Intact cell MALDI-TOF mass spectrometry, a promising proteomic profiling method in farm animal clinical and reproduction research

Laura Soler, Svetlana Uzbekova, Elisabeth Blesbois, Xavier Druart, Valérie Labas

Revised

Intact cell MALDI-TOF Mass Spectrometry, a promising proteomic profiling method in farm animal clinical and reproduction research

Laura Soler[1,2], Svetlana Uzbekova[2], Elisabeth Blesbois[2], Xavier Druart[2,3] †and Valérie Labas[2,3]

1) INRA, Université de Toulouse, ENVT, INP-Purpan, UPS, Toxalim (Research Centre in Food Toxicology), Toulouse, France

2) INRA UMR85-CNRS UMR7247, Université de Tours, IFCE Physiologie de la Reproduction et des Comportements, Nouzilly, France

3) INRA, Université de Tours, CHU de Tours, Plate-forme de Chirurgie et d'Imagerie pour la Recherche et l'Enseignement (CIRE), Pôle d'Analyse et d'Imagerie des Biomolécules (PAIB), Nouzilly, France

† Corresponding author.

Abstract

The objective of this review is to provide new insights into the possible use of a proteomic method known as Intact Cell Matrix-Assisted Laser Desorption-ionization Time-Of-Flight Mass Spectrometry (ICM-MS) in animal clinical research. Here, we give an overview of the basics of this technique, its advantages and disadvantages compared with other proteomic approaches, past applications and future perspectives. A special emphasis on its implementation in animal reproduction science is given, including examples of the reliable use of ICM-MS on fertility screening. In mammals, the ICM-MS profiles from pig epididymal spermatozoa reflect the proteome changes that they undergo during epididymal maturation and could be associated with the acquisition of fertilizing ability. In chicken, using adequate pre-processing and bioinformatics analysis tools, sperm ICM-MS profiles showed characteristic spectral features that allowed their classification according to their actual

fertilizing ability. The association of ICM-MS and Top-down proteomic strategies allowed the identification of chicken fertility biomarkers candidates such as protein vitelline membrane outer layer protein 1 (MVO-1) and avian beta-defensin 10 (AVB10). In female reproduction, a similar approach on ovarian follicular cells allowed the identification of specific markers of oocyte maturation in the oocyte and surrounding cumulus cells. Altogether, these results indicate that ICM-MS profiling could be a suitable approach for molecular phenotyping of male and female gametes.

1. Clinical proteomics in animal sciences

Clinical proteomics is a field of proteomics whose objective is the identification of biomolecules showing differential abundance in a certain body fluid, cell and/or tissue associated with a defined clinical condition. These molecules, usually denominated “biomarkers”, can be thereafter used in clinical applications for diagnosis, stratification of patients for specific treatment, or therapy monitoring. A combination of different proteomic biomarkers or “biomarker profile” can also be employed, giving a combined measure associated with the studied condition [1,2]. Biomarker profiles are potentially better at defining a complex patho-physiological condition than single biomarkers, and clinical proteomics is moving its focus to the identification of such fingerprints panels that can define a diagnostic phenotype [1]. In animal sciences, clinical proteomics has not been as widely used as in human medicine, but there is growing interest in identifying and implementing the use of disease biomarker profiles [3]. However, applications of veterinary clinical proteomics go beyond, and also include food security, quality assessment of animal products (meat, milk, eggs, wool), generation of model organisms and phenotyping of production traits [4,5]. One good example of the latter is the application of clinical proteomics in animal reproduction science, and more specifically in searching for gamete quality biomarkers. Optimization of reproduction management is actually essential in farming since keeping high productivity often relies on adequate selection of breeders as well as the use of optimal gamete handling protocols (sperm

cryopreservation, oocyte in vitro maturation, in vitro fertilization, etc) [6]. Actually, one of the main problems in animal breeding is that the selection for production traits, such as fast growth, egg-laying capacity or milk productivity, usually has a detrimental effect in reproductive traits [7]. Breeders are therefore interested in the accurate phenotyping of the fertility of their animals to maintain high productivity [7].

2. Technical approaches employed in clinical proteomics: the emergence of protein profiling platforms

Traditionally, biomarker research using proteomics starts with an initial discovery step, whose objective is untargeted screening of the whole proteome of a certain biological system in defined conditions, in search of biomarkers [8]. The initial discovery approach is usually based on the orthogonal separation of whole proteins contained in the sample either using gel electrophoresis or liquid chromatography followed by a "bottom-up" protein analysis (Figure 1). In this strategy, isolated proteins or protein mixtures are digested with a protease to obtain peptides. These fragments of proteins can be analyzed directly by mass spectrometry (MS) and identified by a peptide mass fingerprint (PMF) strategy [9]; and/or using tandem mass spectrometry (MS/MS) using a peptide fragment fingerprint (PFF) in order to obtain sequence peptides [10]. However, this method is cumbersome and complicated to be implemented in a clinical setting for routine diagnostics, since it involves complex sample preparation procedures, analytic processing of technical complexity and limited throughput. This workflow typically results in a long list of candidates (peptides/proteins showing differential relative abundance) to be integrated in order to construct a biomarker profile. These need to be verified and validated following a second analysis, this time using large cohorts, and high-throughput targeted analysis through the application of multiplexed assays [11]. One of the most recent and efficient techniques developed for the latter is known as MRM (Multiple Reaction Monitoring), which shows several advantages such as being accurate, fast and very sensitive [12]. However, the applicability of this technology in clinical settings remains to be

seen, since the development of such tests depends on the use of pre-existing quality MS data, is expensive, as well as the specific detection and/or quantification of some of the desired targets can be challenging [13].

Behind the tremendous progress and possibilities in mass spectrometry, “Bottom-up” proteomics has been and is still the driving force in proteome analysis. However, the proper assignment of proteoforms with their pattern of modifications remains a challenge for bottom-up proteomics. The real goals in proteomic analyses are knowing the functions of the intact proteins in biological processes with or without their different modifications, their locations and interactions. Many of these aspects may get lost when breaking down proteins to their corresponding peptide fragments. Beside this conventional strategy, the top-down mass spectrometry constitutes an alternative (Figure1) [14, 15]. In this approach, intact proteins are analysed by MS and MS/MS in order to obtain structural informations allowing both the identification of the proteins and the characterization of the post-translational modifications. To perform Top-down proteomics, protein fractionation combined to high-resolution mass spectrometry and specific bioinformatics tools are necessary. Despite the technical difficulty of proteome-wide analysis at the intact protein level, Top-down proteomics offers an alternative to digestion-based approaches. For several years, there is growing interest in developing untargeted methodologies based on the use of peptide/protein profiling platforms, of which the most important are based on the use of Surface Enhanced Laser Desorption Ionization Time-Of-Flight Mass Spectrometry (SELDI-TOF MS), Capillary Electrophoresis Mass Spectrometry (CE-MS, [16]) or Matrix-Assisted Laser Desorption/Ionization Time Of Flight Mass Spectrometry (MALDI-TOF MS). One of the first approaches to peptide/protein profiling is SELDI-TOF MS, which employs active coatings/surfaces with affinity for specific biomolecules. Liquid samples (biological fluids or tissue extracts) are incubated with these surfaces, thus reducing sample complexity and allowing the acquisition of readable peptide/protein profiles by applying MS analysis directly to the surface-bound molecules. These profiles are later compared between conditions, in search for diagnostic patterns [16]. Several studies have explored the possible application of SELDI-

TOF profiling in veterinary sciences in the past, e.g for the detection of porcine diseases [17] or bovine paratuberculosis [18]. However, it is nowadays generally accepted that its low reproducibility (inter-laboratory CV% values above 40%) and lack of analytical robustness (difficulties with calibration and mass accuracy) impair its routine use in the clinic [13]. Capillary electrophoresis mass spectrometry (CE-MS) appeared later as an alternative to SELDI-TOF MS profiling using biofluids. This technique employs biomolecule separation through capillary electrophoresis as a method to reduce sample complexity. Separated molecules are then directly analyzed by MS to obtain a profile [19]. While CE-MS shows improved reproducibility and sensitivity compared to SELDI-TOF MS, it requires laborious sample preparation and shows low throughput, which are negative aspects in clinical proteomics [20]. CE-MS has been successfully applied in animal sciences [21,22], although it is recognized that its application to clinical proteomics is hampered by its low throughput, the need to standardize and simplify the sample preparation, as well as the development of accurate methods for calibration and data extraction/processing .

Along with CE-MS, Matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) profiling is an emerging approach in clinical proteomics. This technique has theoretically a higher potential for clinical application than CE-MS due to its high throughput, ease of operation, simple sample preparation and the existence of established data-processing platforms . MALDI-TOF MS has the ability to detect the most intense intact molecular ions present directly in the biological sample over a relatively wide mass range (typically 1,000 – 30,000 Da), with a high sensitivity (from the picomolar to the femtomolar concentrations) and a high tolerance for contaminants such as salts [23]. Consequently, MALDI-MS constitutes a powerful tool for the direct analysis of complex peptide/protein mixtures within biological specimens such as crude extracts and intact cells. It is in fact, the possibility of using whole, untreated cells what makes this technique particularly promising for clinical applications. In spite of this, this platform has been rarely used in medicine and even less in animal sciences. In this review, we will focus on the technical aspects of the platform MALDI-TOF MS using intact cells (known as Intact cell MALDI-TO MS or ICM-

MS), as well as the past applications and possibilities of this technique especially in animal sciences. Particularly, we will provide with a more in-deep overview on the past use of this platform in animal gametes.

3. Intact cell MALDI-TOF MS: the basics

In the following paragraphs we will describe the typical ICM-MS protocol workflow (Figure 2).

3.1. Sample preparation

An important application of ICM-MS is to detect endogenous peptides and proteins (intact forms unmodified by any chemical treatments) directly in whole cells using MALDI-TOF profiling without previous sample pre-fractionation, protein extraction or treatment. In this context, "intact cell" means that whole cells are subjected to analysis without any preparatory steps, although they can be cryopreserved or kept frozen prior to analysis. Direct analysis of cells by MALDI-TOF MS thus avoids the problems associated with extraction procedures (lost or degraded biological material), avoids the variability introduced by sample preparation and allows significant increase of a number of conditions to compare (up to several hundred). Samples susceptible of being analyzed through ICM-MS must consist of homogenous cellular populations, thus derived from cell culture or being isolated from a cell mix. Pre-treatment of such cells only requires the removal of the biological matrix where cells are contained, such as biological fluids or culture media. Usually, cells are harvested, rinsed and centrifuged, discarding after the supernatant containing biological matrix proteins. Wash buffers must keep cell integrity, promote crystallization and ionization, and ideally chelate salts. Good examples of the latter are Tris HCl-sucrose buffer (osmolarity at ≈ 300 mOsm/L) [24,25] or ammonium acetate [26,27]. It is important to avoid buffers containing phosphate-buffered saline (PBS), since the salts and the negative charge provided by PBS can interfere with the analysis.

3.2. Matrix deposition

Although sample deposition on the MALDI plate is a simple process, the protocol followed at this stage can have a significant effect on the outcome of the analysis [28], and so this has to be optimized for each sample type. A known amount of cells (usually 2×10^3 to 1×10^6 cells) is deposited on the target of a conductive MALDI metallic plate and mixed with an adequate matrix solution. The matrix consists of crystallized molecules, the three most commonly used being 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid) [24-27,29-33], α-cyano-4-hydroxycinnamic acid (CHCA, alpha-cyano or alpha-matrix) [34,35] and 2,5-dihydroxybenzoic acid (DHB) [36]. A solution of one of these molecules is made, often in a mixture of highly purified water and UV-absorbing weak organic acid such as acetonitrile (ACN), which is present in a concentration ranging 40-60%. A counter ion source such as Trifluoroacetic acid (TFA) is usually added to generate the ions, in a concentration ranging 0.1-3%. The ratio of analyte to matrix is important, and must be optimized for each sample type. There are multiple ways of performing the mix of samples and matrix. Among these methods, the 'dried droplet' method is the most usually employed [24-26, 29, 32-34,36]. It consists of depositing the sample and subsequently the matrix on the plate and then allowing co-crystallization. However, the thin layer [30-31] (a saturated matrix layer + mix of sample/matrix) or the sandwich [27] (a saturated matrix layer + dried droplet) methods can provide better results on a qualitative level. The resulting mixture is then allowed to completely dry at room temperature.

3.3. Spectra acquisition

Once the MALDI sample plate is prepared and placed within the mass spectrometer, a laser pulse hits the cells-matrix mixture. The matrix absorbs the laser energy, causing its own desorption and protonation, and subsequent sample protonation and vaporization. The generated ions are then accelerated by an applied high voltage (15-25 kV), separated in a linear field-free tunnel and detected as an electrical signal at the end of the TOF analyzer. Peptides and proteins of different masses and charges (mostly under the single-charged ions form) then fly at different speeds through the TOF analyzer. Their arrival times at the detector are distributed according to the square root of

the mass-to-charge ratio ($\sqrt{m/z}$). The result is a spectrum in the range of 1,000-30,000 m/z, whose components (peaks) show a degree of intensity that correlates with their abundance in the original sample. Each spectrum is produced by averaging several hundred to thousand laser shots that are spread over the entire sample spot surface. Several replicate spectra (a minimum of 6, up to 12) are taken for each biological sample under the same experimental conditions and a minimum of three reads per spot are recommended to be performed. Spectral acquisition is performed using an automated MS mode, thus increasing the speed and throughput of this technique. External mass calibration must be performed, using a peptide and protein known mixture whose masses must cover the analytical mass range. A high degree of mass accuracy is desired, and typical calibration error should not be higher than 0.5%. In order to optimize mass accuracy, it is recommended to apply an additional internal calibration (by superposing sample and calibration standard) as well as a lock-mass correction using a high-intensity mass at the middle of the mass range (error of 0.05%).

3.4. Spectra pre-processing

Raw MALDI data usually contain a large set of spectra, each containing thousands of measurements composed by m/z-intensity pairs, most of which correspond to noisy signals. In order to distinguish true from noise signals as well as to reduce experimental variance within the data set, several spectral treatments (known as pre-processing) must be carried out. This pre-processing includes the following steps: (1) internal lock-mass calibration using a mass of known or unknown identity, to provide with a higher mass accuracy, (2) baseline correction to remove the effect of noise introduced by the matrix, (3) filtering to smooth the signal, (4) alignment of all spectra using common m/z, (5) automatic peak detection on average spectra (area and/or intensity), and (6) normalization. The criteria followed for this pre-processing can drastically change the outcome of analysis, and so it is essential to pay attention to this stage in order to perform a rigorous study [37].

3.5. Statistical analysis and pattern matching algorithms

Once this is accomplished, applying statistical methods and/or machine-learning techniques is often desired in order to identify those spectral features (peaks) associated with a given condition (candidate biomarkers or biomarker profiles), sample grouping (through hierarchical classification or principal component analysis) and classification (through the constructions of class-predictive models based on the use of data-dependent algorithms or through comparison with reference spectra contained in a database). While processing and data analysis were initially based on the use of general tools like R or SPSS (Statistical Package for the Social Sciences) softwares, requiring programming skills and implying a lack of uniformity in the criteria followed for this process, a number of user-friendly tools specific for MALDI profiling data are currently available. In general, there are two kinds of tools regarding the type of post-processing analysis: those based on the use of a reference spectra database to which test spectra are compared, and those in which the user sets examples of the different studied classes. The first group of tools includes commercial software packages containing reference databases, and its use is mainly restricted to microorganism identification (e.g. MALDI Biotyper from Bruker Daltonics or Anagnostec SARAMIS from Shimadzu). The second group entails tools that allow both supervised and unsupervised techniques to analyze data following a clinical approach. In general, the biomarker research-oriented analysis starts with performing a statistical analysis (parametric or non-parametric) using quantitative values (normalized peak height or area). The m/z peaks presenting significant differences between the different studied conditions are selected and relative data are then utilized to perform additional analysis such as Receiver Operating Characteristic curves, cluster analysis, Principal Component Analysis, etc. It is important to set a statistical stringency adequate to the desired statistical power (usually set at 0.8), especially in studies using a reduced sample number. This statistical analysis can be completed with machine-learning techniques. In this case, the diagnostic performance of the ICM-MS-based predictive model for a given condition should be compared with that of traditional diagnostic tests of the same condition. In the recent years, different open-source tools (e.g. Mass-up, MALDIquant) and

commercial tools (e.g. ProgenesisMALDI-Nonlinear Dynamics, ClinProTools-Bruker Daltonics. MarkerView™- AB Sciex) specifically designed for MALDI profiling analysis have been released.

4. Top-Down MS for identification of proteins detected by ICM-MS

The progress in the general use of MS-based profiling has been somewhat hampered by the difficulties in confidently identifying diagnostic MS signatures [38]. ICM-MS analyses (in general all profiling methods) are ideally completed with the identification of the m/z peaks present in the spectra. As aforementioned, these masses correspond to native, endogenous molecules present in the sample. Of all the mass spectrometry-based identification approaches currently available (Figure 3), Top-Down proteomics represents the best approach to identify endogenous peptidoforms and proteoforms, as involves the detection of an intact molecular species and its direct fragmentation, thus providing with a complete description of the primary structure of the protein and of its modifications [39].

As it is generally desired to employ the same platform for both profiling and sequencing, following an on-line approach, there is big interest in the scientific community in advancing in technology and data-processing systems that would allow MALDI Top-Down analysis in complex samples (such as protein extracts, whole cells or tissue slides). Post-Source (PSD) [40] or In-Source Decay (ISD) [41] MS methodologies can be used for de novo sequencing of the whole proteins, with the condition of these being enriched or purified, since their main limitation is the inability of dealing with complex samples, and the MS/MS strategy is routinely restricted to small peptides with m/z < 5,000 [42,43]. Thus, the *in-situ* targeted identification of peaks of interest observed by MALDI remains challenging. Sample complexity can be decreased through fractionation (e.g. using chromatography, gel separation, depletion of abundant proteins, etc), and fractions are later re-analyzed by MALDI-TOF/TOF. This approach indeed increases the number of positively identified biomolecules, but still shows a low throughput. Alternatively, Top-Down proteomics combining liquid chromatography separations and high resolution tandem mass spectrometry (LC-HR-MS/MS) with

electrospray (ESI) source seems to be the most appropriate approach to identify endogenous peptido- and proteoforms [36,41]. ESI generates multiply charged ions and is the preferred MS method for the analysis of both peptides and intact proteins, especially on a proteome-wide scale (global approach). Nowadays, Top-Down proteomic studies have largely been implemented using ESI coupled to either Fourier transform ion cyclotron resonance (FT-ICR) or Orbitrap mass analyzers to obtain high resolution (R> 100,000) and mass accuracy (< 5 ppm). Thus, a large number of whole, intact biomolecules can be directly identified, whose masses can be matched with MALDI m/z peaks, with a mass accuracy tolerance < 0.05%. As described in the literature, the combination of ICM-MS and Top-Down HRMS is therefore a convenient analytical strategy [44, 45].

5. Intact Cell MALDI-TOF Mass Spectrometry – from bacteria to mammalian cells

The use of ICM-MS was firstly reported in 1996 by Holland et al., and was applied to bacterial chemotaxonomy [46]. In this study, it was described that ICM-MS protein profiles contained species-specific signatures, so they could be used as a taxonomic tool [47]. In effect, the analysis of endogenous peptidoforms and proteoforms can be considered as a functional genomics analysis and therefore useful for phenotyping [48]. ICM-MS has been applied in different studies for bacterial identification, aiding at the development of the field of phyloproteomics [46,49,50]. Identification of the microorganisms was based on the identification of species-specific signatures in the ICM-MS spectra using bioinformatic tools and in the subsequent construction of reference spectral databanks [51]. Thereafter, comprehensive and curated databases were constructed using rigorous standardization procedures and commercialized together with tools that allowed a robust identification analysis [52, 53]. Nowadays, ICM-MS has become a valuable tool for the identification and the classification of microorganisms, including bacteria [47] or yeast [54-56], and it has been introduced in clinical routine microbiological diagnostics. Thus, this approach assists clinicians to provide a fast, accurate diagnosis associated with an adequate antimicrobial treatment in infectious diseases. Furthermore, this approach can also be applied to detect specific bacterial biomarkers, such

as antibiotic resistance, and may potentially predict infectious complications or/and being used for the development of novel virulence factor-based therapies [57].

Theoretically, ICM-MS could be applied to any cell type including superior eukaryotes, with the sole condition of cells being isolated (not in a tissue). Even if this approach is not yet routinely applied to the analysis of mammalian cells, there is sufficient evidences in the literature showing that ICM-MS spectra from “whole cells” or “intact cells” are sufficiently characteristic to address biological questions in human and animal clinical research (Table 1). In fact, ICM-MS has been applied mostly in cultured cells as a phenotyping tool in order to differentiate similar cell lines including glial [27] or cancer [58] lines. It has also been employed to evaluate the response of established cell lines to different treatments, like toxic chemicals [30]. Another series of studies have isolated different circulating white blood cells like monocytes or neutrophils to evaluate their differential response to different conditions [29, 59-62].

6. ICM-MS proteomic analysis in animal sciences

ICM-MS is a promising diagnostic tool, showing several advantages such as being cost-effective, simple and showing a high throughput. As it is based on the comparison of peptide and small proteins profiles rather than the comparison of single biomarkers, it is theoretically more prone to identify phenotypic differences associated with a certain physiological, pathological or experimental condition [5]. Bearing that in mind, multiple applications of ICM-MS in animal sciences could be foreseen. As it could be easily applied in cultured cells, it might be useful in screening studies in toxicology and infectious diseases, thus helping evaluate in a fast way the effect of single and combined substances (toxic compounds, pathogens, immuno-reactive substances, etc) in order to establish similarities, for instance, in their mechanisms of action, evaluating the nature of their response, etc. There are several examples in the literature describing how ICM-MS can be easily implemented in cellular toxicology tests. For instance, it has been shown in a fish cell model that proteome changes induced by copper sulfate, acridine, and β-naphthoflavone were reflected in the

316 ICM-MS profiles [30]. Similar results were obtained in human cell lines treated with hydrogen
317 peroxide and aflatoxin B1 [26].

318 The use of ICM-MS for disease diagnostics in animals could be implemented in blood
319 circulating cells, where the presence of pathognomonic ICM-MS spectra patterns could be
320 researched. This could be particularly interesting in the course of infections often leading to sub-
321 clinical symptoms (e.g. paratuberculosis), and/or those involving intracellular pathogens in blood
322 cells (ehlichiosis, anaplasmosis, babesiosis etc.). Several studies carried out in human samples have
323 actually evidenced that ICM-MS analysis of human monocytes evidenced proteome changes
324 occurring upon microbial stimulation [29]. ICM-MS spectra from M1 and M2 macrophage subtypes
325 were also significantly different , as well as those obtained from macrophages in contact with
326 extracellular (group B Streptococcus, *Staphylococcus aureus*) and intracellular (*Mycobacterium bovis,*
327 *Orientia tsutsugamushi, Coxiella burnetii*) bacteria [60]. The ability of ICM-MS-based analysis to
328 detect molecular differences linked to specific phenotypes has allowed the discovery of several
329 markers. For instance, whole cell MALDI profiling analysis of peripheral chicken cells evidenced that
330 thymosin beta-4 was an abundant peptide present in avian monocytes/macrophages (and not in
331 granulocytes) and that it was regulated by bacterial lipopolysaccharides and peptidoglycans [63]. The
332 same approach helped to put in evidence that β-defensin 2 is an important anti-microbial peptide in
333 chicken and turkey heterophils [63].

334 However, clinical research in animal science is not only focused in finding disease biomarkers,
335 but also in finding molecular tools that could help in animal production. Reproduction management
336 represents one of the most important aspects of animal production, and very few molecular tools are
337 available to monitor the implementation of new management methods in farm animals. In this
338 sense, ICM-MS has been proven useful to follow cell-specific peptidome/proteome changes in the
339 cells related to the animal reproductive system. In the following sections, we will discuss how ICM-
340 MS profiling has been implemented for the analysis of male and female gametes in farm animals.

341

7. ICM-MS proteomic analysis of male gametes

Evaluation of sperm quality is usually based on the assessment of individual sperm parameters (motility, morphology), which is cumbersome and sometimes not well correlated with the actual fertilizing ability of tested sperm. Thus, the development of a fast, accurate tool for fertility screening is largely needed in human and veterinary medicine. From all possible molecular biology techniques applicable, proteomics seems to be one of the most adequate to study sperm biology. Numerous proteomic studies aiming to identify sperm markers of fertility were recently performed in several mammalian species (reviewed in [64-66]). However these approaches are not directly applicable for their routine use and the use of individual biomarkers has not been yet successfully implemented in fertility clinics to stratify sperm according to their quality. Sperm peptide and small protein profiling through ICM-MS has been recently proven useful to identify proteome patterns linked with fertility in chicken [45, 67]. The method included only a few and simple preparatory steps namely cell counting, washing sperm cells to eliminate seminal plasma proteins, and resuspension in the same buffer at a known concentration. The plating procedure was simple: sperm suspension was spotted onto a MALDI plate, overlaid with an adequate matrix and allowed to dry. Further MS analysis can be automated, this increasing the throughput of the method. Using adequate pre-processing and bioinformatics analysis tools, sperm ICM-MS profiles showed characteristic spectral features that allowed their classification according to their actual fertilizing ability [67]. This method was successfully employed in different chicken genetic lines, as well in other livestock species, thus showing a high versatility. Moreover, several peaks of chicken sperm ICM-MS spectra corresponding to peptidoforms (due to specific or non specific proteolytic activities) or proteoforms were identified using HRMS, including some fertility biomarkers candidates, such as protein vitelline membrane outer layer protein 1 or different forms of the avian beta-defensin 10. It has been also shown that the ICM-MS profiles from pig epididymal spermatozoa reflect the proteome changes that they undergo during epididymal maturation, mostly those related with a differential activity of sperm intracellular proteases and peptidases [68]. Moreover, it was observed

368 that the analysis of whole cells allowed a better discrimination of different maturation stages
369 compared with sperm cell extracts as obtained with neutral detergents [68]. In short, these results
370 indicate that ICM-MS profiling could be a suitable approach for molecular phenotyping of male
371 gametes.
372
373 **8. ICM-MS proteomic analysis of female gametes**
374 Today, it is possible to initiate early embryonic development of just about any mammalian
375 species by co-incubating oocyte with spermatozoa following a protocol known as in vitro fertilization
376 (IVF). The original process includes the administration of high levels of gonadotropins to stimulate
377 oocyte maturation and ovulation [69]. Oocyte maturation is a transition from immature germinal
378 vesicle (GV) transcriptionally silenced oocyte to metaphase-II stage (Meta-II) which is accompanied
379 by polar body extrusion, reorganization of the organelles through the ooplasm and molecular
380 modifications including changes in protein abundance and post-translational modifications. Oocyte
381 maturation plays an important role in the acquisition of oocyte developmental competence, referred
382 as oocyte quality, which, at the same time, is crucial for embryo quality [70]. Maturation can be
383 accomplished in vitro starting from the oocytes recovered from small antral follicles at the stage prior
384 to selection and dominance, by using specific culture media, following a protocol known as *in vitro*
385 *maturation* (IVM). IVM requires little or no use of in vivo-administered gonadotropins and has been
386 proposed as an alternative to conventional IVF with ovulated oocytes in humans since it reduces the
387 primary adverse effects caused by controlled ovarian stimulation, including the ovarian
388 hyperstimulation syndrome [71]. IVM is actually a routine technique in animal reproduction
389 biotechnologies, used especially in dairy cows of high genetic merit, where the success rate after
390 artificial insemination is often low. In cattle, immature oocytes can be obtained from the ovaries
391 either by ovum pick up or after slaughter, then subjected to IVM, fertilized and developed to
392 transferable embryos. This technique of *in vitro* embryo production allows the use of less semen (this
393 is important in the case of expensive and rare sperm), as well as increasing the progeny of oocyte

394 donors [72]. However, proteomic changes that occur in the oocyte during maturation and that define
395 the quality of *in vitro* matured oocytes are mostly unknown, and so there is no accurate way of
396 evaluating/monitoring how different IVM protocols can affect the process. Classic proteomics has
397 been employed in the past with this objective. In fact, by using 2D-DIGE or 2DE coupled with nano-
398 liquid chromatography and high resolution tandem MS, proteomes were established for bovine,
399 porcine and murine oocytes identifying numerous proteins including putative markers of oocyte
400 developmental competence including proteins needed for fertilization, reprogramming, embryo
401 genome activation and first cleavages during early embryogenesis [73-75]. However, these proteomic
402 approaches require a high amount of oocytes to be performed. In cows, only 8-10 oocytes fully-
403 grown oocytes can be collected per ovary, so the amount of material is limited. The bovine oocyte is
404 about 120 µm in diameter, but it contains a relatively low intra-cellular protein content, about 80-
405 100 ng per oocyte. ICM-MS was recently adapted to bovine ovarian cells from the individual antral
406 follicles to obtain the protein/peptide signatures (<17kDa) of single oocytes, surrounding cumulus
407 cells and granulosa cells [76]. It was demonstrated, that this technology could be applied to single
408 bovine oocytes, and that single-oocyte ICM-MS spectra are characteristic of each stage of oocyte
409 meiotic maturation from prophase-I in immature oocytes to metaphase-II in mature ones [76]. In
410 this study, ICM-MS coupled to an optimized Top-down HR-MS proteomic approach on ovarian
411 follicular cells allowed the identification of specific markers of oocyte meiotic maturation including
412 IGF2 binding protein 3 and hemoglobin B in the oocyte and thymosins beta-4, and beta-10, histone
413 H2B and ubiquitin in surrounding cumulus cells.
414

415 **9. Conclusions**

416 ICM-MS is a peptide/protein profiling platform widely used for phyloproteomics that has been
417 recently revealed as a promising tool for molecular phenotyping in clinical proteomics. The
418 applicability of this technique has been already shown in clinical microbiology and it is starting to
419 bring some attention in biomedicine. Indeed, specific characteristic features present in ICM-MS

profiles in a given condition can be considered as biomarker profiles, and individual differential features can be identified following Top-Down proteomics. Although still not widely used, ICM-MS profiling can have multiple applications in animal research, ranging from basic sciences - as for example, in newly established primary cell lines phenotyping or in vitro toxicological studies - to applied sciences for use in disease diagnostics, male fertility classification according to semen profiling or monitoring of oocyte quality, among others.

Acknowledgements

This work was supported by the French National Infrastructure of Research CRB anim funded by "Investissements d'avenir", ANR-11-INBS-0003 and from the French National Institute of Agronomic Research. The high-resolution mass spectrometer was financed (SMHART project) by the European Regional Development Fund (ERDF), the Conseil Régional du Centre, the French National Institute for Agricultural Research (INRA) and the French National Institute of Health and Medical Research (Inserm). Laura Soler has received the support of the EU in the framework of the Marie-Curie FP7 COFUND People Programme, through the award of an AgreenSkills fellowship (under grant agreement n° 267196).

References

[1] V. Thongboonkerd. Clinical proteomics: Towards diagnostics and prognostics.Blood, 2007, 109, 5075–5076.

[2] H. Mischak, G. Allmaier, R. Apweiler, T. Attwood, M. Baumann, A. Benigni, S. E. Bennett, R. Bischoff, E. Bongcam-Rudloff, G. Capasso, J. J. Coon, P. D'Haese, A. F. Dominiczak, M. Dakna, H. Dihazi, J. H. Ehrich, P. Fernandez-Llama, D. Fliser, J. Frokiaer, J. Garin, M. Girolami, W. S. Hancock, M. Haubitz, D. Hochstrasser, R. R. Holman, J. P. A. Ioannidis, J. Jankowski, B. A. Julian, J. B. Klein, W. Kolch, T. Luider, Z. Massy, W. B. Mattes, F. Molina, B. Monsarrat, J. Novak, K. Peter, P. Rossing, M. Sánchez-Carbayo, J. P. Schanstra, O. J. Semmes, G. Spasovski, D. Theodorescu, V. Thongboonkerd, R. Vanholder, T. D. Veenstra, E. Weissinger, T. Yamamoto and A. Vlahou. Recommendations for biomarker identification and qualification in clinical proteomics. Sci. Transl. Med., 2010, 2, 46ps42.

[3] F. Ceciliani, D. Eckersall, R. Burchmore and C. Lecchi. Proteomics in veterinary medicine: applications and trends in disease pathogenesis and diagnostics. Vet. Pathol., 2014, 51, 351–62.

[4] P. D. Eckersall, A. M. de Almeida and I. Miller, J. Proteomics, a new tool for farm animal science. Proteomics, 2012, 75, 4187–9.

[5] E. Bendixen, M. Danielsen, K. Hollung, E. Gianazza and I. Miller, J. Farm animal proteomics--a review. Proteomics, 2011, 74, 282–93.

[6] A. M. de Almeida and E. Bendixen, J. Pig proteomics: a review of a species in the crossroad between biomedical and food sciences. Proteomics, 2012, 75, 4296–314.

[7] J. L. Williams. Genetic Control of Meat Quality Traits, in Meat Biotechnology, ed. F. Toldrá, Springer New York, New York, NY, 2008, pp. 21–60.

[8] W. H. McDonald and J. R. Yates. Shotgun proteomics and biomarker discovery. Dis. Markers, 2002, 18, 99–105.

[9] Mann M, Hojrup P, Roepstorff P. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. Biol Mass Spectrom. 1993 Jun; 22(6):338-45.

[10] Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003; 422:198–207.

[11] B. Hernández, A. Parnell and S. R. Pennington. Why have so few proteomic biomarkers "survived" validation?(Sample size and independent validation considerations). Proteomics, 2014, 14, 1587–92.

[12] T. Shi, D. Su, T. Liu, K. Tang, D. G. Camp, W.-J. Qian and R. D. Smith. Advancing the sensitivity of selected reaction monitoring-based targeted quantitative proteomics. Proteomics, 2012, 12, 1074–92.

[13] H. Mischak, A. Vlahou and J. P. A. Ioannidis. Technical aspects and inter-laboratory variability in native peptide profiling: the CE–MS experience. Clin. Biochem., 2013, 46, 432–43.

[14] J. P. Savaryn, A. D. Catherman, P. M. Thomas, M. M. Abecassis and N. L. Kelleher. The emergence of top-down proteomics in clinical research. Genome Med., 2013, 5, 53.

[15] Kelleher NL. Top-down proteomics. Anal Chem. 2004; 76:196 A–203 A.

[16] A. Albalat, H. Husi, J. Siwy, J. E. Nally, M. McLauglin, P. D. Eckersall and W. Mullen. Capillary electrophoresis interfaced with a mass spectrometer (CE-MS): technical considerations and applicability for biomarker studies in animals. Curr. Protein Pept. Sci., 2014, 15, 23–35.

[17] M. G. J. Koene, H. A. Mulder, N. Stockhofe-Zurwieden, L. Kruijt and M. A. Smits. Serum protein profiles as potential biomarkers for infectious disease status in pigs. BMC Vet. Res., 2012, 8, 32.

[18] L. Zhong, D. L. Taylor and R. J. Whittington. Proteomic profiling of ovine serum by SELDI-TOF MS: optimisation, reproducibility and feasibility of biomarker discovery using routinely collected samples. Comp. Immunol. Microbiol. Infect. Dis., 2010, 33, 47–63.

[19] A. Stalmach, H. Husi, K. Mosbahi, A. Albalat, W. Mullen and H. Mischak. Methods in capillary electrophoresis coupled to mass spectrometry for the identification of clinical proteomic/peptidomic biomarkers in biofluids. Methods Mol. Biol., 2015, 1243, 187–205.

[20] L. Molin, R. Seraglia, A. Lapolla, E. Ragazzi, J. Gonzalez, A. Vlahou, J. P. Schanstra, A. Albalat, M. Dakna, J. Siwy, J. Jankowski, V. Bitsika, H. Mischak, P. Zürbig and P. Traldi, A comparison between MALDI-MS and CE-MS data for biomarker assessment in chronic kidney diseases. J. Proteomics, 2012, 75, 5888–97.

[21] R. Mansor, W. Mullen, A. Albalat, P. Zerefos, H. Mischak, D. C. Barrett, A. Biggs and P. D. Eckersall. A peptidomic approach to biomarker discovery for bovine mastitis. J. Proteomics, 2013, 85, 89–98.

[22] J. E. Nally, W. Mullen, J. J. Callanan, H. Mischak and A. Albala. Detection of urinary biomarkers in reservoir hosts of leptospirosis by capillary electrophoresis-mass spectrometry. Proteomics. Clin. Appl., 2015, 9, 543–51.

[23] A. Mehta and L. P. Silva. MALDI-TOF MS profiling approach: how much can we get from it? Front. Plant Sci., 2015, 6, 184.

[24] J. F. Povey, C. J. O'Malley, T. Root, E. B. Martin, G. A. Montague, M. Feary, C. Trim, D. A. Lang, R. Alldread, A. J. Racher and C. M. Smales. Rapid high-throughput characterisation, classification and selection of recombinant mammalian cell line phenotypes using intact cell MALDI-ToF mass spectrometry fingerprinting and PLS-DA modelling. J. Biotechnol., 2014, 184, 84–93.

[25] V. Labas, L. Spina, C. Belleannee, A.-P. Teixeira-Gomes, A. Gargaros, F. Dacheux and J.-L. Dacheux. Analysis of epididymal sperm maturation by MALDI profiling and top-down mass spectrometry. J. Proteomics, 2015, 113, 226–43.

[26] N. H. L. Chiu, Z. Jia, R. Diaz and P. Wright. Rapid Differentiation of In Vitro Cellular Responses to Toxic Chemicals by Using MALDI-TOF Mass Spectrometry. Environ. Toxicol. Chem., 2015, 34, 161–6.

[27] J. Hanrieder, G. Wicher, J. Bergquist, M. Andersson and A. Fex-Svenningsen. MALDI mass spectrometry based molecular phenotyping of CNS glial cells for prediction in mammalian brain tissue. Anal. Bioanal. Chem., 2011, 401, 135–47.

[28] G. M. Toh-Boyo, S. S. Wulff and F. Basile. Comparison of sample preparation methods and evaluation of intra- and intersample reproducibility in bacteria MALDI-MS profiling. Anal. Chem., 2012, 84, 9971–80.

[29] D. Portevin, V. Pfluger, P. Otieno, R. Brunisholz, G. Vogel and C. Daubenberger. Quantitative whole-cell MALDI-TOF MS fingerprints distinguishes human monocyte sub-populations activated by distinct microbial ligands. BMC Biotechnol, 2015, 15, 24.

[30] S. L. Kober, H. Meyer-Alert, D. Grienitz, H. Hollert and M. Frohme. Intact cell mass spectrometry as a rapid and specific tool for the differentiation of toxic effects in cell-based ecotoxicological test systems. Anal. Bioanal. Chem., 2015, 407, 7721–31.

[31] B. Munteanu, B. Meyer, C. von Reitzenstein, E. Burgermeister, S. Bog, A. Pahl, M. P. Ebert and C. Hopf. Label-free in situ monitoring of histone deacetylase drug target engagement by matrix-assisted laser desorption ionization-mass spectrometry biotyping and imaging. Anal. Chem., 2014, 86, 4642–7.

[32] G. Vogel, A. Strauss, B. Jenni, D. Ziegler, E. Dumermuth, S. Antz, C. Bardouille, B. Wipf, C. Miscenic, G. Schmid and V. Pflüger. Development and validation of a protocol for cell line identification by MALDI-TOF MS. BMC Proc., 2011, 5 Suppl 8, P45.

[33] H. Feng, L. C. Sim, C. Wan, N. S. C. Wong and Y. Yang. Rapid characterization of protein productivity and production stability of CHO cells by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom., 2011, 25, 1407–1412.

[34] R. Ouedraogo, A. Daumas, E. Ghigo, C. Capo, J. L. Mege and J. Textoris. Whole-cell MALDI-TOF MS: a new tool to assess the multifaceted activation of macrophages. J Proteomics, 2012, 75, 5523–5532.

[35] A. Karger, B. Bettin, M. Lenk and T. C. Mettenleiter. Rapid characterisation of cell cultures by matrix-assisted laser desorption/ionisation mass spectrometric typing. J. Virol. Methods, 2010, 164, 116–21.

[36] L. Bai, E. V Romanova and J. V Sweedler. Distinguishing endogenous D-amino acid-containing neuropeptides in individual neurons using tandem mass spectrometry. Anal Chem, 2011, 83, 2794–2800.

538 [37] H. López-Fernández, H. M. Santos, J. L. Capelo, F. Fdez-Riverola, D. Glez-Peña and M. Reboiro-
539 Jato. Mass-Up: an all-in-one open software application for MALDI-TOF mass spectrometry knowledge
540 discovery. BMC Bioinformatics, 2015, 16, 318.

541 [38] P. Findeisen and M. Neumaier. Mass spectrometry-based clinical proteomics profiling:
542 current status and future directions. Expert Rev. Proteomics, 2009, 6, 457–9.

543 [39] D. R. Ahlf, P. M. Thomas and N. L. Kelleher. Developing top down proteomics to maximize
544 proteome and sequence coverage from cells and tissues. Curr. Opin. Chem. Biol., 2013, 17, 787–794.

545 [40] P. Chaurand and R. M. Caprioli. Direct profiling and imaging of peptides and proteins from
546 mammalian cells and tissue sections by mass spectrometry. Electrophoresis, 2002, 23, 3125–3135.

547 [41] D. Debois, V. Bertrand, L. Quinton, M.-C. De Pauw-Gillet and E. De Pauw. MALDI-in source
548 decay applied to mass spectrometry imaging: a new tool for protein identification. Anal. Chem.,
549 2010, 82, 4036–45.

550 [42] D. Suckau and A. Resemann. MALDI Top-Down sequencing: calling N- and C-terminal protein
551 sequences with high confidence and speed. J. Biomol. Tech., 2009, 20, 258–62.

552 [43] Z. Liu and K. L. Schey. Optimization of a MALDI TOF-TOF mass spectrometer for intact protein
553 analysis. J. Am. Soc. Mass Spectrom., 2005, 16, 482–90.

554 [44] B. Parks, L. Jiang, P. M. Thomas, C. D. Wenger, M. J. Roth, M. T. Boyne, P. V Burke, K. E.
555 Kwast and N. L. Kelleher. Top-down proteomics on a chromatographic time scale using linear ion trap
556 fourier transform hybrid mass spectrometers. Anal. Chem., 2007, 79, 7984–91.

557 [45] V. Labas, I. Grasseau, K. Cahier, A. Gargaros, G. Harichaux, A.-P. Teixeira-Gomes, S. Alves, M.
558 Bourin, N. Gérard and E. Blesbois. Qualitative and quantitative peptidomic and proteomic
559 approaches to phenotyping chicken semen. J. Proteomics, 2015, 112, 313–35.

560 [46] R. D. Holland, J. G. Wilkes, F. Rafii, J. B. Sutherland, C. C. Persons, K. J. Voorhees and J. O. Lay.
561 Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser
562 desorption/ionization with time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom., 1996,
563 10, 1227–32.

564 [47] M. Welker. Proteomics for routine identification of microorganisms. Proteomics, 2011, 11,
565 3143–3153.

566 [48] W. P. Blackstock and M. P. Weir. Proteomics: quantitative and physical mapping of cellular
567 proteins. Trends Biotechnol., 1999, 17, 121–7.

568 [49] M. A. Claydon, S. N. Davey, V. Edwards-Jones and D. B. Gordon. The rapid identification of
569 intact microorganisms using mass spectrometry. Nat. Biotechnol., 1996, 14, 1584–6.

570 [50] T. Krishnamurthy and P. L. Ross. Rapid identification of bacteria by direct matrix-assisted
571 laser desorption/ionization mass spectrometric analysis of whole cells. Rapid Commun. Mass
572 Spectrom., 1996, 10, 1992–1996.

573 [51] C. J. Keys, D. J. Dare, H. Sutton, G. Wells, M. Lunt, T. McKenna, M. McDowall and H. N. Shah,
574 Compilation of a MALDI-TOF mass spectral database for the rapid screening and characterisation of
575 bacteria implicated in human infectious diseases. Infect Genet Evol, 2004, 4, 221–242.

576 [52] K. Sogawa, M. Watanabe, K. Sato, S. Segawa, C. Ishii, A. Miyabe, S. Murata, T. Saito and F.
577 Nomura. Use of the MALDI BioTyper system with MALDI–TOF mass spectrometry for rapid
578 identification of microorganisms. Anal Bioanal Chem, 2011, 400, 1905–1911.

579 [53] A. Zavalin, E. M. Todd, P. D. Rawhouser, J. Yang, J. L. Norris and R. M. Caprioli. Direct imaging
580 of single cells and tissue at sub-cellular spatial resolution using transmission geometry MALDI MS. J.
581 Mass Spectrom., 2012, 47, i.

582 [54] J. Qian, J. E. Cutler, R. B. Cole and Y. Cai. MALDI-TOF mass signatures for differentiation of
583 yeast species, strain grouping and monitoring of morphogenesis markers. Anal Bioanal Chem, 2008,
584 392, 439–449.

585 [55] L. G. Stevenson, S. K. Drake, Y. R. Shea, A. M. Zelazny and P. R. Murray. Evaluation of matrix-
586 assisted laser desorption ionization-time of flight mass spectrometry for identification of clinically
587 important yeast species. J Clin Microbiol, 2010, 48, 3482–3486.

588 [56] N. Fatania, M. Fraser, M. Savage, J. Hart and A. Abdolrasouli. Comparative evaluation of
589 matrix-assisted laser desorption ionisation-time of flight mass spectrometry and conventional
590 phenotypic-based methods for identification of clinically important yeasts in a UK-based medical
591 microbiology laboratory. J. Clin. Pathol., 2015, 68, 1040–2.

592 [57] J. Gagnaire, O. Dauwalder, S. Boisset, D. Khau, A.-M. Freydiere, F. Ader, M. Bes, G. Lina, A.
593 Tristan, M.-E. Reverdy, A. Marchand, T. Geissmann, Y. Benito, G. Durand, J.-P. Charrier, J. Etienne, M.
594 Welker, A. Van Belkum and F. Vandenesch, Detection of Staphylococcus aureus delta-toxin
595 production by whole-cell MALDI-TOF mass spectrometry. PLoS One, 2012, 7, e40660.

596 [58]. Povey JF, Saintas E, Aderemi AV, Rothweiler F, Zehner R, Dirks WG, Cinatl J, Racher AJ, Wass
597 MN, Smales CM, Michaelis M. Intact-Cell MALDI-ToF Mass Spectrometry for the Authentication of
598 Drug-Adapted Cancer Cell Lines. Cells. 2019 ,2;8(10)

599 [59] R. Ouedraogo, C. Flaudrops, A. Ben Amara, C. Capo, D. Raoult and J.-L. Mege. Global analysis
600 of circulating immune cells by matrix-assisted laser desorption ionization time-of-flight mass
601 spectrometry. PLoS One, 2010, 5, e13691.

602 [60]. Ouedraogo R, Textoris J, Gorvel L, Daumas A, Capo C, Mege JL. Analysis and Characterization of
603 Immune Cells and Their Activation Status by Whole-Cell MALDI-TOF Mass Spectrometry. Methods
604 Mol Biol. 2019; 2024:339-351

605 [61] L. Kannan, N. C. Rath, R. Liyanage and J. O. Lay Jr. Direct screening identifies mature β-
606 defensin 2 in avian heterophils. Poult Sci, 2009, 88, 372–379.

607 [62] L. Kannan, N. C. Rath, R. Liyanage and J. O. Lay Jr.. Thymosin beta in macrophage. Ann. N. Y.
608 Acad. Sci., 2007, 1112, 425–434.

609 [63] L. Kannan, R. Liyanage, J. O. Lay Jr. and N. C. Rath. Evaluation of beta defensin 2 production
610 by chicken heterophils using direct MALDI mass spectrometry. Mol Immunol, 2009, 46, 3151–3156.

611 [64] Druart, X., Rickard, J.P., Tsikis, G., de Graaf, S.P., 2019. Seminal plasma proteins as markers of
612 sperm fertility. Theriogenology 31, 30172-30174.

613 [65] Dipresa, S., De Toni, L., Foresta, C., Garolla, A., 2018. New Markers for Predicting Fertility of the
614 Male Gametes in the Post Genomic Age. Protein Pept Lett 18, 092986652566180418120635.

615 [66] Jodar, M., Soler-Ventura, A., Oliva, R., 2017. Semen proteomics and male infertility. J Proteomics
616 162, 125-134.

[67] Soler L, Labas V, Thélie A, Grasseau I, Teixeira-Gomes AP, Blesbois E. Intact cell MALDI-TOF MS on sperm: a molecular test for male fertility diagnosis. Mol Cell Proteomics. 2016, 15(6):1998-2010.

[68] V. Labas, L. Spina, C. Belleannee, A.-P. Teixeira-Gomes, A. Gargaros, F. Dacheux and J.-L. Dacheux. Analysis of epididymal sperm maturation by MALDI profiling and top-down mass spectrometry. J. Proteomics, 2015, 113, 226–43.

[69] B. G. Brackett, D. P. Wolf and M. Zelinski-Wooten. Advances in Animal In Vitro Fertilization. 2001, 21–51.

[70] G. Marteil, L. Richard-Parpaillon and J. Z. Kubiak. Role of oocyte quality in meiotic maturation and embryonic development. Reprod. Biol., 2009, 9, 203–24.

[71] E. M. Chang, H. S. Song, D. R. Lee, W. S. Lee and T. K. Yoon. In vitro maturation of human oocytes: Its role in infertility treatment and new possibilities. Clin. Exp. Reprod. Med., 2014, 41, 41–6.

[72] L. Zhang, R. S. Denniston and R. A. Godke. A simple method for in vitro maturation, in vitro fertilization, and co-culture of bovine oocytes. J. Tissue Cult. Methods, 1992, 14, 107–112.

[73] G. J. Arnold and T. Frohlich. Dynamic proteome signatures in gametes, embryos and their maternal environment. Reprod. Fertil. Dev. 2011, 23, 81–93.

[74] I. Virant-Klun and J. Krijgsveld. Proteomes of animal oocytes: what can we learn for human oocytes in the in vitro fertilization programme? Biomed Res. Int., 2014, 2014, 856907.

[75] Q. Fu, Z.-F. Liu, Y.-L. Huang, Y.-Q. Lu and M. Zhang. Comparative proteomic analysis of mature and immature oocytes of the swamp buffalo (Bubalus bubalis). Int. J. Mol. Sci., 2015, 17.

[76]. Labas V, Teixeira-Gomes AP, Bouguereau L, Gargaros A, Spina L, Marestaing A, Uzbekova S. Intact cell MALDI-TOF mass spectrometry on single bovine oocyte and follicular cells combined with top-down proteomics: A novel approach to characterise markers of oocyte maturation. J Proteomics. 2018 . 20;175:56-74.

[77] L. F. Marvin-Guy, P. Duncan, S. Wagniere, N. Antille, N. Porta, M. Affolter and M. Kussmann. Rapid identification of differentiation markers from whole epithelial cells by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry and statistical analysis. Rapid Commun. Mass Spectrom., 2008, 22, 1099–1108.

[78] L. Kannan, N. C. Rath, R. Liyanage and J. O. Lay Jr.. Thymosin beta in macrophage. Ann. N. Y. Acad. Sci., 2007, 1112, 425–434.

[79] C. M. Buchanan, A. S. Malik and G. J. S. Cooper. C. M. Buchanan, A. S. Malik and G. J. S. Cooper, Rapid Commun. Mass Spectrom., 2007, 21, 3452–3458.

[80] X. Zhang, M. Scalf, T. W. Berggren, M. S. Westphall and L. M. Smith. Identification of mammalian cell lines using MALDI-TOF and LC-ESI-MS/MS mass spectrometry. J Am Soc Mass Spectrom, 2006, 17, 490–499.

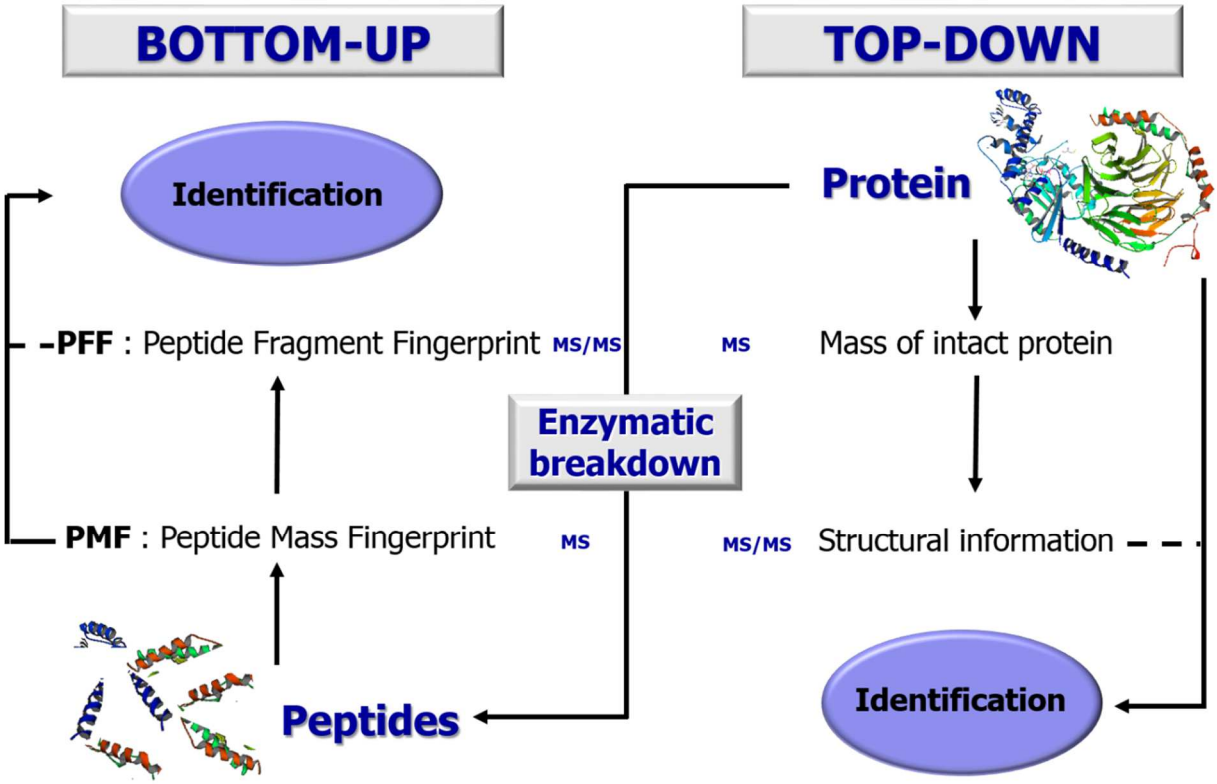
[81] J. Bergquist. Cells on the target matrix-assisted laser-desorption/ionization time-of-flight mass-spectrometric analysis of mammalian cells grown on the target. Chromatographia 1999, 49, 41–48.

Figure Legends

Figure 1: Proteomic approaches distinguishing traditional bottom-up strategies and no-classical top-down strategy

Figure 2: Typical ICM-MS workflow

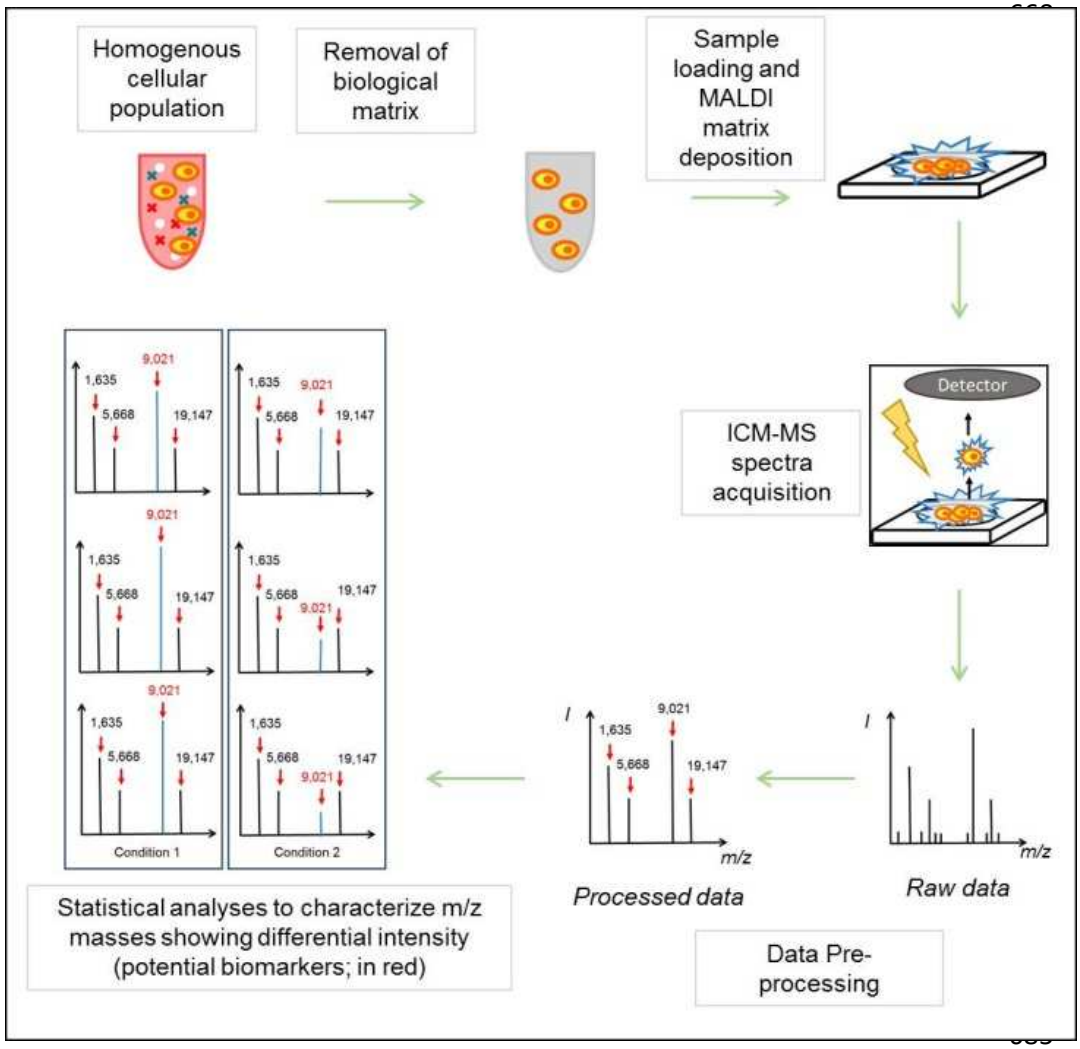
Figure 3: Top-Down protein identification workflow adapted for ICM-MS spectra characterization. 1. Fractionation can be done using chromatographic separation, gel separation or any other sample treatment aiming at decreasing sample complexity. 2. Top-Down protein identification can be performed from whole cells, crude protein extracts, or pre-fractionated protein extracts. Analysis can be based on the use of MALDI (A,B) or ESI (C,D) source ionisation m and fragmentation can be performed using PSD/ISD (A,B), or HCD (B,C,D).



665

666

667 Figure 1.



686 Figure 2

687

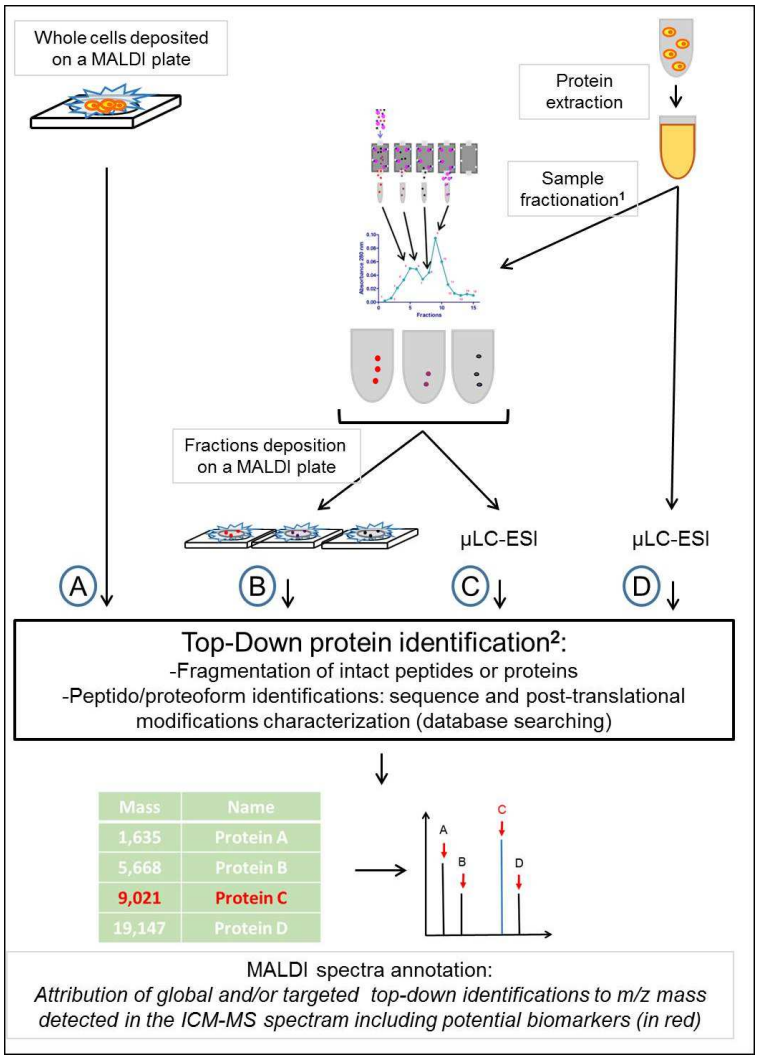

Whole cells deposited on a MALDI plate
Protein extraction
Sample fractionation[1]
Fractions deposition on a MALDI plate
µLC-ESI
µLC-ESI
A
B
C
D
Top-Down protein identification[2]:
-Fragmentation of intact peptides or proteins
-Peptido/proteoform identifications: sequence and post-translational modifications characterization (database searching)
Mass
Name
1,635
Protein A
5,668
Protein B
9,021
Protein C
19,147
Protein D
MALDI spectra annotation:
Attribution of global and/or targeted top-down identifications to m/z mass detected in the ICM-MS spectram including potential biomarkers (in red)

Figure 3

Year	Model	Objectives	Proteomic approach	m/z mass identification approach	Reference
2019	Neuroblastoma cell lines	Authentication of Drug-Adapted Cancer Cell Lines.	Targeted	no	58
2019	macrophages	Characterization of activation of macrophages	Targeted	no	60
2018	Bovine oocytes	Impact of meiotic maturation on single oocytes, cumulus cells and granulosa cells.	Global	Top-Down HRMS	76
2016	Chicken spermatozoa	Diagnosis of avian male fertility	Global	Top-Down HRMS	67
2015	Chicken spermatozoa	Phenotyping of semen to discriminate avian males on their reproductive capacity	Global	Top-Down HRMS	45
2015	Porcine Spermatozoa	Characterization of epididymal sperm maturation process	Global	Top-Down HRMS	68
2015	Human monocytes	Distinction of human monocyte subpopulations and monitorization of activation profiles of monocytes exposed to distinct microbial ligands	Global	No	29
2015	Fish cell line RTL-W1	Characterization of toxic effects and identification of toxins in complex environmental samples	Global	No	30
2015	HepG2 and THP-1 cells	Characterization of *in vitro* cellular responses to toxic chemicals	Global	No	26
2014	Cancer cells (chronic myelogenous leukemia cell line K562)	Label-free quantification of histone acetylation to measure the tumor-selective pharmacodynamic responses to drug in a mouse model of gastrointestinal cancer	Targeted	No	31
2012	Human macrophages	Monitorization of the various activation patterns of immune cells (macrophage)	Global	No	34
2012	Human monocytes and polymorphonuclear neutrophils	Phenotyping of mammalian primary cells and cell lines (HL-60, THP-1 and Hela S3)	Global	No	62
2011	Mammalian and insect cell lines	Cell line identification	Global	No	[29]
2011	Rodent glial cells	Brain cells phenotyping (astroglial, microglial, and oligodendroglial cells)	Global	Bottom-up using one-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis followed by LC-MS/MS	27
2011	Chinese hamster ovary cells	Selection of stable and high-producing cell lines for monoclonal antibody (mAb) IgG	Global	No	[33]
2011	Sea slug neurons	Characterization of D-amino acid-containing peptides	Targeted	Top-Down MALDI-TOF-TOF	6

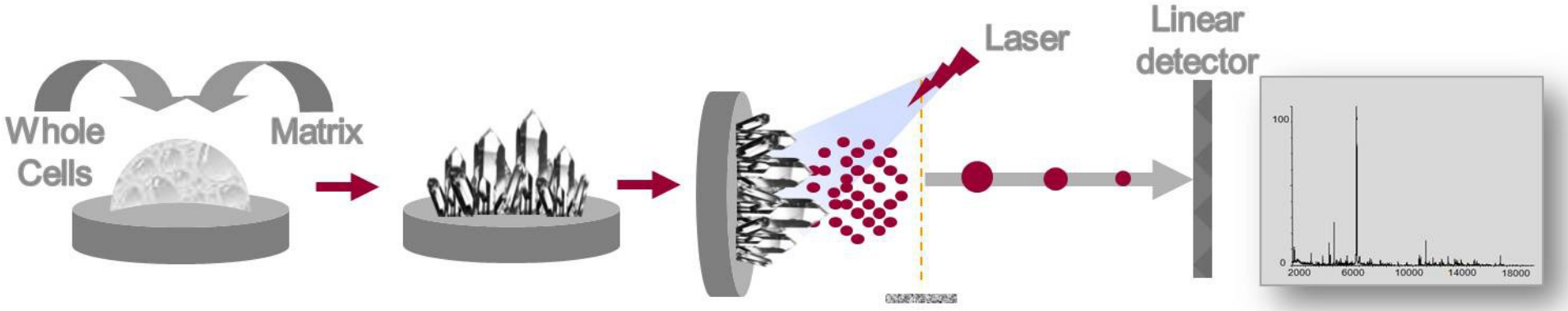
2010	Human and murine mmune cells	Immune cells (monocytes, T lymphocytes and polymorphonuclear cells) phenotyping	Global	No	58
2010	Insects and animal cells	Cell lines phyloproteomic s	Global	No	35
2010	Chinese hamster ovary cells	Characterization of low/high-producing cell lines for the production of recombinant biotherapeutics	Global	No	33
2009	Avian heterophils	Direct identification of mature β-defensin 2 in avian heterophils	Targeted	Edman sequencing and bottom-up	63
2008	Human colon epithelial carcinoma T84 cell line	Identification of cell differentiation markers	Global	Bottom-up using GeLC-MS/MS	77
2007	Chicken macrophages	Identification and characterization of thymosin beta-4	Target	Purification by HPLC and Bottom-up using MALDI-TOF	78
2007	Mouse pancreatic islet alpha and beta cells	Identification of secretory products (hormones) in endocrine cells	Global	No	79
2006	Human myelomonocytic and lymphoblast, rodent BHK-21	Identification of mammalian cell lines	Global	Top-Down HRMS, MALDI-TOF-TOF and PSD, bottom-up using nanoLC-MS/MS	80
1999	Rat adrenal pheochromocytoma PC 12 cells)	Profiling of mammalian cell line after stimulation with neurotrophic factor (nerve-growth factor)	Global	No	81

720

721 Table 1. Literature on superior eukaryotic Intact Cell MALDI-TOF MS studies.

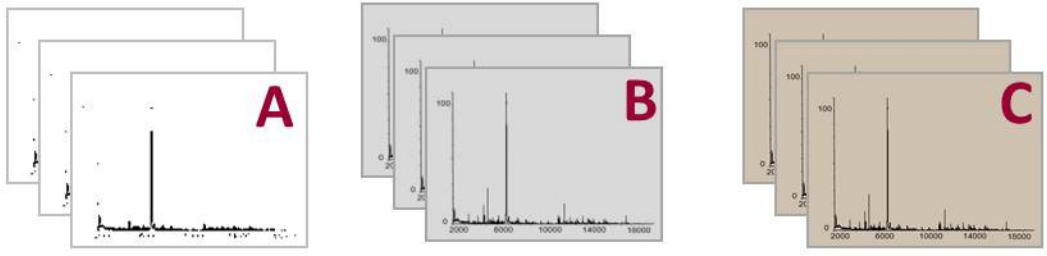
722

Intact Cell MALDI-TOF Mass Spectrometry



Co-crystallization | MALDI-TOF MS | Proteomic profile

Differential analysis



Informatic treatments

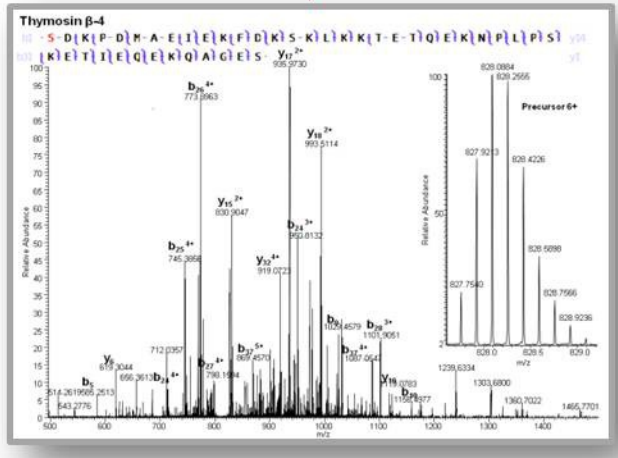
Internal calibration, Baseline subtraction, Smooth, Spectra alignment, Peak detection, Normalization

Statistical analyses for marker characterization

ANOVA, PCA, hierarchical clustering…

Top Down MS

Fragmentation of intact endogenous peptides/proteins



Sequence and PTM identifying pepti/proteoforms of interest