Benchmark of Micro-flow Chromatograph for Robust Proteomics Analysis

[1]Yang Liu, [1]Amirmansoor Hakimi, [1]David M Horn, [2] Runsheng Zheng, [2] Oleksandr Boychenko, [1]Steven L. Reeber, [1]Daniel Lopez-Ferrer
[1]Thermo Fisher Scientific, San Jose, CA [2] Thermo Fisher Scientific, Germering, Germany

ABSTRACT

Purpose: To demonstrates the outstanding performance and the broad range of proteomics applications of the micro-flow liquid chromatography-mass spectrometry workflow.

Methods: LFQ and TMT-labeled samples were separated on a Thermo Scientific™ Vanquish™ Neo UHPLC system and were analyzed on Thermo Scientific™ Orbitrap Exploris™ mass spectrometers. Data was processed by Thermo Scientific™ Proteome Discoverer™ 3.0 software.

Results: We have demonstrated the micro-flow LC-MS setup to be highly reproducible and robust without sacrificing performance in proteomics applications for both discovery and quantitation.

INTRODUCTION

Liquid chromatography-mass spectrometry (LC-MS) has been a powerful analytical tool in protein identification and quantification. In the past few decades, nano-flow LC-MS has been the primary approach due to its high sensitivity. However, challenges arise from the need to satisfy objectives of high throughput, reproducibility and robustness. Here we present a micro-flow LC-MS workflow using a robust setup with a Vanquish Neo UHPLC System coupled to Orbitrap Exploris mass spectrometers. Gas-phase fractionation (GPF) was performed to improve protein and peptide coverage using a Thermo Scientific™ High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) Pro Duo interface. We further applied this platform to scale up the characterization of proteomes using peripheral blood mononuclear cells (PBMCs) across different organisms.

MATERIALS AND METHODS

Sample Preparation

Thermo Scientific™ Pierce™ HeLa Protein Digest Standard was reconstituted to a final concentration of 1µg/ul in 0.1% FA in LC/MS-grade water. Thermo Scientific™ Pierce™ TMT11plex Yeast Digest Standard was reconstituted to a final concentration of 2µg/ul in 0.1% TFA/5% acetonitrile in LC/MS-grade water. PBMC samples from different organisms were purchased from BioIVT (Westbury, NY). After cell lysis, protein amount was measured using Thermo Scientific™ Pierce™ Rapid Gold BCA Protein Assay Kit. Protein digestion and peptide clean-up were performed following Thermo Scientific™ EasyPep™ Mini MS Sample Prep Kit protocol on a platform adapted from Thermo Scientific™ TriPlus™ RSH Autosampler.

Test Method

Different amounts of HeLa protein digest standard and TMT11plex Yeast digest standard were loaded onto a Thermo Scientific™ PepMap™ C18 2µm 150x1mm column coupled to the Vanquish Horizon UHPLC system and separated by 30min and 50min LC gradients. Peptides were analyzed on an Orbitrap Exploris 240 mass spectrometer (ICSW 3.1) in data dependent acquisition mode. PBMC digest peptides were separated by a PepMap C18 2µm 150x1mm column coupled to the Vanquish Neo UHPLC system. A FAIMS Pro Duo interface was installed to provide an additional dimension in peptide separation. MS^2 spectra were collected at data dependent acquisition mode on Orbitrap Exploris 480 mass spectrometer (ICSW 3.1).

Data Analysis

Data was processed in Proteome Discoverer 3.0 software using MSPepSearch against the human NIST Orbitrap HCD library, in parallel with a new search engine, CHIMERYS. Percolator FDR calculation was used to only allow those spectra within 1% FDR rate to be reported.

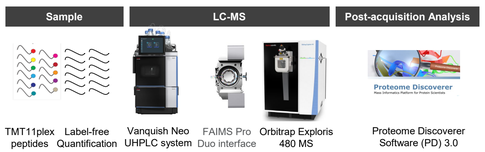


Figure1. Experiment Workflow micro-flow LC-MS data acquisition and data analysis

RESULTS

Label-Free Quantification of HeLa Protein Digest Standard

Label-free proteomics performance on the micro-flow LC-MS setup was evaluated at 3 different loads of HeLa protein digest, 1µg, 5µg, 10µg, separated at 30min and 50min gradient lengths. Each experiment condition was repeated for 5 injections to demonstrate the reproducibility.

The number of identifications increases at a longer LC gradient length or a higher loading amount. Meanwhile, the datasets maintain high reproducibility among replicates as shown in **Figure 1** and **Table 1**. Multiple metrics, such as missed cleavage, alkylation, methionine oxidation and deamidation, were evaluated (**Figure 2**) to show an excellent data quality and outstanding robustness of the micro-flow LC-MS platform.

Figure 1. Number of a) protein groups and b) peptide groups identified from 1µg, 5µg and 10µg of HeLa protein digest standard at 30min and 50min LC gradient. Each bar represents the average IDs out of 5 replicates. Error bar stands for standard deviation. Protein groups were by the combination of MSPepSearch and CHIMERYS in PD 3.0, filtered at 1% FDR.

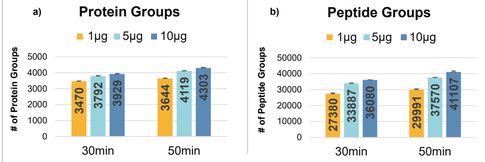


Table 1. Coefficient of variation (CV) of a) protein groups identifications and b) peptide groups identifications among 5 replicates of each dataset.

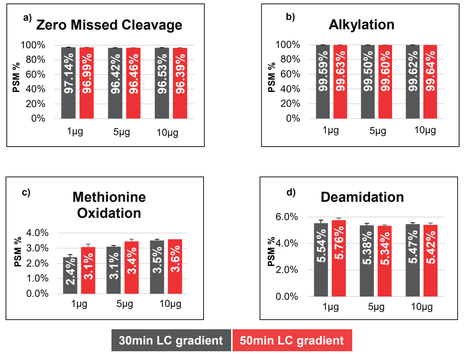
a) Protein Groups IDs Coefficient of Variation

LC Gradient	1µg	5µg	10µg
30min	0.72%	0.92%	0.69%
50min	0.75%	0.49%	0.99%

b) Peptide Groups IDs Coefficient of Variation

LC Gradient	1µg	5µg	10µg
30min	1.87%	1.00%	0.20%
50min	1.87%	0.34%	1.69%

Figure 2. Metrics evaluation of data quality in the aspects of a) missed cleavage b) alkylation rate c) Methionine oxidation rate and d) deamination rate. Each bar represents the average IDs out of 5 replicates. Error bar stands for standard deviation.



In addition to the great performance in identification, the micro-flow LC-MS system exhibited remarkable reproducibility in protein abundance (**Figure 3a, b**). The median of protein abundance CV was smaller than 11% for all conditions. Over 80% of proteins were with an abundance CV less than 20% (**Figure 3c**). Quantitative accuracy across a range of 1:10 ratio was also evaluated (**Figure 4**). Although ratio compression exists at higher ratios, longer gradient resulted in better accuracy compared to shorter gradient.

Figure 3. Label-free quantification analysis protein abundance distribution among 5 replicates at a) 30min b) 50min LC gradient. C) pie chart for the portion of proteins with abundance CV<20%

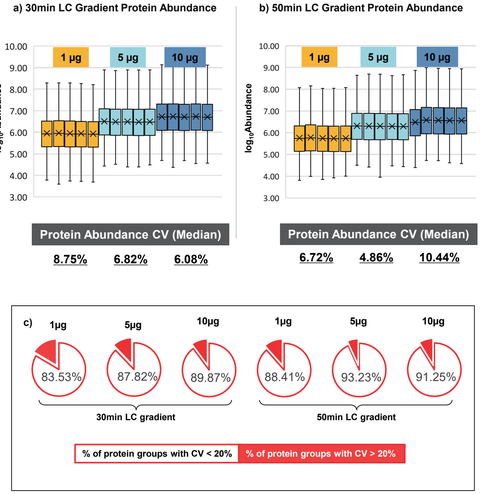
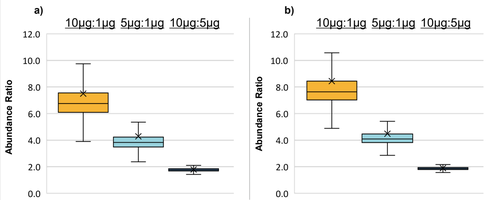


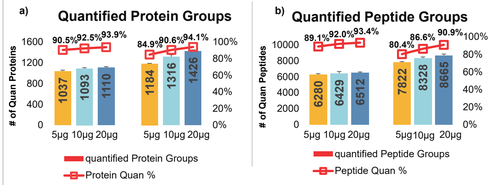
Figure 4. Label-free quantification accuracy of protein abundance at a) 30min b) 50min LC gradient



TMT11plex Yeast Digest Standard

The quantitation capability of the micro-flow LC-MS platform was also evaluated using TMT labeled samples. Triple knock-out standard (Met6: 126, 127N, 127C; His4: 128N, 128C, 129N; Ura2: 129C, 130N, 130C; Eno2: 131N, 131C) was analyzed at different load amounts and LC gradients. As the load amount increases, the percentage of quantified IDs increases. The longer LC gradient tends to quantify more protein/peptide groups, while the shorter gradient resulted in higher percentage of quantified IDs (**Figure 5a, b**). Under all conditions, micro-flow LC-MS maintained great reproducibility in assessing protein abundance among knock-out channels (**Figure 5c**) and labeling efficiency (**Figure 6**). Data analysis was performed in Proteome Discoverer 3.0 software using the Sequest™ HT search engine.

Figure 5. Number of a) protein groups and b) peptide groups quantified from 5µg, 10µg and 20µg of TMT11plex Yeast digest standard at 30min and 50min LC gradient. Each bar represents the average IDs out of 5 replicates. Error bar stands for standard deviation. Protein groups were filtered at 1% FDR. TMT channels were grouped according to knock-out channels. c) Replicates within each channels were evaluated for abundance CV.



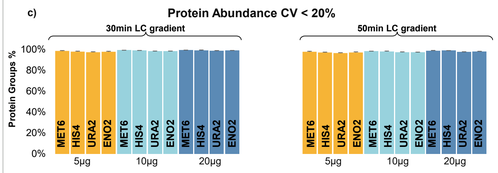
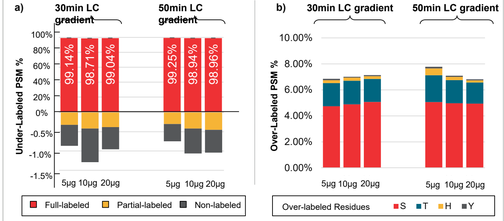


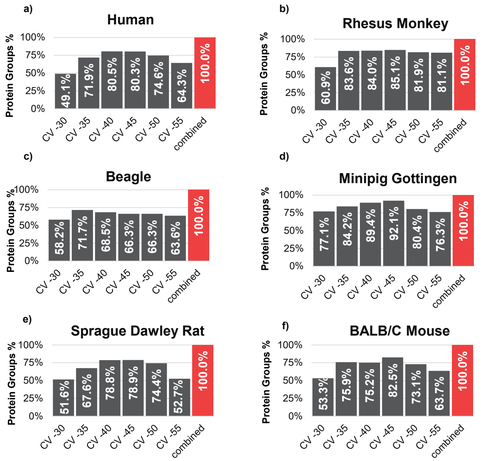
Figure 6. Labeling efficiency evaluated at 30min and 50min LC gradient from 5µg, 10µg and 20µg of TMT11plex Yeast digest standard. a) under-labeled PSMs b) over-labeled PSMs. Each bar represents the average IDs out of 5 replicates. Error bar stands for standard deviation.



Peripheral Blood Mononuclear Cells (PBMCs)

Peripheral blood mononuclear cells (PBMCs) from a variety of animal species (Human, Rhesus Monkey, Beagle, Minipig Gottingen, Sprague Dawley Rat and BALB/C Mouse, all in male gender) were analyzed for in-depth profile discovery. The micro-flow LC-MS system was coupled to the FAIMS Pro DUO interface, achieving an increment in number of identifications while maintaining the robustness through over 100 injections. Data acquisition of the complete dataset was completed within three and half days (48min method duration). Gas phase fractionation provides deep proteome profiling without the need for off-line RPLC fractionation, which generally reduced the overall experiment time. Data was processed on PD 3.0 using CHIMERYS algorithm.

Figure 7. FAIMS Pro DUO multiple compensation voltage proteome mapping from 1ug a) Human b) Rhesus Monkey c) Beagle d) Minipig Gottingen 3) Sprague Dawley Rat and f) BALB/C Mouse peripheral blood mononuclear cells. Each gray bar represents the combination of 3 replicates of a single compensation voltage (n30 stands for FAIMS CV -30). Each red bar represents the combination of 6 single compensation voltages.



CONCLUSIONS

This micro-flow LC-MS setup has been demonstrated to be highly reproducible and robust without sacrificing performance in proteomics applications for both discovery and quantitation. The addition of the FAIMS Pro DUO interface to the micro-flow LC-MS platform further facilitates the deep analysis of proteome profiling.

TRADEMARKS/LICENSING

PO66122 EN0921S