Nanoflow Zeno SWATH data-independent analysis (DIA) for high-sensitivity protein identification and quantification

Using core-shell based nanoflow columns on the ZenoTOF 7600 system

Christie Hunter and Alexandra Antonoplis
SCIEX, USA

Nanoflow chromatography is often used in proteomics workflows to obtain the highest sensitivity, especially when available sample amounts are very low. High-quality chromatographic separations are important for good results, as good peak shape and peak resolution can reduce ion suppression and allow the MS system to sample as many unique peptides as possible.

In other work using microflow chromatography, a series of loading curves were generated using several different gradient lengths to determine the optimal conditions for microflow proteomics analysis using Zeno SWATH DIA.[1] A similar strategy was employed here to both characterize the optimal conditions for nanoflow chromatography and enable a comparison between the flow regimes. Also, a comparison between the Zeno data-dependent analysis (DDA) workflow and the Zeno SWATH DIA workflow showed that >2x more proteins could be identified and quantified using the Zeno SWATH DIA strategy.[1] This work

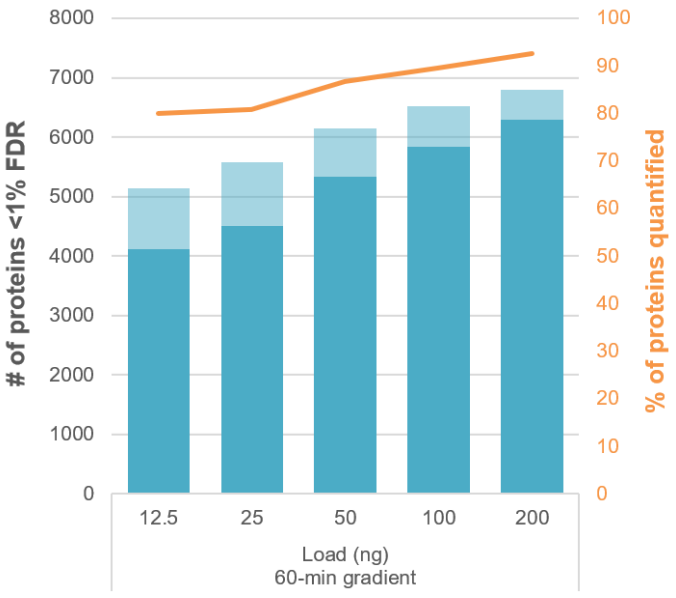

Figure 1. Proteins identified and quantified at different loadings of K562 digest for a 1-hour gradient. The number of proteins identified at <1% FDR (transparent) and quantified at <20% CV (solid) using Zeno SWATH DIA across a range of loadings was studied. At all loads, >80% of proteins identified were also quantified.

focused on characterizing the Zeno SWATH DIA approach with nanoflow chromatography.

Here, long nanoflow columns packed with core-shell chromatographic phase were used in combination with Zeno SWATH DIA to evaluate the power of the ZenoTOF 7600 system for nanoflow protein identification and quantification workflows. A range of sample loadings of a human cell lysate, K562, was acquired and the numbers of proteins identified and quantified were determined (Figure 1). Various library strategies for processing Zeno SWATH DIA data with DIA-NN software were also assessed.

Key features of nanoflow LC with Zeno SWATH DIA

- The ZenoTOF 7600 system with Zeno SWATH DIA provides high-quality proteomics quantification on many proteins and peptides in a single run
- The OptiFlow Turbo V ion source enables the user to easily switch between nanoflow and microflow chromatography[3]
- Phenomenex columns packed with core-shell C18 phase provide robustness and separation power
- Zeno SWATH DIA enables the identification and quantification of large numbers of proteins from a human cell lysate (Figure 1)
- From 200 ng of sample on column and a 1-hour gradient, ~6800 proteins were identified and ~6300 proteins were quantified (92%)

Methods

Sample preparation: Digest of human K562 cell lysate was used from the SWATH Acquisition Performance Kit (SCIEX). Sample loadings ranged from 12.5–200 ng on column.

Chromatography: A nanoflow column packed with core-shell bioZen 5 µm Peptide XB-C18 (0.075 x 500 mm, 5 µm, P/N 00J-4792-AW-SX) was used for sample analysis. A NanoLC 425 system (SCIEX) configured for direct injection was used for all liquid chromatography in contact closure mode. A flow rate of 400 nL/min was used with a 1-hour gradient.

Mass spectrometry: Data were acquired using a ZenoTOF 7600 system with the OptiFlow Turbo V ion source installed in nanoflow configuration.[3] Zeno SWATH DIA was performed on all sample loads using 100 variable Q1 windows and an accumulation time of 25 msec. Data were acquired in triplicate with the Zeno trap enabled in MS/MS mode for all acquisitions.

Data processing: Zeno SWATH DIA data were processed using DIA-NN software using a spectral ion library previously generated from 2 human cell lysates.[6] Data from each fractionation experiment of each cell line were processed into a single search result using the ProteinPilot app in OneOmics suite. The search results for each cell line were then merged using the Extractor application to create a final ion library for SWATH DIA processing. Processing settings were described previously in a community post.[5] Data were also processed using the library-free process in DIA-NN software.[1,4] 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.

Using core-shell columns

When high sensitivity is required for proteomics analysis, nanoflow chromatography is typically used. There is a wide range of column phases and column lengths available from various suppliers that includes longer columns and columns packed with small particles to maximize separation power. As these columns get longer or the particle size gets smaller, the back pressure can increase, making the columns more difficult to use routinely and robustly. Phenomenex has developed core-shell based phases that provide higher resolution and better efficiencies at higher flow rates, without increasing the back pressure. Long, 50 cm columns packed with the bioZen 5 µm Peptide XB-C18 were used in this study at a flow rate of 400 µL/min. Under these conditions, the back pressure was ~4000 psi. Typical separation for the 1-hour linear gradient is shown in Figure 2 for the 50 ng sample load.

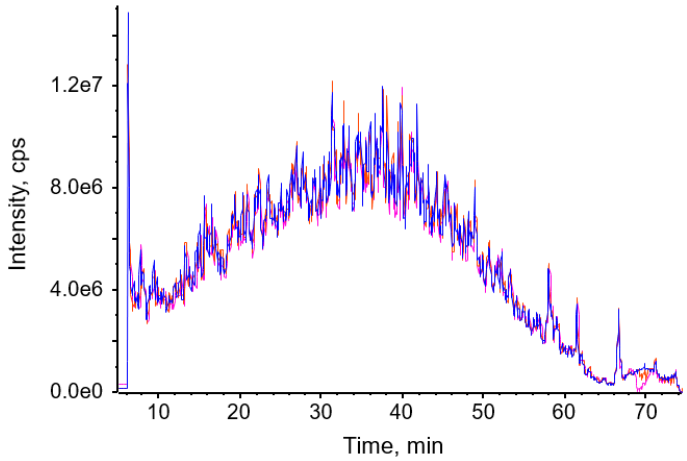

Figure 2. Reproducible separations with the Kinetex 50 cm nanoflow column. Triplicate injections were performed at each sample load to obtain quantitative reproducibility measurements. Shown here is the TOF MS total ion chromatogram (TIC) for 50 ng of K562 digest on column with a 1-hour linear gradient.

High numbers of proteins identified and quantified

Using the 1-hour linear gradient, a range of loadings of digested cell lysate was analyzed with Zeno SWATH DIA using nanoflow chromatography (Figure 1). Data were processed using DIA-NN software and a large library created on a ZenoTOF 7600 system, from Zeno DDA data collected on 2 fractionated cell lysates.[6] As expected, as the sample loading increased, the numbers of proteins identified also increased. The quantitative reproducibility was assessed from the triplicate injections to determine the number of proteins at <1% FDR that were also quantified with <20% CV. When 200 ng of sample was loaded on column, ~6800 proteins were identified and ~6300 proteins were quantified (92%).

Performance of different library strategies

Interest has grown in using DIA data for protein identification workflows and protein quantification work. Therefore, the library-free approach was also tested on nanoflow data. Using a library generated *in silico* from a human FASTA file, the same Zeno SWATH DIA data were processed and compared to the data from the experimentally generated library (Figure 3). Very similar results were observed between the 2 approaches for the numbers of proteins and peptides identified and quantified. Similar results were also achieved when the SWATH DIA data were processed with the publicly available Pan Human Library.

Figure 3. Comparing protein identification results from 3 different library approaches. The nanoflow Zeno SWATH DIA data from the 100 ng sample load was processed using 3 different libraries (Lib Free, using an *in silico* library from a FASTA file; PHL, Pan Human library; ZT Lib, large library from a Zeno DDA dataset). (Top) The numbers of proteins identified at <1% FDR (transparent) and quantified at <20% CV (solid) were measured. (Bottom) The proteins identified using the library-free approach or the ZenoTOF 7600 library were compared. Very good overlap was observed in the detected proteins.

To further test the fidelity of the library-free approach, the proteins identified using the *in silico* generated library were compared to the ZenoTOF 7600 system library. The Venn diagram (Figure 3, bottom) shows that very similar lists of proteins were generated by these approaches, adding further confidence to the library-free approach.

Peptide precursors identified and quantified

Visualizing load curves is helpful when determining the correct LC strategy and sample load for a specific sample set. The number of peptide precursors detected and quantified is shown in Figure 4. Evaluation of the peptide and protein level load curves (Figure 1) suggests that a sample load of 100-200 ng on column provides the best depth of proteome coverage for this chromatographic strategy when using Zeno SWATH DIA.

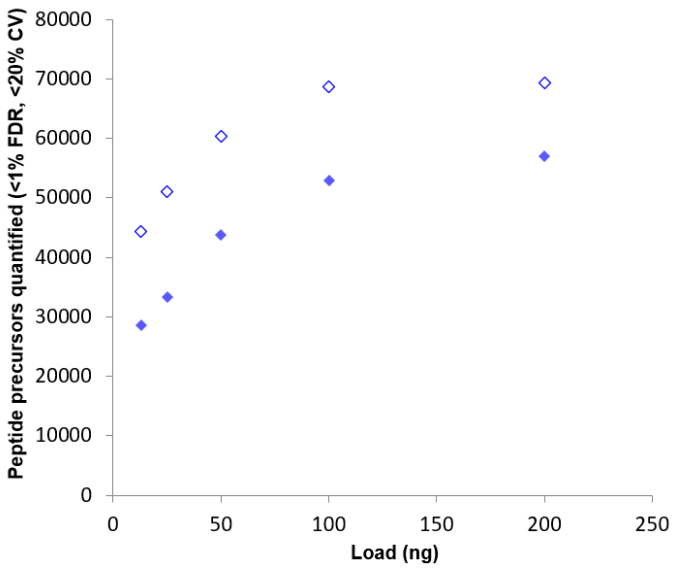

Figure 4. Comparing peptide precursors quantified at the various loading amounts. As expected, the number of peptide precursors identified (open) and quantified (filled) from the 1-hour nanoflow gradient increased as the loading amount increased. Nearly 67,000 peptides were identified and ~80% of them quantified at the 100 ng load.

Conclusions

Zeno SWATH DIA on the ZenoTOF 7600 system can be used to achieve large gains in the number of proteins and peptides that can be quantified from a proteomics sample.

- Nanoflow Zeno SWATH DIA enabled the identification and quantification of many proteins from human cell lysate samples, with ~6800 proteins identified and ~6300 proteins quantified (92%) from 200 ng of sample on column and a 1-hour gradient
- With Zeno MS/MS, more peptides and proteins were reproducibly quantified with <20% CV
- DIA-NN software provided a robust processing engine to identify and quantify proteins from nanoflow Zeno SWATH DIA data acquired on the ZenoTOF 7600 system using both experimentally generated libraries and the library-free approach

References

1. Going library-free for protein identification using Zeno SWATH DIA and in silico-generated spectral libraries. SCIEX technical note, RUO-MKT-02-14675-A.
2. Zeno MS/MS with microflow chromatography powers the Zeno SWATH DIA workflow for more proteins quantified. SCIEX technical note, RUO-MKT-02-14668-A.
3. Easy switching of sources and LC flow regimes on the ZenoTOF 7600 system. SCIEX community post, RUO-MKT-18-14697-A.
4. Creating a library from a FASTA file for library-free data analysis. SCIEX community post, RUO-MKT-18-14611-A.
5. Processing ZenoTOF 7600 system data with DIA-NN software. SCIEX community post, RUO-MKT-18-14611-A.
6. Large-scale protein identification using microflow chromatography on the ZenoTOF 7600 system. SCIEX community post, RUO-MKT-18-14415-A.
7. Rosenberger G *et al* (2014) A repository of assays to quantify 10,000 human proteins by SWATH-MS. *Scientific data.* **1,** 140031.

Headquarters
500 Old Connecticut Path | Framingham, MA 01701 USA
Phone 508-383-7700
sciex.com

International Sales
For our office locations please call the division headquarters or refer to our website at
sciex.com/offices