Optimized spectral library generation for HRM/DIA/SWATH as implemented within Spectronaut

Tejas Gandhi; Roland M. Bruderer; Claudia Escher; Magdalena Bober; Vito Zanotelli; Oliver M. Bernhardt; Oliver Rinner; Lukas Reiter

Biognosys, Zurich, Switzerland

Introduction

Data-independent acquisition (DIA) with targeted analysis offers new possibilities for highly multiplexed peptide and protein quantification experiments. A high-quality spectral library, acting as a template for the target signals, is a prerequisite for successful targeted analysis. However, the process of generating a refined library is currently a bottleneck. Here we investigate several important parameters of spectral library design for maximum identification power and present its implementation within the Spectronaut software. Specifically, we considered the effect of 1) no fractionation as compared to SDS-Page and strong anion-exchange (SAX) fractionation, 2) accurate retention time, and 3) post processing of fragment ions on both absolute identification of unique peptides and proteins and its reproducibility across replicates of DIA runs.

Methods

All data generated here was based on protein extracts of the human HEK-293 cell line. 24 unfractionated replicates and 6 fractions each using in-gel digestion with fractionation of the SDS-Page lane and strong anion-exchange (SAX) fractionation were measured on a Thermo Q Exactive instrument in shotgun mode. Furthermore, triplicates of unfractionated samples were measured in Hyper Reaction Monitoring (HRM), our DIA workflow. The samples were spiked with the Biognosys HRM Calibration Kit prior to measurement. The shotgun data was searched using MaxQuant 1.3 with a false discovery rate (FDR) of 1% at PSM and protein level. Libraries with varying attributes were created with Spectronaut 5.0 and compared based on the numbers of identified peptides from HRM data with an FDR of 1%.

Figure 1. Spectronaut software's Prepare perspective offers a user-friendly spectral library generation pipeline.

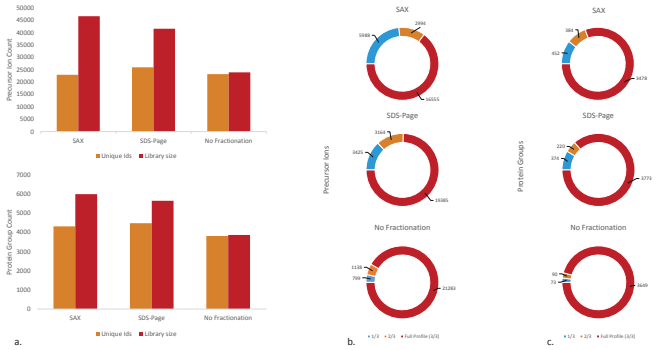

Figure 2.
a. Comparison of precursors and protein groups identification in 3 DIA runs of unfractionated sample using a spectral library derived from 6 shotgun runs prepared with SAX, SDS-Page, and no fractionation each.
b, c. Reproducibility of precursor ions and protein group identification in the 3 DIA runs of unfractionated sample based on the fractionation strategy used during library generation. Full profile indicates identification in all three DIA runs.

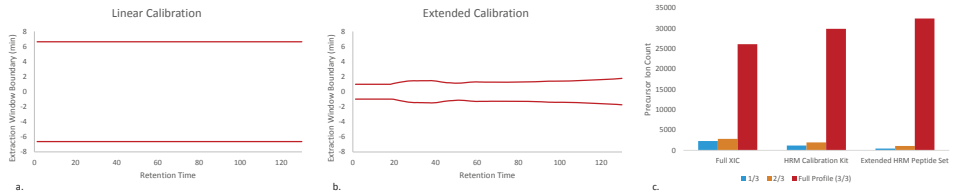

Figure 3.
a, b. Comparison of ion trace extraction windows in 3 DIA runs using a library generated with 1) linear retention time calibration using HRM Calibration Kit peptides, 2) a non-linear calibration using an Extended HRM Peptide Set for HEK-293 cell line consisting of abundant precursors routinely identifed using shotgun.
c. Finally, the resulting precursor identifications of these two libraries are compared with when not using a HRM Calibration Kit and performing a full ion trace extraction. All three libraries were generated from 24 shotgun runs with only the precursor iRT information differing amongst them.

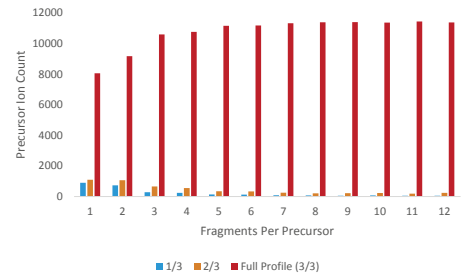

Figure 4.
Comparison of the effect of the number of fragments per precursor ion in a spectral library on identification in DIA runs. A spectral library of exactly 12 most intense fragments was generated from 24 shotgun runs. Subsequently, 11 more libraries were generated by subtracting the lowest intense fragment each time till a library of exactly 1 fragment per precursor remained. This resulted into 12 libraries with same precursors but differing number of fragments per precursor. 3 DIA runs were used to compare the performance of these libraries in terms of identification.

Conclusion

A spectral library generation pipeline involves multiple steps ranging from sample preparation to post processing. The take home messages presented here are as follows:

1) Our results show that unfractionated library not only leads to a comparable number of identifications (Figure 2a), but the identifications are more reproducible (Figure 2b, 2c). As such, performing fractionation during library generation is not necessarily beneficial if the sample for the DIA runs is also not fractionated. We argue that amount of extra effort required for fractionation could be instead used for performing more unfractionated replicates. As a rule of thumb, sample preparation for library generation should be similar to the DIA runs.

2) Using iRT scale in a spectral library outperforms full ion trace extraction in terms of identification by returning 24% more full profiles when using the Extended HRM Peptide Set (Figure 3c). In general, HRM Calibration Kit plays a critical role in the performance of a spectral library by allowing for a more finely tuned ion trace extraction window.

3) At least 3 fragments per precursor ion is recommended for full profile identification with gains flattening out when using more than 8 fragments per precursor ion (Figure 4).

4) Spectronaut software offers a powerful and intuitive solution for generating spectral libraries for DIA experiments (Figure 1). It supports the following key features:

- Automated determination of iRT
- Normalization of relative fragment intensities across replicates and selection of best fragments
- Selection of the best peptide spectrum match across replicates

References

[1] Bernhardt O. M. et al. "Spectronaut A fast and efficient algorithm for MRM-like processing of data independent acquisition (SWATH-MS) data." ASMS 2012.

[2] Escher C, Reiter L et al. "Using iRT, a normalized retention time for more targeted measurement of peptides." Proteomics. 2012 Apr;12(8):1111-21.