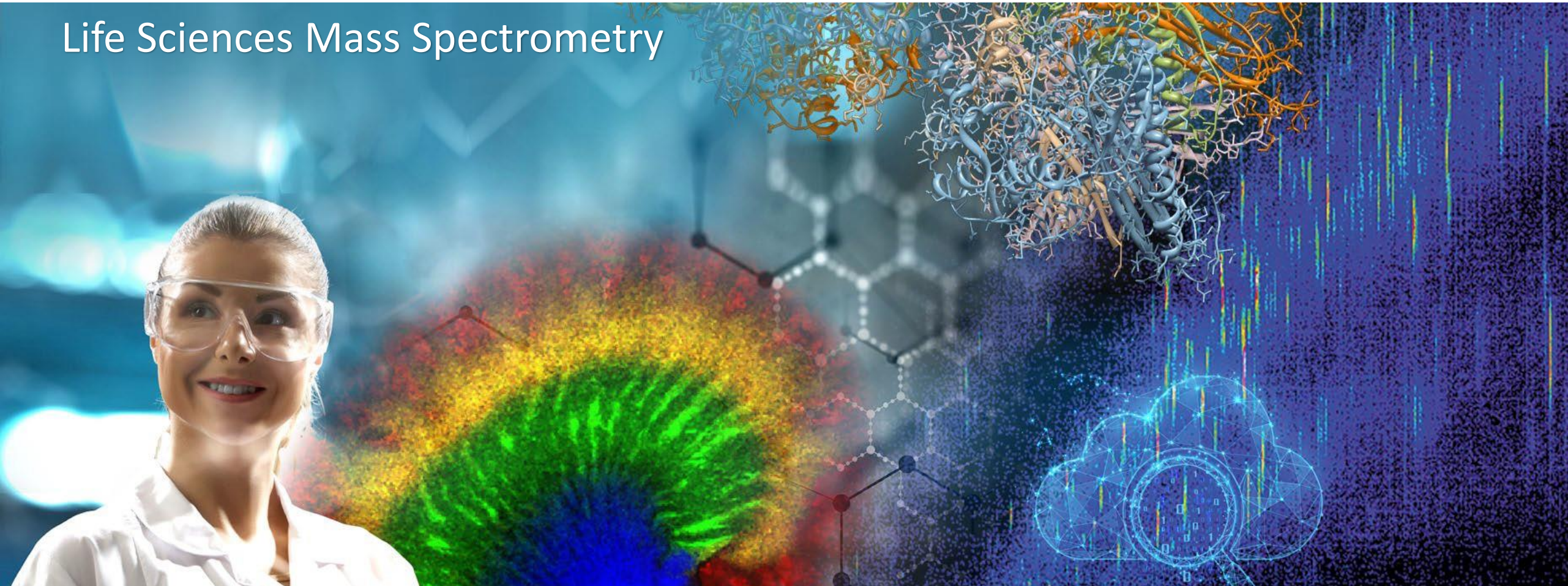
High throughput proteomics Application of dia-PASEF and Evosep One for short gradients

Stephanie Kaspar-Schoenefeld[1], Markus Lubeck[1], Thomas Kosinski[1], Scarlet Koch[1], Oliver Raether[1], Gary Kruppa[1]
[1]Bruker Daltonik GmbH, 28359 Bremen, Germany



High throughput proteomics – Application of dia-PASEF and Evosep One for short gradients

Stephanie Kaspar-Schoenefeld[1], Markus Lubeck[1], Thomas Kosinski[1], Scarlet Koch[1], Oliver Raether[1], Gary Kruppa[1]

[1]Bruker Daltonik GmbH, 28359 Bremen, Germany

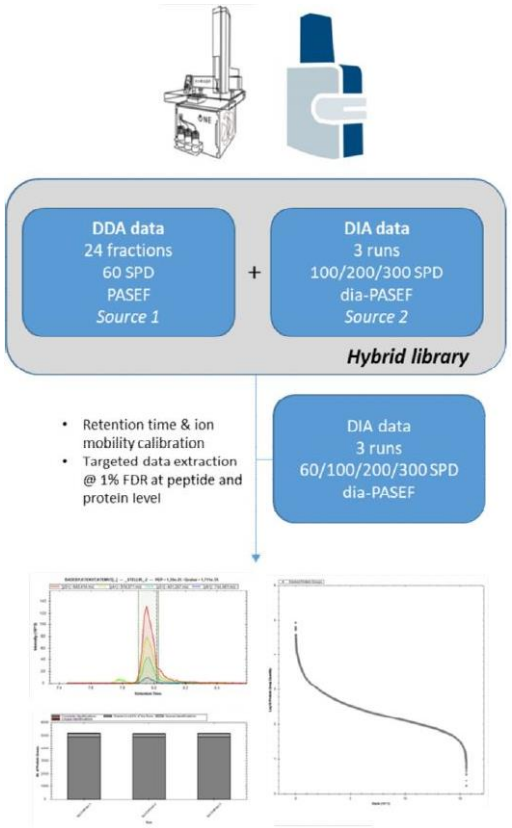
Introduction

Data-independent acquisition (DIA) promises reproducible and accurate protein identification and quantification across large sample cohorts by using wide selection windows, rather than selecting individual peptides in DDA, to ensure that all precursor ions are fragmented in every sample. Ion mobility separation provides an additional dimension for separation of complex proteomics samples, that can also be used for alignment of precursor and fragment information. Here, we combine the PASEF technology (Meier et al., MCP, 2018, 17(12):2534-2545) with a DIA approach, called dia-PASEF (Meier et al., 2019), and investigate the potential for complex proteomics samples using very short gradients.

Methods

An in-house tryptic digest of HeLa was analyzed by coupling an Evosep One system (Evosep Biosystems) online to a trapped ion mobility spectrometry – quadrupole time of flight mass spectrometer (timsTOF Pro, Bruker Daltonics). A dia-PASEF scheme optimized for the short gradient methods has been used targeting +2 and +3 ions in a three-window method (100ms).

Eight of these scans (resulting in 24 windows, each using 25 Da window size) covers an m/z range from 400 to 1000. Each dia-PASEF cycle includes 1 MS scan (100ms) resulting in a total cycle time of 900ms per dia-PASEF cycle. Data processing was done using Spectronaut 14 (Biognosys, Figure 1).



Results

DIA workflows rely on spectral libraries for the correlation of quantitative data from fragment ion spectra with peptide identifications. We used PASEF in DDA mode and fractionated samples to assemble the resource-specific library using Spectronaut software. The library comprised 8,381 protein groups and 93,301 peptide sequences. Data acquisition time for library generation was ~10 hours using the 60 SPD (samples per day) method. We applied the hybrid library approach supported in Spectronaut (Muntel et al., Mol. Omics, 2019, 15, 348) for the shorter gradient data (100, 200, 300 SPD) by combining the resourcespecific library with a project-specific library obtained from corresponding dia-PASEF runs using directDIA processing approach. This workflow allows keeping retention time precision of the shorter gradients for targeted data extraction.

Fig. 1: Workflow for dia-PASEF data processing. Resource-specific library data was acquired using the 60 SPD method in combination with PASEF technology. Raw data from 24 fractions were analyzed with Pulsar. Project-specific libraries have been created for the 100,200 and 300 SPD methods using dia-PASEF data to allow for source-specific retention time alignment. Both libraries have been combined to so-called "Hybrid Libraries" used for targeted data extraction from the dia-PASEF data.

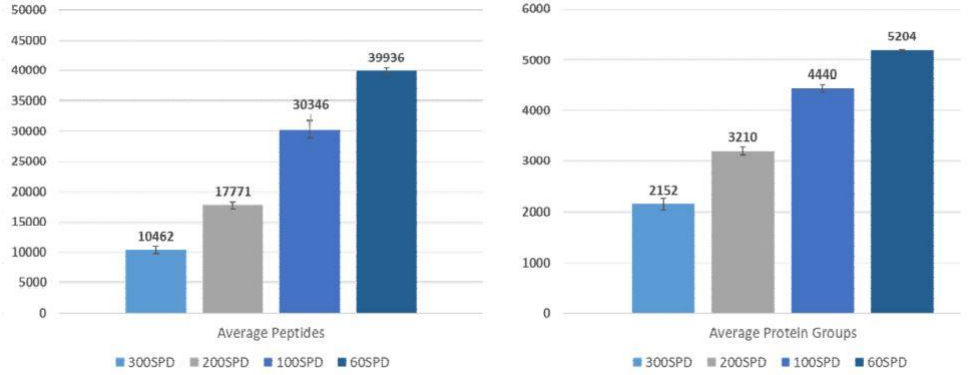


Fig. 2: Average number of identified peptides and protein groups (1% FDR) from triplicate injections of 200ng HeLa cell digest for the different gradients (60, 100, 200, 300 SPD)

Using the comprehensive libraries, we could identify and quantify on average 5,204 protein groups and 39,936 peptide sequences using 60 SPD method at a 1% FDR (Figure 2). The identified proteins cover a dynamic range of around 5 orders of magnitude. The median coefficient of variation was at 8.2% for peptide and 5.9% for protein group level for the triplicate injections. We further decreased run time to improve throughput using the 100, 200, and 300 SPD methods on the Evosep system. When analysing HeLa sample using the highest throughput method (300SPD), we were able to identify more than 2,100 protein groups and 10,462 peptide sequences on average in just 3 min gradient time (5 min total run time). In summary, application of dia-PASEF for very short results in excellent proteome coverage for HeLa digest.

Conclusion

- timsTOF Pro using dia-PASEF in combination with the Evosep One system is ideally suited for highthroughput proteomics:
  - On average 5,200 protein groups can be identified and quantified in single-shot analysis using a short LC gradient (60 samples per day)
  - Using a 300 SPD method (3 min gradient, 5 min total run time) dia-PASEF identifies and quantifies more than 2,000 protein groups per run.

Questions and Answers