TECHNICAL NOTE 000127

DIA performance in discovery and quantitation analysis on the Orbitrap Exploris 240 mass spectrometer

Authors: Yang Liu[1], Khatereh Motamedchaboki[1], Joshua Nicklay[2], Sega Ndiaye[3], Aman Makaju[1], Amarjeet Flora[4], Ryan Bomgarden[4], Daniel Lopez-Ferrer[1]

[1]Thermo Fisher Scientific, San Jose, CA, USA

[2]Thermo Fisher Scientific, Somerset, NJ, USA

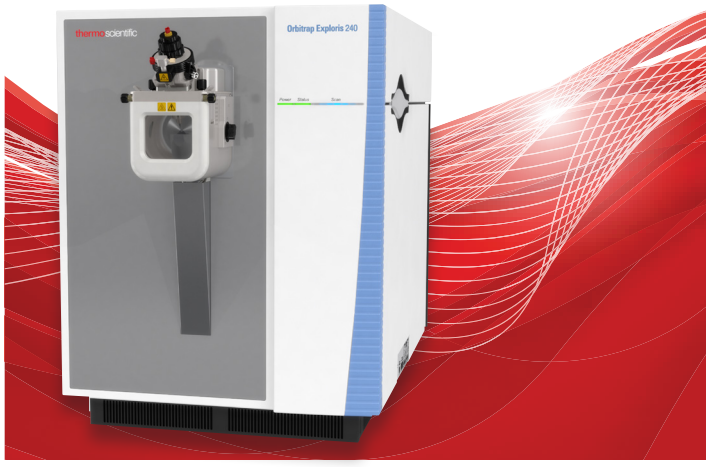
[3]Thermo Fisher Scientific, Courtaboeuf, France

[4]Thermo Fisher Scientific, Rockford, IL, USA

Keywords: Data-independent acquisition, DIA, Orbitrap Exploris 240, label-free quantitation, LFQ, gas-phase fractionation, HeLa

Goal

To evaluate data-independent acquisition performance on the Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer for quantitative discovery proteomics analysis.



Introduction

Label-free quantitation (LFQ) is the simplest way to explore the proteome. This approach offers deep proteome coverage, quantitative accuracy, minimal sample preparation, and the analysis of relatively large numbers of samples from diverse origins. There are two types of LFQ approaches, and they differ in how the identification (MS^2) data is acquired. Data-dependent acquisition (DDA) is ideal when analytical robustness and the highest level of data precision are the main priorities. It is also suited to working with small to medium sample sizes. In contrast, data-independent acquisition (DIA) can rapidly identify and reproducibly quantify all ions within liquid chromatography-mass spectrometry (LC-MS) analysis, making this process

ThermoFisher SCIENTIFIC

uniquely suited for large-scale proteome profiling and quantitative studies. In the DIA strategy, precursor ions are fragmented regardless of their peak intensity across wider isolation windows to co-fragment all the co-eluted peptides. This approach eliminates the intensity bias associated with DDA-based MS^2 acquisition, leading to reproducibility improvements and overcoming the limitation of under-sampling. In the DIA strategy, peptides may be identified by matching MS^2 data against a spectral library. The libraries are generated from a set of DDA experiments consisting of the same proteome as the DIA analysis. Typically, the sample has undergone high pH reverse-phase offline fractionation to ensure deep proteome coverage to generate the spectral libraries.

Here we present a mass spectrometry (MS)-only workflow that combines gas-phase fractionation (GPF) and DIA acquisition, saving significant experiment time while maintaining very high reproducibility. This strategy generates a chromatogram spectral library with GPF deep scanning experiments, which consist of replicate injections of the pool proteome modifying the MS acquisition mass range for each of the injections. This method operates in scanning mode, employing 80 amu windows at the MS^1 level and 3 amu at the MS^2 level. For each of the injections, the windows in the *m/z* range progress to higher values with only 5 amu of overlap at the MS^1 level. Data were integrated with a predicted spectral library to create an empirically corrected library for DIA library search. The GPF-DIA method was carried out on the Orbitrap Exploris 240 mass spectrometer to highlight the performance in both discovery identification and LFQ, as well as the simplicity and robustness of the method to enable laboratories to tackle large-scale, standardized proteomics projects.

Experimental

Materials and samples

Protein	Thermo Scientific™ Pierce™ HeLa Protein Digest Standard (Cat. #8828) BL21-Gold (DE3) *E. coli* competent cells (Agilent Technologies, Cat. #230132) Thermo Scientific™ EasyPep™ Mini MS Sample Prep Kit (A40006)
Column	IonOpticks™ Aurora™ column, 25 cm × 75 µm ID, 1.6 µm C18, (P/N AUR2-25075C18A)
Column oven	Sonation Column Oven (PRSO-V2™, Sonation GmbH) operating at 40 °C
Source	Thermo Scientific™ Nanospray Flex™ ion source (Cat. #ES071)
Mobile phase A	0.1 formic acid in water
Mobile phase B	0.1 formic acid in 80% acetonitrile, 20% water
Mass spectrometer	Orbitrap Exploris 240 MS
Liquid chromatography	Thermo Scientific™ EASY-nLC™ 1200 system (Cat. #LC140)
Software	Spectronaut™ 14.0 software (Biognosys)

Sample preparation

Pierce HeLa Protein Digest Standard was dissolved in 0.1% formic acid (FA) with 30 seconds of vortexing. BL21-Gold (DE3) *E. coli* competent cells were prepared using the EasyPep Mini MS Sample Prep Kit. *E. coli* digested peptides were added to a fixed amount of HeLa digest peptides (800 ng) at a ratio of 0 ng: 24 ng: 48 ng: 96 ng: 192 ng (0:1:2:4:8), and the HeLa-only sample was used as the control (Figure 2B). Mixed digested peptides were transferred to the autosampler vial for LC-MS analysis.

LC-MS method

HeLa/*E. coli* digested peptides were loaded on to a 25 cm Aurora Series emitter column (25 cm × 75 µm ID, 1.6 µm C18) coupled to a PRSO-V2 Sonation column oven. Samples were separated at 300 nL/min flow rate in direct injection mode on an EASY-nLC 1200 system with a 90 min LC gradient before being injected onto the Orbitrap Exploris 240 mass spectrometer using instrument control software 1.1 (Figure 1).

The HeLa/*E. coli* sample at the highest protein amount (E4, Figure 2B) was used for GPF-DIA, which was performed at different mass ranges with a narrow isolation window. Single-injection DIA of HeLa/*E. coli* sample (E0-E4, Figure 2B) was performed at a wide *m/z* isolation window on the same instrument and matrix (Figure 2C).

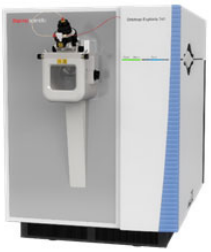
LC conditions

Retention (min)	Flow (nL/min)	Mixture (%B)
00:00	300	3
01:00	300	3
76:00	300	24
91:00	300	40
95:00	300	98
104:00	300	98

25 cm Aurora Series emitter column

PRSO-V2 Sonation column oven

EASY-nLC 1200 LC system



Orbitrap Exploris 240 mass spectrometer

Data processing software

Figure 1. LC-MS setup for DIA performance evaluation

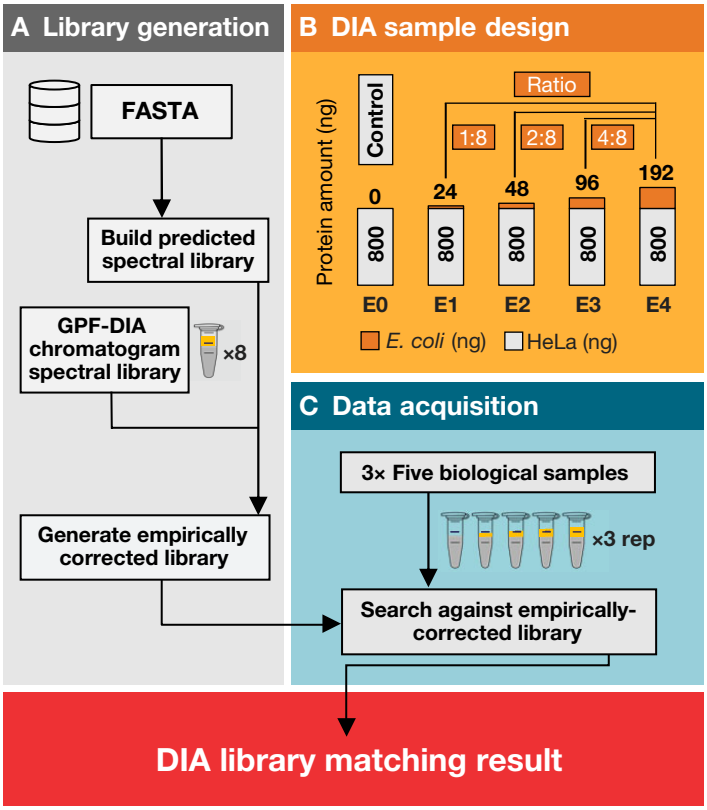


Figure 2. DIA experimental workflow

MS global settings

Spray voltage	1,800 V (adjustment needed depending on column)
Ion transfer tube temperature	275 °C
Sweep gas	0
Expected peak width	15 s
Advanced peak detection	True

MS parameters-Single DIA injection

MS^1 resolution	120,000
MS^1 scan range	380–985 *m/z*
MS^1 AGC	300% (3e6)
MS^1 max injection time	100 ms
RF lens	70%
NCE	28%
Cycle time	3 s
DIA resolution	15,000
DIA scan range	145–1450 *m/z*
DIA AGC	200% (2e6)
DIA max injection time	22 ms
DIA isolation window	10 *m/z*
DIA window overlap	1 *m/z*

MS parameters-GPF DIA

MS¹ resolution	60,000
MS¹ scan range	380–460 *m/z*, 455–535 *m/z*, 530–610 *m/z*, 605–685 *m/z*, 680–760 *m/z*, 755–835 *m/z*, 830–910 *m/z*, 905–985 *m/z*
MS¹ AGC	300% (3e6)
MS¹ max injection time	100 ms
RF lens	70%
NCE	28%
Cycle time	3 s
DIA resolution	30,000
DIA scan range	145–1450 *m/z*
DIA AGC	200% (2e6)
DIA max injection time	22 ms
DIA isolation window	3 *m/z*
DIA window overlap	1 *m/z*

Data analysis

Predicted tryptic peptides spectral libraries for HeLa and *E. coli* were generated from the FASTA database using the Prosit™ prediction tool. The GPF-DIA chromatogram spectral library was then matched to the predicted library and built into an experiment-specific, empirically corrected library (Figure 2A). Single-injection DIA data were searched against the empirical library using Spectronaut 14.0 software.

Results and discussion

DIA performance in identifications for discovery experiments on an Orbitrap Exploris 240 mass spectrometer

To evaluate DIA performance, 992 ng of HeLa/*E. coli* digest (E4, Figure 2B) was analyzed on an Orbitrap Exploris 240 mass spectrometer using the GPF-DIA method. The GPF acquisitions were searched against the predicted library of HeLa and *E. coli* tryptic peptides to generate an empirical chromatogram library. Samples E0-E4 were loaded according to the amount indicated in Figure 2B and were analyzed using a single-injection DIA method. The acquired DIA data were searched against the empirical libraries.

The Orbitrap Exploris 240 mass spectrometer had excellent performance for protein identifications. Over 9,000 proteins were identified from ~1 µg of the mixed proteome sample in a 90 min gradient run (Figure 3A). This method provided excellent data quality with a minimum number of missing values. Over 75% of the proteins were quantified in every LC-MS run (Figures 3B and 3C). In addition, the Orbitrap Exploris 240 mass spectrometer achieved identifications over 6 orders of magnitude in protein dynamic range (Figure 3D).

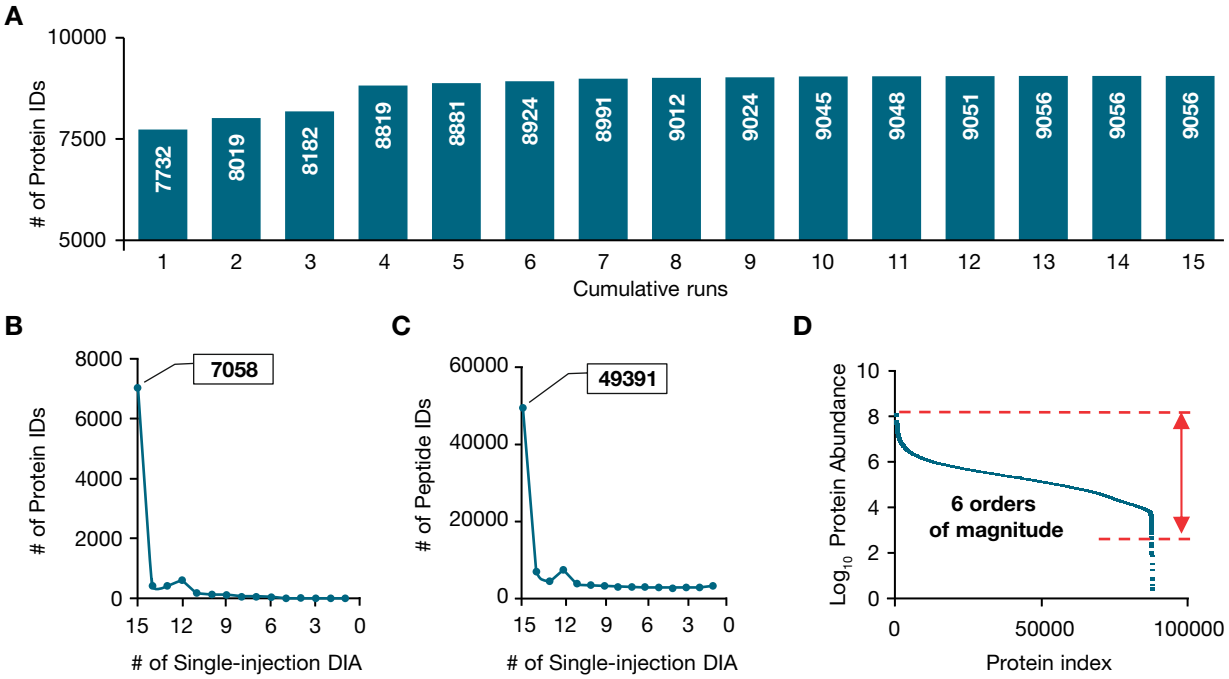


Figure 3. Protein identifications (A), data completeness (B and C) and the abundance dynamic range (D) of DIA experiments on an Orbitrap Exploris 240 MS

DIA performance on an Orbitrap Exploris 240 mass spectrometer at different sample loads

Each HeLa/*E. coli* sample (E0-E4) contains different protein amounts, ranging from 800 ng to 992 ng (Figure 2B). In the descriptions below, DIA experiments using full load (800–992 ng protein amount) are shown as "1,000 ng DIA" and half load (400–496 ng protein amount) experiment are shown as "500 ng DIA".

The chromatogram spectral libraries were generated for both 500 ng and 1,000 ng HeLa/*E .coli* on the Orbitrap Exploris 240 mass spectrometer using the GPF-DIA method. They were matched to the same HeLa/*E. coli* predicted spectral library and built into the empirical libraries for 500 ng and 1,000 ng sample load. Single-injection DIA of 500 ng HeLa/*E. coli* samples (E0-E4) were performed in triplicate and then searched against both the 500 ng and 1,000 ng DIA libraries (Figure 4).

By using the same experiment condition and sample matrix, the empirically corrected library built with the 1,000 ng protein amount contained 11.6% more proteins and 20.4% more peptides (Figure 5A). When 500 ng single-DIA data were searched against the 1,000 ng DIA library, there were 20% more peptides identified compared to being searched against the 500 ng DIA library (Figure 5B). Cscore is a Spectronaut identification score indicating closeness of match. Higher quality spectra tend to have a greater score value and yield more confident identifications. Figure 5C shows the Cscore distribution of the proteins identified by searching against different DIA libraries. Although the difference in protein identifications were only 3% between 500 ng single-injection DIA searching against 500 ng and 1,000 ng empirical libraries, the identifications from the higher load library were at a higher Cscore range, which indicates better quality than those from the lower load library. We have also compared the Cscore of common proteins found in both DIA libraries. Over 98% of the proteins were associated with a higher Cscore when being identified from the 1,000 ng library relative to the same identifications identified from the 500 ng DIA library (Figure 5D).

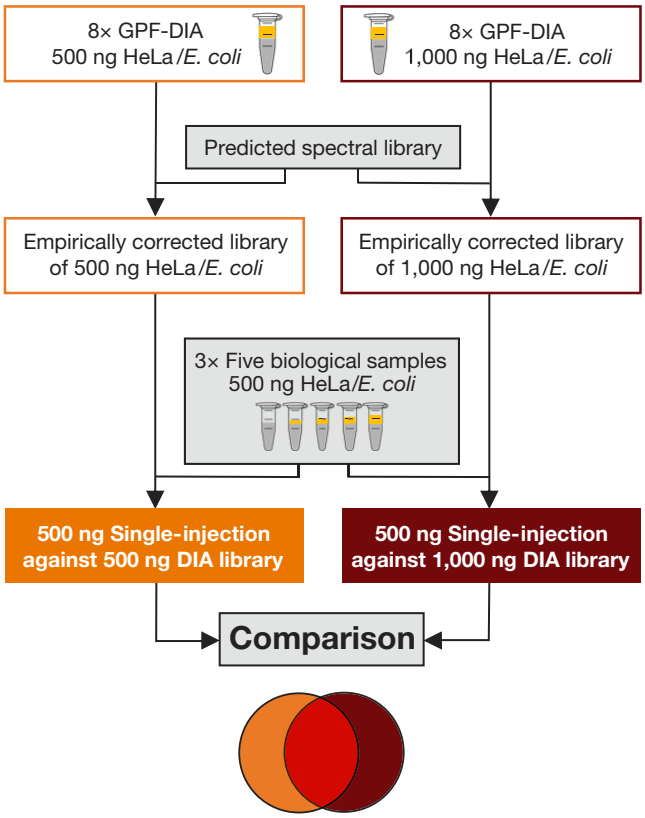


Figure 4. Workflow of DIA experiments at different sample loads

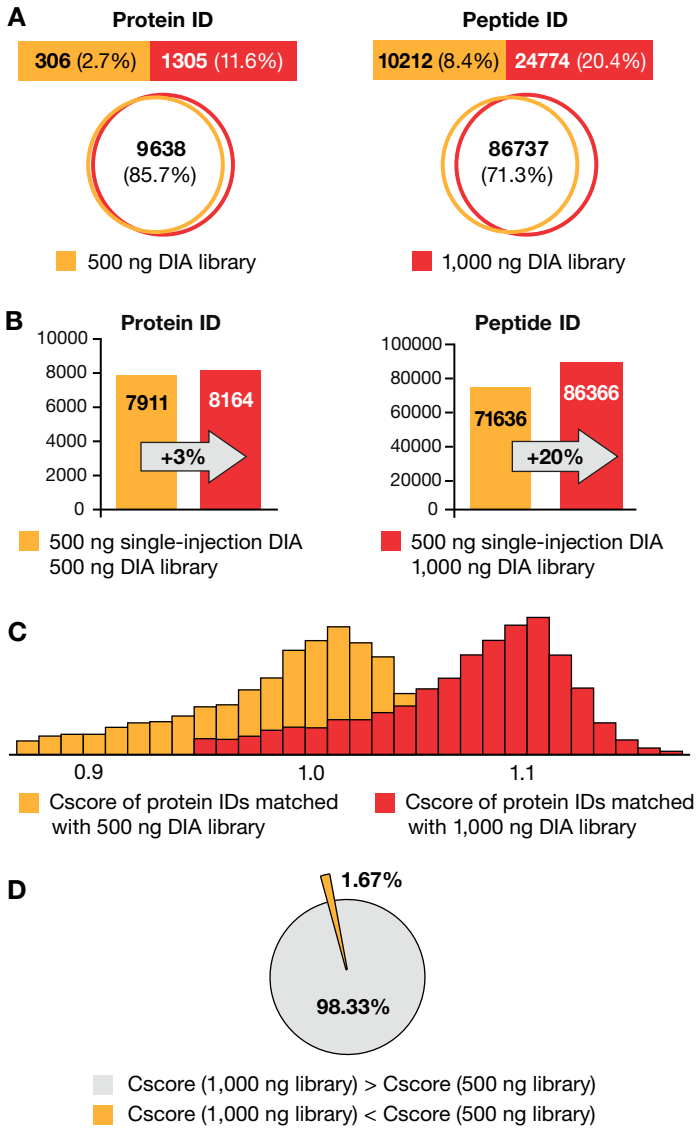
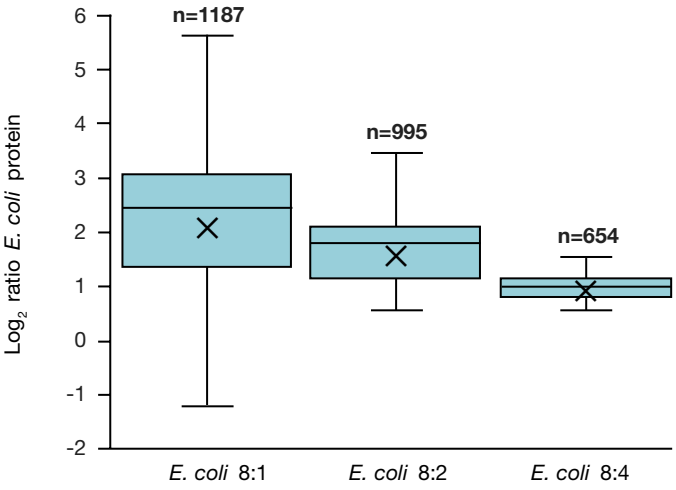


Figure 5. DIA performance in ID discover at different sample loads

thermo scientific



	Log$_2$ (8:1)	Log$_2$ (8:2)	Log$_2$ (8:4)
240 MS	2.44	1.81	1.00

Figure 6. DIA-LFQ quantification performance of Orbitrap Exploris 240 MS (n represents number of quantified proteins)

These results demonstrate that the empirical chromatogram library built at a higher load could be universally used to match a single-injection DIA at lower sample loads, which would save experimental time for generating multiple libraries and would improve the number and the quality of identifications.

DIA LFQ on an Orbitrap Exploris 240 mass spectrometer

800–992 ng of HeLa/*E. coli* samples E0-E4 were analyzed on an Orbitrap Exploris 240 mass spectrometer using the single-injection DIA method. The acquired DIA data was then searched against their respective GPF-DIA library on Spectronaut 14.0 software for MS[1] quantitation. The quantified protein identifications were filtered by "absolute average $\log_2$ ratio ≥ 0.58" and "Qvalue ≤ 0.05" (Figure 6). The Orbitrap Exploris 240 mass spectrometer provided outstanding performance in DIA based LFQ.

Conclusion

- The Orbitrap Exploris 240 MS is demonstrated to be a powerful mass spectrometer for DIA analysis.
- GPF is an ideal option to create large proteome datasets to explore the proteome. In this case, we have used it to create DIA libraries, thereby minimizing sample handling steps compared to traditional chromatography-based methods while maximizing robustness, ease-of-use, standardization and proteome depth.
- The Orbitrap Exploris 240 mass spectrometer achieves excellent performance, quantifying robustly more than 9,000 proteins in 90 min and with minimal missing values, ensuring full data completeness in over 80% of proteins quantified in every run.
- An empirical chromatogram library built at high sample load could be universally matched with single-injection DIA acquisitions at lower load of the same sample matrix.

References

1. Searle, B.C., Swearingen, K.E., Barnes, C.A. et al. *Nat Commun* **2020**, *11*, 1548.
2. Gessulat, Schmidt et al. *Nat Methods* **2019**, *16*, 509–518.

Find out more at **thermofisher.com/orbitrapexploris240**

TN000127-NA-EN 0721S

ThermoFisher SCIENTIFIC