Evaluation of dia-PASEF for short gradient using library and library free approaches

Diego Assis[1], Elizabeth Gordon[1] and Matthew Willetts[1]

1. Bruker Daltonics, Billerica, MA, USA

ASMS 2021, MP 024

Introduction

dia-PASEF (Meier et.al.,2020) takes advantage of the additional dimension of separation provided by trapped ion mobility for the analysis of complex proteomics samples by data independent analysis (DIA). Additionally TIMS separation increases selectivity, excludes singly charged precursors from fragmentation and cleans up the sample by concentrating signals from noise. Making use of the correlation of molecular weight and CCS coded information from the dual-TIMS funnel, dia-PASEF enables most confident compound identification. Over the entire LC-MS/MS dia-PASEF runs a perfect data cuboid is created containing m/z, ion mobility (CCS), retention time and intensity. Here, to evaluate these benefits of dia-PASEF, we compared gradient lengths and results from two independent software platforms which can process native dia-PASEF data using spectral libraries or a library-free approach.

Methods

K562 tryptic digest (Promega) was analyzed by coupling an EVOSEP One (EVOSEP) or nanoElute (Bruker) system to a trapped ion mobility spectrometry – quadrupole time of flight mass spectrometer (timsTOF Pro 2). Data were acquired in DDA-PASEF mode and dia-PASEF using gradient lengths of 300, 200, 100, 60, 50, 30 and 15 samples per day (SPD) (fig. 1). dia-PASEF schemes were optimized for the short and long gradients covering m/z range from 400-1000 and distinct mobility ranges (fig 2). DDA-PASEF data were processed with PASER (Bruker). Spectronaut 15 (Biognosys) and DIA-NN 1.8 (Demichev et.al, 2021) were used for DIA data processing using both spectral library and library-free approaches.

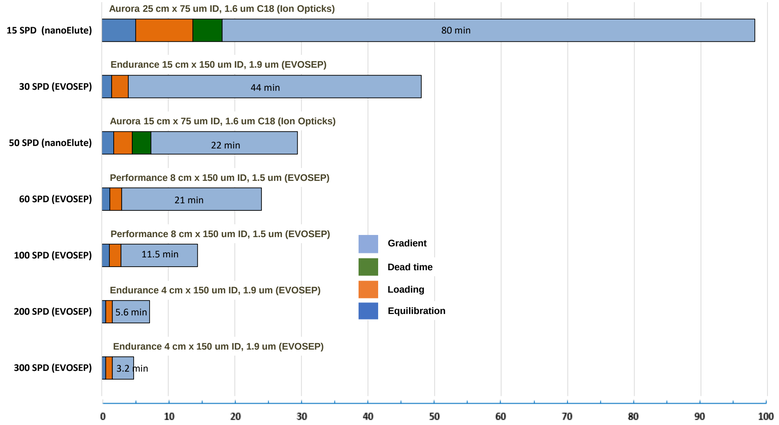

Figure 1. LC methods showing columns, equilibration, loading and gradient time.

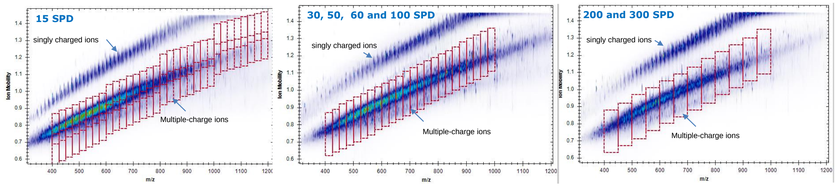

Figure 2. dia-PASEF windows scheme (m/z versus mobility) for three different methods. Ion clouds showing different charge states. Graphics generated on Spectronaut 15.

Results

Initially, a K562 cell line spectral library was created by collecting multiple DDA runs totaling 8,018 protein groups and 116,870 peptides with Spectronaut. Then, the number of identified proteins were compared using different gradients, software, and library on dia-PASEF and DDA mode (fig. 3).

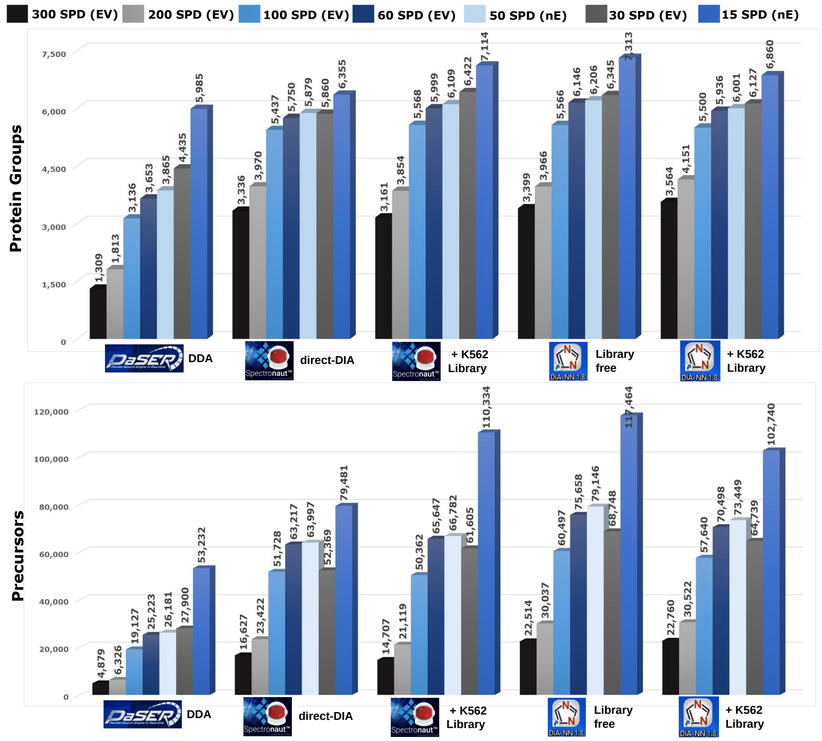

Figure 3. Identified protein groups and precursors (different gradients, approaches and software).
*All data were collected in quadruplicates with CVs < 3% but plots show average values. **EV** – Evosep One, **nE** - nanoElute

Compared to DDA mode, dia-PASEF resulted in a higher number of identifications even using very short gradients. Figure 4 demonstrates similar performance for DIA and DDA at the protein level for the longest gradient (ratio close to 1). However, drastically reducing the gradient time (blue trace), dia-PASEF identifications showed a 2.5-fold increase for the shortest gradient (orange trace on the left plot). At the peptide level, the number of identified precursors is doubled with 15SPD when dia-PASEF is used and increases to a ratio of 4.2 with 300 SPD method.

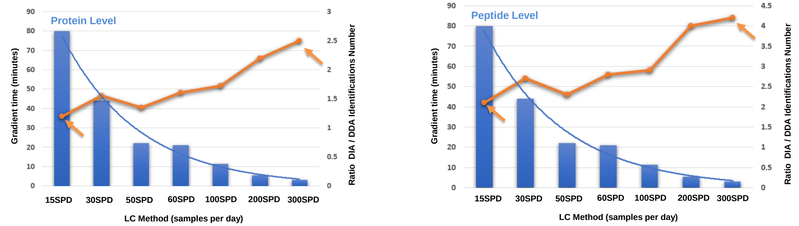

Figure 4. Identification ratio (orange trace) for dia-PASEF vs DDA as a function of gradient length (blue trace).

Conclusions

- Results show dia-PASEF maximizes the number of identifications. Even using very short gradient of 3.2 minutes the timsTOF Pro 2 was able to identify more than 3,500 protein groups and 22,000 precursors. Number of identified precursors is at least doubled when compared to DDA.
- *Library free* and *library approach* performed similarly on both software platforms. Thus, the library free workflow seems to be a good alternative in experiments where library creation is not possible.
- With *library free at the protein level*, DIA-NN and Spectronaut were similar. However, at the peptide level DIA-NN identified 20% more precursors.
- Using the *library approach at the protein level*, DIA-NN and Spectronaut were similar. *At the peptide level* with gradients of 15, 30, 50 and 60 SPD, numbers of IDs are similar but interestingly, when moving from the longest to shorter gradients, specially 100, 200 and 300 SPD, DIA-NN showed a higher number of identified precursors (additional 14, 42 and 50%, respectively).
- Quantitative experiments using these gradients are being performed to evaluate accuracy, coefficient of variations and number of datapoints per peak. Initial results suggests an accurate quantification for all methods on both software.
- Data presented demonstrates the benefits of using dia-PASEF acquisition for high-throughput, deep proteome studies, using both library and library-free approaches.

timsTOF Pro 2