



Comparison of S-trap and Direct Trypsinization (DTR) Protein Extraction Methods for Proteomic Study

Mohammed Fadhil , Basima Q. Hassan , Abdul Hussein M. Al Faisal

Genetic Engineering and Biotechnology Institute / University of Baghdad, Iraq.

Received: August 25, 2020 / **Accepted:** September 20, 2020 / **Published:** October 1, 2020

Abstract: The proteomic analysis of formalin fixed paraffin embedded FFPE tissues often employs technologies such as Liquid Chromatography (LC) and Mass Spectrometry (MS), which are considered as important tools in protein biomarker research. Fixation of tissue samples using formalin leads to extensive inter- and intramolecular crosslinking among proteins in these tissues, which hampers the proteome analysis of these samples. Therefore, a substantial need exists to develop methods and procedures for this technology to discover new protein biomarker candidates with diagnostic or therapeutic potential. **Methods:** Two protocols were used in this study for protein extraction from cervicitis FFPE samples. First protocol used xylene in deparaffinization and rehydrated samples with graded series of ethanol. Then, samples were incubated in different temperatures, lysed with buffer and subjected to sonication and centrifugation. Afterwards, samples were reduced and alkylated then followed by trypsin digestion. The second protocol used S-trap buffer directly on FFPE slices then followed by series of incubation and sonication. Eventually, reduction and alkylation of samples were performed to be followed by trypsin digestion. Finally, the protein extracts were subjected to LC-MS/MS analysis. The data were analyzed by scaffold and PANTHER tools. **Results:** The number of proteins that shared by these protocols were 734 and the S-trap method was higher in unique protein identifications (n= 338) than the direct trypsinisation (DTR) method (n= 276). Most of the proteins that missed in DTR methods were mainly located in organelle, protein-containing complex, and membrane. While there were no highly differences between the numbers of extracellular proteins from both protocols. **Conclusion:** The S-trap method can be considered as a promising method in FFPE proteomic studies because there is no need for highly experience and required short time to be performed. More important proteins such as membrane proteins were less missed in this protocol which is considered as a pivotal method in membrane proteomic study.

Key words: FFPE, Proteomic, S-trap, LC-MS/MS, cervicitis.

Corresponding author: (Email: mohaf1988@gmail.com).

Introduction

Use of FFPE tissues for proteomics was traditionally thought to be too problematic, given the presence of formalin-induced protein cross-links and modifications, which create problems for separating, visualizing and characterizing individual proteins (1). Recently, it was suggested that formalin-fixed paraffin-embedded

(FFPE) tissue, the golden standard for long-term preservation of tissue worldwide, can serve as a valuable alternative for fresh tissue in proteomic study (2). Since FFPE tissue is routinely prepared for pathological research, millions of samples are available in hospital archives and large numbers of samples can be collected in short periods of time (3). Thus, the ability to retrospectively analyze documented

archival FFPE cases, with known diagnosis, prognosis, response to therapy, and outcome, represents a significant resource for biomarker discovery (4).

However, in the last few years several groups described successful protein extraction from FFPE tissues (5-9). Many FFPE proteomics studies are published in the literature. This FFPE research can be mainly classified into three categories: studies investigating FFPE protein extraction methods very often accompanied by a fresh frozen tissue versus FFPE tissue comparison (10,11), studies determining the suitability of FFPE tissue in biomarker (12) and studies looking for the influence of (pre-analytical factors (13). All these studies indicate that multiple methods are used among different research groups. This complicates rigorous comparison and divergent results are obtained. Therefore, there is a high need for a rigorous comparison of multiple protein extraction methods with the same proteomic workflow.

In this study, the most recent protein extraction method called S-trap is going to be compared with previous method of direct trypsinization (DTR) using a standardized LC-MS/MS workflow in order to determine which method is optimum and can be used for proteomic study.

Materials and methods

Protein extraction methods

Two extraction methods was used in this study one of these methods called direct trypsinization (DTR) as shown in protocol 1 while the S-trap

method explained in protocol 2. The samples were collected from medical city teaching laboratory / Baghdad 2019/2018.

Protocol 1 (3, 14)

Two cervicitis samples were used in this protocol. The five slices of 10 μ m FFPE sections for each sample were deparaffinized by xylene for 10 min, followed by centrifugation at 10 000 xg for 3 min. Graded series of ethanol (100%, 95% and 70%) were used to rehydrate the tissue pellets. The hydrated FFPE samples were suspended in 20 mM Tris HCl pH 8.8, 200 mM DTT, 2% SDS and 1% protease inhibitor (Complete cocktail, Roche, Penzberg, Germany). After suspension in lysis buffer, samples were incubated for 20 min at 98 °C, followed by incubation at 80 °C for 2 h. Then, samples sonicated with a microprobe (Sonic Dismembrator 550, Fisher Scientific) 2 x 30s at power 2.5 and centrifuged at 14 000 xg for 30 min at 4 °C, the supernatant was transferred to a new tube. Afterwards, Digestion of proteins was performed by reducing with 20mM DTT for 10min at 95°C and alkylated with 40mM IAA (iodoacetamide) for 30min at room temperature in the dark. Sequencing grade trypsin (Worthington, Lakewood, NJ, USA) was added in a ratio (1:50). For trypsinization, the sample was incubated at 37 °C overnight. Reduction and halted trypsinization were performed by adding 1 μ L of formic acid. Any precipitate was removed by centrifugation at 19,000 xg for 10 min. 15 μ g of peptides per sample were desalted using self-packed C18 STAGE tips (Empore, St. Paul, MN, USA).

Peptides amount was measured by nanodrop at 205nm.

Protocol 2 (9)

Cervicitis FFPE tissue sections were sectioned as previously mentioned and resuspended in 1ml of S-trap (Protifi) lysis buffer (5% SDS, 50mM triethylammonium bicarbonate pH 7.55) then heated at 50°C for 10 min, sonicated with a microprobe (Sonic Dismembrator 550, Fisher Scientific) 2 x 30s at power 2.5 then heated at 80°C for 60 min followed by sonication of 1 x 30 s and finally heated at 80°C for 60min. The extract was centrifuged at 16,000g for 15min and the supernatant was acetone precipitated and then resuspended in 50ul of S-trap Lysis buffer (as above).

Digestion of proteins was performed by reducing with 20mM DTT for 10min at 95°C and alkylated with 40mM IAA (iodoacetamide) for 30min at room temperature in the dark. S-Trap (Protifi) protocol was followed for the next steps, briefly phosphoric acid was added at a final concentration of 1.2%. The solution was mixed with the S-Trap binding buffer (90% methanol in 100mM TEAB pH 7.1) and transferred in a S-Trap spin column then centrifuged and washed. Trypsin (1:50) in 50mM TEAB was added and incubated for 1hr at 47C. Peptides were eluted with 50mM TEAB then with 0.2% formic acid and dried down. Peptides amount was evaluated with the nanodrop at 205nm.

LC-MS/MS

Samples were analyzed by nanoLC-MS/MS. For each injection, 1 µg of

peptide digest were separated by using Dionex UltiMate 3000 nanoRSLC chromatography system (Thermo Fisher Scientific / Dionex Softron GmbH, Germering, Germany) connected to an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) driving with Orbitrap Fusion Tune Application 2.0 and equipped with a nanoelectrospray ion source.

Peptides were trapped at 20 µL/min in loading solvent (2% acetonitrile, %0.05TFA) on a 5 mm × 300 µm C18 PepMap cartridge pre-column (Thermo Fisher Scientific/Dionex Softron GmbH, Germering, Germany) during 5 min. Then, the pre-column was switch online with a 50 cm length, 75 µm ID Acclaim PepMap 100 C18 analytical column (Thermo Fisher Scientific/Dionex Softron GmbH, Germering, Germany) and the peptides were eluted with a linear gradient from 5 to 40% solvent B (A: %0.1formic acid, B: 80% acetonitrile, 0.1% formic acid) in 30 min, at 300 nL/min.

Mass spectra were acquired using a data dependent acquisition mode using Thermo XCalibur software version 3.0.63. Full scan mass spectra (350 to 1800 m/z) were acquired in the orbitrap using an AGC target of 4e5, a maximum injection time of 50 ms and a resolution of 120,000. Internal calibration using lock mass on the m/z 445.12003 siloxane ion was used. Each MS scan was followed by acquisition of fragmentation MSMS spectra of the most intense ions for a total cycle time of 3 s (top speed mode). The selected ions were isolated using the quadrupole analyser in a window of 1.6 m/z and fragmented by Higher energy Collision-induced Dissociation (HCD) with 35%

of collision energy. The resulting fragments were detected by the linear ion trap in rapid scan rate with an AGC target of $1e4$ and a maximum injection time of 50 ms. Dynamic exclusion of previously fragmented peptides was set for a period of 20 s and a tolerance of 10 ppm.

Data analysis

Spectra was acquired scaffold 4.11.0 and protein identified by using Mascot database (Matrix Science, London, UK; version 2.5.1). Mascot was set up to search the contaminants_thegpm_20170713.fasta; REF_HomoSapiens_ci_9606_up000005_640_20180425 database (93675 entries) and digestion enzyme was trypsin. Mascot was searched with a fragment ion mass tolerance of 0.100 Da and a parent ion tolerance of 0.100 Da. Carbamidomethyl of cysteine was specified in Mascot as a fixed modification. Deamidated of asparagine and glutamine and oxidation of methionine were specified in Mascot as variable modifications.

Scaffold (version Scaffold_4.11.0, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 93.0% probability to achieve an FDR less than 1.0% by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 85.0% probability to achieve an FDR less than 5.0% and contained at least 1 identified peptide. Protein probabilities were

assigned by the Protein Prophet algorithm (15).

Functional enrichment analysis was performed by using by PANTER tools\ Gene list analysis (16).

Results and discussion

Protein extraction method

Formalin fixed paraffin embedded samples considered as the most challenging samples to work with in proteomic study especially in quantitative study. For this reason we optimized the method of protein extraction from these samples to reach better results to run TMT experiment in further steps.

We did use two methods for protein extraction as previously mentioned in materials and methods chapter and we named every method as FFPE. The first protocol from (3, 14) named as FFPE 1 and the second protocol from (9) called FFPE2. The data viewed in Scaffold software to find the number of identified proteins.

The files from LC-MS\MS contain only raw MS peaks which need blasting onto database of MS spectra. Mascot database was used to align our spectra into this database to determine the sequence of peptide resulted from these spectra. Tubulin beta-2A chain was taken as protein example from hundreds of proteins. After spectra alignment, the spectrum would be annotated by scaffold and the spectra of peptides colored in regard to protein entry in database because there is more than one entry for same protein in database Figure (1).

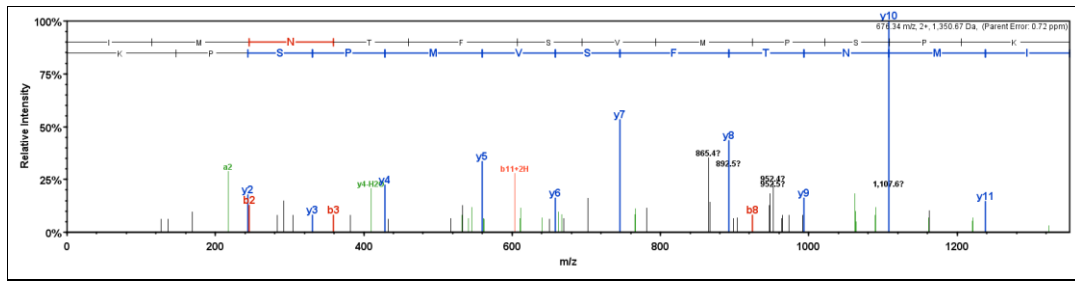


Figure (1): The spectrum of unique peptide sequence of Tubulin beta-2A chain.

Because there are many same peptides shared by different proteins, thus each protein identified according a unique peptide which not shared by other proteins. The same protein example, Tubulin bet-2A chain, was

used and as shown in Figure (2). There are many peptides not unique and cannot be used in identification while only one peptide was unique which is labeled in red color.

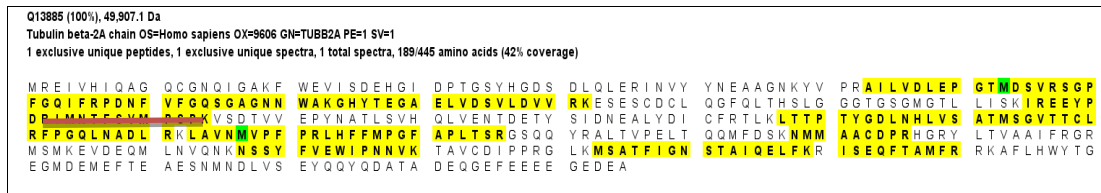


Figure (2): Shows the unique peptide sequence of Tubulin beta-2A chain.

The number of identified proteins, total unique peptides and total unique spectra by these two methods were remarkable different. Generally, there were some proteins shared by these two methods but mainly the FFPE2 was the higher one in protein number. Total unique protein identification in FFPE2

was 338 and 276 in FFPE1 while the shared proteins were 734. However, the total unique peptides and spectra in FFPE2 were approximately twice higher than FFPE1 and they shared 1594 and 1614 proteins respectively Figure (3).

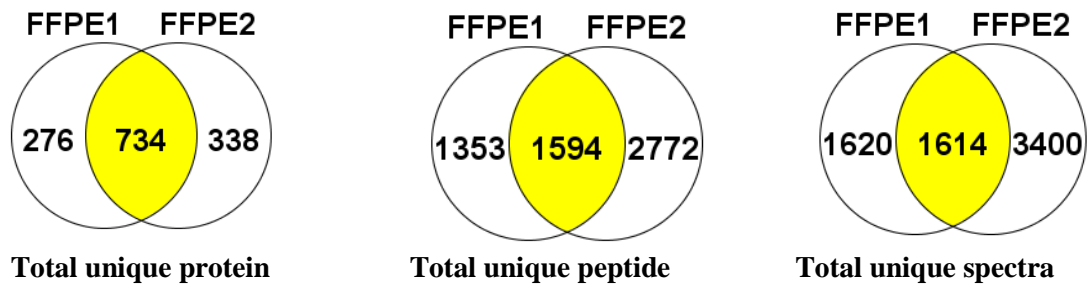


Figure (3): Reveals the differences and similarities between these two methods in total proteins, total unique peptides and total unique spectra.

The functional enrichment analysis was performed by using PANTER tools\ Gene list analysis (16) for

determining the location of proteins that missed between the two extraction methods FFPE1 and FFPE2. The results

showed that large number of proteins were missed from FFPE1 and detected by FFPE2 and mainly located in organelle, protein-containing complex,

and membrane. While there were no highly differences between the number of extracellular proteins from both FFPE1 and FFPE2 Table (1).

Table (1): Shows the locations of proteins in both FFPE1 and FFPE2.

Type of extraction	Cellular location	Number of proteins	Frequency
FFPE2	cell (GO:0005623)	359	20.90%
	cell part (GO:0044464)	359	20.90%
	organelle (GO:0043226)	237	13.80%
	protein-containing complex (GO:0032991)	160	9.30%
	organelle part (GO:0044422)	138	8.00%
	extracellular region part (GO:0044421)	116	6.80%
	extracellular region (GO:0005576)	116	6.80%
	membrane (GO:0016020)	110	6.40%
	membrane part (GO:0044425)	55	3.20%
	membrane-enclosed lumen (GO:0031974)	27	1.60%
	supramolecular complex (GO:0099080)	24	1.40%
	cell junction (GO:0030054)	8	0.50%
	synapse (GO:0045202)	5	0.30%
	synapse part (GO:0044456)	4	0.20%
FFPE1	cell (GO:0005623)	297	21.10%
	cell part (GO:0044464)	297	21.10%
	organelle (GO:0043226)	187	13.30%
	protein-containing complex (GO:0032991)	119	8.50%
	organelle part (GO:0044422)	107	7.60%
	extracellular region part (GO:0044421)	99	7.00%
	extracellular region (GO:0005576)	99	7.00%
	membrane (GO:0016020)	86	6.10%
	membrane part (GO:0044425)	43	3.10%
	supramolecular complex (GO:0099080)	29	2.10%
	membrane-enclosed lumen (GO:0031974)	24	1.70%
	cell junction (GO:0030054)	13	0.90%
	synapse (GO:0045202)	4	0.30%
	synapse part (GO:0044456)	3	0.20%

The efficiency of protein extraction methods from FFPE samples is highly variable in proteomic analysis and depends on different factors. Fixation time was reported as one of these factors that effect on number of identified proteins. Sprung et al., 2009 (17) found the fixation for four days had negatively impact on the number of identified protein groups. Even the yield of total proteins strongly decreased in 192 hours of fixation time (18). The age of FFPE tissue block is another factor

may effect on proteomic experiment. The protein yields decreased 42% when comparing FFPE blocks from 2010 and 1990 respectively (19).

The type of extraction method in proteomic experiment had a major role in determining the number of identified proteins. The deparaffinization process for FFPE samples and the type of buffer were considerably studied in regard to number of protein identifications. Hood *et al.* (2005) (20) conducted an earlier proteomic study on FFPE samples of

prostate cancer tissue and they used SubX with a graded ethanol series for deparaffinization and Liquid Tissue™ as lysis buffer. The number of protein identifications was 1156 and 702 for prostate cancer samples and benign prostate samples respectively. Another method, that used octane/methanol for deparaffinization and heat induced antigen retrieval (HIAR) for protein extraction from human renal carcinoma samples, found the number of protein identifications was quietly higher than previous study (n=1830) (21). The other study on FFPE nephrectomy tumor samples identified 283 protein groups by using xylene for deparaffinization and RapiGest as lysis buffer (22). Another commercial kit, Qproteome™, was used to extract protein from FFPE samples and the number of identified proteins that achieved in this proteomic experiment was 1812 (23). The variation among these methods came from different factors and one of these factors is the buffer that used in each work.

A remarkable study compared the differences in protein identifications among direct trypsinisation method, in solution digestion method and FASP method. The number of protein identifications for the FASP method was 1353 and direct trypsinisation was 1126 while in solution digestion was 1124. As clearly, the high number was achieved from FASP method because this method devised to remove the SDS which effect on MS analysis with little loose for sample (24).

The pre-fractionation step such as 1-DE, IEF, SCX, or MudPIT, prior to LC-MS/MS step highly improves the number of protein identifications and these techniques should be unified in

comparison of protein extraction methods (25).

Our data which compared between S-trap and general protein extraction method showed one of the highly affected types of proteins was membrane proteins and this is previously investigated and proved. The heat-induced antigen retrieval (HIAR) method of protein extraction from FFPE tissue sections produced a heavily depletion in basic proteins and specifically in membrane proteins because of formalin-fixation process that promote the HIAR-induced aggregation (26,27). In contrast, the in solution digestion and FASP methods produced higher yields of hydrophobic and membrane proteins (24).

Conclusion

To best of our knowledge this is the first study compare between S-trap method and DTR method. The proteome coverage that obtained from S-trap method was higher than DTR method which means the FFPE samples can be easily handled in extraction with high yield of proteins. The highly number of proteins that found in S-trap method and missed in DTR method made this method more efficient in overcoming the problem of formalin fixation. This study used a few technical replicates which need to be expanded in further studies with more representing for extraction buffers and longer elution time.

Acknowledgement

We would like to thank Dr. Sylvie Bourassa /proteomic center of Laval University/Canada for undertaken all proteomic work.

References

- Palmer-Toy, D.E.; Krastins, B.; Sarracino, D.A.; Nadol, J.B.J. and Merchant, S.N. (2005). Efficient method for the proteomic analysis of fixed and embedded tissues. *J Proteome Res.*;4(6):2404-2411.
- Magdeldin, S. and Yamamoto, T. (2012). Toward deciphering proteomes of formalin-fixed paraffin-embedded (FFPE) tissues. *Proteomics*.;12(7):1045-1058.
- Maes, E.; Valkenburg, D.; Mertens, I.; Inge, M.; Valérie, B.; Geert, B., *et al.* (2013). Proteomic analysis of formalin-fixed paraffin-embedded colorectal cancer tissue using tandem mass tag protein labeling. *Mol Biosyst.*;9(11): 2686-2695.
- Tanca, A.; Pagnozzi, D. and Addis, M.F. (2012). Setting proteins free: progresses and achievements in proteomics of formalin-fixed, paraffin-embedded tissues. *Proteomics Clin Appl.*; 6(1-2): 7-21.
- Ikeda, K.; Monden, T.; Kanoh, T.; Tsujie, M.; Izawa, H.; Haba, A., *et al.* (1998). Extraction and analysis of diagnostically useful proteins from formalin-fixed, paraffin-embedded tissue sections. *J. Histochem Cytochem* 46: 397–403.
- Gillespie, J.W.; Best, C.J.; Bichsel, V.E.; Cole, K.A.; Greenhut, S.F.; Hewitt, S. M. *et al.* (2002). Evaluation of non-formalin tissue fixation for molecular profiling studies. *Am J Pathol* 160: 449–457.
- Chung, J.Y.; Lee, S.J.; Kris, Y.; Braunschweig, T.; Traicoff, J.L. and Hewitt, S.M. (2008). A well-based reverse-phase protein array applicable to extracts from formalin-fixed paraffin-embedded tissue. *Proteomics Clin Appl.*;2(10-11):1539-1547.
- Maes, E.; Broeckx, V.; Mertens, I.; Sagaert, X.; Prenen, H.; Landuyt, B., *et al.* (2013). Analysis of the formalin-fixed paraffin-embedded tissue proteome: pitfalls, challenges, and future perspectives. *Amino Acids* 45, 205–218.
- Marchione, D.M.; Ilieva, I.; Garcia, B. A.; Pappin, D. J.; Wilson, J. P. and Wojcik, J. B. (2019). HYPERsol: flash-frozen results from archival FFPE tissue for clinical proteomics. *BioRxiv*, 632315.
- Fowler, C.B.; O'Leary, T.J. and Mason, J.T. (2014). Improving the Proteomic Analysis of Archival Tissue by Using Pressure-Assisted Protein Extraction: A Mechanistic Approach. *J. Proteomics Bioinform.*; 7(6): 151-157.
- Davalieva, K.; Kiprijanovska, S. and Polenakovic, M. (2014). Assessment of the 2-d gel-based proteomics application of clinically archived formalin-fixed paraffin embedded tissues. *The Protein Journal. Apr.*, 33(2): 135-142.
- Matsukawa, S.; Morita, K.; Negishi, A.; Harada, H.; Nakajima, Y.; Shimamoto, H., *et al.* (2014). Galectin-7 as a potential predictive marker of chemo- and/or radiotherapy resistance in oral squamous cell carcinoma. *Cancer Medicine. Apr.*, 3(2): 349-361.
- Kokkat, T.J.; Patel, M.S.; McGarvey, D.; LiVolsi, V.A. and Baloch, Z.W. (2013). Archived formalin-fixed paraffin-embedded (FFPE) blocks: A valuable underexploited resource for extraction of DNA, RNA, and protein. *Biopreserv Biobank.*; 11(2): 101-106.
- Föll, M.C.; Fahrner, M.; Oria, V. O.; Kühn, M.; Biniossek, M. L.; Werner, M., *et al.* (2018). Reproducible proteomics sample preparation for single FFPE tissue slices using acid-labile surfactant and direct trypsinization. *Clinical Proteomics*, 15(1): 1–15.
- Nesvizhskii, A.I.; Keller, A.; Kolker, E. and Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* 75:4646-4658.
- Mi, H.; Muruganujan, A.; Ebert, D.; Huang, X. and Thomas, P.D. (2019). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.*;47(D1): D419-D426.
- Sprung, R. W.; Brock, J. W. C.; Tanksley, J. P.; Li, M.; Washington, M. K.; Slebos, R. J. C., *et al.* (2009). Equivalence of protein inventories obtained from formalin-fixed paraffin-embedded and frozen tissue in multidimensional liquid chromatography-tandem mass spectrometry shotgun proteomic analysis. *Molecular and Cellular Proteomics*, 8(8): 1988–1998.
- Tanca, A.; Pagnozzi, D.; Falchi, G.; Tonelli, R.; Rocca, S.; Roggio, T., *et al.* (2011). Application of 2-D DIGE to formalin-fixed, paraffin-embedded tissues. *Proteomics*, 11(5): 1005–1011.
- Wolff, C.; Schott, C.; Porschewski, P.; Reischauer, B. and Becker, K. F. (2011). Successful protein extraction from over-

- fixed and long-term stored formalin-fixed tissues. *PLoS ONE*, 6(1): 1–8.
20. Hood, B. L.; Darfler, M. M.; Guiel, T. G.; Furusato, B.; Lucas, D. A.; Ringeisen, B. R., *et al.* (2005). Proteomic analysis of formalin-fixed prostate cancer tissue. *Molecular and Cellular Proteomics*, 4(11), 1741–1753.
 21. Shi, S. R.; Liu, C.; Balgley, B. M.; Lee, C. and Taylor, C. R. (2006). Protein extraction from formalin-fixed, paraffin-embedded tissue sections: Quality evaluation by mass spectrometry. *Journal of Histochemistry and Cytochemistry*, 54(6): 739–743.
 22. Nirmalan, N. J.; Hughes, C.; Peng, J.; McKenna, T.; Langridge, J.; Cairns, D. A., *et al.* (2011). Initial development and validation of a novel extraction method for quantitative mining of the formalin-fixed, paraffin-embedded tissue proteome for biomarker investigations. *Journal of Proteome Research*, 10(2): 896–905.
 23. Paulo, J. A.; Lee, L. S.; Banks, P. A.; Steen, H. and Conwell, D. L. (2012). Proteomic analysis of formalin-fixed paraffin-embedded pancreatic tissue using liquid chromatography tandem mass spectrometry. *Pancreas*, 41(2): 175–185.
 24. Tanca, A.; Abbondio, M.; Pisanu, S.; Pagnozzi, D.; Uzzau, S. and Addis, M. F. (2014). Critical comparison of sample preparation strategies for shotgun proteomic analysis of formalin-fixed, paraffin-embedded samples: Insights from liver tissue. *Clinical Proteomics*, 11(1).
 25. Gustafsson, O. J. R.; Arentz, G. and Hoffmann, P. (2015). Proteomic developments in the analysis of formalin-fixed tissue. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1854(6): 559–580.
 26. Nirmalan, N. J.; Harnden, P.; Selby, P. J. and Banks, R. E. (2009). Development and validation of a novel protein extraction methodology for quantitation of protein expression in formalin-fixed paraffin-embedded tissues using western blotting. *Journal of Pathology*, 217, 497–506.
 27. Tanca, A.; Addis, M.F.; Pagnozzi, D.; Cossu-Rocca, P.; Tonelli, R.; Falchi, G., *et al.* (2011). Proteomic analysis of formalin-fixed, paraffin-embedded lung neuroendocrine tumor samples from hospital archives. *Journal of Proteomics*. Mar; 74(3): 359-370.