thermo scientific

POSTER NOTE 72267

Deep and reproducible human proteome profiling with novel Nano Flow LC technology and HRAM mass-spectrometry

Authors

Alexander Boychenko[1], Stephan Meding[1], Wim Decrop[1], Mike Baynham[2], Martin Ruehl[1], Jenny-Marie T. Wong[1], Martin Markus[1], and Remco Swart[1]

[1]Thermo Fisher Scientific, Germering 82110, Germany; [2]Thermo Fisher Scientific, Runcorn WA7 1TA, United Kingdom

ABSTRACT

The complexity of the human proteome requires highly efficient separation of peptides generated during enzymatic protein digestion. Additionally, sensitive MS detectors with a wide dynamic range are needed to find significant differences in protein abundances. In this study we combined the new Thermo Scientific™ UltiMate™ 3000 RSLCnano system with Thermo Scientific™ ProFlow™ technology with the high-resolution accurate-mass Thermo Scientific™ Q Exactive™ HF quadrupole-Orbitrap mass spectrometer to achieve high analytical depth for human cell proteome profiling. We investigated the effect of column length and protein loading amount on the number of protein and peptide groups identified in one dimensional experiments. More than 5000 proteins were identified in a single run with 240 min gradient time using a 75 cm Thermo Scientific™ EASY-Spray™ nano flow column. Further increase in protein identifications was achieved by increased sample loading. The deepness of proteome profiling was assessed. The range of emPAI scores extends over 4 orders of magnitude for DDA acquisitions. This high level of data quality enables large comparative studies for discovery of regulated peptides and proteins where several thousand proteins are simultaneously compared.

INTRODUCTION

The reliable and deep mapping of the human proteome in large sample cohorts is urgently needed. It will help to understand the relation between genomes and phenotypes, increase the efficiency of developing new biotherapeutics, and allow early stage diagnosis and disease prognosis. Label-free and stable isotope labeling quantification are popular approaches to identify differentially expressed proteins in complex biological samples. In both cases the high LCMS uptimes, reproducibility of chromatographic profiles and proteome deepness are essential parameters to obtain high quality data. We describe here a nanoLC-MS platform that combines the advantages of the novel ProFlow technology for precise nano flow control with state-of-the-art High-Resolution Accurate-Mass (HRAM) Mass Spectrometry and 75 cm long EASY-Spray columns. This nanoLC-MS platform allows confident identification of >5000 human proteins in a single run and provides high reproducibility of results between replicates.

ThermoFisher SCIENTIFIC

MATERIALS AND METHODS

Sample Preparation

Pierce HeLa Protein Digest Standard (P/N 88329) was diluted with 0.1% formic acid solution and used for all further experiments.

Liquid Chromatography

HeLa protein digest was analyzed using an UltiMate 3000 RSLCnano system with ProFlow flow meter that was coupled to a Q Exactive HF mass spectrometer (Fig. 1) and controlled with Thermo Scientific™ SII 1.3 and Xcalibur 4.0. The separations were performed on EASY-Spray C18 (2 µm, 75 µm ID) columns with 25 cm or 75 cm length packed with Thermo Scientific™ Acclaim PepMap™ particles. Flow rate was 300 nL/min on the 75 µm x 25 cm column and 250 nL/min on the 75 µm x 75 cm column. Column temperature was at 60° C. The gradient conditions are described in Table 1.

Table 1. Gradient conditions used for nanoLC-MS analysis

The loading phase was extended for 3 min when using the 75 µm x 75 cm EASY-Spray column. The sample loop was switched off-line after the gradient preparation step. Solvent A was 0.1 % formic acid (FA) in water; solvent B was 0.1 % FA in acetonitrile

	30 min		90 min		120 min		180 min	240 min
Gradient phase	**B, %**	**time**	**B, %**	**time**	**B, %**	**time**	**time**	**time**
Sample loading	**2**	0-13	**2**	0-14	**2**	0-14	0-14	0-14
Gradient preparation	**6**	13-17	**5**	14-17	**4**	14-17	14-17	14-17
Peptides separation step 1	**6-20**	17-50	**5-16**	17-80	**4-16**	17-100	17-140	17-140
Peptides separation step 2	**20-40**	50-54	**16-28**	80-110	**16-28**	100-140	140-200	140-200
Column washing	**72**	55-58	**72**	113-116	**72**	143-146	203-206	263-266
Column equilibration	**2**	59-70	**2**	119-135	**2**	149-165	209-225	269-285

Mass Spectrometry

A Q Exactive HF instrument was used for MS/MS analysis in data-dependent acquisition mode (DDA). The following MS settings were used: MS^1 resolution 60,000, AGC target: $3 \cdot 10^6$, maximum IT 50 ms, scan range 350-1500 *m/z*, MS^2 resolution 15,000, AGC target $2 \cdot 10^5$, TOP 20, isolation window 1.4 *m/z*.

MS2 maximum IT and dynamic exclusion (DE) were varied with gradient length to reflect the reduction of peak height: 30 min: IT 19 ms, DE 10 s; 90 min: IT 35 ms, DE 20 s; 120 min: IT 50 ms, DE 30 s; 180 min: IT 75 ms, DE 40 s; 240 min: IT 100 ms, DE 60 s.

Data Analysis

Thermo Scientific™ Proteome Discoverer™ 2.1 was used for data analysis and data base search with SEQUEST® HT against the human protein database SwissProt (11-05-2016). Peptides were generated from a tryptic digestion allowing for up to two missed cleavages, carbamidomethylation (+57.021 Da) of cysteine residues was set as fixed modification, and oxidation of methionine residues (+15.995 Da) and acetylation of protein N-terminus (+42.011 Da) as variable modifications. Precursor mass tolerance was set to 10 ppm and fragment mass tolerance to 0.02 Da. The matches were validated using Percolator with 1% FDR at protein and peptide level, respectively.

RESULTS

NanoLC on reversed-phase columns coupled with HRAM MS is currently the gold standard for shotgun proteomics experiments due to unprecedented sensitivity and wide elution window for peptide separation. In order to achieve high peak capacity peptides separation is done on long nano columns with high loadability. We compared the performance of EASY-Spray nano columns with 25 cm and 75 cm length for shotgun proteomics experiments using 30 to 240 min long gradients.

Figure 1. NanoLC-MS setup with EASY-Spray column and EASY-Spray source

UltiMate 3000 RSLCnano

Q Exactive HF

Chromatographic performance

We selected 6 high abundant peptides from HeLa protein digest that are equally distributed over the elution gradient to compare the separation performance of 25 and 75 cm EASY-Spray columns (Fig. 2, Table 2) in terms of peak width and peak capacity.

Figure 2. Distribution of 6 peptides from HeLa protein digest over 240 min gradient on 75 µm x 75 cm EASY-Spray column

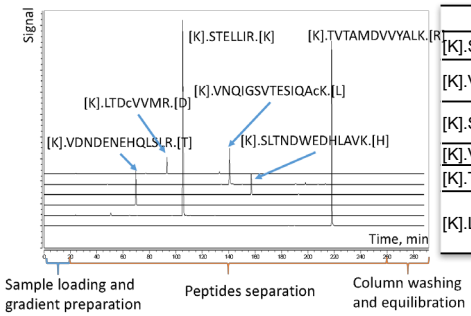


Table 2. Peptides from HeLa protein digest that were selected to compare the performance of 25 and 75 cm long EASY-Spray columns

Peptide	Protein	m/z, Da	Charge
[K].STELLIR.[K]	Histone H3.2, Q71DI3	416.252	2
[K].VDNDENEHQLSLR.[T]	Nucleophosmin, P06748	523.584	3
[K].SLTNDWEDHLAVK.[H]	Heat shock protein HSP 90-beta, P08238	509.921	3
[K].VNQIGSVTESIQAcK.[L]	Beta-enolase, P13929-1	817.417	2
[K].TVTAMDVVYALK.[R]	Histone H4, P62805	655.858	2
[K].LTDcVVMR.[D]	Heterogeneous nuclear ribonucleoproteins A2/B1, P22626	497.248	2

The average PWHM (peak width half maximum) for 6 peptides obtained on the 25 cm column was approximately 2 times higher than on the 75 cm long EASY-Spray nano column for all separation gradients (Figure 3). The significant reduction of peak width using 75 cm nano columns allows to reach a peak capacity of 700 (Fig. 4). We observed linear increase of PWHM as well as PW at base (data not shown) with gradient length that translates into the decrease of peak height (Fig. 5). At the same time peak area did not change significantly with the length of the gradient (Fig. 5). The peak height is extremely important to ensure sufficient ion sampling for MS^2 fragmentation and peptide identification. The better performance of the 75 cm column results in 2 to 3 times higher peaks in comparison with the 25 cm long EASY-Spray column. In order to take into account the effect of gradient length on peak widths and peak height we proportionally increased MS^2 ion transfer time and the dynamic exclusion window with gradient length to obtain high-quality MS^2 spectra.

Figure 3. Average PWHM for 5 peptides as a function of gradient length for 25 cm and 75 cm x 75 µm EASY-Spray columns

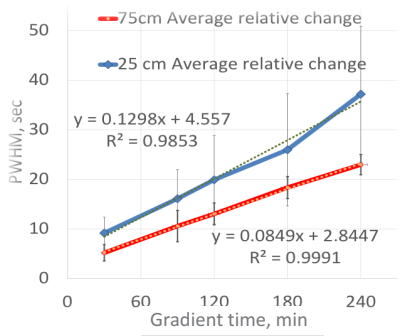


Figure 4. Estimated peak capacity as a function of gradient length for 25 cm and 75 cm x 75 µm EASY-Spray columns

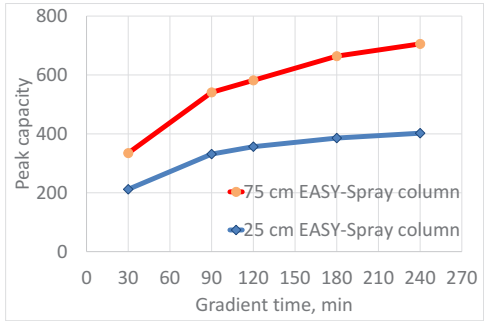
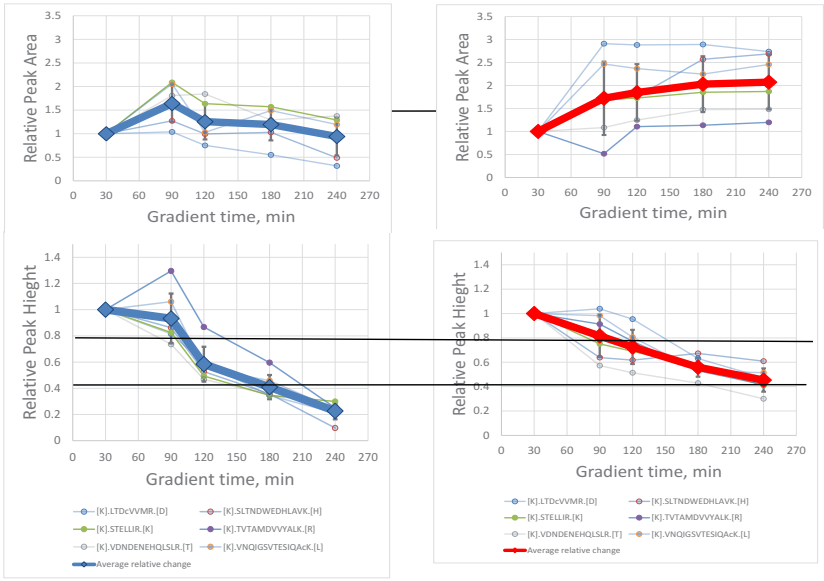


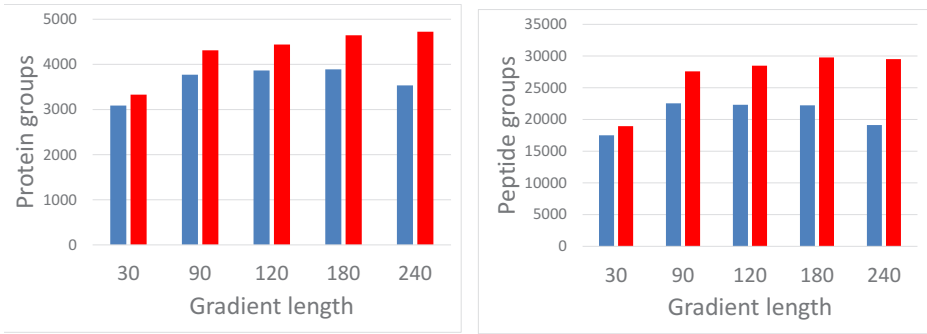
Figure 5. The relative change of peak area and peak height with gradient length for 75 µm x 25 cm (on the left) and 75 µm x 75 cm (on the right) EASY-Spray columns



Shotgun proteomics

The performance of 75 cm long EASY-Spray column is significantly better than its shorter counterpart. The number of protein and peptide groups that were identified with the 75 cm nano column increases with gradient time (Fig. 6). In contrast, the maximum number of peptide and protein groups using the 25 cm long column was achieved with the 90 min gradient. Further increasing the gradient length did not result in a higher number of identifications and showed decreased ratio of Peptide Spectrum Matches (PSMs) to MS^2 events. The presented data unambiguously shows that 75 cm columns allow deeper one dimensional proteome profiling especially with longer gradients (Table 1).

Figure 6. Protein and peptide groups identified at different gradient length using 75 µm x 25 cm (blue) and 75 µm x 75 cm (red) EASY-Spray columns and 1 µg of HeLa protein digest



thermoscientific

Boost protein identifications with higher sample loading

Higher sample loading amounts lead to additional protein and peptide identifications, especially with long gradients. We observed almost a linear relationship between the loading amount and the number of additionally identified proteins on the 75 cm column (Fig. 7). We were able to confidently identify >5300 protein groups when loading 3 µg of HeLa protein digest. At the same time the increased sample amount loaded onto the column did not negatively impact on peak properties (Fig. 8)

Figure 7. Number of protein and peptide groups identified at different loading amounts of HeLa protein digest on the 75 cm column (240 min gradient)

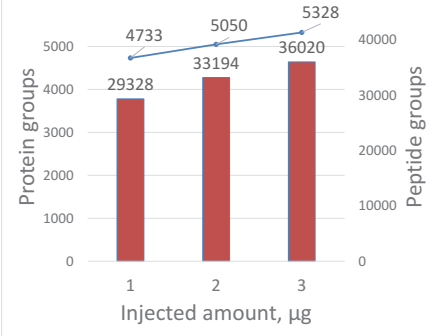
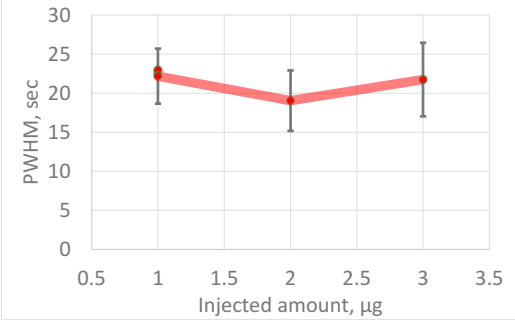


Figure 8. Dependence of the average PWHM on the loading amount of HeLa protein digest for the 75 cm column (240 min gradient)

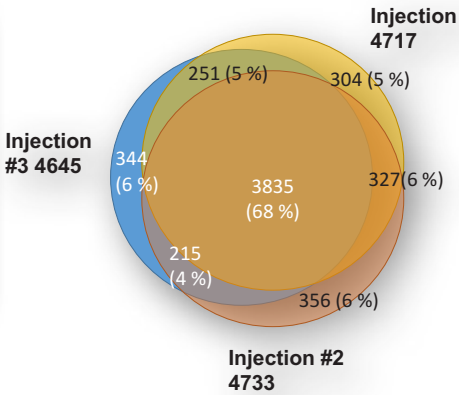


Proteome deepness and reproducibility of results

Combining the results for 3 consecutive replicates using 240 min gradients and 1 µg sample resulted in the identification of 5632 protein groups that covered 4 orders of magnitude in abundance levels based on their emPAI scores (Fig. 9). Among them 3835 proteins were identified in all 3 replicates (Fig. 10).

Figure 9. Sorted exponentially modified protein abundance index (emPAI) of proteins identified in 3 replicates of HeLa protein digest obtained with 240 min gradient.

Figure 10. Venn diagram showing the overlap among 3 technical replicates of HeLa protein digest (1 µg) obtained with 240 min gradient. In total 5632 protein groups were identified.



CONCLUSIONS

The UltiMate 3000 RSLCnano ProFlow nano LC system in combination with HRAM Q Exactive HF and 75 cm long EASY-Spray columns is a powerful platform for shotgun proteomics experiments.

- The use of 75 cm nano column results in deep and reproducible proteome coverage with more than 5000 proteins identified in a single LC-MS/MS experiment
- The high sample loading capacity of 75 cm EASY-Spray column boosts the number of protein identifications that cover more than 4 orders of magnitude in abundance levels.

Find out more at **thermofisher.com**

ThermoFisher SCIENTIFIC