*Manuscript
Click here to view linked References

Protein identification in mass spectrometry imaging

Nadine E. Mascini, Ron M. A. Heeren*

FOM Institute AMOLF, Science Park 104, 1098 XG Amsterdam, The Netherlands

*Corresponding author.
Tel.: +31 20 7547100
Fax: +31 20 7547290
E-mail: R.Heeren@amolf.nl

Abstract

With a rapidly growing number of biomedical applications of mass spectrometry imaging (MSI) and expansion of the technique into the clinic, spectrum annotation is an increasingly pressing issue in today's MSI field. Although identification of the species of interest is the key to answering biomedical research questions, only few of the hundreds of observed biomolecular signals in each MSI spectrum can be easily identified or interpreted. So far no standardized protocols exist that resolve this issue.

Present strategies for protein identification in MSI, their limitations, as well as future developments will be the scope of this review. We will discuss advances in MSI technology, workflows and bioinformatic tools to improve the confidence and number of protein identifications within MSI studies.

Keywords: Bioinformatics, Biomedical research, *In situ* analysis, Mass spectrometry, Mass spectrometry imaging, Molecular imaging, Tissue analysis, Protein identification, Proteomics

Abbreviations: CID, collision induced dissociation; ETD, electron transfer dissociation; ESI, electrospray ionization; FTICR, Fourier transform ion cyclotron resonance; IHC, immunohistochemistry; IMS, ion mobility spectrometry; LC, liquid chromatography; LCM, laser capture microdissection; MALDI, matrix-assisted laser desorption ionization; MS, mass spectrometry; MSI, mass spectrometry imaging; PTM, posttranslational modification; MS/MS, tandem mass spectrometry; ROI, region of interest; TOF, time-of-flight

1. Introduction

In recent years, technological and methodological advances have brought mass spectrometry imaging (MSI) to the biomedical field. MSI allows for the analysis and visualization of peptides, proteins, lipids, metabolites and pharmaceuticals directly from biological tissues and cell samples [1, 2]. The technique uses a surface sampling process in which mass spectra are collected at discrete locations according to a predefined Cartesian grid. In this way, the distribution of ions of interest can be mapped.

MSI has several advantages compared to other imaging techniques such as immunohistochemistry (IHC) or positron emission tomography: it has the capability to simultaneously detect hundreds of (unknown) compounds in one molecular imaging experiment, allowing for multiplexed analysis and discovery-based research. As MSI requires no target-specific labeling, unmodified species can be studied. Importantly, in contrast to standard mass spectrometric analysis, which requires tissue homogenization, MSI leaves the molecular distributions in the tissue intact: it can be utilized to assess molecular differences between specific cellular regions within tissues.

An increasing number of studies report on applications of MSI in the biomedical field. MSI is used in distribution studies of pharmaceutical compounds and their metabolic products for drug evaluation [3, 4] and in (clinical) proteomics applications [5, 6]. MSI has already been employed to assist in diagnosis, prognosis and biomarker discovery: the technique is utilized to construct protein profiles that predict a disease status or progression, to identify molecular patterns for disease prognosis and to assess molecular markers in treatment response studies [7-9]. Not only can a better fundamental understanding of the molecular processes underlying disease be acquired using MSI, but this knowledge can also aid in the development of new drugs and treatments. The study of the molecular basis of intratumor heterogeneity, for

example, is not only expected to lead to improved understanding of tumor biology, but also fits in the trend towards personalized medicine [10, 11].

The most widely used ionization technique for MSI is matrix-assisted laser desorption ionization (MALDI) [12]. MALDI-MSI was introduced in 1997 by Caprioli and coworkers and utilizes a matrix, typically an acidic aromatic compound [13]. As the matrix compound absorbs energy at the wavelength of the laser, exposure of the crystals to laser pulses results in desorption and ionization of the sample. Ions are separated based on their mass-to-charge (*m/z*) ratio, usually by a time-of-flight (TOF) mass analyzer which is high-throughput, sensitive and has a broad mass range [14].

Despite the fast developments in MSI technology and workflows, several challenges still need to be addressed for MSI to become an established tool in the biomedical research environment. Apart from the need for improved mass resolution, spatial resolution and sensitivity of the instruments used for MSI, an important limitation is that only a few of the hundreds of observed signals in each mass spectrum can be easily identified or interpreted. Annotation of ions of interest requires an additional step in the experimental workflow and so far no standardized protocols exist that solve this issue. Identification might be hampered by (unknown) modifications, even when the compound class is known. This holds for example for ions derived from proteins, where posttranslational modifications (PTMs), protein isoforms and chemical modifications resulting from sample preparation or proteolysis can hinder interpretation.

From the point of view of a biomedical researcher, however, identification of the species of interest is an essential step to solve a biomedical research question. Although recently studies have been published in which statistical data analysis tools were used to annotate tissues solely based on their mass spectrometric profiles [10, 11, 15], MSI data needs to be

complemented with information on the nature of the biomolecular species to access its full potential [16].

With a rapidly growing number of biomedical applications and expansion of MSI into the clinic, spectrum annotation is an increasingly pressing issue in today's MSI field. Present strategies to provide annotation of MSI spectra, their limitations, as well as newly developed identification strategies will be the scope of this review. As proteins are the biomolecules most often probed by MSI in a biomedical context this review will focus on protein identification in MSI. Confident chemical assignment of any biomolecular species in MSI spectra, however, faces similar challenges and some of the approaches described here could also be employed in that context.

2. MS-based protein identification methods

Mass spectrometry (MS) is an established analytical technique for protein characterization both at the species level and at the level of the proteome. Numerous, often very sophisticated, methods of MS-based protein identification have been developed [17, 18]. Current MSI has implemented MS-based protein identification methods in its workflows according to the needs and constraints posed by the technique.

In general, two approaches exist for MS-based protein identification: in a top-down experiment identification is performed through intact mass measurement followed by MS/MS analysis. Sequence-specific fragmentation patterns are used for identification through database searching, in which the experimentally obtained fragments are compared with theoretical fragments. A top-down approach in MALDI-MSI works best for small to medium sized proteins up to 7-10 kDa, because large singly charged molecules will not easily dissociate.

In a bottom-up experiment, a protein or protein mixture is first enzymatically digested. The resulting proteolytic peptides are analyzed by MS (so called peptide mass fingerprinting), and by tandem MS (MS/MS) in case of a protein mixture. In a bottom-up imaging approach, multiple peptide matches per protein are required for confident identification of the protein. An ideal MSI experiment consists of automatically triggered tandem MS experiments on proteins or peptides directly from tissue, thereby combining the localization of species with their identification within a single experiment. Low sensitivity, however, seriously hampers the identification as compared to standard identification approaches using protein extraction followed by gel-based separation or liquid chromatography (LC) coupled to electrospray ionization (ESI) MS/MS. The low sensitivity is caused by ion suppression effects due to its complex molecular composition. In addition, ions generated by MALDI typically have only unit charge. The resulting inefficient ion activation of larger ions renders intact proteins too big for direct identification through fragmentation. As a result efficient tandem MS can only be performed in a mass range of 500-3500 Da on the majority of mass spectrometers used for MSI. Figure 1 summarizes the protein identification workflows utilized in MSI.

2.1 Top-down approaches in MSI

Few examples of a top-down approach used in MALDI-MSI can be found in literature: Minerva and coworkers identified several endogenous peptides up to 3.5 kDa using MALDI-TOF/TOF directly on mouse pancreatic tissue [19]. Alternative identification strategies are employed to annotate larger masses, which combine fractionation of tissue extracts by LC, MALDI MS for fraction selection, followed by ESI-MS/MS [20-22]. In this way, identification of the 8.4 kDa cysteine-rich intestinal protein 1 in breast cancer tissue was demonstrated [20]. Top-down analysis of the 14+ charge state resulted in identification of the protein, which was found to be correlated with human epidermal growth factor receptor 2, an

important marker for treatment response prediction. These classical identification strategies are labor intensive, require the extraction of the protein of interest, and remain thus limited to only a few identifications per study.

An alternative approach to standard fragmentation techniques in MSI such as TOF/TOF and collision induced dissociation (CID) is in-source decay, where ions are fragmented in the source region before extraction [23]. This technique, however, suffers from the lack of precursor ion selection, which makes the mass spectra hard to interpret. In-source decay is therefore only rarely utilized for protein identification. A recent trend is the development of electron-based MS/MS techniques. Electron-induced dissociation of singly-charged peptides has already been demonstrated [24].

Despite the limited utility of a top-down approach due to technical and practical constraints, it should be kept in mind that by studying intact proteins, not only information on the complete amino acid sequence is retained, which allows for high confidence protein assignment, but also on protein state. Cazares and coworkers for example, identified specifically the fragment of the MEKK2 protein to discriminate tumor from normal tissue [25]. This type of information typically cannot be obtained using IHC or a bottom-up approach (section 2.2).

2.2 Bottom-up approaches in MSI

In a bottom-up approach proteins are digested on-tissue while their spatial distribution is preserved. Trypsin is the enzyme of choice for digestion and can be applied by automated spotting devices. These devices deposit picoliter droplets in an array with a spot size of 100-200 μm. After incubation, matrix can be deposited onto the tissue using the same device [26]. The resulting tryptic peptides are subjected to tandem MS directly on-tissue [27, 28]. This *in situ* digestion approach is often considered the method of preference for MSI studies, because

it facilitates on-tissue fragmentation, hence peptide identification within the imaging experiment itself. An additional advantage is that on-tissue digestion can be used to "unlock" proteins from the formalin-fixed paraffin-embedded tissues widely used in bio(medical) research, the proteome of which would otherwise have remained inaccessible for MSI analysis [29-32].

On the other hand, the on-tissue digestion approach also suffers from considerable background signal from the tissue. This results in identification of mainly highly abundant proteins and only a limited number of peptides per identified protein. The limited amino acid sequence coverage per protein may also result in loss of information on protein state, for example on type and location of PTMs.

Several improvements in sample preparation and instrumental set-up have found implementation in bottom-up MALDI-MSI workflows. The addition of the detergent n-octylglucoside to the trypsin buffer solution was found to increase the number of peptide signals as well as their signal intensities and led to enhanced detection of lipophilic proteins [28]. Second, on-tissue chemical peptide derivatization strategies were developed for enhanced identification. Franck and coworkers showed that on-tissue derivatization of tryptic peptides is compatible with an *in situ* digestion approach [33]. MS/MS spectra recorded on conventional MALDI-TOF instruments are often difficult to interpret due to the different types of ion series generated. This results in only small sequence tags available for identification. The addition of a N-terminal negative charge by derivatization with sulfonation agents generated (almost) complete y-ion series and even allowed for *de novo* sequencing (i.e. without the help of a protein database) of tryptic peptides.

A recent advancement is the combination of ion mobility spectrometry with MSI [26, 28, 31, 34, 35]. Ion mobility spectrometry (IMS) is a gas-phase chromatographic technique which separates ions based on their collision cross-section (i.e. size and shape). The extra dimension

provided by the post-ionization ion mobility separation allows for separate inspection of isobaric contributions to a spectrum. This is especially useful for complex spectra resulting from the analysis of *in situ* digested tissues, which show unresolved peaks from overlapping species, as for example isotopic distributions of peptide species, lipids and matrix ions. In this way, the complexity of tandem MS spectra is reduced. Figure 2 shows how ion mobility separation prior to fragmentation of two singly charged tryptic peptide ions that both have a molecular weight of *m/z* 1039 resulted in their identification (from tubulin and ubiquitin respectively) [34]. An MS/MS database search using the Mascot engine without ion mobility separation resulted in a score which was too low for confident identification of either of the two tryptic peptides and the proteins they originate from.

2.3 Indirect identification approaches

Indirect approaches for protein identification are often used to avoid sensitivity issues with on-tissue fragmentation. The aim of these strategies is to eliminate ion suppression effects, which introduce ionization bias in the MSI analysis, and to increase the dynamic range of the analysis.

In short, MSI data is matched with data generated using complementary methods which include a fractionation step (mostly LC-MS) [36]. This is not a trivial undertaking as ESI, the most commonly used ionization technique for LC-MS due to the simplicity of the interface, favors the ionization of different peptides as compared to MALDI. LC-MALDI is utilized less frequently [37], and LC coupled to secondary ion mass spectrometry still has to prove its utility [38]. In an indirect approach, independent experimental data serves as a tissue-specific reference database which can be searched to identify peptides in the MSI data (ure 1).

The capability of MSI to measure complex samples might be further enhanced by targeting specific cell populations from tissues by laser capture microdissection (LCM) [39]. LCM is an

especially useful enrichment technique for tissues showing a high degree of heterogeneity as for example breast cancer tissues [40].

Mass correlation between MSI data and data from independent experiments requires extensive (manual) data interpretation, often combined with prior knowledge of the species of interest. Wide mass tolerance windows for mass matching of up to ±2 Da are reported [41]. Masses with tolerance windows of this size theoretically match thousands of possible peptides. In these cases additional validation is an absolute necessity to prevent erroneous protein identification.

A MSI study of tumor margins in renal cell carcinoma reports on the use of an additional peptide characteristic to eliminate false positive protein identifications [42]. Tryptic peptides from tissue extract were isoelectrically focused using an immobilized pH gradient strip to provide additional information, the peptide's isoelectric point, to match the experimental with theoretical peptides.

A recent paper by Schober and coworkers describes an improved indirect strategy which combines MALDI-MSI with complementary off-line LC coupled to ESI-MS/MS [43, 44]. All results were based on accurate mass measurements recorded on Fourier transform MS instruments which allowed for improved quality and quantity of peptide identifications. Fourier transform ion cyclotron (FTICR) and Orbitrap mass spectrometers have mass accuracies in the low or sub ppm range instead of the at most 10-50 ppm mass accuracy obtained by using TOF systems. The high mass accuracy and mass resolving power of these instruments provide enhanced means to resolve the complexity of biological samples, but the use of these mass analyzers for protein identification in MALDI-MSI is still limited to only a few examples due the limited sensitivity of FTICR, limited mass range of Orbitrap, and the relatively long measurement time needed to obtain high accuracy [45, 46]. Moreover, an

indirect accurate mass approach only works if both the MALDI-MSI data and the LC-ESI data are recorded with high mass accuracy.

3. The importance of mass accuracy for protein identification

Mass accuracy can be defined as the degree of similarity between a measured value and its theoretical value. If multiple species are assessed usually the root mean square or mass measurement error is used. Precision is defined as the degree to which a measured value is similar during a (series of) experiment(s). In MS-based protein and proteomics research statistical tools have already been widely implemented to assess accuracy, precision hence also confidence of identification [47, 48].

MS-based identification can be extremely accurate and precise because it uses the intrinsic property of a species (i.e. its mass) which can be measured by mass spectrometry in an unbiased way. A highly accurate monoisotopic mass (sub ppm for a peptide of 1 kDa) provides information on mass defect, isotopic distribution and can even specify the elemental composition. Accurate mass measurements lead therefore to improved confidence in protein assignment.

In addition, an important parameter in peptide identification, directly related to mass accuracy, is the threshold value (or mass tolerance window) used in a database search. The set threshold is a tradeoff between maximum specificity and maximum sensitivity. At high thresholds true positives are potentially rejected, whereas at low thresholds the mass resolution of the recorded data might not be fully used.

Until now little emphasis has been placed on assessment of peptide annotation reliability in MSI studies. Instead, orthogonal validation methods such as IHC are employed. Recent high

mass resolution MSI studies show, however, that results can largely vary depending on the used mass bin width for ion selected images, which is directly dependent on the mass accuracy and mass resolution of the data, as exemplified in Figure 3 [49].

In a 2011 study, assessment of the mass accuracy of MALDI-MSI data is described [50]. I{Minerva, 2011 #480;= }n this paper, MSI data is linked to tandem MS data from independent experiments by employing an intermediate step using accurate mass data from FTMS measurements. The mass accuracy of the recorded MALDI-MSI data was found to decrease with increasing mass range and the applied mass tolerance window for mass correlation was adjusted accordingly. These examples demonstrate the trend to use accurate mass data to improve the number and confidence of peptide annotations.

4. Data analysis

Protein identification also heavily depends on the data processing and mining strategy chosen, the quality of the protein database and database searching algorithms used. As MSI data analysis utilizes the bioinformatic tools and databases developed for MS-based protein and proteomics research, the challenges in data analysis show large overlap. An introduction to the already well described problems associated with protein identification from tandem mass spectra can be found in [51]. Existing data analysis tools, however, are often specifically designed for (ESI)-MS data and might thus not perform optimally on MSI data. Occasionally, in-house developed algorithms are reported [44, 50]. In a MALDI-MSI study of the obese mouse pancreas, for example, tandem MS datasets were clustered to allow for the identification of structurally related peptides [50].

Importantly, the large datasets generated in MSI experiments create new bioinformatic challenges. Data are processed and mined to reduce the influence of technical and analytical variation and extract information relevant to the biological problem, either with standard

software (e.g. Biomap, ClinProTools [52]) or in-house developed algorithms [53]. The interested reader is referred to www.maldi-msi.org that provides a concise overview of available MSI software.

Key here is the identification and extraction of relevant spatial and chemical features. Mass spectra generated from different locations on a tissue probe differences in molecular make-up of that tissue and can thus be used for clustering or classification. Data mining methods allow for the identification of signature masses for specific tissue regions or tissue states, which can be assigned using either protein extraction or *in situ* digestion approaches. In this way, MSI enables the targeted analysis of relevant species for biomarker discovery.

Supervised methods make use of prior knowledge about the tissue, and typically use histological images to define different regions of interest (ROIs) [25, 30, 54-56]. These approaches are referred to as "histology-directed" MSI [54, 57]. Differentially expressed peaks between the ROIs are identified using statistical tests and used to generate classification models. A recent report showed that histology-directed classification of MALDI-MSI data led to the identification of differentially expressed modified protein species in skin cancer, as for example multiply acetylated forms of histones H4 and H2A [54]. As PTMs reflect the actual biological state of proteins, they can be highly relevant for biomarker discovery.

Unsupervised methods, including multivariate methods as Principal Component Analysis and Hierarchical Clustering, can reveal histology-independent regions [11, 58]. Combinations of (un)supervised methods and newly developed strategies are reported as well [10, 59, 60]. The use of multiple multivariate techniques on one MSI dataset can provide a more accurate description of regions with distinct mass spectrometric profiles, but this type of study is still in a developmental stage [61].

Furthermore, the large datasets generated by MSI have a big impact on the computational infrastructure necessary for data processing and analysis [53]. Further advancements in high throughput analysis are crucial to improve speed and reliability of protein annotation.

5. Validation

In the field of MSI the use of additional validation methods is a necessity. As explained in section 2.3 one should be careful with inferring identities between (imaging) datasets. Once MSI has pinpointed towards interesting species, they are typically further investigated using techniques from biochemistry such as IHC [31, 55] and *in situ* hybridization [56], and histology [20]. The use of standards, common practice in drug distribution studies, is seldom used for independent validation of peptides or proteins, because the preparation of (synthetic) isotope-labeled peptides or protein standards is often far from straightforward.

Also targeted chemical labeling of proteins for direct protein identification from tissue is reported. The addition of a tag allows for enhanced detection of specific species in an MSI experiment, but requires prior knowledge of the protein of interest. The advantage of such an approach is that it allows for multiplexed analysis of preselected proteins, usually a problem when using IHC. Moreover, it enables the analysis of low abundant and high mass proteins, which are hard to probe by MALDI-MSI.

Thiery and coworkers showed multiplex immunolabeling of proteins, named TAMSIM for TArgeted multiplex MS IMaging [62]. In this strategy, proteins are linked to an antibody with a mass tag, which is released upon laser irradiation and subsequently detected by MSI without the need for matrix addition. Lemaire and coworkers showed the similar concept of “Tag-Mass”: the addition of a probe with a photocleavable tag of known mass linked to mRNA or protein [63, 64]. Proof of principle was shown for the 180 kDa carboxypeptidase D membrane protein from rat brain tissue (Figure 4). As antibodies can show high specificity for their

corresponding antigen, this method allows for specific protein identification in an MSI experiment and can be used for validation, as shown for a new potential biomarker for ovary cancer [65]. The Mass-Tag molecule is now patented for use in quantitative diagnostic assays. All validation methods have one disadvantage in common, namely that they only allow for the validation of a limited number of proteins per study. At the moment it is not possible to validate MSI data on the whole proteome level.

An emerging approach for validation is to study the same sample with different mass spectrometric or spectroscopic techniques as magnetic resonance spectroscopic imaging [66, 67]. In addition to its use for independent verification, a multimodal approach can provide a more comprehensive view of the investigated sample.

6. Conclusions and future perspectives

The greatest strength of MSI is that it can provide relatively unbiased molecular information in an anatomical context. Used as a discovery tool, MSI can highlight interesting species to be further investigated. Alternatively, MSI can be utilized to visualize a biomedical hypothesis. Although MSI has proved its capability in biomedical research, it is not widely adopted yet. This is partly due to a still existing gap between technique-based method development of MSI and the demands from the biomedical research community.

In this review one of the main bottlenecks, namely protein identification in a MSI experiment, is addressed. At the moment, spectral annotation is a laborious and often complicated task for each new set of samples within a lab. The number of annotations therefore remains limited to at most tens per study, while the studied spectra contain easily a tenfold more signals.

An obvious way to improve annotation is to merge MSI data with complementary data as exemplified by the indirect approaches described. This will, however, require a substantial effort in the area of bioinformatics. Algorithms need to be developed to improve data correlation and allow for smarter and faster annotation workflows. Furthermore, the implementation of statistical evaluation methods commonly used in MS-based protein identification are expected to enhance the reliability of a MSI study.

Recently, a case was made for improved identification through community annotation [16]. An online data repository for published MSI datasets should allow for re-mining of data. In this way, one can benefit from annotations by labs with other expertise. This initiative exemplifies the need for improved annotation within the MSI community, but will face several challenges, least of all standardization of data formats.

Databanks with patient data combined with biomaterials already exist to facilitate (bio)medical research. Although the set-up and use of these databases is governed by strict guidelines from medical ethics, we are convinced that MSI data integration with these "biobanks", but also with imaging data from other imaging modalities and "-omics" data will greatly improve our capability to mine and annotate MSI data.

Recently, the technique of laserspray ionization (LSI) was applied to protein analysis directly from tissue sections [68]. Although this technique is not yet suited for imaging experiments, the generation of multiply charged ions directly from tissue, combined with high mass resolution and mass accuracy, might facilitate protein imaging and identification in the future. The generation of multiply charged species in LSI makes electron transfer dissociation (ETD) fragmentation on tissue possible. Formation of c- and z-ions during ETD provides complementary fragmentation information and also suggests the possibility to study posttranslational modifications retained during ETD fragmentation but lost during CID typically used in MSI.

For MSI to become a standard technique in biomedical research it is of the utmost importance that MSI workflows for protein identification are further developed to provide useful information to the (bio)medical research community. As researchers extend MSI technology to study more complex biological problems, there will be an increasing need for (bioinformatic) tools that improve the confidence and number of identified proteins within these studies. Ongoing efforts to embed MSI into the interdisciplinary world of life sciences will move the field into the next decade.

Acknowledgments

This work is part of the research program of the Foundation for Fundamental Research on Matter (FOM) which is financially supported by the Netherlands Organization for Fundamental Research (NWO). Authors acknowledge funding through the Dutch national program COMMIT.

References

[1] D.S. Cornett, M.L. Reyzer, P. Chaurand, R.M. Caprioli, MALDI imaging mass spectrometry: molecular snapshots of biochemical systems, Nature methods, 4 (2007) 828.

[2] K. Chughtai, R.M.A. Heeren, Mass Spectrometric Imaging for Biomedical Tissue Analysis, Chem Rev, 110 (2010) 3237.

[3] D.M. Drexler, T.J. Garrett, J.L. Cantone, R.W. Diters, J.G. Mitroka, M.C. Prieto Conaway, S.P. Adams, R.A. Yost, M. Sanders, Utility of imaging mass spectrometry (IMS) by matrix-assisted laser desorption ionization (MALDI) on an ion trap mass spectrometer in the analysis of drugs and metabolites in biological tissues, J Pharmacol Toxicol Methods, 55 (2007) 279.

[4] Y. Hsieh, J. Chen, W.A. Korfmacher, Mapping pharmaceuticals in tissues using MALDI imaging mass spectrometry, J Pharmacol Toxicol Methods, 55 (2007) 193.

[5] S. Rauser, S.O. Deininger, D. Suckau, H. Hofler, A. Walch, Approaching MALDI molecular imaging for clinical proteomic research: current state and fields of application, Expert Rev Proteomics, 7 (2010) 927.

[6] J. Franck, K. Arafah, M. Elayed, D. Bonnel, D. Vergara, A. Jacquet, D. Vinatier, M. Wisztorski, R. Day, I. Fournier, M. Salzet, MALDI imaging mass spectrometry: state of the art technology in clinical proteomics, Mol Cell Proteomics, 8 (2009) 2023.

[7] B. Balluff, C. Schone, H. Hofler, A. Walch, MALDI imaging mass spectrometry for direct tissue analysis: technological advancements and recent applications, Histochem Cell Biol, (2011).

[8] L.A. McDonnell, G.L. Corthals, S.M. Willems, A. van Remoortere, R.J. van Zeijl, A.M. Deelder, Peptide and protein imaging mass spectrometry in cancer research, J Proteomics, 73 (2010) 1921.

[9] E.H. Seeley, R.M. Caprioli, MALDI imaging mass spectrometry of human tissue: method challenges and clinical perspectives, Trends Biotechnol, 29 (2011) 136.

[10] S.M. Willems, A. van Remoortere, R. van Zeijl, A.M. Deelder, L.A. McDonnell, P.C. Hogendoorn, Imaging mass spectrometry of myxoid sarcomas identifies proteins and lipids specific to tumour type and grade, and reveals biochemical intratumour heterogeneity, J Pathol, 222 (2010) 400.

[11] S.O. Deininger, M.P. Ebert, A. Futterer, M. Gerhard, C. Rocken, MALDI imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers, J Proteome Res, 7 (2008) 5230.

[12] M. Karas, D. Bachmann, F. Hillenkamp, Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules, Anal. Chem., 57 (1985) 2935.

[13] R.M. Caprioli, T.B. Farmer, J. Gile, Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS, Anal Chem, 69 (1997) 4751.

[14] P. Chaurand, S.A. Schwartz, M.L. Reyzer, R.M. Caprioli, Imaging mass spectrometry: principles and potentials, Toxicol Pathol, 33 (2005) 92.

[15] B. Balluff, M. Elsner, A. Kowarsch, S. Rauser, S. Meding, C. Schuhmacher, M. Feith, K. Herrmann, C. Rocken, R.M. Schmid, H. Hofler, A. Walch, M.P. Ebert, Classification of HER2/neu status in gastric cancer using a breast-cancer derived proteome classifier, J Proteome Res, 9 (2010) 6317.

[16] J.D. Watrous, T. Alexandrov, P.C. Dorrestein, The evolving field of imaging mass spectrometry and its impact on future biological research, J. Mass Spectrom., 46 (2011) 209.
[17] B. Domon, R. Aebersold, Mass spectrometry and protein analysis, Science, 312 (2006) 212.
[18] J.R. Yates, C.I. Ruse, A. Nakorchevsky, Proteomics by mass spectrometry: approaches, advances, and applications, Annu Rev Biomed Eng, 11 (2009) 49.
[19] L. Minerva, S. Clerens, G. Baggerman, L. Arckens, Direct profiling and identification of peptide expression differences in the pancreas of control and ob/ob mice by imaging mass spectrometry, Proteomics, 8 (2008) 3763.
[20] S. Rauser, C. Marquardt, B. Balluff, S.O. Deininger, C. Albers, E. Belau, R. Hartmer, D. Suckau, K. Specht, M.P. Ebert, M. Schmitt, M. Aubele, H. Hofler, A. Walch, Classification of HER2 receptor status in breast cancer tissues by MALDI imaging mass spectrometry, J Proteome Res, 9 (2010) 1854.
[21] M. Lagarrigue, M. Becker, R. Lavigne, S.O. Deininger, A. Walch, F. Aubry, D. Suckau, C. Pineau, Revisiting rat spermatogenesis with MALDI imaging at 20-microm resolution, Mol Cell Proteomics, 10 (2011) M110 005991.
[22] H. Meistermann, J.L. Norris, H.R. Aerni, D.S. Cornett, A. Friedlein, A.R. Erskine, A. Augustin, M.C. De Vera Mudry, S. Ruepp, L. Suter, H. Langen, R.M. Caprioli, A. Ducret, Biomarker discovery by imaging mass spectrometry: transthyretin is a biomarker for gentamicin-induced nephrotoxicity in rat, Mol Cell Proteomics, 5 (2006) 1876.
[23] D. Debois, V. Bertrand, L. Quinton, M.C. De Pauw-Gillet, E. De Pauw, MALDI-in source decay applied to mass spectrometry imaging: a new tool for protein identification, Anal Chem, 82 (2010) 4036.
[24] Y.M. Fung, C.M. Adams, R.A. Zubarev, Electron ionization dissociation of singly and multiply charged peptides, J Am Chem Soc, 131 (2009) 9977.
[25] L.H. Cazares, D. Troyer, S. Mendrinos, R.A. Lance, J.O. Nyalwidhe, H.A. Beydoun, M.A. Clements, R.R. Drake, O.J. Semmes, Imaging mass spectrometry of a specific fragment of mitogen-activated protein kinase/extracellular signal-regulated kinase kinase kinase 2 discriminates cancer from uninvolved prostate tissue, Clin Cancer Res, 15 (2009) 5541.
[26] J.A. McLean, W.B. Ridenour, R.M. Caprioli, Profiling and imaging of tissues by imaging ion mobility-mass spectrometry, J. Mass Spectrom., 42 (2007) 1099.
[27] M.R. Groseclose, M. Andersson, W.M. Hardesty, R.M. Caprioli, Identification of proteins directly from tissue: in situ tryptic digestions coupled with imaging mass spectrometry, J. Mass Spectrom., 42 (2007) 254.
[28] M.C. Djidja, S. Francese, P.M. Loadman, C.W. Sutton, P. Scriven, E. Claude, M.F. Snel, J. Franck, M. Salzet, M.R. Clench, Detergent addition to tryptic digests and ion mobility separation prior to MS/MS improves peptide yield and protein identification for in situ proteomic investigation of frozen and formalin-fixed paraffin-embedded adenocarcinoma tissue sections, Proteomics, 9 (2009) 2750.
[29] R. Lemaire, A. Desmons, J.C. Tabet, R. Day, M. Salzet, I. Fournier, Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections, J Proteome Res, 6 (2007) 1295.
[30] M.R. Groseclose, P.P. Massion, P. Chaurand, R.M. Caprioli, High-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue microarrays using MALDI imaging mass spectrometry, Proteomics, 8 (2008) 3715.
[31] M.C. Djidja, E. Claude, M.F. Snel, P. Scriven, S. Francese, V. Carolan, M.R. Clench, MALDI-ion mobility separation-mass spectrometry imaging of glucose-regulated

protein 78 kDa (Grp78) in human formalin-fixed, paraffin-embedded pancreatic adenocarcinoma tissue sections, J Proteome Res, 8 (2009) 4876.
[32] M. Ronci, E. Bonanno, A. Colantoni, L. Pieroni, C. Di Ilio, L.G. Spagnoli, G. Federici, A. Urbani, Protein unlocking procedures of formalin-fixed paraffin-embedded tissues: application to MALDI-TOF imaging MS investigations, Proteomics, 8 (2008) 3702.
[33] J. Franck, M. El Ayed, M. Wisztorski, M. Salzet, I. Fournier, On-tissue N-terminal peptide derivatizations for enhancing protein identification in MALDI mass spectrometric imaging strategies, Anal Chem, 81 (2009) 8305.
[34] J. Stauber, L. MacAleese, J. Franck, E. Claude, M. Snel, B.K. Kaletas, I.M. Wiel, M. Wisztorski, I. Fournier, R.M. Heeren, On-tissue protein identification and imaging by MALDI-ion mobility mass spectrometry, J Am Soc Mass Spectrom, 21 (2010) 338.
[35] L.M. Cole, M.C. Djidja, J. Bluff, E. Claude, V.A. Carolan, M. Paley, G.M. Tozer, M.R. Clench, Investigation of protein induction in tumour vascular targeted strategies by MALDI MSI, Methods, 54 (2011) 442.
[36] J. Stauber, R. Lemaire, J. Franck, D. Bonnel, D. Croix, R. Day, M. Wisztorski, I. Fournier, M. Salzet, MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting, J Proteome Res, 7 (2008) 969.
[37] E.B. Monroe, S.P. Annangudi, N.G. Hatcher, H.B. Gutstein, S.S. Rubakhin, J.V. Sweedler, SIMS and MALDI MS imaging of the spinal cord, Proteomics, 8 (2008) 3746.
[38] L. MacAleese, M.C. Duursma, L.A. Klerk, G. Fisher, R.M. Heeren, Protein identification with Liquid Chromatography and Matrix Enhanced Secondary Ion Mass Spectrometry (LC-ME-SIMS), J Proteomics, 74 (2011) 993.
[39] D. Mustafa, J.M. Kros, T. Luider, Combining laser capture microdissection and proteomics techniques, Methods Mol Biol, 428 (2008) 159.
[40] R.B. Braakman, M.M. Tilanus-Linthorst, N.Q. Liu, C. Stingl, L.J. Dekker, T.M. Luider, J.W. Martens, J.A. Foekens, A. Umar, Optimized nLC-MS workflow for laser capture microdissected breast cancer tissue, J Proteomics, (2012).
[41] J.O. Gustafsson, M.K. Oehler, S.R. McColl, P. Hoffmann, Citric acid antigen retrieval (CAAR) for tryptic peptide imaging directly on archived formalin-fixed paraffin-embedded tissue, J Proteome Res, 9 (2010) 4315.
[42] S.R. Oppenheimer, D. Mi, M.E. Sanders, R.M. Caprioli, Molecular analysis of tumor margins by MALDI mass spectrometry in renal carcinoma, J Proteome Res, 9 (2010) 2182.
[43] A.F.M. Altelaar, I.M. Taban, L.A. McDonnell, P.D.E.M. Verhaert, R.P.J. de Lange, R.A.H. Adan, W.J. Mooi, R.M.A. Heeren, S.R. Piersma, High-resolution MALDI imaging mass spectrometry allows localization of peptide distributions at cellular length scales in pituitary tissue sections, Int. J. Mass spectrom., 260 (2007) 203.
[44] Y. Schober, T. Schramm, B. Spengler, A. Rompp, Protein identification by accurate mass matrix-assisted laser desorption/ionization imaging of tryptic peptides, Rapid Commun Mass Spectrom, 25 (2011) 2475.
[45] A. Rompp, S. Guenther, Y. Schober, O. Schulz, Z. Takats, W. Kummer, B. Spengler, Histology by mass spectrometry: label-free tissue characterization obtained from high-accuracy bioanalytical imaging, Angew Chem Int Ed Engl, 49 (2010) 3834.
[46] I.M. Taban, A.F. Altelaar, Y.E. van der Burgt, L.A. McDonnell, R.M. Heeren, J. Fuchser, G. Baykut, Imaging of peptides in the rat brain using MALDI-FTICR mass spectrometry, J Am Soc Mass Spectrom, 18 (2007) 145.

[47] R. Zubarev, M. Mann, On the proper use of mass accuracy in proteomics, Molecular & Cellular Proteomics, 6 (2007) 377.
[48] T. Liu, M.E. Belov, N. Jaitly, W.J. Qian, R.D. Smith, Accurate mass measurements in proteomics, Chem. Rev., 107 (2007) 3621.
[49] Y. Schober, S. Guenther, B. Spengler, A. Rompp, High-resolution matrix-assisted laser desorption/ionization imaging of tryptic peptides from tissue, Rapid Commun. Mass Spectrom., 26 (2012) 1141.
[50] L. Minerva, K. Boonen, G. Menschaert, B. Landuyt, G. Baggerman, L. Arckens, Linking mass spectrometric imaging and traditional peptidomics: a validation in the obese mouse model, Anal Chem, 83 (2011) 7682.
[51] W.S. Noble, M.J. MacCoss, Computational and statistical analysis of protein mass spectrometry data, PLoS Comput Biol, 8 (2012) e1002296.
[52] R. Ketterlinus, S.Y. Hsieh, S.H. Teng, H. Lee, W. Pusch, Fishing for biomarkers: analyzing mass spectrometry data with the new ClinProTools software, BioTechniques, Suppl (2005) 37.
[53] L.A. McDonnell, A. van Remoortere, N. de Velde, R.J. van Zeijl, A.M. Deelder, Imaging mass spectrometry data reduction: automated feature identification and extraction, J Am Soc Mass Spectrom, 21 (2010) 1969.
[54] W.M. Hardesty, M.C. Kelley, D. Mi, R.L. Low, R.M. Caprioli, Protein signatures for survival and recurrence in metastatic melanoma, J Proteomics, 74 (2011) 1002.
[55] J. Le Faouder, S. Laouirem, M. Chapelle, M. Albuquerque, J. Belghiti, F. Degos, V. Paradis, J.M. Camadro, P. Bedossa, Imaging mass spectrometry provides fingerprints for distinguishing hepatocellular carcinoma from cirrhosis, J Proteome Res, 10 (2011) 3755.
[56] J. Bruand, S. Sistla, C. Meriaux, P.C. Dorrestein, T. Gaasterland, M. Ghassemian, M. Wisztorski, I. Fournier, M. Salzet, E. Macagno, V. Bafna, Automated querying and identification of novel peptides using MALDI mass spectrometric imaging, J Proteome Res, 10 (2011) 1915.
[57] D.S. Cornett, J.A. Mobley, E.C. Dias, M. Andersson, C.L. Arteaga, M.E. Sanders, R.M. Caprioli, A novel histology-directed strategy for MALDI-MS tissue profiling that improves throughput and cellular specificity in human breast cancer, Mol Cell Proteomics, 5 (2006) 1975.
[58] M. Hanselmann, M. Kirchner, B.Y. Renard, E.R. Amstalden, K. Glunde, R.M.A. Heeren, F.A. Hamprecht, Concise Representation of Mass Spectrometry Images by Probabilistic Latent Semantic Analysis, Anal Chem, 80 (2008) 9649.
[59] D. Bonnel, R. Longuespee, J. Franck, M. Roudbaraki, P. Gosset, R. Day, M. Salzet, I. Fournier, Multivariate analyses for biomarkers hunting and validation through on-tissue bottom-up or in-source decay in MALDI-MSI: application to prostate cancer, Anal. Bioanal. Chem., 401 (2011) 149.
[60] G. McCombie, D. Staab, M. Stoeckli, R. Knochenmuss, Spatial and spectral correlations in MALDI mass spectrometry images by clustering and multivariate analysis, Anal Chem, 77 (2005) 6118.
[61] E.A. Jones, A. van Remoortere, R.J.M. van Zeijl, P.C.W. Hogendoorn, J.V.M.G. Bovee, A.M. Deelder, L.A. McDonnell, Multiple Statistical Analysis Techniques Corroborate Intratumor Heterogeneity in Imaging Mass Spectrometry Datasets of Myxofibrosarcoma, Plos One, 6 (2011).
[62] G. Thiery, M.S. Shchepinov, E.M. Southern, A. Audebourg, V. Audard, B. Terris, I.G. Gut, Multiplex target protein imaging in tissue sections by mass spectrometry--TAMSIM, Rapid Commun Mass Spectrom, 21 (2007) 823.

[63] R. Lemaire, J. Stauber, M. Wisztorski, C. Van Camp, A. Desmons, M. Deschamps, G. Proess, I. Rudlof, A.S. Woods, R. Day, M. Salzet, I. Fournier, Tag-mass: specific molecular imaging of transcriptome and proteome by mass spectrometry based on photocleavable tag, J Proteome Res, 6 (2007) 2057.

[64] J. Stauber, M.E. Ayed, M. Wisztorski, M. Salzet, I. Fournier, Specific MALDI-MSI: Tag-Mass, Methods Mol Biol, 656 (2010) 339.

[65] R. Lemaire, S.A. Menguellet, J. Stauber, V. Marchaudon, J.P. Lucot, P. Collinet, M.O. Farine, D. Vinatier, R. Day, P. Ducoroy, M. Salzet, I. Fournier, Specific MALDI imaging and profiling for biomarker hunting and validation: fragment of the 11S proteasome activator complex, Reg alpha fragment, is a new potential ovary cancer biomarker, J Proteome Res, 6 (2007) 4127.

[66] E.R. Amstalden van Hove, T.R. Blackwell, I. Klinkert, G.B. Eijkel, R.M. Heeren, K. Glunde, Multimodal mass spectrometric imaging of small molecules reveals distinct spatio-molecular signatures in differentially metastatic breast tumor models, Cancer Res, 70 (2010) 9012.

[67] K. Chughtai, L. Jiang, T.R. Greenwood, I. Klinkert, E.R. Amstalden van Hove, R.M. Heeren, K. Glunde, Fiducial markers for combined 3-dimensional mass spectrometric and optical tissue imaging, Anal Chem, 84 (2012) 1817.

[68] E.D. Inutan, A.L. Richards, J. Wager-Miller, K. Mackie, C.N. McEwen, S. Trimpin, Laserspray ionization, a new method for protein analysis directly from tissue at atmospheric pressure with ultrahigh mass resolution and electron transfer dissociation, Mol Cell Proteomics, 10 (2011) M110 000760.

Figure captions

Figure 1.
Workflow for MS-based protein identification. Identification can be performed within the MSI experiment itself (direct identification) or by using independent MS/MS data followed by mass correlation (indirect identification).

Figure 2.
Ion mobility separation combined with MSI. On-tissue tandem MS spectra of ion mobility separated tryptic peptides (*m/z* 1039), identified as tubulin and ubiquitin fragments (A). Ion images of separated tubulin and ubiquitin fragments (B). Without ion mobility separation (no drift time selection), the ion image would have corresponded to the superposition of the two images. Adapted from [34] with permission.

Figure 3.
High mass resolution MSI not only improves the reliability of peptide assignment, but also the spatial distribution information. Zoomed Orbitrap mass spectrum of a MALDI-MSI analysis of mouse brain (A). An overlay of ion images generated with an *m/z* bin width of ±0.01 shows the different spatial distributions of a myelin tryptic peptide (red) and a phospholipid (green) at *m/z* 726.405 and *m/z* 726.515, respectively (B). An ion image generated with a larger bin width of ±0.1 leads to a superposition of the two images, hence the spatial distribution of the two ions is not resolved (C). An overlay of ion images of a tryptic peptide of SNAP-91 (green) and a myelin peptide isotopomeric peak (red) at *m/z* 727.315±0.01 and *m/z* 727.405±0.01 (D) and an ion image at *m/z* 727.4±0.1 (E) show the same effect. Reprinted from [49] with permission.
[70]

Figure 4.
Example of the Tagg-Mass concept. MALDI mass spectra from adjacent rat brain tissues after IHC experiment against carboxypeptidase D (CPD), with untagged (A) and tagged (B) secondary antibody. Two characteristic signals for the Mass-Tag (P-PC) were observed (*m/z* 1686.43 and *m/z* 1703.23). Corresponding ion image at *m/z* 1686.43 (C). Rat brain tissue before analysis (D). Similar results were obtained with secondary antibody detection with fluorescence (E) or peroxidase staining using 4-chloronaphtol (F). Reprinted from [63] with permission.

Figure 1
Click here to download high resolution image

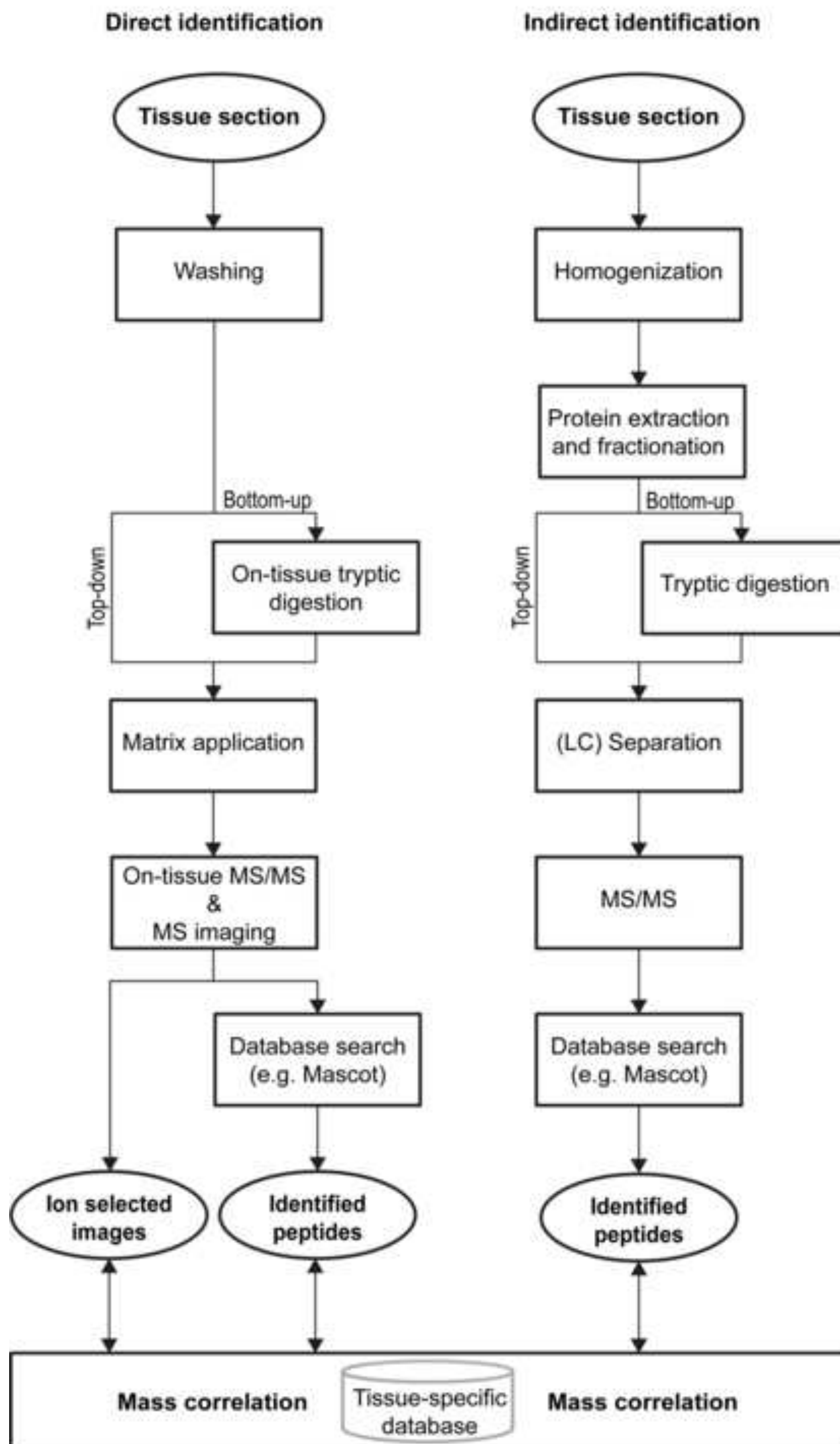


Figure 2A

Click here to download high resolution image

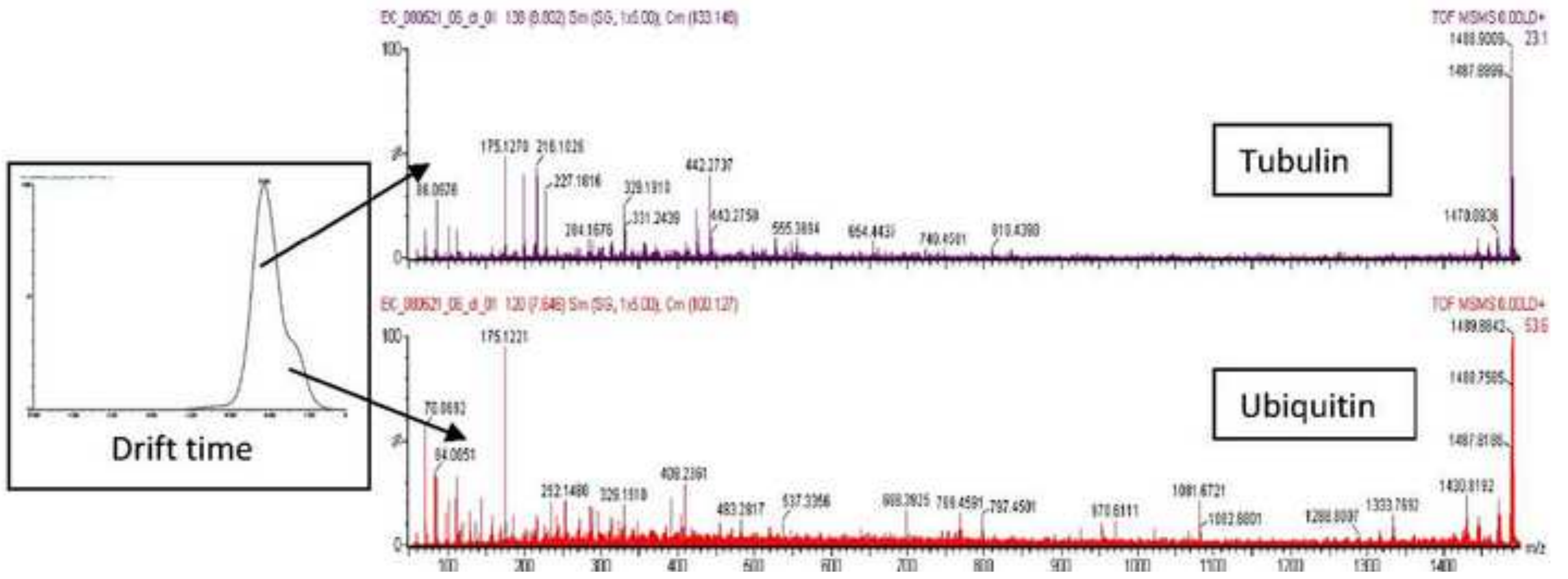
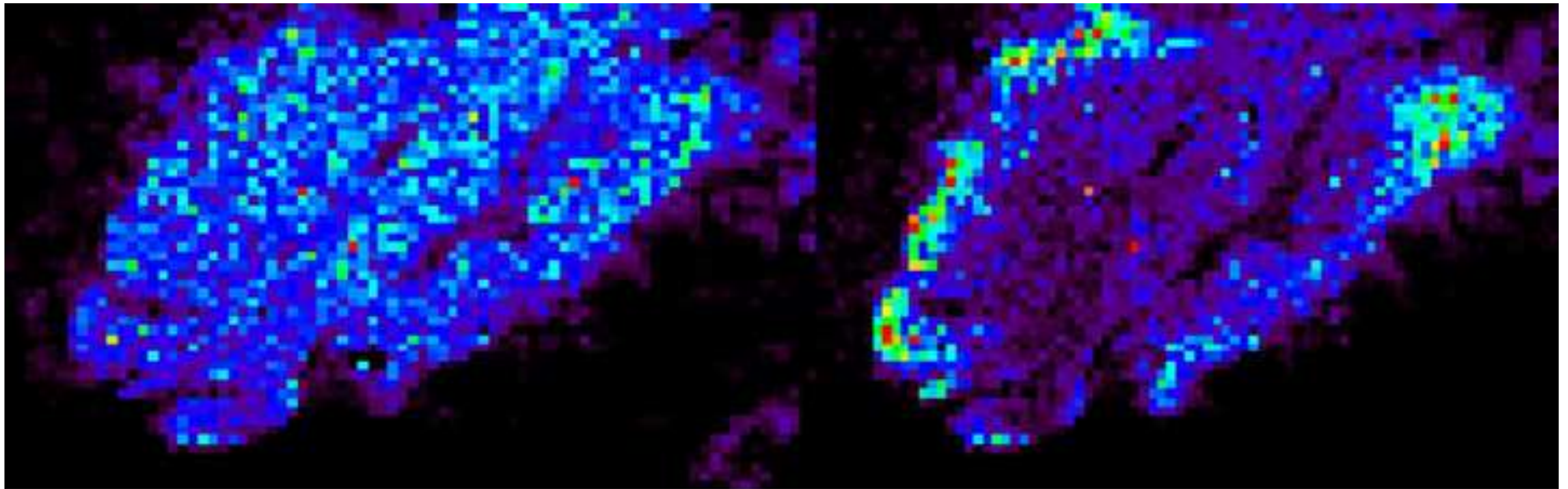


Figure 2B

Click here to download high resolution image



High mobility peptide conformation

Tubulin

Low mobility peptide conformation

Ubiquitin

Figure 3
Click here to download high resolution image

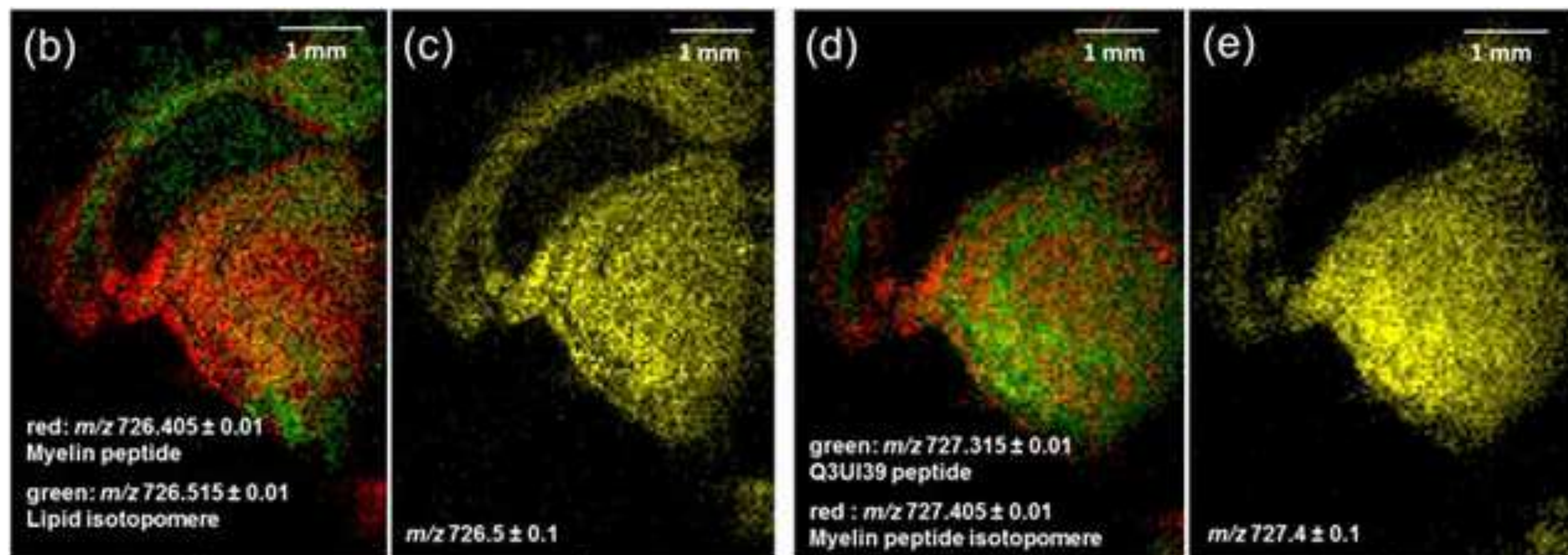


Figure 4

Click here to download high resolution image

Revised version with changes highlighted
Click here to download Supplementary material: IDinMSI_revised_with changes.docx