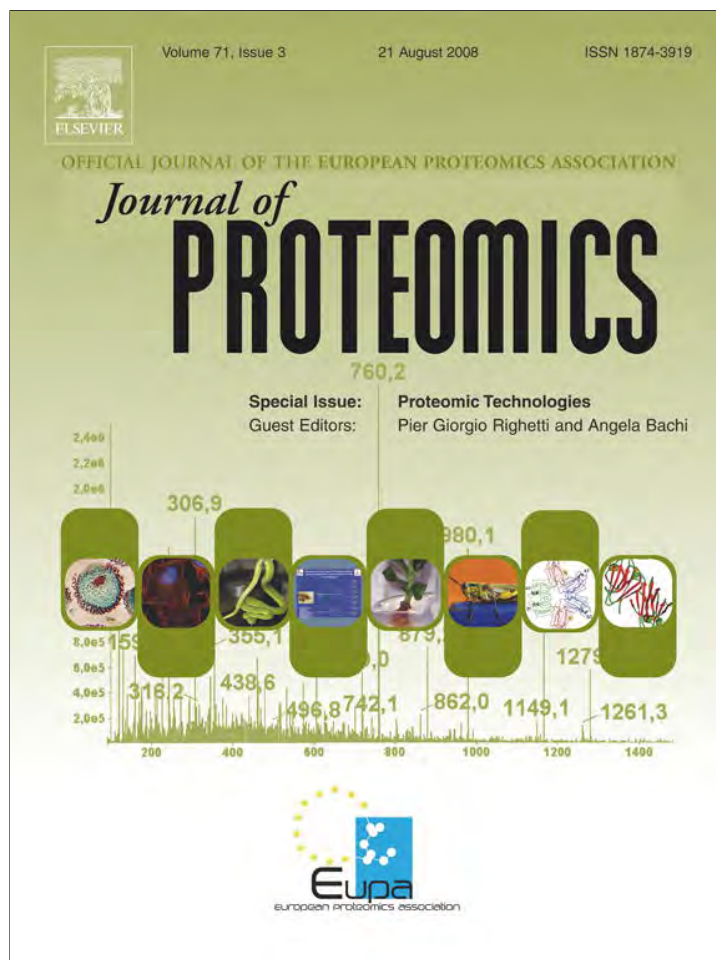


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

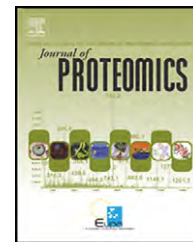


This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>

available at www.sciencedirect.comwww.elsevier.com/locate/jprot

MITICS (MALDI Imaging Team Imaging Computing System): A new open source mass spectrometry imaging software

Olivia Jardin-Mathé^a, David Bonnel^a, Julien Franck^a, Maxence Wisztorski^a,
Eduardo Macagno^b, Isabelle Fournier^{a,*}, Michel Salzet^{a,*}

^aUniversity of Lille1, CNRS FRE 2933, MALDI Imaging team, Laboratoire de Neuroimmunologie des Annélides, Villeneuve d'Ascq, France

^bSection of Cell and Developmental Biology, Division of Biological Sciences, University of California San Diego, La Jolla, CA 92093, USA

ARTICLE INFO

Keywords:

MALDI imaging
Mass spectrometry
Software

ABSTRACT

MITICS is a new software developed for MALDI imaging. We tried to render this software compatible with all types of instruments. MITICS is divided in two parts: MITICS control for data acquisition and MITICS Image for data processing and images reconstruction. MITICS control is available for Applied BioSystems MALDI-TOF instruments and MITICS Image for both Applied BioSystems and Bruker Daltonics ones. MITICS Control provides an interface to the user for setting the acquisition parameters for the imaging sequence, namely set instruments acquisition parameters, create the raster of acquisition and control post-acquisition data processing, and provide this settings to the automatic acquisition software of the MALDI instrument. MITICS Image ensures image reconstruction, files are first converted to XML files before being loaded in a database. In MITICS image we have chosen to implement different data representations and calculations for image reconstruction. MITICS Image uses three different representations that have shown to ease extraction of information from the whole data set. It also offers image reconstruction base either on the maximum peak intensity or the peak area. Image reconstruction is possible for single ions but also by summing signals of different ions. MITICS was validated on biological cases.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Mass spectrometry analyzes proteins based on the mass-to-charge (m/z) ratio of ionized molecules. A typical mass spectrometer consists of an ionization source, a mass analyzer, and a detector for counting the number of analytes at each m/z ratio. In addition, mass spectrometers are commonly coupled with separation devices, such as liquid chromatography instrumentation. Aside from providing rapid data on various proteins present in biological samples, mass spectrometry also offers specific information about post-translational modifications associated with, for example, a particular disease state. Recent instrument developments and electronic improvements now place mass spectrometry at the

forefront of many biological applications. In particular, this technology has been crucial in efforts to identify the human proteome. In combination with classical biochemical methods, mass spectrometry has been largely responsible for revealing an expanded biological diversity at the protein level, due to alternative splicing and post-translational modifications that could not be predicted by genome sequence analysis alone.

Although mass spectrometry technologies are efficient in the identification of novel proteins, until recently they did not allow the determination of chemicals distribution or the regulation of peptide/proteins expressed in tissues or single cells. With new emerging mass spectrometry technologies, it has now been demonstrated that direct tissue analysis is

* Corresponding authors.

E-mail addresses: Isabelle.fournier@univ-lille1.fr (I. Fournier), michel.salzet@univ-lille1.fr (M. Salzet).

feasible using matrix-assisted laser desorption/ionization (MALDI) sources. A major advantage of direct MALDI analysis is to avoid time-consuming extraction, purification or separation steps, which also have the potential for producing artifacts. Therefore, direct MALDI analysis in tissue sections enables the acquisition of cellular expression profiles while maintaining cellular and molecular integrity [1–7]. With automation and the ability to reconstruct complex spectral data with imaging software, it is now possible to produce multiplex imaging maps of selected bio-molecules within tissue sections [8,9]. Thus, direct MALDI analysis obtained from tissue sections is convertible into imaging maps, a method now known as MALDI imaging. MALDI imaging combines the power of mass spectrometry, namely exquisite sensitivity, with unequivocal positional information within a normal morphological context [8,9]. Every MALDI imaging experiment is comprised of two steps: (a) sample preparation and treatment, and (b) data acquisition and processing. Given the nature and magnitude of the physical measurements, computationally powerful and transparent data processing and displaying tools are a critical requirement for the successful analysis of the large number of spectra obtained in any MALDI imaging experiment.

Currently, four principal MALDI imaging software packages are available, some commercially (Table 1). The original one, MMSIT, was written by M. Stoeckli for use on Applied Biosystems MALDI-TOF and MALDI-TOF/TOF instruments [8]. This software is available at <http://www.maldi-msi.org>. For Applied Biosystems, BioMap and more recently Tissue view 1.0 have been developed for 4700 and 4800 MALDI-TOF/TOF™ Analyzers, QStar Pulsar System, QSTAR® XL Hybrid LC/MS/MS System and the QSTAR. Bruker, Shimadzu and Thermo Fisher have developed their own software's namely FlexImaging, Axima2 Analyze and ImageQuest. The latest ones, Create Target and Analyze this! are two new software assisting MALDI-IMS on Bruker Reflex IV and Ultraflex II instruments [10]. We have developed MITICS to have a software that is as possible workable with all type of data whatever the instrument used for both data acquisition when necessary and for data processing. Data processing for MALDI imaging is a difficult part that is always a compromise between finding the processing that better extracts real biological data and processing that could generate non relevant biological information as translate into image representations. Here, we have been interested by adding new features to the software for trying to extract as many data as possible. We have also aim in releasing software that is open and implementable by who ever wants to. MITICS is composed of two parts namely MITICS Control and MITICS Image which are respectively dedicated to data acquisition and to data processing and image reconstruction. MITICS Control is actually available for MALDI-TOF Applied BioSystems instruments. MITICS Image is actually interfaced with Applied Biosystems and Bruker Daltonics data files format. MITICS converts all reading files, such as .dat files in Data Explorer for Applied Biosystems, to .xml files, the format used by FlexAnalysis for Bruker Daltonics. MITICS uses two different interfaces, one to load all the spectra from each run into a single database that is readily available for repeated analysis, and the second to reconstruct and display interactively the molecular images of interest. The

software also offers two alternative computational tools to create molecular maps, using the maximum intensity value of a peak to show where the sample is present, or use the integrated area of the peak above a set threshold. This presents an interest for MALDI-TOF data since peaks are broadenings and decreasing in intensity with increasing m/z values. Thus as a general trend, maximum intensity of a peak shows to be of interest for smaller m/z values whereas peak area is better fitting for higher m/z ratios. Threshold values for minimum and maximum have also been included to take into account for signal/noise ratio or intensity variations between experiments. This allows for defining what is the minimum intensity or peak area to be taken into account and considered as a real signal and what is the maximum intensity of a peak excluding intensity variations between collected data at different coordinates. MITICS offers also three different representations of the data collected, the classical averaged spectrum as well as two new representations that are the sum of all the collected mass spectra as well as a spectrum displaying for each m/z the maximum intensity as found in the data set. These different representations have shown to enlarge the number of signals that have a biological significance, i.e. better exploitation of data can be taken out of using these three representations while avoiding to browse data set m/z by m/z . MITICS is presented here on two different biological applications for validation of the software running. MITICS was applied to a case of biomarker following in brain tissues of rats treated with 6-OHDA, fixed in formalin and embedded for the past 12 years in paraffin, and to peptide mapping in dissected leech embryo whole mounts.

2. Experimental

2.1. Materials

α -cyano-4-hydroxycinnamic acid (HCCA), ammonium bicarbonate, trisma base, toluene, ethanol (HPLC grade), acetonitrile (HPLC grade), methanol (HPLC grade), Trifluoroacetic acid, aniline, Angiotensin II, Des-Arg-Bradykinin, Substance P, ACTH 18–39, ACTH 7–38 and bovine Insulin were obtained from Sigma–Aldrich and used as described. Trypsin was obtained from Promega.

2.2. Animal surgery and treatment

Wistar male rats (180–200 g) were anesthetized with pentobarbital (50 mg/kg, i.p.) and infused over 8 min with 6-OHDA (8 mg in 4 ml of 0.05% ascorbic acid in saline) at coordinates A 5 23.8 mm, L 5 1.5 mm, H 5 28.5 mm [11].

Leech embryos were obtained from a laboratory colony. Prior to dissection, they were anesthetized in 8% Ethanol Ringer's, then opened dorsally and pinned flat in a Sylgard-covered dish. After removal of the yolk, the embryos were briefly submerged in Methanol, transferred to metal-coated (ITO) slides and gently dried on a hot plate at ~45 °C.

2.3. Tissue fixation

Rats were sacrificed 4 h after the last infusion of 6-OHDA. Their brains were removed and fixed for 40 min at 48 °C in 4%

Table 1 – Comparison table between all imaging mass spectrometry software already known

Software	Type	Different choices for calculations	Input format	Output format	Spectrometers compatibility	Use of databank	Requirements	Maximum data points	Processing Time	Availability	Free	Authors	Code
MMSIT	Acquisition	No	Instrument format .dat	Analyze 7.5	AB MALDI/TOF	Yes	Aquiris DP211 digitizer board	Unknown	Few seconds	www.maldi-si.org	Yes	M. Stoeckli	VB6 and assembler C#
4700/4800 Imaging	Acquisition	No	Instrument format	Analyze 7.5	AB MALDI-TOF/TOF	Unknown	Instruments software's	32 768 per spectrum	-	www.maldi-msi.org	Yes	M. Stoeckli	
Create Target AND Analyze This!	Acquisition/data conversion	No	Binary	Analyze 7.5	BD MALDI/TOF and MALDI-TOF/TOF	No	Instruments software's and Biomap	Theoretically 1 million per picture	25 mn for .img conversion	www.maldi-msi.org	Yes	C. Ceuppens et al.	Visual Basic 6
Axima2	Acquisition/data conversion	No	Instrument format	Analyze 7.5	SZU	Unknown	Windows	Unknown	-	www.maldi-msi.org	Yes	M. Setou et al	Unknown
oMALDI Server 5.1	Acquisition/Image reconstruction	No	Unknown	Analyze 7.5	AB/MDS SCIEX oMALDI QSTAR Hybrid LC/MS/MS	Unknown	Instruments software's	Unknown	Unknown	AB	No	AB/MDS SCIEX	C++
Flex Imaging	Acquisition/Image reconstruction	No	Instrument format .img files	bmp, jpeg, jpg, tif, tiff	BD MALDI/TOF and MALDI-TOF/TOF	Unknown	Instruments software's	Unknown	Few seconds	BD	No	BD	Unknown
Tissue view 1.0	Acquisition/Image reconstruction	No	Instrument format	tiff, jpeg, Analyze 7.5	AB/MDS SCIEX MALDI-TOF/TOF and QSTAR	Yes	Instruments software's	Unknown	Few seconds	AB	No	AB/MDS SCIEX	IDL
Image Quest	Acquisition/data conversion	No	Instrument format	Unknown	TS MALDI LTQ XL, LTQ Orbitrap hybrid series	Unknown	Instruments software's	Unknown	Unknown	TS	No	TS	Unknown
BIOMAP	Images reconstruction	No	Analyze 7.5, .img, .tiff, .dicom, .pnp	tiff, jpeg, Analyze 7.5	AB MALDI-TOF and 4700 and 4800 MALDI-TOF/TOF, MALDI Micro MX(tm) (Waters)	Yes	IDL RT 6.1 or higher	32 768 per spectrum for older versions	Few seconds	www.maldi-msi.org	Yes	M. Rausch with additions by M. Stoeckli	IDL
MITICS	Acquisition/Image reconstruction	Maximum intensity or Area	.xml	bmp, jpeg, jp2, gif, pict, png, tif	AB MALDI/TOF BD MALDI/TOF TOF/TOF	Yes, MySQL	Instruments software's, IDL VN and Data miner	Theoretically 1 million per picture	45 min/16000 points	Upon email	Yes	O. Jardin-Mathé et al.	C++, IDL, Visual Basic 6

Table has been created from data collected via different sources and may not be exhaustive or may contain uncompleted data and could be submitted to changes.
 AB: Applied Biosystems; BD: Bruker Daltonics; SZU: Shimadzu; TS: Thermo Scientific.

paraformaldehyde. The accuracy of the 6-OHDA lesion was checked by immunocytochemistry with an anti-tyrosine hydroxylase antibody according to a previously published procedure [11]. These brain tissue blocks were then stored in a box at room temperature 12 years before being used in this experimental plan procedure.

2.4. Tissue dewaxing

Rat tissue sections of 10 μm thickness were obtained using a microtome (E. Leitz, Westlart, Germany) and mounted on ITO (Indium Teen Oxided) one side coated conductive glass slides. Paraffin was removed by 2 baths of 5 min in toluene and lightly rehydrated with graded ethanol (100°, 96°, 70° and 30°) baths before drying at room temperature.

2.5. In situ trypsin digestion for MALDI imaging of rat brain tissue sections

Trypsin deposition (20 $\mu\text{g}/\text{mL}$ in water) was performed using an automatic vibration vaporization system (ImagePrep, Bruker Daltonics, Bremen, Germany). The quantity of trypsin was set, after optimization, to 1 mL over 30 cycles in order to cover the whole surface of the tissue. The system settings were: Spray power 30%, Spray modulation 40%, time spray on 2 s, incubation time: 30 s and dry time 30 s. HCCA/Aniline solid ionic matrix was used for these experiments [12]. The matrix preparation was described previously. Briefly, just prior to use, 4.8 μL of Aniline (1 eq.) was added to the matrix solution (HCCA 10 mg/mL in ACN/TFA 0.1%, 6:4, v/v) and sonicated for 15 min. HCCA/Aniline was then deposited on the tissue section using the vibration vaporization system according to the same parameters as described above.

2.6. Mass spectrometry

2.6.1. MALDI MS direct analysis of tissue sections after in situ automatic trypsin digestion

Trypsin fingerprint experiments on FFPE tissue sections after in situ digestion of the whole tissue section using the vibration sprayer system were performed on an Ultraflex II TOF–TOF instrument (Bruker Daltonics, Bremen, Germany) equipped with LIFT III cell and a Smartbeam laser with a repetition rate up to 200 Hz. For MS/MS experiments, parameters were set as follow: laser repetition rate was 100 Hz with 33% attenuation, ion source voltages were respectively 8 kV and 7.3 kV on a MALDI sample plate and 1st electrode; LIFT cell was pulse from ground for electrode 1 and 2 to 19 kV and in the last step electrode 3 was decrease to 3.2 kV; reflector end voltage was set to 29.5 kV and mid-grid to 13.85 kV. Trypsin fingerprint protein identification in databanks were performed using the Biotools 3.0 software (Bruker Daltonics, Bremmen, Germany) connected to Mascot search engine and interrogating the NCBI, Swissprot and ESTs databanks with Oxydation as Variable modification, Monoisotopic as mass values, a peptide mass tolerance of ± 0.1 Da with a peptide charge state of 1, a number of maximum missed cleavages of 2 and number of queries of 300.

2.7. MALDI MSI

MALDI-MSI of rat brain tissue sections was performed on an Ultraflex II TOF–TOF (Bruker Daltonics, Bremen, DE). After dewaxing, images were obtained in positive reflector mode after trypsin digestion and matrix deposition. 1 mL of the mix was applied onto the tissue using the ImagePrep sprayer as explained above and dried at room temperature. Acquisition was performed using a Smartbeam laser; with a repetition rate of 100 Hz and a spatial resolution of $100 \times 100 \mu\text{m}$. Images were reconstructed using either MITICS or FlexImaging v2.0 software (Bruker Daltonics, Bremen, Germany). Average mass spectra were recorded in positive mode for each of 12919 spots, covering the whole tissue section with an average of 500 laser shots per position.

MALDI-MSI was also performed on several dissected leech embryos (see above) at the E12 stage of development, mounted either with the inside of the embryo or the epidermis facing the laser beam. In one case the endodermal layer was removed during dissection. For reference, histological images were recorded prior to covering the tissues with a layer of ionic matrix.

3. Results and discussion

3.1. Software development

The new MITICS MS imaging software is divided in two parts, namely “MITICS Control” and “MITICS Image” (Fig. 1). MITICS Control is dedicated to create an interface for data acquisition on the instrument and thus requires that the instruments acquisition software be run in parallel, while MITICS Imaging is dedicated to post-acquisition data processing and image reconstruction. MITICS Control has been programmed in C++ for faster execution, whereas MITICS Image has been coded under IDL (Interactive Data Language) v 6.0 (ITT VIS, Paris, France), since IDL offers a powerful computing environment for data analysis and visualization. MITICS needs the IDL Virtual machine and the IDL Dataminer module to run. A temporary licence to test it can be obtained by emailing RSI support.

3.2. Data acquisition using MITICS: MITICS control

Almost all MALDI instruments now provide an automatic acquisition mode for high throughput analysis of samples in standard proteomics workflows. In the automatic acquisition mode, the software uses a table that defines the positions of the spots to be analyzed. MITICS provides a simple and easy interface with the automatic acquisition software of the instrument for defining the image raster and the acquisition parameters of the instrument. For MALDI-TOF Voyager series Applied Biosystems instruments, MITICS provides to the acquisition software (Voyager Instrument Control Panel) all parameters that are required in the MALDI-IMS sequence. For this purpose, the interface creates three files: .bic, .plt and .seq that are given to the instrument for data acquisition. The .bic file of MITICS Control substitutes itself for the acquisition parameters normally administered by the Voyager Instrument

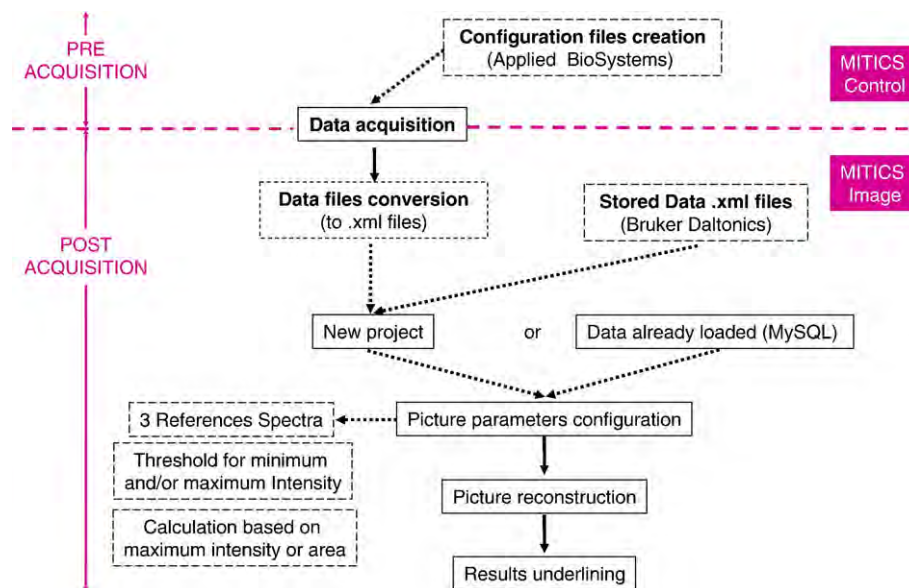


Fig. 1 – Scheme of the MITICS software global workflow including MITICS Control and MITICS Image parts.

acquisition software. The analysis parameters *i.e.* laser intensity, number of shots to be acquired per position, acceleration voltage on target and extraction grids, guide wire voltage, extraction delay time, mass scale, and low mass threshold are define and registered in the .bic file at the tab “BIC file” of the graphic interface (GUI). By contrast, the .plt and .sql files replace the instrument software for adjustment automatic mode values, *i.e.* the definition of the spots that will be sampled by the laser. Based on user-provided information, the interface calculates a table that is stored in the .plt files and given to the instrument during the automatic acquisition. For raster creation, the user defines Start and End coordinates and either a step size between spots or the number of spots per line or per column. The software then calculates automatically either the number of points or the step size, respectively. The X and Y dimensions have been decoupled, so that interval size can be set to a different value for each dimension. The third file (.sql) contains all the file information used by MITICS Control for data acquisition by providing these information's to the automatic acquisition software, namely the name and the location of the .bic and .plt files, and the recorded data files. MITICS Control also allows the application of post-acquisition macros to the recorded data. For example, the user can define an automatic baseline correction to be performed on each spectrum, or apply a smoothing or noise removal function to the data. It is also possible to enter the name of the suitable function in the text area provided for defining such manipulations prior to acquiring the spectra. A similar interface is currently under development for Bruker Daltonics MALDI instruments.

3.3. Data processing and image reconstruction: MITICS Image

Both Biomap and FlexImaging softwares provide sets of analytical tools for image reconstruction and data or image processing, including display an image of an ion of interest,

review of spectra behind each pixel of an image, definition of regions of interest (ROI) for comparison and statistical analysis, or co-registration of images with tissue section pictures. However, the first step for an external reconstruction tool is to import and convert the data collected during the IMS experiments. Due to the differences between the formats of the recorded data depending on the instrument, different GUIs are required. For MITICS, in order to speed up the calculation and optimize the browsing of the spectra, we have chosen to import the data files from the instruments, convert them (if necessary) to XML file format, use XML parsing to collect data in the files and store them in a databank that is powered by MySQL. Spectra in XML format files are easily obtained thanks to existing command lines that run in FlexAnalysis (SaxBasic coding) for Bruker Daltonics and to a new macro that has been written in DataExplorer (Visual Basic) for Applied Biosystems devices. MySQL has been chosen for the database mainly because of its accuracy but also for its speed and ease of use.

The initial loading of IMS data requires XML files to be written and data to be transferred to the database and stored. At the second loading of the data, the data already stored in the database are extracted, which increase the speed of processing. This allows the user to store data in order to rebuild pictures later, facilitates search and interrogation of the spectra and enables the delocalization of the data bank to another computer on the network. Actually, it is not necessary that the database reside on the computer where the software is running. Moreover, since it is possible to link several databases to MITICS, as long as the ODBC sources are correctly configured, the data can be recorded and stored on several computers. This is a great advantage, knowing a picture composed of about 7000 points with a mass range from 600 to 24000 Da (a whole acquired data folder of 4,33 Go) requires a database size of ~1.2 GB. Conversely, if MITICS is running on several computers on a network, several pictures can be reconstructed at the same time from a single data set of the

database, allowing for high throughput data processing and image reconstruction. Because MALDI-IMS data contains up to hundreds of different peaks, i.e., potential images, this can be a great advantage.

3.3.1. XML files

MITICS is designed to operate on .xml files, but in some instruments the XML files may not be automatically created during spectra acquisition. To solve this potential difficulty, a macro can be added to MITICS control. This function will be executed for every spectrum during the data acquisition, so as to get all the files available in XML as soon as the acquisition processes completed. If a picture has to be created from spectra previously acquired, XML files can still be produced by running a batch process in *FlexAnalysis* for an analysis realised on a Bruker Daltonics device, or by executing a visual basic macro in *DataExplorer* for an analysis performed on an Applied Biosystems instrument. In MITICS macro is available that allows the selection of an appropriate macro and its application to all the files of a folder in order to transform .dat file to .xml files. Fig. 2 presents a typical XML file as obtained from Bruker Daltonics instruments after data exportation. Each collected XML file corresponds to a collected mass spectrum of the image data. The file name includes the (X,Y) coordinates at which this

average spectrum was collected. The data itself is composed of blocks that give all relevant information for image reconstruction on a precise peak defined by its m/z value. An image sequence is then represented by a collection of XML files. Originally XML files were found to be the easiest way to easily obtain compatible data from various instruments. Although, recently on the initiative of HUPO a new international standard format for imaging data “i-mzML” as been proposed (<http://www.i-mzml.org/>) as an extension of the “mzML” format also proposed as international standard for MS data in general. This means that spectra in this format should be available for all instruments soon. “i-mzML” format are composed of two files one is an XML file that contains informations relative to the experiment. The second file is a binary file containing of the recorded mass spectra. We propose in the future to replace our XML conversion and develop the software to use these standard file formats since this will give more simplicity and unity to the software while allowing to easily getting through data of all instruments.

3.3.2. Processing time

Tests were performed on a computer running Windows XP with an Intel Pentium 4 (3.2 GHz) processor and 1.00Go of RAM. In these conditions, it took 45 min to first load and store in the

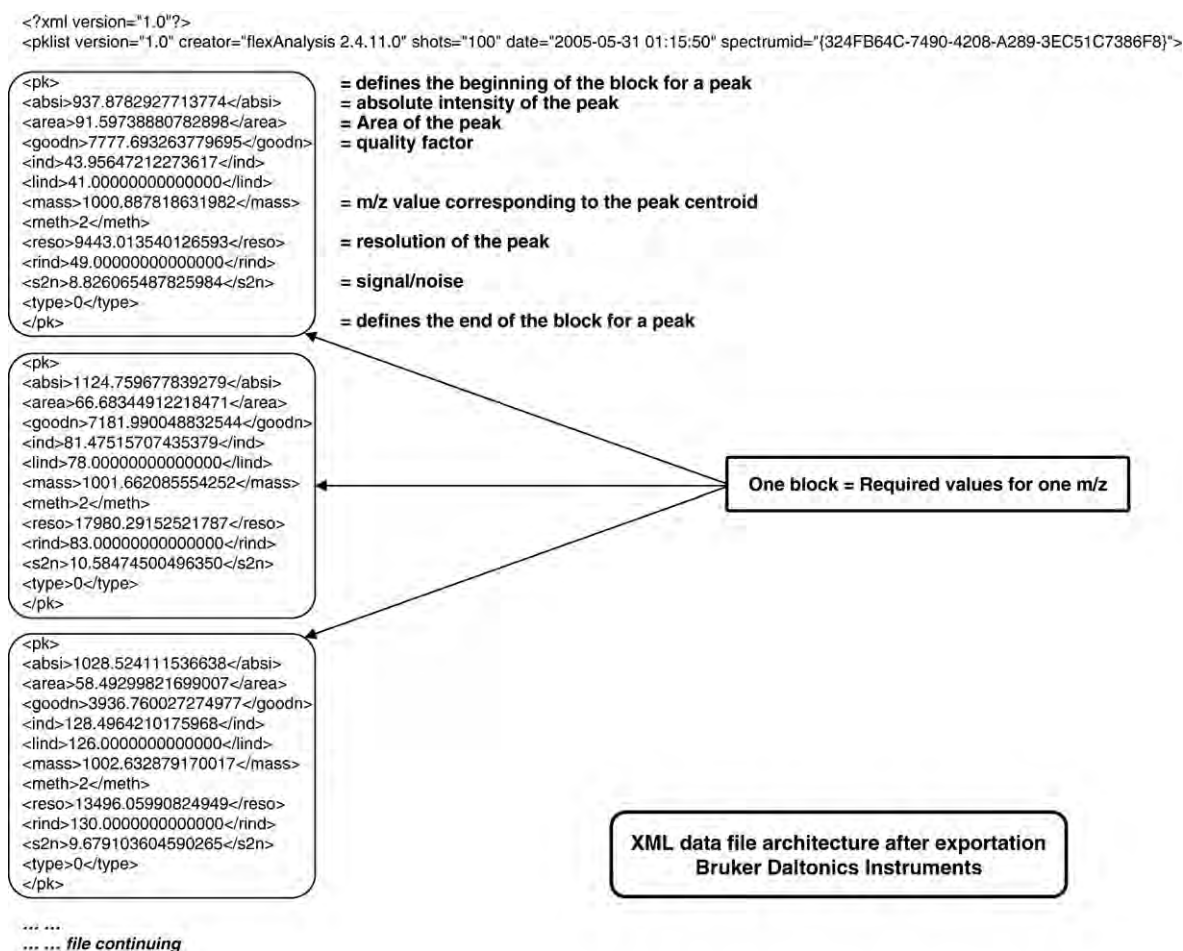


Fig. 2 – Example of the architecture of an XML file as obtained after exportation in batched process from a Bruker Daltonics instrument and prior to its loading into the Database.

database a 16000 point IMS data set. Then, reloading of the data directly from the databank took less than 1 min. Image reconstruction from firstly opened or reloaded data took also less than 1 min.

3.3.3. Data loading and databank storage

When data are first loaded on MITICS, the user has as first choice to define the MS system used for acquisition, e.g., an Applied Voyager DE-STR MALDI-TOF button or a Bruker

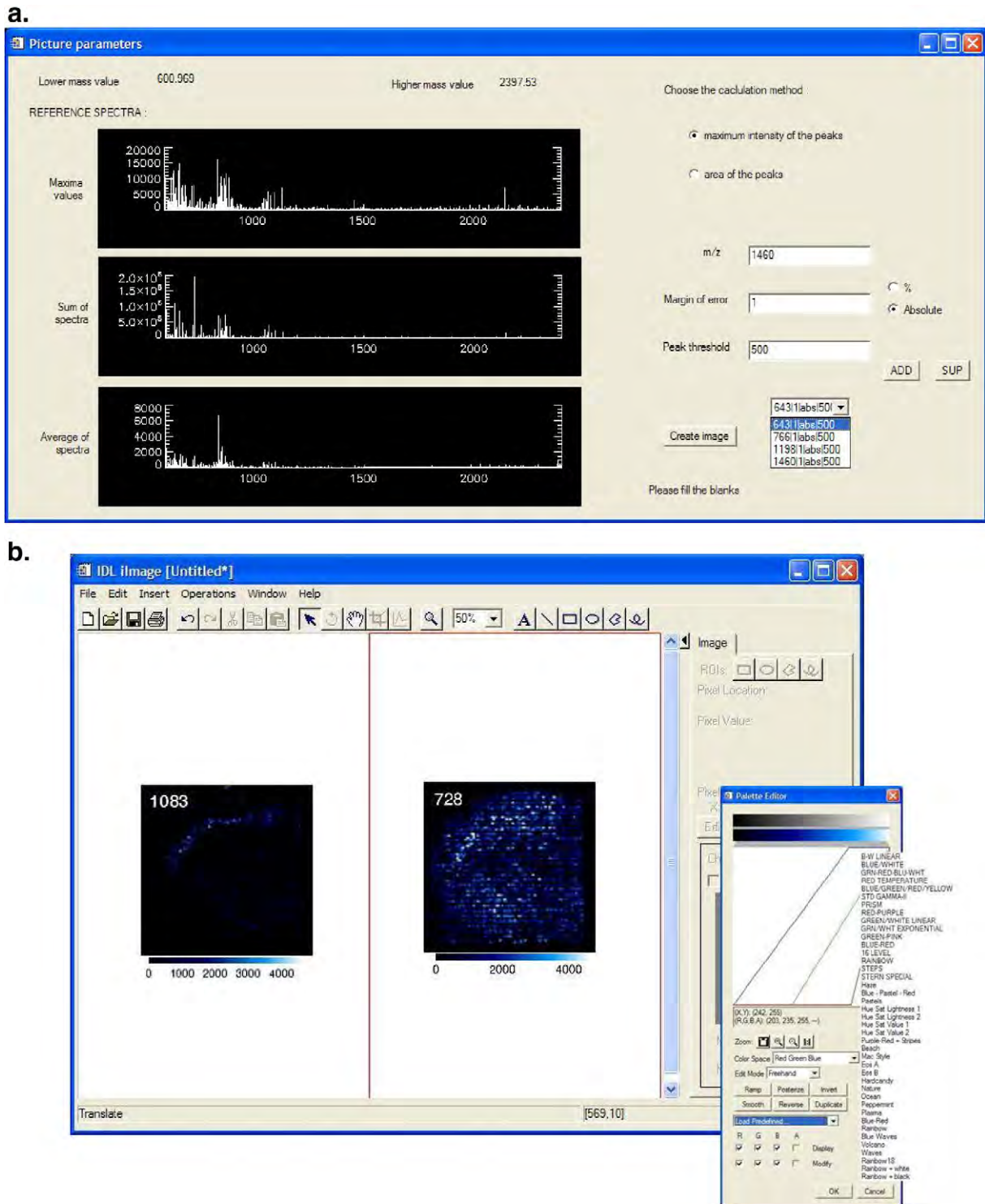


Fig. 3 – Snapshots presenting the interface proposed by MITICS Image for data processing and Image reconstruction. a) Presents the interface for image parameters setting including the three different representations of the image data as well as the choice for using maximum peak intensity or peak area for image reconstruction. b. Presents the graphic interface after image reconstruction and the palette editor with the predefined color scales.

UltraFlex MALDI-TOF-TOF button. For data originating from MITICS control the .plt file (containing the coordinates of the picture's pixels) has to be provided. For all data, the location of the IMS data must be given. While processing the data, a continuously updated message appears at the bottom of the window to inform the user about the progress in the analysis. At the end of processing, a new window appears asking the user to define the parameters for image reconstruction. The XML files are loaded into the database the first time. The database itself is composed of three tables that respectively include the name of the project, the acquisition position (image rastering) and the spectral information's (m/z , intensity, area,...). These three tables are linked together and associate to a project a specific set of positions and for each of these positions spectral data are linked. After loading in the database the data are immediately available for image reconstruction. They are kept there for faster reloading. This allows also as mentioned to perform image reconstruction from the network on different computers and process several

imaging sequences at the same time from delocalized computers.

One part of the software is dedicated to the configuration of the databank used. The software can check automatically on the computer where the available ODBC sources are located in the Administration Tools of the Control Panel. In this manner, when a new project is created, collected spectral data can be loaded into the right database.

3.3.4. Image reconstruction

Fig. 3 presents the user interface of MITICS for image reconstruction. Image reconstruction is the most difficult part to be settled to find what the best representation of the data is without inducing false representation. For image reconstruction we have decided to implement new representations of the MS data collection and 2 different types of calculations for image reconstruction. For image reconstruction, the user must provide an m/z value within a margin error. Threshold values (minimum and maximum) can also be defined for taking into

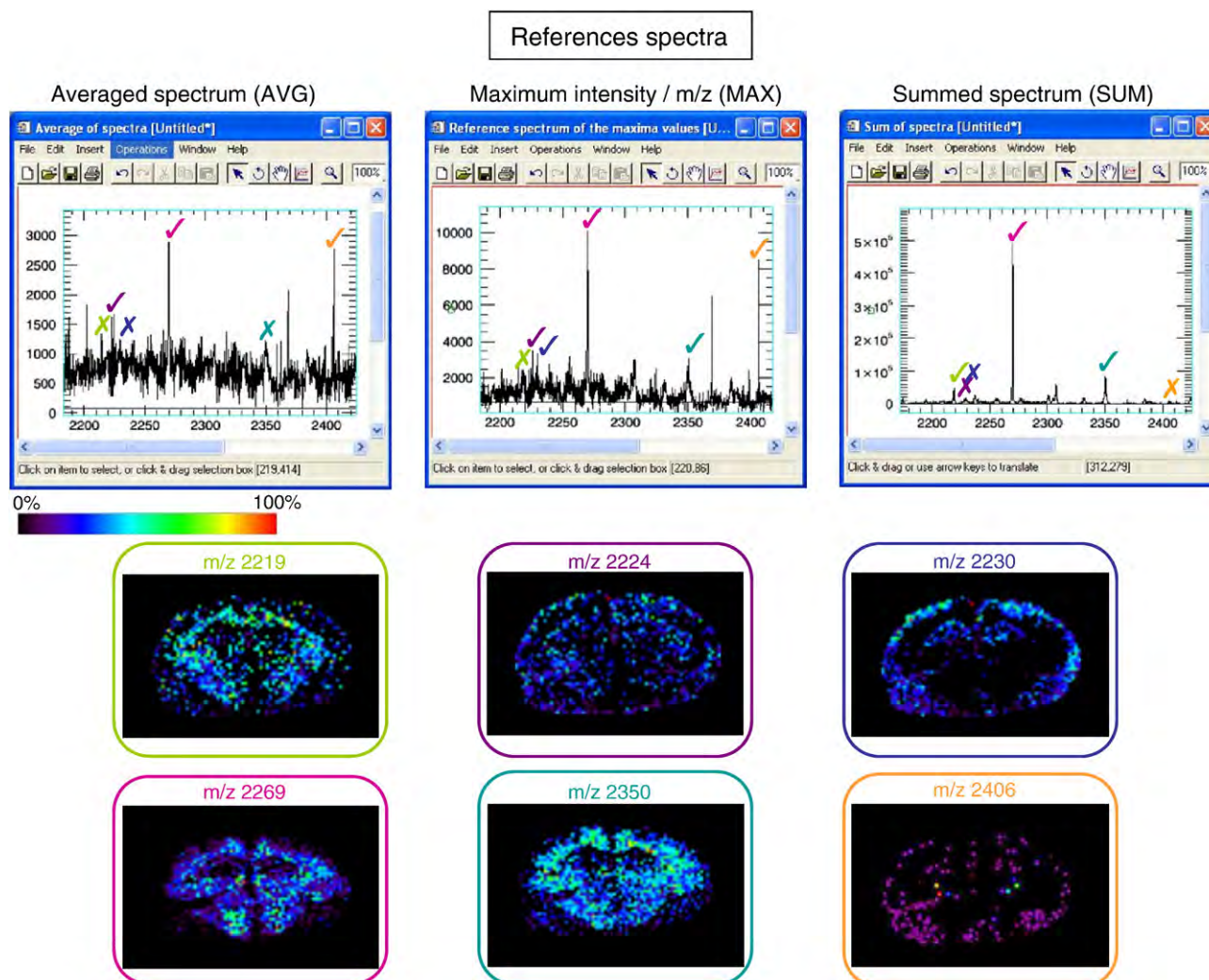


Fig. 4– Presentation of the three representations of image data proposed by MITICS and their interest for image reconstruction. Each spectrum correspond to one of the different representation (average, sum or maximum intensity) and images derived from the peaks observed in each of the representation according to the color code/ It is clear that all peaks are not displayed by all representations.

account signal variability (according to the position on the tissue). Minimum intensity threshold allows for defining the minimum value for intensity or area to consider that the signal is a peak. If the value is under the threshold then the value is set to 0 (no signal). The maximum value algorithm considers signal variability from one spot to another and decides that above a certain threshold values will be set to 100% intensity. After parameter adjustments, clicking the “add” button includes the selected ion in a list of ions that have been analyzed. Regarding the error margin, it can be defined either in percent of the total mass range or with an absolute value to add and subtract from the m/z value to obtain the boundaries of the accepted interval.

For data representation we have decided to had to the classical reconstructed averaged mass spectrum of all data (AVG), a mass spectrum corresponding to the sum of all mass spectra from the data collection (SUM) and a representative mass spectrum obtained by plotting for each m/z the maximum intensity as found in the data collection (MAX). This was motivated by the fact that significant biological information can be lost in the classical averaged spectrum representation of the data collection. This is for example the case for molecules that are either intense but much localized and for which the signal will be diluted by averaging or for molecules which are present in most of the tissue but at a quite low level of intensity. This well illustrated on Fig. 4 on the image reconstruction using MITICS for data obtained on a rat brain tissue section. The top part of the figure displayed the three different representations (AVG, SUM and MAX) of the same data set and the bottom part of the figure the

reconstructed molecular images using MITICS. It is clear from the different spectrum representations that peaks observables are different in each of the representation. Logically, the AVG representation gives the lowest signal intensity and the SUM representation reaches the highest level which as the convenience of increasing signal/noise ratio. To summarize, it can be shown that a peak that is quite intense with a large representation on the tissue section will be found in the three representations (e.g. m/z 2269, pink). On the contrary a peak with a large representation on the tissue section but with a weak intensity is better depicted by the SUM or MAX spectra (e.g. m/z 2350, pale blue). Peaks with weak intensity and more localized representation on the tissue section are only found in the MAX spectrum. This clearly illustrates the interest of adding different representation of the image data set for faster access to relevant information while avoiding browsing the entire m/z range to find hidden data in the averaged representation. All of the displayed representation allow for zooming in the corresponding window for ease of visualization.

We also choose to propose two methods of calculations either measuring the peak maximum intensity or its area. Because on MALDI-TOF instruments the resolution is not constant over the whole mass range (TOF peaks become broader with increasing m/z because of variations in initial positions and velocities), the maximum peak intensity does not always accurately describe the data. For this reason, measurement of the peak area instead of the maximum intensity can be displayed in MITICS Imaging. The user can then choose the calculation method. Peak intensity is better

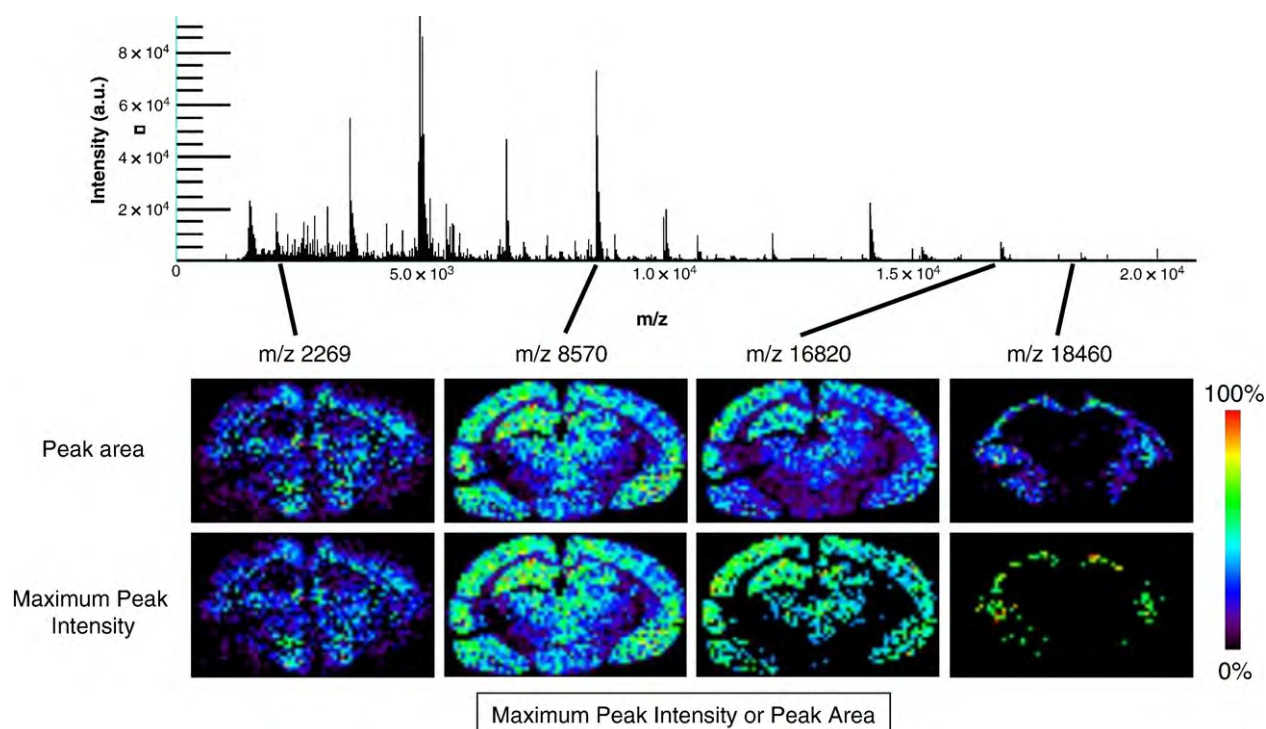


Fig. 5—Molecular Image reconstructed with MITICS of 4 peptides/proteins at respectively m/z 2269, 8570, 16 820 and 18 460 compared for the two calculations methods (maximum peak intensity or peak area). Data collected after an imaging sequence in positive mode on a rat brain tissue section using HCCA/Aniline as matrix.

Table 2 – Specific 6OHDA biomarkers and their amino acid sequence found after trypsin digestion by comparing 6-OHDA FFPE tissues to control FFPE tissues of rat brains after nanoLC-nanoESI-IT MS and MS/MS sequences; and their in silico trypsin digestion (Biotools 3.1, Bruker Daltonics, Bremen), including sequences and the [M+H]⁺ values of the digested peptides found after MALDI direct analysis of in situ digested tissues (in sequence box, matched peptides are indicated by a bar and are underlined in red)

Name mass (Da)	Sequence coverage	Peptides match	Digestion peptide	
			[M+H] ⁺ m/z	Sequences
Ubiquitin M= 8 559.62 Da	43.4%	10 20 30 40 50 60 70 80 MQIFVKLTG KITLEVEPSS DTIENVKAKI <u>QDKEGIPPDQ</u> <u>QLLIFAGKQL</u> EDGRTLS ^{DN} YV <u>IQKSTLHLV</u> <u>LRLRGG</u>	1067.62	ESTLHLVLR
		1081.56	TLS ^{DN} YNIQK	
		1336.67	ESTLHLVLR ^{LR}	
		1523.84	IQDKEGIPPDQQR	
α-Enolase M=47 098.24 Da	16.8%	50 60 70 80 90 100 110 <u>REILKIHARE</u> <u>IFDQKQVY</u> <u>EVKLTFTAGL</u> FRAAVFSGAS TGITALELR INDKTIFRQK GVS ^{KA} VENIN <u>WETAPALVK</u> KLVAVGQSKI <u>DQLRIEMDC</u> <u>IQKQKFGANA</u>	591.08	MSILK
		766.46	EIFDSR	
		786.47	IAKAAGEK	
		932.53	SFRNPLAK	
		976.49	TGAPCRSER	
		1143.64	IGAEVYHNLK	
		1636.85	IDQLMIEMDGTENK	
		2153.16	EIFDSRGNPTVEVDLYTAK	
		591.08	ITSDR	
		766.46	IVAPPGGR	
CRMP2 M=62 238.59 Da	15.0%	10 20 30 40 50 60 70 80 90 100 110 <u>RYQKQENIP</u> <u>PITLDELLIK</u> GQKIVNDQGS FTADITREDS LIKQIGENLI VPGQVETEA HSRVIFQGI DVHTRFQMPD QQRTADFFQ QGTRALAGS TTMIDRQVP	591.08	ITSDR
		766.46	IVAPPGGR	
		841.52	MSYQGGK	
		844.48	SRLAELR	
		1143.64	SAAEVIAQARK	
		1460.79	VMSKSAAEVIAQAR	
		2169.17	NLHQSGFSLGSAQIDDNIPR	
		2406.33	GLYDGPVCEVSTVPTKTVTPASSAK	
		788.47	SVDEALR	
		1636.85	GVLRQITVNDLPVGR	
Peroxidoxin M=21 770.06 Da	11.1%	10 20 30 40 50 60 70 80 90 100 110 RASONARQK PAFDPTQAV VEGAFKELK SDYDQKTVL FFFPLNFTV CPTETIAPSD HAEQFRLQK EVLGVSDSQ FTHLAVINFP RREGGLQENL IPLLAVTPE	788.47	SVDEALR
		1636.85	GVLRQITVNDLPVGR	
Profilin M=14 977.52 Da	40.7%	10 20 30 40 50 60 70 80 90 100 110 <u>RAWNAIYIS</u> <u>LNAICTQGA</u> <u>AINVQDEDS</u> <u>WMAAVVQTF</u> VSIPTAEVQV LQKDESSFF VNSGLTQQR CSVPSLQD DGEFTDLEK ESTGQAPTF <u>WVITRTAKTL</u>	1052.62	ESVHGGLINK
		3959.72	MAGWNAIYIDSLMADGTQDAAIVGYKDSPS VVAAVVPGK	
Ferritin Heavy Chain M=20 982.27 Da	26.0%	10 20 30 40 50 60 70 80 90 100 110 TTASPCVQK HYDQKQKAL NQKINLQVLS STVYLSRQY FRRDQVLLRN FAKYFLQQR EEPENALRN ELQVQVQRI FLDQKQKPR DVEVQSLAR KCLLELEVD	1481.8	SIKELGDHVTNLR
		4105.41	QNYHQDSEAINRQINLEYASYVYLSMSCY FDR	
Apolipoprotein E M=35 731.33 Da	25.3%	10 20 30 40 50 60 70 80 90 100 110 KRALWALLV PTLTQALQK ELEVTDLQK QSQQVQKAL HRFHDTLQV QTLSPQVQK LQSQVQKEL VLNREYTSK VKATKKELE QLQVAKETR <u>ARLAREVQAL</u>	726.23	LSTHLR
		888.44	GVS ^{AI} REK	
		968.53	LGPLVEQGR	
		1162.14	MEEQ ^T QQR	
		1183.80	LAKEVQAAQAR	
		1523.84	TANLGAQAQPLRGWFEPLVDMQR GWFEPLVDMQR QRTANLGAQAQPLR	
HSP27 M=22 807.57 Da	20.5%	10 20 30 40 50 60 70 80 90 100 110 <u>MTERRVPSL</u> <u>LESFSEPPFR</u> DVPYANSLF DQAFGPRFF DENWQFSSA GNPQVYRFLP AATAEGFAV TLR ^{AF} SRAL NQLSSGVSE <u>IRQTADRVV</u> <u>SLVNHFAPE</u>	591.080	QTADR
		932.532	QTADRWR	
		988.549	RVFSLLR	
		1358.716	AQGGPESEQSAK	
		1481.856	EGVVEITCKHEER	

for lower m/z ratio or for peaks very closed in m/z almost overlapping and peak area give better description for higher m/z signals if no overlapping of peaks is observed. Fig. 5 proposed the comparison for an imaging sequence recorded from a rat brain tissue section (same data collection as Fig. 4) of the two calculations methods. As expected the molecular images are almost identical by the two calculation methods for lower m/z ratio (e.g. m/z 2269 or m/z 8570). However, for higher m/z reconstructed images based on the peak area give better description of the compound localization and better fit the morphology of the tissue (ions at m/z 16820 and m/z 18460). Image reconstruction based on the peak area is not always an absolute requirement but can benefit to images for higher mass proteins.

Finally, the software can reconstruct several images at once if the m/z , error margins and threshold values have been settled. The software also gives the possibility of adding the signal of several ions and displayed the sum as one unique image. This was added in order to take into account the possibility to sum several ions corresponding to a same molecule i.e. for taking into account the presence of protonated or cationized ions of a same compound, different charge states for higher mass compounds or to sum all signals of different peptides generated by enzymatic digestion but belonging to the same original protein. Using this way we aimed in increasing the signal level (i.e. images with higher contrast) as well as averaging spot to spot variations due for example to a higher abundance of salts in some area leading to increase of cationized ion signals at the expense of protonated ions or to differences in enzyme activity depending the molecular environment at a precise location.

After image reconstruction, the software allows defining regions of interest, moving or turning the picture, resizing it, or

adding some geometrical shapes on it. It is also possible to save the image or print it, or to open a saved file or a new empty one. Some copy and/or paste or zoom functions are available as well. Another picture can be opened in the same window, making comparisons easier to do and some operations are available to work on the image, like filters for example, which permits getting a smoother picture compared to the initial simple hard pixel display. It is possible, moreover, to mark the results with a wide range of colors and gradations of basic tints or pre-specified multichromatic sets which are adaptable according to the wanted intensities. Lastly, the picture can be reprocessed after it has been displayed without the need to initiate another computation from the beginning — for example, maximum and minimum thresholds can be changed easily thanks to the Window Level Control in the Image Panel.

3.4. Application and validation of MITICS on biological cases

3.4.1. Applications to biomarkers using Parkinson's model animals

As previously demonstrated [13,14], brain FFPE tissue blocks of ten 6-OHDA treated versus ten untreated rats were systematically studied and compared for biomarker hunting using trypsin digestion followed by nanoLC-ion-trap identification. Several biomarkers have been identified by this strategy to be either up or down regulated between control and injected animals especially Ubiquitin, Alpha enolase, CRMP2, Peroxidoxin 6 and F1 ATPase were found to be up-regulated. Some of these proteins and others (Profilin, Ferritin Heavy Chain, HSP27, Apolipoprotein E, Vimentin) were also previously found by other proteomics approaches [15] or direct tissue analysis [16,17] to be up-regulated.

To access the efficiency of MITICS software on a real biological case we have used it to reconstruct molecular

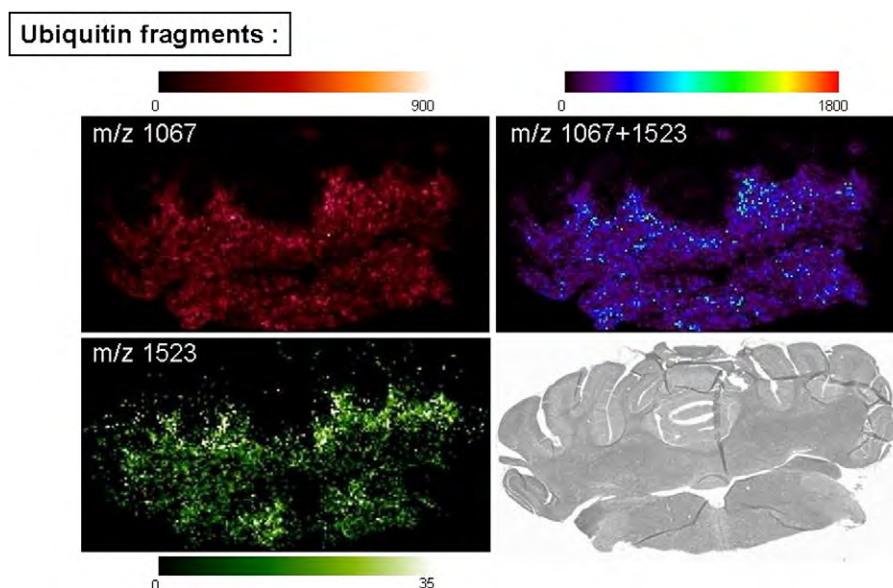


Fig. 6 – Molecular images reconstructed using MITICS software of m/z 1067 and 1523 ions and the composite image summing these two ions that correspond to two trypsin digestion fragments of Ubiquitin proteins as found by MALDI imaging in the reflectron positive mode on a FFPE rat 6-OHDA rat brain tissue section after tissue dewaxing and in situ trypsin digestion. Last panel presents photography of the tissue section prior to the experiment.

images from FFPE 6-OHDA rat brain tissue sections. We have performed, after dewaxing and rehydration, on tissue enzymatic digestion with trypsin of the tissue sections using a micro-spotting device and searched for fragments of the already known biomarkers by comparison with *in silico* digestions of these biomarkers (Table 2). Several digestions peptides of the markers of interest were found in the MALDI mass after enzymatic digestion and are given Table 2. These fragments were used to obtain molecular images of their repartition in the 6-OHDA injected rat brain sections. Fig. 6 presents the reconstructed images of two fragments (m/z 1067 and m/z 1523) of ubiquitin. The images show that the two fragments do present the same localization in the medulla oblongata whereas no signal is found in the cerebellum part of the section by comparison to the tissue section photography taken before the experiment where morphology is well apparent. The composite image of the sum of these two fragments is presented in the last panel. As previously mentioned possibility to sum different ions corresponding to a same molecule leads to a more contrasted image with higher signal intensity and fewer variations inter spots. Molecular image for m/z 1523 ion presents a more diffuse image than those of m/z 1067. This is not due to the software treatment but is probably due to the presence of a second smaller peak under the envelope of the peak corresponding to the ubiquitin digestion fragment that do not exactly present the same distribution as our fragment. Such effect is removed by using the composite image visualization, showing again the interest of using images based on several ion signals. Apolipoprotein E fragments were also found in the images data. The molecular images for two different digestion peptides of Apolipoprotein E respectively m/z 726 and 1523 are shown Fig. 7. Here, molecular images reconstructed either on the peak maximum

intensity or area, are compared. Again we observed that in some cases area value give better description of the data than the peak maximum intensity.

3.4.2. *Applications to peptides/proteins localization in the leech*
A second application was performed using MITICS for validation by studying peptides and proteins located in the CNS of the leech comparing leech embryos and adults. For adults, a metamer of the animal was removed and open on top of the dorsal part to obtain a strip of a metamer. An imaging sequence was performed on this leech segment to image the dorso-ventral orientation of molecules in the animal. Fig. 8 shows the difference for the m/z range 3200–3600 of the two representations AVG and MAX. In this m/z range only few and weak intensity peaks are observed in the AVG representation whereas more intense and numerous peaks are depicted in the MAX spectrum. By selecting on of these peaks (m/z 3507 \pm 0.5) it is possible to reconstruct the molecular image for this specific protein. It appears that it is highly localized in the nervous system (here a nervous ganglia). This validates again the interest of representation image data under several representations to extract interesting biological features. Whole mounted embryos were also studied using MALDI imaging. From the collected data set molecular images were reconstructed for two peptides at m/z 3507 and 6443 using MITICS but also using FlexImaging and Biomap software's (Fig. 9). For ion at m/z 3507 the three software's give very similar results especially for Biomap and MITICS that exactly provide the same color scale. Slight differences for FlexImaging are due to the color scale and the normalization of the data. For ion at m/z 6443 more differences are observed. The signal of this ion is quite weak and the ion localization is hardly observed in the nephridia of the animal. It is a bit clearer on the molecular

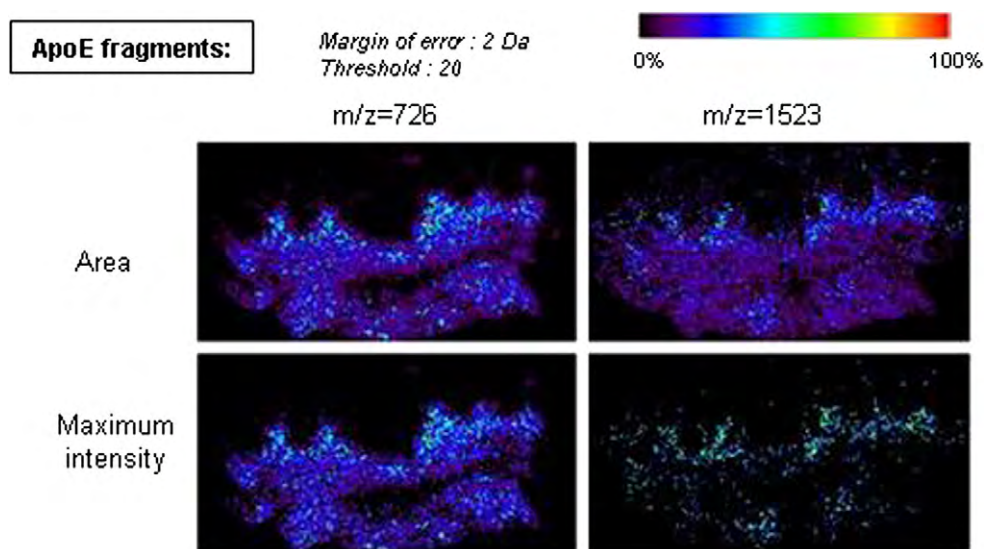


Fig. 7 – Comparison of the images reconstructed with MITICS based on Apolipoprotein E (Apo E) trypsin specific peptides m/z 726 and 1523 using either the peak maximum intensity or the peak with a threshold of 20 and a margin of error of 2 Da for comparison of the two calculation methods. Data have been obtained after an imaging sequence in the positive reflectron mode on a FFPE 6-OHDA rat brain tissue section after tissue dewaxing and *in situ* trypsin digestion.

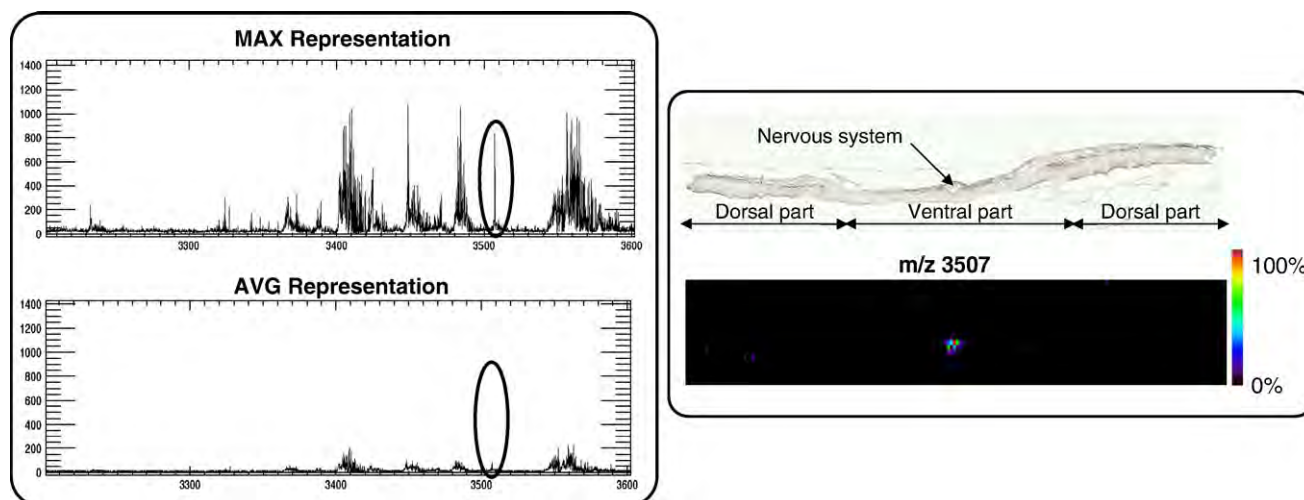


Fig. 8–Comparison of two representations proposed by MITICS (average and maximum of intensity per m/z) for a MALDI imaging sequence performed on a leech segment. Data show that images of the ion at m/z 16 620 are equivalents even if the peak of interest is only observable from the representation of maximum intensity per m/z showing the interest of displaying several representations. Data have been obtained by an imaging sequence in the linear positive mode after dissection of one leech segment and its opening on top of the dorsal part to obtain a dorso-ventral leech strip where the nervous ganglia of this peculiar segment is observable in the middle of the ventral part. HCCA/aniline was used as matrix for this imaging sequence.

image from Fleximaging due to normalization of the data. As observed here normalization of data find interest by enhancing signal weakest signal but as a tendency to increase images noise. However, this is something not yet implemented in MITICS that would be required in the next future to increase performances of the software.

4. Conclusion

MITICS is novel integrated software for IMS that is designed for compatibility with all MALDI instruments. MITICS provide for the moment completely developed solutions for processing data from Applied Biosystems and Bruker Daltonics MALDI-TOF instruments, but it is aimed to evolve to be

compatible with SCIEX or Shimadzu or even other instruments. MITICS includes two parts, namely MITICS Control for data acquisition on the instrument and MITICS Image for data processing and image reconstruction. Data are processed following XML conversion and data extraction using an XML parser before storage in a database. Addition of databanks to the software allows for simultaneous reconstruction of numerous images in the goal of high throughput imaging to be developed. MITICS provides different views of data than other available software, namely different spectral representations for ions of interest, and offers the flexibility to use peak area instead of maximum intensity as the measurement basis for image reconstruction, which has proved to be better in describing biological informations for higher m/z ratio. Finally, MITICS tends to be universal software, easily modified for

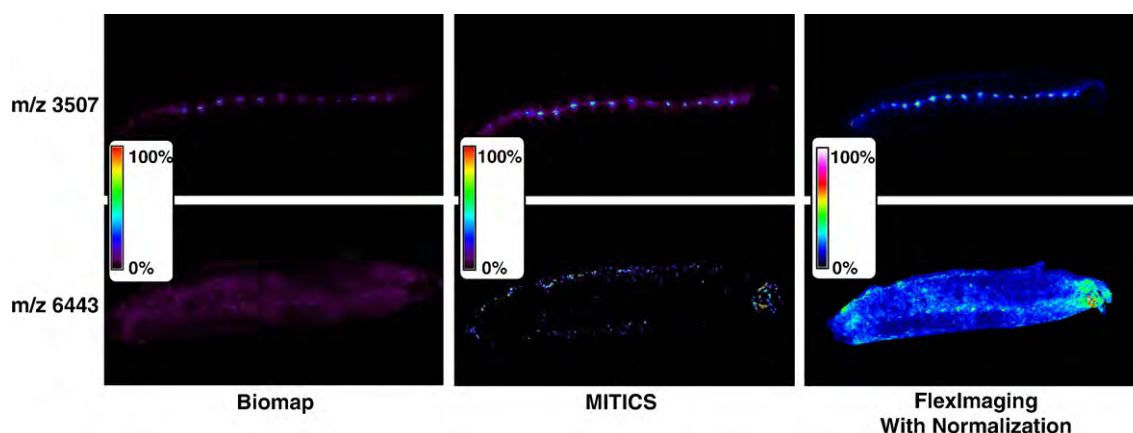


Fig. 9–MALDI molecular images reconstructed with Biomap, MITICS and FlexImaging for validation of MITICS software of two ions m/z 3507 and 6443 found after a MALDI imaging sequence of dissected stage E12 specimens leech embryos in whole mount. MALDI images were reconstructed from the mass spectra collected in the linear positive mode from 20000 locations covering the embryos completely in a rectangular raster of points $60 \mu\text{m}$ apart using HCCA/aniline as matrix.

integration of new tools and is available on request on our web site (<http://www.maldi-imaging.com>).

Acknowledgements

Supported by grants from Centre National de la Recherche Scientifique (CNRS), Ministère de L'Éducation Nationale, de L'Enseignement Supérieur et de la Recherche (ACI Jeunes Chercheurs ACI JC4074 to I. Fournier), ANR PCV (to I. Fournier). Also supported by collaboration Bruker Daltonics and the laboratory for the MALDI imaging technology development.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jprot.2008.07.004](https://doi.org/10.1016/j.jprot.2008.07.004).

REFERENCES

- [1] Fournier I, Day R, Salzet M. Direct analysis of neuropeptides by in situ MALDI-TOF mass spectrometry in the rat brain. *Neuro Endocrinol Lett* 2003;24(1–2):9–14.
- [2] Li L, Garden RW, Romanova EV, Sweedler JV. In situ sequencing of peptides from biological tissues and single cells using MALDI-PSD/CID analysis. *Anal Chem* 1999;71(24):5451–8.
- [3] Chaurand P, Stoeckli M, Caprioli RM. Direct profiling of proteins in biological tissue sections by MALDI mass spectrometry. *Anal Chem* 1999;71(23):5263–70.
- [4] Redeker V, Toullec JY, Vinh J, Rossier J, Soyez D. Combination of peptide profiling by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and immunodetection on single glands or cells. *Anal Chem* 1998;70(9):1805–11.
- [5] Jimenez CR, Li KW, Dreisewerd K, Spijker S, Kingston R, Bateman RH, et al. Direct mass spectrometric peptide profiling and sequencing of single neurons reveals differential peptide patterns in a small neuronal network. *Biochemistry* 1998;37(7):2070–6.
- [6] Dreisewerd KKR, Geraerts WPM, Li KW. Direct mass spectrometric peptide profiling and sequencing of nervous tissues to identify peptides involved in male copulatory behavior in *Lymnaea stagnalis*. *Int J Mass Spectrom Ion Process* 1997;169/170:291–9.
- [7] Li KW, Hoek RM, Smith F, Jimenez CR, van der Schors RC, van Veelen PA, et al. Direct peptide profiling by mass spectrometry of single identified neurons reveals complex neuropeptide-processing pattern. *J Biol Chem* 1994;269(48):30288–92.
- [8] Stoeckli M, Farmer TB, Caprioli RM. Automated mass spectrometry imaging with a matrix-assisted laser desorption ionization time-of-flight instrument. *J Am Soc Mass Spectrom* 1999;10(1):67–71.
- [9] Caprioli RM, Farmer TB, Gile J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* 1997;69(23):4751–60.
- [10] Clerens S, Ceuppens R, Arckens L. CreateTarget and Analyze This!: new software assisting imaging mass spectrometry on Bruker Reflex IV and Ultraflex II instruments. *Rapid Commun Mass Spectrom* 2006;20(20):3061–6.
- [11] Tonnaer JA, Lammers AJ, Wieringa JH, Steinbusch HW. Immunohistochemical evidence for degeneration of cholinergic neurons in the forebrain of the rat following injection of AF64A-picrylsulfonate into the dorsal hippocampus. *Brain Res* 1986;370(1):200–3.
- [12] Lemaire R, Tabet JC, Ducoroy P, Hendra JB, Salzet M, Fournier I. Solid ionic matrixes for direct tissue analysis and MALDI imaging. *Anal Chem* 2006;78(3):809–19.
- [13] Stauber J, Lemaire R, Franck J, Bonnel D, Croix D, Day R, et al. MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting. *J Proteome Res* 2008;7(3):969–78.
- [14] Lemaire R, Desmons A, Tabet JC, Day R, Salzet M, Fournier I. Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. *J Proteome Res* 2007;6(4):1295–305.
- [15] Basso M, Giraudo S, Corpillo D, Bergamasco B, Lopiano L, Fasano M. Proteome analysis of human substantia nigra in Parkinson's disease. *Proteomics* 2004;4(12):3943–52.
- [16] Pierson J, Norris JL, Aerni HR, Svenningsson P, Caprioli RM, Andren PE. Molecular profiling of experimental Parkinson's disease: direct analysis of peptides and proteins on brain tissue sections by MALDI mass spectrometry. *J Proteome Res* 2004;3(2):289–95.
- [17] Pierson J, Svenningsson P, Caprioli RM, Andren PE. Increased levels of ubiquitin in the 6-OHDA-lesioned striatum of rats. *J Proteome Res* 2005;4(2):223–6.