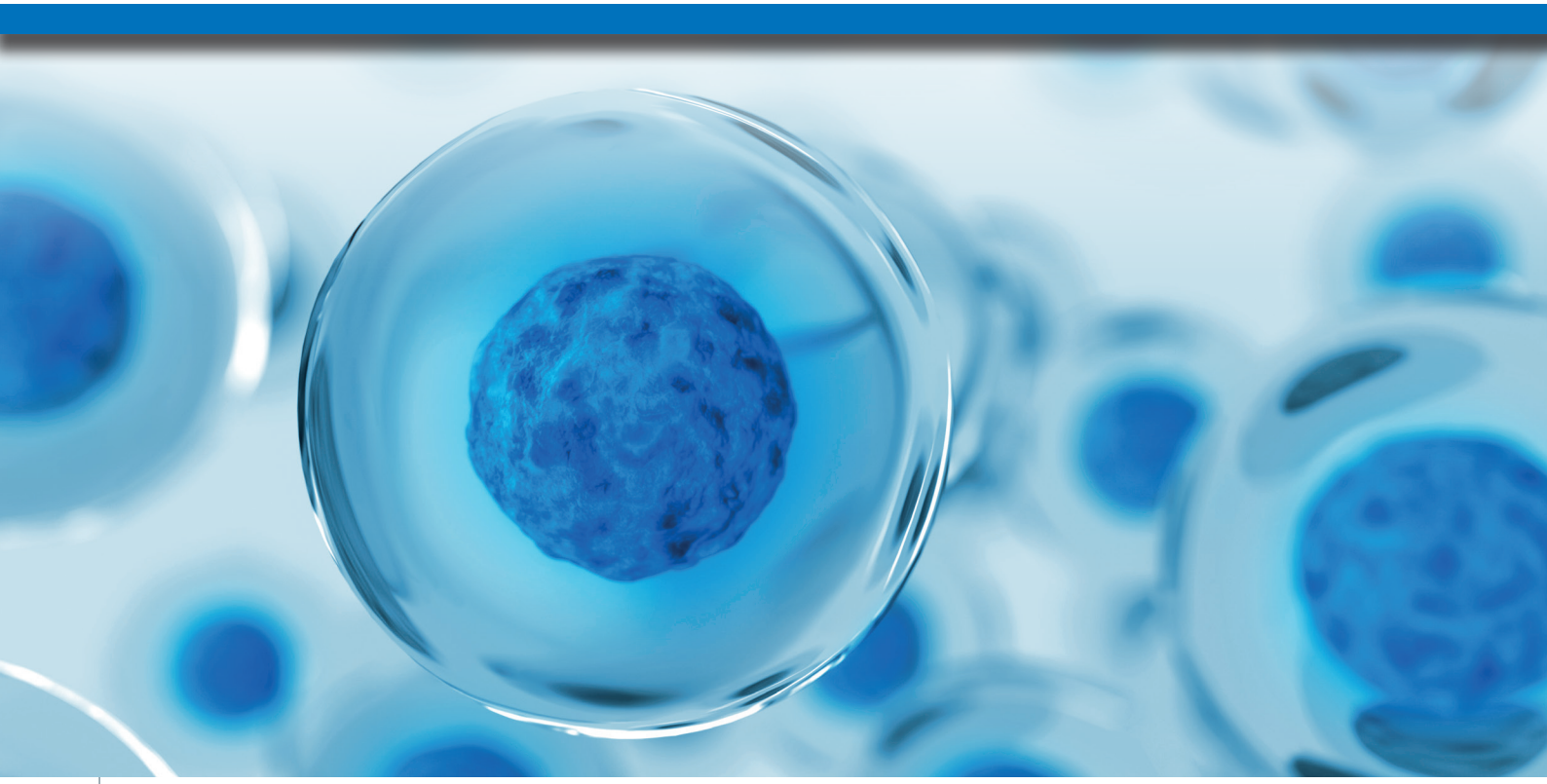

Parallel accumulation - serial fragmentation combined with data-independent acquisition (diaPASEF)

The timsTOF Pro with diaPASEF enables deeper proteome coverage in a single 4D shotgun proteomics experiment, with highly reproducible qualitative and quantitative results – making it a near-ideal mass analyzer for proteomics.

Abstract

Data-independent acquisition (DIA) workflows have gained in popularity as they overcome the issue of stochastic selection of peptide precursors encountered in typical data-dependent approaches (DDA). The success of DIA relies on key instrumental capabilities, namely: resolution, sensitivity, accuracy and dynamic range uncompromised by a fast-spectral acquisition rate (>40 Hz) demanded by DIA. Trapped ion mobility spectrometry adds an additional dimension for separation of complex proteomics samples. In addition, the collisional cross section (CCS) term allows for unbiased alignment of precursor and fragment information. In this application note, we combine Parallel Accumulation Serial Fragmentation (PASEF, [1, 2]) with a DIA approach, called

Keywords: Proteomics, diaPASEF, timsTOF Pro, PASEF, DIA

Authors: Stephanie Kaspar-Schoenefeld[1], Florian Meier[2], Andreas-David Brunner[2], Max Frank[3], Annie Ha[3], Markus Lubeck[1], Oliver Raether[1], Ben Collins[4], Ruedi Aebersold[4, 5], Hannes Rost[3], Matthias Mann[2, 6];
[1]Bruker Daltonik GmbH, Bremen, Germany; [2]Max Planck Institute of Biochemistry, Martinsried, Germany; [3]Donnelly Centre for Cellular and Biomol. Research, Toronto, Canada; [4]ETH Zurich, Zurich, Switzerland; [5]University of Zurich, Zurich, Switzerland; [6]NNF Center for Protein Research, Copenhagen, Denmark.

diaPASEF, to investigate its potential for the in-depth analysis of complex proteomics samples. In brief, more than 7,000 proteins could be reliably identified and quantified at a false discovery rate of 1% in each of triplicate samples consisting of a HeLa protein digest. For each injection 200 ng total peptide mass was loaded on column and 120 minute LC-MS runs were applied. Fragment ion-based quantification was very reproducible with a median coefficient of variation (COV) of 10%.

Introduction

Mass spectrometry (MS)-based proteomics has become a powerful technology for the identification and quantification of thousands of proteins [3]. However, the coverage of complete proteomes remains very challenging. State-of-the-art mass spectrometers efficiently transfer ions into the vacuum, but only use a small fraction of the ion beam for mass analysis. The recently launched timsTOF Pro achieves a nearly 100% duty cycle by parallel ion storage and sequential release from a trapped ion mobility (TIMS) device into a quadrupole time of flight mass analyzer. Synchronizing the release of ions from the TIMS device with the quadrupole (PASEF) increases the MS/MS sequencing speed more than 10-fold while simultaneously boosting sensitivity in online DDA experiments [1, 2]. Nonetheless, even with the high >100 Hz MS/MS speed of the timsTOF Pro, not all peptides can be targeted for MS/MS identification. Due to the high complexity of proteomic samples further complicated by biological variation, the same peptides may not be targeted in all runs, exacerbating the so-called "missing value problem" when performing quantitative proteomics.

DIA is an alternative approach which promises reproducible and accurate protein quantification across large sample cohorts by using wide selection windows. The mass spectrometer typically cycles through many isolation windows, seamlessly covering the m/z range of interest. Narrow isolation windows increase specificity, but also entail longer cycle times when sampling the same mass range. To date the method has been

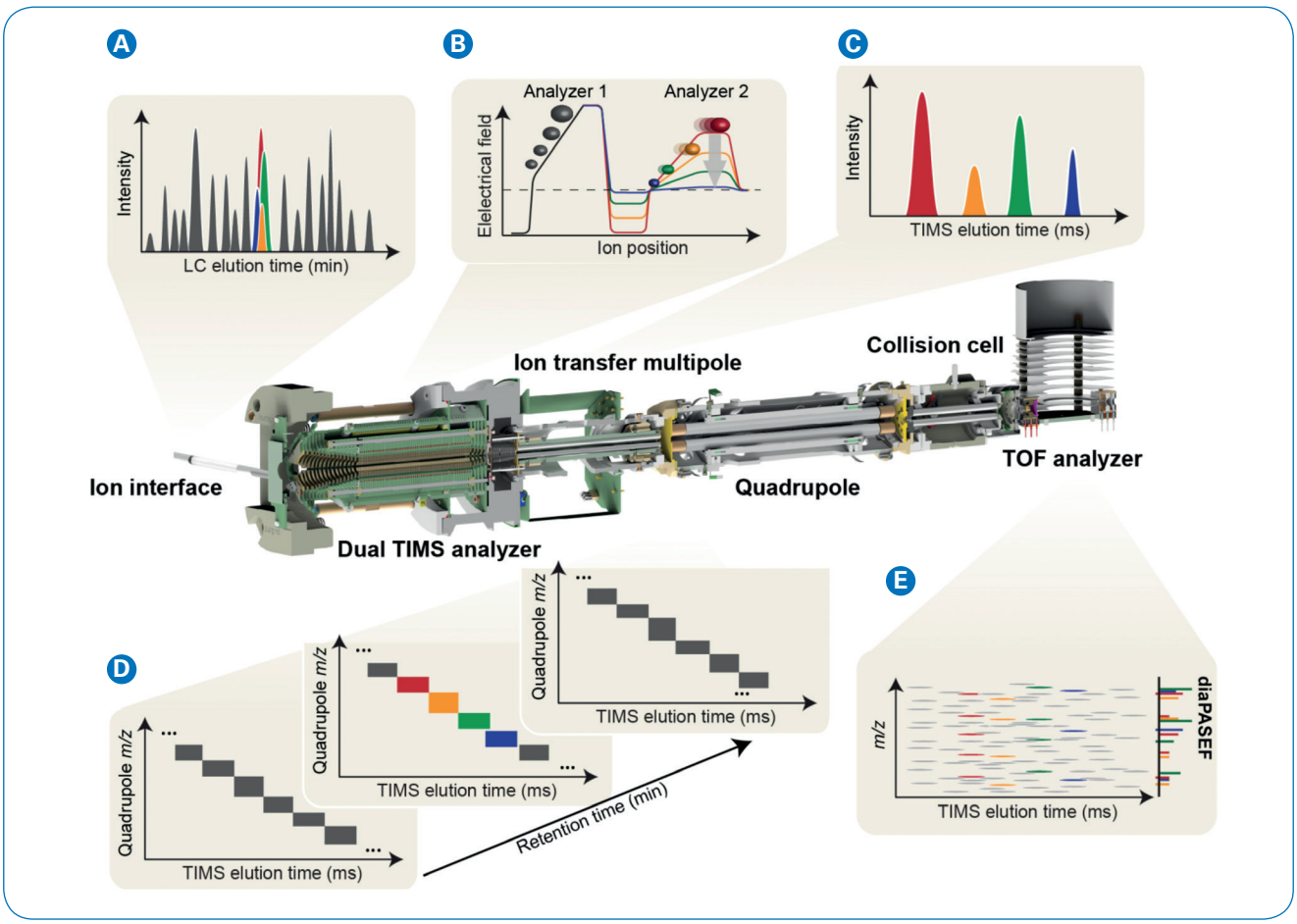

Figure 1: Overview diaPASEF workflow. **A** *Peptides eluting from the chromatographic column are ionized and enter the mass spectrometer through a glass capillary.* **B** *In the dual TIMS analyzer, the first TIMS section traps and stores ion packets, and the second resolves them by ion mobility.* **C** *Ion mobility separated ions are released sequentially from the second TIMS analyzer as function of decreasing electrical field strength.* **D**, **E** *For diaPASEF we couple DIA isolation windows to the precise ion mobility elution of the corresponding ions. Within a single TIMS separation multiple precursor windows are set.*

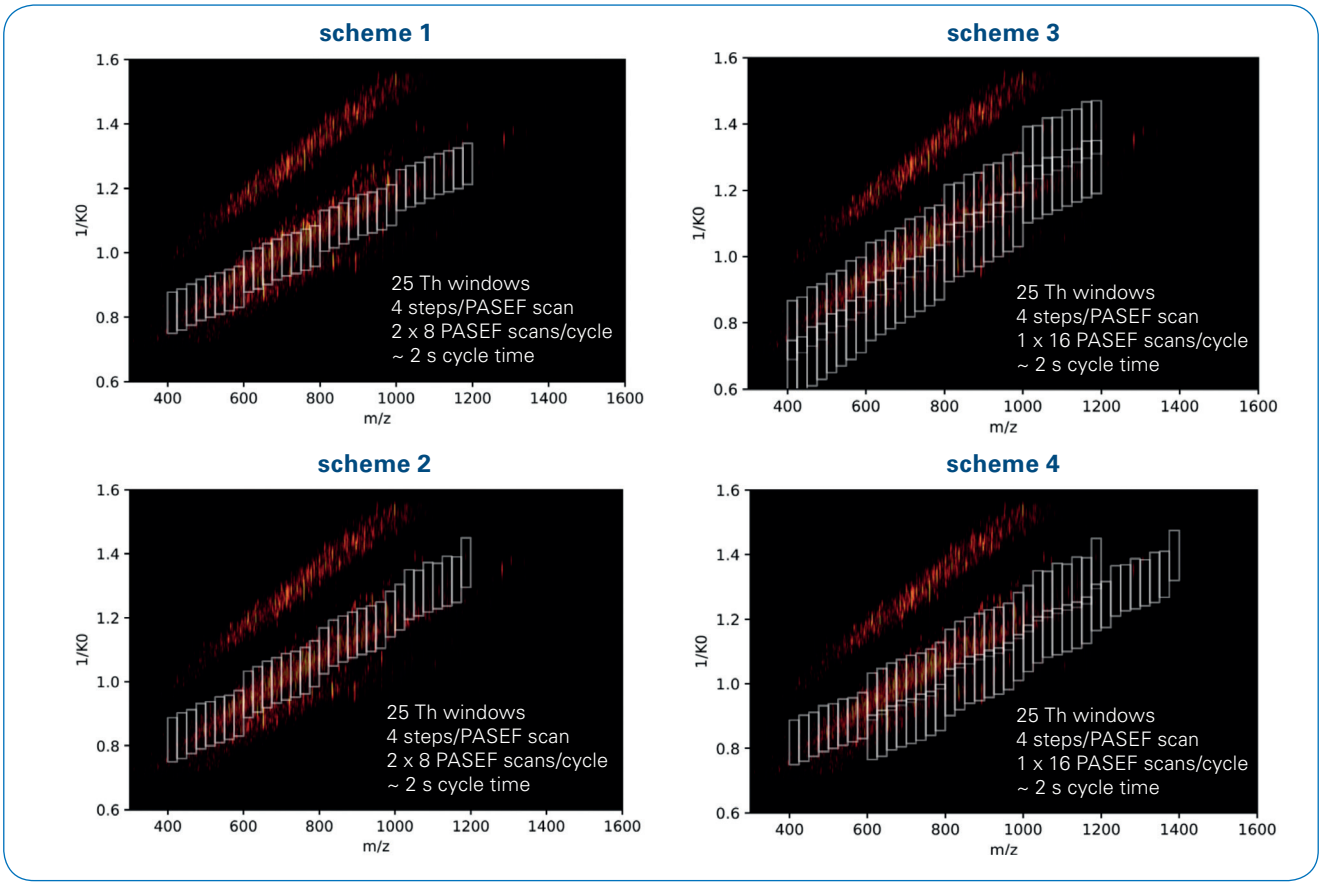

Figure 2: Different isolation schemes have been developed and tested. All presented schemes (1, 2, 3, and 4) use 25 Th windows, 4 steps per PASEF scan and result in a cycle time of ~2 seconds. The schemes 1 and 2 mainly differ from 3 and 4 in the fact that either one row (scheme 1 and 2) of DIA windows is measured twice or two rows covering a broader mobility range are measured once.

largely implemented in QTOF and trapping instruments [4]. We now combined the PASEF principle with DIA resulting in a new acquisition method, called diaPASEF. diaPASEF benefits from the reproducibility of the CCS values which makes library-based approaches very attractive. We explored multiple schemes for the window placement to balance specificity and duty cycle with the aim of increasing the number and reproducibility of identifiable and quantifiable peptides and proteins in complex samples exemplified by whole-cell proteome digests.

Methods

Whole-cell proteomes were extracted from a human cancer cell line (HeLa S3, ATCC, USA) and digested with LysC and trypsin (1:100 enzyme:protein (wt/wt) for both) according to [2]. A nanoflow liquid chromatography system was coupled online to a trapped ion mobility spectrometry – quadrupole time of flight mass spectrometer (timsTOF Pro, Bruker Daltonics). Peptides were separated on a reversed-phase column (50 cm x 75 µm i.d.) with a pulled emitter tip, packed with 1.9 µm C18-coated porous silica beads (Dr. Maisch, Germany) using a linear gradient from 5 to 30% B (80/20/0.1% ACN/water/formic acid) within 95 minutes followed by an increase to 60% B within 5 minutes. We adapted the instrument firmware to perform data-independent isolation of multiple precursor windows within a single TIMS separation (100 ms). This new scan mode is termed diaPASEF. We tested multiple schemes for precursor selection window size and placement in the m/z-ion mobility plane. For a single TIMS scan, we defined up to 32 DIA windows as a function of the TIMS scan time.

Analysis of the four-dimensional data space (intensity, retention time, ion mobility and mass) was performed using Mobi-DIK (Ion Mobility DIA Analysis Kit, http://openswath.org/en/latest/docs/mobi-dik.html). Protein false discovery rate was controlled < 1% FDR with PyProphet. For ion mobility-aware targeted data extraction, we generated a spectral library from our previously published dataset (PRIDE PXD010012, [2]) of 48 high-pH reverse-phase peptide fractions acquired with PASEF technology. MS data were collected over an m/z range of 100 to 1700. During each MS/MS data collection, each TIMS cycle was 1.1 seconds and included 1 MS and 10 PASEF MS/MS scans, with an average of 12 precursors selected for each PASEF MS/

MS scan, resulting in an MS/MS data acquisition rate of 109 Hz. Raw files were processed using MaxQuant (v1.6.5.0) using default timsTOF Pro parameters except for the main search maximum precursor mass tolerance, which was set to 20 ppm and deisotoping of fragment ion spectra which was deactivated.

Results and Discussion

We investigated the potential of diaPASEF for in-depth proteomics of whole human cell lysates. diaPASEF combines the well-known PASEF method, which exploits trapped ion mobility spectrometry to achieve high duty cycle, efficient ion usage and improved peptide identification rates and reproducibility of identification across multiple sample with DIA (Figure 1). By coupling DIA isolation windows to the precise ion mobility elution of the corresponding ions, diaPASEF allows multiplexing of DIA windows in a single 100 ms precursor ion mobility separation.

Mobi-DIK is a novel software based on the OpenMS package [5] capable of analyzing highly multiplexed diaPASEF data. The software generates ion mobility-enabled spectral libraries directly from highly fractionated DDA PASEF runs from the MaxQuant output files (msms.txt, evidence.txt and allpeptides.txt) and stores them in the standardized TraML format [6]. The DDA PASEF library resulting from the analysis of 48 high-pH reverse phase peptide fractions consists of 8,938 protein groups and 107,067 precursors. Different diaPASEF acquisition settings were investigated with the aim of increasing library coverage in a single run. As CCS and mass are correlated, different window placement schemes following diagonal lines in the m/z-CCS space were investigated (Figure 2), with an emphasis on specificity and duty cycle. Depending on the width of the isolation window (in m/z and CCS), the large majority of the multiply charged ion species could be sampled in a single TIMS scan, maintaining a near 100% duty cycle. With a 120 minute LC-MS run, a cycle time of 2 seconds was used to obtain sufficient data points over the chromatographic peak to accurately reconstruct fragment ion chromatograms while maintaining specificity and sensitivity. The high acquisition speed of the timsTOF Pro allowed us to cover the mass range of interest twice (e.g. 1 MS and 2 diaPASEF cycles; scheme 1 and 2) or to increase mobility range coverage by using 2 different diagonal lines (scheme 3 and 4) within the 2 second MS cycle time. In theory scheme 3 and 4 should have the highest coverage of the library, with 97% and 95% theoretical coverage of doubly and triply charged ions, respectively.

Resulting data sets have been processed with Mobi-DIK, which automatically calibrates mass (non-linear), retention time (non-linear)

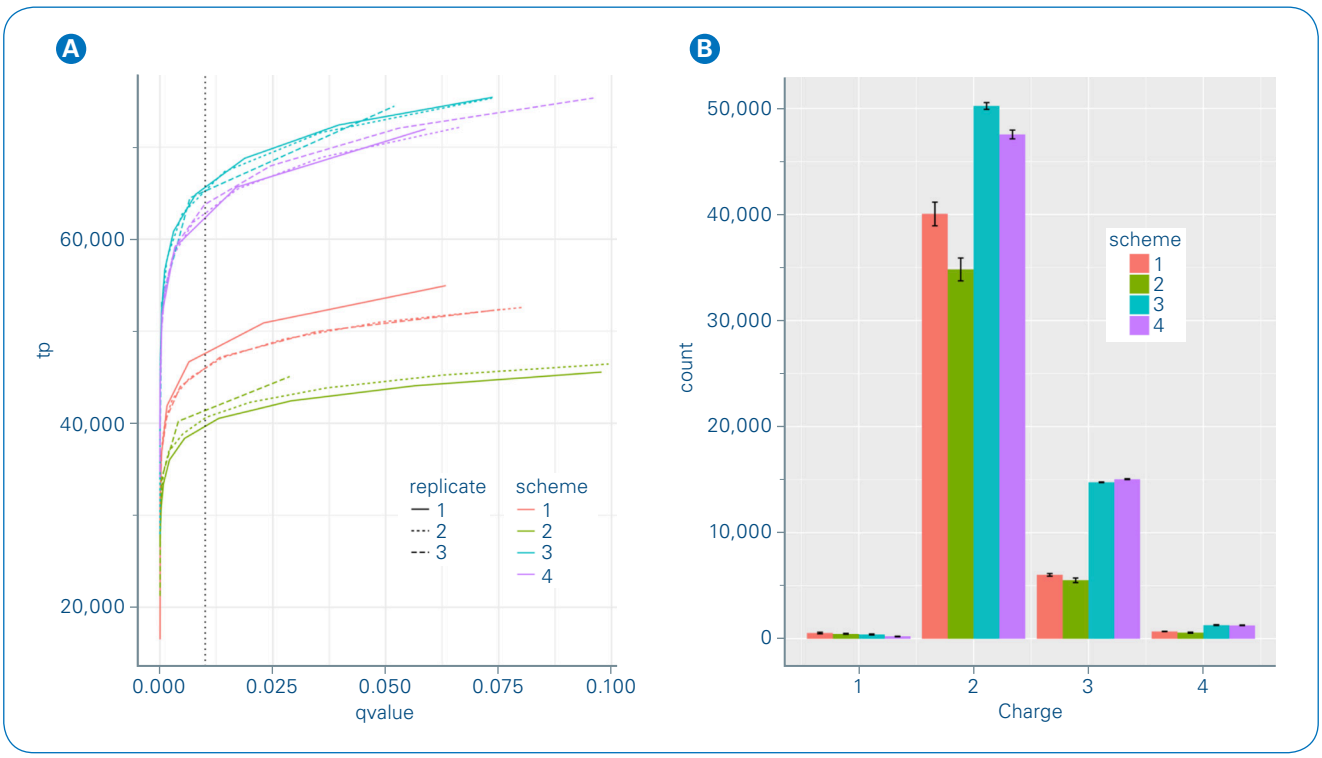

Figure 3: Mobi-DIK results for the different isolation schemes. A Number of identified peptides at different q-values for the technical replicates of the different schemes. B The average charge distribution for the 4 different acquisition schemes shows a clear improvement in the number of identified doubly and triply charged peptides for scheme 3 and 4 compared to 1 and 2.

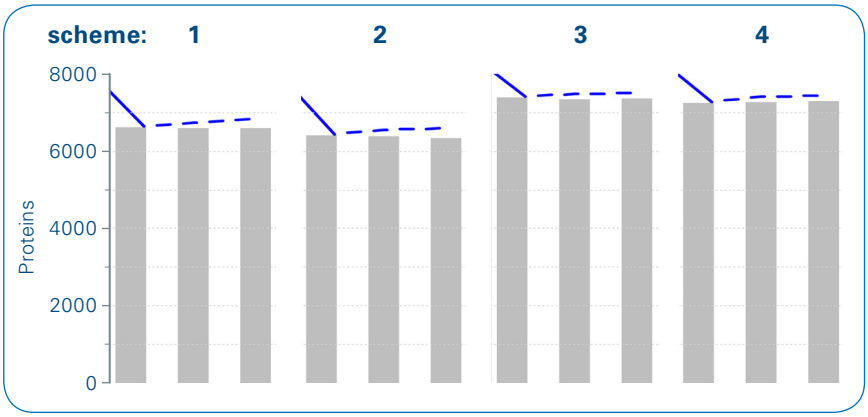

Figure 4: Number of inferred proteins for the different acquisition schemes tested (protein FDR = 1%). Proteins were inferred from proteotypic peptides only. Results of triplicate injections and their cumulative profile are shown.

and mobility (linear) between the generated assay library and experimental diaPASEF runs. After the linear alignment, CCS values extracted from the diaPASEF runs deviated less than 2% from the library and scoring the accuracy of the measured CCS improved the separation of targets and decoys. MS/MS spectra were additionally cleaned up by removing interfering ions from mobility-separated precursors of similar mass that fall into the same quadrupole isolation windows. In the analysis of triplicate 120 minute LC-MS runs of 200 ng HeLa digest, the scheme 3 yielded the highest identification rates, followed by scheme 4 (Figure 3 A). At a q-value of 0.01 scheme 1 and 2 resulted in significantly lower number of detected peptides (< 50,000) compared to scheme 3, where more than 60,000 peptides were quantified. Sampling triply charged ions (Figure 3B) by scanning two diagonal lines across the m/z-CCS space as opposed to just one showed a significant increase with an additional 10,000 peptides detected per run. In addition, the number of doubly charged peptides was also significantly increased with schemes 3 and 4 over 1 and 2. Protein identification rates were very high across triplicate injections for all methods (Figure 4).

In total, with our optimized method we were able to identify on average 7,371 proteins (and 61,503 peptide peak groups) at a 1% false discovery rate in single diaPASEF runs using 200 ng HeLa (Table 1). This equates to ~83% of proteins present in the library in a single-shot diaPASEF run. The remaining 17% that were not detected in diaPASEF were often proteins identified as single peptide hits or by singly charged peptides in the library, which were not covered in our acquisition schemes. The great advantage of DIA approaches in general and (very specific of diaPASEF) lies not only in the very reproducible identification of 1,000s of proteins but also in a high quantitative consistency and accuracy. With diaPASEF quantification at the MS/MS-level was very precise. A median coefficient of variation of 10% and a pairwise mean Pearson correlation > 0.96 for triplicate runs was detected. This very high quantitative reproducibility enables usage of diaPASEF for large-scale quantitative studies.

Table 1: Summary of different isolation schemes tested. Scheme 3 results in highest identification rates on both the peptide and protein level while also giving lowest standard deviation values and highest pairwise mean Pearson correlation

Scheme	1	2	3	4
Peptide peak groups (total)	52,198	44,899	71,143	69,069
Peptide peak groups (median)	42,567	37,055	61,503	59,097
Proteins (total)	6,844	6,607	7,516	7,445
Proteins (median)	6,610	6,382	7,371	7,292
Median protein CV (med. norm.) %	12.6	12.4	10.1	10.3
Mean protein CV (med. norm.) %	20.4	20.3	16.9	17.0
Median protein Pearson R	0.954	0.972	0.973	0.961
Median log2 protein Pearson R	0.949	0.948	0.962	0.961

Conclusions

- diaPASEF captures and utilizes a very large proportion of the ion current.
- More than 7,000 proteins were reproducibly identified and quantified in single runs of 200 ng HeLa digest, taking advantage of the new features of the timsTOF Pro including high ion usage efficiency from the TIMS device, and correlation with CCS values in the library.
- With the excellent identification rates in combination with robust and reproducible quantification diaPASEF helps to go one step further towards the ideal mass analyzer.

Learn More

You are looking for further Information?
Check out the link or scan the QR code for more details.

www.bruker.com/timstofpro

References

[1] Meier F. et al. (2015), J Proteome Res, doi: 10.1021/acs.jproteome.5b00932

[2] Meier F. et al. (2018) Mol Cell Proteomics, doi:10.1074/mcp.TIR118.000900

[3] Aebersold R. & Mann M. (2016) Nature, doi: 10.1038/nature19949

[4] Gillet LC et al. (2016) Annu Rev Anal Chem, doi: 10.1146/annurev-anchem-071015-041535

[5] Röst et al. (2016) Nat Methods, doi: 10.1038/nmeth.3959

[6] Deutsch EW et al.(2012) Mol Cell Proteomics, doi:10.1074/mcp.R111.015040

For Research Use Only. Not for Use in Clinical Diagnostic Procedures.

Bruker Daltonik GmbH

Bremen · Germany
Phone +49 (0)421-2205-0

Bruker Scientific LLC

Billerica, MA · USA
Phone +1 (978) 663-3660

ms.sales.bdal@bruker.com – www.bruker.com