $m/z$  2003.4, 2159.4, and 2296.6, respectively. As previously shown, great similarities in the distributions of these ions made us think that they were originating from the same protein. The extracted mass

spectrum is given in Figure 5E. Unless the background noise is very high, we were able to create two tags again: EDKNFQGRHY DSDCDC and PVDWGAAS. The protein with highest score (59) and best  $E$ -value ( $2 \times 10^{-9}$ ) was the bovine  $\beta$ -crystallin S. Here again, even if there is a similarity for the tag and the bovine primary sequence, there is a mass difference of approximately 30 Da between them for the c-ion series.

For the z-ion series, this difference is even higher since the mass difference reaches approximately 200 Da. These differences show that a part of the primary sequence of the  $\beta$ -crystallin S is common for pig and beef but not the N- and C-termini of the proteins. As these data were acquired during an imaging (and so, automatic) run, no T<sup>3</sup>-sequencing was performed.

Two more proteins have been identified with this “imaging strategy” (data not shown). The  $\alpha$ -crystallin A (Swiss Prot entry P02475) was identified without any doubt since this protein has already been sequenced for the pig (score 74 and  $E$ -value  $6 \times 10^{-14}$ ). On another pixel, the bovine  $\beta$ -crystallin B2 was identified with a score of 39 and an  $E$ -value of  $2 \times 10^{-3}$ .

## CONCLUSIONS

Protein identification is a crucial issue in MALDI mass spectrometry imaging. Some years ago, Chaurand et al. already used ISD to characterize spermine-binding protein but this study was carried out after mouse prostate lobes blotting on a polyeth-

ylene membrane.<sup>37</sup> Our work is thus the first report on the use of in-source decay directly on a tissue section. ISD on tissue allows for *de novo* sequencing of unknown proteins and the N- and C-terminus parts can even be determined thanks to T<sup>3</sup>-sequencing, provided that ISD fragments with sufficient intensity are generated. Ions accumulation in a trap before mass measurement will allow overcoming limitations in the other case. Associated with MALDI imaging, ISD should become a powerful tool as the mapping of all ISD fragments present on the average mass spectrum (i) gives an idea of which fragments are coming from the same protein as they exhibit very close localizations and (ii) allows extracting the MS spectrum of one pixel for which the intensity is high for a majority of fragments, facilitating the data interpretation. Moreover, ISD is fast and easy to perform since no special sample treatment is required, which is an advantage, as the workflow of an imaging experiment is already quite long. Nevertheless, it has to be admitted that the absence of precursor ion selection is problematic. Indeed, the presence of different proteins within the same pixel (100 μm wide approximately) leads to complicated mass spectra from which a manual exploitation is fastidious. The solution could be to operate with smaller pixels (30 μm, for example)

(37) Chaurand, P.; DaGue, B. B.; Ma, S.; Kasper, S.; Caprioli, R. M. *Biochemistry* 2001, 40, 9725–9733.

to limit the number of proteins but in this case, we may wonder about the sensitivity of the instrument. We think that adequate sample treatments allied to improvements of the bioinformatics tools used to interpret data should greatly enhance the capabilities of the technique. Finally, MALDI-ISD imaging should become a new tool for biomarkers identification, avoiding the time-consuming extraction/purification/separation steps which are necessary to identify a potential biomarker candidate.

#### ACKNOWLEDGMENT

D.D is a postdoc fellow of the ULg Research Council. V.B. is indebted to Lionel Project (Grant Number R.RWAL.1793) for a Ph.D. fellowship. L.Q. is a postdoc fellow of the FNRS (Fonds National de la Recherche Scientifique, Belgium). The FNRS and Walloon Region contributed to the mass spectrometry funding.

#### SUPPORTING INFORMATION AVAILABLE

Additional information as noted in text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Received for review December 17, 2009. Accepted March 23, 2010.

AC902875Q