Routine Single-Shot Identification of >9K Proteins and >100K Peptides with the Next-Generation Low-Flow UHPLC Coupled to HRAM MS

ThermoFisher SCIENTIFIC

Runsheng Zheng[1], Karel Stejskal[2], Bernard Delanghe[3], Christopher Pynn[1], Kean Woodmansey[4], Theresa Riley[5], Karl Mechtler[2], Alexander Boychenko[1]
[1]Thermo Fisher Scientific, Germering, Germany; [2]IMP, Vienna, Austria; [3]Thermo Fisher Scientific, Bremen, Germany; [4]Thermo Fisher Scientific, Hemel Hempstead, United Kingdom; [5]Thermo Fisher Scientific, Sunnyvale, CA

ABSTRACT

Purpose: Demonstrate the best-in-class performance for nanoLC-MS single shot bottom-up proteome profiling with Thermo Scientific™ Vanquish™ Neo next-generation low-flow UHPLC system when coupled to a Thermo Scientific™ Orbitrap Exploris™ 480 mass spectrometer using the 75 µm I.D. × 75 cm Thermo Scientific™ EASY-Spray™ PepMap Neo column.

Methods: Data-dependent acquisition for nanoLC-MS analysis of HeLa protein digest and data processing with 2-step Sequest HT and INFERYS rescoring node in Proteome Discoverer 2.5 software or spectral library search with the ability to resolve "chimeric" MS/MS spectra in Thermo Scientific™ Proteome Discoverer™ 3.0.

Results: More than 9000 protein groups and 100K peptides can be identified in single-shot nanoLC-MS analysis using 180 and 240 min gradients and 2 µg sample load

INTRODUCTION

Bottom-up research proteomics seeks to both identify and quantify the complete proteome within a cell, tissue, or organism. As the depth of proteome profiling coverage increases, so does our insight into complex physiological processes and their effect on phenotype, potentially leading to advancements in fields including biomarker discovery and precision medicine. The limitations of complex proteome quantitative profiling are linked to high complexity and a wide dynamic range of proteins that require better separation and detection. Here we demonstrate the latest advances in nanoLC-MS technology and data processing pipelines used to enhance peptides separation for the deepest label-free proteome profiling in HeLa cell protein digest.

MATERIALS AND METHODS

Sample Preparation

Thermo Scientific™ Pierce™ HeLa Digest/PRTC Standard (A47996, 10 µg/vial) was reconstituted by adding 50 µL of 0.1% formic acid (FA) in water. The vial was subsequently sonicated for 2 min, followed by aspirating and releasing 10-times with a pipette to fully reconstitute the sample. The sample was subsequently transferred to another vial of HeLa Digest/PRTC Standard (A47996), which was again sonicated for 2 mins and mixed in the same way to produce a final sample concentration of 400 ng/µL HeLa containing 200 fmol/µL PRTC.

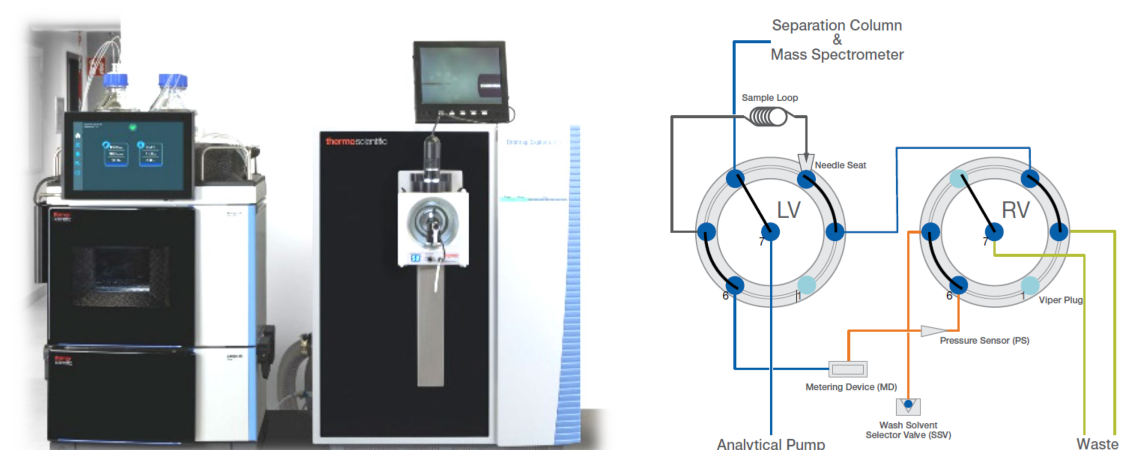
MS acquisition parameters

MS data were recorded with an Orbitrap Exploris 480 mass spectrometer in DDA mode after peptides separation with Vanquish Neo UHPLC system (Figure 1). An example of the MS data acquisition TIC profiles is shown in Figure 2. All MS acquisition settings are available for download in the AppsLab library for all the methods reported here.

Data Analysis

Data were acquired using the SII for Xcalibur software package version 1.5.1 in data-dependent acquisition (DDA) mode, followed by .raw file processing with Thermo Scientific™ Proteome Discoverer™ 2.5 software using a 2-step Sequest™ HT search algorithm and INFERYS rescoring node. The data were also searched with the beta version of the software enabling the resolution of chimeric spectra. The false discovery rate (FDR) was set below 1% at the peptide and the protein level.

Figure 1. Vanquish Neo system coupled with the Orbitrap Exploris 480 mass spectrometer and scheme of direct injection workflow

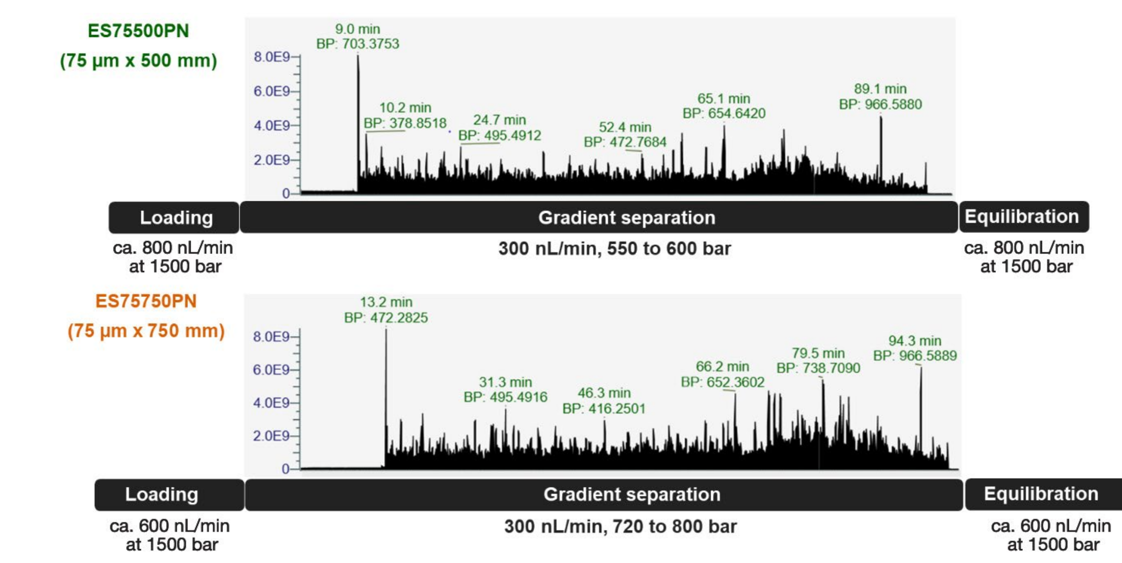


RESULTS

Enhancing separation with 75 cm nanoLC columns

Commercially available 75 µm I.D × 50 cm columns are commonly associated with discovery proteomics applications on nanoLCMS platforms because they deliver high peak capacity in single-shot proteome profiling for both direct injection as well as trap-and-elute workflows. The 75 µm I.D. × 75 cm columns, on the other hand, have hitherto received less acclaim despite their capacity for deeper proteome profiling. We found that the 75 µm I.D x 75 cm column outperforms the 75 µm I.D. 50 cm in a 90-min gradient by reducing the FWHM by 2 seconds for 50% of the peptide peaks. the employment of the long column on the Vanquish Neo UHPLC system affords distinct performance advantages for routine bottom-up proteomics research without compromising MS utilization time.

Figure 2. The Vanquish Neo UHPLC system enables fast sample loading and equilibration on 75 µm I.D. x 50 cm and 75 µm I.D. x 75 cm columns



Optimizing nano-flow rate for maximum performance

The wide flow-pressure footprint and reproducible gradient delivery of the Vanquish Neo system permits the study of the influence of changing flow rates on ESI-MS sensitivity as well as upon overall LCMS performance in nanoLC proteomics experiments. The backpressure capabilities of the system make it possible to deliver flow rates of up to 500 nL/min on 75cm long columns where the pressure reaches 1450 bar during the gradient and 1500 bar during sample loading and column equilibration. The improved ionization efficiency in ESI-MS afforded by the lower flow rates results in up to a 40% sensitivity gain (Figure 3). The increase in MS1 intensity, however, had no significant impact on either the number of MS/MS events, or the number of peptide-spectrum matches (PSMs), or peptide, and protein identifications (Figure 4). This suggests that the MS1 intensity levels required for the maximum possible protein coverage were already met at the elevated nano-flow rates thanks to the high sensitivity of Orbitrap Exploris 480 mass spectrometer under the separation conditions employed in these experiments.

Figure 3. The typical nanoLCMS profiles for HeLa protein digest separated on 75 µm I.D. . 75 cm column with flow rates from 200 to 500 nL/min and pressure during the gradient separation from 560 to 1450 bar

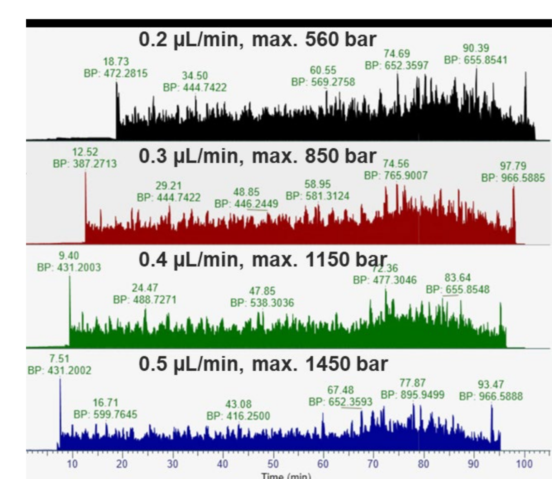
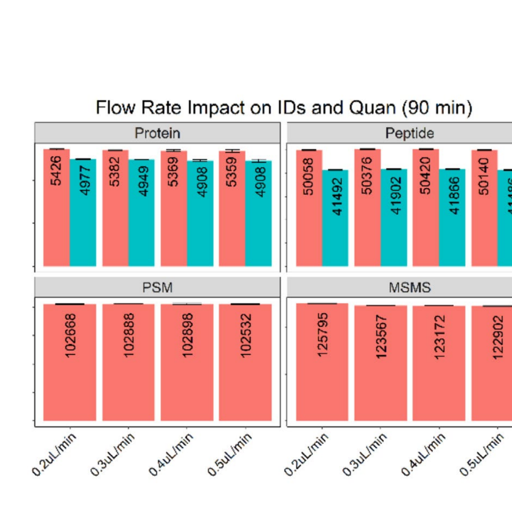


Figure 4. The effect of flow rate from 0.2 to 0.5 µL/min on the number identified and quantified peptides and proteins with 90-min gradient and 1 µg of HeLa protein digest.



LONG GRADIENTS FOR MAXIMUM IDS

Single shot nanoLC-MS analyses require the highest possible separation efficiency to maximize the number of identified peptides and proteins. The increase of the gradient length by using long columns provides higher peak capacity, the theoretical number of fully resolved peaks, as it is proportional to gradient time divided by peak width. the sensitivity losses incurred through the corresponding peak broadening with increased gradient length can be compensated by loading larger sample amounts in order to increase peak height (Figure 6). The detrimental effects of column overloading on FWHM, however, means that there are limits to the sensitivity improvements that can be practically achieved (Figure 6). Interestingly no significant improvement of peptide and protein IDs was observed with increased loading amount (Figure 5).By employing a 240-min gradient and loading 1 µg HeLa protein digest, more than 7,100 proteins (1% FDR) with ca. 80K peptides were successfully quantified after separation on a 75 cm nano column (Figure 7, 8).

Figure 5.The number of peptides (A) and proteins (B) that were confidently identified and quantified using gradient length from 90 to 240 min and loading amount from 1 to 4 µg.

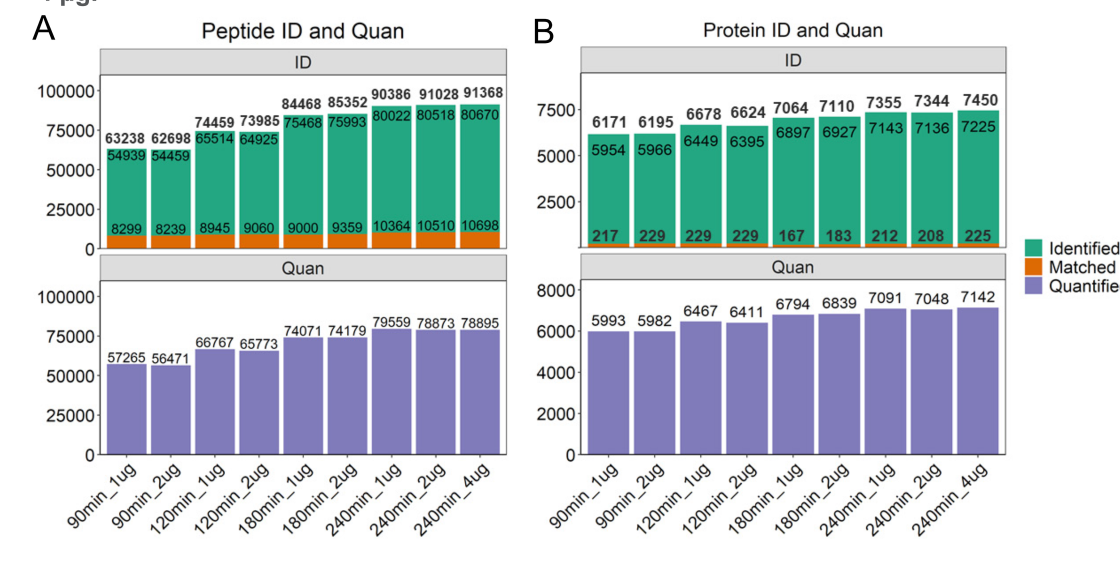


Figure 6. The dependency of FWHM and peak height on the gradient length using the constant flow rate (250 nL/min) and different loading amounts.

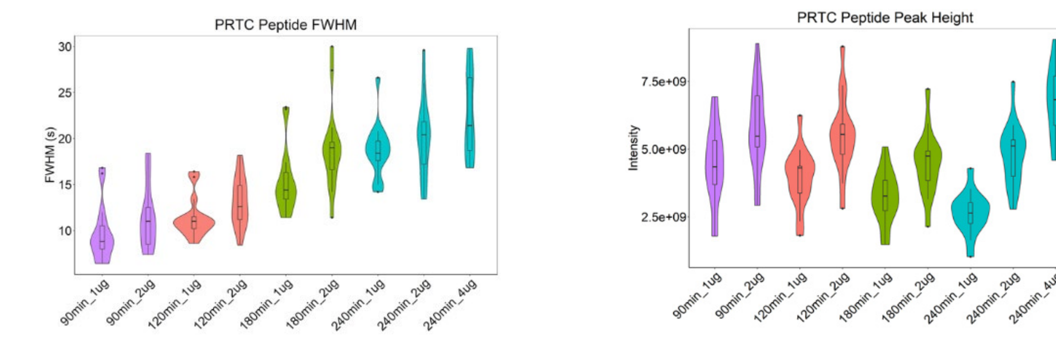
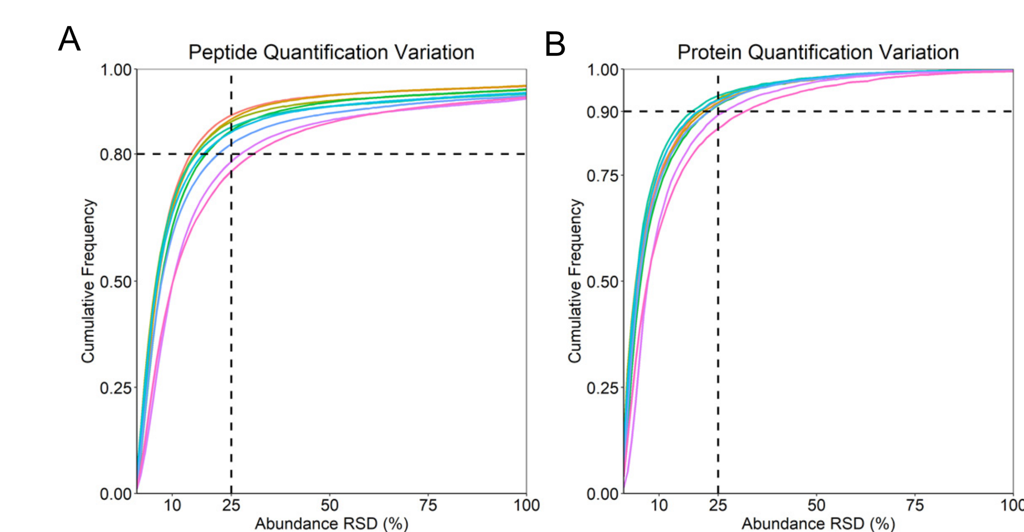


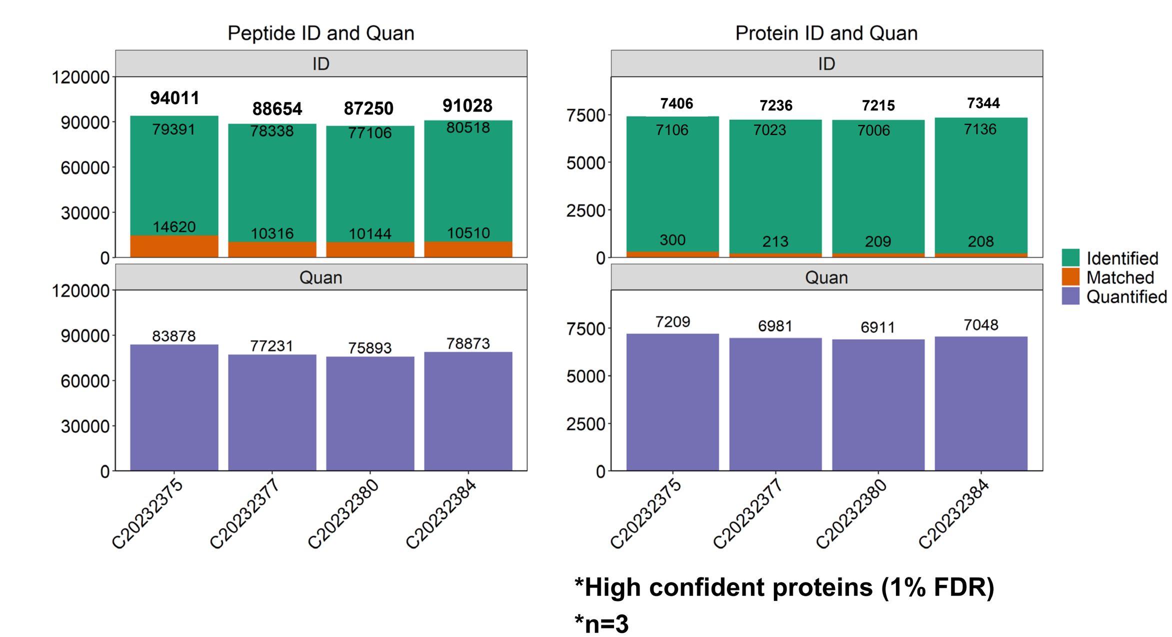
Figure 8. The cumulative frequency of peptides (A) and proteins (B) that were quantified with the specified level of variation measured as RSD, % vs. gradients and sample amount. Around 80% of peptides and 90% proteins are quantified with RSD <25% over the different



REPRODUCIBLE TOP PERFORMANCE

The reproducibility of results obtained with multiple columns is essential for the analysis of large sample cohorts. It is also critical for continuous quality control of results and for monitoring of the system performance over time. The higher packing pressure conditions adopted for PepMap Neo columns permits reproducible elution profiles for complex peptide digests over multiple columns, thus affording a high level of data reproducibility. We evaluated the column impact on HeLa proteome coverage under identical analysis conditions. The four 75 µm I.D. x 75 cm columns were used for HeLa proteome profiling with a 4-hour gradient (240 min). The column-to-column variation for peptide and protein identifications was below 5% and 1%, respectively (Figure 9). Similar results were obtained with the number of peptides and proteins that can be quantified in each sample (Figure 9). Thus, considering the reproducible performance on LC and MS hardware PepMap Neo columns showed excellent column-to-column reproducibility even for very long gradients and operation at a maximum pressure of 1500 bar during sample loading and fast column equilibration.

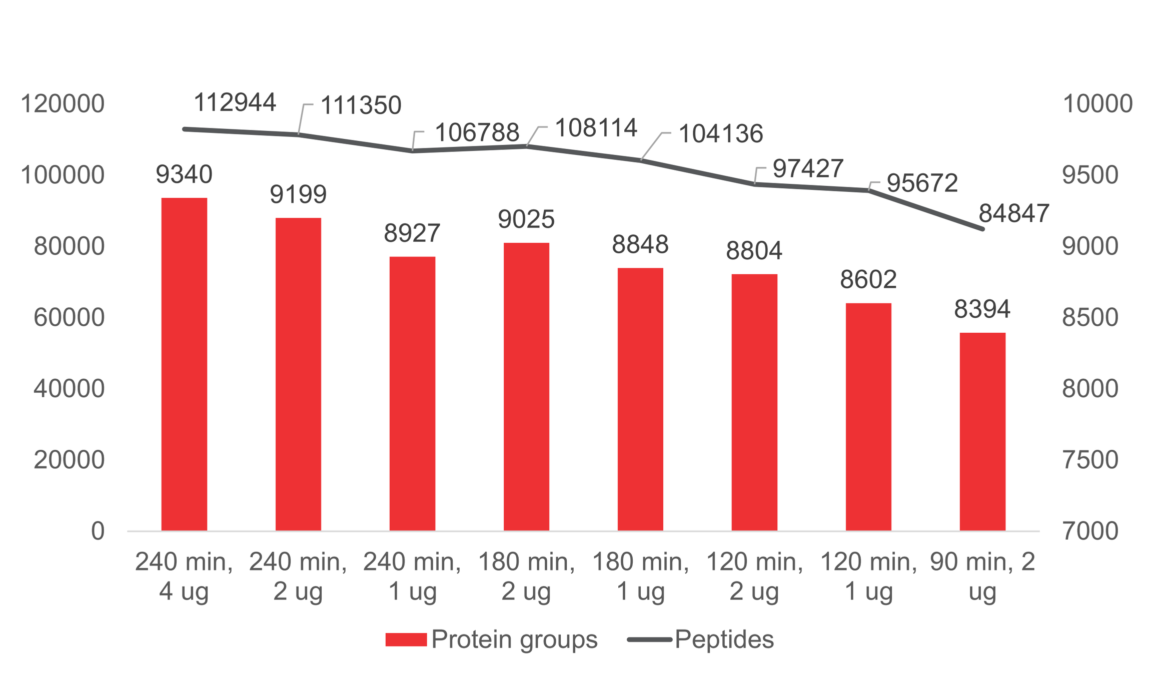
Figure 9. Reproducible identification and quantification of HeLa peptides and proteins over 4 EASY-Spray PepMap Neo columns while using Vanquish Neo UHPLC system coupled with the Orbitrap Exploris 480 mass spectrometer.



Chimeric search results in deepest proteome profiling

The reprocessing of obtained results with spectra library search and resolution of chimeric spectra significantly boosted the number of peptide and protein identifications and allowed to identify above 9000 protein groups and 100K peptides in a single run with 180 min and 240 min gradients (Figure 10).

Figure 10. The results of single shot HeLa proteome DDA profiling w/o match between runs obtained after data processing using spectra library and chimeric spectra resolving



CONCLUSIONS

Experiments employing nanoLCMS-based analytical systems continue to play a pivotal role in deep-dive discovery proteomics. Here we evaluated the performance of the Vanquish Neo UHPLC system coupled with an Orbitrap Exploris 480 mass spectrometer for bottom-up proteome profiling. The Vanquish Neo boosts chromatographic performance while employing long columns and long gradients, permits maximal MS utilization even for direct injection workflows, and provides versatility to separate peptides at low or elevated flow rates to optimize ESI-MS sensitivity. Furthermore, its ultra-high pressure capacity creates flexibility to explore the separation power of ultra-long columns and achieve deep proteome profiling in a single-shot DDA run.

REFERENCES (if necessary)

1. Zheng, R. et al. Vanquish Neo UHPLC system sets new performance standards for single-shot nanoLCMS bottom-up proteomics. TN74152 (2021)

TRADEMARKS/LICENSING

© 2021 Thermo Fisher Scientific Inc. All rights reserved. Sequest is a registered trademark of the University of Washington. All other trademarks are the property of Thermo Fisher Scientific and its subsidiaries. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.

PO66094 EN0921S

thermo scientific