Single shot deep DIA methods with optimal coverage, reproducibility and quantification precision

Roland Bruderer[1], Oliver M. Bernhardt[1], Tejas Gandhi[1], David Gomez-Varela[2] & Lukas Reiter[1]

[1]Biognosys AG, 8952 Zürich-Schlieren, Switzerland. [2]Signaling and Systems Biology, Max Planck Institute, Goettingen, Germany.

Introduction

Targeted analysis of DIA is a powerful mass spectrometric approach for comprehensive, reproducible and precise proteome quantitation. It provides valuable insight into biological processes or enables the discovery of novel biomarkers. Today, identification of the majority of the expressed proteins can be achieved. With this achievements in the identification, the reproducibility and quantitative accuracy and precision have become increasingly important. Here, we optimize data-independent acquisition (DIA) on a Q Exactive HF mass spectrometer for identification and quantification enabling deep proteome coverage. We optimized methods by varying data point per peak, MS1 and MS2 resolutions. We apply the results of the method optimization to a DIA experiment with a comprehensive spectral library generated from published data. Additionally, murine barrel cortex samples were profiled in distinct stages of sensory development.

Experimental Summary

Peptide samples for HeLa and barrel cortex were prepared using the FASP protocol (Wisniewsk et al.). Biognosys iRT Kit was spiked into the samples before injection.
The sample were acquired using a EASY-nLC 1200 system equipped with either an analytical column (Reprosil Pur C18-AQ, 50cm (or where specified 1m)) or a Thermo Scientific™ EASY-Spray™ column of 75µm ID and 25cm length, packed with 2µm C18 particles, coupled to a Q Exactive HF mass spectrometer. Peptides (2 µg) were separated by segmented gradients.
For the DIA acquisitions, the m/z range covered was 350 - 1650 Th. The number of DIA windows are calculated based on the LC peak width. Full scan AGC target was set to 3e6, IT to 50ms. For DIA windows, stepped NCE was 27+/-10%, the target value 3e6 and maximal injection time set to "auto". For DDA for the spectral library, the instrument was operated in TOP15 mode with the following settings: m/z range 350 -1650 Th; resolution for MS1 scan 60,000; resolution for MS2 scan 15,000; isolation width 1.6 m/z; NCE 27, dynamic exclusion was set to 20s.

Data Analysis

The DIA data were analyzed with Spectronaut (Bruderer et al.), a mass spectrometer vendor independent software from Biognosys, using default settings and 1% FDR. The DDA spectra were analyzed with MaxQuant analysis software and were filtered to satisfy FDR of 1 % on peptide and protein level.

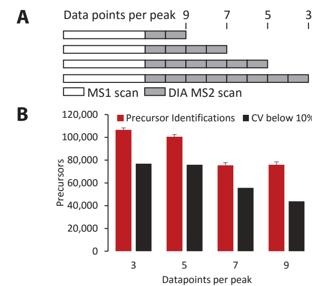
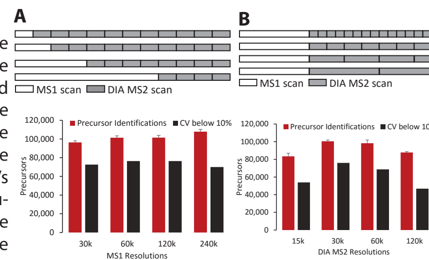


Figure 1: Data points per peak optimization
Sampling of the chromatographic peaks is a prerequisite for robust and high quality identification and quantification. The sampling of the peaks (FWHM 6.6s for a 2h gradient) was performed with 3, 5, 7 and 9 data points per peak (dppp) on average. A spectral library based on HPRP fractionation of HeLa was used for the analysis, it contained 203,435 peptide precursors. **A** Schematic of the DIA methods with varying cycle times. **B** Peptide precursor identification and coefficients of variation (CV) were analyzed. As a benchmark for quantification, all CVs below 10% were counted. For identification, the mean and the standard deviations were plotted.

Figure 2: MS1 and MS2 resolution optimization
The influence of scan resolutions on the performance of DIA-methods was evaluated for 2h gradients. **A** The MS1 scan resolutions were varied from 30k to 240k and the DIA-MS2 resolution was kept constant at 30k. The number of MS2 scans was adjusted to keep the cycle time of the method constant (2.3s and 5 dppp). The means for the identifications and the number of CVs below 10% were plotted. **B** The DIA-MS2 scan resolution was varied at constant 120k MS1 resolution. The identification and quantification performances were compared as before.



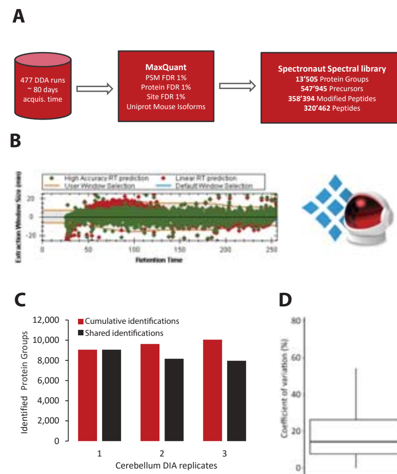


Figure 3: Maximal identification DIA with a resource spectral library
A A comprehensive spectral library for mouse samples was generated from the published resources of Azimifar et al. 2014 and Sharma et al. 2015. These studies used a similar, consistent LC-MS setup as used in this work. The experiments were performed using Reprosil Pur resin with 4h acquisitions on Q Exactive (HF) mass spectrometers. **B** An optimized DIA method was generated for a 4h acquisition using a 1m nanoflow column. The method contained two MS1 segments and 44 MS2 segments. Using a murine Cerebellum sample, triplicate injections were performed and the identifications and CVs were calculated. The median extraction windows was about 11% of the gradient. **C** Cumulative protein identifications increased to 10,058 protein identifications. For ~8000 proteins quantitative values in all three runs were obtained. For example, 515 transcription factors were identified (estimated are about 1,400 in the whole human genome by Vanquerizas et al. 2009). **D** The coefficient of variation on precursor level was 13,7% in median for 101,345 precursors.

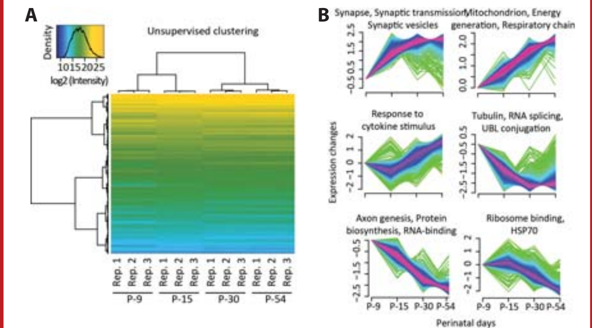


Figure 4: Barrel cortex perinatal development DIA analysis
Five relevant perinatal stages of murine barrel cortex development of mice were selected to monitor proteomic changes. The DIA experiment was performed with 90min gradients and resulted in profiling of over 6,000 proteins and median CVs of 9-14% (sample preparation). **A** Unsupervised clustering was performed on sample and protein level and visualized by a heat map. **B** The quantitative protein data was subjected to c-means clustering based on six clusters. The clustered groups were subjected to GO enrichment analysis.

Conclusions

- DIA Method development for the Q Exactive HF resulted in optimal resolutions of 120k for MS1, 30k for MS2 and 5 dppp for 2h gradients. These acquisitions resulted in 101,000 precursor identifications (6,500 protein groups) for Hela samples.
- Over 9,000 protein groups could be identified and quantified from murine Cerebellum samples using a high quality, comprehensive spectral library from published data.
- Profiling of murine barrel cortex development at perinatal stages enables the quantitative profiling of over 6,000 protein groups and gives insights in the proteomic changes.

References:

- Azimifar SB., Nagaraj N., et al. (2014) Cell-Type-Resolved Quantitative Proteomics of Murine Liver Cell Metabolism 6, 20, 1076-1087
- Bruderer, R., Bernhardt OM., et al. (2015) Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen treated 3D liver microtissues. Mol. Cell. Proteomics, 5, 14, 1400-1410
- Sharma K., Schmitt S., et al. (2015) Cell type– and brain region–resolved mouse brain proteome Nat. Neuroscience, 12, 18, 1819-1830
- Wisniewski, JR., Zougman, A., et al. (2009) Universal sample preparation method for proteome analysis Nat Meth, 6, 359-362
- Vaquerizas, JM., Kummerfeld, SK., et al, 2019 A census of human transcription factors: function, expression and evolution Nature Reviews Genetic, 10, 252-263