PEAKS™ Software for the processing of timsTOF Pro PASEF data: identification and label-free quantitation

Bioinformatics Solutions Inc.

BRUKER

ASMS 2018, TP367

Markus Lubeck(1), Heiner Koch(1), Scarlet Beck(1), Paul Savage(2), Oliver Raether(1),Pierre-Olivier Schmit(2) Paul Shan (3) and Gary Kruppa (4)
(1) Bruker Daltoniks GmbH, Bremen-Germany, (2) Bruker France SA, Wissembourg, France, (3) Bioinformatics Solutions, Waterloo, Canada, (4) Bruker Daltonics Inc, Billerica, USA

Introduction

The "Parallel Accumulation Serial Fragmentation" (PASEF) method for trapped ion mobility spectrometry (TIMS) quadrupole time of flight (QTOF) instruments described previously enables faster acquisition of fragment ion spectra from isolated precursor ions improving sensitivity without sacrificing spectral quality (Figure 1). The outcome of a PASEF acquisition is a collection of high resolution (50000@ 1222 m/z), high sensitivity MS/MS spectra acquired at very high speed (> 100Hz MS/MS) in a Data Dependent Analysis (DDA) after a TIMS separation event. The combination of the volume of data that is generated and the addition of an extra separation dimension can constitute a challenge for the software. Here we describe how a modified version of the well-known PEAKS™ proteomics processing software suite can be used for the processing of PASEF datasets.

Methods

The performance of a timsTOF Pro instrument with PASEF for shotgun proteomics (Bruker Daltonics) has been evaluated using tryptic digests of human cancer cell lysates (HeLa) (Pierce) and mouse heart cell digests. Chromatographic separations were performed on a nanoElute (Bruker) nano-UHPLC using a 25 cm, 75 µm ID Odyssey C18 nano column with integrated emitter (IonOpticks, Australia) at 400 nl/min and column oven at 50 °C with direct loading. Data was analyzed using multiple interations of the PEAKS (www.bioinfor.com) and Perseus (Jürgen Cox, Max Planck Institute of Biochemistry) software. Results were normalized to <1% FDR on PSMs.

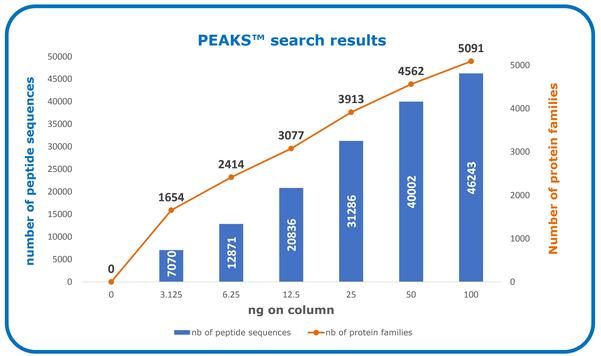


Fig1: **PEAKS™ protein identification from PASEF acquisitions**
Increased amount of Hela cell digest, separated with a 60 min gradient (3.1-100ng). Searches have ben performed using the standard PEAKS ID workflow (DeNovo + Database searches) maintaining a 1% PSM FDR, 2 miscleavages and Human taxonomy.

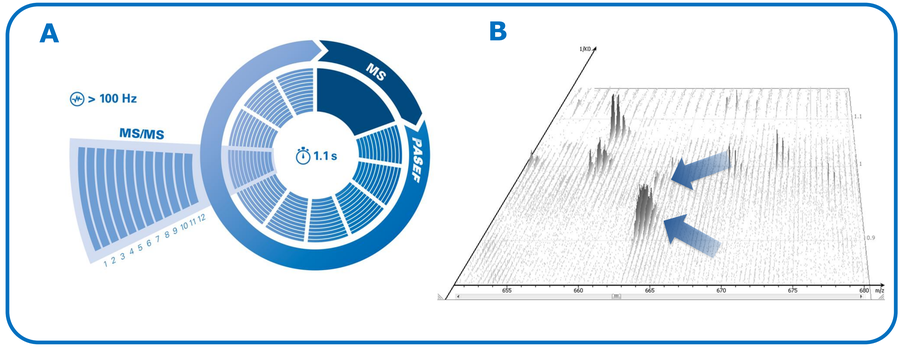


Fig1: **Overview of a PASEF acquisition cycle and PEAKS™ 3D display**
A – The PASEF acquisition method: A PASEF cycle (total: 1.1s) consists of 1 TIMS MS scan (100 ms, dark blue) and 10 PASEF MS/MS scans (each 100 ms, light blue). The signal recorded during the TIMS MS scan is used for precursor selection, while taking into account dynamically excluded ions. Per PASEF scan on average 12 different precursors are selected for MS/MS resulting in a sequencing speed of > 120 Hz.
B – PEAKS zoomed display on one of the TIMS MS scan acquired during the PASEF analysis of a mouse heart digest. This heat map displays m/z vs mobility vs intensity. The near isobaric ions highlighted by the blue arrows and separated in the ion mobility dimension are the precursors for the MS/MS displayed in Fig.3

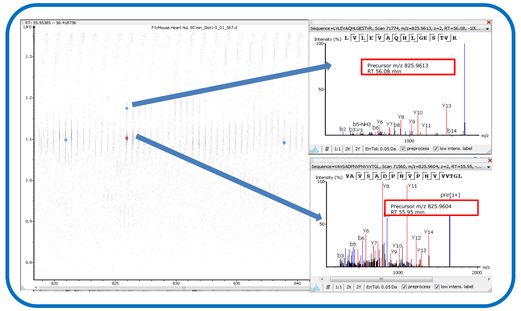


Fig3: **PEAKS™ display of PASEF ion-mobility separation** Data from mouse heart digest acquired at a 90 min gradient with timsTOF Pro and PASEF enabled. The two co-eluting parent ions (0.009 Da apart) have been separated in the ion mobility dimension prior to quadrupole isolation and fragmentation, revealing two distinct peptides which are successfully identified.

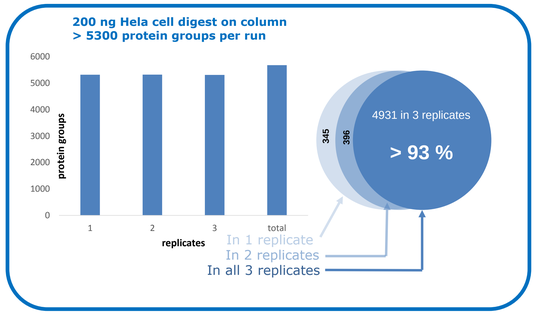
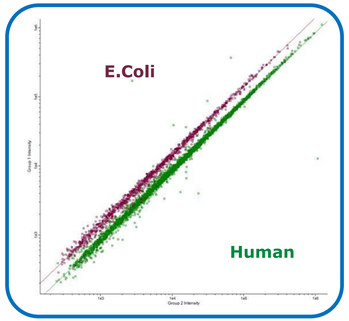


Fig4: **Assaying PASEF identification reproducibility**
Triplicate injection of a Hela cell digest separated with a 90 min gradient and analyzed in PASEF mode before being processed with PEAKS™. >93% of the identified protein groups are commonly identified in the three samples. PASEF enables reproducible identification of more than 90% of detected proteins, substantially better than other data dependent approaches.

Results & Discussion

A Peaks database search, performed on data generated with a standard 1.1 second PASEF acquisition cycle, which performs over 120 MS/MS at high sensitivity, identifies over 1654 protein groups from a 3,125 ng injection of a HeLa cell digest separated with a 90 min gradient, and over 5000 protein groups from a 100 ng injection with the same gradient (Figure 2). In parallel, the ion mobility separation enables the separation of isobaric co-eluting peptides prior to fragmentation, further increasing the ID rate (Figure 3). The increased MS/MS sequencing speed reduces a typical drawback of Data Dependent Analysis, the stochastic element of precursor selections, as the parent ion selection becomes highly reproducible (Figure 4). Consequently, a very large majority of identified proteins can be quantified, while maintaining selectivity (as a combination of ion mobility separation and parent ion isolation) which has advantages over Data Independent Acquisition approaches. The combination of those features allows for good Label-Free Quantitation (LFQ) performance on a mixture of 2 proteomes at a 0.5 ratio, as illustrated in Figure 5.

Fig5: **Label-free quantitation from PASEF acquisition**
Enhanced sensitivity, speed and reproducibility in ID rates are pre-requisites for enhanced Label-Free Quantitation (LFQ) performance. Hela and E.Coli proteome digests have been mixed in a ½ ratio prior to separation with a 90 min gradient, PASEF analysis and PEAKS™ data processing. The correlation diagram illustrates the capacity of the system to deliver an accurate and reproducible relative quantitation information despite the challenge of a 0.5 proteome mixed ratio.



Conclusions

- The adapted version of PEAKS™ enables identifications and label-free quantification from PASEF runs
- Fast, sensitive label-free quantitation is delivered with enhanced selectivity
- TIMS PASEF is therefore very well suited for large cohort (e.g. clinical proteomics), low abundancies (e.g biopsies) and high complexity (e.g PTM screening) shotgun proteomics studies

timsTOF Pro