

Pushing the boundaries of sensitivity and depth-of-coverage for nanoflow proteomics

Featuring Zeno SWATH data-independent acquisition using IonOpticks Aurora SX columns

Patrick Pribil, Stephen Tate and Ihor Batruch
SCIEX, Canada

High-resolution mass spectrometry (MS) has become a standard technology for the characterization of proteins, particularly using data-independent acquisition (DIA). The ability to detect and quantify proteins across a wide dynamic range requires achieving both high sensitivity and maximum depth of proteome coverage. For this purpose, nanoflow liquid chromatography (LC) is often used. Maximizing MS performance can be realized through improvements in nanoflow chromatographic separation, specifically with narrower LC peak shapes and broader separation of complex mixtures of peptides.

This technical note describes the use of Zeno SWATH DIA on the ZenoTOF 7600 system in combination with nanoflow chromatographic separation using IonOpticks Aurora SX series columns. Maximum sensitivity is demonstrated with the detection of >2,300 protein groups (>10,000 precursors) in low column loads (250 pg) of a complex digest standard, approximating single-cell protein levels. To show the depth of proteome coverage and quantitative fidelity, mixtures of digests from 3 separate organisms (human, yeast and *E. coli*) at different ratios were analyzed, whereby >19,000 protein groups were detected in the mixtures with >94% showing quantitative reproducibility of <20% CV and protein ratios between different mixtures

accurately quantitated using a label-free quantitation (LFQ) strategy with Zeno SWATH DIA.

Key features of Zeno SWATH DIA performance using IonOpticks Aurora SX nanoflow columns

- **Unlocking the power of label-free quantitation of complex proteome mixtures:** >19,000 protein groups (>145,000 precursors) were identified from 500 ng mixtures of human/yeast/*E. coli* digests using 90-minute nanoflow LC gradients, with >94% of the protein groups and >88% of the precursors having CVs <20%, as shown in Figure 1. Excellent quantitative results were achieved from the LFQ protein ratios between mixtures
- **Sensitivity at single-cell level column loads:** >2,300 protein groups (>10,000 precursors) were detected from 250 pg of human K562 digest using 30-minute nanoflow LC gradients
- **Pushing the boundaries of protein detection at higher loads:** 8,400 protein groups (80,000 precursors) were detected from 200 ng of human K562 digest using 90-minute nanoflow LC gradients

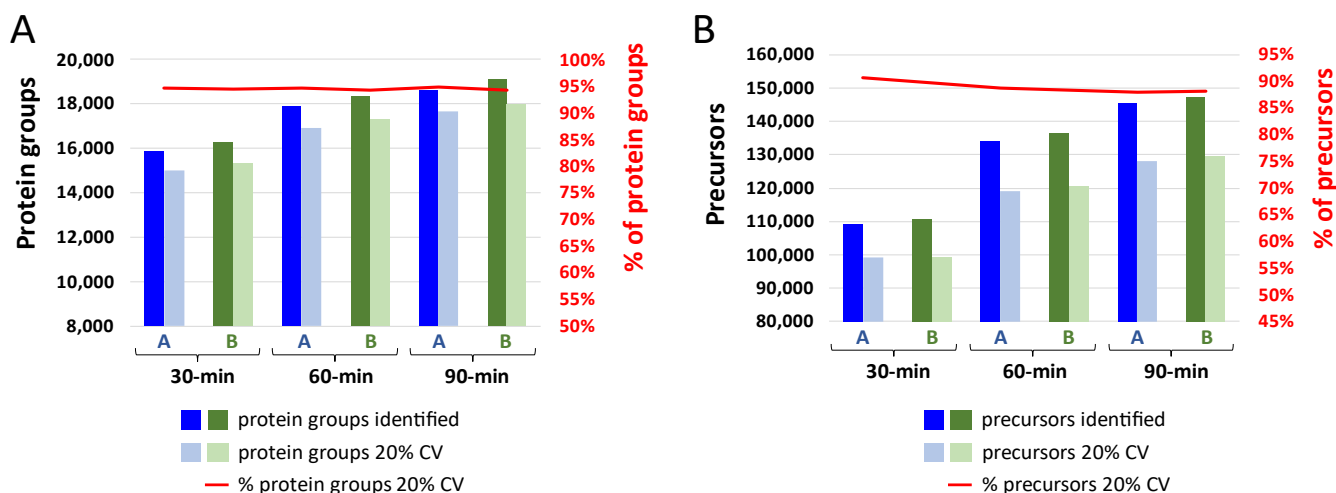


Figure 1. Protein groups (A) and precursors (B) detected and quantified from 500 ng of 3-organism mixtures using Zeno SWATH DIA. Sample A contains 65% human, 30% yeast and 5% *E. coli* digest, while sample B contains 65% human, 15% yeast and 20% *E. coli* digest. Samples were analyzed on an IonOpticks Aurora UltimateSX nanoflow column with the indicated gradient lengths. Data was processed with DIA-NN software against a spectral library made from the UniProt FASTA protein sequences from human, yeast and *E. coli*.

Introduction

The central goal of proteomics is the characterization of proteins, specifically their biological functions within an organism and the roles they play in both normal cellular function and disease states. This requires the ability to detect proteins and quantify their relative abundances or abundance changes in a dynamic environment. DIA methodologies have become well-established for proteomics research, as a high degree of qualitative and quantitative information can be obtained using an efficient method without prior knowledge about a sample. Zeno SWATH DIA is a powerful technique for the detection and quantitation of proteins across various sample loads and gradient lengths, due to the speed of acquisition as well as the sensitivity that Zeno trap pulsing provides in MS/MS.^{1,2} Speed and throughput, generally associated with higher flow rate techniques such as microflow LC, are typically preferred by users, and the robustness of Zeno SWATH DIA proteomics on the ZenoTOF 7600 system using microflow LC has been recently shown.³ However, there are instances where nanoflow LC is required, specifically when extremely low amounts of samples are being analyzed or if maximum proteome depth is needed. Improvements in nanoflow column performance in recent years, namely LC peak shape/width, peak capacity and separation, as well as minimalization of dead volumes, all directly impact the quality of MS data acquisition. This work demonstrates the combination of nanoflow chromatography (using IonOpticks Aurora SX nanoflow columns, Figure 2) with Zeno SWATH DIA. Sensitivity for detecting and quantitating proteins at low column loads are shown using a 15 cm column with 30-minute LC gradients. With higher column loads of complex digest samples, very high numbers of protein and peptide identifications were achieved using a 25 cm column with 30-, 60- and 90-minute gradients. Together, these results reach new maxima for protein ID and quantitative performance and highlight the capabilities of the ZenoTOF 7600 system for the extreme challenges of proteomics research.

Methods

Sample preparation: Lyophilized human K562 (Promega), yeast (Promega) and *E. coli* (Waters) tryptic digests were reconstituted in water containing 5% acetonitrile and 0.1% formic acid to create stock solutions with concentrations of 0.5 µg/µL.

K562 dilutions: The K562 stock solution was further diluted to 200, 50, 25, 10, 5, 1, 0.5 and 0.25 ng/µL using the same dilution solution as above. For dilutions of 5, 1, 0.5 and 0.25 ng/µL, bovine serum albumin was added to the dilution solution at a final concentration of 5 fmol/µL as a background carrier.

3-organism mixtures: K562, yeast and *E. coli* stock solutions were mixed at two different ratios (samples A and B) as previously described.⁴ The final mixture ratios for each organism (A/B) were 1:1 for human, 2:1 for yeast and 1:4 for *E. coli*, with the final concentration of total protein in each sample being 500 ng/µL.

Chromatography: A Waters M-Class LC system was used for all LC-MS experiments. In all cases, injections of 1 µL were done using a direct-inject LC scheme.

K562 dilutions: Samples (50 ng/µL dilutions and lower) were analyzed on an IonOpticks Aurora EliteSX C18 column (75 µm x 15 cm) with a 30-minute gradient (70-minute total run time). The gradient was applied at a flow rate of 150 nL/min, while loading and re-equilibration of the column were done at 350 nL/min.

3-organism mixtures: Samples were run on an IonOpticks Aurora UltimateSX C18 analytical column (75 µm x 25 cm) with either 30-minute gradients (110-minute total run time), 60-minute gradients (140-minute total run time) or 90-minute gradients (170-minute total run time). For all 3 gradients, the flow rate was kept constant at 150 nL/min throughout the LC run. Column heating for all experiments was set to 50°C.



Figure 2. IonOpticks Aurora SX nanoflow columns connected to the OptiFlow TurboV ion source on the ZenoTOF 7600 system.

Mass spectrometry: The ZenoTOF 7600 system was operated in Zeno SWATH DIA using the OptiFlow TurboV ion source in nanoflow configuration. Zeno SWATH DIA methods consisted of either 38 variable-width windows (Q1 mass range 400-900 Da) with MS/MS accumulation times of 32 ms, or 85 variable-width windows (Q1 mass range also 400-900 Da) with MS/MS accumulation times of 18 ms. Both Zeno SWATH DIA methods consisted of a TOF-MS survey scan from 400-1,500 Da with an accumulation time of 50 ms. The MS/MS mass range for all methods was 140-1,750 Da, using CID fragmentation with dynamic collision energies and Zeno trap pulsing turned on. The ion source conditions for all experiments were as follows: ion spray voltage = 1,500 V, curtain gas = 25, nebulizing gas (gas 1) = 10, and interface temperature = 300°C. Calibration of the ZenoTOF 7600 system for both TOF MS and MS/MS was done approximately every 6 hours with separate injections of 20 fmol of PepCalMix standard (SCIEX), using a 15-minute gradient (55-minute total run time).

Data processing: All Zeno SWATH DIA datasets were processed using DIA-NN software (version 1.8.1).⁵

K562 dilutions: Results were processed against a spectral library created using data-dependent acquisition results from high-pH fractionated K562 and HeLa digests (11,269 protein groups and 169,395 peptide precursors).⁶ Default processing settings were used, with match between runs (MBR) checked. The pg.matrix.tsv and pr.matrix.tsv output files were used for reporting protein groups and precursors, respectively, and for calculating identifications at 20% CV thresholds.

3-organism mixtures: Zeno SWATH DIA data was processed against a spectral library created from the combined UniProt FASTA canonical and isoform protein sequences for human, yeast and *E. coli* (119,434 protein groups, 5,597,011 peptide precursors). Default processing settings were used, with MBR checked and protein inference turned off. Protein groups and precursors identified and quantified in each sample were determined using the output pg.matrix.tsv and pr.matrix.tsv files, respectively. The LFQ area ratio plots were generated from the MaxLFQ values (calculated from protein groups and precursors detected in all 3 replicates for each sample) using Python software.⁷

Pushing the boundaries of depth-of-coverage using Zeno SWATH DIA

Two separate samples with varying mixtures of human, yeast and *E. coli* digests (500 ng total sample injected) were analyzed using 30-, 60- and 90-minute gradients on an IonOpticks Aurora UltimateSX nanoflow column. The resulting protein group and

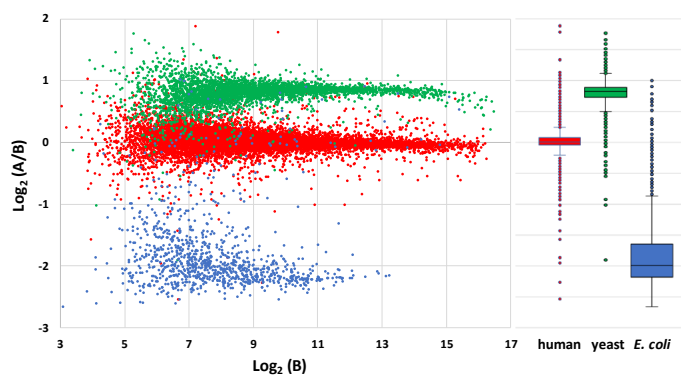


Figure 3. Label-free quantitation of the ratios of protein groups by species using Zeno SWATH DIA. Two sets of mixtures from 3 different organism digests were prepared and analyzed with Zeno SWATH DIA. Replicate injections of 500 ng of each sample were analyzed using 90-minute gradients. Ratios were calculated from the resulting MaxLFQ outputs generated from DIA-NN software processing. The expected ratios for human (1:1, $\text{Log}_2(A/B) = 0$), yeast (2:1, $\text{Log}_2(A/B) = 1$), and *E. coli* (1:4, $\text{Log}_2(A/B) = -2$) are shown. The box-and-whisker plots to the right show the high degree of precision and accuracy.

precursor identifications are shown in Figure 1. Using 30-minute nanoflow gradients, 15,853 and 16,260 protein groups could be identified from samples A and B, respectively (Figure 1A). Increasing the gradient length to 60 minutes or 90 minutes increased the protein group identifications to 17,900 and 18,622 for sample A and to 18,363 and 19,105 for sample B, respectively. The percentage of protein groups with CVs <20% remained very constant at 95% for all cases, showing the extremely high quantitative reproducibility of the analysis. Figure 1B shows a similar high detection rate for precursors: 109,120 and 110,673 precursors were detected in samples A and B, respectively, using 30-minute gradients. Using 60-minute or 90-minute gradients increased the precursor identifications to 134,123 and 145,572 for sample A and to 136,390 and 147,248

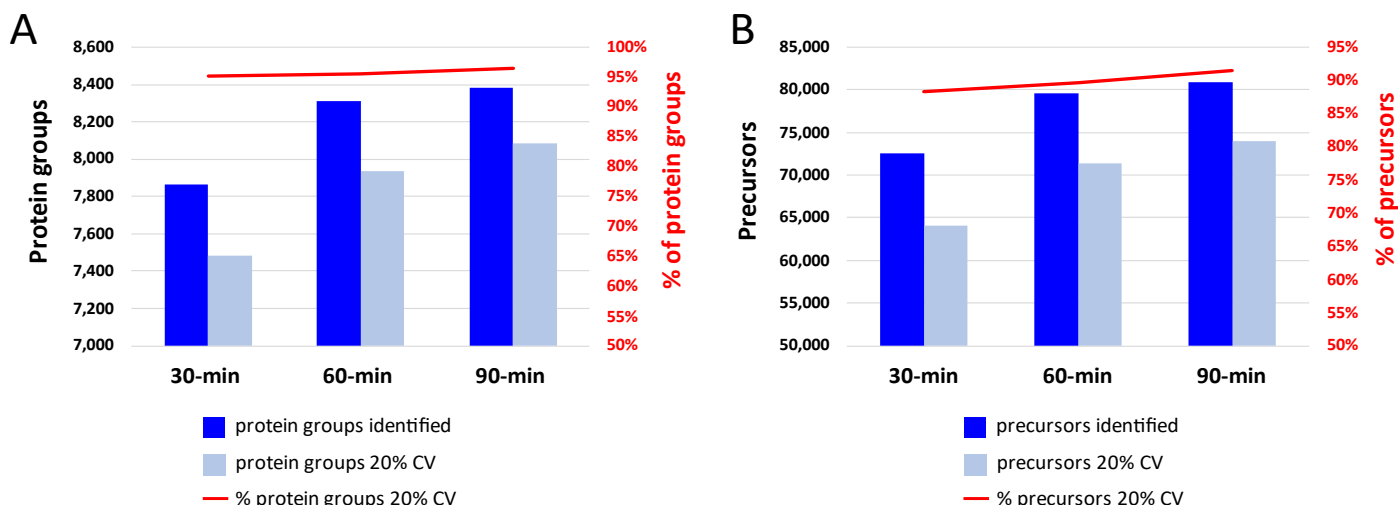


Figure 4. Protein groups (A) and precursors (B) detected and quantified with <20% CV from 200 ng of human K562 digest using Zeno SWATH DIA. Samples were analyzed using nanoflow chromatography on an IonOpticks Aurora UltimateSX column with the indicated gradient lengths. Data was searched with DIA-NN software against an in-house K562/HeLa spectral library.

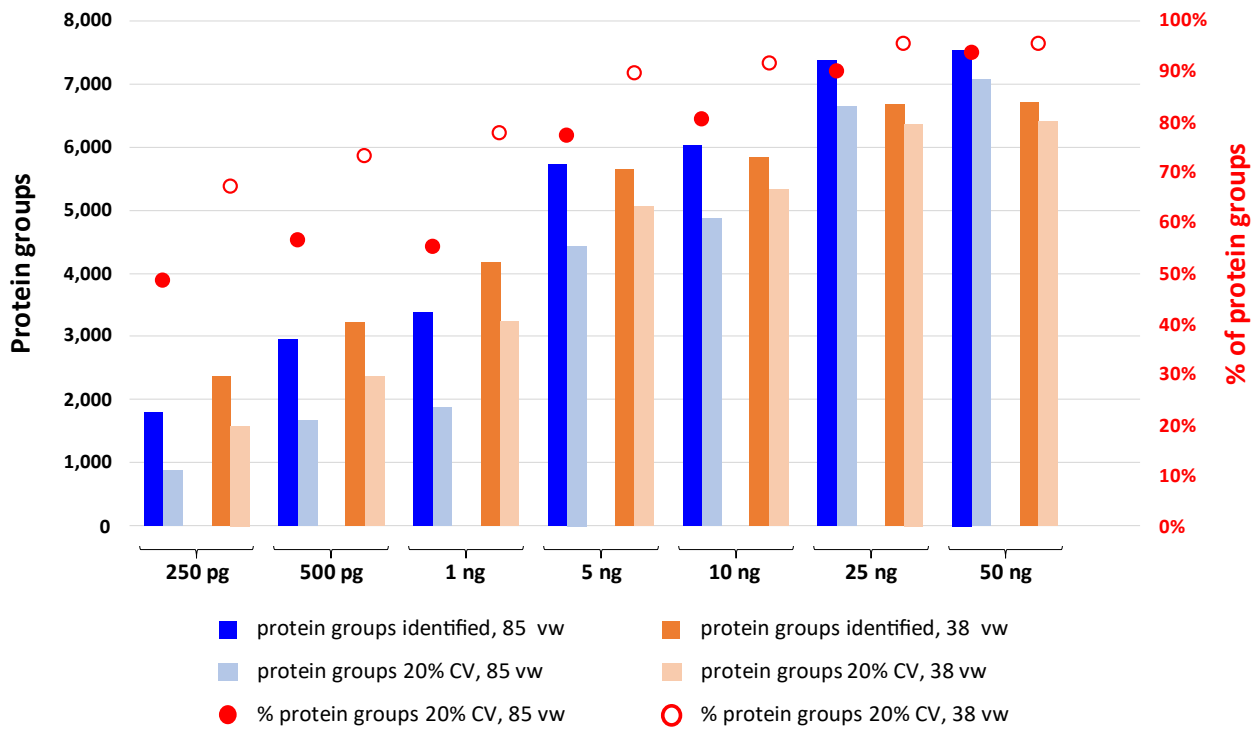


Figure 5. Protein groups identified and quantified with <20% CV from different on-column loadings of human K562 digest using Zeno SWATH DIA. Samples were analyzed using nanoflow chromatography on an IonOpticks Aurora EliteSX column with 30-minute gradients. The Zeno SWATH DIA methods (85 vw or 38 vw) are indicated. Data was searched with DIA-NN software against an in-house K562/HeLa spectral library.

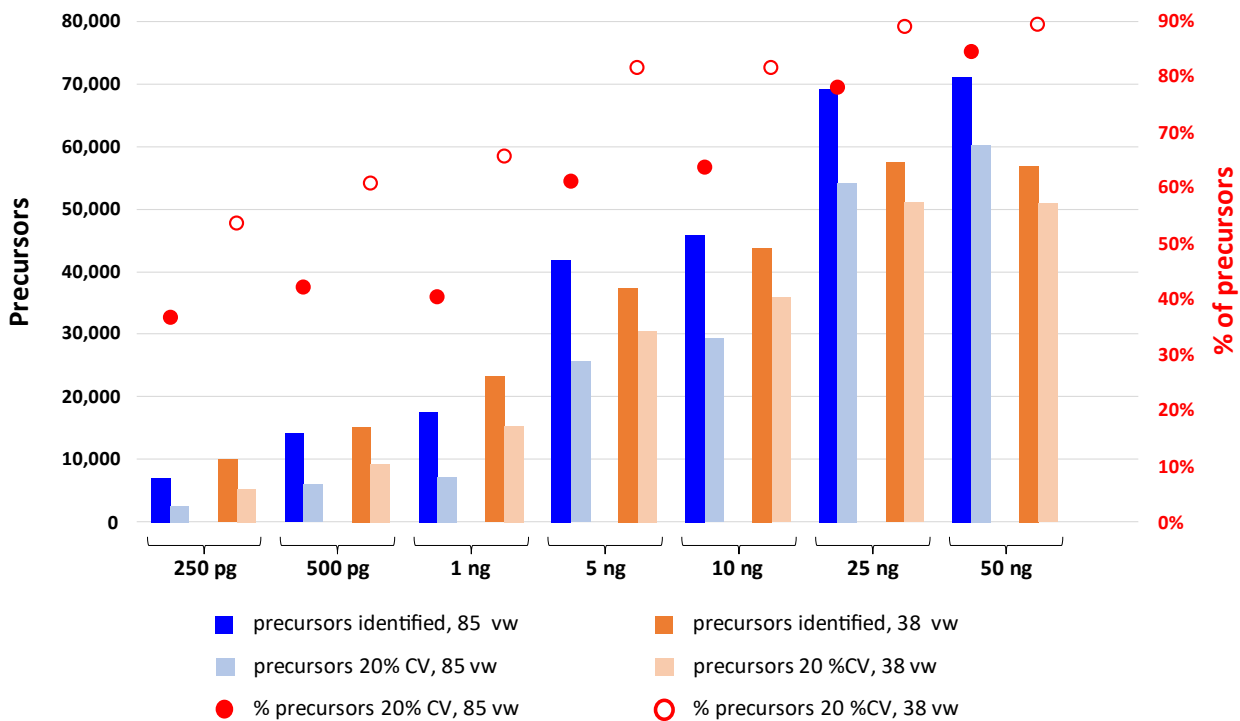


Figure 6. Precursors identified and quantified with <20% CV from different on-column loadings of human K562 digest using Zeno SWATH DIA. Samples were analyzed using nanoflow chromatography on an IonOpticks Aurora EliteSX column with 30-minute gradients. The Zeno SWATH DIA methods (85 vw or 38 vw) are indicated. Data was searched with DIA-NN software against an in-house K562/HeLa spectral library.

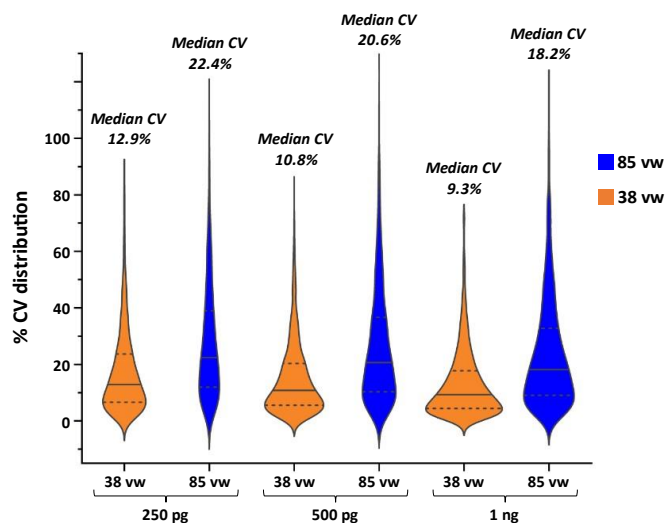


Figure 7. %CV distributions of protein groups identified from low on-column loadings of human K562 digest using Zeno SWATH DIA. Samples were analyzed using nanoflow chromatography on an IonOpticks Aurora EliteSX column with 30-minute gradients. The Zeno SWATH DIA methods (85 vv or 38 vv) are indicated. Data was searched with DIA-NN software against a K562/HeLa spectral library. The median CV values for each condition are shown and are indicated by the solid bars in the violin plots, where dashed lines indicate the interquartile levels.

for sample B, respectively. The percentage of precursors with <20% CV also remained constant, between 88-90% in all cases. Transitioning from 30-minute to 60-minute gradients increased protein group identifications by 12.9%, while transitioning from 60-minute gradients to 90-minute gradients only increased protein group identifications by 4%. Although this indicates that the upper limit of usable gradient length may have been reached for the 25 cm nanoflow column with this amount of sample, it also highlights the speed of the ZenoTOF 7600 system for Zeno SWATH DIA analysis.

Using the output MaxLFQ results from the 90-minute gradient data, the calculation of the protein group ratios for each species is summarized in Figure 3, whereby the Log₂ (MaxLFQ area ratio for samples A/B) is plotted against the Log₂ (MaxLFQ area for sample B) for each species. Only protein groups having MaxLFQ values for all three replicates in both samples and with CVs <20% were used for the plots. A total of 17,879 protein groups (13,187 for human, 3,705 for yeast and 987 for *E. coli*) are plotted in the figure. As shown, the observed ratios match well with the expected ratios for each species due to the quantitative precision and accuracy provided by Zeno SWATH DIA.

To further test the depth of coverage, 200 ng injections of K562 digest were analyzed using 30-, 60- and 90-minute gradients

with the IonOpticks Aurora UltimateSX columns and Zeno SWATH DIA. Protein group and precursor identification and quantitation are shown in Figure 4. After processing the data with DIA-NN software against a K562/HeLa spectral library, 7,862, 8,312 and 8,383 protein groups were identified with the respective gradients, with 95-96% of protein groups having CVs <20% (Figure 4A). Likewise, 72,594, 79,552 and 80,866 precursors were identified with the 30-, 60- and 90-minute gradients, respectively, with 88-92% of the precursors having CVs <20% (Figure 4B).

Together, this data eclipses the previous boundaries of total protein group and precursor identifications using Zeno SWATH DIA on the ZenoTOF 7600 system and demonstrates the quantitative power made possible using this technique when coupled with high-performance chromatographic separation.

Extending the limits of sensitivity of protein detection at low column loads with Zeno SWATH DIA

Various applications, such as single-cell proteomics, require the utmost sensitivity levels with low amounts of samples analyzed. To demonstrate the performance of Zeno SWATH DIA at such

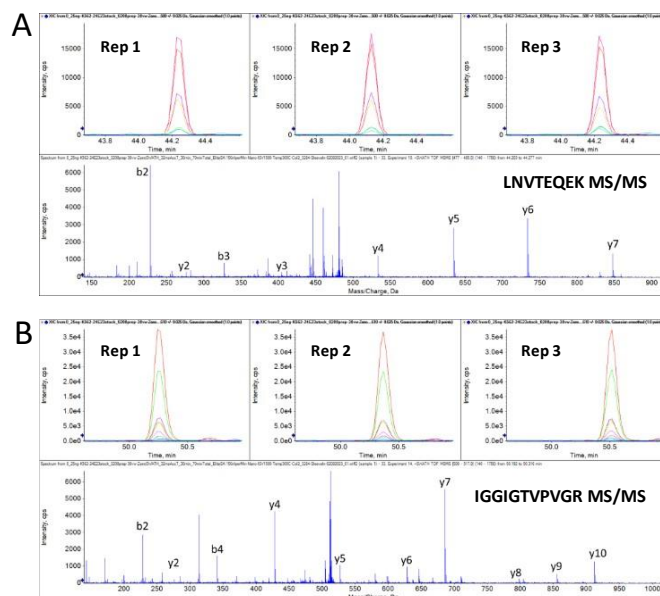


Figure 8. MS/MS sensitivity at 250 pg on-column loadings of K562 digest using Zeno SWATH DIA. Two representative peptides from K562 tryptic digest (A, LNVTEQEK from ENO_HUMAN; B, IGGIGTVPVGR from EF1A3_HUMAN) are shown. The top panes for both A and B show the overlaid XICs of the full y-ion fragment series for each peptide across three replicate injections. The bottom panes for both A and B show the Zeno SWATH DIA MS/MS spectra for each peptide, with major fragment ions indicated. Excellent MS/MS spectral quality and reproducibility across technical replicates were observed at this load level.

Protein	Peptide	Precursor	Fragment	Rep 1	Rep 2	Rep 3	Mean	%CV
ENO_HUMAN (P06733)	LNVTEQEK	480.754 (2+)	y6 (733.373)	8.22E+04	7.79E+04	8.10E+04	8.04E+04	2.7
			y5 (634.304)	7.16E+04	7.12E+04	7.40E+04	7.23E+04	2.1
			y8 (847.416)	3.41E+04	3.25E+04	3.13E+04	3.26E+04	4.3
EF1A3_HUMAN (Q5VTE0)	IGGIGTVPVGR	513.309 (2+)	y7 (685.399)	2.47E+05	2.36E+05	2.35E+05	2.40E+05	2.7
			y4 (428.262)	1.55E+05	1.52E+05	1.56E+05	1.54E+05	1.1
			y10 (912.526)	5.10E+04	4.68E+04	4.67E+04	4.81E+04	5.1

Figure 9. Fragment ion peak area reproducibility at 250 pg on-column loadings of K562 tryptic digest using Zeno SWATH DIA. The XIC peak areas for representative fragment ions of two peptides from K562 digest (LNVTEQEK from ENO_HUMAN; IGGIGTVPVGR from EF1A3_HUMAN, see Figure 8) are summarized. The mean and %CV for each fragment ion XIC area were calculated across the three replicate injections.

low loadings, a series of serially diluted K562 digest was analyzed with 30-minute nanoflow gradients using an IonOpticks Aurora EliteSX column. Two different Zeno SWATH DIA methods were tested, varying in the number of variable windows (vw): a 38 vw Zeno SWATH DIA method using 32 millisecond accumulation times per MS/MS spectrum, versus an 85 vw Zeno SWATH DIA method using 18 millisecond accumulation times per MS/MS spectrum. The resulting protein group identifications are shown in Figure 5, while the precursor identifications are summarized in Figure 6. The figures indicate that at lower column loadings (≤ 1 ng on column), the 38 vw Zeno SWATH DIA method results in increased numbers of identified protein groups, and with a higher percentage of those protein groups having CVs $<20\%$. At 250 pg loadings, 2,361 protein groups were identified (1,587 protein groups with CV $<20\%$) using this method. At higher on-column loadings (≥ 5 ng), the 85 vw Zeno SWATH DIA method resulted in higher numbers of protein groups identified. Although the 38 vw Zeno SWATH DIA method yielded lower protein group identifications in these cases, the proportions of identified protein groups with CV $<20\%$ were still higher, showing that the longer accumulation times per MS/MS spectrum provided better overall quantitative data quality. This fact is also shown in Figure 7, where the %CV distributions for the lower column loadings (250 pg, 500 pg and 1 ng) are plotted for both Zeno SWATH DIA methods. The median CVs improve in all cases as the on-column loadings increase. However, the 38 vw Zeno SWATH DIA method provides a clear improvement in overall quantitative performance at all loadings, coupled with the protein group identification gains at these low loads.

To further enhance the confidence of the identifications at low load levels, extracted ion chromatograms (XICs) of fragment ions for two different peptides (LNVTEQEK from ENO_HUMAN protein, and IGGIGTVPVGR from EF1A3_HUMAN protein) are shown in Figure 8. The XICs for the entire y-ion series are shown for both peptides across the three replicates, showing excellent reproducibility. Both peptides' MS/MS spectra are also shown, with the major y- and/or b-ion fragments highlighted. It should be noted that neither peptide was detected in blank injections run immediately before the dilution series. In fact, DIA-NN

processing of the blank injections did not reveal any protein or peptide matches (data not shown). The quantitative reproducibility of the peak areas from the fragment ion XICs is summarized in Figure 9. The top 3 fragment ion XIC peak areas for each peptide are summarized in the figure. The mean peak areas were calculated for each fragment ion across the 3 replicate injections, and the %CV values are shown. The CV values are all $\leq 5\%$ for all transitions, highlighting the quantitative capabilities of Zeno SWATH DIA even at these extremely low levels.

Conclusions

- $>19,000$ protein groups ($>145,000$ precursors) were identified from 500 ng of a human/yeast/*E. coli* digest mixture using IonOpticks Aurora SX series nanoflow columns and Zeno SWATH DIA
- Excellent quantitative performance was achieved with mixtures of varying species, both at the reproducibility level (i.e. proportion of protein groups and precursors with CVs $<20\%$) as well as the protein ratios calculated between mixture samples using label-free quantitation with Zeno SWATH DIA
- $\approx 8,400$ protein groups ($\approx 80,000$ precursors) were identified from 200 ng of human K562 digest using IonOpticks Aurora nanoflow columns and Zeno SWATH DIA, with $>95\%$ of the protein groups having CVs $<20\%$
- $>2,300$ protein groups ($>10,000$ precursors) were identified from 250 pg of human K562 digest using 30-minute nanoflow gradients with Zeno SWATH DIA, highlighting the sensitivity of the ZenoTOF 7600 system for single-cell proteomics applications

References

1. Flexibility, speed, and throughput for high proteome coverage using Zeno SWATH data-independent acquisition (DIA) coupled with the Evosep One system. [SCIEX technical note, RUO-MKT-02-15461](#).
2. Quantifying 1000 protein groups per minute of microflow gradient using Zeno SWATH DIA on the ZenoTOF 7600 system. [SCIEX technical note, RUO-MKT-02-15429-A](#).
3. Assessment of ZenoTOF 7600 system robustness for quantitative proteomics workflows. [SCIEX technical note, MKT-28411-A](#).
4. Label-free protein quantitation of protein mixtures using Zeno SWATH DIA. [SCIEX technical note, MKT-28641-A](#).
5. Demichev V et al. (2019) DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. [Nature Methods, 17, 41-44](#).
6. Large-scale protein identification using microflow chromatography on the ZenoTOF 7600 system. [SCIEX technical note, RUO-MKT-02-14415-A](#).
7. Cox, J et al. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. [Mol. Cell. Proteomics 13, 2513-2526](#).

The SCIEX clinical diagnostic portfolio is For In Vitro Diagnostic Use. Rx Only. Product(s) not available in all countries. For information on availability, please contact your local sales representative or refer to <https://sciex.com/diagnostics>. All other products are For Research Use Only. Not for use in Diagnostic Procedures.

Trademarks and/or registered trademarks mentioned herein, including associated logos, are the property of AB Sciex Pte. Ltd. or their respective owners in the United States and/or certain other countries (see www.sciex.com/trademarks).

© 2023 DH Tech. Dev. Pte. Ltd. MKT-29730-A



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