Methods in Molecular Biology 2228

Springer Protocols

Katrin Marcus
Martin Eisenacher
Barbara Sitek *Editors*

Quantitative Methods in Proteomics

Second Edition

Humana Press

Chapter 20

Application of Label-Free Proteomics for Quantitative Analysis of Urothelial Carcinoma and Cystitis Tissue

Kathrin E. Witzke, Frederik Großerueschkamp, Klaus Gerwert, and Barbara Sitek

Abstract

A label-free approach based on a highly reproducible and stable workflow allows for quantitative proteome analysis. Due to advantages compared to labeling methods, the label-free approach has the potential to measure unlimited samples from clinical specimen monitoring and comparing thousands of proteins. The presented label-free workflow includes a new sample preparation technique depending on automatic annotation and tissue isolation via FTIR-guided laser microdissection, in-solution digestion, LC-MS/MS analyses, data evaluation by means of Proteome Discoverer and Progenesis software, and verification of differential proteins. We successfully applied this workflow in a proteomics study analyzing human cystitis and high-grade urothelial carcinoma tissue regarding the identification of a diagnostic tissue biomarker. The differential analysis of only 1 mm^2 of isolated tissue cells led to 74 significantly differentially abundant proteins.

Key words Label-free proteomics, FTIR imaging, Laser microdissection, Urothelial cell carcinoma (UCC), Bladder cancer, AHNAK2

1 Introduction

Label-free mass spectrometry (*see* also Chaps. 8, 16, 21–24), as the name implies, does not use any labeling strategies, making it very cost-effective, but only quantifies by matching identical peptides over several runs. Advantages beside cost-effectiveness and less sample preparation steps are high proteome coverage and high dynamic range. The disadvantages here are, therefore, high measurement times, as each sample or condition has to be measured separately and, furthermore, the separate handling of each sample from acquisition to measurement [1, 2].

In label-free proteomics, there are mainly two approaches of protein quantification, which are spectral counting and ion intensity-based quantification. While in spectral counting, as the name implies, the number of MS/MS fragment ion spectra that were

Katrin Marcus et al. (eds.), *Quantitative Methods in Proteomics*, Methods in Molecular Biology, vol. 2228,
https://doi.org/10.1007/978-1-0716-1024-4_20,

obtained for the peptides of a protein are counted and compared, the second type of label-free quantification measures the chromatographic peak areas of peptide precursor ions. Both strategies are possible, as both the number of MS/MS spectra of a peptide increase with the amount of the corresponding protein [3] and the areas under the curves (AUC) of chromatographic peptide peaks correlate linearly with the corresponding protein abundance [4, 5].

Nowadays, the main approach of protein quantification is ion intensity-based quantification as it relies on measuring physical data and not simply on counting the acquired spectra. In this approach, raw MS data have to be further processed for analysis. This includes, for example, feature detection, retention time alignment, intensity normalization, and peak picking [4].

The sample preparation is one very important step in proteomics for obtaining high-class quantification results. Especially in tissue proteomics, the right sample preparation is crucial due to tissue heterogeneity [6]. Here, a new strategy for automated annotation and isolation of regions of interest (ROI) has been used [7–9]. Conventionally, histological stainings or pathological annotations are necessary to detect ROIs in tissue samples, which are then transferred to unstained adjacent sections for LCM. The transmission of ROIs to adjacent slides, though, implies insurmountable deviations to annotated ROIs. The novel strategy of label-free automated tissue annotation and subsequent isolation via FTIR (Fourier transform infrared)-guided laser capture microdissection was coupled to subsequent label-free LC-MS/MS proteome analysis. By combining these techniques, very homogeneous samples can be obtained that are very accurately annotated, as the same tissue section annotated via FTIR imaging can be used for proteome analysis [10].

In label-free proteomics, the reproducibility and stability of the workflow are of highest importance due to the high sample complexity and separate measurements for each sample. Therefore, all steps of the label-free approach have to be optimized for best results, which includes the before-mentioned sample preparation with protein extraction and digestion, peptide separation by liquid chromatography, and data analysis including identification, quantification, and statistical analysis. One major advantage of label-free proteomics is its compatibility with high-throughput analyses that allow for processing of large numbers of biological samples required for statistically significant quantification.

We describe the application of a label-free approach for the identification of biomarker candidate proteins in the context of urothelial carcinoma diagnosis. We used fresh-frozen human tissue of patients with an inflammation of the bladder (cystitis) in comparison with high-grade urothelial carcinoma and performed label-free tissue annotation via FTIR (Fourier transform infrared) imaging with guided automated laser microdissection for unbiased

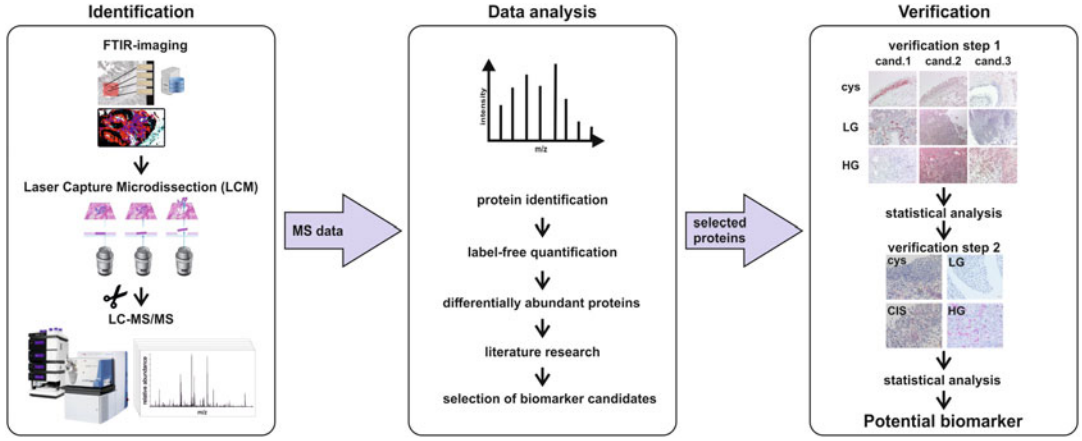

Fig. 1 Workflow of the label-free proteomics approach coupled with FTIR-guided LCM. First the tissue is annotated label-free and regions of interest isolated. After isolation, proteins will be extracted and digested with trypsin. Peptides will be analyzed via LC-MS/MS and generated data evaluated with Proteome Discoverer for identification and Progenesis QI for quantification. For verification of biomarker candidate proteins, immunohistochemistry will be performed

isolation of the tissue sections of interest only and subsequent label-free LC-MS/MS proteome analysis. For that, we used nano-HPLC coupled to an Orbitrap Elite mass spectrometer for the generation of peptide profiles. For quantitative analysis of the data, the software Progenesis QI for proteomics was used. Altogether, 74 proteins were found to show significant differential abundance between the analyzed groups (FDR-adjusted p-value ≤ 0.05 and absolute fold change ≥ 1.5). Verification was performed in two steps with increasing cohort sizes and the addition of more urothelial carcinoma groups (low grade and carcinoma in situ). From three tested candidates in the first step, AHNAK2 was selected for further verification and proposed as a biomarker candidate for the differentiation between cystitis and several subgroups of urothelial carcinoma (Fig. 1).

2 Materials

2.1 FTIR Imaging and Laser Microdissection

1. HM550 cryostat (Thermo Fisher Scientific, Waltham, MA, USA).
2. PET (polyethylene terephthalate) frame slides (Leica, Wetzlar, Germany).
3. Cary 620 IR microscope equipped with a 128 × 128 pixel liquid nitrogen-cooled mercury cadmium telluride (MCT) focal plane array (FPA) (Agilent Technologies, Santa Clara, CA, USA).

4. Cary 670 spectrometer (Agilent Technologies, Santa Clara, CA, USA).
5. Parker Balston AirDryer Assembly 75-62 (Parker Hannifin Corporation, Lancaster, NY, USA).
6. PALM Microbeam Laser microdissection (LMD) microscope (Carl Zeiss Microscopy GmbH, Jena, Germany).
7. MATLAB (MathWorks, Natick, MA, USA) or equivalent (e.g., R or Python).

2.2 Sample Preparation and Digestion

1. Ultrasonic bath (VWR, Darmstadt, Germany).
2. Centrifuge (Eppendorf, Hamburg, Germany).
3. Lysis buffer: 50 mM ammonium bicarbonate with 0.1% RapiGest SF surfactant (Waters GmbH, Eschborn, Germany).
4. Digestion: 20 mM dithiothreitol (DTT), 100 mM iodoacetamide (IAA), 33 ng/μL trypsin, trifluoroacetic acid (TFA).

2.3 Liquid Chromatography

1. Ultimate 3000 RSLCnano high-performance liquid chromatography system (Dionex, Idstein, Germany).
2. Trap column: Acclaim PepMap100 C18 Nano-Trap column (C18, 100 μm × 2 cm, particle size 5 μm, pore size 100 Å; Thermo Fisher Scientific, Bremen, Germany).
3. Nano column: Acclaim PepMap RSLC Nano Viper C18 analytical column (C18, 75 μm × 50 cm, particle size 2 μm, pore size 100 Å; Thermo Fisher Scientific).
4. Loading solvent: 0.1% (v/v) TFA (MS grade).
5. Gradient solvent A: 0.1% (v/v) Formic acid (FA) (MS grade).
6. Gradient solvent B: 0.1% (v/v) FA (MS grade), 84% (v/v) acetonitrile (ACN) (MS grade).

2.4 Mass Spectrometry

1. LTQ Orbitrap Elite with an online nano-ESI source (Thermo Fisher Scientific).
2. Pico Tip™ emitter Silica Tip™ (New Objective, Woburn, USA).
3. Collision gas: nitrogen.

2.5 Data Analysis

1. Proteome Discoverer v.1.4 (Thermo Fisher Scientific).
2. Mascot v.2.5 (Matrix Science, London, UK).
3. Progenesis QI v.2.0 (Nonlinear Dynamics, Durham, NC, USA).
4. R v.3.4.0 (Free Software Foundation).

3 Methods

3.1 General Practice

Human tissue of patients with an inflammation of the bladder (cystitis) or high-grade UCC was collected during cystectomy surgery according to standard operation procedure. Tissue was washed with isotonic saline solution, slowly frozen on the surface of liquid nitrogen within 8 min and stored at −80 °C. Frozen tissue was sectioned with an HM550 cryostat (Thermo Fisher Scientific, Waltham, MA) at −20 °C, and 10 μm sections were collected on polyethylene terephthalate (PET) frame slides.

3.2 Automatic Annotation Via FTIR Imaging and Laser Microdissection

1. Take the tissue thin section mounted on a PET frame slide and place it under the FTIR imaging microscope (here Cary 620) (*see* also **Note 1**).
2. Thaw the tissue sample under dry air in the FTIR system. At the same time, the system is stabilized with the dry air.
3. Select a clean background position on the PET slide and collect spatially resolved IR spectra in the wave number region 3700–950 cm^{-1} at a spectral resolution of 4 cm^{-1} with co-added 128 scans. Use a 15× objective resulting in a pixel resolution of ~5.5 μm and a field of view (FOV) of 715 μm^2 per FPA field.
4. Select the region of interest (ROI) on the PET slide and collect spatially resolved IR spectra in mapping mode. This allows imaging of larger regions than the FOV by stitching the collected FPA fields afterwards.
5. Stitch the collected spectral hypercubes in MATLAB (or equivalent software) and pre-process the data. The first step is a quality test based on the integral of the amide I band and the signal-to-noise ratio (noise, 2100–2000 cm^{-1}; signal, 1600–1500 cm^{-1}). Then subject all spectra to extended multiplicative scattering correction-based Mie and resonance-Mie scattering correction from 2300 to 950 cm^{-1}.
6. The dataset is now prepared for multivariate data analysis or machine learning algorithms. For bladder, pre-train and use a random forest (RF) classifier to annotate the tissue (*see* also **Note 2**).
7. Select the ROIs from the label-free annotated IR image for isolation via LMD.
8. Select three reference points at the IR imaging system and then transfer the tissue section to the LMD.
9. Find the three selected reference points at the LMD. They are needed for the coordinate transfer.

10. Transfer the coordinates of the ROIs selected from the FTIR imaging results to LMD by two-dimensional Helmert transformation based on three reference points in MATLAB or equivalent software (*see* also **Note 3**).
11. Collect the needed tissue area. For bladder, 10 μm sections were used, and regions of 1 mm^2 were collected in lysis buffer (20 μL/1 mm^2 cells).

3.3 Sample Preparation and Enzymatic Digestion

1. Lyse the cells in lysis buffer (20 μL/1 mm^2 cells) and sonicate the samples upside down on ice for 1 min and finally centrifuge the samples in the upright position for 1 min to transfer them from the lid to the vial itself.
2. Normally, it is necessary to know the concentration of the samples. However, due to LCM isolation of only 1 mm^2 tissue, digest all.
3. Perform a tryptic in-solution digest for the proteolysis of proteins. For reduction, add 3.7 μL DTT (20 mM) to the samples and incubate for 30 min at 60 °C. Afterwards alkylate with 2.2 μL IAA (100 mM) for 30 min at room temperature in the dark. Add trypsin (0.02 μg) to digest the proteins overnight (max. 16 h) at 37 °C. Stop the digestion by adding 1.3 μL 10% TFA to the solution, incubate for 30 min at 37 °C. Afterwards centrifuge the samples (10 min, 16,000 × *g*) and transfer the supernatant to a glass vial (*see* **Note 4**).
4. Before performing the LC-MS analysis, dry the samples in a vacuum centrifuge and dissolve them in 17 μL 0.1% TFA. Use a sample amount of the whole 1 mm^2 tissue area for one LC-MS/MS measurement (*see* **Note 5**).

3.4 Peptide Separation with Reversed Phase High-Performance Liquid Chromatography

1. For the separation of the digested proteins, perform a reversed phase high-performance liquid chromatography with the Ultimate 3000 RSLCnano high-performance liquid chromatography system (Dionex). Within this, use a system containing a nano-trap column (C18) and a nano-analytical column (C18). The columns need to be heated to 60 °C to allow high flow rates of 400 nL/min at acceptable pressure (*see* **Note 6**).
2. Use a sample volume of 15 μL for injection. First, peptides are pre-concentrated on the trap column for 7 min, while detergents and salts are washed away. Use a flow rate of 30 μL/min for loading.
3. The gradient for peptide separation works as follows: (a) linear gradient from 5 to 40% solvent B over 98 min, followed by (b) 95% B in 2 min, (c) constant 95% B for 7 min, and finally (d) 5 min at 5% B for equilibration. Set the gradient pump flow rate to 400 nL/min.

3.5 Detection of Separated Peptides with Mass Spectrometry

1. Operate the Orbitrap Elite™ mass spectrometer (Thermo Fisher Scientific) in the data-dependent mode to automatically switch between MS and MS/MS acquisition.
2. Set the mass range for survey full scan MS spectra and MS/MS spectra to m/z 350–2000.
3. For fragmentation, apply collision-induced dissociation (CID) with nitrogen as collision gas and normalized collision energy of 35. For MS/MS measurements, use a top 20 method based on intensity (*see* **Note 7**). The minimal required signal for precursor ions is 500 counts, and the isolation width is 2 ppm.
4. Reject charge states 1+ and prefer charge states 2+, 3+, and 4+ for precursor ion isolation.
5. Utilize dynamic exclusion with an exclusion duration of 30 s and one repeat count within 30 s. Use exclusion list size of 500 precursor ions with an exclusion mass width of 10 ppm.
6. In the end, export generated data as Thermo .raw file format.

3.6 Identification of Measured Proteins

1. For protein identification, use the software Proteome Discoverer version 1.4 (Thermo Fisher Scientific). Search spectra against the UniProtKB/Swiss-Prot database using the Mascot search engine version 2.5 (Matrix Science).
2. Create a Proteome Discoverer workflow. Use the following parameters: (a) taxonomy setting, homo sapiens; (b) enzyme, trypsin, (c) missed cleavages, allow up to one; (d) dynamic modification, oxidation (methionine); (e) static modification, carbamidomethyl (cysteine); (f) precursor mass tolerance, 5 ppm; (g) fragment mass tolerance, 0.4 Da; (h) false discovery rate (FDR), via p function (identifications with FDR >1% are rejected).
3. Import .raw data files into the Proteome Discoverer Daemon and start the search with the created workflow.
4. Open the results with the Proteome Discoverer Viewer and export results in Excel format.

3.7 Quantitative Proteome Analysis

1. As software for quantitative proteome analysis, use Progenesis QI for proteomics (Nonlinear Dynamics). First, import the LC-MS analysis .raw data files into the program.
2. Select the reference run that all your other runs are aligned to. This can either be done by yourself, or there is the option to let the program select the reference run, either out of all runs or out of a selection you make. An optimal reference run should have the greatest similarity to all other runs (*see* **Note 8**).
3. The alignment step is most important for label-free quantification. Therefore it is necessary to have a very accurate alignment result. First, apply the automatic alignment, but be sure to

check the alignment carefully when it is finished. If the result of the alignment is not good enough, you can often alter the result for the better by manually adding vectors to align specific runs.

4. Progenesis automatically performs feature detection, normalization, and quantification. Always check results and exclude or include features from analysis results. Exclude retention times from washing/equilibration. Include only ion charge states of 2+, 3+, and 4+ with a minimum of three isotope peaks to exclude contaminations from the analysis (*see* **Note 9**).
5. Create the experimental design, in our case cystitis vs. high-grade UCC. Runs which, e.g., did not align properly can be excluded.
6. Identify the quantified proteins. Therefore, import Excel results of priorly obtained identification via Proteome Discoverer. Consider all non-conflicting peptides for protein quantification.
7. Differential analysis is also performed by Progenesis. You can filter based on *p*- and *q*-value (≤ 0.05) and fold change (≥ 1.5) and tag them for further analysis.
8. You can perform principal component analysis (PCA) to check if runs cluster based on experimental grouping. Also, check regulation profiles of interesting candidates (*see* **Note 10**).
9. Export results of quantified proteins in Excel format.

3.8 Statistical Analysis

1. Despite the results of the differential analysis by Progenesis, perform a separate statistical analysis via R. Arcsinh-transform normalized protein abundances obtained from Progenesis and use those for *t*-test calculations. Adjust test *p*-values for FDR control with the method of Benjamini Hochberg.
2. Use normalized protein abundances obtained from Progenesis for fold change calculation.
3. Consider proteins significantly differentially abundant between experimental groups if they have an absolute fold change ≥ 1.5 and an FDR-corrected *p*-value ≤ 0.05.

4 Notes

1. You can use other FTIR or IR imaging systems also. The procedure is equivalent. Furthermore, if other options in data collection are more convenient for you, change them. The values given here are only recommendations that have worked well for our work on bladder cancer.

2. For analysis of the pre-processed FTIR imaging datasets, you can use multivariate data analyses like *k*-means clustering or hierarchical cluster analysis. This way, you select your ROIs on spectral similarity. A better and more precise way is to train a supervised cluster algorithm like random forests (RF). Therefore, you need a spectral database for different tissue types that is created by measuring known tissue samples previously. Then the classifier (RF) can be trained, and afterwards it is possible to annotate unknown spectral datasets by the use of this classifier.
3. The coordinate transfer is the same as the transfer of coordinates in geographic sciences. Beside the Helmert transformation, you can use other methods also. It is always helpful to test the accuracy of the transfer with a test target previously.
4. The centrifugation step after stopping the tryptic digest is crucial here to remove excess RapiGest from the samples that could potentially damage your LC system.
5. The workflow can also be applied for body fluids and cell culture experiments. An adapted sample processing could also be carried out successfully for FFPE tissue.
6. On the one hand, heating of columns is needed to reduce the pressure of the system and to get a better peptide separation; however, heated columns need higher flow rates for acceptable pressure. Higher flow rates mean worse sensitivity. Thus, a compromise between separation and sensitivity is necessary.
7. In a top 20 method, the 20 most abundant peptide ions of the full scan are selected for fragmentation and measured for tandem mass spectra in the linear ion trap.
8. Where possible, master mixes of all samples can be used as optimal reference runs for alignment, as they combine features of all runs.
9. Tryptic peptides have charge states between 2+ and 4+, while contaminations mostly have charge states of 1+. If another protease than trypsin is used, other charge states might apply.
10. Checking regulation profiles offers easy insight into the quality of candidate proteins. It can be observed that sometimes only few but high differences can boost statistical significance.

Acknowledgments

This work was supported by the Ministry of Innovation, Science and Research of North-Rhine Westphalia, Germany. The authors would like to thank Lidia Janota, Kristin Fuchs, Stephanie Tautges, and Birgit Zülch for their excellent technical assistance.

References

1. Bantscheff M, Schirle M, Sweetman G et al (2007) Quantitative mass spectrometry in proteomics: a critical review. Anal Bioanal Chem 389(4):1017–1031. https://doi.org/10.1007/s00216-007-1486-6
2. Megger DA, Pott LL, Ahrens M et al (2014) Comparison of label-free and label-based strategies for proteome analysis of hepatoma cell lines. Biochim Biophys Acta 1844 (5):967–976. https://doi.org/10.1016/j.bbapap.2013.07.017
3. Liu H, Sadygov RG, Yates JR 3rd (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 76(14):4193–4201. https://doi.org/10.1021/ac0498563
4. Megger DA, Bracht T, Meyer HE et al (2013) Label-free quantification in clinical proteomics. Biochim Biophys Acta 1834(8):1581–1590. https://doi.org/10.1016/j.bbapap.2013.04.001
5. Bondarenko PV, Chelius D, Shaler TA (2002) Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. Anal Chem 74(18):4741–4749. https://doi.org/10.1021/ac0256991
6. Mukherjee S, Rodriguez-Canales J, Hanson J et al (2013) Proteomic analysis of frozen tissue samples using laser capture microdissection. Methods Mol Biol 1002:71–83. https://doi.org/10.1007/978-1-62703-360-2_6
7. Grosserueschkamp F, Kallenbach-Thieltges A, Behrens T et al (2015) Marker-free automated histopathological annotation of lung tumour subtypes by FTIR imaging. Analyst 140 (7):2114–2120. https://doi.org/10.1039/c4an01978d
8. Miller LM, Dumas P (2006) Chemical imaging of biological tissue with synchrotron infrared light. Biochim Biophys Acta 1758 (7):846–857. https://doi.org/10.1016/j.bbamem.2006.04.010
9. Ooi GJ, Fox J, Siu K et al (2008) Fourier transform infrared imaging and small angle x-ray scattering as a combined biomolecular approach to diagnosis of breast cancer. Med Phys 35(5):2151–2161. https://doi.org/10.1118/1.2890391
10. Grosserueschkamp F, Bracht T, Diehl HC et al (2017) Spatial and molecular resolution of diffuse malignant mesothelioma heterogeneity by integrating label-free FTIR imaging, laser capture microdissection and proteomics. Sci Rep 7:44829. https://doi.org/10.1038/srep44829