Impacts of nanoflow column length and core-shell based chromatographic phases on protein and peptide identifications in DDA workflows

Alexandra Antonoplis[1], Jason Anspach[2], Christie Hunter[1]
[1]SCIEX USA, [2]Phenomenex, USA

INTRODUCTION

Data dependent acquisition (DDA) is one of the main discovery workflows in mass spectrometry-based proteomics for the identification of peptides in digested protein samples. During data acquisition, a survey TOF MS scan is first performed, and the highest-ranking peptide ions that pass user-specified criteria in each MS cycle are sent for MS/MS analysis. Database search algorithms enable identification of the peptides in the sample based on the MS/MS fragmentation patterns as well as inference of the original digested proteins. Nanoflow chromatography is often used in proteomics DDA workflows order to obtain the highest sensitivity. High quality chromatographic separations are critical to success, as peak shape and resolution are critical to allow the MS system to sample as many unique peptides as possible. Here, new long nanoflow columns packed with core-shell chromatographic phase were evaluated for the impact on protein identification workflows.

MATERIALS AND METHODS

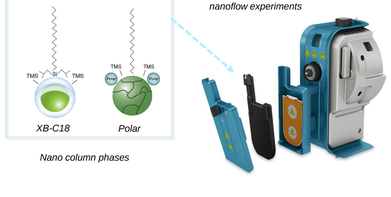
Sample preparation: A 500 ng/µL solution of digested K562 human cell lysate was prepared in water with 0.1% formic acid. The amount loaded on column for all experiments was 1 µg. PepCalMix, a mixture of 20 heavy-labeled peptides, was used to track instrument performance throughout the study, with 20 fmol loaded on column.

Chromatography: In the current study, three different nanoflow column lengths (15, 25, and 50 cm) were packed with core-shell Kinetex C18 phase and used in DDA workflows (Figure 1). The particle sizes were 2.6 µm for the 15 and 25 cm column lengths and 5.0 µm for the 50 cm column. Normal porous C18 phase columns with 15 and 25 cm lengths were also tested. Several linear gradients ranging from 60 to 180 minutes in duration were evaluated across the column types for DDA workflows. A nanoLC 425 system was used for all liquid chromatography in trap-elute configuration. Mobile phase A was water with 0.1% formic acid and mobile phase B was acetonitrile with 0.1%. Samples were trapped at 2 µL/min for 10 min at 100% mobile phase A. A flow rate of 300 µL/min was used for all analytical gradients tested.

Mass spectrometry: Data dependent acquisition data was collected using a SCIEX TripleTOF 6600 system equipped with an OptiFlow ion source. For the TOF MS survey scan, the mass range was 400-1250 Da with an accumulation time of 250 ms. For TOF MS/MS scans, the mass range was 100-1500 Da with an accumulation time of 50 msec, and high sensitivity mode was used. Only precursors with charge states between 2-5 with intensities greater than 150 cps were selected for fragmentation. The maximum number of candidate ions per cycle was 30, and the mass tolerance was 100 ppm. An exclusion time of 15 msec was used and rolling collision energy was implemented for fragmentation.

Data processing: Cloud-based data processing was used with the ProteinPilot app in the OneOmics suite. All searches performed were thorough searches, with >800 biological modifications included. The resulting proteins and peptides identified at <1% global FDR were used in data analysis.

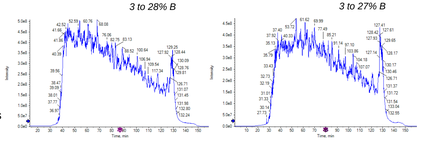
Figure 1. Overview of nano column testing for protein and peptide identification workflows. Nanoflow columns in cartridge format ranging from 15-50 cm in length were used in proteomics experiments. Two different phases (core-shell Kinetex XB-C18 (inset, left), and normal porous nano polar C18 (inset, right) were tested. The cartridges integrate with the column heater in the OptiFlow source. The nanoLC 425 system and TripleTOF 6600 system were used for all testing, with data processing performed using the ProteinPilot app in the OneOmics suite.



OPTIMIZATION OF LC GRADIENTS FOR DDA WORKFLOWS

Prior to collection of protein identification data, LC gradients were adjusted for each of the four durations and columns tested to achieve even elution of peptides during DDA experiments. The shape of the TIC was used to monitor elution of peptides. Linear LC gradients were used in all testing, with the %B range adjusted and effects on the TIC monitored (Figure 2). The pressure traces for each of the column lengths tested were also monitored (Figure 3).

Figure 2. Example gradient optimization for a 120 min gradient with 50 cm Kinetex column. The upper %B limit was optimized for each gradient duration to maximize distribution of peptides eluting across the gradient. Left: Asymmetric elution of peptides was observed with a 3-28%B linear gradient. The gradient was adjusted to 3 to 27% B for further DDA experiments.



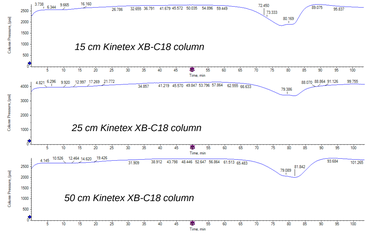
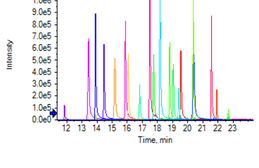


Figure 3. Sample pressure traces for each of the column lengths tested. The analytical pressure traces were monitored during experiments for each of the three column lengths tested. For the 15 cm Kinetex XB-C18 column (2.6 µm particle size), pressures remained less than 3000 psi for all experiments. For the 25 cm column (2.6 µm particle size, the analytical pressure increased, maximizing around 4000 psi. The 50 cm column exhibited reduced pressures due to its 5 µm pore size. The maximum pressure observed with the 50 cm column in experiments was ~2500 psi.

To track LC-MS performance throughout the study, PepCalMix, a mixture of 20 heavy labeled peptides, was used. 20 fmol of PCM was injected on column and both MS peak areas and MS/MS peak areas were monitored to ensure sensitivity and detection of early eluting and late eluting peptides. The PCM mixture was also used for TOF MS calibration throughout the study.

Figure 4. PepCalMix nanoflow system suitability test was used to track LC-MS performance. Viewing extracted ion chromatograms for PepCalMix enables monitoring of peak shapes and elution of the 20 heavy labeled peptides in the sample. TOF MS and MS/MS peak areas were also tracked according to the nanoflow PepCalMix system suitability test from the SWATH acquisition performance kit protocol.[1] A sample PCM XIC for a 15 cm Kinetex X-B column is shown to the right.



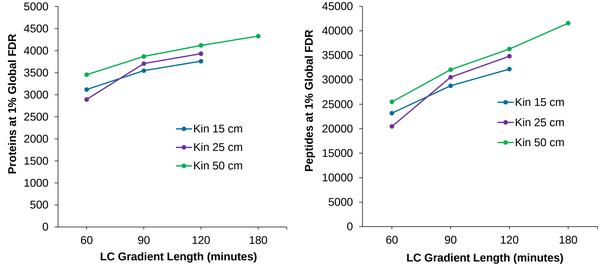


Figure 5. Protein and peptide Identifications at 1% FDR using Kinetex XB-C18 columns. Left: Protein identifications at 1% FDR were plotted as a function of gradient length according to column length tested. Right: Peptide identifications at 1% FDR were plotted as a function of gradient length. All protein and peptide identifications at 1% FDR were determined using the ProteinPilot app in OneOmics suite, with each data point representing an average of three technical replicates. A 180 min gradient length was tested with the 50 cm column only.

PROTEIN AND PEPTIDE IDENTIFICATIONS AS A FUNCTION OF EACH LC GRADIENT LENGTH AND COLUMN TESTED

Several different nanoflow columns were evaluated in combination with optimized gradient lengths for DDA workflows (Figures 5 and 6). Protein and peptide identifications increased as the gradient length increased and column length increased, with optimal identifications observed with a 180 min gradient and the 50 cm column length (4329 proteins and 41,448 peptides at 1% FDR with a 1 µg K562 digest load). Interestingly the core-shell packed columns provided more protein and peptide identifications at all gradient lengths, with average increases in protein IDs of 10-20% depending on column length.

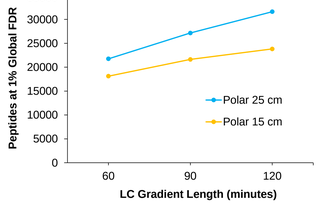


Figure 6. Evaluation of polar nanoflow columns for proteomics DDA workflows. Two polar column lengths were tested, 15 and 25 cm, in nanoflow DDA protein identification experiments with gradient durations ranging from 60-120 minutes. The 25 cm polar column resulted in increased protein and peptide identifications as compared to the 15 cm column. As the gradient duration increased, protein and peptide identifications also increased. The maximum number of peptides observed was 31,624 with the 25 cm column and 120 min gradient. Each data point represents an average of three technical replicates

CLOUD-BASED DATA PROCESSING

The ProteinPilot app in OneOmics suite enables rapid processing and visualization of nanoflow DDA results. The application includes a dashboard of tables and plots illustrating protein identifications at 1% FDR, peptide identifications @ 1% FDR, along with chromatography-related metrics. These plots can be used during a method optimization study to assess the performance of nanoflow columns in DDA workflows and guide further experimentation

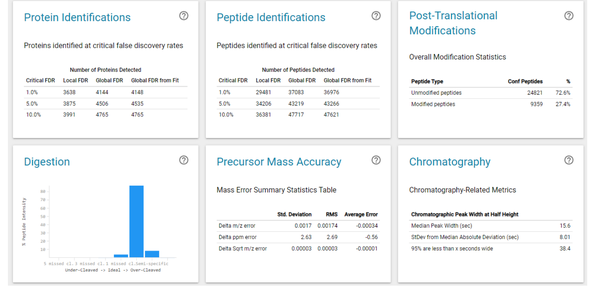


Figure 7. The ProteinPilot app in OneOmics suite facilitates nanoflow DDA method optimization. The performance of nanoflow columns was evaluated by determining protein and peptide identifications at 1% FDR. Here a sample dashboard is shown for the 50 cm Kinetex column with a 120 min gradient.

CONCLUSIONS

Nanoflow chromatography is used when highest sensitivity is required and is often done with long columns and long gradients. As columns get longer, back pressure can increase, making the columns more difficult to use routinely and robustly. The core-shell based phases were developed in order to provide higher resolution and better efficiencies at higher flow rates. These phases produce lower backpressure than traditional phases, and in this study back pressure did not exceed 4500 psi for any of the columns tested.

Here, the impacts of column phase, column length and gradient duration on protein and peptide identifications were assessed. As expected, protein and peptide identifications increased as the gradient length increased and column length increased, with optimal identifications observed with a 180 min gradient and 50 cm column length

REFERENCES

1. SWATH acquisition performance kit protocol RUO-IDV-05-3363-A.

TRADEMARKS/LICENSING