Scalable and Automated Plasma Workflow Based on the Thermo Scientific Q Exactive HF-X MS Platform

Jing Wang[1]; Sarah Trusiak[1]; Ryan D. Bomgarden[2]; Sergei Snovida[2]; Emily I. Chen[1], [1]Thermo Fisher Scientific Precision Medicine Science Center, Cambridge, MA; [2]ThermoFisher Scientific, Rockford, IL

ABSTRACT

Purpose – This poster presents a high throughput serum and plasma proteomics analysis workflow for large population cohorts.

Methods – To reduce analytical variability of plasma sample preparation for LC-MS/MS analysis, we automated Thermo Scientific™ EasyPep™ Mini MS Sample Prep Kit using the Hamilton Microlab STARlet liquid handling system with [MPE]^2 positive pressure and evaporation modules. The Evosep LC system was used to run high throughput and automated LC methods. Thermo Scientific™ Q Exactive™ HF-X MS and data-dependent acquisition (DDA) were used to generate quantitative LFQ plasma proteome data. Skyline was used for retention time analysis and Thermo Scientific™ Proteome Discoverer™ 2.3 software was used for database search and post-data analysis.

Results – The automated sample preparation can process 96 samples within 4 hours with ~80% recovery. The throughput of LC analysis for the standardized workflow is > 50 samples per day with 10% overhead to minimize sample carrying over. Around 150 and 200 core proteins (high confident) could be reproducibly identified and quantified for the undepleted serum and the depleted plasma samples, respectively. An example of this workflow applied to small scale depleted plasma lung cancer samples is presented in this poster as well.

INTRODUCTION

Variabilities and dynamic range of protein abundance substantially influence the human plasma proteome analysis. To develop novel markers indicative of diseases using proteomics-based approaches, the plasma workflow has to be high-throughput and robust for hundreds of runs to make a reliable conclusion out of a clinical study. In this study, we developed a standardized high throughput (HT) plasma proteomics analysis workflow focusing on balancing the depth identification and scalability for sampling large population cohorts. The workflow consists of an automated sample preparation method and an Evosep LC system coupled to an Q Exactive HF-X platform. The automated sample preparation method digests and purified peptides from 96 well samples in less than 4.5 hours. The hand-free design of the sample preparation method dramatically increases the throughput and minimizes systematic errors due to pipetting. The positive air pressure mechanism [MPE]^2 is introduced to replace the centrifugation-based mechanism and is optimized for peptide recovery. Pre-set gradients of the Evosep One system allows the user to run 30 samples a day, 60 samples a day or over 100 samples a day. QC sample and system suitability control are included throughout the sample sequence to monitor and assess the instrument performance. Also, a calibration standard peptide mix is spiked in every sample to monitor retention time drifts and peak area variations.

MATERIALS AND METHODS

Material

Reagents and HPLC grade buffers used for proteomics analyses are from Thermo Fisher Scientific. Pierce™ retention time peptide standard (PRTC), HeLa protein digest standard, and EasyPep™ MS sample prepare kit are from Thermo Fisher (Rockford, IL).

Sample Preparation

A commercial human serum sample pooled from several individuals purchased from BioreclamationIVT was processed using the EasyPep Mini MS Sample Prep Kit (A40006) and used as the system suitability control. The lung cancer samples were processed using the Thermo Scientific™ Top 12 Abundant Protein Depletion Spin Columns (85165), followed by the EasyPep Mini MS Sample Prep Kit.

LC-MS Analysis

Peptides from digested samples (HeLa digest standard, serum, and plasma) were separated using a 8cm Evosep column (EV-1064) and the Evosep one LC system. A Thermo Scientific™ EASY-Spray™ adapter (EV-1072) with a stainless steel emitter (EV-1086) was connected to an EASY-Spray ion source equipped Q Exactive HF-X. The mobile phase A is composed of 0.1% formic acid in water (HPLC grade), and the mobile phase B is composed of 0.1% FA in acetonitrile. Peptides were loaded on Evotips based on the manufacture protocol. Data dependent acquisition (DDA) and Evosep manufacture gradient methods were used to acquire LC-MS data.

Data Analysis

Proteome Discoverer 2.3 software was used for searching acquired MS2 spectra against human protein database (UniProt reviewed, December 2018) and post-data analysis (heatmap and PCA plots). 1% FDR was set as the filtering threshold for protein and peptide identification. Skyline was used for retention time and peak area analysis.

RESULTS

1. WORKFLOW DEVELOPMENT

Figure 1. An illustration of the high throughput (HT) plasma workflow with the Evosep LC coupled to a Q Exactive HF-X mass spectrometer.

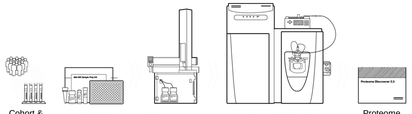


Figure 2. Thermo Scientific 96-well format EasyPep MS Sample Prep on the Hamilton Microlab STARlet liquid handling system.

The deck layout for the workflow is presented as the following. A & B) Troughs contain the EasyPep digestion reagents and peptide clean-up buffers. C) A C-Pac chiller unit holds plasma/serum samples at 4°C. D) A heater/shaker incubates the digestion reaction at 37°C. E) Location houses the 96-well EasyPep column. F) Hamilton's [MPE]^2 positive pressure for peptide purification. G & H) Locations hold pass-through supernatant. I) Evaporator unit dries down eluted peptides.

Digestion set-up 25 mins ⟹ *Digestion incubation* 3 hrs ⟹ *Peptide clean-up* 25 mins ⟹ *Evaporation* 40 mins = ***Total time 4hr 30 mins***

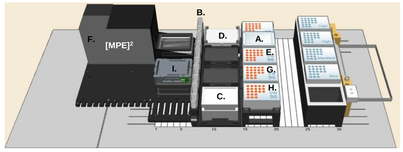


Table 1. Comparison of peptide recovery from the EasyPep MS Mini Sample Prep spin column format vs automated 96-well EasyPep Sample Prep column format.
A.) Undepleted pooled serum was used to compare the peptide recovery efficiency by using the spin column format (manual centrifugation) and automated 96 well format (positive pressure). Triplicate analyses were carried out per method. B) To determine well to well variations of the 96 well column format, 48 undepleted pooled serum (technical replicates) were processed using the automated protocol and the percentage of peptide recovery as well as % CV of column to column and row to row variations are listed in red.

A.

	Peptide clean-up only			Full workflow		
	Peptide Input	Recovery	% CV	Protein Input	Recovery	% CV
Manual spin column	20µg	68.5%	5.7%	45µg	77.9%	10.7%
Automated 96-well column on Hamilton	20µg	72.1%	3.7%	45µg	80.7%	7.2%

B.

	1	2	3	4	5	6	% CV
A	74.0%	74.3%	72.6%	72.4%	83.7%	76.9%	5.6%
B	72.5%	73.8%	66.7%	68.3%	73.8%	82.5%	7.6%
C	68.9%	72.2%	69.7%	70.3%	77.3%	78.6%	5.7%
D	72.0%	71.7%	69.2%	101.7%	76.2%	74.4%	15.6%
E	74.3%	75.1%	72.2%	67.4%	73.1%	81.6%	6.2%
F	70.4%	75.3%	74.3%	67.9%	74.2%	73.8%	4.0%
G	75.6%	73.1%	72.6%	72.8%	77.5%	80.1%	4.1%
H	70.0%	72.5%	76.8%	71.7%	77.1%	77.1%	4.3%
% CV	3.2%	1.8%	4.4%	15.4%	4.4%	4.1%	7.4%

Figure 3 – Precision and variability of liquid dispensing using the Hamilton MicroLab STARlet system versus manual pipetting.

50µl of lysis solution was dispensed by either the Hamilton liquid handling system (n=48) or a manually (n=16). In the box plot, the red line shows the mean, each black dot denotes the volume of lysis solution measured by weight, and the blue lines represent one standard deviation above and below the mean.

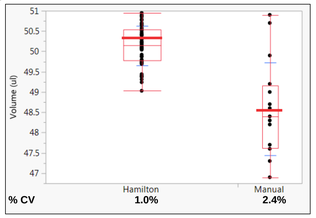


Figure 4 – Analysis of HeLa Protein Digest Standard with the 8cm Evosep Column.

Peptides from 50ng, 100ng, and 250ng of HeLa digest standard were analyzed using the 8cm Evosep column coupled to a Q Exactive HF-X mass spectrometer at the throughput of 30, 60, and 100 samples per day. Bar graphs showed means of protein (1% FDR) and peptides (1% FDR) identified from each concentration and standard deviations from at least 3 runs per method.

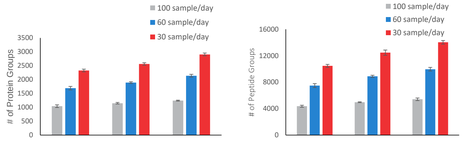


Figure 5. Reproducibility and robustness of the HT plasma workflow with the retention time peptide standard and undepleted pooled serum.

(A & B) 48 runs of undepleted pool serum (500ng) with spiked-in Pierce PRTC standard (50fmol) were evaluated for retention time drift and %CV of the peak area. (C & D) Bar graphs showed means of protein (1% FDR) and peptides (1% FDR) identified from each concentration and standard deviations from 6 runs per method.

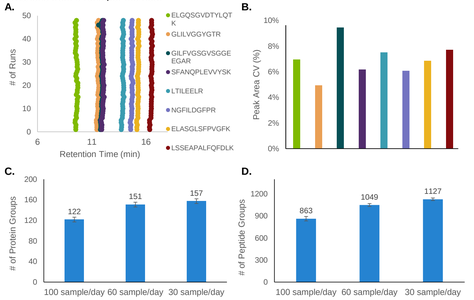
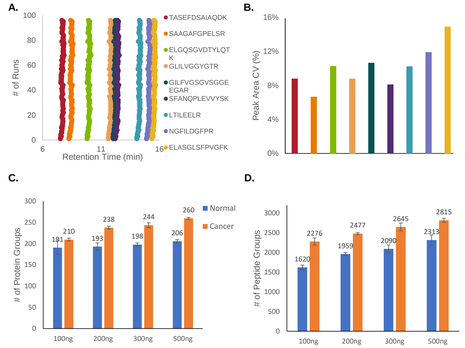


Figure 6 – Characterization of Depleted Plasma with Evosep Column

(A & B) 100 runs of crudely depleted normal and lung cancer plasma (500ng) with spiked-in Pierce PRTC standard (30fmol) were evaluated for retention time drift and %CV of the peak area. (C & D) Bar graphs showed means of protein (1% FDR) and peptides (1% FDR) identified from each concentration and standard deviations from 5 runs per method.



2. WORKFLOW APPLICATION:

Figure 7 – Similarity of protein expression among depleted normal and lung cancer plasma.

Box-and-whisker plots show median and maximum 1.5 interquartile range (IQR) of protein (1% FDR) and peptides (1% FDR) identified from 4 normal plasma replicates (n=10 each replicate) and 5 cancer plasma replicates (n=10 each replicate). Graphic representation of proteomics profiles from 500ng digested depleted plasma using unsupervised hierarchical clustering showed clustering between normal replicates and cancer replicates. Manhattan distance function (the sum of the differences) with a complete linkage method were used to generate the heatmap (C). PCA shows a reduced dimension of data among normal and cancer replicates (D).

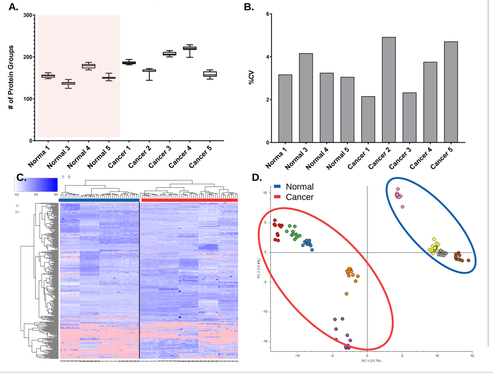
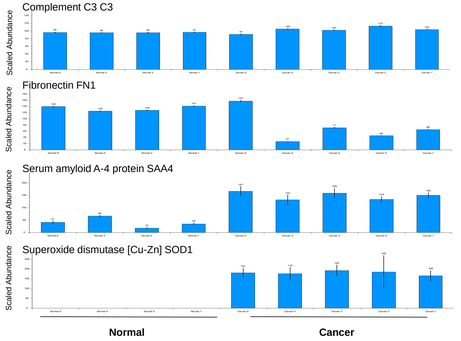


Figure 8 – Example of four proteins detected in the depleted normal and lung cancer plasma.

PD2.3 applies normalization of the total abundance values for each run across all files, equalizing the total abundance between different runs. After aggregating all the normalized abundance values per sample, PD scales the abundance values of each sample so that the average of all samples is 100. The quantitative value for each protein is expressed as scaled abundance. C3 is one of the most abundant proteins in the dataset and show equal abundance among normal and cancer samples. FN1, SAA4, and SOD1 were found vastly reduced in either cancer or normal plasma.



CONCLUSIONS

- The automated EasyPep MS Sample Prep protocol can process 96 serum or plasma samples in 4.5 hours or less with peptide recoveries and CVs on par or better than the spin column format. A large sample run (n=48) using the automated protocol yielded on average 74% peptide recovery rate and 7.4% well to well variations.
- We have evaluated the performance of the Evosep One with the QE HF-X MS with >100 repeated measurements of standard peptides and showed system robustness (3% CV for retention time shift and <15% CV for peak area quantification) while maintaining the sensitivity of current nanoflow LC instruments with HeLa digest standard (1).
- Using the HT plasma workflow, we were able to achieve <6% variations among technical replications of depleted normal and lung cancer plasma samples and identified relevant lung cancer biomarkers (2,3).

REFERENCES

1. Bache, N. et al., *Mol. Cell. Proteomics,* mcp.TIR118.000853
2. W C S Cho et al., Serum amyloid A is elevated in the serum of lung cancer patients with poor prognosis, Br J Cancer. 2010 Jun 8; 102(12): 1731–1735.
3. Papa L. et al., SOD1, an unexpected novel target for cancer therapy, Genes Cancer. 2014 Jan; 5(1-2): 15–21.

ACKNOWLEDGEMENTS

We would like to thank Igor Gojkovic from Hamilton Robotics for scripting the EasyPep Mini MS Sample Prep protocol on the STARlet and for his help with sample testing.

For Research Use Only. Not for use in diagnostic procedures.

TRADEMARKS/LICENSING

PO65497-EN0519S

ThermoFisher SCIENTIFIC