P-II-0437

Versatile and bidirectional use of microfluidic LC columns, optimized configuration for achieving significant increases in proteome coverage at low nanoLC flow rates

Jeff Op de Beeck[1],Natalie Van Landuyt[1], Robert Van Ling[2], Bo Claerebout[1], Paul Jacobs[1],
[1]Thermo Fisher Scientific, Ghent, Belgium, [2]Thermo Fisher Scientific, Breda, The Netherlands

Abstract

Purpose: Evaluate the impact of using an optimized LC-MS set-up with reduced post column volumes for both routine nanoflow as for high sensitivity low nanoflow proteomics experiments.

Methods: Dilution series of human cell lysate protein digest standards were separated using nanoflow LC on a Thermo Scientific™ 50 cm µPAC™ Neo High column hyphenated to Orbitrap MS for detection. Data was acquired in data independent acquisition.

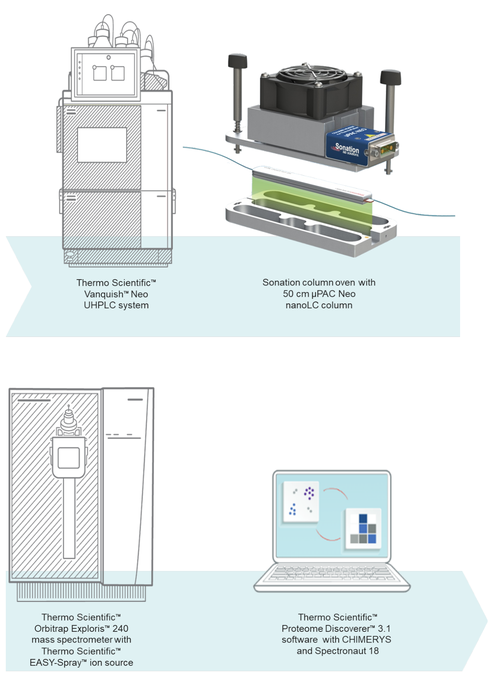
Results: Reduction of median eluting peak widths by 20% resulted in an increase in proteome coverage of up to 27%.

Introduction

LC-MS-based proteomics relies on HRAM mass spectrometers for resolution, speed, and sensitivity, critical for accurate identification and quantification after LC separation. To match the increasing speed of MS instruments, nanoLC column development has focused on reducing dead volume and ensuring optimal elution, particularly at ultra-low flow rates. Pillar array technology, with its precise order and high permeability, demonstrates outstanding performance, although connectivity quality can significantly impact results. An optimal strategy is proposed to enhance connectivity and diminish post-column dead volumes, resulting in notable performance enhancements and enabling more thorough proteomic analysis. Moreover, its unique capabilities facilitate high-throughput, ultra-sensitive analysis by swiftly adapting to programmed flow rates, enabling 100 samples per day at elution rates as low as 100 nL/min

Materials and methods

Figure 1. Experimental set-up



Sample preparation

Mammalian cell digests were resuspended in either 0.1 % TFA or 0.1% TFA 1% DMSO to obtain stock solutions of respectively 200 ng\µl and 25 ng/µL.

LC-MS configuration

The separation was performed on a Vanquish Neo UHPLC instrument using a 50 cm long µPAC Neo column in a direct injection workflow. The LC column was positioned in an external heating device (Sonation, PRSO-V2_PF) directly mounted onto a EasySpray source that was coupled to a Orbitrap Exploris 240 mass spectrometer. Variable flow rate LC methods with initial flow at 750 nL/min and eluting flow rates of 250 (routine methods) and 100 nL/min (low input methods) were used. Data independent acquisition (DIA) parameters used are listed below.

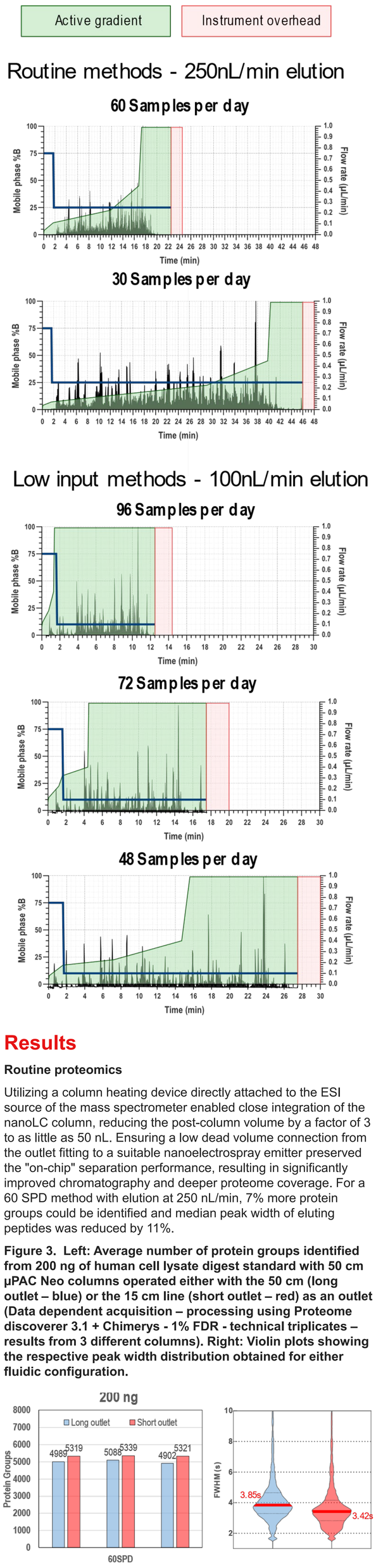
Table 1. MS acquisition settings

	Full Scan		MS² scan			
	Resolution	Max IT	**Resolution**	**Precursor mass range**	Isolation window	Max IT
Routine	45k	Auto	15k	525-825 m/z	6 Th	Auto
Low input	120k	Auto	60k	375-6755 m/z	30 Th	118 ms

Data analysis

LC-MS data were analyzed either using Proteome Discoverer 3.1 software with Chimerys or with Spectronaut® 18. Results shown have been filtered to a 1% FDR.

Figure 2. Gradient and flow rate profiles used for routine (Top) and low input (Bottom) LC-MS evaluation. 60 and 30 sample per day (SPD) throughput with elution at 250 nL/min, basepeak chromatograms of 100 ng human cell lysate digest overlaid. 96, 72 and 48 samples per day throughput with elution at 100 nL/min, basepeak chromatograms of 1 ng human cell lysate digest overlaid.



Results

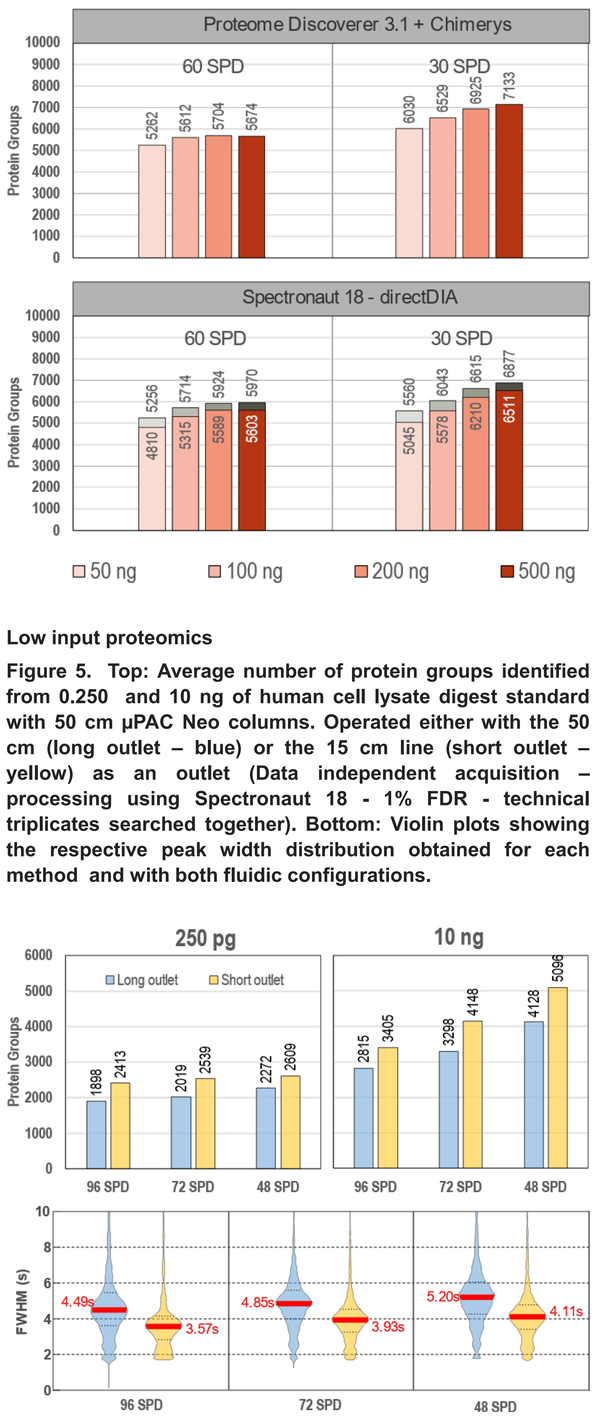
Routine proteomics

Utilizing a column heating device directly attached to the ESI source of the mass spectrometer enabled close integration of the nanoLC column, reducing the post-column volume by a factor of 3 to as little as 50 nL. Ensuring a low dead volume connection from the outlet fitting to a suitable nanoelectrospray emitter preserved the "on-chip" separation performance, resulting in significantly improved chromatography and deeper proteome coverage. For a 60 SPD method with elution at 250 nL/min, 7% more protein groups could be identified and median peak width of eluting peptides was reduced by 11%.

Figure 3. Left: Average number of protein groups identified from 200 ng of human cell lysate digest standard with 50 cm µPAC Neo columns operated either with the 50 cm (long outlet – blue) or the 15 cm line (short outlet – red) as an outlet (Data dependent acquisition – processing using Proteome discoverer 3.1 + Chimerys - 1% FDR - technical triplicates – results from 3 different columns). Right: Violin plots showing the respective peak width distribution obtained for either fluidic configuration.

We explored the benefits for conventional bottom-up proteomics sample loads (50-500 ng of tryptic digest injected) using data independent acquisition and sample throughput rates of 60 and 30 SPD. At 60 SPD, 5970 protein groups could be identified from technical triplicates. And out of these, 94% (5603) could be accurately quantified with a coefficient of variation below 20%. By extending the gradient length and operating at a slightly reduced throughput rate of 30 SPD, we achieved an increased coverage of 6877 protein groups, again with excellent quantification metrics (95% - 6511 proteins quantified with CV ≤ 20%).

Figure 4. Proteins identified and quantified at CV ≤ 20% from human cell lysate digest standards using LC methods with respective sample throughput of 60 and 30 samples per day. Top: Average number of protein groups identified using Proteome discoverer 3.1 and Chimerys. Bottom: Average number of protein groups identified and quantified at CV ≤ 20% using a directDIA approach in Spectronaut 18.



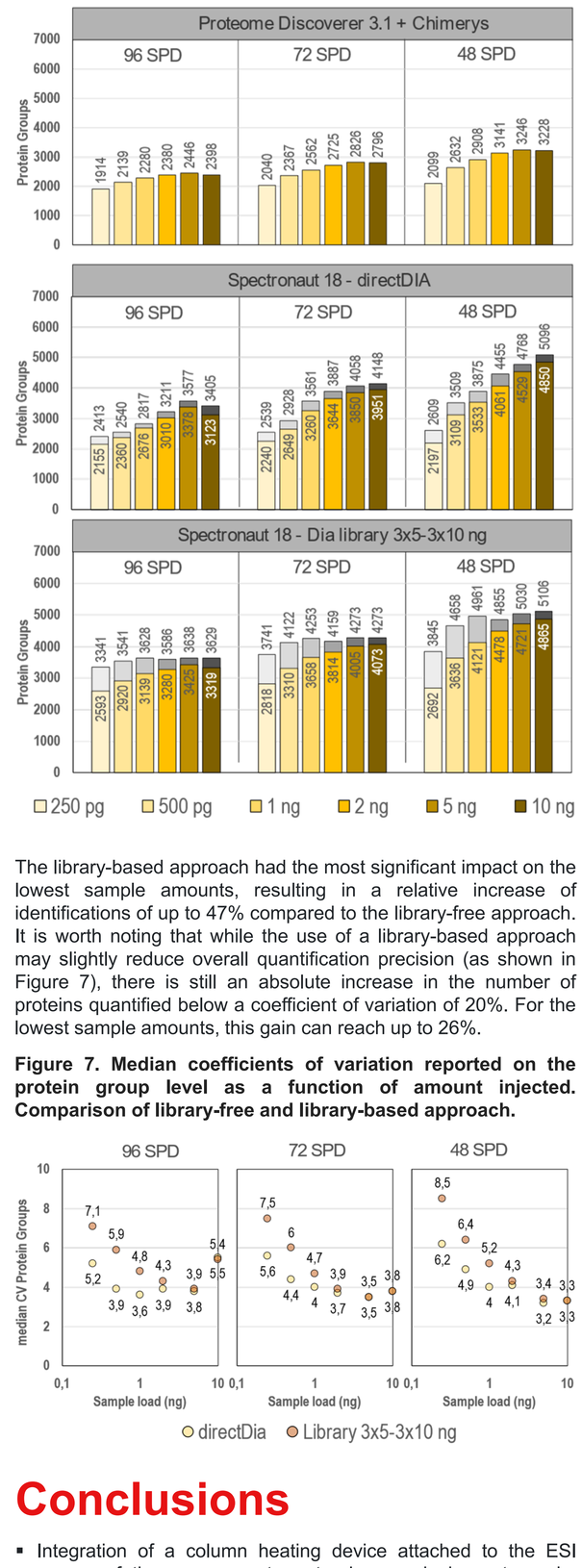
Low input proteomics

Figure 5. Top: Average number of protein groups identified from 0.250 and 10 ng of human cell lysate digest standard with 50 cm µPAC Neo columns. Operated either with the 50 cm (long outlet – blue) or the 15 cm line (short outlet – yellow) as an outlet (Data independent acquisition – processing using Spectronaut 18 - 1% FDR - technical triplicates searched together). Bottom: Violin plots showing the respective peak width distribution obtained for each method and with both fluidic configurations.

Anticipating greater performance gains at lower flow rates, we conducted benchmarking with sample loads ranging from 250 pg to 10 ng. Methods were optimized for elution at 100 nL/min, achieving sample throughput rates of 96, 72, and 48 samples per day with instrument productivities of 70%, 77.5%, and 85%, respectively. Post-column volume reduction led to a performance increase of up to 20% (reducing average FWHM by approximately 1s) and increasing proteome coverage by up to 27% on the protein group level and identifying up to 2609 protein groups from as little as 250 pg using a library free data processing approach.

When dealing with low sample amounts, it was reported that utilizing a project (and method)-specific library generated from these limited samples produced more favorable outcomes compared to a library-free approach. By employing a spectral library-based approach, specifically utilizing a DIA library consisting of 3 x 5 ng and 3 x 10 ng DIA runs (totaling 47,236 precursors), we were able to identify 3845 protein groups from 250 pg HeLa protein digests. Furthermore, when working with sample loads above 1 ng, we successfully identified nearly 5,000 proteins using this approach.

Figure 6. Proteins identified and quantified at CV ≤ 20% from human cell lysate digest standards using LC methods with respective sample throughput of 96, 72 and 48 samples per day. Top: Average number of protein groups identified using Proteome discoverer 3.1 and Chimerys. Middle: Average number of protein groups identified and quantified at CV ≤ 20% using a directDIA approach in Spectronaut 18. Bottom: Average number of protein groups identified and quantified at CV ≤ 20% using a library-based search approach in Spectronaut 18.



The library-based approach had the most significant impact on the lowest sample amounts, resulting in a relative increase of identifications of up to 47% compared to the library-free approach. It is worth noting that while the use of a library-based approach may slightly reduce overall quantification precision (as shown in Figure 7), there is still an absolute increase in the number of proteins quantified below a coefficient of variation of 20%. For the lowest sample amounts, this gain can reach up to 26%.

Figure 7. Median coefficients of variation reported on the protein group level as a function of amount injected. Comparison of library-free and library-based approach.

Conclusions

- Integration of a column heating device attached to the ESI source of the mass spectrometer improved chromatography and proteome coverage by reducing post-column volume and preserving separation performance.
- Peak widths could be reduced by 11 and 20% at eluting flow rates of respectively 250 and 100 nL/min.
- The impact on proteome coverage was the highest for the lowest flow rates, with up to 27% increase in the number of proteins identified.
- Up to 2609 protein groups could be identified from 250 pg of HeLa digest using a library free processing approach, and up to 3845 protein groups could be identified using a project specific library constructed from 3 x 5 and 3 x 10 ng DIA runs.

Trademarks/licensing

.