

Evaluation of regular UPLC/MS system for experimental and clinical proteomics

Engin KOÇAK¹ , Sacide ALTINÖZ^{1*} 

¹ Department of Analytical Chemistry, Faculty of Pharmacy, Hacettepe University, Ankara, Turkey.

* Corresponding Author. E-mail: saltinoz@hacettepe.edu.tr (S.A.); Tel. +90-546-720-23-52

Received: 23 December 2019 / Revised: 14 April 2020/ Accepted: 26 April 2020

ABSTRACT: The objective of this study was to investigate if UPLC/MS system without chip or Jetstream technology could be used in proteomics. A regular UPLC column (2.1 mm × 150 mm, 1.7 μm) was used to separate peptides at a flow rate of 0.200 mL/min. First, bovine serum albumin samples were analyzed with intra- and inter-day (11 days) experiments. In *Escherichia coli* experiments, 518 proteins were identified and subsequently investigated with PANTHER gene list and KEGG pathway. Results demonstrated that UPLC/MS provides detailed information regarding cellular process. Finally, we analyzed human plasma sample to observe if UPLC/MS system could be used to analyze body fluids. We identified 87 proteins in the plasma sample, and the identified proteins were investigated in terms of molecular biology. The results showed that UPLC/MS system could provide promising results for experimental and clinical proteomics.

KEYWORDS: Chromatography; maxquant; proteomics; *E. coli*; plasma.

1. INTRODUCTION

A post-genomic science, proteomics aims to analyze cellular processes at the level of proteome, enabling the study of phenotypes, which depends on the genotype and the environment of the organism. In recent years, several analytical methods such as NMR and mass spectroscopy (MS) have been developed to investigate structure, modifications, expression level, and interactions of proteins in organisms [1-3].

MS, which provides greater accuracy and sensitivity than other analytical techniques, has emerged as the primary technique for proteomics analyses [4-6]. MS-based proteomics approach, which constitutes protein digestion with proteolytic enzymes, such as trypsin, separation of peptides, and MS analysis, offers opportunity for the analysis of hundreds of proteins within a short duration [7-10].

Liquid chromatography (LC) enables efficient separation of peptides before the MS analysis and provides greater sensitivity in proteomics; Thus, the less abundant peptides could be analyzed using MS [10-13] In particular, nano-LC, which is a miniature form of a regular HPLC, has served as an important tool in protein analysis for 20 years [14, 15]. Nano-LC systems are compatible with the Nano-ESI interface; and the flow rate of this combined system can be reduced to 200 nL/min with a 75-μm inner diameter LC column. The primary advantage of Nano-LC is that it provides high sensitivity and is thus useful for the analyses of the less abundant proteins in complex mixtures [16, 17]. Despite the advantage of high sensitivity, Nano-LC systems present certain issues, which include leakage, high backpressure, and dead volume. These problems lead to low reproducibility and robustness [18, 19]. Several methods have been tested to overcome these challenges, and as an alternative to Nano-LC systems, ultra-performance liquid chromatography (UPLC)-based proteomics analysis with jet stream and chip technology has been considered. Studies have confirmed that the yield of the said UPLC systems could be comparable with that of Nano-LC systems when analyzing unlimited samples [20- 22]. Also, UPLC/MS combined analytical systems are used more commonly than nano-LC/MS in analytical chemistry laboratories around the world; and effective use of UPLC/MS system in proteomics will encourage more scientists to explore proteomics.

In this study, a regular UPLC/MS system without jet stream or chip technology was evaluated for proteomics. There were three steps in the experimental procedure. In the First step, bovine serum albumin (BSA) was taken as the standard protein to understand the system's suitability for proteomics. Tryptic peptides of BSA separated in UPLC system, then analyzed in MS. The retention times of peptides, matching scores,

How to cite this article: Koçak E, Altinöz S. Evaluation of regular UPLC/MS system for experimental and clinical proteomics. J Res Pharm. 2020; 24(4): 582-592.

peptide intensities, and BSA intensity were analyzed with intra-day and inter-day experiments to observe the reproducibility of the system. In second step, we tried to ascertain if UPLC/MS system could be employed to analyze complex proteome structure cell or bacteria; and for this purpose, *Escherichia coli* (*E. coli*) proteome was analyzed with intra- and inter-day experiments. Protein identification and quantification yield were evaluated for UPLC/MS system. Furthermore, the identified proteins were classified in terms of molecular function, cellular component, and biological process by using panther gene list analysis tool. The final step was the analysis of human plasma proteome without any depletion method. Human plasma is the most significant specimen for conducting clinical studies as well as for observing the performance of UPLC/MS system. Most abundant proteins in plasma were identified without difficulty; also, this method allowed the identification of several low-abundance proteins. All identified proteins were analyzed with gene list, then the yield of UPLC/MS system was evaluated. Significantly, intensities of proteins were calculated with label-free quantification algorithm, and the results demonstrated that UPLC/MS method has high reproducibility for proteomics analysis of plasma samples.

The results of the experiment showed that UPLC/MS has high reproducibility for the analysis of single proteins or complex protein mixtures, but UPLC/MS system has lower sensitivity than nano-LC/MS system. We believe that this study will contribute to further proteomics studies.

2. RESULTS

2.1. BSA analysis

2.1.1. Peptide identification

In our work, we focused peptide identification process as a first step of proteomics analysis. In this process, the amino acid sequences first, then peptide structure were identified by using *in silico* libraries. In Figure 1, there is an example of the identification of a peptide. Recorded MS/MS data was compared with the database to match b- and y-fragment ions from N-terminus and C-terminus. Each peptide has a matching score, which is important parameter to show identification quality.

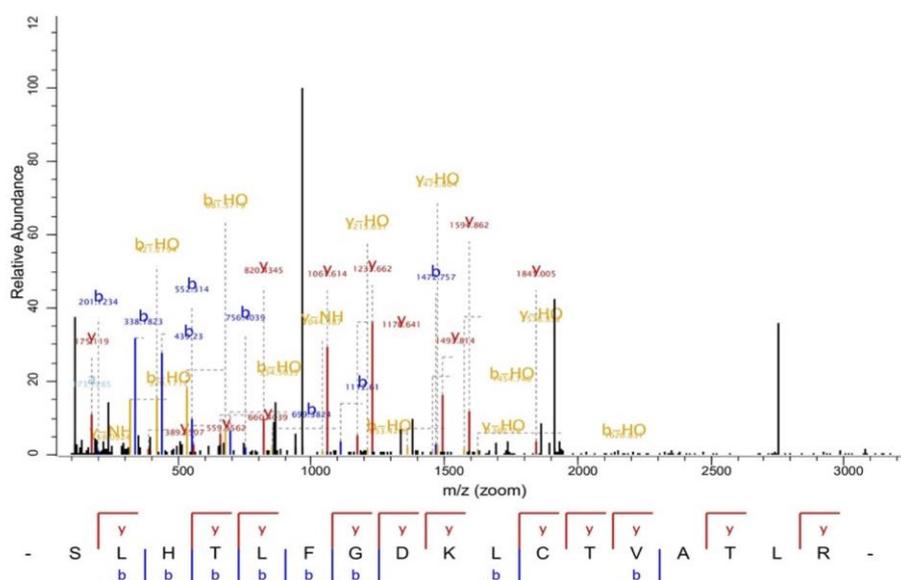


Figure 1. Recorded MS/MS data was compared with *in silico* library. Identified -b and -y ions were showed in red and blue.

2.1.2. Gradient time

In BSA analysis, two different gradient systems were evaluated (45 min and 70 min). As expected, the longer gradient system (70 min) yielded a greater peptide number (34 peptides) and protein coverage (56%) of BSA. The 45 min gradient yielded 26 peptides, and the identified BSA coverage was 41%. In other experiments, we used 70 min gradient system (Figure 2).

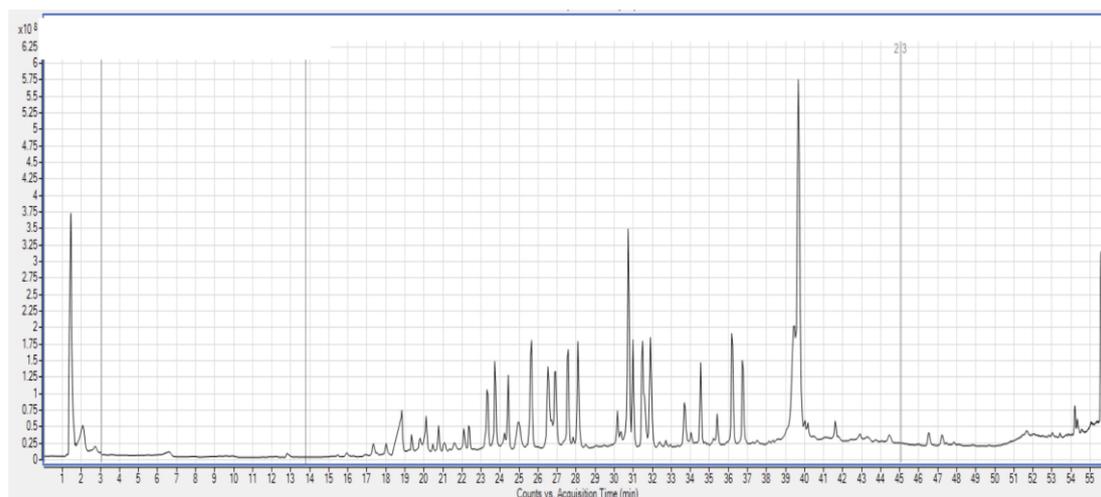


Figure 2. Chromatogram of tryptic peptides of BSA in 70 min gradient.

2.1.3. Loading capacity

To obtain optimum loading capacity for BSA analysis, 0.5, 1, 5, 10 and 20 μg tryptic peptides were injected with three replicates. Average number of identified peptides were given in supplementary information (S1). Result showed that 34 and 40 peptide identified for 10 and 20 μg tryptic peptide loading. Moreover, 10 and 20 μg tryptic peptide loading provided more than 50% BSA coverage.

2.1.4. Reproducibility of BSA experiment

In reproducibility experiments, number of identification peptide, retention time and matching scores were observed beyond 11 days. In our experiments, 10 μg tryptic peptides of BSA were injected and analyzed for each parameter.

Identified number of peptides were analyzed to observe reproducibility of identification process. we found that 30 common peptides were identified beyond 11 days. These peptides constituted 50.1% of BSA coverage. We also evaluated retention times of 30 peptides every day. The average relative standard deviation was calculated as 0.18%. In addition, matching scores of peptides were observed beyond 11 days and average matching score was calculated as 72.6. Average relative standard deviation was found as 15.42%.

Peptide and protein intensities were also calculated by using Maxquant label free quantification algorithm. Average relative standard deviation of 30 peptides were calculated as 16.70%. The relative standard deviation of BSA intensity was found as 14.27%.

2.2. *E. coli* proteome analysis

We tried to identify *E. coli* proteins, which is a good Prokaryote model for proteomics. 20, 40 and 80 μg of tryptic peptides of *E. coli* were injected to UPLC/MS system. For accurate identification, at least two unique peptide should be detected and also proteins should be identified at least two of three replicates.

We identified 224, 382, 518 proteins by loading 20, 40 and 80 μg tryptic peptides. In our experiments, we tried to load more than 80 μg tryptic peptides but observed carryover problems; consequently, we accepted that 80 μg is the maximum level of tryptic peptides and used the same for inter day experiments. In Figure 3, total ion chromatography of *E. coli* peptides were shown for 80 μg tryptic peptides. Average matching scores of 518 proteins was calculated as 46.2.

We investigated 518 proteins by using Panther gene list analysis tool to observe the yield of UPLC/MS system in terms of molecular biology. The proteins were grouped according to protein class, biological process, and cellular component. The result showed that we could identify 15 different classes of proteins (Figure 4a). We analyzed proteins in terms of biological process in which the proteins were involved (Figure 4b). Results showed that we could mostly identify proteins that are in metabolic and cellular process. Finally, we identified proteins according to their location in the cell (Figure 4c).

In this study, the identified proteins were assessed with KEGG pathway analysis tool (Supplementary information S2).

We also investigated whether UPLC/MS-based proteomic results could be used for semi-quantitative studies. We used label-free quantification (LFQ) algorithm, which is a general algorithm for relative

quantification, to calculate intensity. Multi scatter plot analysis was employed to observe reproducibility of the quantification process. We calculated and evaluated LFQ intensity of each protein for three technical replicates by multi scatter plot analysis. R squared values in multi scatter plot analysis were calculated as 0.985 between injections 1 and 2, 0.929 between injections 1 and 3, and 0.950 between injections 2 and 3. These results were satisfactorily good and showed high reproducibility for protein quantification (Figure 5).

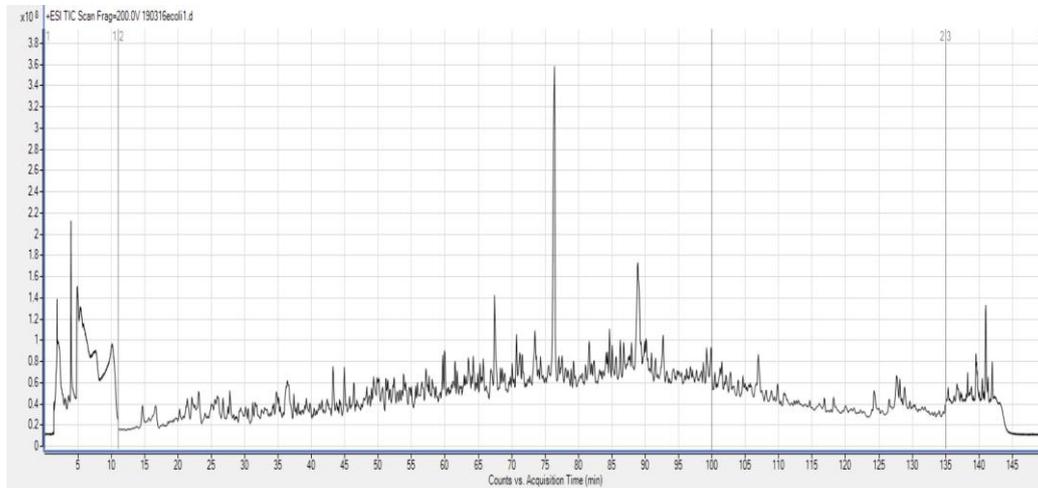


Figure 3. Total ion chromatogram of 80 µg of *E. coli* tryptic peptides.

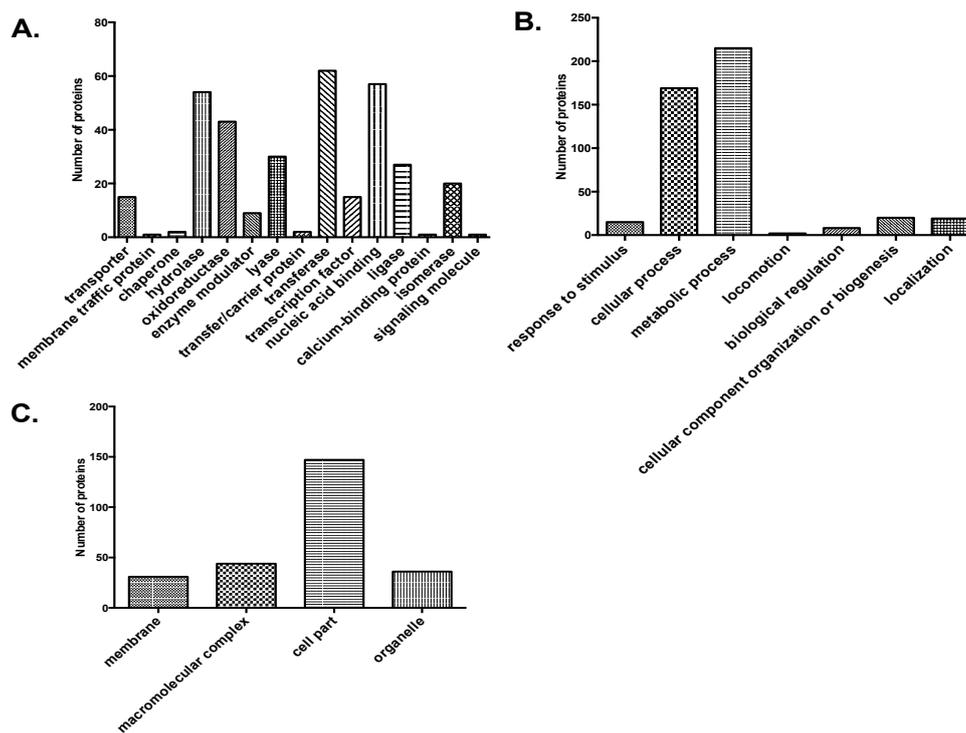


Figure 4. Identified proteins were analyzed by Panther gene list analysis tool in terms of protein classes (a), biological process (b), and cellular part (c).

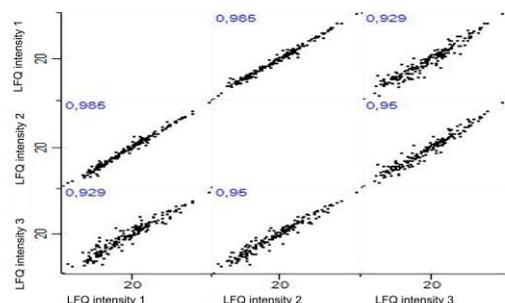


Figure 5. Multi scatter plot analysis of protein intensities and R^2 values for three replicates.

Inter-day experiments with *E. coli* were conducted for three days. Each day, 80 μg tryptic peptides were analyzed with three technical replicates, and the results showed that 497 common proteins were identified over three days. This was a sign of high reproducibility and robustness of UPLC/MS system.

2.3. Human plasma analysis

We analyze human plasma proteome by using UPLC/MS system. 80 μg tryptic peptide of plasma was analyzed with three technical replicates. The total ion chromatogram of peptides of plasma proteins are shown in Figure 6. For accurate identification, at least two unique peptide should be detected and also proteins should be identified at least two of three replicates as well as *E. coli* proteomics.

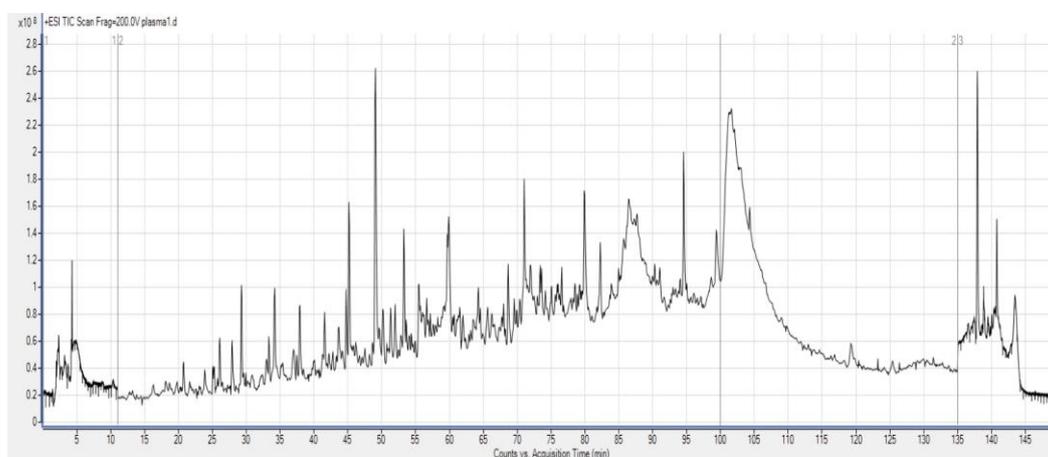


Figure 6. Total ion chromatogram of peptides of plasma proteins.

Proteins were identified by the process of matching with UniprotK Human database in fasta format, and 87 proteins were identified. In the supplementary data (S3), the identified proteins were assigned peptide numbers.

Identified proteins were evaluated by gene list analysis. Results showed that 13 different protein classes could be analyzed, and these proteins were involved in 10 different biological processes. (Figure 7).

Protein intensities were calculated by using Maxquant. The reproducibility of this process was evaluated with multi scatter plot analyses (Figure 8). The results showed high R^2 value and reproducibility between technical replicates.

Inter-day experiments were conducted for three days. 61 proteins were identified every day.

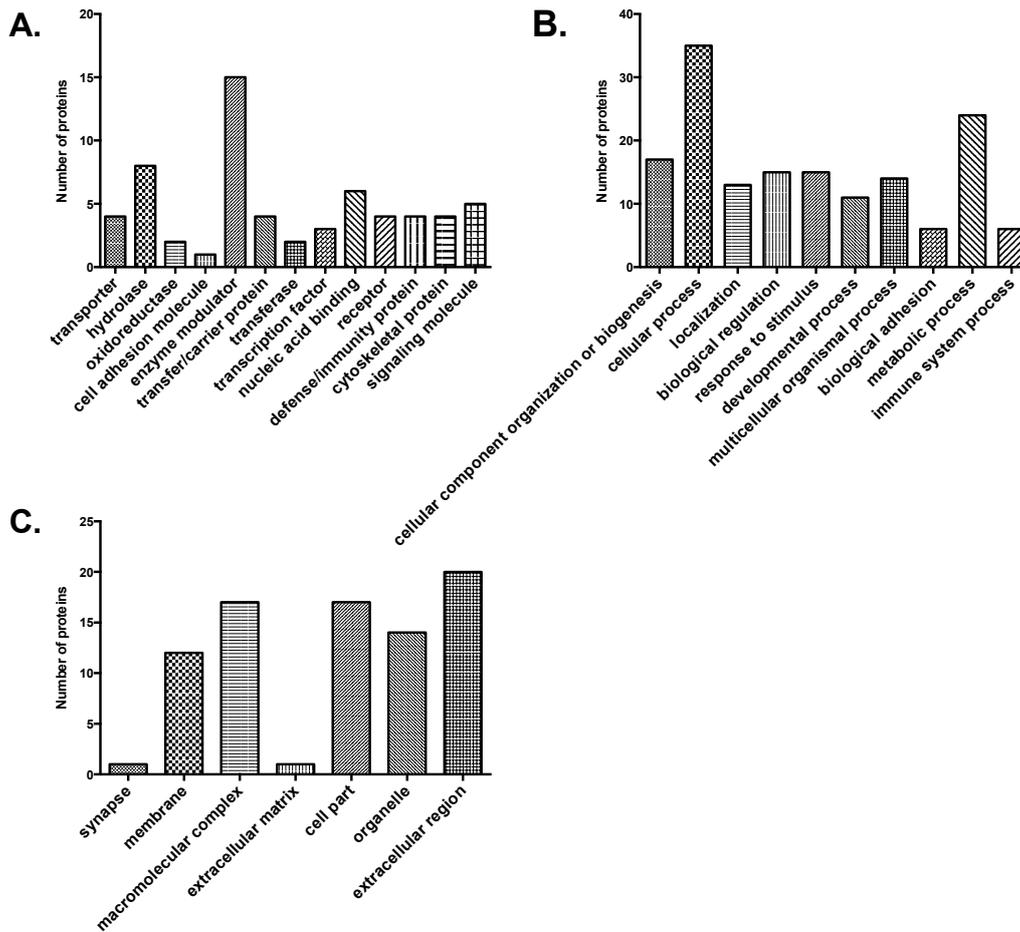


Figure 7. Identified proteins were analyzed by Panther gene list analysis tool in terms of protein classes (a), biological process (b), and cellular part (c).

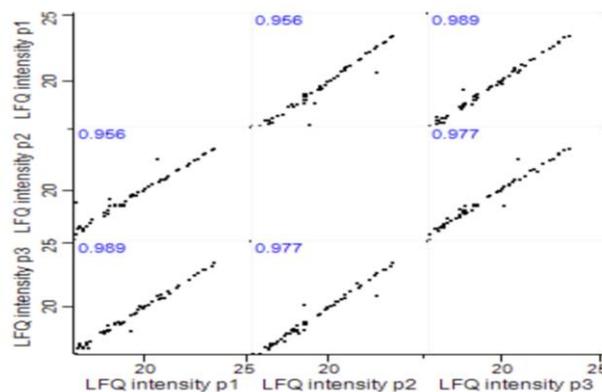


Figure 8. Multi scatter plot analysis of protein intensities and R² values.

3. DISCUSSION

In recent years, proteomics has been emerged as a key discipline to understand the genetic background of cellular processes and also environmental factors that influence organism phenotypes. Proteomics has especially drawn attention for clinic studies. In clinical studies, proteomics have been used to interpret the pathological mechanisms of diseases and to determine the therapeutic mechanisms of these active agents. Moreover, the detection of biomarkers in body fluids for early diagnosis of diseases such as cancer is still an

important research area in proteomics. For this purpose, UPLC/MS system was chosen to analyze standard BSA protein, *E. coli*, and human plasma proteome to evaluate its suitability for experimental and clinical proteomics.

3.1. BSA analysis

BSA has been considered as the golden standard for proteomics. Therefore, we used BSA in this study to evaluate the reproducibility and reliability of UPLC/MS system. First, we tried to observe matching of MS/MS data with theoretical *in silico* data. We found b- and y- fragments ion with 20 ppm mass tolerance (Figure 1).

In our work, we evaluate several optimization parameter. We used two different gradient system and observed 70 min gradient gives better results in terms of identified peptide numbers. Loading capacity is among the key parameters for proteomics. Nano-LC/MS system possesses high sensitivity, and nano- and pico-gram levels are sufficient for reliable protein analysis. UPLC/MS system has lower sensitivity as compared to nano LC; therefore, the amount of tryptic peptides that were injected UPLC/MS system should be increased. 10 and 20 µg tryptic peptide loading yielded more identified peptides, allowing us to identify more than 50% of BSA (Supplementary information S1). On this basis, 10 µg of BSA peptides was used for the reproducibility analysis.

Reproducibility is the principle parameter for proteomics as LC/MS systems have complex instrumentation that can cause variability; moreover, sample complexity is another variability source that affects the identification and quantification of peptides and proteins. Therefore, we evaluated reproducibility parameter, firstly with BSA peptides, then with more complex samples (*E. coli* proteome and human plasma). To understand system reproducibility in BSA experiments, we analyzed 10 µg tryptic peptides of BSA for 11 days. First, we evaluated the identification process. The Results showed that we could identify 30 common peptides beyond 11 days and the peptides constituted 50.1% of BSA coverage. Retention time reproducibility is an important parameter for reliable protein identification and quantification, and it was investigated with 30 peptides, which were recorded daily beyond 11 days. The average relative standard deviation of the retention times of 30 peptides was calculated as 0.18%. These results showed that UPLC system possesses high reproducibility for peptide separation. The peptide matching score is a sign of identification quality. The peptide matching scores were analyzed beyond 11 days. Average matching scores were calculated as 72.6. Furthermore, average relative standard deviation of matching scores of 30 common peptides that were identified was calculated as 15.42%.

Peptide and protein intensities were used for semi quantitative proteomics and critical point for clinical studies to find biomarkers and disease mechanisms. In our work, Maxquant tool calculated the intensities of the peptides. The average relative standard deviation of the peptide intensities of 30 common peptides that were identified was calculated as 16.70%.

Reproducibility was also investigated at the protein level, and the calculated intensity of BSA was observed for 11 days. The Maxquant tool was used to determine the protein intensity based on the intensities of the peptides that belong to the protein. The relative standard deviation of the protein intensity for 11 days was calculated as 14.27%.

3.2. *E. coli* proteome analysis

Cell, bacteria, tissue, and body fluid are the most important samples for proteomics: These samples contain thousands of proteins. Simultaneous analysis of all proteins requires high sensitivity and repeatability. Nano-LC/MS system has high sensitivity and can be used by injecting nano-gram level. In UPLC/MS based proteomics, the injected amount of tryptic peptides should be increased to obtain comparable results with nano-LC/MS system.

In *E. coli* proteomics analysis, we loaded 20, 40 and 80 µg tryptic peptides. Results showed that identified protein numbers increased dramatically from 20 to 80 µg. We identified 224, 382, 518 proteins by loading 20, 40 and 80 µg tryptic peptides. Reproducibility of identification process were evaluated with inter day experiments beyond three days. We observed that 497 common proteins were identified over three days.

In gene ontology analysis, we focused on yield of UPLC/MS system for molecular biology. Identified proteins were classified in terms of protein classes, biological process and cellular component. We found that identified proteins were belonged to 15 different classes of proteins especially transferase, hydrolase and transferase proteins were identified commonly (Figure 4a). In biological process classification, we observed that most of the identified proteins involved in metabolic and cellular process. Biological regulation and

biogenesis are other important biological processes for cell proliferation; and we could identify proteins in these biological processes. One of the most interesting results in this study was the analysis of proteins that were in response to stimulus process. This was significant for further studies in microbiology to aid the understanding of the resistance mechanism. The most of identified proteins were in the cell part (cytoplasm). The identification of membrane proteins was the critical point for this study as the analysis of membrane proteins is difficult in proteomics studies. The results showed that UPLC/MS system could provide detailed information to understand cellular process at the level of protein and could be used for experimental proteomics (*in vitro*).

In recent years, pathway analysis has become a key step for interpreting proteomic results. Results showed that cell metabolism (glycolysis, fatty acids, and pentose phosphate pathways) could be investigated in depth by UPLC/MS. Amino acid biosynthesis and ribosome pathways – which are key cellular processes for bacterial proliferation and important targets for a large number of antibiotics – can be easily analyzed with proteomics data. We observed that 21 proteins that are ABC transporters were identified by UPLC/MS. ABC transporters are key elements for exporting or importing molecules. A critical point in pathway analysis was the identification of 14 proteins that are involved in quorum sensing process. Quorum sensing is a process of cell-cell communication that allows bacteria to share information.

Today, semi quantitative proteomics approach is very important to find cellular processes at different conditions. Especially in clinical studies, semiquantitative proteomics is used to find new biomarker and discover pathogenesis of diseases at protein level. In present work, we evaluated UPLC/MS based proteomics if it is suitable for semiquantitative proteomics analysis. We used label free quantification algorithm to find protein intensity. We compared three replicates by using multiscatter plot analysis to observe reproducibility of quantification process. Results showed that UPLC/MS gives highly reproducible results.

3.3. Human plasma analysis

Human plasma, which contains thousands of proteins that come from nearly every cell in the body, is the most important specimen for proteomic analysis. Human plasma has been used in general to find biomarkers for early detection of diseases such as cancer and Alzheimer's syndrome.

In present work, we evaluated UPLC/MS based proteomics for plasma proteomics. We did not use any depletion method before proteomics analysis. Our aim was to analyse most abundant proteins and also less abundant proteins in plasma. Identified proteins list has been given in supplementary information. We identified most abundant proteins like serum albumin, globulins, complement C3, transferrin, fibrinogen. These proteins constitute 90% of all proteome in plasma. Also, we identified low abundant proteins. Inter day experiments were carried out for three day to observe reproducibility of identification process. Most abundant proteins were identified easily every day. Moreover 61 common proteins were identified everyday and most of them were less abundant proteins.

Identified proteins were evaluated with gene ontology analysis in terms of protein classes, biological processes and cellular components. Identified proteins belonged to 13 different protein classes and these proteins involve in different biological process like cellular component organization process, metabolic process, biological regulation. When we evaluated identified proteins according to their cellular location, we found that most of them were in extracellular region, cellular part, macromolecular complex. These results showed that UPLC/MS based proteomics gives important information about the plasma proteome structure.

Finally we evaluated protein quantification processes as well as *E. coli* experiments. Results showed that label free quantification method was very reproducible for plasma proteomics analysis. It is very important for clinical experiments.

4. CONCLUSION

Proteomics is a growing area in post-genomic studies useful for the characterization and quantification of systems of interest at a molecular level in the health and clinical fields. These areas require a fast and reliable analysis of complex systems; therefore, LC/MS is the most adequate methodology for the proteomics studies. In this study we tried to investigate suitability of UPLC/MS method for proteomics analysis. Results showed that UPLC/MS system has low sensitivity than nano-LC/MS systems and this caused identification of less number of proteins. Although there is a sensitivity problem, UPLC/MS system offers high reproducibility and reliable analysis. As a result UPLC/MS system could be used as an alternative system for proteomics studies. We believe that this work will contribute further proteomics studies

5. MATERIALS AND METHODS

5.1. Sample preparation for proteomics analysis

BSA (Sigma Aldrich) was dissolved in 100 mM ammonium bicarbonate (Sigma Aldrich) that contained 20% methanol (Sigma Aldrich) (v/v), pH: 8.5. Disulfide bonds in BSA were reduced using 200 mM DDT (Sigma Aldrich) at 56°C for 15 min. Iodoacetamide (Bio-Rad) was used for alkylation of BSA in the dark at room temperature for 30 min. BSA was digested by adding trypsin (Sigma Aldrich) at a 1:100 trypsin: protein ratio (w/w), incubating for 16 h at 37 °C. The samples were mixed by gentle vortexing, and the peptide mixtures were dried using a vacuum centrifuge. Finally, the peptides were dissolved in acetonitrile with 0.1% (v/v) formic acid (Sigma Aldrich).

E. coli proteins, commercial sample (Bio-Rad), were suspended in 100 mM ammonium bicarbonate solution. 200 mM DTT and 100 mM iodoacetamide were used as a reducing and alkylation agents. Proteins were incubated with trypsin at 1:100 (w/w) for 16 hours.

300 µg protein containing commercial plasma sample (Sigma) was mixed with 400 µL methanol (Sigma). The second step was to add 100 µL chloroform (Riedel-dehaen). Finally, 300 µL water was added to the mixture. The mixture was vortexed for 2 minutes, then centrifugation was performed for 5 minutes at 14000 g. This extraction method produced a 3-phase system. Polar metabolites were separated at upper phase (methanol and water); non-polar lipids were separated at the lower organic phase. Protein pellet was observed between two phases. Upper phase was removed, and 400 µL methanol was added, then centrifugation was performed for 3 minutes at 14000 g. Protein pellets were accumulated at the bottom of the tube. The supernatant was removed, and the protein pellet was left at room temperature to completely remove the organic solvents.

The proteins were suspended in 100 mM ammonium bicarbonate solution. 200 mM DTT and 100 mM iodoacetamide was used as reducing and alkylation agents. Finally, the proteins were incubated with trypsin at 1:100 ratio (w/w).

5.2. Separation of Peptides Using UPLC system

BSA, *E. coli* and plasma peptides separated in reverse phase LC system. Peptides were separated in C18 column (Agilent, Zorbax, 2.1 mm × 150 mm, 1.7 µm, 300 Å, flow rate 0.2 mL/min). Mobile phase A consisted of 0.1% formic acid in water, and mobile phase B consisted of 0.1% formic acid in acetonitrile. The peptides were separated using the Agilent 1290 LC system.

In BSA experiments, two different gradients were taken to determine the optimum separation condition for BSA peptides. Gradient 1 was 45 min, and gradient 2 was 70 min.

In 45 min gradient, mobile phase B (acetonitrile) content was increased from 3% to 55% over 35 minutes and then increased up to 85% over 2 minutes, held for 2 min, then followed by a ramp back to 3% in 1 min. In 70 gradient, Mobile phase B content was increased from 3% to 55 over 50 minutes and then increased to 85% over 5 minutes and held for 3 min, then followed by a ramp back to 3% in 2 min.

In *E. coli* experiments, 20, 40 and 80 µg of tryptic peptides were injected to LC/MS system with three technical replicates. Mobile phase A consisted of 0.1% formic acid in water, and mobile phase B consisted of 0.1% formic acid in acetonitrile. In the 150 min gradient system, the mobile phase B (acetonitrile) content was increased from 3% to 55% over 130 minutes, then increased up to 85% over 5 minutes and held for 2 min, then followed by a ramp back to 3% in 1 min [23].

80 µg tryptic peptides of plasma sample were injected to LC/MS system with three technical replicates. The same LC/MS experimental procedure, which was used for *E. coli* peptides, was used for separation and analysis of peptides in LC/MS system.

5.3. Q-TOF-MS analysis of the BSA peptides

An Agilent Q-TOF-MS 6530 system was used to analyze the peptides. Once separated, they passed through the ESI source and were ionized in the positive mode. The capillary voltage was adjusted at 4000 V with a drying temperature of 350°C. Auto MS/MS data were recorded between 300 and 1400 m/z above the 1500 count threshold. The MS-MS cycle time was 3.63 s, and the maximum ion number was selected to be 3. Ion charge states were +2, +3 and +4, and the fragmentation energy was adjusted to 45 V.

5.4. Maxquant data analysis

In BSA experiments, The recorded MS/MS data were investigated by using the Maxquant proteomics tool. For the qualitative analysis, the UniProtK bos-taurus database was downloaded, and Trypsin was

selected as the digestion enzyme. As fixed modification, carbamidomethylation of cysteine (C) was selected; and for the variable modification, oxidation on methionine (O) was selected. For the first matching process, mass tolerance was fixed to 20 ppm; and for the second matching process, 20 ppm was selected. The FDR ratio for peptides was fixed at 0.01.

In *E. coli* experiments, MS and MS-MS parameters, which are used to analyze BSA peptides, were used for peptide analysis. For data processing, Maxquant software was used to identify proteins and calculate label-free protein intensity. UniProtK *E. coli* database was used for protein identification. As fixed modification, carbamidomethylation of cysteine (C) was selected; and for the variable modification, oxidation on methionine (O) was selected. The mass tolerance was fixed to 20 ppm for the first and primary matching process, and the FDR ratio for peptides was fixed at 0.01.

In plasma sample analysis, Maxquant proteomics tool was used for protein identification, and quantification. UniprotK Human database was used for the matching process as well. In protein identification process, carbamidomethylation of cysteine (C) and oxidation on methionine (O) was selected as fixed and variable modifications. FDR value was adjusted at 0,01. 20 ppm mass tolerance was used for the first and main matching process. Gene list analysis (Panther software) was used to observe the yield of the system in terms of molecular biology.

5.5. Loading capacity optimization for BSA

Increasing amounts of protein (0.5, 1, 5, 10 and 20 µg) were injected with three technical replicates. The identified peptide numbers and the BSA content were calculated to find optimum loading capacity.

5.6. Reproducibility

Reproducibility parameter was studied at the level of peptides and proteins with intra- and inter-day experiments. Identified peptide numbers, retention times, matching scores, and intensities were evaluated to understand reproducibility and the quality of system at peptide level.

Reproducibility was also analyzed at protein level. BSA intensity as well as the relative standard deviation was calculated. Inter-day experiments were conducted for 11 days.

Acknowledgements: This study has been supported Hacettepe University BAP coordination unit (THD-2017-15330).

Author contributions: Concept – E.K., S.A.; Design – E.K., S.A.; Supervision – S.A., E.K.; Resources – E.K., S.A.; Materials – E.K., S.A.; Data Collection and/or Processing – E.K., S.A.; Analysis and/or Interpretation – E.K., S.A.; Literature Search – S.A., E.K.; Writing – S.A., E.K.; Critical Reviews – S.A., E.K.

Conflict of interest statement: The authors declared no conflict of interest.

REFERENCES

- [1] Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. *Annu Rev Biochem.* 2001; 70: 437-473. [\[CrossRef\]](#)
- [2] Phizicky E, Bastiaens PI, Zhu H, Snyder M, Fields S. Protein analysis on a proteomic scale. *Nature.* 2003; 422(6928): 208-215. [\[CrossRef\]](#)
- [3] Wright PC, Noirel J, Ow SY, Fazeli A. A review of current proteomics technologies with a survey on their widespread use in reproductive biology investigations. *Theriogenology.* 2012; 77(4): 738-765. [\[CrossRef\]](#)
- [4] Mann M. Origins of mass spectrometry-based proteomics. *Nat Rev Mol Cell Biol.* 2016; 17(11): 678-685. [\[CrossRef\]](#)
- [5] Doerr A. Mass spectrometry-based targeted proteomics. *Nat Methods.* 2013; 10(1): 23. [\[CrossRef\]](#)
- [6] Resing KA, Ahn NG. Proteomics strategies for protein identification. *FEBS Lett.* 2005; 579(4): 885-889. [\[CrossRef\]](#)
- [7] Sidoli S, Lin S, Karch KR, Garcia BA. Bottom-Up and Middle-Down Proteomics Have Comparable Accuracies in Defining Histone Post-Translational Modification Relative Abundance and Stoichiometry. *Anal Chem.* 2015; 87(6): 3129-3133. [\[CrossRef\]](#)
- [8] Olsen JV, Mann M. Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. *Proc Natl Acad Sci U S A.* 2004; 101(37): 13417-13422. [\[CrossRef\]](#)
- [9] Liu J, Bell A, Bergeron J, Yanofsky C, Carrillo B, Beaudrie C, Robert E. Methods for peptide identification by spectral comparison. *Proteome SCI.* 2007; 5. [\[CrossRef\]](#)

- [10] Rabilloud T, Lelong C. Two-dimensional gel electrophoresis in proteomics: A tutorial. *J proteomics*. 2011; 74(10): 1829-1841. [\[CrossRef\]](#)
- [11] Galeva N, Altermann M. Comparison of one-dimensional and two-dimensional gel electrophoresis as a separation tool for proteomic analysis of rat liver microsomes: Cytochromes P450 and other membrane proteins. *Proteomics*. 2002; 2(6): 713-722. [\[CrossRef\]](#)
- [12] Chenau J, Michelland S, Sidibe J, Seve M. Peptides OFFGEL electrophoresis: a suitable pre-analytical step for complex eukaryotic samples fractionation compatible with quantitative iTRAQ labeling. *Proteome SCI*. 2008; 6(1): 1-8. [\[CrossRef\]](#)
- [13] Fukao Y, Yoshida M, Kurata R, Kobayashi M, Nakanishi M, Fujiwara M. Peptide Separation Methodologies for In-Depth Proteomics in Arabidopsis. *Plant cell physiol*. 2013; 54(5): 808-815. [\[CrossRef\]](#)
- [14] Saz JM, Marina ML. Application of micro- and nano-HPLC to the determination and characterization of bioactive and biomarker peptides. *J Sep Sci*. 2008; 31(3): 446-458. [\[CrossRef\]](#)
- [15] Wilson SR, Vehus T, Berg HS, Lundanes E. Nano-LC in proteomics: recent advances and approaches. *Bioanalysis*. 2015; 7(14): 1799-1815. [\[CrossRef\]](#)
- [16] Waanders LF, Chwalek K, Monetti M, Kumar C, Lammert E, Mann M. Quantitative proteomic analysis of single pancreatic islets. *Proc Natl Acad Sci USA*. 2009; 106(45): 18902-18907. [\[CrossRef\]](#)
- [17] Qian WJ, Jacobs JM, Liu T, Camp DG, Smith RD. Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Mol Cell Proteomics*. 2006; 5(10): 1727-1744. [\[CrossRef\]](#)
- [18] Noga M, Sucharski F, Suder P, Silberring J. A practical guide to nano-LC troubleshooting. *J Sep Sci*. 2007; 30(14): 2179-89. [\[CrossRef\]](#)
- [19] Kocher T, Pichler P, Swart R, Mechtler K. Analysis of protein mixtures from whole-cell extracts by single-run nanoLC-MS/MS using ultralong gradients. *Nat Protoc*. 2012; 7(5): 882-890. [\[CrossRef\]](#)
- [20] Percy AJ, Chambers AG, Smith DS, Borchers CH. Standardized protocols for quality control of MRM-based plasma proteomic workflows. *J Proteome Res*. 2013; 12(1): 222-233. [\[CrossRef\]](#)
- [21] Gonzalez Fernandez-Nino SM, Smith-Moritz AM, Chan LJ, Adams PD, Heazlewood JL, Petzold CJ. Standard flow liquid chromatography for shotgun proteomics in bioenergy research. *Front Bioeng Biotechnol*. 2015; 3: 44. [\[CrossRef\]](#)
- [22] Krokhin OV. Sequence-specific retention calculator. Algorithm for peptide retention prediction in ion-pair RP-HPLC: application to 300- and 100-Å pore size C18 sorbents. *Anal Chem*. 2006; 78(22): 7785-7795. [\[CrossRef\]](#)
- [23] Koçak E. PhD Thesis. Liquid Chromatography-Mass Spectroscopy Based Proteomics Studies on Caco-2 Colon cancer cells. Analytical Chemistry Department, Faculty of Pharmacy, Hacettepe University; 2017.

This is an open access article which is publicly available on our journal's website under Institutional Repository at <http://dspace.marmara.edu.tr>.