Ultra-Fast Capillary-Flow LC-MS Profiling of Complex Biological Matrices: Applicable to Large Sample Cohorts

Oleksandr Boychenko[1]; Jenny Ho[2]; Christopher Pynn[1]
[1]Thermo Fisher Scientific, Germering, Germany; [2]Thermo Fisher Scientific, Hemel Hempstead, UK

ABSTRACT

Purpose: Whilst LC-MS has matured into a technology capable of both high proteome depth and specificity, its widespread adoption has been limited by the low-throughput and insufficient robustness of the nanoLC-MS tools typical to research. Here we demonstrate a new set of capillary-flow LC-MS (capLC-MS) methods capable of large sample cohort analysis using a Thermo Scientific™ UltiMate™ 3000 RSLCnano system coupled to a Thermo Scientific™ Q Exactive™ HF-X Hybrid Quadrupole-Orbitrap™ mass spectrometer.

Methods: We optimized five low-flow LC-MS methods capable of throughputs of 180, 100, 60, 30, 24 samples per 24 hours and MS utilization from 75 to more than 90% respectively. The methods were validated using HeLa protein and crude plasma digests and showed excellent long term reproducibility and good protein coverage with more than 1000 protein groups identified with 8 min LC-MS method and 3400 protein groups within one 60 min LC-MS run.

Results: The developed low-flow LC-MS methods provide sharp and symmetric peaks that, in combination with the high sensitivity and fast acquisition speed of the Q Exactive HF-X mass spectrometer result in robust, fast and deep profiling of biological samples including cell lysate and crude plasma protein digests. Consistent results are generated over hundreds of replicate injections proving that the methods are suited to the analysis of large sample cohorts.

INTRODUCTION

The availability of blood and its plasma and serum constituents, combined with its comparative ease of sampling makes it ideal for clinical research. Proteins, in particular, as major players in catalytic and structural as well as many signaling functions, are essential research targets. Whilst antibody based assays currently dominate, issues concerning their specificity, the challenges related to multiplexing and in particular, the high cost in the development of new assays pose significant challenges.

LC-MS based technologies, whilst enabling precise detection and quantification of a multitude of targets from a single analysis, have nevertheless yet to fulfill their potential in the field of plasma proteomics. Blocking points include challenges due to the complexity and dynamic concentration range of the constituent proteins, which demand both high resolution and robustness from the analytical LC columns employed, coupled with high throughput to cope with the samples involved but without sacrificing sensitivity or robustness. Whilst sensitivity can, to an extent, be addressed using conventional nanoLC, this comes at the expense of both robustness and speed.

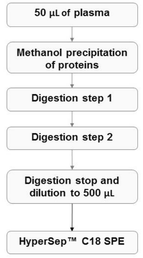
Here we demonstrate the efficacy of highly robust high throughput capillary-flow LC-MS methods, capable of the analytical depth and quantitative dynamic range which is well suited to the field of plasma or cell proteomics. The range of developed methods provides the flexibility to analyze samples of different complexity and balance throughput with the required proteome depth.

MATERIALS AND METHODS

Sample Preparation

Thermo Scientific™ Pierce™ HeLa protein digest (P/N 88328, 20 µg/vial) was reconstituted to a final concentration of 200 ng/µL in loading buffer. One microliter of HeLa protein digest solution was injected onto the column. Crude plasma samples were processed and enzymatically digested according to the scheme described in Fig. 1. One microliter of plasma protein digest solution was injected onto the column.

Figure 1. Simple sample workup for crude plasma extract analysis



Instrumentation

The UltiMate 3000 RSLCnano system was coupled with a Q Exactive HF-X mass spectrometer equipped with a Thermo Scientific™ EASY-Spray™ source operated in either Full MS or DDA+ modes (Fig. 2). Measurements were carried out using an UltiMate 3000 RSLCnano system which was configured in pre-concentration mode using the Thermo Scientific™ EASY-Spray™ connection kit (P/N 6720.0395) as described in the UltiMate 3000 RSLCnano Standard Application Guide [1] and a Thermo Scientific™ EASY-Spray™ column (75 µm x 150mm, 3µm, P/N ES800) (Fig. 3). The details concerning the LC-MS system configuration and solvents are described in [2].

Data Analysis

Data were acquired using SII for Thermo Scientific™ Xcalibur™ software. DDA data were processed with Thermo Scientific™ Proteome Discoverer™ software. The false discovery rate (FDR) was below 1% at the peptide and protein level.

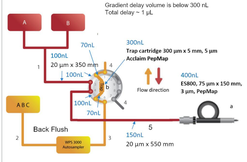
OPTIMIZED HIGH-THROUGHPUT LC-MS METHODS

Versatile low-flow LC-MS setup

Standard pre-concentration onto a trap cartridge UltiMate 3000 RSLCnano setup (P/N 6720.0310) is usually used for nanoLC-MS experiments with long columns to dive deep into the proteome. However, the versatility of this fluidics configuration also allows developing robust high-throughput methods that can be easily adjusted to any sample type or application. This LC setup comprises: low gradient delay volume and delay time; fast and precise gradient formation; no hardware changes required to switch between "deep dive" proteomics and high-throughput low-flow LC-MS (Fig. 2 and 3).

Figure 2. Instrument platforms used to develop ultra-fast capLC-MS methods for bottom-up proteomics.

Figure 3. Fluidic configuration for the high-throughput low-flow LC-MS analysis



Low-flow LC-MS methods for analysis of large sample cohorts

Multiple LC method components were optimized to develop a set of fast low-flow LC-MS methods with no sample loss, negligible carryover, robust operation and maximum MS utilization (Table 1). We defined LC method parameters for

- High speed sample loading and desalting using the integrated loading pump at high flow rate
- Fast autosampler injection routines, affording injection times of less than 1 minute without compromising injection reproducibility
- UHPLC optimized gradient conditions
- Independent washing of the trap and analytical column to ensure fast comprehensive washing of both
- Elevated flow rates during washing and equilibration steps for fast cycle times

Table 1. The main characteristics of developed low-flow LC-MS methods for high-throughput shotgun analysis of tryptic protein digests

	Flow, µL/min	Samples per 24 hours	MS utilization, %	Average PWHM, sec	Average PW base, sec	Asymmetry
60 min	0.300	24	95	10	19	1.23
48 min	0.600	30	90	9	18	1.21
24 min	0.800	60	87	7	13	1.17
14.4 min	1.000	100	85	4	7	1.13
8 min	1.500	180	75	3	6	1.16

Figure 4. The number of samples that can be analyzed with each high-throughput LC-MS method per 24 hours

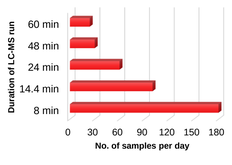
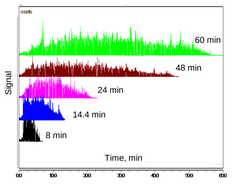


Figure 5. The typical TIC profiles for HeLa protein digest using LC-MS methods with 8, 14.4, 24, 48 and 60 min total analysis time



LOW-FLOW LC-MS ANALYSIS OF HELA PROTEIN DIGEST

Chromatographic performance

Chromatographic performance was assessed using EICs for 6 selected peptides from HeLa protein digest spanning the entire elution profile. The optimized methods bear witness to the following attributes:

- Robust retention times (SD < 0.1 min)
- Sharp peaks (Table 1) with PWHM from 3 sec for 8 min method to 10 sec for 60 min method
- Good peak shape with average Peak Asymmetry 1.15
- High MS utilization

The careful optimization of flow rates during the gradient allowed to compensate peak height reduction for longer gradients with improved ionization efficiency observed for reduced flow rates [3]. The obtained TIC profiles for HeLa protein digest (Fig. 5) show good distribution of peptides over the entire gradient affording a maximum number of MS/MS events. The conversion rate for MS/MS events to peptide spectrum matches was between 40 and 60%.

Figure 6. The number of protein groups identified with 1 % FDR in 200 ng of HeLa protein digest with optimized LC-MS methods

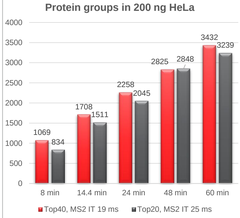
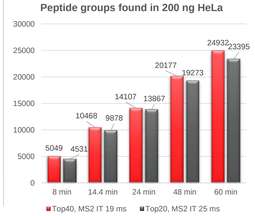


Figure 7. The number of peptide groups identified with 1% FDR in 200 ng of HeLa protein digest



MS performance and results of shotgun proteomics experiments

The number of protein and peptide identifications increased almost linearly with the length of low-flow LC-MS method. More than 1000 protein groups and 5000 peptide groups were identified with 8 min methods and above 3400 protein groups and 24000 peptide groups with the 60 min method (Fig. 6). We also confirmed that the improved speed of Q Exactive HF-X instrument significantly boosts the number of identifications, especially for short gradients (Fig. 6, 7). The plot of PSMs vs. retention time (Fig. 8) shows that peptides are eluted within the entire .raw file length with the exception of approximately 1 min at the start and 1 min at the end of the run (Fig. 8). The constructed Venn diagrams for proteins identified using each of the 5 methods shows that the longer methods almost completely encompass the results of the shorter LC-MS runs (Fig. 9), thus confirming the reliability of the obtained data.

Figure 8. The number of PSMs vs. retention time for optimized LC-MS DDA methods that shows high MS utilization

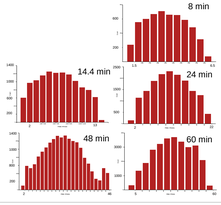
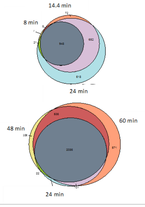


Figure 9. Venn diagram for protein groups identified in HeLa protein digest using LC-MS methods comprising 8, 14.4, 24, 48 and 60 min.



ROBUST ANALYSIS OF CRUDE PLASMA SAMPLES

Fast low-flow LC-MS profiling of plasma samples

LC-MS analysis of body fluids is extremely difficult due to the high dynamic range of proteins, presence of high-abundant proteins that cover > 99% of total proteome amount in biofluids and the significant amount of hydrophobic small molecules, e.g. phospholipids that can irreversibly absorb to the column stationary phase. The developed methods were applied to the analysis of crude plasma samples (Fig. 10) that were digested using a simple procedure that doesn't include reduction and alkylation of Cysteine residues and can be used for automated sample preparation (Fig. 1). We identified 150 protein groups and above 1000 peptide groups within 8 min. More than 250 protein groups and 2700 peptide groups were identified with the 60 min method (Fig. 11, 12). The quantitative performance was assessed from 100 different samples analyzed using the 14.4 min method. The method showed: robust quantitation with 90% of proteins quantified with CVs < 20% and broad concentration range of quantifiable proteins spanning 5 orders of magnitude (Fig. 13, 14).

Figure 10. The typical TIC profiles of crude plasma protein digests analyzed with 8, 14.4, 24, 48 and 60 min LC-MS methods

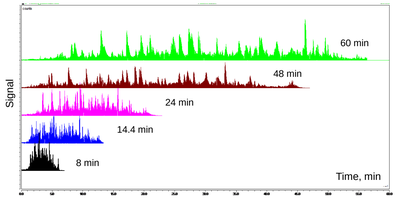


Figure 11. The number of protein groups identified with 1 % FDR in crude plasma protein digest using optimized LC-MS methods

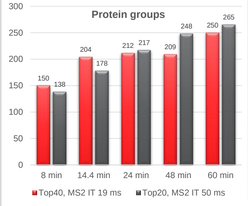


Figure 12. The number of peptide groups identified with 1 % FDR in crude plasma protein digest using optimized LC-MS methods

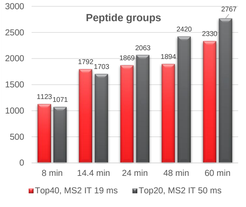


Figure 13. Typical peak characteristics for eluted peptides (A) and data sampling rate for label free quantification (B)

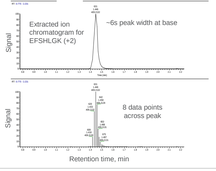
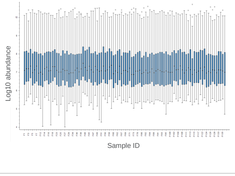


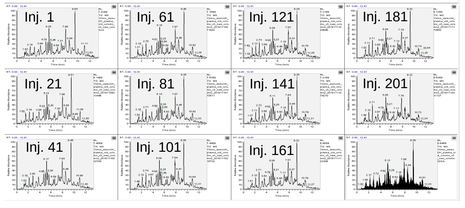
Figure 14. Protein abundances (normalized) based on label free protein quantification



REPRODUCIBILITY OF LC-MS METHODS FOR ANALYSIS OF LARGE SAMPLE COHORTS

The method durability and robustness was proven by LC-MS analysis of > 200 crude plasma samples in less than 2 days using the 14.4 min method. No loss in sensitivity or chromatographic performance was observed for any of the injections (Fig. 15).

Figure 15. Representative TIC Chromatograms for the consecutive measurement of > 200 crude plasma samples in < 3 days. Last Figure represents LC-MS/MS profile.



CONCLUSIONS

The requirement for high throughput multianalyte assays is of primary importance for clinical scientific research (including the rapidly emerging field of personalized medicine). Although immunoaffinity based assays yield unsurpassed levels of sensitivity, specificity multiplexing capabilities and development costs continue to pose a significant challenge.
Here we introduce a set of novel high-throughout capLC-MS methods for the analysis of biological samples including biofluids for research use only which yield the following attributes:

- high MS utilization > 80%
- negligible carryover facilitating continuous sample analysis
- multiple analytes quantitation capabilities
- broad analyte concentration range spanning 5 orders of magnitude
- compatibility with challenging matrices (e.g. crude plasma extracts)
- versatility of low-flow LC-MS methods that can be easily adjusted to any sample type and target(s)

REFERENCES

1. RSLCnano pre-concentration nano LC kit (P/N 6720.0310) in the Ultimate 3000 RSLCnano Standard Applications Guide. Hyperlink
2. Boychenko, A.; Pynn, C.; van den Berg, B.; Arrey, T. N., Baynham, M.; Decrop, R.; Ruehl, M. High-throughput capillary-flow LC-MS proteomics with maximum MS utilization. TN 72227. Hyperlink
3. Boychenko A., Meding S., Decrop W., Ruehl R., Swart R. Capillary-flow LC-MS: combining high sensitivity, robustness, and throughput. TN 72277. Hyperlink

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

TRADEMARKS/LICENSING

PO65495-EN0519S

ThermoFisher SCIENTIFIC