Rapid and sensitive label-free proteome quantification on the timsTOF Pro 2

Shotgun proteomics is commonly used in discovery proteomics for unbiased quantification from a variety of sample sources.

Abstract

Shotgun proteomics capacity to deliver in-depth information on the identification, the relative amount, and the modification patterns of ever larger fraction of the proteins present in a sample has made it an essential approach in most biological research areas. The recent introduction of Parallel Accumulation and SErial Fragmentation (PASEF) technology and the corresponding development of 4D-Proteomics™ approaches has really boosted the potential of these techniques by eliminating the usual need to compromise selectivity or sensitivity for throughput and robustness, by simultaneously improving all these parameters. These features are now enhanced with the introduction of the timsTOF Pro 2 based proteomics solution, which now enables the measurement of over 7000 proteins in a 70 min LC separation while using a standard 4D-Proteomics™ approach.

Keywords: 4D-Proteomics, PASEF, dia-PASEF, data completeness, PasER, timsTOF Pro 2

Authors: Pierre-Olivier Schmit, Gary Kruppa, Thomas Kosinski, Markus Lubeck, Sven Brehmer, Tharan Srikumar and Nagarjuna Nagaraj; Bruker Daltonics GmbH & Co. KG, Bremen, Germany

Introduction

Untargeted proteome-wide quantification is the most commonly used strategy in proteomics. Data dependent and now increasingly data-independent acquisition strategies are used for routine quantification of samples from cell lines, tissues and for affinity-based interaction proteomics. The key attributes for achieving the best and high-throughput proteome quantification include very fast MS/MS data acquisition speed, sensitivity, accuracy, and robustness. While most technological approaches require a compromise of one of these features for another, the introduction of the timsTOF Pro platform improved performance of all these attributes simultaneously and the timsTOF Pro has also proven to be a robust instrument for running hundreds to thousands of samples without service intervention. PASEF on the timsTOF Pro platform has two key advantages, namely space and time focusing of the ions, thus boosting the sensitivity and efficient precursor scheduling for fragmentation resulting in MS/MS spectral acquisition rates >100 Hz [1,2]. On top of these benefits, trapped ion mobility spectroscopy (TIMS) measurements provide collisional cross-section (CCS) values and separation of isomeric species that are mobility offset but mass aligned (MOMA) and alleviates ratio compression in multiplexed quantification approaches. The introduction of these 4D-Proteomics capabilities has bridged the gap between the requirements of the most demanding proteomics approaches (e.g clinical research proteomics, companion diagnostics research and personalized medicine research) and the solutions effectively available on the market. The new timsTOF Pro 2 platform enhances the current benefits with a modified TIMS cartridge, along with further optimized ion optics and RF fields that effectively increase the robustness, capacity, and sensitivity of the front end. The result is a system that is capable of delivering in-depth quantitative proteomics information in a shorter time, thus increasing the number of samples that can be included in a cohort with no need for potentially biased and expensive labelling solutions. Here we report on the label free proteomics performance of timsTOF Pro 2 on a cell line – a common sample type routinely studied in any proteomics or cell biology lab.

Methods

Human embryonic kidney cell lines were lysed and digested in a buffer containing sodium deoxycholate or acetonitrile derived from published protocols described elsewhere [3,4]. Peptides were separated over a

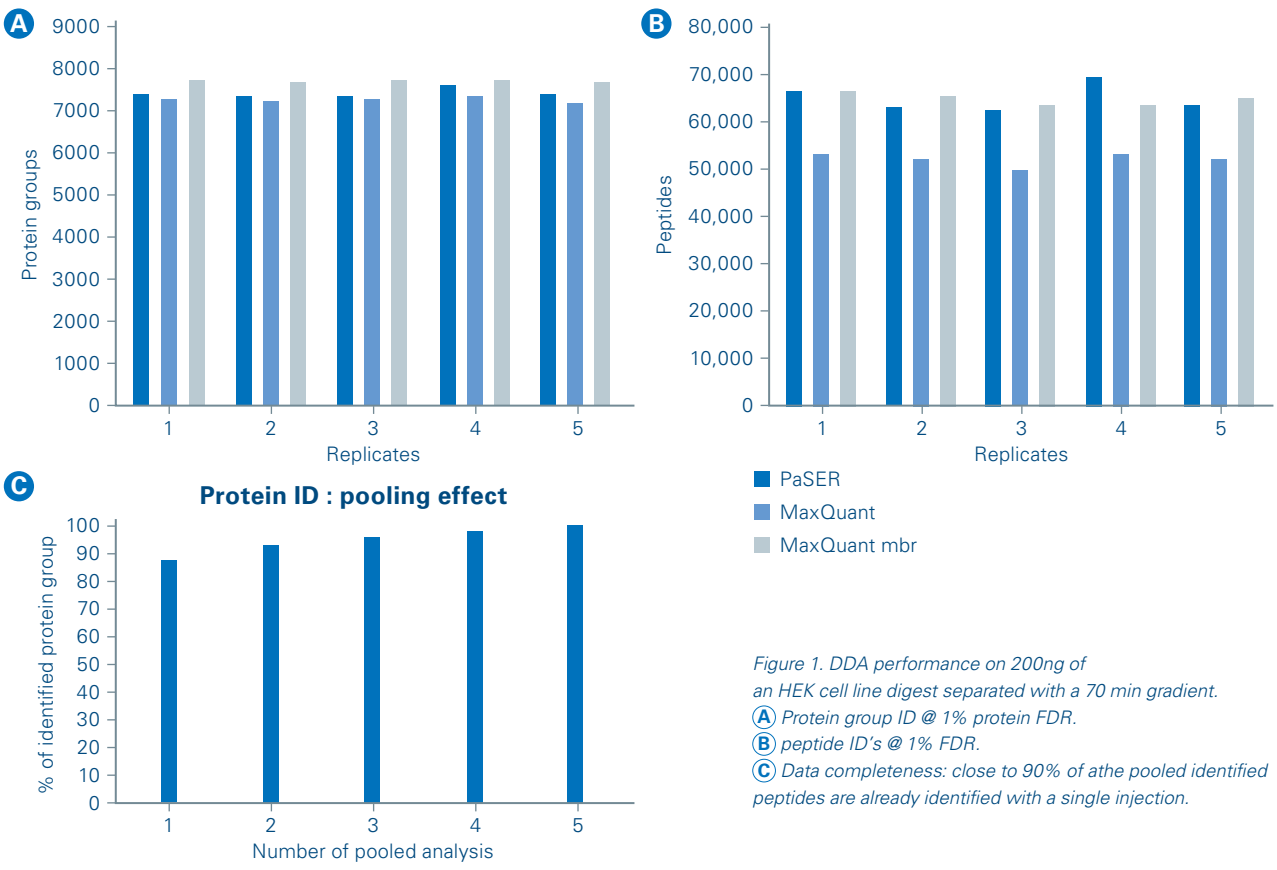

Figure 1. DDA performance on 200ng of an HEK cell line digest separated with a 70 min gradient.
(A) Protein group ID @ 1% protein FDR.
(B) peptide ID's @ 1% FDR.
(C) Data completeness: close to 90% of athe pooled identified peptides are already identified with a single injection.

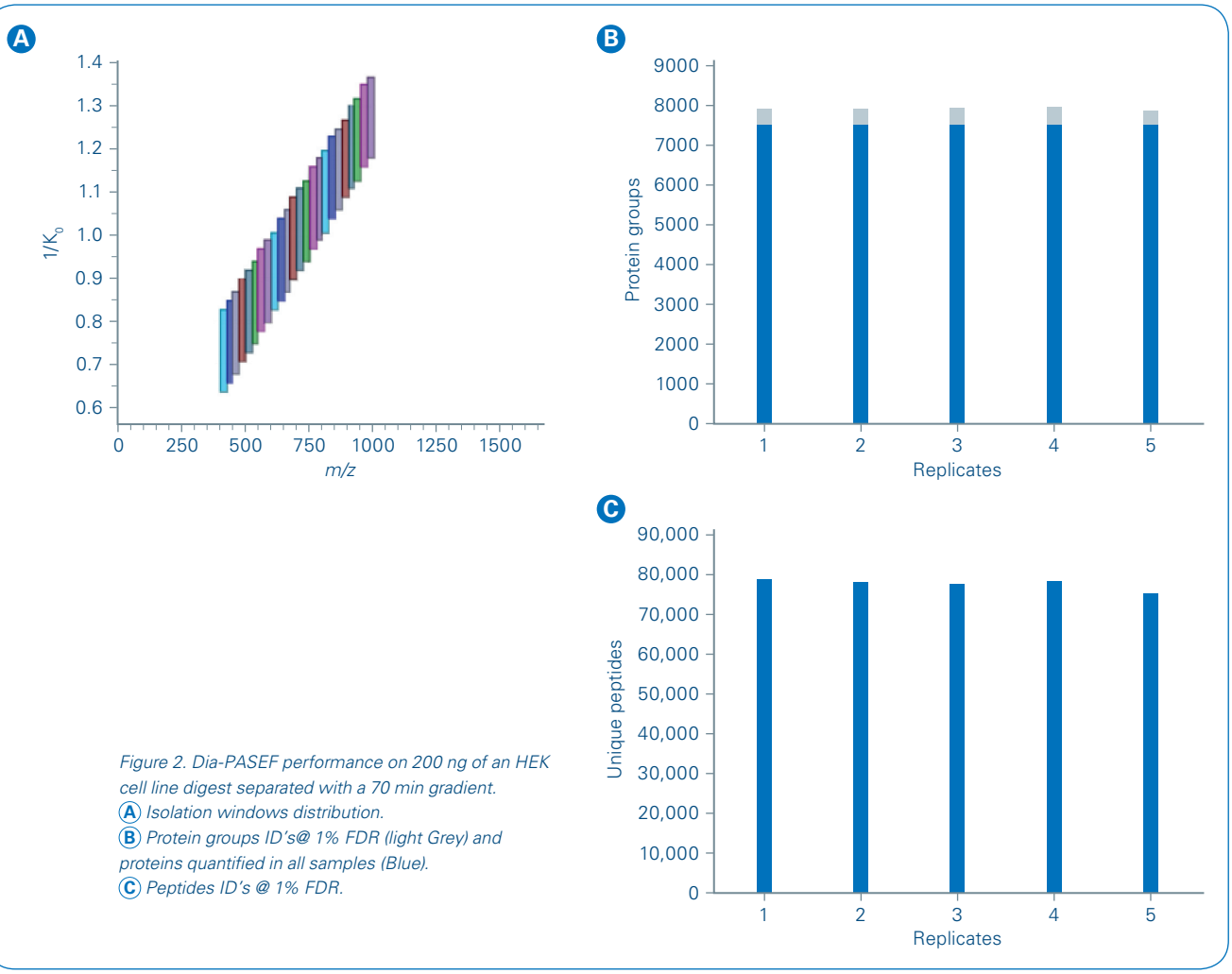

Figure 2. Dia-PASEF performance on 200 ng of an HEK cell line digest separated with a 70 min gradient.
(A) Isolation windows distribution.
(B) Protein groups ID's@ 1% FDR (light Grey) and proteins quantified in all samples (Blue).
(C) Peptides ID's @ 1% FDR.

70 minute gradient (80 minute runtime) on an Aurora-25 cm column (IonOpticks Australia) using a nanoElute UHPLC coupled to a timsTOF Pro 2 mass spectrometer via a CaptiveSpray ionization source. The gradient was delivered at a flow rate of 250 nL/min and washout was performed at 300 nL/min. For DDA, the full scans were recorded from 100 to 1700 *m/z* spanning from 1.35 to 0.85 Vs/cm^2 in the mobility ($1/K_0$) dimension. Up to 10 PASEF MS/MS frames were performed on ion-mobility separated precursors, excluding singly charged ions which are fully segregated in the mobility dimension, with a threshold and target intensity of 1750 and 14,500 counts, respectively. For 20 ng peptide loads, the same method as above was used except that the accumulation time was set to 166 ms and the threshold was lowered to 1000 counts. The 20 ng peptides were measured with a 35 minute gradient (47 minute runtime) with a constant flow rate of 250 nL/min and the high-sensitivity detection was enabled in the timsControl acquisition software. PaSER was utilized for real-time database searches during acquisition [5]. MS1 precursor tolerance of 20 ppm, MS2 tolerance of 20 ppm with fully-tryptic digestion allowing 2 missed cleavages, fixed modification of cystine [+57.02] and oxidation of methionine and deamination of asparagine and glutamine were specified. All raw data were also processed using MaxQuant computational platform version 1.6.17.0 with default parameters for LFQ [6] and match between runs.

For dia-PASEF analysis a window placement scheme consisting of 24 isolations and cycle time, including MS1 frame, of 0.9 seconds was used (illustrated in Figure 2C) [7]. All dia-PASEF data were processed in Spectronaut version 14 using default parameters. A hybrid library was constructed using the DIA data together with a search archive containing a 24-fraction HeLa library.

Results and Discussion

In the proteomics field it is generally recognized that the deepest coverage of the proteome can be achieved by first fractionating the peptides using an orthogonal separation technique such as high pH or anion exchange chromatography. However, shotgun proteomics workflows employing

fractionation typically require 12-24 or more hours to complete. Without fractionation, single shot proteomic experiments offer deep coverage of proteomes in a relatively short measurement time, typically 60 minutes to a few hours. In recent years, the introduction of the timsTOF Pro has resulted in a paradigm shift toward lower peptide loading amounts and shorter gradient lengths for single shot proteomics experiments. For cellular proteomics experiments on the timTOF Pro, 200 ng of peptides are typically loaded and measured in a 90 minute gradient, yielding very good proteome coverage in a relatively fast experiment. With the timsTOF Pro 2 platform we introduce more robust ion optics together with further optimized acquisition methods. Here we report our recent observations from using the new methods on the timsTOF Pro2 platform. In addition to optimizing the methods the gradient length was also reduced from 90 to 70 minutes to further exploit the sheer sequencing (MS/MS) speed of the instrument.

From 200 ng of peptides analyzed using a 70 minute gradient we could identify more than 7000 protein groups in single injection from both ProluCID (PaSER) and Andromeda (MaxQuant) search engines (Figure 1). These protein groups were covered by more than 60,000 unique peptide

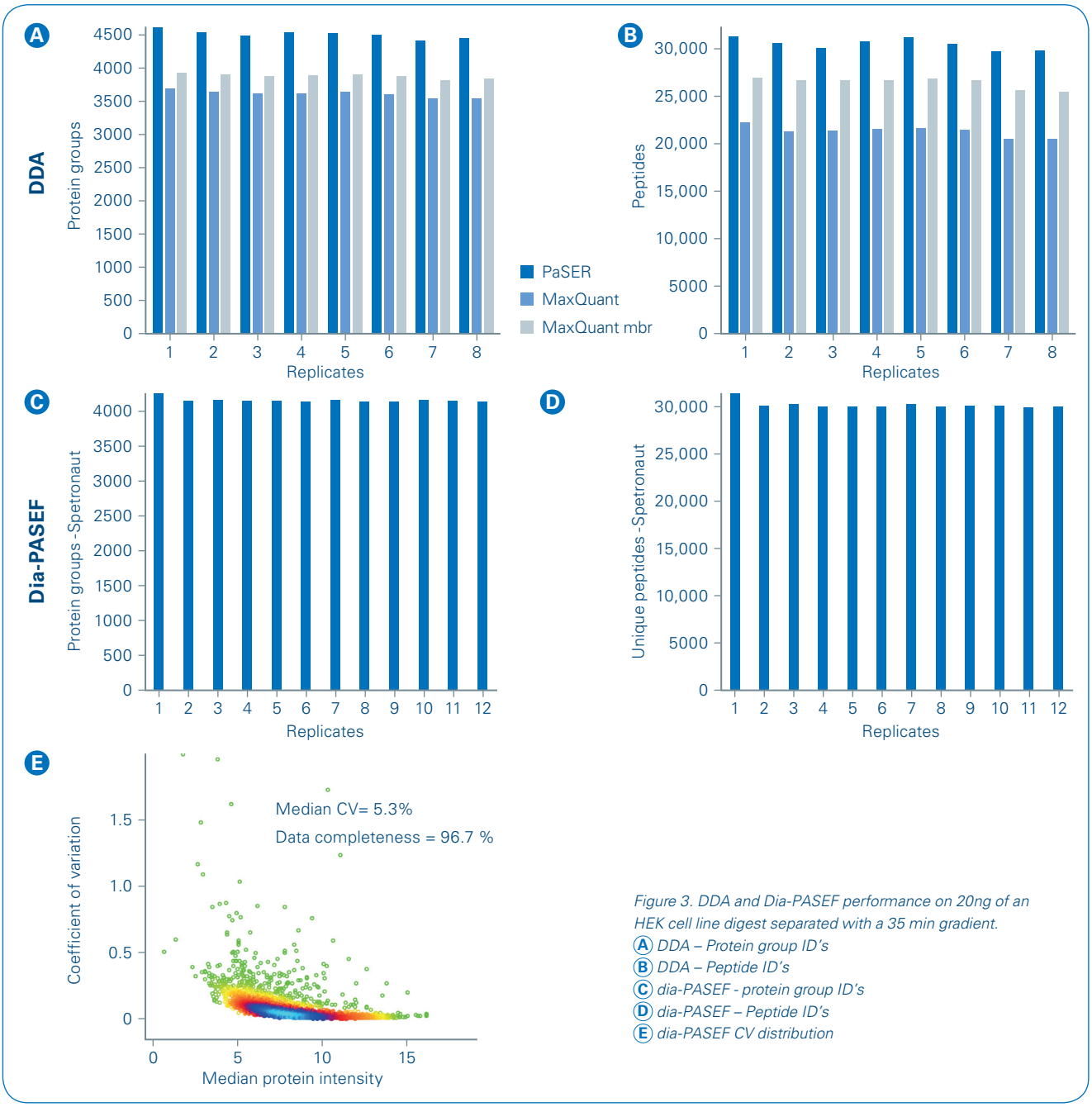

Figure 3. DDA and Dia-PASEF performance on 20ng of an HEK cell line digest separated with a 35 min gradient.
(A) DDA – Protein group ID's
(B) DDA – Peptide ID's
(C) dia-PASEF - protein group ID's
(D) dia-PASEF – Peptide ID's
(E) dia-PASEF CV distribution

sequences reported by PaSER while MaxQuant reported consistently above 50,000 unique peptide sequences. This unprecedented proteome coverage is further improved by enabling the match between runs feature in MaxQuant (without any external spectral library) providing close to 7700 protein groups and more than 64,800 unique peptide sequences. Interestingly in addition to the deep proteome coverage from single shot analyses, the five injections in total led to 8147 protein groups with intensity values in MaxQuant with a data completeness of 87.6 % (7142 protein groups). The median CV for all protein groups quantified was 0.16 and Pearson correlations among the replicates were on average above 0.97.

The same samples were also analyzed using dia-PASEF mode on the timsTOF Pro 2 platform, using a window scheme consisting of 24 isolation windows with a width of 25 Da each-(Figure 2A). From these single shot injections, using a library containing more than 11,000 protein groups, we could identify more than 7800 protein groups. Among all the injections 7493 protein groups were reproducibly quantified corresponding to a data completeness of 91.9% with an average of 7892 protein groups and 77,283 peptides identified per replicate (Figure 2B and 2C). The median CV for protein quantification in this dataset was 11%.
It should be noted that the results presented here for the 70-minutes gradient are for an in-house extraction and digest of proteins isolated from HEK cell lines. At least some of the gain in performance results from the optimizations we have performed on the sample prep (manuscript in progress). However, even for commercially available standard digests of proteins from K562 or HeLa cell lines, the new timsTOF Pro 2 typically achieves in the range of 6000 protein groups, and nearly 60,000 unique peptides in a 60 minute gradient.

Lastly, we analyzed 20 ng of commercially available HeLa cell digest in PASEF mode on the timsTOF Pro 2 platform using a 35 minute LC gradient separation. 4500 protein groups, represented by nearly 30,000 peptides, were identified by ProluCID (PaSER). The raw data processed with MaxQuant identified more than 3500 protein groups, supported by more than 25,000 peptides sequences on average and 3877 protein groups and 26,450 peptide sequences across all 8 injections using the match between runs feature (Figure 3). Even at lower samples loads, the timsTOF Pro 2 platform showed great sensitivity and reproducibility, as evidenced by

greater than 83.4% (3387 protein groups) identified protein groups across all 8 injections. In addition, using dia-PASEF resulted in average 4156 protein groups and more than 30,100 peptides sequences with a data completion of 96 % and an excellent median CV of 5.3%.

Conclusion

The timsTOF Pro ushered in the age of 4D-Proteomics and a trend to high-throughput larger cohort studies without sacrificing sensitivity, accuracy, or robustness. The timsTOF Pro 2 extends this legacy further by capitalizing on the improved TIMS cartridge and frontend ion optics to deliver >70,000 unique peptides from >7700 protein groups from 70 minutes of LC separation. The timsTOF Pro 2 not only delivers high identification rates, but high quantification rates with low variability (i.e. low missing values with exceptional accuracy and sensitivity) in both DDA and DIA modes. These results show coverage of a substantial portion of the number of proteins generally expected to be found in a single cell in the range of 50-80%, in a single shot experiment, and in the range of 70-80% of what can be found with time consuming fractionation approaches. The continued improvement in speed and depth of coverage on the timsTOF Pro 2 enables dramatic improvements in performance for research involving the comparison of cell cultures under different conditions, an important workflow in biological and pharmaceutical research. The depth of coverage achievable with sample amounts down to 20 ng means that limited sample amounts can be used for tissue proteomics, workflows using enrichment for PTM studies, and immunopeptidomics, all important for research in biomarker discovery and personalized medicine research.

Learn More

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

www.bruker.com/massspectrometry

References

[1] Meier F et al. (2015). J. Proteome Res.

[2] Meier F et al. (2018). Molecular & Cellular Proteomics.

[3] Strader MB et al. (2006). Anal. Chem.

[4] Lin Y et al. (2008). Analytical Biochemistry.

[5] Xu T et al. (2015). Journal of Proteomics.

[6] Prianichnikov N et al. (2020). Molecular & Cellular Proteomics.

[7] Meier F et al. (2020). Nat Methods.

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

Bruker Daltonics GmbH & Co. KG

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