High-throughput nanoLC-MS for Sample-Limited Proteomics

Alec Valenta[1], Christopher Pynn[1], Dominich Hoch[2], Manuel Matzinger[3], Rupert Mayer[3], Karl Mechtler[3], Runsheng Zheng[1]
[1]Thermo Fisher Scientific, Germering, Germany; [2]Thermo Fisher Scientific, Reinach, Switzerland; [3]Institute of Molecular Pathology, Vienna, Austria

INTRODUCTION

Nano-flow liquid chromatography-mass spectrometry (nLC-MS) is often utilized for bottom-up proteomics of extremely limited samples, down to individual cells. This is due to the large increase in ionization efficiency achieved at LC flow rates < 300 nL/min. Generating reproducible and timely results at ultra-low flow rates requires state-of-the-art LC instrumentation and optimization of LC-MS parameters, particularly with respect to the separation and electrospray ionization (ESI) interface. Here we present 6 high-throughput nanoLC-MS methods for the bottom-up proteomics label-free quantitation (LFQ) analysis of limited samples in order to balance sample throughput and proteome depth. Several data acquisition strategies were compared including data-dependent (DDA), wide-window (WW-DDA), and data-independent (DIA) acquisition. Lastly, a modified trap-and-elute method was applied to single-cell proteomics (SCP) profiling.

MATERIALS AND METHODS

Sample preparation

The Thermo Scientific™ Pierce™ HeLa Digest/PRTC Standard was prepared in water (0.1% FA, v/v) comprising of 1 ng/µL HeLa digest with 0.5 fmol/µL PRTC. Sample mass was varied by injecting volumes from 0.25 – 10 µL of HeLa digest onto the column. Single-cell samples were prepared in a 384 well plate using a label-free, one-pot workflow.[1]

Instrument configuration

Method optimization and single-cell measurements were performed on a Thermo Scientific™ Vanquish™ Neo UHPLC system coupled to a Thermo Scientific™ Orbitrap Exploris™ 480 mass spectrometer equipped with a Thermo Scientific™ FAIMS Pro™ interface. Samples were injected onto a Thermo Scientific™ Acclaim™ PepMap™ 100 C18 50 µm I.D. column and separated at 100 nL/min (**Figure 1**). Both direct injection and trap-and-elute workflows were explored for balancing throughput and sensitivity (**Table 1**). For the trap-and-elute workflow, a 300 µm x 5 mm trap column was operated in backward flush mode. Mobile phase A and weak wash liquid were water with 0.1% FA while mobile phase B and strong wash liquid were 80% acetonitrile with 0.1% FA. The column was heated to 50 °C. The column outlet was connected to a 10 um µm I.D. x 5 cm emitter (Fossillion technology, LOTUS).

Figure 1. nano-flow LC-MS configuration

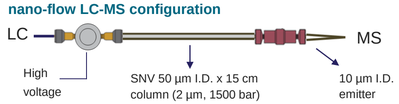

Figure 1. Optimized fluidic configuration for separations at 100 nL/min.

Table 1. Six nanoLC-MS methods for sample-limited proteomics

Samples/Day	Cycle Time (min)	Elution Window (min)	Sample Injection & Loading (min)*	MS utilization (%)
100	14.4	10	1.4	69%
72 *	20	11	5.4	55%
60	24	15	5.4	63%
48	30	21	5.4	70%
36	40	31	5.4	78%
24	60	51	5.4	85%

*100 samples/day using trap-and-elute workflow

Table 1. Method details for the analysis of 250 pg of HeLa digest analyzed using the six nanoLC-MS methods operated at 100 nL/min. Sample throughput ranged from 24-72 SPD for direct injection, with an increase to 100 SPD using trap-and-elute. Sample injection/loading required 5.4 min of cycle time for DI and 1.4 min for D&E.

Figure 2 shows chromatograms for 250 pg HeLa digest using two of the nanoLC-MS methods operated at 100 nL/min: 72 samples/day (direct injection) and 100 samples/day (trap-and-elute). Fast sample loading, column washing, and equilibration enable 55% and 69% mass spectrometer utilization, respectively.

Figure 2. Example Chromatograms for 72 and 100 samples/day

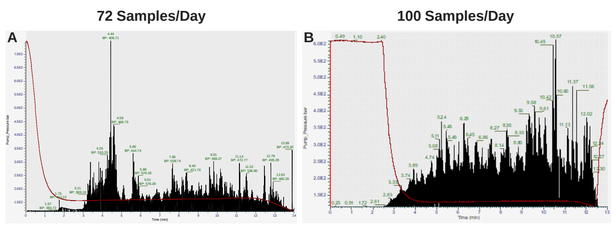

Figure 2. Total ion chromatograms and pressure traces for 250 pg HeLa digest using 72 (A) and 100 (B) samples/day methods in the direct injection and trap-and-elute workflows, respectively. Both methods were operated at a separation flow rate of 100 nL/min.

Data analysis

DDA and WW-DDA datasets were processed with Thermo Scientific™ Proteome Discoverer™ 2.5 software using a 2-step SEQUEST™ HT search algorithm and INFERYS™ rescoring node. DDA chimeric spectra were searched using the CHIMERYS™ algorithm in Proteome Discover 3.0 while DIA files were submitted to Spectronaut 17 (SN17) for peptide and protein ID and quantification. The false discovery rates (FDR) were all set below 1% at both the peptide and the protein levels.

Results

New benchmark in LFQ-DDA

Using LFQ-DDA in the direct injection workflow with a 2-step SEQUEST search + INFERYS rescoring, a linear increase was observed in protein IDs from 250 pg to 5 ng HeLa digest (**Figure 3**), suggesting method suitability for sample-limited analysis (e.g., SCP). We confidently ID ~1,500 protein groups from 250 pg HeLa digest without match-between-runs, which to the authors knowledge, represent the most comprehensive DDA data to date.[2]

Figure 3. LFQ-DDA Low Sample Input Evaluation

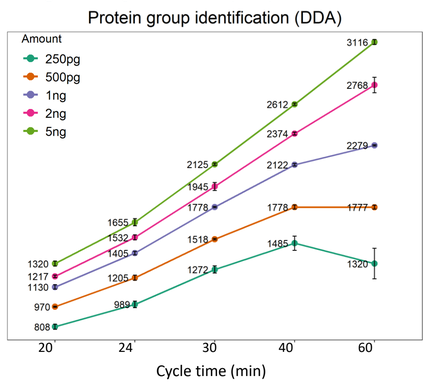

Figure 3. HeLa digest samples from 0.25 - 5 ng were analyzed in the direct injection workflow at 100 nL/min by varying the injection volume (n = 3). Mass spectrometer MS1 and MS2 resolution were set to 120K and 6K, respectively.

DDA Chimeric spectral deconvolution

CHIMERYS, an AI-driven algorithm for chimeric spectral deconvolution, was employed in order to explore the effect of MS1 acquisition window width on proteome depth for the 10-min and 30-min gradient methods. This WW-DDA strategy enabled the identification of up to ~1,800 protein groups from 250 pg HeLa digest using the 20-min method while the 40-min method yielded 2,000 protein group IDs (**Figure 4**). Results indicate an optimal isolation window of 10-12 m/z for both method lengths

Figure 4. WW-DDA window evaluation

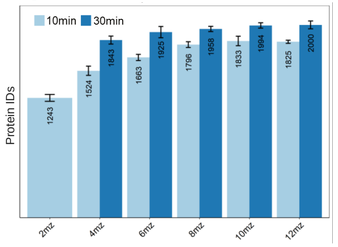

Figure 4. Evaluation of WW-DDA acquisition window sizes from 2-12 m/z for 10 and 30-min gradients (n = 3).

LFQ-DIA performance

A systematic evaluation of the MS1 open window in DIA yielded as much as 3,000 protein IDs from 250 pg HeLa digest using the 20 min method (**Figure 5A**), covering > 4 orders of magnitude of the protein abundance dynamic range (**Figure 5B**).

Figure 5. Proteome coverage in LFQ-DIA: Direct injection

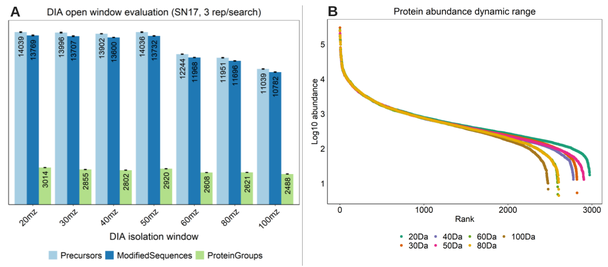

Figure 5. Proteome coverage (A) and dynamic range (B) in the direct injection workflow using LFQ-DIA when varying the MS1 isolation window. SN17 was used for data analysis (n = 3).

Increasing sample throughput for LFQ-DIA profiling

To accelerate sample loading and eliminate the potential negative impact of impurities and detergent on electrospray ionization, we employed a trap column operated in a backward flush mode to maintain peak shape, successfully decreasing the method cycle time to 14.4 min (100 samples/day) for a 10-min gradient at 100 nL/min (~70% MS utilization, Figure 6). Method performance enabled identification of >2,200 protein groups in 250 pg and >1,100 protein groups in a little as 60 pg.

Figure 6. Proteome coverage in LFQ-DIA: Trap-and-Elute

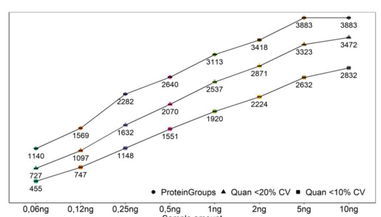

Figure 6. Proteome coverage in LFQ-DIA using a trap-and-elute workflow for increased sample throughput (100 samples/day). Mass was varied from 0.06 - 10 ng by adjusting the injection volume (n=3).

LFQ-DIA profiling of single-cells

To validate the 100 SPD trap-and-elute workflow for single-cell proteomics, the method was applied to HeLa and K562 single-cell QC samples followed by individual HeLa and K562 cells. QC results were compared across multiple sites and provided > 1,200 protein IDs , despite multiple labs, LC-MS systems, operators, columns, etc. (**Figure 7A**). Performance was then demonstrated on 10 individual HeLa cells where ~1,700 proteins were identified per cell (**Figure 7B**).

Figure 7. Single-cell profiling performance

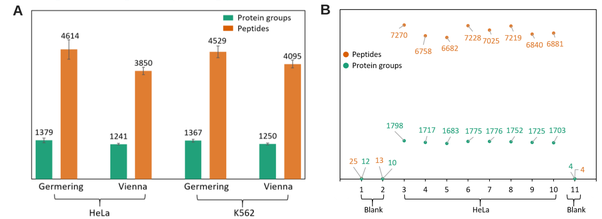

Figure 7. Evaluation LFQ-DIA for single-cell QC samples (A) and individual cells (B). For single cells, each raw file was searched individually.

CONCLUSIONS

We developed a high-sensitivity and high-throughput nano-LCMS platform that affords the analysis of 24-100 samples/day with industry leading protein coverage in both LFQ-DDA & LFQ-DIA modes. Using a 20 min (72 SPD) label-free direct injection method we were able to identify > 800 protein groups (DDA +SEQUEST+INFERYS), >1,800 protein groups (DDA+CHIMERYS), and >2,600 (DIA+SN17). Throughput was further enhanced by utilizing a trap-and-elute workflow for 100 SPD, which provided >2,200 protein groups (DIA+SN17). Lastly, ~1,700 proteins were identified in individual HeLa cells using the 100 SPD method.

REFERENCES

1. Manuel Matzinger, et al. Anal. Chem. 2023, 95 (9), 4435–4445
2. Karel Stejskal, et al. Anal. Chem. 2021, 93, 25, 8704–8710

ACKNOWLEDGEMENTS

We would like to thank Xuefei Sun, Cornelia Boeser, Jeff Op de Beeck, Tabiwang Arrey and Dominic Hoch from Thermo Fisher Scientific for the technical supports and fruitful discussion.

TRADEMARKS/LICENSING

PO2023-13EN