International Journal of Mass Spectrometry 452 (2020) 116336

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

Effects of electron-transfer/higher-energy collisional dissociation (EThcD) on phosphopeptide analysis by data-independent acquisition

Thierry Schmidlin [1], Maarten Altelaar*

Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University and Netherlands Proteomics Centre, Padualaan 8, 3584 CH, Utrecht, the Netherlands

ARTICLE INFO

Article history:
Received 11 March 2020
Accepted 25 March 2020
Available online 31 March 2020

ABSTRACT

In quantitative proteomics the use of data-independent acquisition (DIA) methods is getting increasingly popular and is applied to a growing number of sample types ranging from peptidomics to the analysis of post-translational modifications. Several of these sample types have been previously shown to profit from electron-based fragmentation methods such as ETD or EThcD. These fragmentation methods have so far not been implemented in DIA analysis. Here, we show the feasibility of combining DIA with EThcD fragmentation and provide insights into its performance. We show how EThcD can be used to increase peptide coverage, reaching similar success rates during targeted data extraction as standard HCD-based DIA. Furthermore, we illustrate how robust MS1 and LC results can be exploited to circumvent detrimental effects of less efficient fragmentation during targeted data extraction. Ultimately, our data demonstrates how EThcD-based MS/MS spectra can be connected to standard HCD-based DIA analyses, enabling the future use of decision-tree based spectral libraries to query large DIA data sets.

1. Introduction

Mass Spectrometry-based proteomics has taken up an increasingly important role in biochemical research throughout the last decades. Constant improvements in instrument technology and acquisition methods has led to ever increasing proteome coverage since then. A major step in this process was the development of data-independent acquisition methods, which omit the intensity-based selection of ions subjected to fragmentation [1]. Initially introduced in the mid 2000s [2], DIA type of approaches have later been successfully applied to various samples such as small molecules [3] and peptides [4]. In 2012 the DIA methodology gained additional momentum when Gillet et al. introduced the concept of sequential window acquisition of all theoretical spectra (SWATH) in conjunction with targeted data extraction as a mean of DIA data analysis [5]. This opened new avenues for consistent and accurate proteome analysis by creating a 'permanent quantitative digital proteome maps', which can be repeatably analyzed and reanalyzed [6].

In contrast to the more commonly used data-dependent acquisition (DDA) analysis, SWATH-type DIA consecutively cycles through a series of wide isolation windows within a chromatographic time scale, giving rise to highly multiplexed MS2-maps. These can be queried for fragment ion traces for numerous peptides by using pre-acquired spectral libraries. This strategy results in quantitative consistency and accuracy comparable to those of Selected Reaction Monitoring while providing proteomic depth comparable to DDA LC-MS [7].

Several studies have since then successfully applied the DIA technology for the analysis of various sample types, ranging from patient biopsies [6] and plasma [8] up to the analysis of N-glycosylation [9] and protein phosphorylation [10,11]. At the same time, extensive method development has been performed, seeking for improvements in terms of specificity and sensitivity. The proposed solutions range from the use of variable Q1 isolation window sizes [12] up to the simultaneous analysis of multiple Q1 isolation windows in combination with subsequent data deconvolution [13].

Surprisingly however, no in-depth investigation of using alternative peptide fragmentation methods such as EThcD in DIA has been presented thus far. This despite the known benefit of electron driven fragmentation methods in many aspects of proteomics such as peptidomics [14], glycoproteomics [15] and phosphoproteomics [16]. Therefore, the present study sets out to investigate different

* Corresponding author.
E-mail address: m.altelaar@uu.nl (M. Altelaar).
[1] present address: Nuffield Department of Medicine, University of Oxford, Oxford, OX3 7BN, United Kingdom.

https://doi.org/10.1016/j.ijms.2020.116336
1387-3806/

aspects of DIA in terms of EThcD peptide fragmentation and data analysis. In this study, we show the principal feasibility of combining EThcD with DIA for the analysis of phosphopeptides.

2. Materials and methods

2.1. Cell cultures

HeLa cells were grown in Dulbeco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum and 10 mM glutamine (all from Lonza, Braine-l'Alleud, Belgium). Six hours before harvesting, the medium was replaced by fresh medium. Cells were harvested and the cell pellets were immediately washed two times with phosphate-buffered saline buffer (PBS) and stored at −80 °C until further usage.

2.2. Protein lysis and digestion

Cells were lysed, reduced, and alkylated in lysis buffer (1% sodium deoxycholate (SDC), 10 mM tris(2-carboxyethyl)phosphine-hydrochloride (TCEP)), 40 mM chloroacetamide (CAA), and 100 mM TRIS, pH 8.0 supplemented with phosphotase inhibitor (PhosSTOP, Roche) and protease inhibitor (cOmplete mini EDTA-free, Roche). Cells were heated for 5 min at 95 °C, sonicated with a Bioruptor Plus (Diagenode) for 15 cycles of 30 s and cell debris was removed by centrifugation at 20 000 g for 10 min. A Bradford protein assay was used to quantify protein amount. Prior to digestion, samples were diluted 1:10 with 50 mM ammoniumbicarbonate, pH 8.0. Proteins were digested overnight at 37 °C with trypsin (Sigma-Aldrich) with an enzyme/substrate ratio of 1:50 and lysyl endopeptidase (Wako) with an enzyme/substrate ratio of 1:75. SDC was precipitated with 2% formic acid (FA) and samples were desalted using Sep-Pak C18 cartridges (Waters), dried *in vacuo* and stored at −80 °C until further use.

2.3. High pH-fractionation

A total of 4 mg HeLa cell digest was fractionated on a high-pH (HpH) reversed-phase C18 column (Gemini 3u C18 110 Å, 100 × 1.0 mm, Phenomenex) coupled to an Agilent 1100 series (Agilent Technologies) as previously described [17]. In brief samples were reconstituted in 10 mM ammonium hydroxide, pH 10 and loaded directly on the column by the pump. Peptides were concentrated on column with buffer A (10 mM Ammonium Hydroxide, pH 10) at 100 μL/min for 2 min. For the subsequent fractionation the gradient was initiated as follows: 5% solvent B (10 mM ammonium Hydroxide in 90% ACN, pH 10) to 30% B in 53 min, 70% B in 7 min and increased to 100% B in 3 min at a flow rate of 100 μl/min. In total 67 fractions of 1 min were collected by an Agilent 1260 Infinity fraction collector and pooled into 5 combined fractions using a concatenation strategy previously described [17]. The pooled samples were subsequently dried *in vacuo* and stored at −80 °C.

2.4. Phosphopeptide enrichment

Phosphopeptide enrichment was performed as previously described [18]. In brief phosphorylated peptides were enriched using Fe(III)-NTA 5 μL (Agilent technologies) in an automated fashion using the AssayMAP Bravo Platform (Agilent Technologies). Fe(III)-NTA cartridges were primed with 200 μL of 0.1% TFA in ACN and equilibrated with 250 μL of loading buffer (80% ACN/0.1% TFA). Samples were dissolved in 200 μL of loading buffer and loaded onto the cartridge at a loading speed of 5 μL/min. Loading amounts per tip were 250 μg for unfractionated HeLa lysate. Each HpH fraction was divided into two samples enriched on one cartridge each. After sample loading columns were washed with 250 μL loading buffer and eluted with 35 μL of 10% ammonia directly into 35 μL of 10% formic acid. Samples were dried down and stored at −80 °C until LC–MS analysis.

2.5. LC-MS setup

Unless otherwise indicated, all experiments were performed on an Orbitrap Fusion mass spectrometer (Thermo Scientific) coupled to an Agilent 1290 Infinity System (Agilent Technologies) adapted to nanoflow conditions by using a split flow setup as described in Ref. [19]. The system was operated with in-house packed trap column (Dr. Maisch Reprosil C18, 3 μm, 2 cm × 100 μm) and analytical column (Agilent Poroshell 120 EC-C18, 2.7 μm, 50 cm × 75 μm). The split flow was adapted to achieve 300 nl/min flow at the front end of the column upon applying a flow rate of 0.2 mL/min. 0.6% acetic acid in water (Milli-Q, Millipore) was used as buffer A and 0.6% acetic acid, 80% ACN was used as buffer B.

2.6. Data acquisition

All samples were reconstituted in 10% formic acid in water and injected at a volume of 10 μL. Upon injection peptides were trapped for 5 min at a 5 μL/min. Subsequently peptides were chromatographically separated on a 100 min gradient running from 8% B to 32% B followed by column washing (ramping up to 100% B for 3 min followed 100% B for 1 min) and equilibration (ramping down to 0% B for 1 min followed by 0% B for 10 min). Parameters used for data-dependent acquisition were set as follows: Survey scans were acquired in the Orbitrap at 60K resolution spanning the 375-1500 *m/z* range using an AGC target of 4e5 and a maximum fill time of 50 ms. Data-dependent MS2 scans were acquired for charge states 2–6 using an isolation width of 1.6 Th and a 12 s dynamic exclusion. HCD and EThcD fragmentations were used respectively. All MS2 scans were recorded in the Orbitrap in centroid mode, scanning the 350–1000 Th range at 30K resolution using an AGC target of 5e4 and a maximum fill time of 35 ms. Parameters used for data-independent acquisition were set as follows: MS scans were acquired in the Orbitrap at 120K resolution covering the 400-1000 *m/z* range. Orbitrap filling was controlled by an AGC target of 2e5 and a maximum injection time 100 ms. DIA scans were recorded using quadrupole isolation of 30 equal sized windows throughout the 400-1000 *m/z* range. All MS2 scans were recorded in the Orbitrap measuring the 350-1000 *m/z* range in profile mode at 30K resolution. Orbitrap fill time was controlled through an AGC target of 5e4 and a maximum fill time of 100 ms. Two different fragmentation modes were used, HCD with stepped collision energy (25% +/- 5), and EThcD (stepped CE 5%).

2.7. Data analysis

Raw files from data-dependent acquisition analyses were searched against a concatenated database consisting of the SwissProt database (version 56.2) and the Biognosys HRM Calibration Kit peptide sequences using SEQUEST accessed by Proteome Discoverer (version 1.4). Parameters were set to tryptic digest, allowing for up to three missed cleavages, using carbamidomethyl cysteine as fixed modification and allowing for serine/threonine/tyrosine phosphorylation and methionine oxidation. Precursor mass tolerance and MS/MS tolerance were set to 50 ppm and 0.05 Da respectively. Results were filtered using Percolator [20] to an FDR below 1%. PhosphoRS [21] was used to evaluate phosphosite localization. Search results were subsequently used to create spectral libraries in Skyline [22]. These spectral libraries were used

for subsequent DIA data extraction. Unless otherwise stated we attempted to extract all identified peptides within the DIA *m/z* range from the DIA runs alongside with an equally sized decoy group. MS1 filtering was set to include 3 isotope peaks using a resolving power of 60 000 at 400 *m/z*. MS/MS filtering was used in DIA mode, following multiple isolation modes as specified in detail for each experiment, using a resolving power of 60 000 at 400 *m/z*. mProphet as implemented in the Skyline advanced peak picking option was used for FDR control of the targeted data extraction ($q \leq 0.01$ was used as significance threshold throughout the whole study). Quantitative analysis was performed by exporting peak area values from Skyline. For MS1 quantification the peak areas for M, M+1 and M+2 ions were summed up, for MS2 quantification the peak areas of all fragments were summed up. Non-integrated ion traces were considered 0 and peptides with total intensity of 0 were excluded from the analyses. CVs and mean quantitative values were calculated and plotted in Excel.

3. Results and discussion

3.1. Spectral library acquisition for targeted data extraction

Here, we set out to investigate the contribution of EThcD fragmentation compared to the more commonly used HCD fragmentation for the analysis of phosphopeptides by DIA. As a model system we chose HeLa cell lysates enriched for phosphopeptides by Fe(III)-IMAC. Key step to most DIA analyses is the acquisition of reference spectral libraries. This typically entails fractionation of the sample to obtain substantial proteomic depth [23,24]. Here, HeLa lysates were fractionated by high-pH reverse phase (HpH) in combination with a concatenation strategy as described by Batth et al. [17]. Five concatenated fractions were each enriched for phosphopeptides by Fe(III)-IMAC. Each of the 5 fractions was analyzed in DDA mode by HCD and EThcD separately (Fig. 1A). This enabled us to create spectral libraries containing spectra for 6624 (HCD) and 5318 (EThcD) phosphopeptides respectively, counting potential phosphosite localization isomers individually if present. Fig. 1B depicts the overlap between the 2 spectral libraries. Representative MS2 spectra of both methods are depicted in Supporting Information Fig. S1.

3.2. Method development for DIA methods using HCD and EThcD

For DIA analyses, HeLa samples were subjected to phosphopeptide enrichment without prior fractionation and directly analyzed by DIA using a 2h LC separation. Two different DIA methods were tested, scanning the 400–1000 Th range in 30 steps of 20 Th, using HCD and EThcD-fragmentation, respectively (Fig. 1A). DIA by nature relies on high speed instrumentation, capable of providing a good balance between a short cycle time and an appropriate dwell time for a high number of isolation windows. Compared to flow-through instruments such as Q-TOFs, iontrap based instruments provide less control over cycle time. This is primarily due to the dynamic regulation of the Orbitrap fill time, which is controlled by an AGC target and a maximum fill time. Additionally, electron-based fragmentation methods are slower than HCD and can thus substantially contribute to an increased cycle time. Due to these effects we set out to determine the actual cycle times of our methods empirically. We observed a cycle time of roughly 3 s for HCD fragmentation and 6 s for EThcD fragmentation.

The longer duty cycle of EThcD in comparison to HCD is a clear disadvantage for DIA analysis. However, given the average peak width at base of 45-60 s obtained on our LC-system, cycle times of 6 s still ensure sufficient data points across the chromatographic peak to perform robust AUC-based quantification [25]. For that reason, differences in cycle time were not considered, instead the effect of EThcD fragmentation on the targeted data extraction was assessed in more detail.

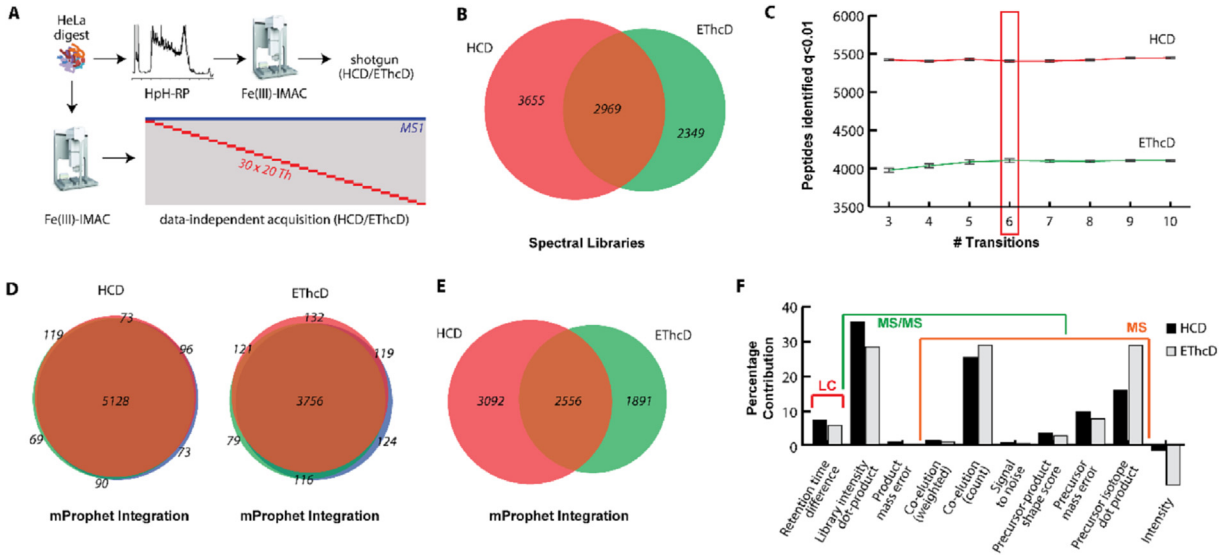

Fig. 1. Comparison of HCD-DIA, EThcD-DIA and ETD-DIA.
(A) Workflow used for the analysis of EThcD- and HCD-based methods for the analysis of phosphopeptides by DIA. Spectral libraries were acquired by subjecting a tryptic digest of HeLa cells to high-pH fractionation collecting 5 concatenated fractions subsequently enriched for phosphopeptides by Fe(III)-IMAC. Each fraction was analyzed twice by shotgun LC-MS using HCD and EThcD respectively (upper panel). For DIA analysis no fractionation was performed prior to phosphopeptide enrichment. DIA methods were acquired using 30 x 20 Th isolation windows and analyzed in triplicates by HCD and EThcD respectively (lower panel). (B) Number of phosphopeptides identified in the HCD-, and EThcD-library respectively, depicted as Venn diagram. (C) Optimization of the data analysis for HCD-DIA and EThcD-DIA in terms of transition numbers used for targeted data extraction. Error bars depict standard deviation Red bar indicates the number of transitions used for subsequent analyses (6). (D) Overlap of successful peak integrations (mProphet q-value ≤ 0.01) for HCD and EThcD-DIA respectively. (E) DIA identification (mProphet q-value ≤ 0.01) for each of the three fragmentation methods depicted as Venn diagram. (F) Contribution of multiple mProphet sub-scores to the composite score used for FDR-controlled re-integration. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article).

3.3. Optimizing data analysis by applying mProphet

First, we performed targeted data extraction for both DIA analyses using their respective spectral library as a reference. Through simultaneous extraction of decoy sequences from the data, this step could subsequently be assessed in an error-controlled manner. The method of choice for this control step is mProphet, a software package that scores multiple data quality criteria, provides an FDR control and can be optimized in a sample specific way by semi-supervised learning [26]. Initially we set out to determine the optimal amount of transitions to be extracted from the MS2-map for each of the two methods individually. Therefore, all peptides present in each library, together with an equal amount of decoy sequences, were extracted from each DIA run using the 3 most abundant isotopic peaks of the precursor ions (M, M+1, M+2) in combination with 3 to 10 fragment ions. For each number of transitions an individual mProphet model was created and extracted peak groups were reintegrated using a $q \leq 0.01$ ($\leq$1% FDR) cutoff. Surprisingly, no drastic changes in successful peak integration were observed upon increasing the number of transitions in both methods (Fig. 1C). Therefore, all subsequent analyses were performed using a standard number of 6 transitions unless otherwise described.

In contrast to DDA, DIA has been shown to be more reproducible across various samples, resulting in less missing values for quantification across large numbers of samples [27]. By comparing successful data extraction from three HCD-DIA injection replicas we could confirm this characteristic, as 91% of the extracted peptides were significantly identified ($q \leq 0.01$) in all three analyses (Supporting Table 1). Interestingly, this reproducibility is slightly worse in EThcD-DIA, displaying an overlap of 84% only (Fig. 1D). Overall retention time correlation plots of peptides detected in multiple runs are depicted in Supporting Fig. S2. Globally, this results in an overall number of 5648 peptides identified ($q \leq 0.01$) in at least one HCD-DIA run and 4447 peptides identified ($q \leq 0.01$) in at least one EThcD-DIA run. Their overlap is depicted in Fig. 1E. Success rates of both DIA methods compared to the library size are fairly similar for both methods ranging from 83.6% for EThcD-DIA (4447 of 5318 peptides) to 85.3% for HCD-DIA (5648 of 6624 peptides).

To obtain a better understanding of the effects of the various MS parameters on DIA data and their analysis, we further investigated mProphet scoring for HCD-DIA and EThcD-DIA. The actual criteria scored for each sub-score can be grouped into LC-dependent (RT-deviation), MS1-dependend (precursor mass error, precursor isotope dot product, etc.), and MS2-dependent (library dotP, product mass error, etc.) as well as combinations thereof. Through semi-supervised learning these sub-scores are weighted differently according to the actual quality of the data at hand [26]. When comparing the two fragmentation methods it is striking that the contribution of the MS2-dependent sub-scores, especially the library dot product, are higher in HCD-DIA than in EThcD-DIA (Fig. 1F). This is likely caused by low fragment ion intensity in EThcD due to the generation of c- and z-ion series in addition to the b- and y-ion series from the same amount of precursor ions. In addition, EThcD is more prone to residual precursor ion signals in the MS2 spectra due to less efficient fragmentation (Illustrated in Supporting Information Fig. S3). To compensate for this, sub-scores relying on MS1 data, especially the precursor isotope dot product, increase dramatically. Hence, the lower intensities in MS2 data of EThcD does not *per se* hamper DIA analysis – although the reliance on MS1 data automatically increases.

3.4. MS1 versus MS2 quantification in DIA

This observation caught our interest as an increasing reliance on MS1 data seems counterintuitive in regards to the central tenet of data-independent acquisition exploiting MS2 quantification because of its increased quantitative accuracy over MS1. We used our dataset to investigate the differences of MS1- and MS2-based quantification in HCD-DIA and EThcD-DIA. To that end we determined quantitative values for each peptide in both methods by summing the area under the curve of all fragment ions and of the three precursor ions M, M+1 and M+2 respectively. Accuracy was assessed by CV distributions across three replicas (depicted as boxplot in Supporting Information Fig. S4). Phosphopeptides not successfully integrated in all three technical replicas were excluded from the analysis, reducing the sample sizes to 4665 and 3360 phosphopeptides for MS1 quantification in HCD-DIA and EThcD-DIA respectively. Unsuccessful integration was less common in the MS2 based quantification resulting in 5049 phosphopeptides in HCD-DIA and 3554 phosphopeptides in EThcD-DIA. CV medians range from 4% (HCD MS1-based quantification) to 8% (EThcD MS2-based quantification). Strikingly EThcD-DIA fragment ion-based quantification seems to result in the highest amount of variability, either on a level of MS/MS acquisition or subsequent data analysis. This seems not surprising *per se*, as it can indeed be an effect of the lower efficiency of the fragmentation method compared to HCD, resulting in lower fragment ion intensities combined with lower quantitative reproducibility. More surprising, however, is the slightly better reproducibility observed for HCD MS1-based quantification compared to HCD MS2-based quantification. This seems to stand in stark contrast to the commonly held idea that MS2-based quantification provides the more accurate results. This might be different in our case due to the increasing resolution obtained from newer generation Orbitrap instrument compared to TripleTOF instruments used by Gillet et al. in 2012 [5]. Alternatively, the presence of MS2 traces might have increased the accuracy of the peak integration used for MS1-based quantification in comparison to peak picking algorithms used for MS1 quantification in DDA data analysis pipelines. Lastly, we observed MS1 based quantification to be more accurate in HCD-DIA than in EThcD-DIA. Likely, this observation is mainly caused by the longer cycle times observed in EThcD resulting in lower number of MS1 measurements across the chromatographic peak.

The lower MS2 quantitative accuracy obtained in EThcD-DIA compared to MS1 quantification also results in a lower quantitative cross method comparability. This is demonstrated in Supporting Information Fig. S5 by aligning mean quantitative values obtained for both methods using MS1-based quantification (Fig. S5A) and MS2-based quantification (Fig. S5B) respectively.

These new insights into the performance of MS1 versus MS2 based quantification strategies primed us to investigate the effect of completely missing MS2-data on the analysis of DIA. Several studies have exclusively employed MS1 data for quantification in DIA. This strategy is based on using spectral libraries to extract MS1 peak traces from MS1-only analyses assisted by high-precision RT analysis [28,29]. In a first step we assessed how a further reduction of transitions affects DIA data analysis. Fig. 2A depicts how the overall identification levels in DIA behave upon reducing the number of transitions stepwise to zero. As expected, the complete omission of fragment spectra affects the analysis of HCD-DIA more drastically than the analysis of EThcD-DIA. While the number of identifications drops by more than 20% for HCD-DIA when compared to the analysis with 3 transitions, the drop in EThcD-DIA is much less dramatic (~6.5%), which can be explained by the lesser reliance of EThcD on MS2-data. In both analyses, however, the standard deviation increases, suggesting a loss of robustness compared to MS2-based data analysis.

This surprisingly small effect of MS2-spectra on DIA data analysis prompted us to further investigate this observation. We built a

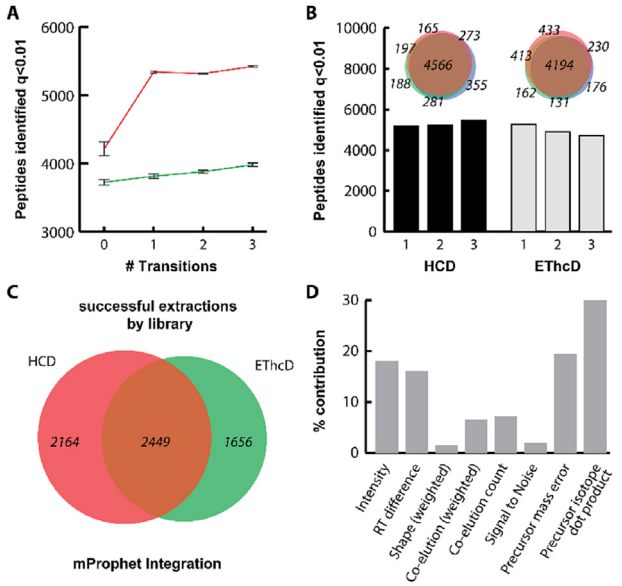

Fig. 2. MS1-based targeted data extraction.
(A) Number of significantly ($q \leq 0.01$) identified phosohopeptides upon stepwise reduction of transition numbers to zero, resulting in exclusively extracting MS1-based data. (B) A combined peptide library comprising DDA runs with HCD and EThcD was used to perform MS1 only targeted data extraction on 6 DIA analyses. This results in the significant ($q \leq 0.01$) identification of 6269 phosphopeptides. Bar plots represent significant identifications per run whilst Venn diagrams show the overlap. (C) Venn diagram grouping the 6269 peptides identified in at least one run to their spectral library of origin. (D) Contribution of multiple mProphet sub-scores to the composite score used for FDR-controlled reintegration of MS1 traces.

combined spectral library comprising identifications from both initial libraries and performed an MS1-only extraction on all 6 DIA runs (HCD and EThcD). A total of 6269 peptides were successfully identified in at least one run, though only 3946 of them in all 6 runs. The number of successful identifications per run ranges from 4732 to 5476 (Fig. 2B). Interestingly, the successful identifications from the three HCD-DIA runs were slightly higher than from the EThcD-DIA runs, which may be caused by the longer cycle time in EThcD or simply because of random fluctuations. For a fair comparison, we therefore performed separate overlap analyses for MS1-extraction for HCD-DIA and EThcD-DIA (Venn diagrams in Fig. 2B). As expected, the overlap of successful identifications between the 3 injection replicas drops compared to MS2-based data analysis, now ranging between 73% (EThcD-DIA) and 75% (HCD-DIA). This is further illustrated in Supporting Information Fig. S6 depicting a detection histogram of peptides across all 6 DIA runs. Success rates of both DIA methods compared to the library size drop to 70% for HCD-DIA and 77% for EThcD-DIA (Fig. 2C). This bias towards EThcD identification might again be caused by the bias towards higher abundant peptides due to the longer duty cycle.

The successful extraction of more than 1600 peptides exclusively from the EThcD DIA data, implies that HCD based DIA analysis alone misses out on a sizeable subset of phosphopeptides. We were thus wondering if these identifications could still be used for targeted data extraction of standard HCD-based DIA analyses. A simple MS1-based targeted data extraction as performed above would be one possible solution. mProphet is capable of relying on MS1 and LC parameters only (Fig. 2D), however, it is unclear whether the lack of MS2 information in the targeted data extraction could compromise confidence in both peptide identification and quantification. We tested this effect by performing a negative control experiment, analyzing a tryptic digest of *E coli* by DIA in triplicate. For the subsequent data analysis, the same mProphet model was used as for the MS1-based analysis of the HeLa phosphopeptides. Fig. 3 depicts the score distributions of target and decoy peptides resulting from this analysis. The data analysis of 3 HeLa DIA samples results in a clearly separated score distribution of target and decoy peptides (Fig. 3A). By including 3 *E coli* DIA measurements to the Hela DIA dataset the mProphet score distribution of the target peptides is greatly affected, increasing the number of target peptides scored in the same range as the decoy peptides. This implies that the MS1 traces from the human phosphopeptide spectral library extracted by the *E coli* DIA sample set, score largely similar to the decoys (Fig. 3B), while maintaining a roughly similar success rate of human derived forward identifications. The analysis of the 3 *E coli* datasets alone results in an almost perfect overlap of the score distribution for target and decoy peptides (Fig. 3C). Using a cutoff of $q \leq 0.01$, mProphet reports only 13–15 significant identification of human phosphopeptides in each *E coli* run. As they are known not be present in the analyzed sample, they can be used as a readout for false discoveries. In comparison to the 5100 peptides identified from the HeLa samples, this corresponds to an FDR of roughly 0.2-0.3%. These analysis together show the strength of mProphet in probabilistic scoring of complex DIA datasets.

This dataset suggests that, MS1-based targeted data extraction appears robust. In itself the approach is also not new *per se* and has been previously used under the name 'accurate mass and time tag' [30]. Also, most of the shotgun data analysis software makes use of linking peptide ID information to MS1 data of different runs (e. g. match-between-runs in MaxQuant [31]). However, these algorithms are usually only reliable within data sets with highly similar LC-MS performance. Other data analysis pipelines such as TIQUAS even expand this further to align DDA and DIA analyses of phosphopeptide samples allowing for the distinction of phosphopeptide isomers whilst being robust to long-term retention time fluctuations [32].

Nonetheless, it has to be thoroughly investigated how much the alignment of MS1 quantification across different experiments can be pushed without losing accuracy. The positive results obtained from the analysis performed here has certainly benefited from the fact that all LC-MS runs were carried out on the same instrument platform within a relatively short time span (3-5 days). In the bigger picture, however, the idea of DIA is to provide the possibility of high throughput. This entails centralized spectral libraries (e. g. online repositories) that can be directly accessed for targeted data extraction of any experiment to avoid the tedious task of generating spectral libraries [33]. Moreover, DIA should enable the assembly of multiple experiments performed over extended time-periods into a single dataset. Thus, the similarities between analyses used for spectral libraries and for the actual quantitative experiment will likely decrease.

3.5. Targeted data extraction of EThcD spectra out of HCD-DIA analyses

An alternative approach to exploit EThcD spectral library data in standard HCD-DIA analyses would be to exclusively extract y- and b-ion series detected in EThcD spectra from HCD-DIA runs. This, however, would require a reasonably high similarity between the relative intensities of the y- and b-ions obtained from EThcD and HCD, which has so far not been investigated in detail. In the following we set out to test this possibility. In a first step, we tested if we could perform targeted data extraction of EThcD library peptides from EThcD-DIA runs by exclusively extracting b- and y-ions. This restriction forced us to reduce the target list by 80 peptides for which the EThcD MS2 spectra did not contain enough b- and y-ions for a successful targeted data extraction (i.e. < 3). For the

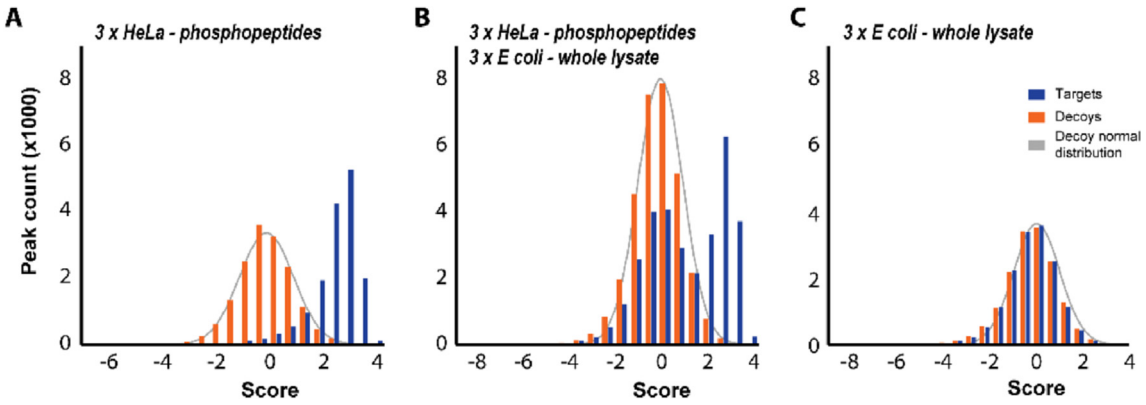

Fig. 3. Negative control for MS1-based targeted data extraction.
mProphet score distribution for the extraction of MS1 ion traces for target and decoy sequences of all HeLa-library phosphopeptides from: (A) HeLa tryptic digests enriched for phosphopeptides analyzed by DIA in triplicates, (B) HeLa tryptic digests enriched for phosphopeptides analyzed by DIA in triplicates and *E. coli* tryptic digest analyzed by DIA in triplicates, (C) *E. coli* tryptic digest analyzed by DIA in triplicates.

remaining peptides, targeted data extraction was performed, including FDR control by mProphet (Fig. 4A). It is noteworthy, that a new mProphet scoring model had to be created as the one trained on extracting both ion series did not reflect the current analysis anymore resulting in a skewed decoy scoring distribution and substantially less significant peptide identifications. Upon retraining, however, the number of significant peptide identifications reaches a similar number as obtained before. Additionally, the overlap of significant peptide identifications between the analysis including c- and z-ions and the analysis relying exclusively on b- and y-ions is very high (>99%). A comparison of the mProphet subscore contributions is provided in Fig. 4B. Not surprisingly, the contribution of the library intensity dot product decreases slightly, whilst the reliance on RT accuracy and precursor isotope dot product increases.

Nonetheless, these results encouraged us to further exploit the use of EThcD b- and y-ion series for targeted data extraction. The possibility of using targeted data extraction to extract peptides identified by EThcD-DDA from HCD-DIA analyses seems intriguing in light of the commonly used data-dependent decision tree methods. To best reflect this situation, we combined the two spectral libraries in a hierarchical manner in which HCD IDs in the libraries are considered preferentially over EThcD IDs (Fig. 5A). By using this setting, the 6 highest abundant b- and y-ions were extracted from all three HCD-DIA runs. The resulting peak integrations were assessed by mProphet and filtered to $q \leq 0.01$. This resulted in additional detection of more than 1500 peptides from the same HCD-DIA files compared to using the HCD-library alone (Fig. 5B). The overlap of successful extractions between the three technical replicas is 90%, similar to the results obtained from using the HCD library alone (Supporting Table 2).

Intrigued by these successful results, we turned our attention to the subset of peptides identified exclusively by EThcD-DDA and not by HCD-DDA. We further reduced this subset by removing peptides that only differ from peptides identified by HCD in terms of phosphosites localization. With this we want to avoid overestimating the success rate due to potential site localization inaccuracies occurring in the DDA database search. This resulted in a list of 1781 phosphopeptides (Supporting Table 3). We analyzed them in terms of reproducible targeted data extraction across the three HCD-DIA injection replicas. A majority of peptides were successfully extracted from either all three replicas or were not identified at all, implying a robust data-extraction process despite the use of different fragmentation techniques (Fig. 5C). Our initial reluctance towards this approach was primarily based on the unknown similarity of the relative intensities for b- and y-ions obtained from EThcD and HCD, which are required for this cross-method comparison. Interestingly, the obtained overall dot products from cross-extracting peptides exceeds the dot products obtained from extracting EThcD spectra from EThcD-DIA analyses. In addition, in

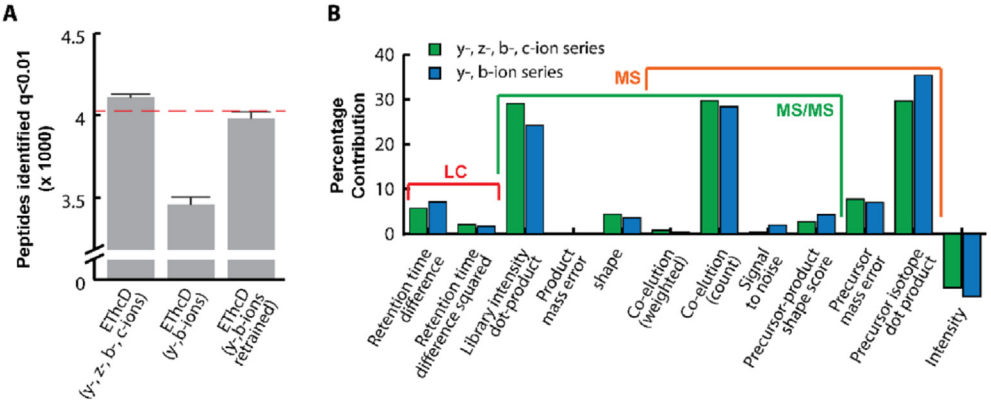

Fig. 4. Targeted data extraction of b- and y-ions in EThcD-DIA.
(A) Number of phosphopeptides extracted from the EThcD-DIA data at a 1% FDR. Bars indicate significant identifications obtained for extracting the 6 most intense b-, y-, c- and z-ions (left) in comparison to extracting the 6 most intense b- and y-ions using the same mProphet integration model (middle) or a newly trained integration model (right). Error bars indicate standard deviation. (B) Contribution of mProphet subscores compared for integration models trained for b-, y-, c- and z-ions (green) and trained for b- and y-ions only (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article).

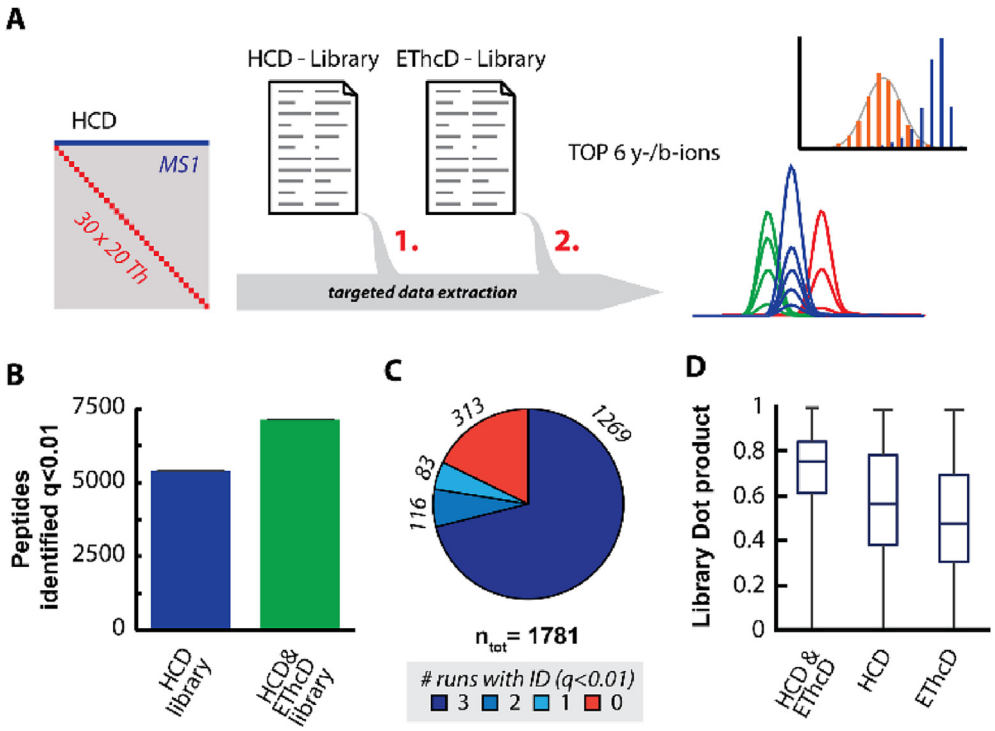

Fig. 5. Targeted data extraction from HCD-DIA using multiple spectral libraries.
(A) Workflow for combined targeted data extraction. Spectral libraries obtained from DDA analyses performed with HCD and EThcD fragmentation were merged and HCD-DIA runs were queried for all library peptides using b- and y-ions only. Hierarchical spectral library handling ensured that EThcD-MS2 spectra were exclusively used when no HCD identification is present. Peak integration by mProphet was used for FDR control. (B) Peptide identifications in HCD-DIA ($q \leq 0.01$) by using HCD-MS2 spectra only compared to using a combined spectral library. Bars indicate average identifications across 3 injection replicas, error bars represent standard deviations. (C) Reproducibility analysis of targeted data extraction for peptides exclusively identified in the EThcD spectral library across the three HCD-DIA injection replicas. (D) Distribution of library intensity dot products for the combined spectral library analysis in comparison to the individual analyses for HCD-DIA and EThcD-DIA.

terms of quantitative reproducibility the peptides detected exclusively in EThcD-DDA show less variation when extracted from HCD-DIA data compared to EThcD-DIA. This is illustrated in Supporting Fig. S7 depicting overall CV distributions of summed fragment intensities obtained across 3 injection replicas per acquisition method. This somewhat surprising observation is likely the result of the MS2 data acquired during EThcD-DIA analysis and can originate from multiple sources. It is possible that the reduced similarity during targeted data extraction is caused by a generally lower MS2 intensity (due to remaining precursor ions), the occurrence of c- and z-ion in addition to b- and y-ions resulting in an increased number of low intensity fragments and/or generally unfavorable ion injection conditions due to constrains in terms of cycle time. This would cause signals to be closer to noise levels and thus increasingly affected by random interferences, resulting in a generally lower spectrum similarity.

3.6. Considerations about assay sensitivity

To further test this hypothesis, we investigated the intensities obtained from EThcD-DIA with HCD-DIA. Indeed, the intensities of the extracted fragment ion traces in an EThcD-DIA method are lower than in HCD. This is likely due to the AGC controlled fill time, keeping total ion numbers injected into the Orbitrap similar. Due to the fact that EThcD gives rise to c- and z-ions in addition to the b- and y-ions common to HCD, the intensity of each ion decreases. This effect is exemplified for the MNTB peptide *DISGP-(pho)S-PSKK* in Fig. 6A. A global analysis of integrated peak areas for precursor and fragment ions can be found in Supporting Information Fig. S8. The same trend of lower MS2 intensity could also be observed upon comparing targeted data extraction using the EThcD spectra library in combination with both HCD-DIA and EThcD-DIA files. Fig. 6B shows DIA traces of a representative peptide identified in DDA by EThcD only but extracted with a higher intensity using the HCD-DIA analysis compared to EThcD-DIA. However, as sensitivity is generally defined by signal to noise ratios and quantitative reproducibility close to the lower limit of detection, further experiments including dilution series of synthetic peptides spiked into alike background samples would be required for a more thorough assessment. Additionally, sensitivity might be highly dependent on peptide sequence, where certain physical-chemical peptides properties favor one fragmentation technique over the other.

3.7. Implications of Orbitrap-based DIA analyses and future directions

By comparing EThcD- and HCD-based DIA it is evident that in a classical DIA method using subsequent isolation windows of the same size, the need for speed becomes predominant. EThcD is known for long reaction times, which increases the cycle time drastically. Additional factors have to be considered when performing the experiments on Orbitrap based instruments. Trap-based instruments are intrinsically slow compared to quadrupole-based fragmentation. Specifically, the requirement of high resolution/accurate mass MS2 data results in a longer cycle time, as each MS2 ion bundle has to be analyzed in the Orbitrap. It is striking that the use of 30 acquisition windows on an Orbitrap Fusion results in a cycle time of 3 s while similar analyses performed on a QTOF instrument can make use of 64 isolation windows in combination with only a slightly longer cycle time of 3.4 s [11]. Reducing these cycle times to a chromatographic time scale, however, entails the use of wider isolation windows, which results in reduced

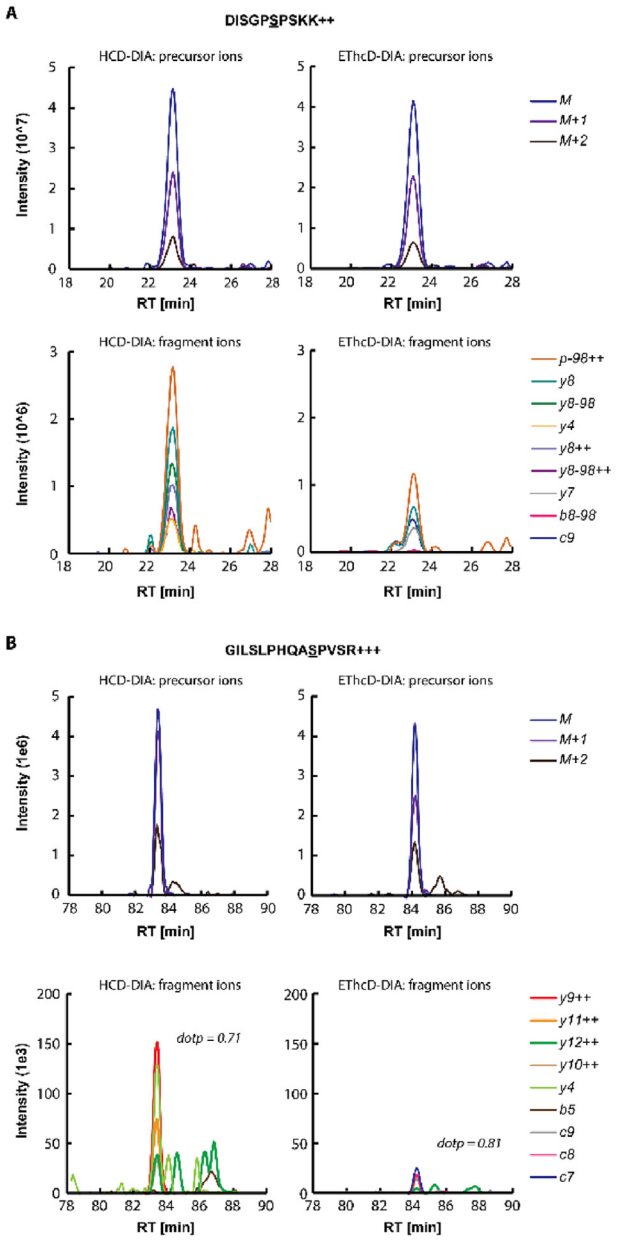

Fig. 6. Comparison of targeted data extraction from HCD-DIA and EThcD-DIA. (A) Targeted data extraction for MNTB peptide DISGP-(pho)S-PSKK from HCD-DIA (left) and EThcD-DIA (right) analyses, based on HCD- and EThcD spectral libraries respectively. Similar XIC intensity for precursor ions (upper panels) were observed, however MS2 XICs intensities differed largely between the HCD-DIA and EThcD-DIA (lower panel). (B) Targeted data extraction of the PIAS1 peptide GILSLPHQA-(pho)S-PVSR from HCD-DIA (left) and EThcD-DIA (right) respectively, using an EThcD-MS2 spectral library. Upper panels show MS1 traces demonstrating equal abundance in both samples. Lower panels show MS2 data extraction using the six most abundant b- and y-ions for HCD and the six most abundant ions overall for EThcD.

specificity. Other workarounds could be to subject the same sample to multiple LC-DIA-MS runs, each analyzing a different, small, MS1 range. This would retain the method specificity by enabling the use of narrow isolation windows, but also substantially increase the total analysis time and sample amount required. An alternative could be the use of MSX methods [13], which are capable of multiplexing several acquisition windows into one single MS2 spectrum and de-multiplex them again at the data analysis level. This should enable the specificity of relatively narrow isolation windows in combination with reasonably short cycle times. Via this approach, the lack of speed during the fragmentation reaction should not affect the interplay of cycle time and specificity. However, it remains to be investigated if high multiplexing affects the sensitivity of the measurement. Essentially, the fill time and the AGC target control the number of ions entering the Orbitrap. In a highly multiplexed method, the composition of each ion bundle will be highly complex and could thus drastically reduce the signal intensity of each individual ion. Increasing the AGC-target might circumvent this effect, but might result in space charge effects, thus compromising the mass accuracy of the acquired MS2-scan [34]. The tight control of the Orbitrap fill time likely also causes the loss of sensitivity for EThcD-DIA compared to HCD-DIA. The fact that each precursor ion gives rise to a z-, and c-ion series in addition to the y- and b-ion series additionally results in a lower intensity of each individual ion originating from the same amount of precursor ions.

Most of these shortcomings are not necessarily unique to DIA analysis, as they can affect DDA alike. Nonetheless, especially EThcD is widely used in shotgun for certain sample types, because it outperforms HCD- or CID-fragmentation. Examples are the analysis of endogenous peptides [14], HLA peptides [35,36] or glycopeptides [15]. It is not inconceivable that the analysis of these samples by DIA is going to be implemented for the quantitative analysis of large cohorts of samples, for instance occurring in biomarker studies or even for diagnostic purposes [7]. Instead of EThcD-DIA, however, it seems more likely to resort to HCD-DIA or MS1-only experiments exploiting EThcD spectral libraries for data analysis.

In line with other studies [28,29] we could demonstrate the robustness of the MS1-based approach within a limited time-span on a single analysis platform. The implementation of this approach over a longer time period or across different instrument platforms and/or laboratories needs to be further investigated. Alternatively, the use of HCD-DIA in combination with decision-tree based acquisition methods for spectral library generation might start to play a bigger role for the reproducible analysis of large numbers of non-standard samples.

4. Conclusions

Altogether this study demonstrates the versatility DIA methods can provide. Through careful optimization of parameters such as peptide fragmentation and library generation, the success rate of DIA experiments can be greatly improved in a sample specific manner. Specifically, on Orbitrap based instruments the balance between speed and sensitivity needs to be carefully tailored and will eventually require rational choices based on factors such as sample type, sample complexity, need for sensitivity and throughput. All of these factors eventually also influence data analysis, as programs such as mProphet can be dynamically adapted. This effect can definitely be exploited as shown here for the analysis of DIA runs with less informative MS2 data. High resolution MS1 spectra in combination with high quality spectral libraries alternatively give increasing credence to the use of MS1 quantification across increasing numbers of samples. Specifically, the suggested combination of decision-tree based library generation and HCD-only DIA analysis seems to open new doors for the robust quantification of non-tryptic peptide samples. Main challenges, however, remain the control of phosphosite localization and phosphoisomers, specifically in the context of MS1 quantification. While MS2-based quantification in EThcD-DIA might bring the potential for a better distinction of phosphoisomers [16], MS1-based quantification will not improve the issue over the current state-of-the-art in shotgun MS. The most crucial aspect in this respect will be the availability of robust software, specifically

tackling this issue in DIA data, such as IPF [37] or TIQUAS [32]. Further investigation needs to focus on quantitative effects of cycle time, as our analysis is currently limited to constant isolation windows. While our data certainly suggests that the lower number of measurement points across the chromatographic peak in EThcD-DIA compared to HCD-DIA results in lower quantitative reproducibility in MS1-quantification, the exact magnitude of this effect has not yet been systematically investigated. It is very likely that this is an effect of the lower number of measurement points across the chromatographic peak, but it seems unclear if this effect only causes the quantitative values to be less reliable or if it also affects the process of peak picking by mProphet, which relies on subscores such as peak shape and coelution between different ions. Therefore, it is also unclear if a reduced cycle time would be able to solve the issue, as peak picking effects caused by EThcD specific characteristics might still be present upon reducing isolation width. Furthermore, reduced cycle time automatically comes with an increased fragment spectra complexity, which will further decrease the method's specificity.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge support from the Netherlands Organization for Scientific Research (NWO) through a VIDI grant for M.A. (Grant 723.012.102) and funding Proteins@Work (project 184.032.201) and X-Omics (project 184.034.019), embedded in The Netherlands Proteomics Centre, financed as part of the National Roadmap Large-Scale Research Facilities of The Netherlands.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijms.2020.116336.

References

[1] C. Ludwig, L. Gillet, G. Rosenberger, S. Amon, et al., Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial, Mol. Syst. Biol. 14 (2018), e8126.
[2] J.D. Venable, M.Q. Dong, J. Wohlschlegel, A. Dillin, J.R. Yates, Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra, Nat. Methods 1 (2004) 39–45.
[3] R.S. Plumb, K.A. Johnson, P. Rainville, J.P. Shockcor, et al., The detection of phenotypic differences in the metabolic plasma profile of three strains of Zucker rats at 20 weeks of age using ultra-performance liquid chromatography/orthogonal acceleration time-of-flight mass spectrometry, Rapid Commun. Mass Spectrom. 20 (2006) 2800–2806.
[4] S.J. Geromanos, J.P. Vissers, J.C. Silva, C.A. Dorschel, et al., The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS, Proteomics 9 (2009) 1683–1695.
[5] L.C. Gillet, P. Navarro, S. Tate, H. Röst, 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 (2012), 016717. O111.
[6] T. Guo, P. Kouvonen, C.C. Koh, L.C. Gillet, et al., Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps, Nat. Med. 21 (2015) 407–413.
[7] S.I. Anjo, C. Santa, B. Manadas, SWATH-MS as a tool for biomarker discovery: from basic research to clinical applications, Proteomics 17 (2017).
[8] Y. Liu, A. Buil, B.C. Collins, L.C. Gillet, et al., Quantitative variability of 342 plasma proteins in a human twin population, Mol. Syst. Biol. 11 (2015) 786.
[9] Y. Liu, R. Hüttenhain, S. Surinova, L.C. Gillet, et al., Quantitative measurements of N-linked glycoproteins in human plasma by SWATH-MS, Proteomics 13 (2013) 1247–1256.
[10] A.M. Zawadzka, B. Schilling, J.M. Held, A.K. Sahu, et al., Variation and quantification among a target set of phosphopeptides in human plasma by multiple reaction monitoring and SWATH-MS2 data-independent acquisition, Electrophoresis 35 (2014) 3487–3497.
[11] T. Schmidlin, L. Garrigues, C.S. Lane, T.C. Mulder, et al., Assessment of SRM, MRM(3) , and DIA for the targeted analysis of phosphorylation dynamics in non-small cell lung cancer, Proteomics 16 (2016) 2193–2205.
[12] Y. Zhang, A. Bilbao, T. Bruderer, J. Luban, et al., The use of variable Q1 isolation windows improves selectivity in LC-SWATH-MS acquisition, J. Proteome Res. 14 (2015) 4359–4371.
[13] J.D. Egertson, A. Kuehn, G.E. Merrihew, N.W. Bateman, et al., Multiplexed MS/MS for improved data-independent acquisition, Nat. Methods 10 (2013) 744–746.
[14] B.L. Parker, J.G. Burchfield, D. Clayton, T.A. Geddes, et al., Multiplexed temporal quantification of the exercise-regulated plasma peptidome, Mol. Cell. Proteomics 16 (2) (2017) 2055–2068.
[15] Q. Yu, B. Wang, Z. Chen, G. Urabe, et al., Electron-Transfer/higher-energy collision dissociation (EThcD)-Enabled intact glycopeptide/glycoproteome characterization, J. Am. Soc. Mass Spectrom. 28 (9) (2017) 1751–1764.
[16] C.K. Frese, H. Zhou, T. Taus, A.F. Altelaar, et al., Unambiguous phosphosite localization using electron-transfer/higher-energy collision dissociation (EThcD), J. Proteome Res. 12 (2013) 1520–1525.
[17] T.S. Batth, C. Francavilla, J.V. Olsen, Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics, J. Proteome Res. 13 (2014) 6176–6186.
[18] H. Post, R. Penning, M.A. Fitzpatrick, L.B. Garrigues, et al., Robust, sensitive, and automated phosphopeptide enrichment optimized for low sample amounts applied to primary hippocampal neurons, J. Proteome Res. 16 (2017) 728–737.
[19] A. Cristobal, M.L. Hennrich, P. Giansanti, S.S. Goerdayal, et al., In-house construction of a UHPLC system enabling the identification of over 4000 protein groups in a single analysis, Analyst 137 (2012) 3541–3548.
[20] L. Käll, J.D. Canterbury, J. Weston, W.S. Noble, M.J. MacCoss, Semi-supervised learning for peptide identification from shotgun proteomics datasets, Nat. Methods 4 (2007) 923–925.
[21] T. Taus, T. Köcher, P. Pichler, C. Paschke, et al., Universal and confident phosphorylation site localization using phosphoRS, J. Proteome Res. 10 (2011) 5354–5362.
[22] B. MacLean, D.M. Tomazela, N. Shulman, M. Chambers, et al., Skyline: an open source document editor for creating and analyzing targeted proteomics experiments, Bioinformatics 26 (2010) 966–968.
[23] J. Zi, S. Zhang, R. Zhou, B. Zhou, et al., Expansion of the ion library for mining SWATH-MS data through fractionation proteomics, Anal. Chem. 86 (2014) 7242–7246.
[24] S. Shao, T. Guo, C.C. Koh, S. Gillessen, et al., Minimal sample requirement for highly multiplexed protein quantification in cell lines and tissues by PCT-SWATH mass spectrometry, Proteomics 15 (21) (2015) 3711–3721.
[25] V. Lange, P. Picotti, B. Domon, R. Aebersold, Selected reaction monitoring for quantitative proteomics: a tutorial, Mol. Syst. Biol. 4 (2008) 222.
[26] L. Reiter, O. Rinner, P. Picotti, R. Hüttenhain, et al., mProphet: automated data processing and statistical validation for large-scale SRM experiments, Nat. Methods 8 (2011) 430–435.
[27] R. Bruderer, O.M. Bernhardt, T. Gandhi, Y. Xuan, et al., Optimization of experimental parameters in data-independent mass spectrometry significantly increases depth and reproducibility of results, Mol. Cell. Proteomics 16 (2017) 2296–2309.
[28] R. Bruderer, O.M. Bernhardt, T. Gandhi, L. Reiter, High-precision iRT prediction in the targeted analysis of data-independent acquisition and its impact on identification and quantitation, Proteomics 16 (2016) 2246–2256.
[29] T.P. Conrads, G.A. Anderson, T.D. Veenstra, L. Pasa-Tolić, R.D. Smith, Utility of accurate mass tags for proteome-wide protein identification, Anal. Chem. 72 (2000) 3349–3354.
[30] R.D. Smith, G.A. Anderson, M.S. Lipton, L. Pasa-Tolic, et al., An accurate mass tag strategy for quantitative and high-throughput proteome measurements, Proteomics 2 (2002) 513–523.
[31] C. Bielow, G. Mastrobuoni, S. Kempa, Proteomics quality control: quality control software for MaxQuant results, J. Proteome Res. 15 (2016) 777–787.
[32] P.R. Cutillas, Targeted in-depth quantification of signaling using label-free mass spectrometry, Methods Enzymol. 585 (2017) 245–268.
[33] G. Rosenberger, C.C. Koh, T. Guo, H.L. Röst, et al., A repository of assays to quantify 10,000 human proteins by SWATH-MS, Sci Data 1 (2014), 140031.
[34] A. Kharchenko, G. Vladimirov, R.M. Heeren, E.N. Nikolaev, Performance of Orbitrap mass analyzer at various space charge and non-ideal field conditions: simulation approach, J. Am. Soc. Mass Spectrom. 23 (2012) 977–987.
[35] G.P. Mommen, C.K. Frese, H.D. Meiring, J. van Gaans-van den Brink, et al., Expanding the detectable HLA peptide repertoire using electron-transfer/higher-energy collision dissociation (EThcD), Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 4507–4512.
[36] F. Marino, G.P. Mommen, A. Jeko, H.D. Meiring, et al., Arginine (Di)methylated human leukocyte antigen class I peptides are favorably presented by HLA-B*07, J. Proteome Res. 16 (2017) 34–44.
[37] G. Rosenberger, Y. Liu, H.L. Röst, C. Ludwig, et al., Inference and quantification of peptidoforms in large sample cohorts by SWATH-MS, Nat. Biotechnol. 35 (2017) 781–788.