Application Note

EVOSEP

Comprehensive proteomes with **8,000 proteins** identified with dia-PASEF using the **30 samples per day method**

1. Introduction

Data-independent acquisition based proteomics measurements in combination with the Evosep One are emerging as a powerful workflow for deep and accurate mapping of proteomes across large sample cohorts. Here, we demonstrate the depth of proteome coverage we can achieve when maximizing chromatographic separation with the performance column (EV1137) matched with the 30 samples per day method. This is accelerated by dia-PASEF on the timsTOF Pro 2 mass spectrometer (Figure 1). The presented workflow extends single-shot proteome coverage to 8,000 proteins identified 30 times a day.

Figure 1: Workflow for deep proteomes with the 30 samples per day method.

2. Method details

HeLa tryptic digest was purchased from Pierce and 400 ng loaded onto Evotip Pure using the standard loading protocol. Samples were analyzed with either the Endurance column (EV1106 at ambient temperature ~22 °C) or the Performance column (EV1137 at 40 °C). This was coupled to a 10 µm Zero Dead Volume Captive Spray Emitter (Bruker #1865691) and a timsTOF Pro 2 mass spectrometer (Bruker), using the 30 samples per day method. MS acquisition was performed by default application methods in timsControl. A deep sample-specific spectral library was generated from 96 high pH reversed-phase fractions of the input sample by dda-PASEF acquisition (DDA PASEF short gradient method with the tims range lower limit set to 0.8). dia-PASEF was carried out in quintuplicates using the 1.1s cycle time dia-PASEF scheme. An m/z range from 400 –1200 m/z and 1/K₀ range of 0.6-1.43 is covered by 16 diaPASEF scans with 2 ion mobility steps and 26 m/z window width per scan, resulting in ~1.1s cycle time (Figure 2).

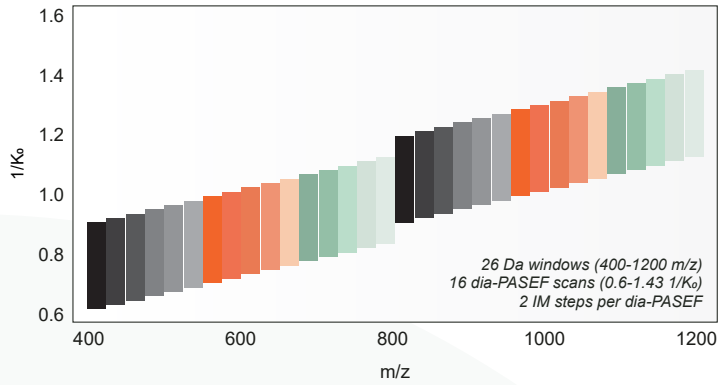

Figure 2: diaPASEF MS acquisition window placement scheme.

3. Data analysis

The spectral library was generated in Spectronaut (v15) using Pulsar with default settings. The UniProt/SwissProt protein sequence database at isoform resolution and a list of common contaminants were used. Carbamidomethylation of cysteine residues was defined as a fixed modification, and methionine oxidation and acetylation of protein N-termini were defined as variable modifications. The FDR was controlled at <1% at both the peptide spectrum match- and the protein level. The diaPASEF data was analyzed with DIA-NN v1.8 in spectral library search mode with default settings. We used the same fixed modifications as before with 2 missed cleavages. Peptide length was adjusted to spectral library of a peptide length 7 to 50. Precursor charge range was set to 2-4.We used a precursor FDR of <1% and fixed mass accuracies of 15.0 ppm on both MS1 and MS2 level. MBR (match-between-runs) was enabled and protein inference was set to 'Off'. The quantification strategy was set to 'Robust LC (high accuracy) and Library generation to 'IDs, RT & IM profiling'. Protein numbers represent unique protein groups from the DIA-NN matrix report.

4. Results

We compared the two columns by analyzing 400 ng HeLa peptide loaded on Evotips in quintuplicates and analyzed these with the 30 samples per day method in dia-PASEF mode on a timsTOF Pro 2. We achieved excellent sequencing depth with 53,000 and 68,000 precursors identified for the Endurance and Performance column respectively (Figure 3A).

This corresponds to 6,382 and 7,209 proteins representing the proteome depth, which is possible to obtain without the need to record a project-specific spectral library (Figure 3B). The quantitative precision across the proteins identified with the Endurance column is good with a median coefficient of variation (CV) of 8.8%. In addition to improved identifications, the precision of quantification was also improved with the Performance column as reduced CVs with a median of 7.2% was calculated from the replicate injections (Figure 3C). Overall, the DIA-NN analysis with the library-free mode provides excellent coverage and quantification with both columns.

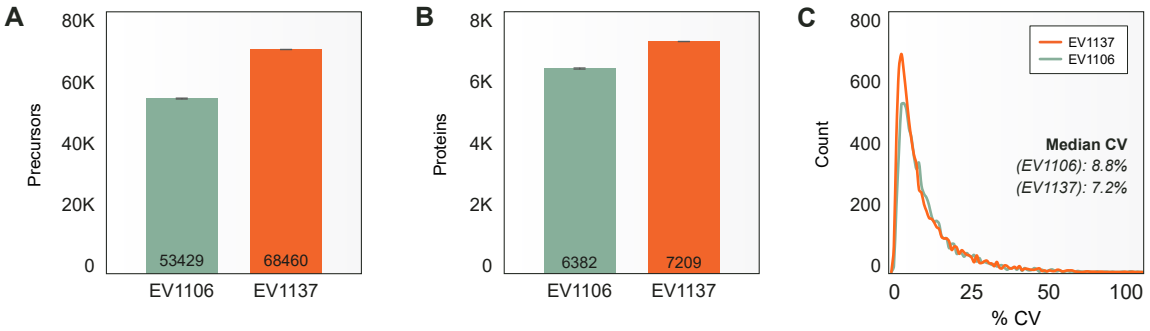

Figure 3: (A) Precursor and (B) protein identifications from library-free analysis of HeLa dia-PASEF runs at 30 SPD. (C) Coefficient of variation distributions for the two columns.

Next, we investigated if we could boost the proteome coverage by using a project-specific spectral library. This was generated by extensive high pH reversed-phase fractionation, where each of 96 fractions were analyzed by PASEF using the same 30 SPD method and the Performance column. We analyzed 30 replicate injections of 400 ng HeLa with the Performance column, requiring one day of instrument time and obtained a coverage of more than 80,000 precursors (Figure 4A) on average corresponding to roughly 7,800 proteins from each injection (Figure 4B). This results in a comprehensive coverage of more than 8,000 proteins total from the 30 replicate injections. This dataset demonstrates a robust readout across the injections and that the proteome coverage is boosted around 8%, when ideal search conditions are used with a project-specific library. Again, the quantitative precision is stable with 5,987 and 3,904 proteins quantified with a CV below 20% and 10% respectively, calculated from proteins identified in all 30 injections (Figure 4C).

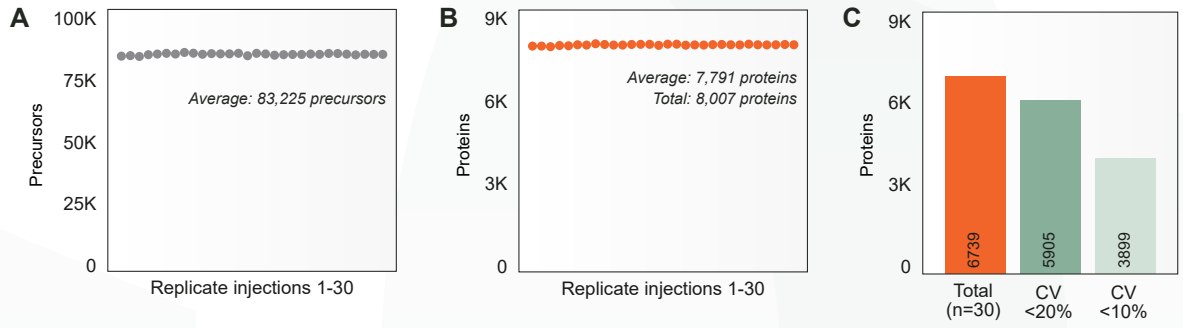

Figure 4: (A) Precursor and (B) protein identifications from a project-specific spectral library with the Performance column and (C) overall precision of quantification calculated from proteins identified in all 30 injections.

5. Conclusion

The Performance column (EV1137) provides a powerful option to improve the chromatographic performance compared to the Endurance column (EV1106). Reduced peak widths lead to increased peak capacity and optimized mapping of deep proteomes using the 30 SPD method. Thus, the single-shot proteome coverage is enhanced to nearly 8,000 proteins identified and quantified in dia-PASEF mode, when a project-specific library is used for the analysis. When comparing this analysis work flow with an easy-to-use library-free analysis, we still cover 90% of the identifications, which is an advantage for projects where it is not possible to acquire a deep project-specific library. Thereby, the standardized 30 SPD method on the Evosep One combined with data-independent acquisition provides a powerful platform for diverse biological applications where great proteome depth can be balanced with high throughput.

References

1. Bache N., Geyer PE., Bekker-Jensen DB., Hoerning O., Falkenby L., Treit PV., Doll S., Paron I., Müller JB., Meier F., Olsen JV., Vorm O., Mann M. (2018) A novel LC system embeds analytes in preformed gradients for rapid, ultra-robust proteomics. Mol Cell Proteomics., mcp.TIR118.000853

Data Courtesy

Marvin Thielert, Andreas-David Brunner and Matthias Mann

Max Planck Insitute for Biochemistry, Munich, Germany

Ordering information

P/N	PART	DESCRIPTION
EV1106	Endurance Column 15 cm x 150 µm ID, 1.9 µm	Used by the 30 SPD & Extended methods. Analytical column with pre- mounted connection fittings. ReproSil-Pur C18, 1.9 µm beads by Dr Maisch.
EV1137	Performance Column 15 cm x 150 µm ID, 1.5 µm	Used by the 30 SPD & Extended methods. Analytical column with pre- mounted connection fittings. ReproSil Saphir C18, 1.5 µm beads by Dr Maisch.