HUPO 2024 P-I-0143

Low-variation proteome profiling across 11 Labs, identifying 7200+ protein groups in 5 minutes with dia-PASEF

Stephanie Kaspar-Schoenefeld[1], Tobias Kroniger[1], Laura Heikaus[1], Ann-Christine König[2], Stefanie Hauck[2], Sebastian Johansson[3], Axel Karger[4], Uli Ohmayer[5], Matteo Pecoraro[6], Stefan Tenzer[7], Ute Distler[7], Sophie Braga-Lagache[8], Jonathan R. Krieger[1], Philip Strohmiedel[1], Lisa Abel[1], Raphael Schuster[1], Georg Kliewer[1], Diego Assis[9], Torsten Mueller[1], Daniel Hornburg[10]

[1]Bruker Daltonics GmbH & Co. KG, Bremen, Germany, [2]Helmholtz Center, Metabolomics and Proteomics Core, Munich, Germany, [3]Preomics GmbH, Planegg, Germany, [4]Friedrich-Loeffler-Institut, Institute of Molecular Virology and Cell Biology, Greifswald, Germany, [5]NEOsphere Biotechnologies GmbH, Planegg, Germany, [6]Università della Svizzera italiana, Bellinzona, Switzerland, [7]University Medical Center of the Johannes-Gutenberg University Mainz, Institute for Immunology, Mainz, Germany, [8]Proteomics and Mass Spectrometry Core Facility, Department for Biomedical Research, University of Bern, Switzerland, [9]Bruker Scientific, Billerica, USA, [10]Bruker Scientific, LLC, San Jose, USA

Introduction

Data independent acquisition (DIA) approaches have attracted attention as these eliminate stochastic changes resulting from data dependent acquisition (DDA) schemes, thereby maximizing reproducibility. In dia-PASEF, all ions in a certain mass and ion mobility window are co-isolated by the quadrupole and hence co-fragmented. It is particularly well-suited for large cohort studies, which require robust and highly reproducible workflows in conjunction with short runtimes and high proteome coverage.

Here, we combined dia-PASEF acquisition with very short gradient lengths (5-minute active gradient) enabling the acquisition of ~140 samples per day and investigated its reproducibility across 11 instruments in different laboratories.

Methods

11 laboratories geographically located throughout Europe and North America were provided with commercial reference sample (K562, Promega) and a protocol (Figure 1A) for sample preparation, instrument set-up and an optimized dia-PASEF window scheme (Figure 1B). Each lab acquired ten replicates. For protein identification, all data files were processed in a library-free (directDIA+) approach with Spectronaut 19 (Biognosys).

Reproducibility across all labs was compared based on number of identified protein groups, identified peptides and respective coefficients of variation (CVs).

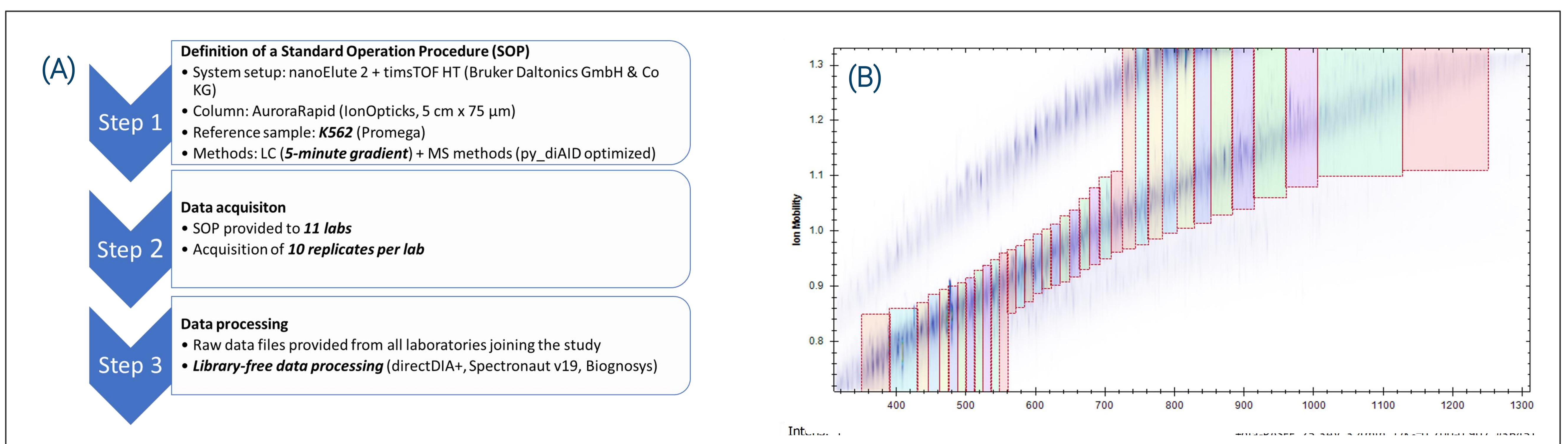

Figure 1: High-throughput proteome profiling in a multi-center study.(A) Detailed study overview. (B) dia-PASEF window scheme consisting of 12 frames with 3 mass windows per frame, optimized with py_diAID.

Results

When running such short gradients, the applied DIA methods' cycle time must match the fast chromatography (narrow peaks of eluting peptides) to sample sufficient data points across the peaks to enable good quantification. Using dia-PASEF on the timsTOF HT platform allows optimal window placing to ensure fragmentation of all theoretical precursors, while maintaining coverage of the chromatographic peak. Here we used py_diAID [1] which adjusts the isolation window width to the precursor density, automatically generating the optimal isolation window scheme in the mass to ion mobility space. By encompassing three mass windows per PASEF frame and implementing brief accumulation and ramp times of 50 ms on the TIMS device, we attained a cycle time of 0.73 s, including 1 MS1 scan. The chromatographic peaks had a median width of about 3 s. Thus, the applied dia-PASEF method resulted in an average peak coverage of 4 to 5 data points in a 5-minute gradient.

Ten replicates were acquired per site resulting in a data set consisting of 110 files. From the 5-minute gradient, on average 7,072 protein groups and 99,835 peptides were identified at 1% FDR using library-free data processing (Figure 2A and C). In total, 7,228 protein groups and 124,064 peptides were identified across the entire data set. The number of proteins identified with CV values below 20% and 10% ranged from 96% to 88%, respectively, across all 110 data sets.

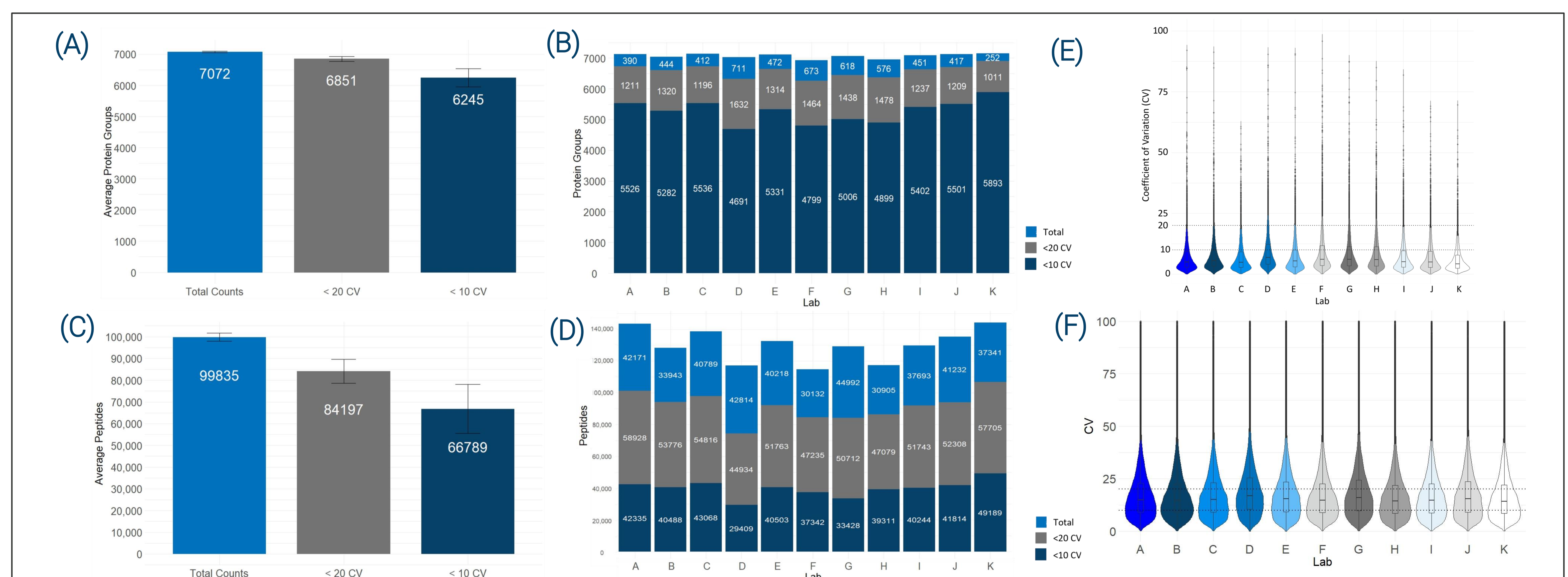

Figure 2: Number of identified protein groups from analysis of 200 ng K562 using a 5-minute gradient on the timsTOF HT. (A) Average number of identified protein groups across all 110 data sets (11 labs, 10 replicates per lab) and number of quantified protein groups with CV values below 20% (grey) and 10% (dark blue). (B) Number of identified protein groups per lab (total, quantified with CV values below 20% and 10%). (C) Average number of peptides across all 110 data sets (11 labs, 10 replicates per lab) and number of quantified peptides with CV values below 20% (grey) and 10% (dark blue). (D) Number of identified peptides per lab (total, quantified with CV values below 20 and 10%). (E) Coefficient of variation for the 10 replicates per laboratory on protein group level. (F) Coefficient of variation for the 10 replicates per laboratory on peptide level

Average CV values for each lab on protein group level are well below 10% for the 10 replicates (Figure 2E). . The median CV value across the complete experiment (including all 110 runs) was at 12.1% on protein group level. In total, 4616 protein groups have been identified in all 10 replicates of all labs participating in this study (corresponding to 110 data sets).

Thus, dia-PASEF results in excellent and reproducible intra- and inter-laboratory performance making the presented workflow ideally suitable for routine proteomics applications.

- Highly **robust** LC-MS system ideal for **high throughput proteomics** and **inter-laboratory studies**
- **dia-PASEF** on the **timsTOF HT** enables **high proteome coverage** and **accurate quantitation** in short gradients of 5 minutes
- More than **7,000 protein groups** from approximately **100,000 peptides** can be **reproducibly** identified on average across 11 different labs

nanoElute 2 + timsTOF HT

Conflict of interest: SK-S, TK, LH, JRK, PS, LA, RS, GK, DA, TM and DH are employees of Bruker Daltonics. ST, UD and Bruker Daltonics are collaboration partners.

1) Skowronek et.al. Mol Cell Proteomics (2022) 21(9) 100279

© 2024 Bruker

For Research Use Only. Not for use in diagnostic procedures.

Innovation with Integrity