bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Parallel accumulation – serial fragmentation combined with data-independent acquisition (diaPASEF): Bottom-up proteomics with near optimal ion usage

Florian Meier[1], Andreas-David Brunner[1], Max Frank[2], Annie Ha[2], Eugenia Voytik[1], Stephanie Kaspar-Schoenefeld[3], Markus Lubeck[3], Oliver Raether[3], Ruedi Aebersold[4,5], Ben C. Collins[4*], Hannes L. Röst[2*] and Matthias Mann[1,6*]

[1] *Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Martinsried, Germany*
[2] *Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Canada*
[3] *Bruker Daltonik GmbH, Bremen, Germany*
[4] *Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland*
[5] *Faculty of Science, University of Zurich, Zurich, Switzerland*
[6] *NNF Center for Protein Research, University of Copenhagen, Copenhagen, Denmark*

*To whom correspondence may be addressed: collins@imsb.biol.ethz.ch, hannes.rost@utoronto.ca or mmann@biochem.mpg.de

ABSTRACT

Bottom-up proteomics produces complex peptide populations that are identified and quantified at the precursor or fragment ion level. Data dependent acquisition methods sequentially isolate and fragment particular precursors, whereas data independent acquisition (DIA) modes isolate and concurrently fragment populations of different precursors by cycling deterministically through segments of a predefined precursor *m/z* range. Although the selection windows of DIA collectively cover the entire mass range of interest, only a few percent of the ion current are sampled due to the consecutive selection of acquisition windows. Making use of the correlation of molecular weight and ion mobility in a trapped ion mobility device (timsTOF Pro), we here devise a novel scan mode that samples up to 100% of the peptide precursor ion current. We analyze the acquired data by extending established targeted data extraction workflow for the analysis of DIA data by the additional ion mobility dimension, providing additional specificity in the precursor identification. Data acquired from simple protein mixtures verify the expected data completeness and data in single runs of a whole proteome digest demonstrate deep proteome coverage and a very high degree of reproducibility and quantitative accuracy, even from 10 ng sample amounts.

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

INTRODUCTION

Mass spectrometry (MS)-based proteomics, like other omics technologies, aims for an unbiased, comprehensive and quantitative description of the system under investigation[1–3]. Although proteomics workflows have become increasingly successful in characterizing complex proteomes in great depth[4,5], application of this technology to large sample cohorts e.g. for systematic screening or clinical applications also requires a high degree of reproducibility and data completeness[6]. In recent years, data independent acquisition (DIA) schemes have become increasingly widespread[7]. Unlike in data dependent acquisition where particular precursors are sequentially selected, in DIA the ions in extended mass ranges are recursively isolated by the quadrupole and concurrently fragmented, thus generating convoluted fragment ion spectra composed of fragments from different precursors[8–10]. Although DIA guarantees that each precursor mass in a predefined mass range is fragmented once per cycle, spectral complexity poses a great challenge to subsequent analysis[11]. Narrow isolation windows reduce spectral complexity, but also increase the cycle times needed to cover the entire mass range. Moreover, as every precursor is only sampled once per cycle, the ion sampling efficiency at the mass selective quadrupole for DIA methods is typically limited to 1-3% with typical schemes of 32 or 64 windows.

Adding ion mobility separation to the chromatographic and mass separation is expected to increase sensitivity and reduce spectral complexity[12–15]. The trapped ion mobility spectrometer (TIMS) is a particularly compact mobility analyzer in which ions are captured in an RF ion tunnel by the opposing forces of the gas flow from the source and opposing counteracting electric field[16–18]. Trapped ions are then sequentially released as a function of their collisional cross section by lowering the electric potential. Resolution depends on the ramp time, which is typically 50 to 100 ms, just between chromatographic peak widths (seconds) and the time of flight spectral acquisition (about 100 µs per scan). In a TIMS-quadrupole-TOF configuration, the release of precursor ions can be synchronized with the quadrupole selection in a method termed parallel accumulation followed by serial fragmentation[19]. PASEF achieves a ten-fold increase in sequencing speed in data dependent acquisition, without the loss of sensitivity that is otherwise inherent to very fast fragmentation cycles[20].

Here we investigate if the PASEF principle can be extended to data independent acquisition, combining the advantages of DIA with the inherent efficiency of PASEF. To realize this vision, we modified the mass spectrometer to support diaPASEF acquisition cycles and build upon open software[21] to perform targeted extraction of fragment ion traces from the four dimensional data space to form peak groups that confirm the identity and indicate the abundance of a specific peptide in the sample. We explore the diaPASEF principle in typical proteomics applications such single run proteome analysis, as well as the in-depth and extremely sensitive characterization of microscopic sample amounts.

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

RESULTS

The diaPASEF principle

With the timsTOF Pro instrument (Bruker Daltonics) peptides separated by liquid chromatography are ionized and introduced into the mass spectrometer and immediately trapped in a first TIMS device (**Fig. 1a**). They are then transferred into TIMS2 from which they are released in reversed order of their ion mobilities, i.e. the largest ions are released first. In parallel, incoming ions are again accumulated in TIMS1, assuring full ion utilization. If operated in MS1 mode, the ion species sequentially mobilized from TIMS2 reach the orthogonal accelerator from which rapid TOF scans result in high resolution mass spectra (> 35,000 over the entire mass range). If operated in MS/MS mode, selected precursors are fragmented in a quadrupole collision cell and the resulting fragment ions analyzed by rapid TOF scans, resulting in fragment ion spectra at the same mass resolution. For peptide ions of a given charge state, ion mobilities and precursor masses are correlated (**Fig. 1b**). We therefore reasoned that this feature can be used to isolate precursor mass windows for DIA in a way that ions that are not selected at any given time are not lost, as is the case in other DIA acquisition schemes. Because high *m/z* (low ion mobility) ions are stored at the end of the

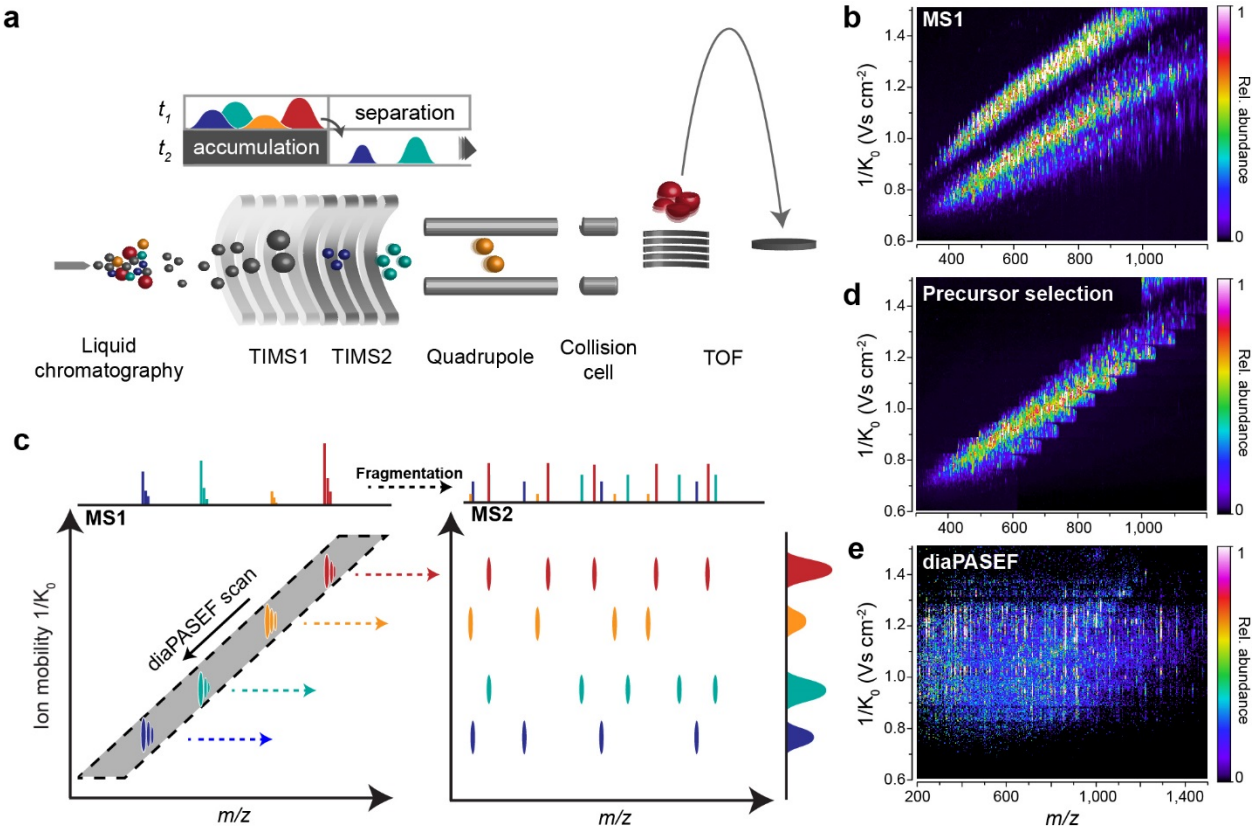

Figure 1 | The diaPASEF acquisition method. a, Schematic ion path of the timsTOF Pro mass spectrometer. **b**, Correlation of peptide ion mobility and mass-to-charge (*m/z*) in a tryptic digest of HeLa cell lysate. **c**, In diaPASEF, the quadrupole isolation window (grey) is dynamically positioned as a function of ion mobility (arrow). In a single TIMS scan, ions from the selected mass ranges are fragmented to record ion mobility-resolved MS2 spectra of all precursors. **d**, Implementation of diaPASEF precursor selection with a stepped quadrupole isolation scheme. **e**, Representative example of a single diaPASEF scan with the precursor selection scheme from d.

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

voltage gradient, they are released first in the TIMS ramp and the quadrupole collision cell therefore needs to be first positioned at high *m/z*. Coordinately with the decreasing *m/z* of the ions mobilized from the TIMS, the quadrupole mass isolation window slides down to lower *m/z* values such that the trapped ion cloud is fully transmitted (grey area in **Fig. 1c**). To approximate this ideal diaPASEF scan, we stepped the isolation window as a function of TIMS release time, covering the vast majority of precursors of the 2^+ and 3^+ charge state, respectively (**Fig. 1d**). Implementation of this principle on the existing hardware required novel firmware to synchronize collision energies with the mass selection in the ramp (**Methods**). Subsequent fragmentation in the collision cell distributes the fragments of each DIA window in the *m/z* dimension at the exact ion mobility position of the precursor (**Fig. 1e**). Over the chromatographic elution of a precursor, the intensities of its fragments follow the precursor intensity in time (z-direction). The signal traced out by the set of fragments of an individual precursor is a set of very flat ellipsoids (x or *m/z* dimension), spreading in ion mobility direction (y-direction) and elongated in the retention time dimension (z-dimension). For the entire liquid chromatography tandem mass spectrometry (LC-MS/MS) run, this leads to a 'perfect data cuboid' in four dimensional space, containing all fragment ions of all precursors over the entire elution time, with signal intensity as the fourth dimension.

Quantifying the increase in data acquisition efficiency

To explore the diaPASEF principle in practice, we measured a tryptic digest of bovine serum albumin (BSA) and compared the signals obtained across DDA, DIA and diaPASEF acquisition methods. As a typical example, the peptide DLGEEHFK eluted over 9 s (**Fig. 2a**). In DDA, the doubly charged precursor was accumulated for fragmentation once for 100 ms at the beginning of the elution peak, corresponding

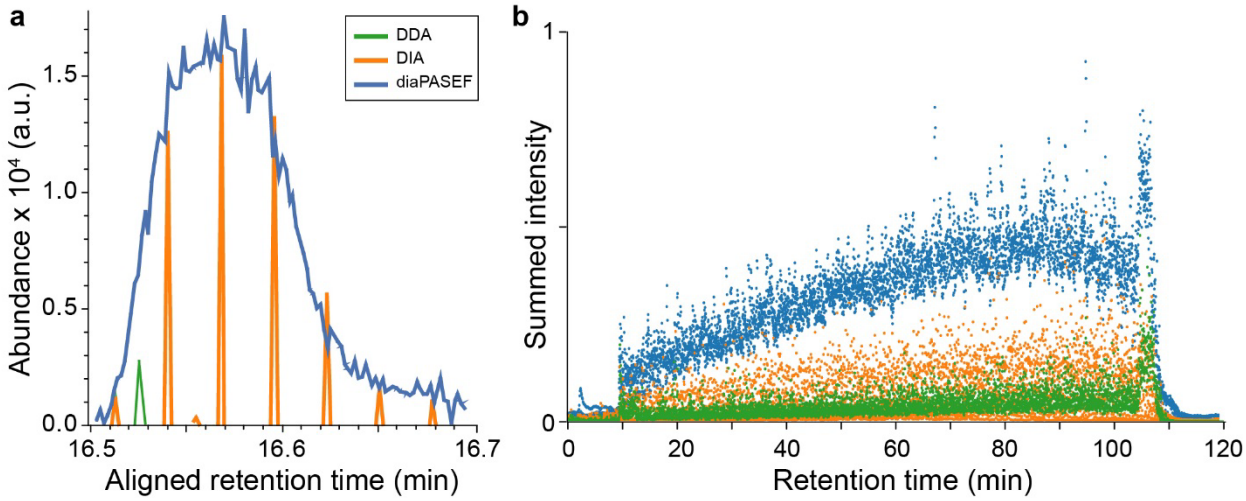

Figure 2 | Efficiency of different data acquisition methods. a, Extracted fragment ion chromatogram of the y_6 ion of the doubly charged DLGEEHFK peptide precursor in an LC-MS analysis of bovine serum album digest acquired with DDA PASEF, DIA and a 100% duty cycle diaPASEF method. **b,** Summed extracted ion chromatograms of multiply charged precursor isolated by the quadrupole in single run analyses of a HeLa digest acquired with DDA PASEF, DIA and a 25% duty cycle diaPASEF method.

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

to about 1% of the total elution time and << 1% of the entire precursor signal as estimated by the relative peak area. In DIA, with a comparably fast cycle time of 1.6 s, the peptide was sampled 7 times over its elution profile, which is sufficient to reconstruct the chromatographic peak shape but still samples a small proportion of the total ion signal. In contrast, diaPASEF sampled the fragments in each TIMS scan and for a total of over 100 times, resulting in a nearly complete record of the fragments at every time point. Total efficiency in terms of acquisition time was 96% rather than 100% because we interspersed full scans.

We next studied the ion sampling efficiency for a standard complex proteome, a HeLa cell tryptic digest. To address the very high density of precursor ion features in the data cuboid, our diaPASEF scheme provides the possibility of balancing the width of the dynamic isolation window (reducing fragment spectral complexity) with the number of TIMS ramps needed to cover the entire precursor space (reducing duty cycle). In practice, we found that just a few ramps are usually sufficient, depending on the experimental requirements (see below). Here we chose four such ramps, each isolating about 1/4th of all precursors (**Fig. 2b**). To assess the relative performance of this scheme, we performed DDA and DIA measurements on aliquots of the same sample, using typical parameters for either method. For comparison, we extracted the total ion current of all isolated precursors in the expected peptide space in the *m/z*-ion mobility plane. The data show that for the DIA measurements the sampled fraction of the ion current was about three-fold higher than for DDA scans, whereas the four-ramp diaPASEF scheme further increased the accumulated peptide ion current by a factor of five compared to the DIA scan. We conclude that the diaPASEF principle yields the expected increase in data acquisition efficiency in both simple and complex proteomes.

Targeted data extraction in four dimensions

To identify and quantify peptides from this novel data structure, we developed Mobi-DIK (Ion Mobility DIA Analysis Kit), a software capable of analyzing the four-dimensional diaPASEF data space (**Fig. 3a**). The workflow is based on the targeted extraction of sets of fragment ions of a specific precursor from the acquired dataset, followed by statistical scoring of the generated peak group with regard to a spectral library. Mobi-DIK extends the targeted data analysis principle for DIA data[22], exemplified in the OpenSWATH software suite[21] to the fourth data dimension generated by diaPASEF. The workflow generates ion mobility-enabled spectral libraries directly from DDA PASEF runs using the MaxQuant[23,24] output and stores them in the standardized TraML[25] format. The spectral library is processed using the OpenMS tools[26,27], which we extended to support ion mobility.

The Mobi-DIK package employs the Bruker interface to query native .tdf files and converts them to mzML files using the vendor API to assign the multiple quadrupole isolation positions in diaPASEF data to individual TOF scans. The software then uses the targeted extraction paradigm for DIA data

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

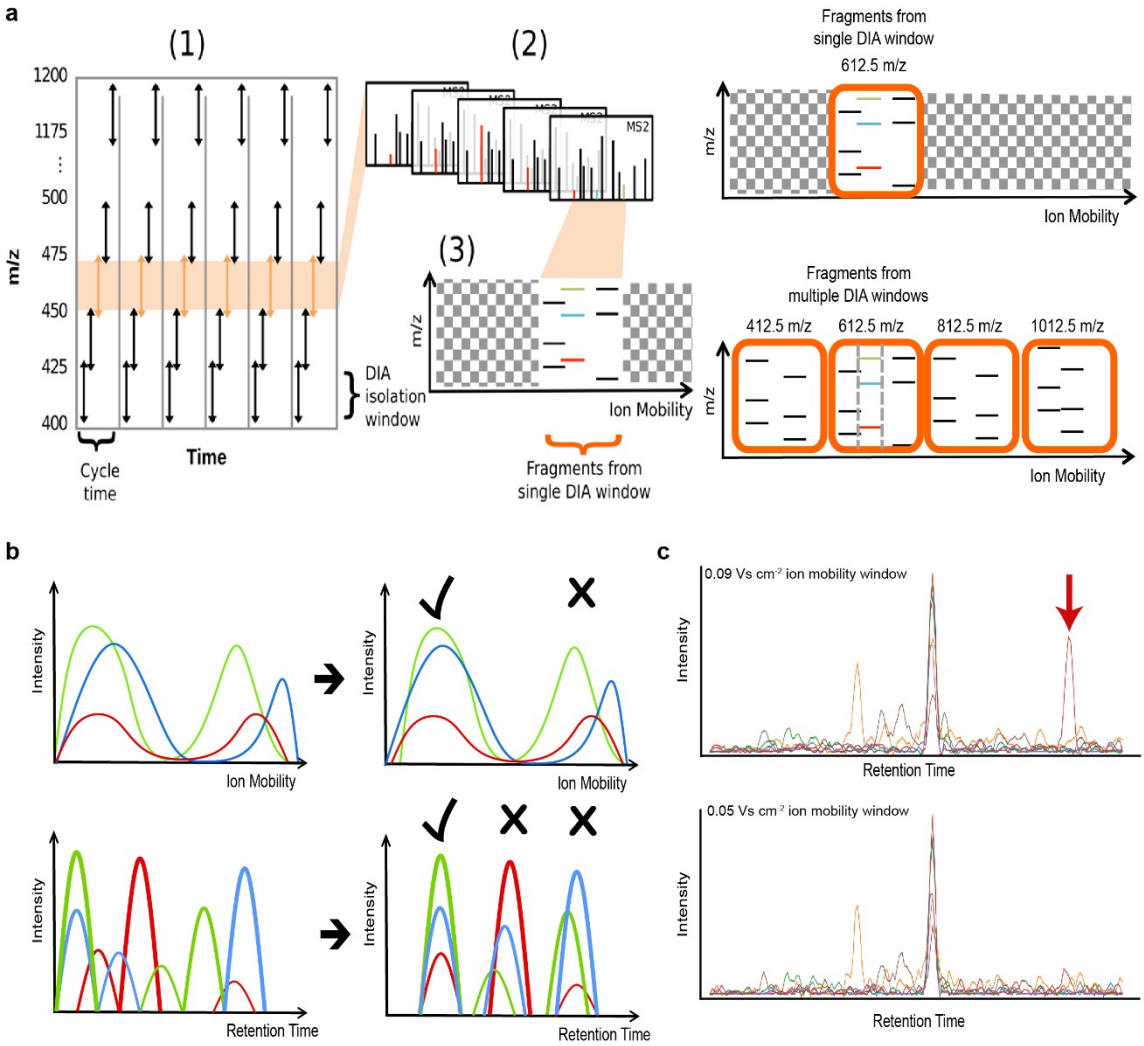

Figure 3 | Ion mobility-aware targeted data extraction. a, Steps in the Mobi-DIK workflow to extract fragment ion chromatograms from multiple DIA windows in a single diaPASEF scan. **b**, Illustration of the data extraction from the four-dimensional data cuboid for a precursor with three transitions: red, blue and green. **c**, Extracted fragment ion chromatograms from a HeLa experiment with and without ion mobility-enhanced extraction to remove interfering signals from co-eluting precursors in the same DIA window.

to extract four-dimensional data cuboids with a user-defined width in *m/z* (ppm), retention time (s) and ion mobility (Vs cm^{-2}). These extractions are projected onto the retention time and ion mobility axes to obtain fragment ion extracted ion chromatograms (XICs) and extracted ion mobilograms for each precursor-to-fragment transition in the spectral library (**Fig. 3b**). Calibration between the assay library and experimental DIA data is automatically performed in the *m/z*, retention time and ion mobility dimensions, based on a user-defined set of high confidence peptides (**Methods**). Targeted extraction in the ion mobility dimension, i.e. restricting the extraction data cube in ion mobility dimension to a user-defined width, improves the signal-to-noise by removing signals from co-eluting peptide

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

species in the very same precursor mass window (**Fig. 3c**).

From the projected precursor-to-fragment transitions, we next pick peak groups along the chromatographic dimension using the OpenSWATH peak picking and scoring modules. This step in the data analysis workflow selects putative peak candidates that are subsequently scored based on their chromatographic co-elution, goodness of library match and correlation with the precursor profile. For Mobi-DIK, we extended these modules to include scoring along the ion mobility dimension. We make use of the precision of TIMS ion mobility measurements (<<1% in replicates of complex samples[20]) to compute a highly discriminatory score based on the difference between the library-recorded and the calibrated experimental ion mobility that is combined with the other scores described above. Furthermore, we extract full ion mobilograms for each fragment ion to score the mobility peak shape as well as the peak consistency between all fragment ions in addition. In the single-run analysis of a whole-cell HeLa digest (see below), targeted extraction in the ion mobility dimension, combined with ion mobility-aware scoring greatly increased peptide identification compared to a naive analysis.

Single run proteome analysis

Having established the diaPASEF acquisition method and the Mobi-DIK analysis workflow, we next investigated the performance of our method in single-run proteome analysis of a human HeLa cancer cell line. First, we built a project-specific library from 24 high-pH reversed-phase peptide fractions with DDA PASEF comprising 135,671 target precursors and 9,140 target proteins. For sample amounts on column of at least 200 ng and 120 min LC-MS runs, we reasoned that a diaPASEF method with a somewhat lower duty cycle, but higher precursor selectivity should be beneficial. We devised a method with four windows in each 100 ms diaPASEF scan. Eight of these scans covered the diagonal scan line for doubly charged peptides in the *m/z*-ion mobility plane and added a second, parallel scan line to ensure coverage of triply charged species with narrow 25 *m/z* precursor isolation windows. For this acquisition scheme, the theoretical coverage of all precursor ions present in the library was 97% and 95% for doubly and triply charged peptides, respectively (**Fig. 4a**).

In triplicate runs, we detected a total of 66,807 peptide precursors (with 1% precursor and protein FDR), and on average 56,125 peptide precursors per run (**Fig. 4b**). The median absolute fragment ion mass accuracy was 3.8 ppm and the median absolute retention time deviation was 20 seconds, which is 0.3% of the total LC-MS runtime. The median deviation of ion mobility values in diaPASEF and library runs was 0.62% for the precursor ions, underlining the benefit of the very high reproducibility of TIMS ion mobility values. The combination of these three values defines the precise position of each precursor and fragment ion in the diaPASEF data cuboid.

Overall, we identified a total of 56,071 unique peptide sequences at 1% FDR, from which we inferred 7,565 proteins at a global protein FDR of 1% (**Fig. 4c**). Proteins were

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

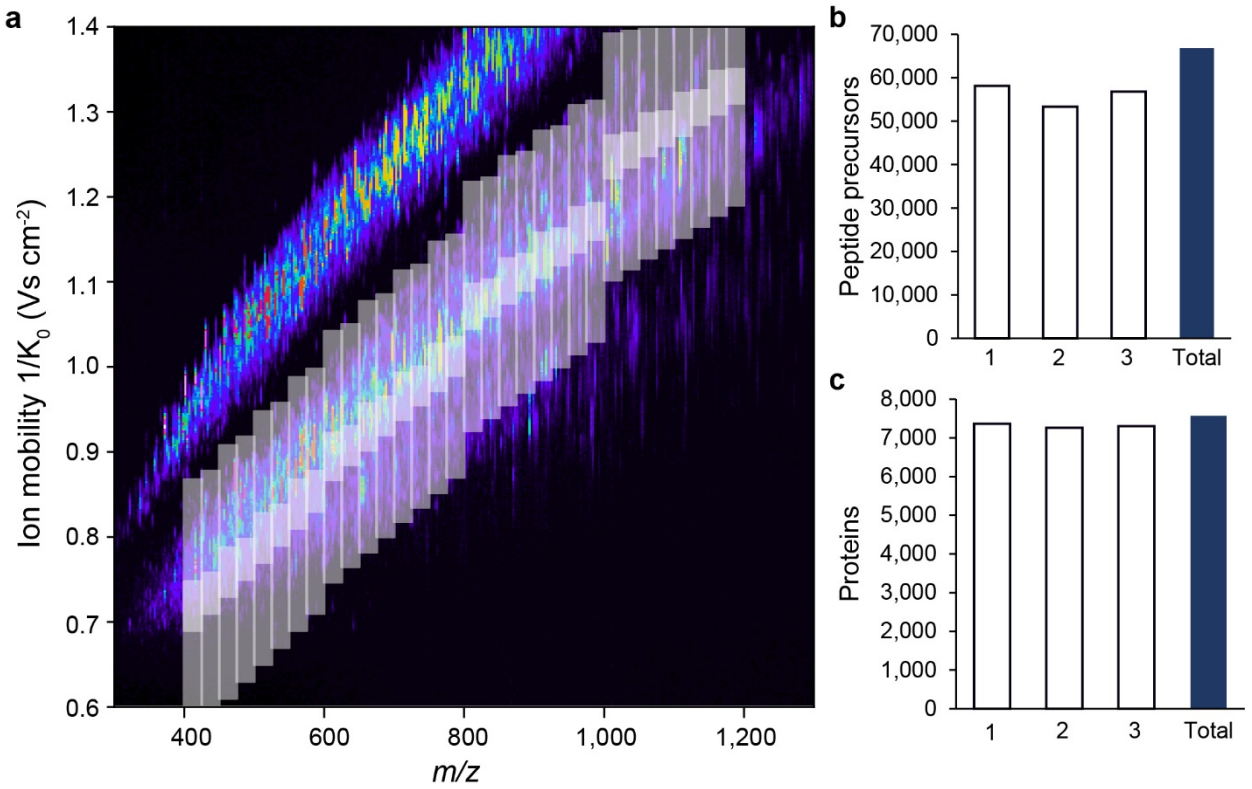

Figure 4 | HeLa proteome analysis with diaPASEF. a, Position of the precursor isolation windows in the diaPASEF acquisition scheme overlaid on the average precursor ion intensity in an 120 min LC-MS experiment. **b**,**c**, Number of b, peptide precursor ions and c, proteins identified in triplicate injections of 200 ng HeLa digest.

inferred directly using only proteotypic peptides as mapped in the low-redundancy SwissProt database without using maxium parsimony or protein grouping. Notably, these are 83% of all proteins in the library and they are covered in single runs without fractionation. Out of these, 6,974 proteins were quantified in all three replicates, 406 in two and only 106 proteins in a single replicate, which translates into a data completeness of 96%. The median coefficient of variation (CV) for fragment-ion based quantification was 14.2% on the precursor, and 11.5% on the protein level after median normalization. The detected proteins spanned a dynamic range of about four orders of magnitude, as estimated by the 'best flier peptide' approach summing the intensities of the top 5 fragment ions from the top 3 peptides (**Methods**).

Very high sensitivity proteomics with diaPASEF

One of the strengths of diaPASEF is that it can be tuned to utilize a very high fraction of the incoming ion beam and still achieve a high precursor selectivity, because single-charge background ions are excluded from the analysis and TIMS separates co-eluting precursors in the same mass window. One application of this principle is the accumulation of very high ion signal for low-abundance precursors. To demonstrate this concept, we analyzed only 10 ng of HeLa digest in triplicate 120 min single runs. In this

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

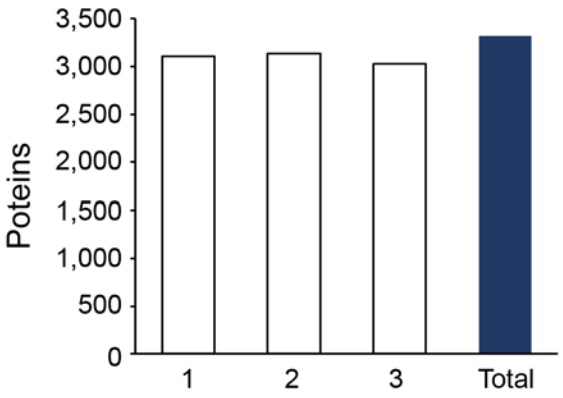

Figure 5 | diaPASEF analysis of 10 ng HeLa digest. Number of proteins identified in triplicate injections.

case, we employed a diaPASEF scheme that samples about 25% of the ion current and covers the most dense precursor ion region.

With this high duty cycle, we were able to identify on average 10,194 precursors and 9,636 unique peptide sequences per run. In total, we identified 3,323 proteins with diaPASEF (**Fig. 5**), which also outperforms our previously published results with DDA PASEF where we identified 2,723 proteins on average[20]. Even with these minimal sample amounts, we achieved a data completeness of 85% and a median CV of 9.4% on the protein level.

DISCUSSION

Here we have developed and demonstrated a DIA workflow for the parallel storage serial fragmentation principle supported in a TIMS-TOF mass spectrometer. Making use of the correlation between the ion mobility and the *m/z* of peptides, our diaPASEF method comes close to maximum ion utilization for mass spectrometry as all the fragments can be stored and analyzed at the same time. This is in contrast to DDA methods, which convert only a very small fraction of the incoming ion beam into fragments (generally <<1%) and even typical DIA workflows, which convert a few percent of the ion beam. On simple mixtures, we achieve close to 100% of the maximum, whereas for more complex mixtures, it was beneficial to use the quadrupole to separately isolate two or a few ranges, at a cost of reducing the fraction of total available ions sampled. To extract information using spectral library-based targeted data analysis, we extended the OpenSWATH tool developed for DIA applications to efficiently make use of the ion mobility dimension for library matching, providing full FDR control and excellent quantification.

Even in this first implementation, we achieved deep proteome coverage of more than 7,000 proteins in single, 2 h LC runs from a total of ca. 200 ng peptide sample on column with high degree of reproducibility. Even more remarkably, we identified over 3,000 proteins from only 10 ng of total peptide mass on column in single runs. This latter result points to a perhaps unexpected advantage of diaPASEF, namely the fact that the extremely high ion sampling also fully translates to higher sensitivity. Likewise, the very fast cycle time of our new scan mode should be very advantageous for short gradients, an increasingly important attribute as large scale biological and clinical studies require very large throughput. For the future, we imagine that both hardware and software can still be greatly optimized to further increase the amount and quality of the information contained in and extracted from the extremely rich four-dimensional diaPASEF data cuboids. Furthermore, we note that application of diaPASEF is not restricted to peptides but could equally well

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

be extended to metabolites, lipids or other compound classes.

METHODS

Sample preparation. The human cancer cell line (HeLa S3, ATCC) was cultured in Dulbecco's modified Eagle's medium with 10% fetal bovine serum, 20 mM glutamine and 1% penicillin-streptomycin. Cells were collected by centrifugation, washed with phosphate-buffered saline, flash-frozen in liquid nitrogen and stored at -80°C. Cell lysis, reduction and alkylation was performed in lysis buffer with chloroacetamide (PreOmics) following our previously published protocol[28]. Briefly, the cell suspension was heated to 95°C for 10 min and subsequently sonificated to further disrupt cells and shear nucleic acids. Proteins were enzymatically cleaved overnight by adding equal amounts of LysC and trypsin in a 1:100 (wt/wt) enzyme:protein ratio. De-salting and purification was performed according to the PreOmics iST protocol on a styrendivinylbenzene reversed-phase sulfonate (SDB-RPS) sorbent. Purified peptides were vacuum-centrifuged to dryness and reconstituted in double-distilled water with 2 vol.-% acetonitrile (ACN) and 0.1 vol.-% trifluoroacetic acids (TFA) for single run LC-MS analysis or fractionation.

High-pH reversed-phase fractionation. Peptides were fractionated at pH 10 with a 'spider fractionator' coupled to an EASY-nLC 1000 chromatography system (Thermo Fisher Scientific) as described previously[29]. Approximately 50 µg purified peptides were separated on a 30 cm C_{18} column within 96 min and automatically concatenated into 24 fractions by shifting the exit valve every 120 s. The fractions were vacuum-centrifuged to dryness and re-constituted in double-distilled water with 2 vol.-% ACN and 0.1 vol.-% trifluoroacetic acids TFA for LC-MS analysis.

Liquid chromatography. Nanoflow reversed-phase chromatography was performed on an EASY-nLC 1200 system (Thermo Fisher Scientific). Peptides were separated within 120 min at a flow rate of 300 nL/min on a 50 cm x 75 µm column with a laser-pulled electrospray emitter packed with 1.9 µm ReproSil-Pur C_{18}-AQ particles (Dr. Maisch). Mobile phases A and B were water with 0.1 vol.-% formic acid and 80/20/0.1 vol.-% ACN/water/formic acid. The %B was linearly increased from 5 to 30% within 95 min, followed by an increase to 60% within 5 min and a further increase to 95% before re-equilibration.

Mass spectrometry. LC was coupled online to a hybrid TIMS quadrupole time-of-flight mass spectrometer (Bruker timsTOF Pro) via a CaptiveSpray nano-electrospray ion source. A detailed description of the instrument is available in ref. [20]. The dual TIMS analyzer was operated at a fixed duty cycle close to 100% using equal accumulation and ramp times of 100 ms each. We performed data-dependent data acquisition in PASEF mode with 10 PASEF scans per topN acquisition cycle. Singly charged precursors were excluded by their position in the *m/z*-ion mobility plane and precursors that reached a 'target value' of 20,000 a.u. were dynamically excluded for 0.4 min. The

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

quadrupole isolation width was set to 2 Th for *m/z* < 700 and 3 Th for *m/z* > 700.

To perform data-independent acquisition, we extended the instrument control software to define quadrupole isolation windows as a function of the TIMS scan time (diaPASEF). The instrument control electronics were modified to allow seamless and synchronous ramping of all applied voltages. We tested multiple schemes for data-independent precursor windows and placement in the *m/z*-ion mobility plane and defined up to 8 windows for single 100 ms TIMS scans as detailed in the main text. In both scan modes, the collision energy was ramped linearly as a function of the mobility from 59 eV at $1/K_0$=1.6 Vs cm^{-2} to 20 eV at $1/K_0$=0.6 Vs cm^{-2}.

Spectral library generation. To generate spectral libraries for targeted data extraction, we first analyzed the high pH reversed-phase fraction acquired in DDA mode with MaxQuant version 1.6.5.0, which extracts four-dimensional features on the MS1 level (retention time, m/z, ion mobility and intensity) and links them to peptide spectrum matches. The maximum precursor mass tolerance of the main search was set to 20 ppm and deisotoping of fragment ions was deactivated. Other than that, we used the default 'TIMS-DDA' parameters. MS/MS spectra were matched against an in-silico digest of the Swiss-Prot reference proteome (human 20,414 entries, downloaded July 2019) and a list of common contaminants. The minimum peptide length was set to 7 amino acids and the peptide mass was limited to 4,600 Da. Carbamidomethylation of cysteine residues was defined as a fixed modification, methionine oxidation and acetylation of protein N-termini were defined as variable modifications. The false discovery rate was controlled <1% at both, the peptide spectrum match and the protein level. Our Mobi-DIK software package builds on OpenMS tools to compile spectral libraries in the standardized TraML or pqp formats from the MaxQuant output tables while retaining the full ion mobility information for each precursor-to-fragment ion transition. Only proteotypic peptides with a precursor *m/z* >400 were included in the library and required to have a minimum of 6 fragment ions with *m/z* >350 and outside the precursor mass isolation range.

Targeted data extraction. To analyze diaPASEF data, we developed an ion mobility DIA analysis kit (Mobi-DIK) that extracts fragment ion traces from the four-dimensional data space as detailed in the main text. Raw data were automatically re-calibrated using 10,000 curated reference values in *m/z*, retention time and ion mobility dimensions. We applied an outlier detection in each dimension before calculating the final fit function to increase robustness. Peak picking and sub-sequent scoring functionalities in the Mobi-DIK software build on OpenSWATH modules. For diaPASEF, we extended these modules to also take into account the additional ion mobility dimension. PyProphet was used to control FDRs <1% at both, the peak group and protein level with a target-decoy approach. In case of two overlapping diaPASEF windows, the analysis was performed separately for the individual windows and for FDR estimations, the highest scoring peak group was selected. Protein abundances were inferred using the 'best flier peptide' approach summing the

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

intensities of the top 5 fragment ions from the top 3 peptides[30,31].

Bioinformatics. Output tables from the data analysis pipeline or MS raw data were further analyzed and visualized in the R statistical computing environment or in Python.

REFERENCES

1. Altelaar, A. F. M., Munoz, J. & Heck, A. J. R. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nat. Rev. Genet.* **14,** 35–48 (2012).
2. Larance, M. & Lamond, A. I. Multidimensional proteomics for cell biology. *Nat. Rev. Mol. Cell Biol.* **16,** 269–280 (2015).
3. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537,** 347–355 (2016).
4. Bekker-Jensen, D. B. *et al.* An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes. *Cell Syst.* **4,** 587–599.e4 (2017).
5. Wang, D. *et al.* A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. *Mol. Syst. Biol.* **15,** e8503 (2019).
6. Rost, H. L., Malmstrom, L. & Aebersold, R. Reproducible quantitative proteotype data matrices for systems biology. *Mol. Biol. Cell* **26,** 3926–3931 (2015).
7. Doerr, A. DIA mass spectrometry. *Nat. Methods* **12,** 35 (2014).
8. Chapman, J. D., Goodlett, D. R. & Masselon, C. D. Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. *Mass Spectrom. Rev.* **33,** 452–470 (2014).
9. Ludwig, C. *et al.* Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* **14,** e8126 (2018).
10. Gillet, L. C., Leitner, A. & Aebersold, R. Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. *Annu. Rev. Anal. Chem.* **9,** 449–472 (2016).
11. Bilbao, A. *et al.* Processing strategies and software solutions for data-independent acquisition in mass spectrometry. *Proteomics* **15,** 964–980 (2015).
12. McLean, J. a., Ruotolo, B. T., Gillig, K. J. & Russell, D. H. Ion mobility–mass spectrometry: a new paradigm for proteomics. *Int. J. Mass Spectrom.* **240,** 301–315 (2005).
13. Distler, U. *et al.* Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* **11,** 167–70 (2014).
14. Helm, D. *et al.* Ion Mobility Tandem Mass Spectrometry Enhances Performance of Bottom-up Proteomics. *Mol. Cell. Proteomics* **13,** 3709–15 (2014).
15. Ewing, M. A., Glover, M. S. & Clemmer, D. E. Hybrid ion mobility and mass spectrometry as a separation tool. *J. Chromatogr. A* **1439,** 3–25 (2016).
16. Fernandez-Lima, F. A., Kaplan, D. A. & Park, M. A. Note: Integration of trapped ion mobility spectrometry with mass spectrometry. *Rev. Sci. Instrum.* **82,** 126106 (2011).

bioRxiv preprint doi: https://doi.org/10.1101/656207; this version posted May 31, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

17. Fernandez-Lima, F., Kaplan, D. a, Suetering, J. & Park, M. a. Gas-phase separation using a trapped ion mobility spectrometer. *Int. J. Ion Mobil. Spectrom.* **14,** 93–98 (2011).

18. Ridgeway, M. E., Lubeck, M., Jordens, J., Mann, M. & Park, M. A. Trapped ion mobility spectrometry: A short review. *Int. J. Mass Spectrom.* **425,** 22–35 (2018).

19. Meier, F. *et al.* Parallel Accumulation–Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device. *J. Proteome Res.* **14,** 5378–5387 (2015).

20. Meier, F. *et al.* Online Parallel Accumulation–Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer. *Mol. Cell. Proteomics* **17,** 2534–2545 (2018).

21. Röst, H. L. *et al.* OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nat. Biotechnol.* **32,** 219–223 (2014).

22. Gillet, L. C. *et al.* Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Mol. Cell. Proteomics* **11,** O111.016717-O111.016717 (2012).

23. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26,** 1367–72 (2008).

24. Prianichnikov, N., Koch, H., Koch, S., Lubeck, M. & Heilig, R. MaxQuant software for ion mobility enhanced shotgun proteomics. *bioRxiv* 1–30 (2019). doi:10.1101/651760

25. Deutsch, E. W. *et al.* TraML—A Standard Format for Exchange of Selected Reaction Monitoring Transition Lists. *Mol. Cell. Proteomics* **11,** R111.015040 (2012).

26. Rosenberger, G. *et al.* Statistical control of peptide and protein error rates in large-scale targeted data-independent acquisition analyses. *Nat. Methods* **14,** 921–927 (2017).

27. Röst, H. L. *et al.* OpenMS: a flexible open-source software platform for mass spectrometry data analysis. *Nat. Methods* **13,** 741–8 (2016).

28. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11,** 319–24 (2014).

29. Kulak, N. A., Geyer, P. E. & Mann, M. Loss-less nano-fractionator for high sensitivity, high coverage proteomics. *Mol. Cell. Proteomics* mcp.O116.065136 (2017). doi:10.1074/mcp.O116.065136

30. Rosenberger, G., Ludwig, C., Röst, H. L., Aebersold, R. & Malmström, L. aLFQ: an R-package for estimating absolute protein quantities from label-free LC-MS/MS proteomics data. *Bioinformatics* **30,** 2511–2513 (2014).

31. Schubert, O. T. *et al.* Absolute Proteome Composition and Dynamics during Dormancy and Resuscitation of Mycobacterium tuberculosis. *Cell Host Microbe* **18,** 96–108 (2015).