LFQ-DIA

High-throughput high-resolution data-independent acquisition workflow on an Orbitrap Exploris 480 mass spectrometer for accurate label-free quantitation

Kevin Yang[1], Julia Kraegenbring[2], Jingjing Huang[1], Julian Saba[3], and Amirmansoor Hakimi[1]

[1]Thermo Fisher Scientific, San Jose, California, USA, [2]Thermo Fisher Scientific, (Bremen) GmbH, Germany, [3]Thermo Fisher Scientific, Winnipeg, MB, Canada

Abstract

Purpose: To develop and assess qualitative and quantitative performance of label-free quantitation (LFQ) with an optimized data-independent acquisition (DIA) method on a Thermo Scientific™ Orbitrap Exploris™ 480 mass spectrometer using a long (60 min active gradient), short (30 min active gradient) and ultra-high throughput (9 min active gradient) gradient for large-scale proteomics analysis.

Methods: DIA workflow investigated in the present study was operated by directly injecting samples onto a 50-cm µPAC Neo™ HPLC column, and the peptides were resolved in a 60-min, 30-min and an ultra-high throughput 9-min gradient operated by a Thermo Scientific™ Vanquish™ Neo UHPLC system. The eluted peptides were analyzed on an Orbitrap Exploris 480 MS operated in DIA mode.

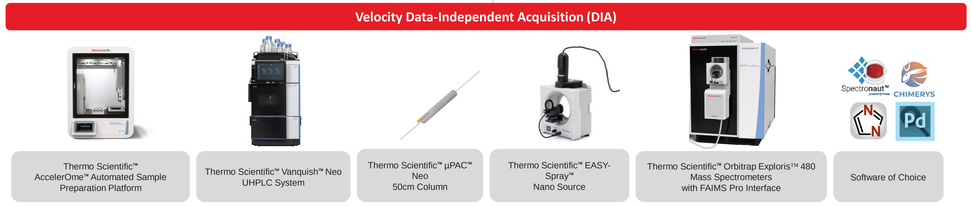
Results: Our results demonstrate that Velocity DIA presented herein is a robust and reproducible workflow on the Orbitrap Exploris 480 mass spectrometer for accurately quantifying and identifying hundreds to thousands of proteins from single cell-line to complex sample mixtures with a high background of human peptides. The Velocity DIA workflow could be coupled with the optional Thermo Scientific™ AccelerOme™ automated sample preparation platform to improve throughput and minimize variations caused by manual sample handling (Figure 1). The LC method and gradient were optimized to fully utilize the capabilities of micropillar array-based separation columns. Similarly, the mass spectrometric method was adopted to account for long (60 min active gradient), short (30 min active gradient), and ultra-high throughput (9 min active gradient) gradient lengths without compromising identification rates at great quantitation accuracy and precision.

Introduction

Quantitative proteomics is an essential tool for understanding global protein expression and the mechanisms of biological processes and disease states. Statistical significance is improved by decreasing variability in measurements and/ or increasing the sample set. However, increasing throughput means decreasing acquisition time, which often comes at a cost to measurement quality. Therefore, acquisition methods must be extensively optimized and validated to ensure that the data will produce meaningful biological insights.

Traditional data-dependent analysis (DDA) approaches have been widely employed for LFQ experiments, but they suffer from run-to-run inconsistencies due to intensity-based stochastic triggering of precursors, often leading to under sampling especially of low-abundant proteins. Missing values become more likely as sample size increases, and so data independent acquisition (DIA) has emerged as a popular technique for large scale quantitative analyses.

Figure 1. Complete 'end-to-end' Velocity DIA workflow.



Velocity DIA workflow on Orbitrap Exploris 480 hybrid MS together with the optional AccelerOme automated sample preparation platform.

Materials and methods

Sample Preparation

Pierce™ HeLa Protein Digest Standard, Waters E. coli MassPREP™ Standard and Promega Mass Spec-Compatible Yeast Protein Digest were dissolved in 0.1% formic acid (FA) with 30 seconds of vortexing. For the three-proteome mix, E. coli peptide digest and yeast peptide digest were added to a fixed amount of HeLa digest (325 ng) at amounts of 100 ng to 25 ng, and 75 ng to 150 ng, respectively, yielding an E. coli peptide ratio of 1:4 and a yeast peptide ratio of 1:0.5.

Test Method(s)

HeLa digest and three-proteome mixtures were loaded onto a 50 cm µPAC Neo HPLC column and separated at a 350 nL/min flow rate in direct injection mode using a Vanquish Neo UHPLC system over 9 min, 30 min, and 60 mins LC gradients, respectively, before being transferred into the Orbitrap Exploris 480 mass spectrometer (Figure 1 and 2).

Source parameters, including spray voltage and ion transfer tube temperature, are tunable parameters and must be optimized for the individual setup.

Data Analysis

Acquired data has been processed by Spectronaut™ software (Biognosys, v18) using a directDIA approach, DIA-NN (v1.8.1) or Thermo Scientific™ Proteome Discoverer™ software (v3.1.0.618) with CHIMERYS™ intelligent search algorithm by MSAID.

Results

Velocity DIA workflow on Orbitrap Orbitrap Exploris 480 mass spectrometers allows for robust protein identification and quantitation

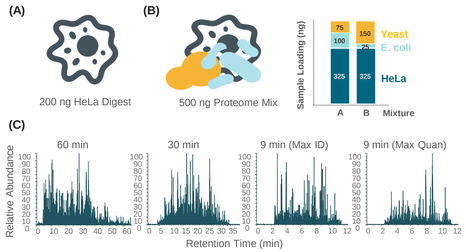
FAIMS Pro interface decreased peptide coverage but increased protein group identifications by an average of 5% (Figure 3). Consequently, the following experiments were all coupled with the FAIMS Pro interface.

With 30 min active gradient, 6,300+ proteins and 40,000+ peptides were identified, along with a protein group CV of approximately 5% (Figure 4), suggesting that the 30 min active gradient method enables relatively high throughput while maximizing identification and quantitative performance. We extended this workflow to 60 min active gradient and successfully identified close to 7,200 proteins and >60,000 peptides, highlighting that deeper proteome coverage can be achieved in the Velocity DIA workflow by using a longer gradient.

We also evaluated a 9 min active gradient to enable ultra-high throughput. We developed two methods to meet different experimental needs: one for maximizing identifications (Max ID) and the other for maximizing quantitative (Max Quan) performance. In the Max ID method, we identify 5,500+ proteins in a 9 min active gradient, highlighting the sensitivity and scanning speed of the Orbitrap mass spectrometer in a high throughput setup. In contrast, the Max Quan method, compromised protein coverage to provide better quantitative performance, as evidenced by ~6% protein group CV in an ultra-high throughput setup (Figure 4).

To understand if a higher load benefits the Velocity DIA workflow in maximizing identification, 500 ng of HeLa digest was analyzed with the 60 min active gradient experiments. The results indicated >7,400 protein groups were identified, with CV being ~ 4% (Figure 5). Thus, the ability of the Orbitrap Exploris 480 MS platform to accommodate sample loads from picograms to micrograms can be leveraged to further increase the depth and quality of measurements.

Figure 2. Experimental sample and active gradient design of the Velocity DIA workflow for label-free quantitation.

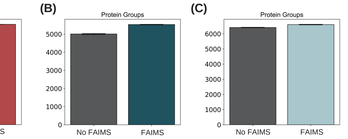


(A-B) Samples to access quantitation performance and proteome coverage. The three-proteome mix contains a medium human background of 325 ng HeLa peptides together with yeast and E. coli peptides digested in ratios of 0.5:1 and 1:4, respectively. (C) Three different active gradient lengths, including 60 min, 30 min, and 9 min, were selected to develop the Velocity DIA workflow.

High accuracy and precision of quantitation

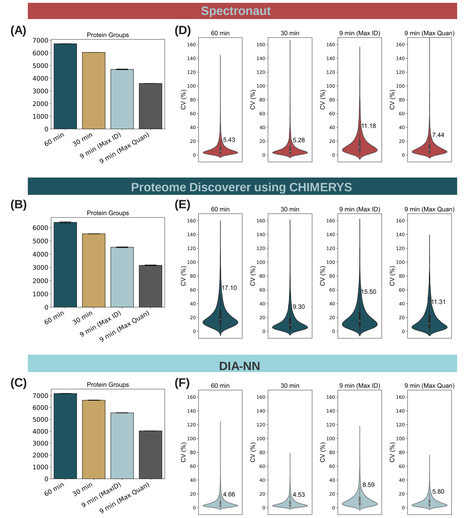
In addition to protein identification, quantitative data is necessary to study biomarkers and get insights into biological pathways. The Velocity DIA workflow yielded excellent quantitative accuracy across a wide dynamic range with median values extremely close to the theoretical ratios, as well as a narrow distribution of all data points around the median values, indicating high quantitative accuracy and precision of the workflow (Figure 6).

Figure 3. Experimental sample and active gradient design of the Velocity DIA workflow for label-free quantitation.



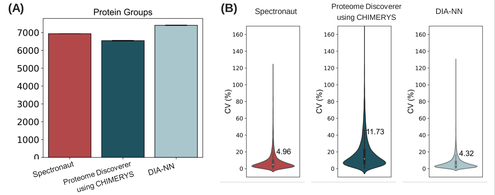
Bar charts depicting the number of proteins identified from 200 ng of HeLa digest in the 30 min method with the FAIMS Pro interface at a CV of -45 V. Data were analyzed with (A) Spectronaut, (B) Proteome Discoverer using CHIMERYS or (C) DIA-NN.

Figure 4. Determination of proteome coverage and quantitation precision of the human proteome.



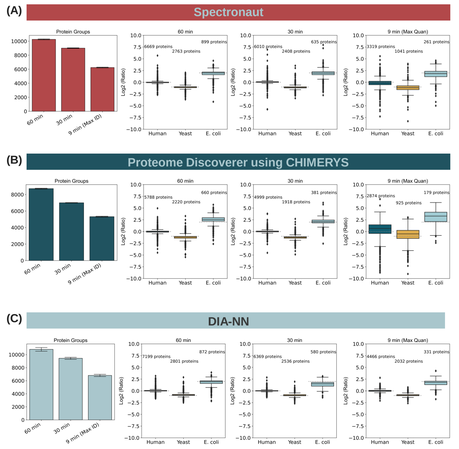
(A-C) Bar charts showing the number of proteins and peptides identified from 200 ng HeLa digests with different active gradients evaluated in the present study. (D-F) Violin plots of all four tested methods reveal high precision of protein quantities in technical replicates. Data was analyzed with Spectronaut (top panel), Proteome Discoverer using CHIMERYS (middle panel) or DIA-NN (bottom panel).

Figure 5. A long gradient coupled with a higher sample load affords the superior performance in the Velocity DIA workflow.



(A) Bar chart of protein groups identified from 500 ng of HeLa digest in the 60 min active gradient method and analyzed with different software demonstrate even deeper proteome coverage. (B) Violin plots reveal high precision of protein quantities in technical replicates. Data was analyzed with Spectronaut (left), Proteome Discoverer software using CHIMERYS (middle), or DIA-NN (right)

Figure 6. Determination of microbial and human protein abundance ratios in a three-proteome mixture.



Bar plots showing the total protein groups identified from three-proteome mix and whisker box plots of protein abundance ratios of all three species demonstrate excellent quantitation accuracy by being consistent with the theoretical ratios (gray dotted line).

Conclusions

The high-resolution DIA workflow for label-free quantitation setup on an Orbitrap Exploris 480 mass spectrometer coupled to a Vanquish Neo UHPLC system running with a 50 cm µPAC Neo UHPLC column was shown to fulfill the following performance criteria:

- Deep (>7000 protein groups) proteome coverage from 200 ng of HeLa standard
- Excellent quantitative accuracy and precision for small amounts of bacterial and fungal proteomes from challenging sample matrices
- Sample throughput and quality of the obtained data while achieving high proteome coverage

Trademarks/licensing

Learn more about Velocity DIA workflow