

# Comparative Analyses of Data Independent Acquisition Mass Spectrometric Approaches: DIA, WiSIM-DIA, and Untargeted DIA

Frank Koopmans,\* Jenny T. C. Ho, August B. Smit, and Ka Wan Li

Data-independent acquisition (DIA) is an emerging technology for quantitative proteomics. Current DIA focusses on the identification and quantitation of fragment ions that are generated from multiple peptides contained in the same selection window of several to tens of  $m/z$ . An alternative approach is WiSIM-DIA, which combines conventional DIA with wide-SIM (wide selected-ion monitoring) windows to partition the precursor  $m/z$  space to produce high-quality precursor ion chromatograms. However, WiSIM-DIA has been underexplored; it remains unclear if it is a viable alternative to DIA. We demonstrate that WiSIM-DIA quantified more than 24 000 unique peptides over five orders of magnitude in a single 2 h analysis of a neuronal synapse-enriched fraction, compared to 31 000 in DIA. There is a strong correlation between abundance values of peptides quantified in both the DIA and WiSIM-DIA datasets. Interestingly, the S/N ratio of these peptides is not correlated. We further show that peptide identification directly from DIA spectra identified >2000 proteins, which included unique peptides not found in spectral libraries generated by DDA.

LC-MS/MS-based quantitative proteomics is the method of choice to measure changes in global protein levels in biological samples. In the past decade, data-dependent acquisition (DDA) has been widely used for this. In DDA, the precursors, usually the top 10–20 peptides per cycle, are sequentially selected from a full mass MS1 scan for fragmentation and acquisition in the


MS/MS mode.<sup>[1,2]</sup> Recently, DDA has been optimized to reveal the comprehensive proteome of a single cell type.<sup>[3]</sup> However, the stochastic precursor selection of DDA leads to inconsistent detection of peptides. In particular, the undersampling of medium to low abundant peptides causes high variation across replicates due to the selection of different subsets of peptides. This results in missing peptide identification, which can be substantial among replicates (>30%) and reduces the number of quantifiable proteins.<sup>[4,5]</sup>

Data-independent acquisition (DIA, also known as SWATH<sup>[6]</sup>) is a recent development in quantitative proteomics. It is mainly performed on the high-resolution high mass accuracy mass spectrometers and has been shown to be superior to DDA<sup>[7]</sup> by producing a higher number of quantified proteins in shorter analysis time, fewer missing

values, and lower coefficients of variation (CoV) across replicates. In DIA, all peptides within a predefined wide selection window, which in the original DIA study spanned a 25  $m/z$  range,<sup>[6]</sup> are simultaneously fragmented. The acquisition is repeated sequentially in stepped selection windows, usually in the 400–1000  $m/z$  range. Generally, the high number of fragments ions generated from multiple peptides contained in the same selection window complicates the analysis in a classical database search strategy. This problem is circumvented by the use of a reference spectral library, which is generated beforehand by an extensive analysis of the same/similar samples by DDA. The information of the elution time of the peptide and its fragment ions stored in the spectral library defines the identity of the peptide measured in a DIA experiment.<sup>[8–11]</sup> Thus, samples not present in a spectral library in principle cannot be analyzed. To circumvent this shortcoming, algorithms have been developed that create a pseudo-DDA dataset from the DIA data (untargeted peptide identification or untargeted DIA<sup>[12,13]</sup>) for subsequent search in way similar to the classical DDA strategy.

An alternative to DIA is a wide selected-ion monitoring, DIA (WiSIM-DIA), which is grossly underexplored. While both DIA and WiSIM-DIA require a spectral library for peptide/protein identification, in contrast to MS2-based DIA method, WiSIM-DIA uses MS1 for quantitation. Previous reports on

F. Koopmans, Prof. Dr. A. B. Smit, Dr. K. W. Li  
Department of Molecular and Cellular Neurobiology  
CNCR  
Amsterdam Neuroscience  
Vrije Universiteit  
Amsterdam, The Netherlands  
E-mail: frank.koopmans@vu.nl  
Dr. J. T. C. Ho  
Thermo Fisher Scientific  
Hemel Hempstead, UK

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/pmic.201700304>

© 2017 The Authors. *Proteomics* Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/pmic.201700304

## Significance of the study

In recent years data-independent acquisition (DIA), the fragment ion-centric approach, is becoming the method of choice for label-free quantification studies. Here, we demonstrated that the precursor ion-centric WiSIM-DIA approach is capable of quantifying >24 000 unique peptides from a neuronal synapse-enriched sample in 2 h analysis time, compared to 31 000 in DIA. This puts WiSIM-DIA as a viable alternative to DIA, especially when interferences of large numbers of fragment ions derived from coeluting peptides is an issue. While data analyses of DIA and WiSIM-DIA generally require a spectral library, recent development of untargeted DIA allows direct interrogation of raw data from a regular DIA experiment. Using this approach, we identified about 2000 proteins, from which 2000 peptides are not represented in the DDA spectral library. Therefore, the output from untargeted DIA can be added to existing spectral libraries to increase peptide identification.

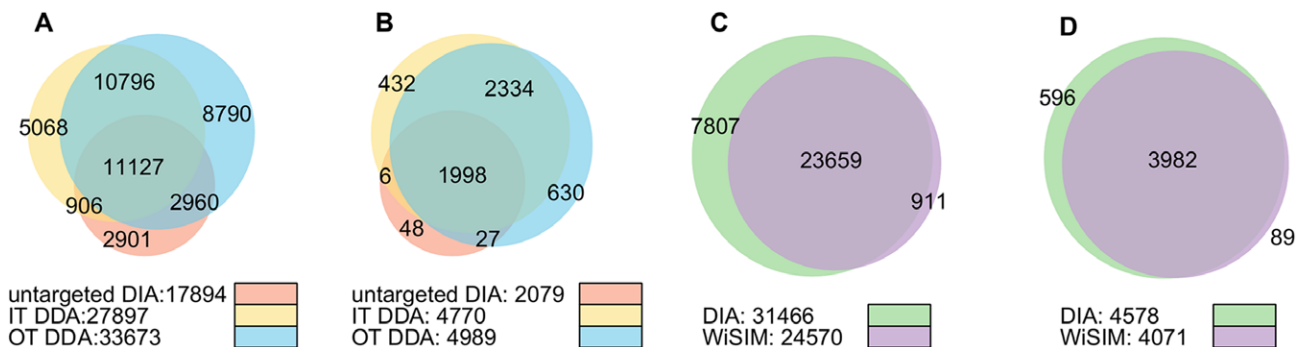
WiSIM-DIA were performed in an Orbitrap (OT) Fusion mass spectrometer.<sup>[14,15]</sup> This method consists of three-stepped SIM scans acquired with 240 000 resolution over a 200  $m/z$  range that covers 400–1000  $m/z$ . In parallel with each SIM scan, peptide fragmentation from selection windows of 12  $m/z$  were acquired in the ion trap (IT), with acquisition repeated with 17 sequential IT MS/MS windows. In comparison, DIA used the OT for high (60 000) resolution MS1 and 17 sequential MS/MS windows in lower (15 000) resolution. So, the quality of MS1 acquisition in WiSIM-DIA was improved compared to DIA by using stepped SIM scans and a higher resolution, while the quality of MS/MS acquisition was favorable for DIA due to the use of the OT (compared to WiSIM-DIA using IT for MS/MS). MS/MS data acquired in the low-resolution IT were used for identification, whereas quantitation was based on the extracted ion chromatogram of the SIM data with a 5 ppm window. Here, the spectral library could be generated with classical DDA where the MS1 full scan is acquired in the high-resolution OT, and the fragment ions in the fast but low-resolution IT. It is proposed that WiSIM-DIA does not suffer from the drawback of DIA, for example, the potential interferences of the large number of fragment ions derived from coeluting peptides. However, the only application published recently reported the quantitation of about 1100 proteins by WiSIM-DIA<sup>[15]</sup>, which seems to be on the lower side acquired by a modern MS. Thus, it has remained unclear whether WiSIM-DIA is a viable alternative to DIA.

In this study, we used an OT Fusion Lumos in DDA mode to generate two spectral libraries from the mouse synaptosome, a preparation enriched for proteins of the neuronal synapse,<sup>[16]</sup> which constitutes the building block of the brain. The tryptic digest of 10  $\mu\text{g}$  synaptosome proteins were fractionated offline using high pH reversed phase cartridges into eight fractions. Each fraction was subjected to DDA by two separate acquisition strategies: (1) MS1 OT with the fast but low-resolution IT for MS/MS (HCD-IT) and, (2) MS1 OT with the high-resolution OT MS/MS (HCD-OT). The data were processed using MaxQuant<sup>[17]</sup> with 1% False Discovery Rate (FDR) at both peptide and protein level.

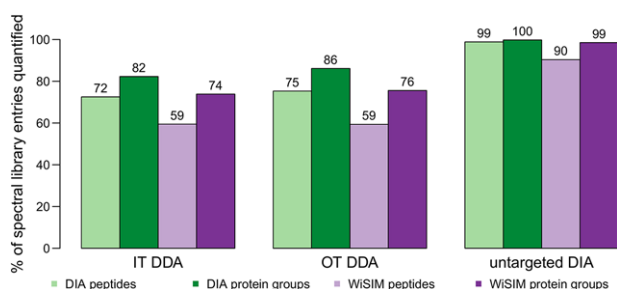
From the same sample, we used 1  $\mu\text{g}$  for DIA with a 2 h LC gradient. Three replicates each for DIA and WiSIM-DIA were performed. Technically, several parameters can be considered to maximize the DIA output. While a cycle scan time is usually fixed around 3–4 s to obtain six to ten measurement points of a peptide that is needed for quantitation, the width of a selection-window, the accumulation time per selection-window, and the whole  $m/z$  range can be varied. The original study opted for a 25  $m/z$  selection window,<sup>[6]</sup> which may cause peptide fragment ion interferences due to their high complexity. In another extreme, a narrow selection window of 3  $m/z$  has been proposed as preference for more comprehensive and in-depth view of protein profiling in a complex sample.<sup>[18]</sup> This is compromised by a shorter acquisition time with potentially reduced sensitivity. Considering the mild protein complexity of the synaptosome fraction of about 5000 proteins contained in the spectral library, we chose the 12  $m/z$  selection window for both DIA and WiSIM-DIA (see also<sup>[19]</sup>). The total mass range covered was 400–800  $m/z$ , which includes the majority of the peptides (Supporting Information, Figure S1).

In addition to the classical DDA-based spectral library, we generated a spectral library from the DIA data using the recently launched Spectronaut Pulsar software (untargeted DIA at 1% peptide and protein FDR, settings analogous to the MaxQuant DDA analysis), which yielded 17 894 unique peptide sequences in 2079 protein groups. This is less than the 27 897 and 33 673 unique peptide sequences and 4770 and 4989 protein groups represented in the IT and OT spectral libraries, respectively, within the 400–800  $m/z$  range (Figure 1A–B, the total number of identified peptides without any  $m/z$  filters in each spectral library is shown in Supporting Information, Figures S1 and S2). Here, we compared the subset of peptides in the 400–800  $m/z$  range to match the DDA spectral library with the acquisition settings for DIA and WiSIM-DIA on the OT Fusion Lumos. The samples used for (untargeted) DIA were not fractionated, in contrast to the extensively fractionated samples used exclusively for DDA spectral library construction, which may account for the overall reduced number of identifications by untargeted DIA. Despite the lack of extensive fractionation, untargeted DIA contributed 2901 unique peptides to the spectral library. Interestingly, most of the peptides exclusively identified by untargeted DIA belong to protein groups that were also identified in the IT or OT libraries (or both). This suggests that the untargeted DIA unique peptides may have been lost in the first dimensional high pH reversed phase HPLC separation used for the OT and IT analyses. Alternatively, respective peptide MS/MS spectra quality could be subpar, be part of mixed chimeric spectra in shotgun MS/MS or it may be due to the nature of stochastic precursor selection of DDA; while these peptides were present, they might not have been selected for MS/MS or the MS/MS could have been triggered far from the peak apex (pseudo MS/MS in untargeted DIA is generated at the apex of the peak). Either way, this argues that DIA data that are generated from routine quantitative analysis might subsequently be added to spectral libraries generated by conventional DDA to increase protein coverage.

The performance of both DIA and WiSIM-DIA was excellent; 31 466 and 24 570 unique peptide sequences contained in the merged spectral library were quantified at 1% peptide-level FDR, respectively, with extensive overlap (Figure 1 C and D). These



**Figure 1.** For the spectral library IT DDA, OT DDA, and untargeted DIA acquisition resulted in A) 27 897, 33 673, and 17 894 unique peptide sequences within the 400–800  $m/z$  range and B) 4770, 4989, and 2079 protein groups, respectively. From the merged spectral library, which contains all peptides identified by either IT, OT, or untargeted DIA; C) 31 466 and 24 570 unique peptide sequences; and D) 4578 and 4071 protein groups were quantified by DIA and WiSIM-DIA, respectively.



**Figure 2.** Fraction of peptide sequences and protein groups from individual spectral libraries quantified by DIA and WiSIM-DIA. DIA quantifies on average 13% more peptides from each spectral library compared to WiSIM-DIA. Although the number of peptides contributed to the merged spectral library is relatively low for Pulsar (Figure 1A and B), their recovery after quantification is remarkably high (note: most likely the peptides with best MS2 abundance profiles are selected for identification). Analogous figures with count data instead of fractions in Supporting Information, Figure S3.

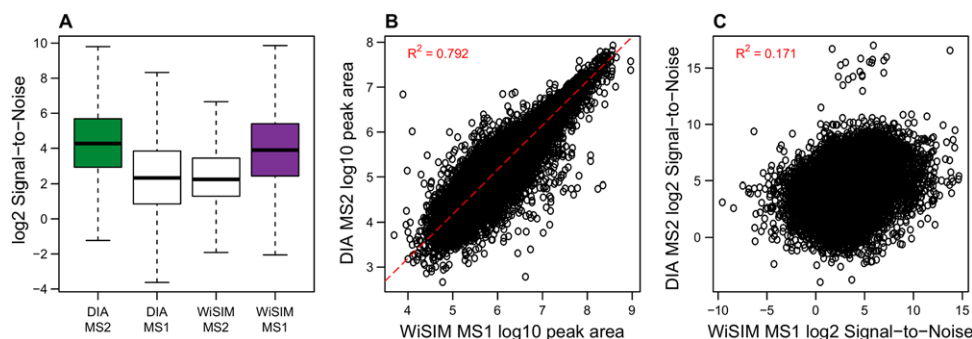
peptides map to >4000 protein groups in the merged spectral library (no protein-level FDR was applied).

From the two spectral libraries generated using MS2 HCD-OT or MS2 CID-IT, we observed a slightly higher number of peptides and protein groups identified by OT (Figure 1), and a slightly higher coverage of the MS2 HCD-OT library following DIA quantification (Figure 2). Thus, despite the slower scan rate of OT its high resolution and mass accuracy favorably affects the population of identified peptides that can be recovered in the DIA data analysis. It may also underlie the fact that both employed the same MS sector, the OT, for the measurement. The coverage of protein groups from the untargeted DIA spectral library is 100% (Figure 2), which reflects the fact that it is generated from the original DIA data. However, using the WiSIM-DIA data, we also quantified 90 and 99% of all peptides and protein groups, respectively, from the untargeted DIA spectral library, which suggests the untargeted DIA approach tends to prioritize peptides that exhibit a clean elution profile with high S/N ratio. The spectral library coverage by WiSIM-DIA is generally lower than that of DIA, which may underlie at least in part that current DIA algorithms are primarily using MS2 fragment intensities to identify spectral library peptides. Algorithm improvements that lead

to better utilization of high-quality MS1 signals with sub-ppm mass error would improve the recovery rate of spectral library peptides by WiSIM-DIA. This approach would be an extension of the previous described “accurate mass tag” strategy in which the identities of the peptides based on the LC-FTICR MS1 measurement were validated by LC-MS/MS analysis on a conventional IT mass spectrometer.<sup>[20,21]</sup> A similar approach has been applied to the analysis of phosphorylated human peptides,<sup>[22]</sup> and HeLa cell proteome.<sup>[23]</sup>

The median coefficient of variation (CoV) for 25 788 peptides quantified in both the DIA and WiSIM-DIA datasets was 9% within three WiSIM-DIA technical replicates and 7% within three DIA technical replicates (Supporting Information, Figure S4). Evaluating all peptides quantified by DIA (35 123) and WiSIM-DIA (26 899) resulted in 8 and 9% median CoV, respectively. For both comparisons, Student’s *t* tests reveal the differences between DIA and WiSIM-DIA CoV was statistically significant ( $p$ -value <  $10^{-16}$ ) albeit with much higher effect size for the former (Cohen’s *d*: 0.29) compared to the latter (Cohen’s *d*: 0.09). Correlation of the abundance values between technical replicates yielded a 0.94  $R^2$  and 0.93  $R^2$  on average for DIA and WiSIM-DIA, respectively. The slightly reduced technical variation of DIA over WiSIM-DIA will likely result in higher sensitivity when performing differential abundance analysis in real-world biological applications.

The S/N ratio is a good indicator of the mass spectrometric measurement quality. In DIA mode, the S/N ratio for fragment ion intensities per precursor is better than those of the corresponding precursor measured in MS1 (Figure 3A). On the other hand, WiSIM-DIA yields a better S/N ratio for the precursor ion measured in MS1 compared to its MS2 S/N ratio. These findings are in accordance to the experimental design that DIA is optimized for MS/MS analysis, and WiSIM-DIA for MS1 measurement. Student’s *t* tests applied to the log-transformed WiSIM-DIA MS1 and MS2 S/N distributions confirmed statistical significance of this comparison ( $p$ -value <  $10^{-16}$ ) with a medium-large effect size (Cohen’s *d*: 0.67). Analogously, the overall DIA MS2 S/N distribution was significantly lower ( $p$ -value <  $10^{-16}$ ) than its WiSIM-DIA MS1 counterpart, but with a relatively small effect size (Cohen’s *d*: 0.34). In addition to comparing the distributions shown in Figure 3A, Student’s *t* tests on log2 DIA MS2 and WiSIM-DIA MS1 S/N values for all



**Figure 3.** Quality of precursor quantification compared between DIA (MS2) and WiSIM-DIA (MS1). A) The S/N ratio is high for both DIA MS2 and WiSIM-DIA MS1. Analogously, the secondary mode of quantification (MS1 for DIA, MS2 for WiSIM-DIA) is more noisy. B) The abundance values of 25 778 precursors quantified by both DIA and WiSIM-DIA are correlated (0.792  $R^2$ ), as expected. The dashed red line shows linear regression. C) Interestingly, the S/N for these peptides is not correlated (0.171  $R^2$ ). A subpopulation of precursors is quantified with higher signal quality by DIA MS2 than WiSIM-DIA MS1, and vice versa.

individual peptides using the triplicate DIA and WiSIM-DIA measurements resulted in 4155 (out of 25 780) significantly different peptides at FDR-adjusted  $p$ -value  $\leq 0.01$ . Of these, 1010 and 3145 peptides showed improved S/N in WiSIM-DIA MS1 and DIA MS2, respectively. We found a strong correlation (0.792  $R^2$ ) between the abundance values of peptides quantified in both the DIA and WiSIM-DIA datasets, as expected (Figure 3B). Interestingly, the S/N ratio of these peptides was not correlated (0.171  $R^2$ ) between DIA and WiSIM-DIA (as compared to 0.80 and 0.59 average  $R^2$  between DIA and WiSIM-DIA replicates, Supporting Information, Figure S5). The lack of correlation in S/N might be explained by the stepped SIM scans in WiSIM-DIA that clean up the spectra of many peptides, but might reduce the signal for already low abundant peptides (Figure 3C). The use of different modes of quantification for WiSIM-DIA (MS1) and DIA (MS2) taken together with their overall similar S/N distributions shown in Figure 3A could give rise to subpopulations of peptides that are quantified with higher signal quality in either WiSIM-DIA or DIA, indicating mutually exclusive benefits. Alternatively, observed differences in peptide subsets could arise by chance. Future research could further investigate this hypothesis using extensive datasets that allow for cross-validation of peptides with stark differences in S/N between WiSIM-DIA and DIA.

We conclude that DIA and WiSIM-DIA can quantify more than 31 000 and 24 000 unique peptides (at 1% peptide-level FDR), respectively, over five orders of magnitude in a single 2 h analysis with nearly no missing values (0.08 and 0.004% missing peptide values between three technical replicates, respectively). The number of peptides from the spectral libraries recovered by WiSIM-DIA will be improved when its high-quality MS1 signal is better taken advantage of by future improvements of analysis software (e.g., by relying on accurate retention time and low precursor mass error for matching precursor peaks to the library in absence of high-quality fragment spectra<sup>[20,21]</sup>). The untargeted DIA spectral library generated from the triplicate 2 h DIA analysis yields nearly 50% of the peptides/proteins contained in the spectral library generated from the  $8 \times 2$  h analysis of the deep MS sequencing of the sample, as well as unique peptides. We anticipate that a narrow selection window of a few  $m/z$  (SWATH-ID of 3  $m/z$ <sup>[18]</sup>) analyzed in a fast machine such as Q-Exactive HF-X

with 40 Hz will generate a untargeted DIA library that might be of competitive quality with the classically generated spectral library, however, with much reduced analysis time, which is also a better match to the subsequent DIA analysis using similar LC-MS/MS parameters.

## Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

## Experimental Section

**Sample Preparation for MS Acquisition:** All animal experiments were performed in accordance with relevant guidelines and regulations of the Vrije Universiteit. The animal ethics committee of the Vrije Universiteit approved the experiments. Hippocampal synaptosomes were prepared from three 3-month-old C57BL6 mice as previously described.<sup>[16]</sup>

The synaptosome was solubilized in 2% SDS, and prepared for MS analysis using the FASP protocol<sup>[1]</sup> and 30K centicon filters from Millipore. Cysteine residues were derivatized by methyl methanethiosulfonate (MMTS). Proteins were digested overnight at 37 °C with Sequence Grade Trypsin/Lys-C from Promega. Peptides were speedvac dried, and redissolved in 0.1% formic acid.

**Mass Spectrometric Acquisition:** Fifteen micrograms of mouse synaptosome was fractionated into eight fractions by high pH reversed-phase cartridges following the protocol included in the kit (Pierce High pH Reversed-Phase Peptide Fractionation Kit). Peptide fractions were dried and then redissolved in 15  $\mu$ L of water containing 0.1% formic acid and spiked with 0.5  $\mu$ L of 10 $\times$  Hyper Reaction Monitoring (HRM) peptides (Biognosys). Five microliters of each fraction was analyzed by nano-LC-MS/MS using the OT Fusion Lumos (Thermo Scientific, San Jose). The spectral libraries were generated by DDA using MS1 OT survey scan and either MS2 HCD with detection in the IT or MS2 HCD with detection in the OT. Peptides were separated by nano-LC (Ultimate 3000 RSLCnano, Thermo Scientific). Peptides were loaded on a  $\mu$ -precolumn (300  $\mu$ m id  $\times$  5 mm, C18 PepMap100, 5  $\mu$ m, Thermo Scientific) at 15  $\mu$ L/min for 3 min using 98/2 water/acetonitrile containing 0.05% TFA. After 3 min, the peptides were separated on an EasySpray column (75  $\mu$ m id  $\times$  50 cm, C18 PepMap, 2 $\mu$ m, Thermo Scientific) at 300 nL/min using water/acetonitrile/formic acid gradient. The gradient consisted of initial step of 3–8% B over 5 min followed

by 8–28% over 90 min, 28–80%B over 7 min, held at 80%B for 4 min and then equilibrated for 15 min at 3% B, where mobile phase A consisted on water containing 0.1% formic acid and mobile phase B consisted of 80/20 acetonitrile/water containing 0.1% formic acid. Separation was performed at 40 °C and the total acquisition time was 150 min. The mass spectrometer was fitted with an EasySpray source (Thermo Scientific) and operated in DDA manner. Each DDA cycle consisted of one OT MS survey scan acquired at 120 000 resolution at  $m/z$  200 and precursors ions meeting user defined criteria such as charge state, monoisotopic precursor selection, intensity, and dynamic exclusion were selected for MS2 based on “most intense.” Precursor ions were isolated using the quadrupole (1.6 Th isolation width) and activated by HCD in the ion routing multipole. In one experiment, fragment ions were detected in the IT in rapid scan and in another experiment fragment ions were detected in the OT at 15 000 resolution (at  $m/z$  200).

One microgram of unfractionated peptides spiked with 1  $\mu$ L of HRM peptides (Biognosys) were analyzed by DIA and WiSIM-DIA by nano-LC-MS/MS using the OT Fusion Lumos. Nano-LC conditions and gradients for DIA and WiSIM-DIA were the same as DDA experiments. DIA on the Fusion Lumos consisted of a MS1 scan at 60 000 resolution at  $m/z$  200 followed by sequential quadrupole isolation windows of 12  $m/z$  for HCD MS/MS with detection of fragment ions in the OT at 15 000 resolution at  $m/z$  200. The  $m/z$  range covered was 400–800 and the Automatic Gain Control (AGC) settings for MS/MS was 5e5 target value and 55 ms maximum injection time. WiSIM-DIA on the Fusion Lumos consisted of four high resolution SIM scans (240 000 resolution at  $m/z$  200) with wide isolation windows of 100  $m/z$  were used to cover all precursor ions of 400–800  $m/z$ . In parallel, each SIM scan, 15 sequential IT MS/MS with 7  $m/z$  isolation windows were acquired to cover the associated 100  $m/z$  SIM mass range. Quantitative information for all precursor ions detected in four sequential SIM scans is recorded in a single run. All ion-trap MS/MS spectra were used to confirm peptide sequences of interest by querying specific fragment ions in the spectral library. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE<sup>[24]</sup> partner repository with the dataset identifier PXD006934.

**Spectral Library Generation:** MS/MS spectra from each DDA dataset were separately imported into MaxQuant<sup>[17]</sup> (version 1.6.0.1) and searched against the Biognosys iRT fasta database and the UniProt mouse proteome (June 2017 release) including both reviewed (Swiss-Prot) and unreviewed (TrEMBL) records of both canonical and isoform sequences. The software does not discriminate between Swiss-Prot and TrEMBL records at any stage; identified protein groups may contain both Swiss-Prot and TrEMBL proteins. Beta-methylthiolation was used as the fixed modification and methionine oxidation and N-terminal acetylation as variable modifications. The minimum peptide length was set to 6, with at most one miss-cleavage allowed. For both peptide and protein identification, a false discovery rate of 1% was set. MaxQuant search results were imported as spectral libraries into Spectronaut with default settings.

The Pulsar search engine integrated in Spectronaut 11 was used to identify peptides and proteins using only the DIA dataset with the exact search engine parameters (fasta database, modifications, peptide length, miss-cleavage, peptide, and protein FDR) as listed above for MaxQuant.

Finally, a merged spectral library based on search engine results from the IT, OT, and DIA datasets was generated using Spectronaut. For each unique precursor, a consensus spectrum of relative fragment ion intensities was composed from all detections of the precursor over all datasets. There was no selection/prioritization by search engine or any other parameters when merging multiple identifications for a precursor. Its consensus iRT (normalized retention time) was computed from the evidence count (number of MS/MS detections) weighted median value. Theoretical  $m/z$  values for fragment ions and precursors were used in all spectral libraries. Protein inference on the merged library was performed on the principle of parsimony using the ID picker algorithm<sup>[25]</sup> as implemented in Spectronaut.

**Analysis of DIA and WiSIM-DIA Data:** Quantitative analysis of DIA and WiSIM-DIA data was performed by using Spectronaut 11 in two separate analyses. Parameter settings of the software were the same for DIA and

WiSIM-DIA and Spectronaut does not perform any computational steps particular to WiSIM-DIA. The generated output of each analysis contains qualitative and quantitative peptide-level data for both MS1 and MS2.

Dynamic retention time prediction was selected to enable nonlinear alignment of precursor retention times between the (iRT, normalized retention time) spectral library and the DIA/WiSIM-DIA data by segmented regression.<sup>[7,8]</sup> Mass calibration was performed by the software to estimate empirical mass accuracy and tolerances used during peak extraction, with initial tolerances set to  $\pm 40$  ppm for OT and  $\pm 0.5$  Th for IT. While matching peptide fragment ions to the spectral library by retention time and  $m/z$ , Spectronaut can additionally use MS1 peptide elution profiles to disambiguate spectral library matches. The peptide identification score FDR,  $Q$ -value in Spectronaut output, was estimated with the mProphet approach<sup>[26]</sup> integrated in Spectronaut using scrambled sequences as decoys (Supporting Information, Figure S6 shows target/decoy spectral library matching score distributions from Spectronaut). This score indicates the preciseness of the observed peptide match and its respective signature in the spectral library and was used as a qualitative metric in our downstream analysis.

The Spectronaut software computed MS1 peptide abundance as the summed precursor XIC (Extracted-Ion Chromatogram, from the monoisotopic precursor ion plus isotopic envelope) and the MS2 peptide abundance as the summation of all selected fragment ions. In downstream analysis, peptide quantification for WiSIM-DIA and DIA was based on MS1 and MS2 abundances, respectively. The S/N ratio was computed using the same XIC profiles and peak integration boundaries; the value for “signal” was defined as the maximum intensity within the peak integration boundary and the “noise” as the average intensity outside of the peak integration boundary for the full width of the extracted XIC.

DIA and WiSIM-DIA analysis results (which contain, among many others; peptide sequence,  $Q$ -value, MS1 and MS2 abundance value, and MS1 and MS2 S/N) were exported as Spectronaut reports and further processed using the R language for statistical computation.<sup>[27]</sup> In downstream analysis we used Spectronaut's peptide  $Q$ -values to discriminate high-confidence peptides within a set of triplicate DIA, or WiSIM-DIA, measurements; only peptides with  $Q$ -value  $\leq 0.01$  in at least two of three replicates were used for quantitative analysis.

## Abbreviations

DDA, data-dependent acquisition; DIA, data-independent acquisition; HCD, Higher-energy collisional dissociation; IT, Ion Trap; OT, Orbitrap; SIM, selected-ion monitoring; WiSIM, wide selected-ion monitoring

## Acknowledgments

F.K. was funded from the Netherlands Organisation for Scientific Research (NWO) Complexity project 645.000.003.

## Conflict of Interest

The authors have declared no conflict of interest.

## Keywords

data-independent analysis, quantitative proteomics, spectral library

Received: August 11, 2017  
Revised: November 4, 2017

- [1] N. J. Pandya, R. V. Klaassen, R. C. van der Schors, J. A. Slotman, A. Houtsmuller, A. B. Smit, K. W. Li, *Proteomics* **2016**, *16*, 2698.
- [2] D. C. Hondius, P. van Nierop, K. W. Li, J. J. Hoozemans, R. C. van der Schors, E. S. van Haastert, S. M. van der Vies, A. J. Rozemuller, A. B. Smit, *Alzheimers Dement.* **2016**, *12*, 654.
- [3] D. B. Bekker-Jensen, C. D. Kelstrup, T. S. Batth, S. C. Larsen, C. Haldrup, J. B. Bramsen, K. D. Sorensen, S. Hoyer, T. F. Orntoft, C. L. Andersen, M. L. Nielsen, J. V. Olsen, *Cell Syst.* **2017**, *4*, 587.
- [4] H. Liu, R. G. Sadygov, J. R. Yates, *Anal. Chem.* **2004**, *76*, 4193.
- [5] A. Michalski, J. Cox, M. Mann, *J. Proteome Res.* **2011**, *10*, 1785.
- [6] L. C. Gillet, P. Navarro, S. Tate, H. Rost, N. Selevsek, L. Reiter, R. Bonner, R. Aebersold, *Mol. Cell. Proteomics* **2012**, *11*, O111.016717.
- [7] R. Bruderer, O. M. Bernhardt, T. Gandhi, S. M. Miladinovic, L. Y. Cheng, S. Messner, T. Ehrenberger, V. Zanotelli, Y. Butscheid, C. Escher, O. Vitek, O. Rinner, L. Reiter, *Mol. Cell. Proteomics* **2015**, *14*, 1400.
- [8] R. Bruderer, O. M. Bernhardt, T. Gandhi, L. Reiter, *Proteomics* **2016**, *16*, 2246.
- [9] C. C. Tsou, C. F. Tsai, G. C. Teo, Y. J. Chen, A. I. Nesvizhskii, *Proteomics* **2016**, *16*, 2257.
- [10] C. C. Tsou, D. Avtonomov, B. Larsen, M. Tucholska, H. Choi, A. C. Gingras, A. I. Nesvizhskii, *Nat. Methods* **2015**, *12*, 258, 257 p following 264.
- [11] A. Keller, S. L. Bader, U. Kusebauch, D. Shteynberg, L. Hood, R. L. Moritz, *Mol. Cell. Proteomics* **2016**, *15*, 1151.
- [12] J. Wang, M. Tucholska, J. D. Knight, J. P. Lambert, S. Tate, B. Larsen, A. C. Gingras, N. Bandeira, *Nat. Methods* **2015**, *12*, 1106.
- [13] Y. Li, C. Q. Zhong, X. Xu, S. Cai, X. Wu, Y. Zhang, J. Chen, J. Shi, S. Lin, J. Han, *Nat. Methods* **2015**, *12*, 1105.
- [14] R. Kiyonami, M. Senko, V. Zabrouskov, A. F. R. Huhmer, Application note Thermo Scientific **2014**, 1.
- [15] L. B. Martin, R. W. Sherwood, J. J. Nicklay, Y. Yang, T. L. Muratore-Schroeder, E. T. Anderson, T. W. Thannhauser, J. K. Rose, S. Zhang, *Proteomics* **2016**, *16*, 2081.
- [16] M. A. Gonzalez-Lozano, P. Klemmer, T. Gebuis, C. Hassan, P. van Nierop, R. E. van Kesteren, A. B. Smit, K. W. Li, *Sci. Rep.* **2016**, *6*, 35456.
- [17] J. Cox, M. Mann, *Nat. Biotechnol.* **2008**, *26*, 1367.
- [18] Y. Kang, L. Burton, A. Lau, S. Tate, *Proteomics* **2017**, *17*, 1500522.
- [19] S. Li, Q. Cao, W. Xiao, Y. Guo, Y. Yang, X. Duan, W. Shui, *J. Proteome Res.* **2017**, *16*, 738.
- [20] R. D. Smith, G. A. Anderson, M. S. Lipton, L. Pasa-Tolic, Y. Shen, T. P. Conrads, T. D. Veenstra, H. R. Udseth, *Proteomics* **2002**, *2*, 513.
- [21] T. P. Conrads, G. A. Anderson, T. D. Veenstra, L. Pasa-Tolic, R. D. Smith, *Anal. Chem.* **2000**, *72*, 3349.
- [22] Y. Mao, L. Zamdborg, N. L. Kelleher, C. L. Hendrickson, A. G. Marshall, *Int. J. Mass Spectrom.* **2011**, *308*, 357.
- [23] M. V. Ivanov, I. A. Tarasova, L. I. Levitsky, E. M. Solovyeva, M. L. Pridatchenko, A. A. Lobas, J. A. Bubis, M. V. Gorshkov, *J. Proteome Res.* **2017**.
- [24] J. A. Vizcaino, A. Csordas, N. del-Toro, J. A. Dienes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-Riverol, F. Reisinger, T. Ternent, Q. W. Xu, R. Wang, H. Hermjakob, *Nucleic Acids Res.* **2016**, *44*, D447.
- [25] B. Zhang, M. C. Chambers, D. L. Tabb, *J. Proteome Res.* **2007**, *6*, 3549.
- [26] H. L. Rost, G. Rosenberger, P. Navarro, L. Gillet, S. M. Miladinovic, O. T. Schubert, W. Wolski, B. C. Collins, J. Malmstrom, L. Malmstrom, R. Aebersold, *Nat. Biotechnol.* **2014**, *32*, 219.
- [27] R. C. Team, R: A language and environment for statistical computing. R Foundation for Statistical Computation. **2015**.