NanoLC-MS coupling of liquid microjunction microextraction for on-tissue proteomic analysis

Jusal Quanico, Julien Franck, Tristan Cardon, Eric Leblanc, Maxence Wiszterski, Michel Salzet, Isabelle Fournier

Université de Lille 1, Inserm U-1192, Protéomique, Réponse Inflammatoire et Spectrométrie de Masse (PRISM), F-59000 Lille, France

Université de Lille 1, Inserm U-1192, Laboratoire de Protéomique, Réponse Inflammatoire, Spectrométrie de Masse (PRISM), Department of Gynecology Oncology, Oscar Lambret Center, Lille, France

A R T I C L E   I N F O

Article history:
Received 30 June 2016
Received in revised form 2 November 2016
Accepted 4 November 2016
Available online xxxx

Keywords:
Liquid microjunction
Microextraction
Microproteomics
Imaging mass spectrometry
Ovarian cancer

A B S T R A C T

Mass spectrometry (MS)-based proteomics on localized regions of tissue sections was achieved by direct coupling of liquid microjunction microextraction with a nanoscale liquid chromatography-tandem MS, resulting in the identification of ~500 protein groups from a region as small as 250 μm in diameter representing only a few hundred of cells. The method was applied on the examination of benign and tumor regions initially defined by imaging mass spectrometry (IMS) analysis of a consecutive high grade serous ovarian tumor tissue section. Results identified the higher abundance of eukaryotic translation initiation factors eIF4A, its isoform eIF4A2, and eIF5A and its isoform eIF5A2, and lower abundance of actin-binding proteins OBSCN, TAGLN and CNN3 on tumor regions, concomitant with previous findings. This demonstrates the use of the method for downstream characterization of distinct regions identified by IMS. This article is part of a Special Issue entitled: MALDI Imaging, edited by Dr. Corinna Henkel and Prof. Peter Hoffmann.

© 2016 Elsevier B.V. All rights reserved.

1. Background

On-tissue proteomic analyses provide a direct means to examine proteomic fluctuations at the cellular level in response to changes in the tissue microenvironment. Its importance is evident in cancer, where proteomic analysis of cultured cell lines was observed to not correlate with results from microdissected cells of the same patient [1].

There is an ongoing effort to develop microscale technologies that can achieve reliable identification and quantification of proteins within an area of the most limited size, and correlate these expression changes with alterations in cell phenotypes and/or biological state. Various direct on-tissue mass spectrometry (MS)-based methods, such as matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) in profiling [2] and imaging modes [3], or desorption electrospray ionization (DESI) [4], have been shown to generate qualitative and quantitative information without a priori knowledge of the components present. But many of these techniques can only provide molecular information on low molecular weight compounds such as metabolites or lipids (DESI for example) and only MALDI gives access to information on the proteins. Despite the ability of MALDI to detect proteins that can be specifically associated to a physiological context, currently its major limitation remains the further identification of these proteins. Indeed, because the quality of direct tissue MS² spectra generated is greatly affected by the complexity of the tissue matrix, these methods suffer from limited identification. At best, these methods can provide hundreds of protein IDs, a tiny amount compared to tens or possibly hundreds of thousands of proteins expressed in cells. Methods involving the microdissection of cells, on the other hand, provide homogeneous cell populations for proteomics analysis. However, the limited number of cells in a particular region of the tissue entails microdissection in other regions to meet the requirements of extraction, thus compromising the importance of performing localized microenvironment analysis [5].
More recently, to meet the possible identification and quantification of small exogenous molecules (namely drugs) from specific tissue areas, another strategy involving microextraction on the tissue surface, was developed [6]. Initially introduced for sampling thin layer chromatography (TLC) plates [7,8], this methodology was further extended for sampling biological tissues. This allowed, for example, the direct absolute quantification of propranolol and its metabolites from mouse whole body sections by direct interface to the electrospray (ESI) source of MS instrument [9] or indirectly through high performance liquid chromatography (HPLC) under LC-MS conditions [10]. Several variants of the same concept have been designed [8] provided they are capable of dispensing and aspirating microliter quantities of extracting solvent in a controlled manner. We further applied this micro-extraction strategy for on-tissue microproteomics and developed a workflow that allows correlation of trypsin-digested protein MS images with protein IDs obtained from shotgun MS analysis of extracts taken from the same regions on adjacent tissue sections [11]. This was accomplished by performing microextraction off-line using a commercially available instrument from a micro-digested tissue area by enzyme microspoting. This allowed us to identify >1000 proteins from a tissue area of about 650 μm in diameter, corresponding to <2500 analyzed cells. This strategy was shown to be highly efficient and robust but relatively long due to sample preparation steps performed post-micro-extraction such as solvent removal, desalting and reconstitution in a solvent more suitable for injection to the LC-MS instrument.

The present work is focused on direct injection of pre-digested protein extracts from tissues into an LC-MS instrument avoiding the time-consuming post-microextraction steps and concomitant sample losses associated with the sample handling involved. We demonstrate that this approach can be applied to the downstream processing of regions of interest (ROIs) generated from matrix-assisted laser desorption/ionization imaging mass spectrometry (IMS) data to provide identification and relative quantification of hundreds to thousands of proteins from micrometer-sized areas.

2. Materials and methods

2.1. Tissue section preparation

Procedures involving the preparation of animal tissue samples were performed in accordance with the Institutional Animal Care and Use Committee (IACUC) of University of Lille 1. Fresh brain tissues were dissected from adult male Wistar rats, flash frozen in liquid nitrogen, and stored in −80 °C until use. 10-μm sections were obtained using a Cryostat (Leica Microsystems, Nanterre, France). These were mounted on indium-tin oxide coated slide for IMS experiments, and the subsequent sections for extraction mounted on polylysine-coated slides. Sections used for staining were cut at 8 μm thickness. The sections were dried under vacuum prior to both IMS and extraction experiments.

2.2. Hematoxylin-phloxine-saffron (HPS) staining

Sections were soaked for 1 min in hematoxylin, then rinsed 3 × in running water. These were then soaked in 0.1% phloxine for 10 s, followed by rinsing 2 × in running water. The sections were then soaked in 50% EtOH once and twice in 100% EtOH, then once in saffron for 5 s. Finally the sections were soaked twice in 100% EtOH and in xylene.

2.3. MALDI imaging

To dried ovarian tumor sections, 2,5-dihydrobenzoic acid (DHB) prepared at a concentration of 15 mg/mL in 70:30 methanol/0.1% trifluoroacetic acid in water was deposited by spraying with a nebulizer (an ion trap electrospray source modified in-house) for 30 min. The nebulizer is connected to a compressed air source operating at <1 bar pressure; a 500-μL syringe is also connected and delivers the matrix at a flow rate of 300 μL/h.

The images were acquired using a MALDI-TOF mass spectrometer (AutoFlex III, Bruker Daltonics, Bremen, Germany) equipped with a Smart Beam laser (Nd:YAG, 355 nm) set at a repetition rate of 2 MHz. The instrument was set to acquire in positive reflector mode, at a mass range of 300–1200 m/z, with the obtained spectra being an average of 500 shots per pixel. This results in a minimum of approximately 5000 spectra for the smallest tissue section. The raster size was set at 50 μm.

The generated images were exported into SCiLS Lab software version 2015b. Baseline removal was performed using the top-hat method and the data were normalized using the root mean square (RMS) method. The maximum peak interval was set at ±0.3 Da and peaks were detected using orthogonal matching pursuit applied on individual spectra. The m/z intervals corresponding to the peaks detected were manually validated to remove those corresponding to isotopic peaks as well as those whose widths do not correspond to the set bin width. Peak alignment was then performed by comparing with a randomly selected mean spectrum from one of the images. Automatic spatial segmentation was then performed using Hierarchical Clustering (HC). This leads to the creation of a dendrogram, with each segment pseudocolored and the colors corresponding to the pixels in the segmentation map. Clusters generated from HC were co-registered with optical images obtained from the HPS staining of the adjacent section to verify if the ROIs generated by the clusters correspond to the tumor regions defined by the pathologist.

2.4. On-tissue digestion

Picoliter quantities of 20 μg/mL sequencing grade trypsin (Promega, Charbonnieres, France) dissolved 50 mM NH₄HCO₃ were deposited on tissue sections using a CHIP 1000 instrument (Shimadzu, Japan). Depending on the experiment, parameters such as dwell voltage, dwell time, quantity of a drop, and iterative printing interval and waiting time were adjusted. The quantity of a drop, iterative printing interval and waiting time determine the size of the droplet that is produced. For example, printing using 50 μL/drop at an interval of 5 drops and waiting time of 5000 ms resulted in a droplet size of 250 μm, while a 200 μL droplet at an interval of 15 drops and waiting time of 10,000 ms resulted in a droplet size of 600 μm. Digestion was effected for 1 h, ensuring that the droplet does not dry out by replenishing with a fresh one after each waiting time, except for ovarian samples where it was performed for 2 h.

2.5. On-tissue microextraction

Extraction of the digested components was performed using the modified FAMOS autosampler set-up. 1 mL and 12 μL of 0.1% FA was placed on reservoirs A and C of the autosampler. The digested regions on the tissue section were marked and the slide was placed on the
autosampler plate. A cut 50-μL Falcon tube placed on top of the tissue section served as a support for the syringe needle during sample extraction. Alignment of the spot with the solvent droplet prior to extraction was done by adjusting the position of the glass slide. Extracts were deposited on reservoir C each time, and once all extractions have been performed, all the liquid in reservoir C was aspirated and taken to the sample loop (25 μL) of the UPLC device. The remaining volume in the sample loop is compensated for by the UPLC pump. Activation of the sample loading sequence of the UPLC was synchronized with the final extraction step of the FAMOS autosampler, whereas the orbitrap MS instruments were set to wait for contact closure before starting the MS acquisition. In the case of the ovarian tumors, the solvent for extraction was replaced with 80-20 ACN/0.1% FA in H2O.

2.6. LC-MS

For analyses performed using LTQ Orbitrap XL, the parameters were as follows. Separation of sample components was done using an online reversed-phase chromatographic system (Thermo Scientific Proxeon Easy-nLC II) equipped with a Proxeon trap column (100 μm ID × 2 cm, Thermo Scientific) and C18 packed tip column (100 μm ID × 15 cm, NikkyoTechnos Co. Ltd.). Elution was carried out using an increasing gradient of ACN (5% to 30% over 10 min, 30% to 90% over 2 min), followed by an isotropic elution at 90% over 8 min. A spray voltage of 2 kV was used to generate the precursor ions, and scans were performed with a range of 300–1600 m/z. The MS instrument was operated in data-dependent acquisition mode performing MS/MS of the top 6 most intense ions from each survey scan. The resolving power of the orbitrap mass analyzer was set to 60,000 FWHM at m/z 500 and a normalized collision energy of 35% with an isolation window of 2 amu was used for the ion trap collision-induced dissociation (CID) fragmentation of precursors whose intensities are over 500 counts. For the MS/MS, a target of 5000 ions and a maximum injection time of 200 ms were used. The dynamic exclusion was set to 30 s.

For analyses using the Q-Exactive instrument, the parameters used were as follows. Separation of sample components was done using an online reversed-phase chromatographic system (Thermo Scientific) equipped with a 75 μm × 2 cm Acclaim PepMap 100 pre-column with nanoViper fittings and a 50 μm ID × 150 mm Acclaim PepMap RS LC analytical column (Thermo Scientific). Unless specified, the peptides were eluted using a 2-h gradient of ACN starting from 5% to 50% over 120 min at a flow rate of 250 nL/min. In the case of the ovarian samples, the gradient was set at 5% to 30% ACN for 90 min, then 30% to 90% for 20 min, followed by a wash out back to 5% ACN for 15 min. The Q-Exactive instrument was set to acquire top 10 MS². The survey scans were taken at 70,000 FWHM (at m/z 400) resolving power in positive mode and using a target of 3E6 and default charge state of 2. Unassigned and +1 charge states were rejected, and dynamic exclusion was enabled for 20s. The scan range was set to 300–1600 m/z. For the MS², 1 microscan was obtained at 17,500 FWHM and isolation window of 4.0 m/z, using a scan range between 200 and 2000 m/z.

2.7. Data analysis

Raw files obtained from nanoLC-MS were processed using SEQUEST version 1.4.114 (Proteome Discoverer, Thermo Fisher Scientific, Bremen, Germany). The protein identifications were obtained using the following parameters for interrogation: parent mass tolerance: ± 10 ppm, fragment mass tolerance: 0.6 Da, max. ΔCn = 0.05, dynamic modification: oxidation/–15.995 Da (methionine), enzyme: trypsin, 2 miscleavages, protein FDR = 0.01. The searches were performed using the UniprotKB/Swiss-Prot database (accessed September 16, 2014) filtered with Rattus norvegicus (41,604 sequences) taxonomy using the SEQUEST HT algorithm. For ovarian biopsies, the database was filtered using the Homo sapiens taxonomy and combined with the alternative open reading frame (AltORF) database previously described [12]. Raw files were also converted to *.mgf using MSConver [13] and submitted to Mascot search when necessary. For the Mascot searches, the following parameters were used: Taxonomy: other mammalia, Database: SwissProt, Enzyme: trypsin, 2 miscleavages, No fixed modifications, Variable modification: oxidation (methionine), Peptide tolerance: ± 10 ppm, MS/MS tolerance: 0.6 Da, Peptide Charge: 2 +, 3 + and 4 +, Instrument: ESI-Trap and with the decoy search activated.

Scaffold version 4.3.4 (Proteome Software Inc., Portland, OR, USA) was used to validate MS/MS-based protein identifications. An X1 Tandem (version CYCLONE, 2010.12.01.1) search was performed using a concatenated version of the database used for the SEQUEST search containing 50% reversed protein sequences as decoys created using COMPASS [14]. The same parameters were used as in the SEQUEST interrogation (oxidation of methionine). Results from the two search engines were used by Scaffold to generate naive Bayes classifiers for estimating the local FDR. Using the landscape FDR feature, optimal peptide and protein FDR thresholds were obtained for each experiment. Unless specified, peptide identifications were accepted if they could be established at a given probability to achieve an FDR < 1.0%, while protein identifications were accepted if they could be established at a given probability to achieve an FDR < 1.0% and contained at least 2 identified peptides. Protein identifications were then grouped using the experiment-wide grouping strategy with protein cluster analysis. Results were then exported into Scaffold perSPECTives version 2.0.6 (Proteome Software Inc., Portland, OR, USA). Quantification was performed by spectral counting (presented as total unique peptide count), with normalization performed using the total unique peptide counts. Comparison of the total unique peptide counts between the benign and tumor regions was performed using Student’s t-test, with the FDR controlled with standard Benjamini-Hochberg procedure. The q value was set to 0.05. Clustering of the protein cluster leading terms generated from any shared evidence during experiment-wide grouping was then performed based on the number of total unique peptides to map protein distribution trends across samples. Single ranked-based Euclidean was used as the distance metric. Protein groups corresponding to benign or tumor-specific clusters were exported to the Protein Analysis Through Evolutionary Relationships (PANTHER) Classification System [15] for gene overrepresentation analysis using the Homo sapiens database [16].

3. Results

3.1. Method optimization

In this work we setup a system using an autosampler allowing for both microextraction from the tissue section and direct injection of the extract into a chromatographic system. Fig. 1 shows a Spark 920 FAMOS autosampler that was modified and interfaced with an UHPLC on-line-coupled to a Q-Exactive MS instrument. In this set-up, the syringe of the autosampler is connected to the valve S of the EASY-nLC device, and it replaces the tubing that connects the valve to the autosampler pump. A 100-μm capillary replaced the tubing in position 1 of the valve and this capillary is connected to the robotic arm that collects the sample from the autosampler tray. The FAMOS device was configured such that the “Template” option is active. Programming a series of steps in the autosampler permitting the automation of the entire liquid microjunction extraction sequence was performed using the “Mix Method” program. A sample sequence for this purpose is shown in Supplementary Table 1.

The performance of the configuration was first evaluated using digested protein samples spotted on a glass slide. First, multiple extractions of a 10 fmol bovine serum albumin (BSA) digest standard was performed. This was achieved using an LTQ orbitrap XL instrument and Mascot as search engine to identify the BSA tryptic peptides (Supplementary Fig. 1A–C). Instead of pooling all the extracts, they were individually injected and the extracted ion chromatogram (EIC)
of a major BSA peptide peak (FWGKYLYEIAR, [M + 2H]+ /m/z 722.32679) was monitored across the runs. The peptide was still observed even on the 4th extraction, and the Mascot score remains significant (i.e. 451). In this case, it can be shown that repeated extraction improves the recovery of the BSA peptides. The setup was then further tested using a standard digest of HeLa cells. Fig. 2A shows the base peak chromatogram of a rat brain extract obtained by multiple extractions of a 500-μm digested spot and using a 2-h gradient to separate the peptides. As expected, the number of protein IDs increased as a function of the elution gradient used (439 for 30 min, 1129 for 1 h or 1351 for 2 h, see Fig. 3A and Supplementary Data). Examination of the Gene Ontology (GO) terms associated with the protein identifications obtained provides an estimate of the kind of proteins extracted after on-tissue digestion. Fig. 3B shows the number of protein IDs for the 3 gradients associated with GO Cellular Component. Majority of the identifications (1197/1360 protein clusters) are associated with the cytoplasm and intracellular organelles, although a large portion (860/1360) are also associated with the cell membrane. Interestingly, a portion (467/1360) associated with the extracellular region have also been identified, showing the potential of the method for examining the cell microenvironment within the specified discrete region of interest. Also, the number of identified proteins with respect to the area sampled here is in line with the results we conventionally obtain using manual processing of the sample and injection into the nLC [11] demonstrating that direct injection without further processing of the sample does not impair the separation quality and results.

We then studied the possible reduction of size of the microdigested area to improve the spatial localization of identified proteins. Reduction of the size of the microdigested area was achieved by setting the CHIP 1000 microspotter to discrete mode, allowing the user to change the number of spots that will be microspotted per region specified (Supplementary Table 2). Fig. 4A shows an example of microdigested regions on the rat brain tissue section, showing that the diameter of microdigested spots can be decreased down to approximately 250 μm following this approach. This corresponds to about 62% reduction in microdigested region diameter, when compared to our previous studies using the liquid microjunction system with a microdigested spot size of 1125 μm.

Please cite this article as: J. Quanico, et al., NanoLC-MS coupling of liquid microjunction microextraction for on-tissue proteomic analysis, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbapap.2016.11.002
650 μm diameter [11], and about 80% reduction when enzymatic microdigestion was not performed (about 1.2 mm spot size) and lipid and native endogenous peptides were washed away.

Comparison of the number of protein IDs when the size of the digested area (respectively 600, 300 and 250 μm) was varied is shown in Fig. 4B (inset) and Supplementary Data. As expected the number of identified proteins decreases with the sampled area respectively (1031 for 600 μm, 700 for 300 μm and 515 protein groups for 250 μm). By a simple calculation, the number of cells sampled in the experiments can be estimated. Since tissue sections of 10 μm thickness are used, they can be approximated as a monolayer of cells. Then if we assume the average diameter of the cells to be 15 μm, we can estimate the numbers of cells sampled to be of about 275 cells for 250 μm, 400 cells for 300 μm and 1600 cells for 600 μm. Thus for 300 μm, or about 400 cells, 700 protein groups are identified, demonstrating the sensitivity of the method.

3.2. Application to IMS-generated ROIs

3.2.1. IMS and spatial segmentation to define ROIs

As a demonstration, the method was applied on a subset of high grade serous ovarian tumors. Ion images were initially acquired from 11 biopsies, 5 of which were grossly identified as benign and 6 as tumor without additional histopathological information, except for sample 1 where the necrotic zone in the tumor section has been clearly discerned and the tumor portion which was actually also present in the benign section has been identified. The ion images for the benign and tumor sections of sample 1 co-localize with the regions defined by HPS staining (Fig. 5). The image files were then exported in SCiLS software and spatial segmentation was performed by HC, as shown in Fig. 6. The individual segments generated by HC are pseudocolored, with the zones in green and blue associated to tumor and benign zones, respectively. It can be seen from the segmentation result that majority of the tumor biopsies display tissue heterogeneity, or that only a limited region in these biopsies actually contain tumor regions. It has to be noted also that several samples annotated as benign contain many regions with spectral profiles resembling the tumor. The red zone which is a daughter segment of the main branch separating the green zone (tumor) from the blue zone (benign) is also possibly associated with the tumor zone, although this has to be further verified. Spatial segmentation can thus be used to discriminate the tumor zones and direct on which regions microextraction should be performed.

3.2.2. Microextraction and shotgun nanoLC-MS analysis

The ROIs identified by IMS were applied on a consecutive tissue section. The sections were washed to remove lipids and the ROIs mapped and subjected to in-situ trypsin digestion. The digested areas were then extracted and the extracts subjected to nanoLC-MS/MS analysis to identify the peptides and proteins present in the ROIs. From the sets of samples used in IMS, a cohort composed of 3 benign and 3 tumor samples were subjected to microextraction. 1148 protein groups with unique accession numbers were identified upon setting both the peptide and protein FDR thresholds to 1% and using a minimum of 2 peptides to identify each protein. Among these, 70 protein groups...
showed significant differences in total unique peptide abundance using Student’s t-test with the q-value set at ≤0.05 when comparing the samples according to benign and tumor categories.

Clustering of the individual protein identifications performed using Scaffold perSPECTives software identified four distinct clusters (clusters 1–4, Fig. 7), three of which (clusters 1–3) contained overexpressed proteins in the tumor samples and 1 (cluster 4) underexpressed. The protein IDs from these clusters were subjected to GO overrepresentation analysis using PANTHER in order to identify if distinct GO annotations can be associated with each cluster and help characterize the proteins therein. The summary of the overrepresentation analysis is shown in Table 1 (for members, see Supplementary Table 3). The first cluster is associated with ribosomal proteins (RPS8, RPS23, and RPS4Y2), evidenced by the related GO term expressions (translation (GO:0006412) and structural constituent of the ribosome (GO:0003735)). Cluster 2 is associated with GO terms describing catalytic activity (GO:0003824) and translation (translation regulator activity (GO:0045182)) and translation factor activity, nucleic acid binding (GO:0008135). Cluster 3 is also associated with GO terms describing translation (GO:0006412 and translation initiation factor activity (GO:0008152)). Both clusters 2 and 3 contain proteins associated with the GO term primary metabolic process (GO:0044238). Cluster 4 on the other hand is associated with muscle contraction (GO:0003735).

4. Discussion

The present system is presented as an alternative means to generate liquid microjunction that can be easily interfaced with a nanoLC MS instrument for direct injection of extracts into the pre-column. Direct injection reduces the preparation steps performed in the previous work in order to replace the highly organic solvent used for extraction with one that is compatible with the pre-column and remove salts using a Ziptip (Fig. 8). The small volumes of solvent used for extraction afforded by using a capillary in the FAMOS autosampler reduce the amount of organic solvent and it is further diluted with aqueous solvent prior to delivery into the sample loop. We did not attempt to compare the number of peptide and protein identifications produced by either approaches directly. However, it is known that material loss is possible during subsequent drying of extracts because the peptides tend to stick to the container walls especially when using non-low-binding tubes [17]. In return, injecting samples without prior clean-up could be detrimental to pre-columns and columns. Further work needs to be done in order to eliminate such effects and improve the robustness of direct injection approaches.

In contrast to discrete droplet-type microextraction, related devices capable of continuous flow extraction, such as the liquid microjunction surface sampling probe (LMJ SSP) [18] and the sealing surface sampling probe [8], have been reported. Recently [19], it has been demonstrated that extraction using the LMJ SSP device is more efficient because the liquid junction can be maintained for extended periods of time and fresh solvent continually replenishes the probe. Its main disadvantage though is that the continuous flow implies dilution of the sample, causing the signal profiles to be maximal during the initial stages of extraction and diminish as a function of extraction time [18]. This was overcome by collecting the extracts in a sample loop prior to injection to the column. Using a 20-μm loop, the authors report extraction times up to 8 min, given a flow rate of 25 μL/min. Meanwhile, to ameliorate the extraction efficiency, multiple cycles of extraction are performed on the discrete droplet-type microextraction. Several reports [20,21] indicate that repeated cycles could lead to increase in the size of the microextracted area, and that the small size of the droplet limits the dissolution of compounds to the liquid interface.

Results of this bottom-up proteomics approach indicate that the latter is not the case, since more than a thousand protein clusters were identified from just five cycles of extraction on a sample surface as small as 600 μm. Also, increases in the size of the microextracted area are not critical because the microdigested region ultimately determines the size of the extractable area and is independently controlled. Perhaps the advantage that can be gleaned from the use of smaller microjunction areas then is that the volume of solvent used will also be small, thus, more cycles of extraction can be performed, if really desired. In the optimization experiments, up to 5 cycles of extraction were done, using 0.4 μL of solvent each time. They were dispensed in a sample reservoir containing 13 μL of 0.1% FA, and after all cycles of extraction have been completed, 13 μL was injected into a 25 μL sample loop. Note that the loop was not filled to capacity. During sample loading, the remaining volume was compensated for by solvent A (0.1% FA) and then injected to the column. If the need arises, additional extraction cycles can be done further since there is still an excess of the loop volume that needs to be filled. Even though the volume is very
small and can thus limit the solubility of peptides, this was overcome because the repeated cycles of extraction leads to wetting of the sampling surface and thus improves the extraction efficiency with each consecutive cycle.

One aspect of the current bottom-up, liquid junction microextraction approach that can further be optimized is the efficiency of the digestion process. Compared to conventional sample digestion techniques, \textit{in situ} digestion has its distinct challenges. For instance, proteins \textit{in situ} are less susceptible to proteolysis compared when they are in solution due to lower accessibility. The lower amount of proteins within micrometer-thick tissue sections also makes them easily overwhelmed by successive addition of proteolytic enzymes leading to autolysis and successive deposition of the proteolytic enzyme can easily lead to loss of spatial resolution by solvent diffusion. The extent of digestion depends on a variety of factors. The presence of disulfide bridges renders them less susceptible to denaturation and subsequently degradation [22]. As such, reduction of these bonds by reducing agents such as DTT, \textbeta-mercaptopethanol, or tris(2-carboxyethyl)phosphine (TCEP) is often incorporated in conventional protein digestion protocols, coupled with cysteine capping using iodoaceticamide to prevent the re-oxidation of these residues. Introduction of chaotropes by heating to denaturing temperatures, addition of high concentrations of urea or guanidinium hydrochloride, addition of organic solvents, etc. are also imperative in the development of protein digestion protocols as the denaturation of
proteins allows for the exposure of otherwise buried lysine and arginine residues thereby facilitating their cleavage. These are aspects that still need to be further investigated to improve on-tissue digestion.

In this work, we demonstrate the use of the liquid microextraction approach as a downstream method to characterize the protein content of the regions defined by segmentation analysis of ovarian cancer IMS

Table 1
Gene Ontology overrepresentation analysis of the four protein clusters described in Fig. 6 using Protein Analysis Through Evolutionary Relationships (PANTHER) Classification System.

<table>
<thead>
<tr>
<th>PANTHER GO-Slim biological process</th>
<th>Homo sapiens (REF) #</th>
<th>#</th>
<th>Expected</th>
<th>Fold enrichment</th>
<th>±</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clusters 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Translation</td>
<td>435</td>
<td>5</td>
<td>0.29</td>
<td>17.09</td>
<td>+</td>
<td>1.51E − 03</td>
</tr>
<tr>
<td>PANTHER GO-Slim molecular function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural constituent of ribosome</td>
<td>180</td>
<td>3</td>
<td>0.12</td>
<td>24.78</td>
<td>+</td>
<td>3.70E − 02</td>
</tr>
<tr>
<td>Clusters 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein metabolic process</td>
<td>2692</td>
<td>10</td>
<td>2.33</td>
<td>4.3</td>
<td>+</td>
<td>4.68E − 03</td>
</tr>
<tr>
<td></td>
<td>Primary metabolic process</td>
<td>6825</td>
<td>14</td>
<td>5.9</td>
<td>2.37</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Metabolic process</td>
<td>8247</td>
<td>15</td>
<td>7.13</td>
<td>2.1</td>
<td>+</td>
</tr>
<tr>
<td>PANTHER GO-Slim molecular function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Translation regular activity</td>
<td>128</td>
<td>3</td>
<td>0.11</td>
<td>27.1</td>
<td>+</td>
<td>2.99E − 02</td>
</tr>
<tr>
<td>Translation factor activity, nucleic acid binding</td>
<td>136</td>
<td>3</td>
<td>0.12</td>
<td>25.51</td>
<td>+</td>
<td>3.57E − 02</td>
</tr>
<tr>
<td>Catalytic activity</td>
<td>5209</td>
<td>12</td>
<td>4.5</td>
<td>2.66</td>
<td>+</td>
<td>3.95E − 02</td>
</tr>
<tr>
<td>Clusters 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Translation</td>
<td>435</td>
<td>5</td>
<td>0.4</td>
<td>12.59</td>
<td>+</td>
<td>8.02E − 03</td>
</tr>
<tr>
<td></td>
<td>Protein metabolic process</td>
<td>2692</td>
<td>11</td>
<td>2.46</td>
<td>4.48</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Primary metabolic process</td>
<td>6825</td>
<td>15</td>
<td>6.23</td>
<td>2.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Metabolic process</td>
<td>8247</td>
<td>17</td>
<td>7.53</td>
<td>2.26</td>
</tr>
<tr>
<td>PANTHER GO-Slim molecular function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Translation initiation factor activity</td>
<td>97</td>
<td>3</td>
<td>0.09</td>
<td>33.88</td>
<td>+</td>
<td>1.57E − 02</td>
</tr>
<tr>
<td></td>
<td>Translation factor activity, nucleic acid binding</td>
<td>136</td>
<td>3</td>
<td>0.09</td>
<td>33.88</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Translation regulator activity</td>
<td>128</td>
<td>3</td>
<td>0.12</td>
<td>25.68</td>
<td>+</td>
</tr>
<tr>
<td>Clusters 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muscle contraction</td>
<td>217</td>
<td>4</td>
<td>0.17</td>
<td>23.98</td>
<td>+</td>
<td>4.30E − 03</td>
</tr>
</tbody>
</table>
data. The combination benefits from the advantages of both approaches in that microscale extraction is done directly on regions of the tissue section where actual phenotype changes associated with cancer are observed, and at the same time is not impeded by protein detection and identification limitations as is typical in IMS experiments. In the absence of additional histopathological information, we used the segmentation results for discrimination of benign and tumor regions in the ovarian biopsies.

Among the proteins identified with significant differential expression in the benign or tumor regions, ribosomal proteins and translation initiation factors comprised a large group of the overexpressed proteins concomitant with increased protein processing in the tumor zones [23]. Translation initiation is the most crucial point of regulation of gene expression at the level of protein synthesis and it is here where effector mechanisms such as extra- and intra-cellular signalling directly impact translation process. It can be, for example, harnessed by cancer cells to initiate and sustain the production of the transformed phenotype for its continued proliferation. EIF4A2, an isoform of eIF4A which is a member of the DEAD-box family of helicases, has been reported to be overexpressed in ovarian as well as other cancers. In fact, only eIF4A2 has been implicated as the required factor in mRNA inhibition by micro RNAs [24]. In ovarian cancer, the Let-7 family of miRNAs are involved in the downregulation of tumor suppressors KRAS, HRAS, C-MYC, HMGA-2, Cyclin A, D1, D2, D3, CDC25 and CDK6 [25,26]. As such, anti-neoplastic agents specifically inhibiting the helicase activity of elf4A2 have been explored [27]. EIF5A and its isoform elf5A2, are hypusine-containing translation initiation factors that have been recently implicated in the elongation phase of protein synthesis [28]. A strong correlation exists between elf5A and cancer, with each isoform being associated with certain malignancies. In contrast to elf5A1 which is ubiquitously expressed at a high level in normal tissues, elf5A2 is less abundant, and has specific overexpression particularly in ovarian and other types of tumors, although elf5A2 gene amplification is often variable suggesting additional mechanisms contributing to its high expression [29]. Both isoforms have been associated with less favorable clinical outcomes, with elf5A2 correlated with advanced stages of cancer and as well as metastasis [30]. Hence its potential as a prognostic marker and therapeutic target [28]. Among the proteins identified to be of lower abundance in the tumor regions, OBSCN [31], TAGLN [32,33], TAGLN3 and CNN3 [34], associated with the GO term “muscle contraction”, are proteins implicated in epithelial-mesenchymal transition (EMT) and consequently tumor progression and metastasis due to the deregulation of actin and microtubule cytoskeleton. The lower abundance of these proteins coincide with the high abundance of the translation initiation factors, which are also implicated in EMT regulation [35].

5. Conclusion

A method providing automatic micro-extraction and injection into the nLC-MS instrument from a tissue surface for Shot-Gun proteomics was developed in this work. By coupling this setup with on-tissue digestion using automatic microspotting of the digestion enzyme, analysis of a very limited area of the tissue section down to 250 μm spot size was realized. We could show that very good protein identification is obtained using the setup without requiring any further sample processing after microdigestion. The method was demonstrated on the examination ROIs defined by IMS experiments, showing its potential application in downstream compound identification which remains limited in the aforementioned method. We expect this method to find important applications in biology and clinics by giving access to the cell microenvironment related to a specific physiological process.

![Fig. 8. Workflow for the direct or indirect injection of liquid microjunction peptide extracts.](image-url)
Funding

This work was supported by the Ministère de l’Enseignement Supérieur et de la Recherche, Fédération de Recherche FR 3688 CNRS FRABIO, and by Siric ONCOLille (Grant INCa-DGOS-Inserm 6041). The authors would also like to thank Clic-Imaging Core Facility.

Transparency document

The Transparency document associated with this article can be found, in the online version.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbapap.2016.11.002.

References