Accelerating DIA studies with fast microflow LC and Zeno SWATH acquisition

Christie Hunter and Alexandra Antonoplis
SCIEX, USA

INTRODUCTION

As the field of quantitative proteomics continues to evolve, larger biological cohorts are being studied, often using precious samples obtained from biobanks or other difficult-to-obtain sources. This creates 2 workflow requirements: the need to acquire quantitative data on the digested samples faster and the need to use smaller amounts of sample. For these types of studies, data-independent acquisition (DIA) continues to grow as the workflow of choice for reproducible quantitative analysis of large numbers of proteins from a proteomic sample.

As such, new workflows and software tools have emerged. Previously, the combination of fast microflow chromatography and SWATH DIA enabled large numbers of proteins to be quantified from complex proteomics samples at very high rates, up to 100 samples per day.[1] Zeno MS/MS on the ZenoTOF 7600 system provides ~5- to 6-fold increase in peptide MS/MS sensitivity and can be used in MRM^{HR}, DDA and DIA workflows.[2,3,4] Also, multiple powerful algorithms have emerged that have enabled more proteins to be identified and quantified from DIA data, such as DIA-NN software.

Here, the improvements in proteins identified and quantified using microflow SWATH DIA coupled with Zeno MS/MS is described. Four different gradient lengths (5, 10, 20 and 45 minutes) were tested to cover a range of application needs. The library-free approach to processing DIA data (using in silico generated spectral libraries) was also evaluated vs. the traditional shotgun proteomics approach with Zeno DDA. Other workflow comparisons were performed to benchmark the workflows. DIA data were processed with DIA-NN software and DDA data were processed with ProteinPilot app in OneOmics suite.

MATERIALS AND METHODS

Sample preparation: Digest of human K562 cell lysate was used from the SWATH Acquisition performance kit (SCIEX). Sample loadings ranged from 12.5–400 ng on column.

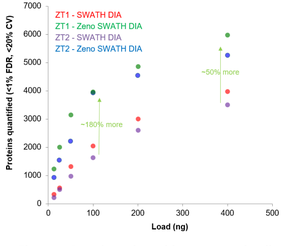
HPLC conditions: Separations were performed using a Waters ACQUITY UPLC M-class system plumbed for microflow chromatography (5 µL/min), operated in trap/elute mode. The trap column was a Phenomenex micro trap (5 µm, 100 Å, 10 x 0.3 mm) and the analytical column was a Phenomenex Kinetex XB-C18 (2.6 µm, 100 Å, 150 x 0.3 mm). Column temperature was controlled at 30 °C. Gradients of 5, 10, 20 or 45 minutes were tested (Table 1).

MS/MS conditions: Data were acquired using 2 ZenoTOF 7600 systems using SWATH DIA mode. Variable window acquisition was used with 40, 60 or 80 windows, and the MS/MS accumulation time was optimized for each gradient length Data were acquired in triplicate with the Zeno trap on and off in MS/MS mode for comparison. In Zeno SWATH DIA mode, the Zeno trap is on continuously for all MS/MS. Scheduled ionization was used to stop ionization during the high flow flush. Data were acquired on the same sample loads and gradient lengths using Zeno DDA for comparison.

Data processing: Data were processed using DIA-NN software using a spectral ion library previously generated from 2 human cell lysates.[6] Processing settings were described previously in a community post.[7] Protein and peptide precursor areas from the *.pr_matrix.tsv and *.pg_matrix.tsv files (<1% FDR filter applied) were copied into Microsoft Excel and the numbers of proteins and peptides quantified at 20% CV were computed.

TRADEMARKS/LICENSING

To receive a copy of this poster:
- Scan the code with your phone camera
- Complete the form



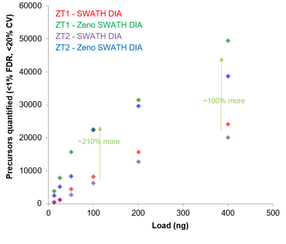


Figure 1. Protein and peptide precursor loading curves for the 45-min gradient with and without Zeno MS/MS. The numbers of proteins (top left) and peptides (top right) quantified increased as sample load was increased on both instruments. The optimal loading amount was 200-400 ng of cell lysate digest on column. The fold change increase (bottom right) in the numbers of proteins (circle) and peptides (diamonds) was computed for datasets collected on each instrument using a 45-min gradient. Larger gains were observed at the lower loads.

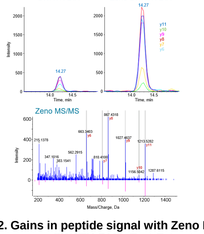
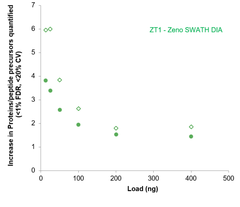


Figure 2. Gains in peptide signal with Zeno MS/MS activated during SWATH DIA. (Top) Extracted ion chromatograms (XICs) of peptide TLGEC[CAM]GFTSQTAR from sp|Q15370|ELOB_HUMAN from the SWATH DIA and Zeno SWATH DIA experiments (100 ng of sample, 45-min gradient). (Bottom) MS/MS spectrum from the Zeno SWATH DIA experiment, with the extracted fragments shown in the Zeno SWATH DIA XIC plot labeled in red.



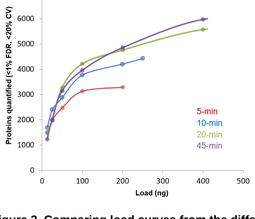


Figure 3. Comparing load curves from the different gradient lengths. Visualizing loading curves allows users to select optimal loading amounts to achieve the best results and system uptime. Longer gradients enable higher loadings until improvements in quantification plateau. For this application, a load of approximately 100 ng of protein on column was optimal for 5- and 10-min gradients, whereas a load of approximately 400 ng of protein was optimal for 20- and 45-min gradients.

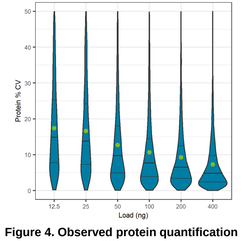


Figure 4. Observed protein quantification variance using a 20-min gradient. Violin plot of the %CV values observed for all detected proteins at each of the loading levels. Green dot is average %CV, thick black line is median. Loadings >100 ng demonstrated %CV distributions well below <20%, with distribution medians of 7.5%, 6.4% and 4.7% for the loads of 100, 200 and 400 ng of protein on column, respectively. Numbers of proteins included in this plot can be visualized in Figure 5.

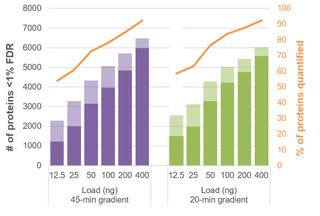
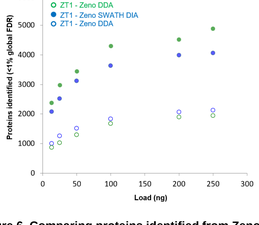


Figure 5. Proteins identified and quantified for 45- and 20-min gradients. The number of proteins identified at <1% FDR (transparent) and quantified at <20% CV (solid) across a range of loadings was studied when using a 45- (left) or 20-min (right) gradient. At the lower loads, ~70% of proteins identified were quantified with <20% CV. At the higher loads, more than 90% of proteins identified were quantified with <20% CV. Data suggest the appropriate loading range for a 300 µm x 15 cm column is 200 – 400 ng on column when using the 45- or 20-min gradients.



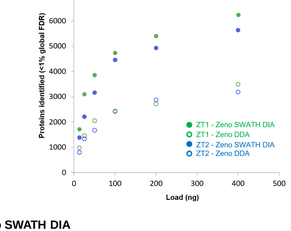
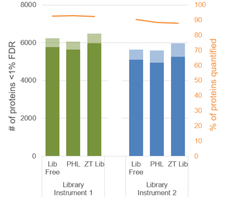


Figure 6. Comparing proteins identified from Zeno SWATH DIA data vs. Zeno DDA data. Zeno SWATH DIA data processed with the *in silico* generated library (filled circles) outperformed Zeno DDA data (open circles) for protein identifications on 2 ZenoTOF 7600 systems (green vs. blue) across a range of sample loadings on column. Two different gradient lengths were evaluated, 45-min gradient (top left) and 10-min gradient (top right). Using the 45-min gradient data collected on instrument 1 for a 400-ng load, nearly all the proteins identified from the Zeno DDA dataset were also identified in the Zeno SWATH DIA dataset (bottom right). The Zeno SWATH DIA approach yielded 2-fold more protein identifications.



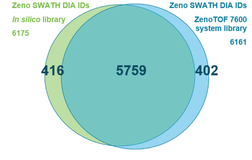


Figure 7. Comparing protein identifications obtained with the library-free approach and 2 experimental libraries. Zeno SWATH DIA data were processed with DIA-NN software using 3 different libraries (Lib Free, using an *in silico* library from a FASTA file; PHL, Pan Human library; ZT Lib, large library from Zeno DDA dataset collected on a ZenoTOF 7600 system). Using a 45-min gradient and a 400-ng sample load, the numbers of proteins identified at <1% FDR (transparent) and quantified at <20% CV (solid) were measured (left). Strong overlap between the proteins identified was observed (right) between the library free approach and the traditional large library approach.

CONCLUSIONS

- The Zeno trap is a key innovation on the ZenoTOF 7600 system that provides large gains in MS/MS sensitivity and enables significant improvements in protein identification workflows,[5] targeted peptide quantification workflows[3] and the SWATH DIA workflow.
- Zeno SWATH DIA provides large gains in the numbers of proteins quantified from human cell lysate samples, with 50% more proteins quantified at high loads and 200-400% more quantified at low loads vs. SWATH DIA without Zeno trap activated (Figure 1)
- Zeno SWATH DIA is readily compatible with fast microflow gradients, enabling up to ~6,000 proteins to be quantified, depending on the selected gradient and sample load, with ~95% of those proteins quantified at <20% CV (Figure 5)
- Using an in silico-generated library for data processing with DIA-NN software, ~2-fold more proteins were identified from a cell lysate using a 45-min LC gradient using Zeno SWATH DIA vs. the traditional shotgun proteomics approach Zeno DDA (Figure 6). Approximately 2.5- to 3-fold more were identified using a 10-min gradient.
- A comparison of the proteins identified using the in silico-generated library to 2 experimentally generated libraries demonstrated high agreement. The library-free workflow using DIA-NN software is an easy and comprehensive alternative for large-scale protein identification experiments (Figure 7).

REFERENCES

1. Accelerating SWATH acquisition for protein quantitation – up to 100 samples per day. SCIEX technical note, RUO-MKT-02-8432-A.
2. Qualitative flexibility combined with quantitative power. SCIEX technical note, RUO-MKT-02-13053-A.
3. Large-scale, targeted, peptide quantification of 804 peptides with high reproducibility, using Zeno MS/MS. SCIEX technical note, RUO-MKT-02-13346-A.
4. Zeno MS/MS with microflow chromatography powers the Zeno SWATH DIA workflow to quantify more proteins. SCIEX technical note, RUO-MKT-02-14668-A.
5. Over 40% more proteins identified using Zeno MS/MS. SCIEX technical note, RUO-MKT-02-13286-B.
6. Large-scale protein identification using microflow chromatography on the ZenoTOF 7600 system. SCIEX technical note, RUO-MKT-02-14415-A.
7. Processing ZenoTOF 7600 system data with DIA-NN software. SCIEX community post, RUO-MKT-18-14611-A.