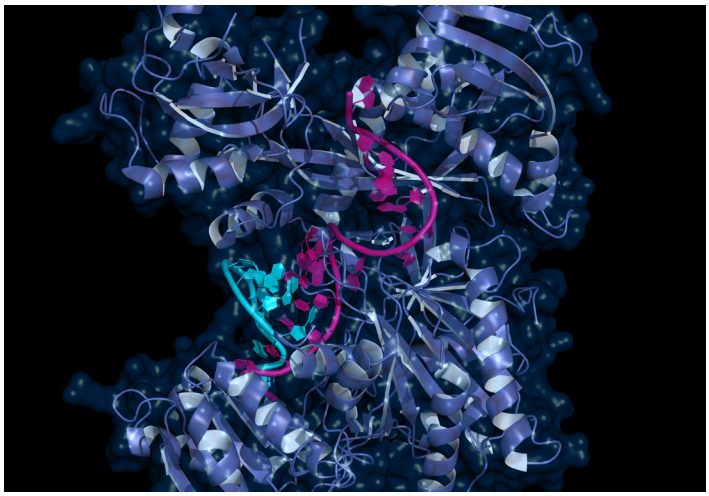
TECHNICAL NOTE 73291

Online 2D-nanoLC hyphenated with high-resolution accurate-mass Orbitrap mass spectrometry for comprehensive and robust proteome profiling

Authors: Runsheng Zheng[1], Alexander Boychenko[1], Tabiwang N. Arrey[2], Christopher Pynn[1], Wim Decrop[1], Peter Jehle[1]

[1]Thermo Fisher Scientific, Germering, Germany; [2]Thermo Fisher Scientific, Bremen, Germany

Keywords: Deep dive proteomics, UltiMate 3000 RSLCnano, Orbitrap Exploris 480, nanoLC-MS, online 2D-nanoLC, peptide fractionation



Goal

Develop and elucidate the benefit of an online 2D-nanoLC-MS platform for deep proteome profiling of HeLa and serum protein digests.

Introduction

In the field of bottom-up proteomics, the profiling of complex samples such as cell lines, tissues, or body fluid proteomes is usually accompanied by the question: how deep is “deep enough”? This inevitably arises due to the necessary trade-off between sample throughput and depth of proteomic coverage, forcing scientists to choose between “long, deep, and slow” or “short, shallow, and fast” analytical methods.

One approach addressing this challenge is the employment of offline fractionation techniques that split samples into several independent nanoLC-MS measurements. Offline high-pH reversed-phase LC (high-pH RPLC) followed by low-pH reversed-phase LC (low-pH RPLC) has gradually replaced the strong cation exchange (SCX) coupled with low-pH reversed-phase LC-MS/MS. This is due to better efficiency of peptide separation under RP conditions, better nanoLC-MS compatibility, and lack of need for an extra desalting step after the first dimension sample fractionation. However, an offline high-pH RP fractionation method still has potential for improvement. For example, it is time-consuming, affording low throughput, and requires large amounts of sample. It is also prone to sample loss during fraction collection and transfer between fractionation devices and analytical instruments.

Here we present a simple-to-use online low-flow high-pH RP x low-pH RP separation platform (termed online 2D-nanoLC) for deep proteome profiling using a Thermo Scientific™ UltiMate™ 3000 RSLCnano system coupled to a Thermo Scientific™ Orbitrap Exploris™ 480 mass spectrometer. This automated approach not only reduces sample complexity but also boosts peptide/protein identifications and quantification without any requirement for extra sample handling procedures between dimensions. The proof-of-principle study presented here permitted the identification of ~7,000 protein and >70,000 peptide groups in HeLa digest within 7 hours, and ~329 protein with ~3,700 peptide groups in serum digest within 4 hours. Even deeper proteome profiling was achieved by applying a spectral library database search to acquired data. Furthermore, the developed methods can easily be adapted to accommodate more fractions for even deeper proteome profiling as well as to provide comprehensive data for enhanced data processing that can boost the number of identifications and improve proteome coverage.

Experimental materials and methods

Sample preparation

Thermo Scientific™ Pierce™ HeLa protein digest (P/N 88328, 20 µg/vial) was reconstituted to a final concentration of 800 ng/µL by consecutively dissolving the pellet with 100 µL 0.1% formic acid (FA) in water.

Serum samples from SeraLab (UK) collected from healthy donors following informed consent were trypsin digested using a simple procedure that does not include reduction nor alkylation of cysteine residues (Figure 1).

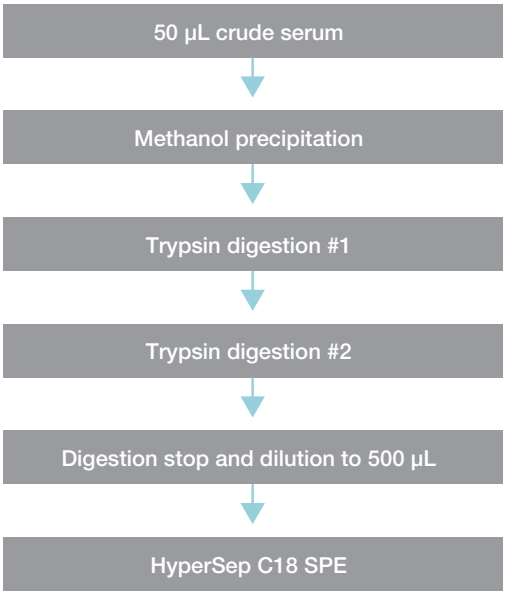


Figure 1. Simple sample preparation procedure for crude serum samples

Briefly, 200 µL methanol was added to 50 µL serum aliquots in an Eppendorf™ LoBind™ 96-well plate for protein precipitation. The plate was vortexed for 5 minutes and subsequently centrifuged (1500 g, room temperature). After discarding the supernatant, the plate was inverted and allowed to dry for 15 minutes. Subsequently, 200 µL digestion mix (pH 7.0) was added to each well and the plate vortexed until the pellet was re-dissolved. Another 250 µL digestion mix was added and digestion allowed to proceed for 3 hours at 37 °C and 1250 rpm. The digestion was quenched using 50 µL of a 10% FA solution.

Finally, 100 µL of the digest was loaded onto a Thermo Scientific™ HyperSep™ C18 cartridge (100 mg bed weight, 1 mL capacity, P/N 60108-302), washed with 400 µL 0.1% FA, eluted with 400 µL 0.1% FA in 50% ACN, dried and resuspended in 200 µL of 0.1% FA. Subsequently, 160 µL was transferred to a polypropylene vial with a glass insert prior to measurement.

Consumables

- Fisher Scientific™ LC-MS grade water (P/N W6-212)
- Fisher Scientific™ LC-MS grade acetonitrile (P/N 10616653)
- Fisher Scientific™ LC-MS grade water with 0.1% formic acid (P/N 85171)
- Fisher Scientific™ LC-MS grade 80% acetonitrile with 0.1% formic acid (P/N 15431423)
- Thermo Scientific™ Pierce™ trifluoroacetic acid (TFA), LC-MS grade, (P/N 13464279)
- Fisher Scientific™ LC-MS grade formic acid (FA) (P/N 10596814)
- Fisher Scientific™ ammonium bicarbonate (P/N 10532775)
- Fisher Scientific™ LC-MS grade isopropanol (P/N 10684355)
- Fluidics and columns used to setup online 2D-nanoLC system (see Table 1 and details in Figure 2b)

Table 1. Fluidics, columns, and consumable accessories for online 2D-nanoLC-MS. The letter designations refer to Figure 2B.

	Description	Part number*
A	Thermo Scientific™ nanoViper™ capillary, 20 µm x 350 mm	6041.5240
B	Thermo Scientific™ Acclaim™ PepMap™ 100 C18 LC columns, 75 µm i.d. x 2 cm (3 µm, 100 Å), nanoViper	164535
C	Thermo Scientific™ nanoViper™ capillary, 20 µm x 550 mm	6041.5260
D	Thermo Scientific™ Viper™ union	6040.2304
E	Thermo Scientific™ EASY-Spray™ column (75 µm x 150 mm, 3 µm)	ES800A
F	Thermo Scientific™ nanoViper™ capillary, 50 µm x 350 mm	6041.5540
G	Tee-piece union	ZT1XCS6-M (VICI)
H	PTFE tubing, 500 µm i.d., 100 cm, used as waste tubing	6720.0077
	1/16" Universal fingertight fitting, one-piece design, extra long thread (4 pieces)	6720.0015
I	Thermo Scientific™ nanoViper™ sample loop 20 µL	6826.2420
J	Thermo Scientific™ nanoViper™ capillary, 20 µm x 750 mm	6041.5260
K	Thermo Scientific™ PepSwift™ monolithic column, 100 µm i.d. x 250 mm, with nanoViper fittings	164543
L	UV-Monitor (45 nL, VF-D4, VWD), with nanoViper fittings	6074.0285
M	Thermo Scientific™ nanoViper™ capillary, 20 µm x 150 mm	6041.5121
Other components	1x Low-Dispersion valve 2pos-10port	6041.0001A
	Solvent lines (NCP-3200RS)	6041.2530
	Rear seal wash tubing (cut to 1.3 m for NCP-3200RS)	6007.9100
	Polypropylene vials 250 µL with glass insert	6820.0027
	Polypropylene caps for WPS vials	6820.0028
	10 mL vials (headspace) with crimp caps and septum (for transport liquid)	6820.0023

* Consumables are all from Thermo Fisher Scientific except for item "ZT1XCS6M" from VICI™-Valco Instruments.

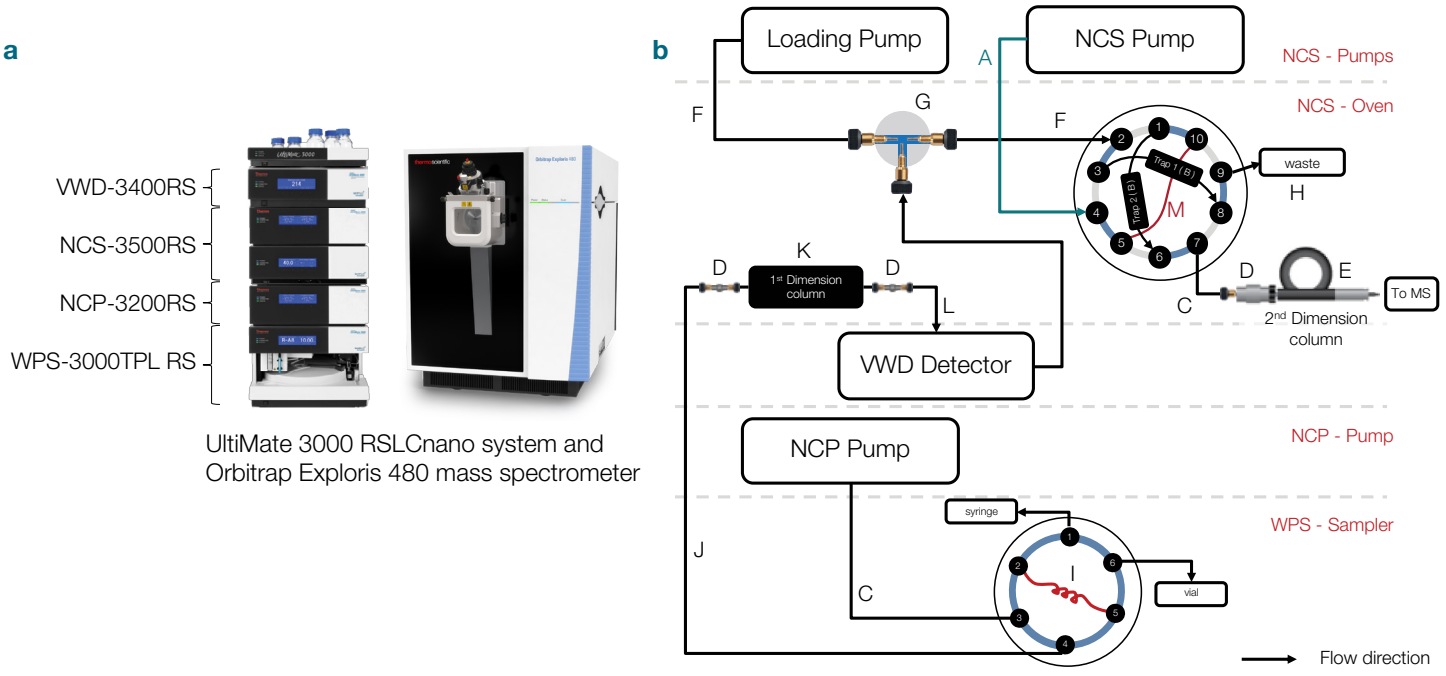


Figure 2. Modules for assembling the online 2D-nanoLC-MS platform (a) and its fluidic connection scheme (b) (denoted letters are described in Table 1)

Online 2D-nanoLC-MS system

The UltiMate 3000 RSLCnano system for online 2D-nanoLC-MS comprises NCS-3500RS, NCP-3200RS, WPS-3000TPL RS, and VWD-3400RS modules, as well as an Orbitrap Exploris 480 mass spectrometer (Figure 2a). Only one 10-port, 2-position switching valve installed in the column compartment is required for this application.

The fluidic configuration is shown in Figure 2b. A PepSwift monolithic capillary column was employed to separate peptides in the first dimension using mobile phases buffered to pH 8.0 (using 50 mM NH_4HCO_3). The loading pump integrated into the NCS-3500RS module was used to deliver 0.2% TFA (pH 1.0) to dilute and acidify the eluate from the first dimension. The eluate was subsequently loaded (concentrated) onto one of two nano trap columns that alternate every 45 min. Peptides from the first dimension were eluted from the trap column onto an EASY-Spray column that was used to separate peptides at pH 3.0 in 45 min cycles.[1] The column was connected to an Orbitrap Exploris 480 mass spectrometer operated in Data Dependent Acquisition (DDA) mode using an EASY-Spray source.

Table 2. Solvents and conditions for online 2D-nanoLC modules

Module	Property	Setting
NCS-3500RS pump	Mobile phase A	H_2O with 0.1% FA (pH ~3.0)
	Mobile phase B	80/20 (v/v) ACN / H_2O with 0.1% FA (pH ~3.0)
Loading pump	Loading buffer (channel A)	H_2O with 0.2% TFA (pH ~1.0)
NCP-3200RS pump	Mobile phase A	H_2O with 50 mM NH_4HCO_3 (pH ~8.0)
	Mobile phase B	80/20 (v/v) ACN / H_2O with 50 mM NH_4HCO_3 (pH ~8.0)
NCS-3500RS; NCP-3200RS; Loading pump	Rear washing buffer	90/10 (v/v) H_2O / Isopropanol with 0.1% FA
WPS-3000TPL RS	Sampler washing buffer	100% ACN with 0.1% FA
WPS-3000TPL RS	Temperature control	5 °C
Column oven		40 °C
EASY-Spray column		50 °C

FA = Formic acid, TFA = Trifluoroacetic acid, ACN = Acetonitrile, NH_4HCO_3 = Ammonium bicarbonate

Software configuration

The modules were configured in the "Chromeleon Instrument Configuration Manager" panel of SII/ Chromeleon as shown in Figure 3a. Standard settings were used for configuring the NCS-3500RS, VWD-3400RS, and WPS-3000TPL RS modules. Please note that the 10-port, 2-position switching valve must be installed in the right valve position (Figure 3b) in order to execute the template methods available for download in Thermo Scientific™ AppsLab. The NCP-3200RS pump was configured by typing the digit "2" after the module name and after the word "pump" for each of the activated channels (Figures 3c and 3d).

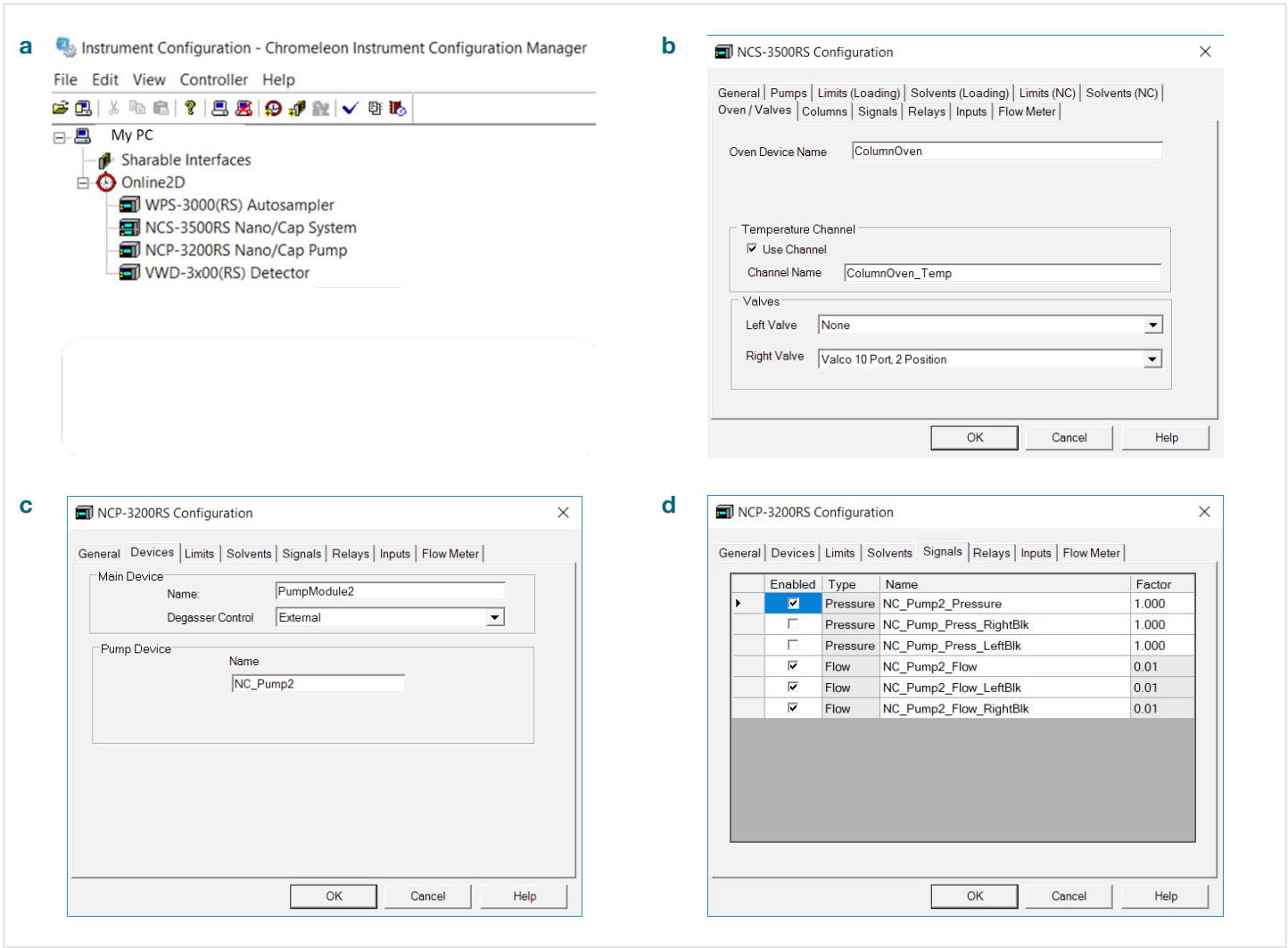


Figure 3. (a) Instrument configuration panel showing the modules required for the online 2D-nanoLC application; (b) configuration of the 10-port switching valve in the column oven; (c) and (d) configuration of the NCP-3200RS pump realized by adding the digit "2" to the name of "Main Device" and "Pump Device" in the "Device" (c) and "Signals" (d) tabs

MS acquisition parameters

MS data were recorded on an Orbitrap Exploris 480 mass spectrometer. The MS tune and data acquisition parameters are shown in Figure 4 and are available for download on AppsLab.

Data acquisition and processing

Data were acquired using Thermo Scientific™ Xcalibur™ software. The Online 2D UltiMate 3000 RSLCnano system was controlled using Standard Instrument Integration (SII). DDA data for HeLa and serum digest were processed with Thermo Scientific™ Proteome Discoverer™ 2.2 software using the SEQUEST™ HT search algorithm. The false discovery rate (FDR) was set below 1% at the peptide and the protein level.

Figure 4. Summary of the MS parameters used for the online 2D-nanoLC-MS platform

Results and discussion

Online 2D-nanoLC-MS platform enables deep proteome profiling

Offline high-pH RP fractionation, although adopted for deep proteome profiling for many years, requires high sample amounts, extra steps for sample reconstitution, and can result in sample loss.

We developed a low-flow online high-pH RP × low-pH RP method where fractions from the first dimension are captured alternately on one of two trapping columns and sequentially analyzed using a high-throughput low-pH RP gradient with an Orbitrap Exploris 480 mass spectrometer. As both the first and second dimension separations are driven by peptide hydrophobicity, it follows that the fractions collected at later time points from dimension one would also elute at higher %B values in dimension two. Thus, to enable a better peptide distribution for each fraction and improve MS utilization, we programmed the 2nd dimension gradients with a staggered increase of mobile phase B portion at the start and end of the gradient for each fraction. Accordingly, 2-, 4-, and 8-fraction methods were developed (Figure 5a) and then successfully applied to the measurement of cell digest (HeLa) (Figure 5b) and body fluid (crude serum prepared without depletion, reduction, and alkylation steps) samples (Figure 5c).

Increased number of fractions as well as higher sample loading (up to 4 µg) in the 1st dimension separation lead to higher proteome depth (up to ~7,000 protein with >70,000 peptide groups) in HeLa digest with the 8-fraction method (6.75 hours) (Figure 6a). The identified peptides are well distributed across eight fractions, with an average of 91% of unique peptides identified in each (Figure 6b). The orthogonality plot of the 8-fraction data (based on peptide hydrophobicity calculation using the R package[2-4]) demonstrates good 1st dimension and 2nd dimension orthogonality (Figure 6c). Furthermore, the correlation (r=0.9985) in the accumulative curve (Figure 6b) reveals a linear gain in peptide identification with an increasing number of fractions.

The positive relationship between increased peptide and protein group identifications with an increased number of fractions was also observed for a serum sample analysis (Figure 7a). Increasing the number of fractions from two to four resulted in 329 protein group identifications, a 30% increase compared to the 2-fraction method. The capacity to reach deeper into the proteome was realized with the 8-fraction method when double the amount of sample was loaded, leading to an increase of 28% and 14% protein and peptide identifications, respectively.

Good orthogonality of the two dimensions was again evidenced by the broad peptide hydrophobicity distribution, which displayed similar features to the HeLa digest (Figures 7b and 7c).

Proteome profiling of serum samples

The requirement for high-throughput multi-analyte assays is of primary importance in clinical research, particularly in the rapidly emerging field of personalized medicine. Although immunoaffinity-based assays yield unsurpassed levels of sensitivity and specificity, shortage in multiplexing capabilities, high development costs, and long development periods of novel antibody invention still pose a significant challenge.

Technical advances in nanoLC-MS technology, enable scientists to perform high-throughput screening to facilitate rapid identification and quantification of potential biomarkers from body fluid samples. Nevertheless, nano-LC-MS analysis still has not reached its full potential in this field of research. Challenges posed by blood proteomics, in particular, include the high dynamic range of proteins in the sample, which severely limits the identifications and analytical depth achievable in a single run. Strategies to overcome these hurdles usually focus on the preparation step (for example depletion of high abundant proteins) or involve offline 2-dimensional fractionation. Both regimes require extensive sample manipulation, which is time-intensive and often results in added sample/protein loss.

We used the online 2D-nanoLC-MS approach to profile the proteome of 15 serum samples with the 4-fraction method in around 57 hours (Figure 8a). The MS1 intensity distribution in each fraction across 15 samples is highly similar, facilitating a reach to a proteome depth of 321 protein and 3658 peptide groups on average (Figure 8b).

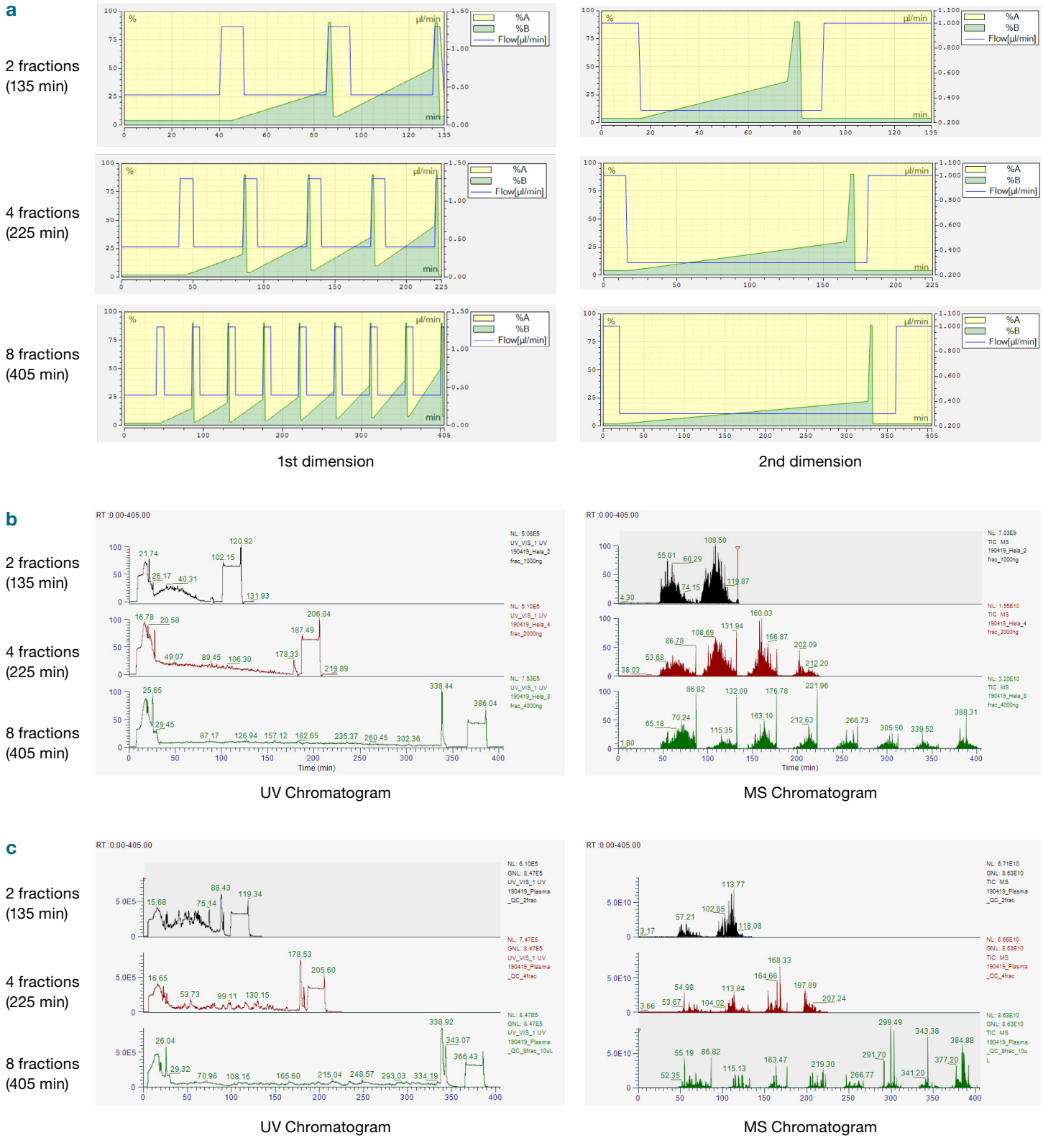


Figure 5. Gradients (a) for high-pH RP (1st dimension) and low-pH RP (2nd dimension) of 2-, 4-, 8-fraction methods, and the corresponding UV and MS chromatogram in HeLa (b) and serum (c) digest

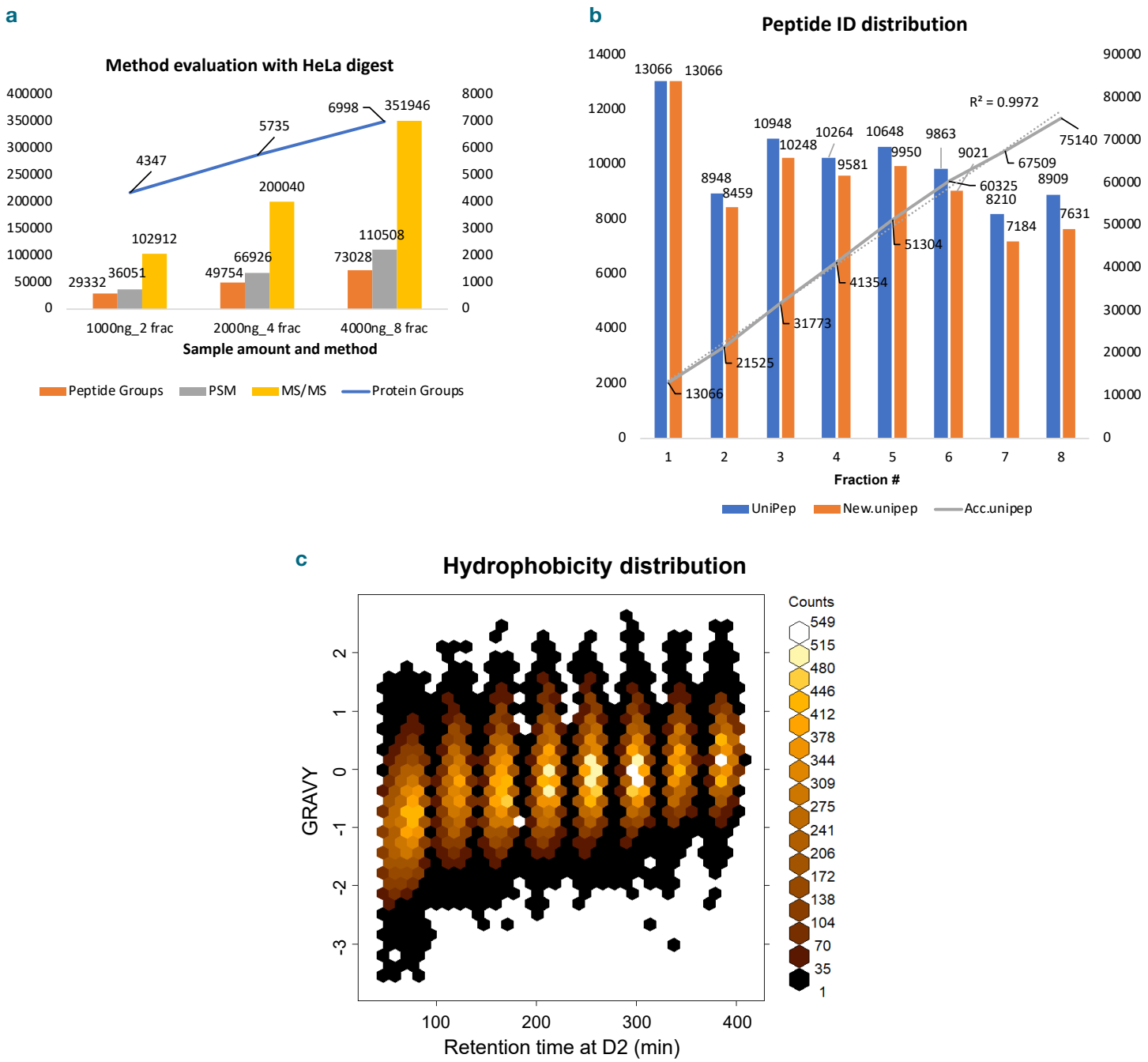


Figure 6. Online 2D-nanoLC-MS enables fast, automated proteome profiling in HeLa digest with 2-,4-, 8-fraction methods. (a) Data quality comparison indicates the benefit of getting more peptide and protein identifications with higher sample loading and an increased number of fractions; (b) peptide identification with the 8-fraction method shows even peptide distribution among fractions and feasibility of method extension for deeper proteome coverage; (c) orthogonality presented as peptide hydrophobicity distribution through eight fractions.

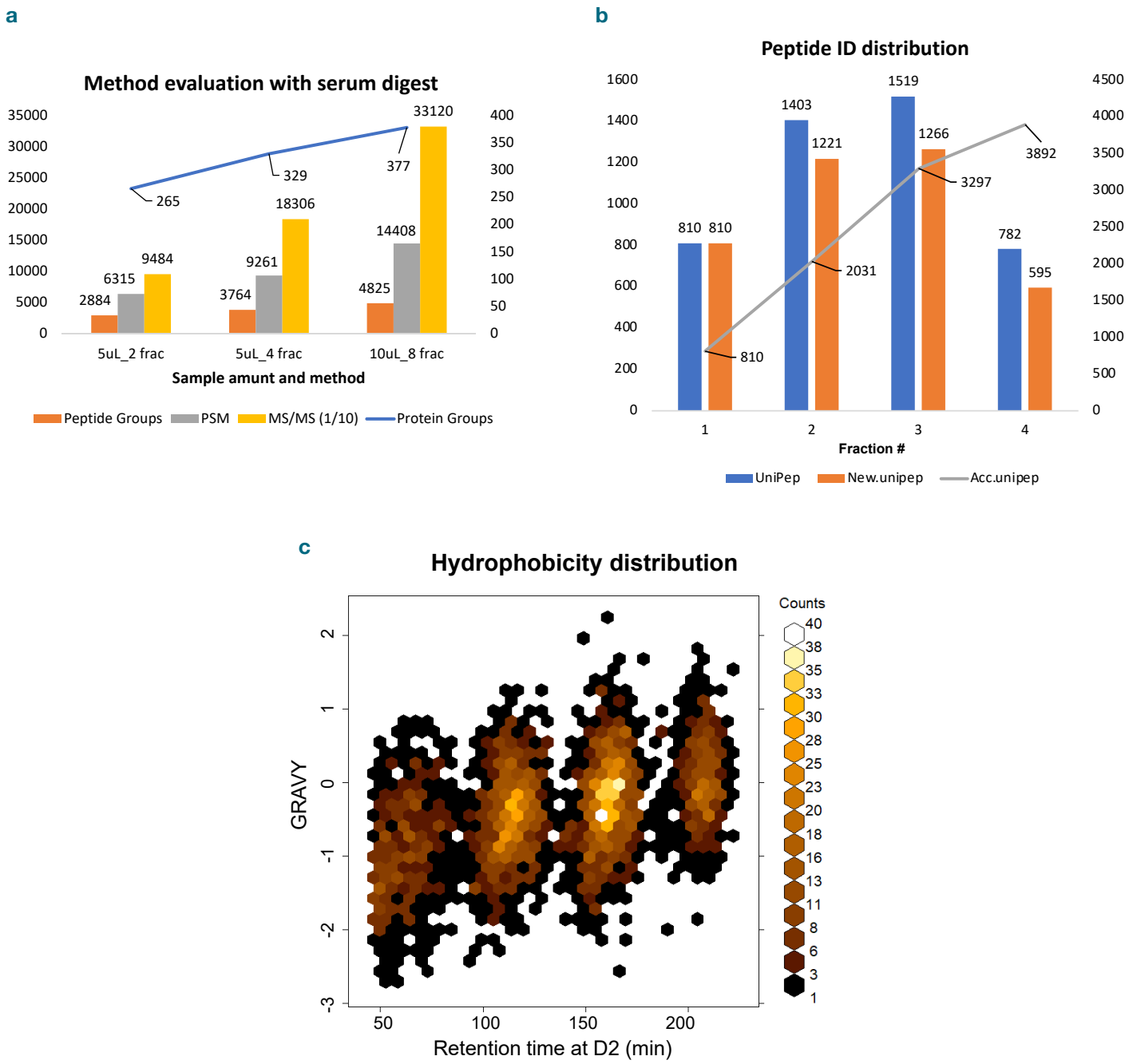


Figure 7. The online 2D-nanoLC-MS system permits fast, automated full proteome profiling in serum digest with 2-,4-, 8-fraction methods. (a) Peptide, PSM, MS/MS and protein group numbers for 2, 4, and 8 fraction methods (equivalent to 0.25 µL crude serum and 0.5 µL serum, respectively) (MS/MS count is shown as 1/10 scale); (b) peptide identification in each fraction; (c) orthogonality presented as peptide hydrophobicity distribution through four fractions.

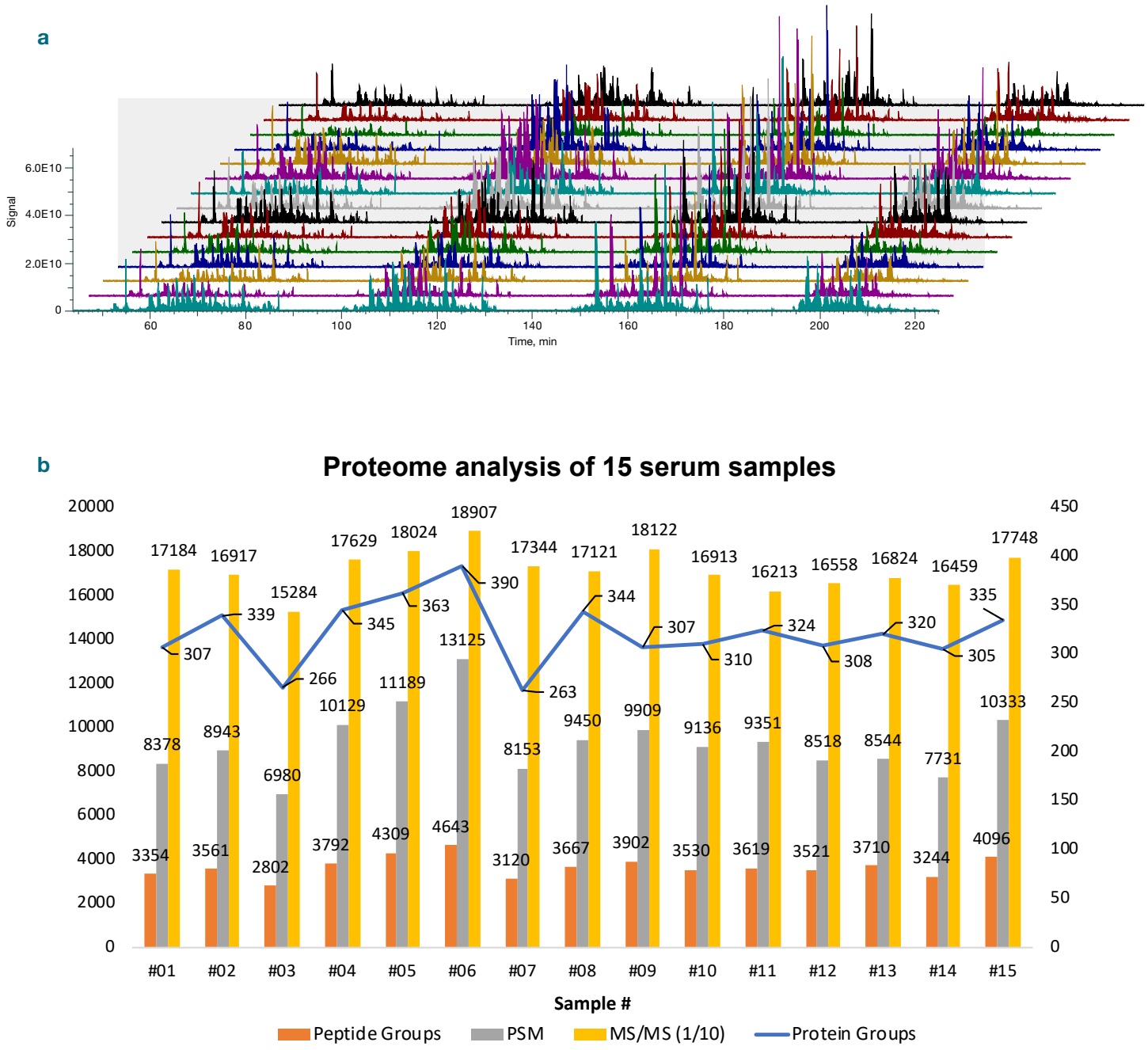


Figure 8. Deep proteome profiling of 15 serum samples with the online 2D-nanoLC-MS platform, showing TIC intensity (a) and peptides, PSM, MS/MS, and protein identifications for further bioinformatic analysis (b) (MS/MS count is shown as 1/10 scale).

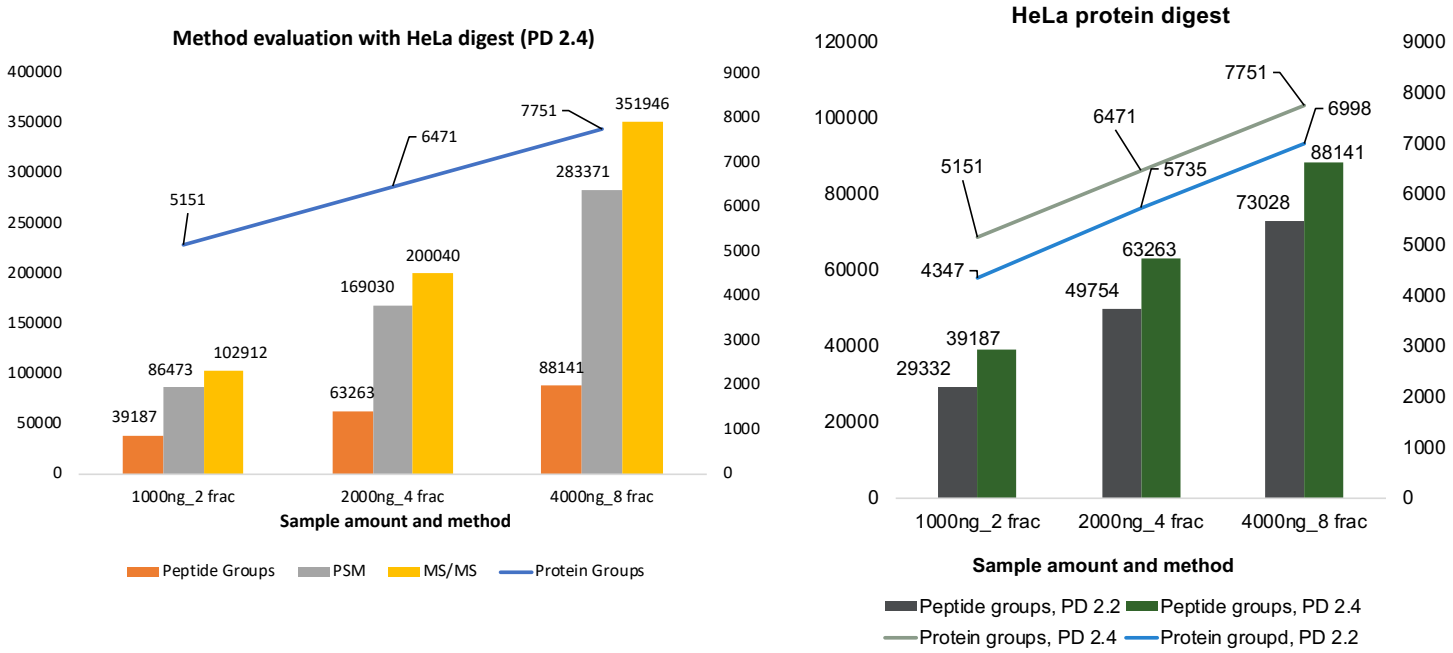


Figure 9. Peptide-spectrum library assists better identification of peptides and proteins in the HeLa digest. (a) Reprocessing of the HeLa dataset with Proteome Discoverer 2.4 software; (b) Proteome Discoverer 2.4 software boosts both the peptide and protein identifications.

Advanced algorithm improves peptide and protein identification

As shown in Figure 6a, the average peptide ID rate (ratio of PSM/MS^2) in the HeLa sample is around 33% among 2-, 4- and 8-fraction methods. This relatively low value typically arises due to the so-called precursor co-isolation phenomenon during peptide fragmentation that leads to the formation of "chimeric" MS/MS spectra, which increase the difficulty for the database search algorithm to correctly identify the peptide. Integration of the "Precursor Detector node" and spectrum library search capabilities into Proteome Discoverer 2.4 (PD 2.4) software significantly improves the peptide and protein identifications while maintaining 1% FDR. Accordingly, the PSMs/MS^2 rate increased to 83% on average (Figure 9a) yielding 27% more peptides and 14% more protein identifications (Figure 9b). As a result, we could identify ~88,000 peptide and ~7,700 protein groups with the 8-fraction method.

Conclusions

We developed a novel and simple-to-use online 2D-nanoLC-MS approach for automated proteome profiling by coupling the versatile UltiMate 3000 RSLCnano system to the state-of-art Orbitrap Exploris 480 mass spectrometer. This approach yields the following attributes:

- No requirement for manual sample manipulation after enzymatic digest
- High orthogonality between 1st and 2nd dimension separations
- High MS utilization (up to 88% in the 8-fraction method)
- Compatibility with challenging sample matrices, e.g. human serum
- A powerful alternative to long one-dimensional separations in shotgun proteomics

This platform permits deep proteome profiling both in cell lysates (e.g., HeLa digest, 6,997 protein groups and 73,028 peptide groups when adopting the classical database search approach, and 7,751 protein groups and 88,141 peptide groups using the peptide library search in only 6.75 h) and body fluid (e.g., human serum, ~320 protein groups in only 3.75 h using the classical database search approach).

Furthermore, the methods can easily be extended to achieve even deeper proteome profiling from more fractions. The online 2D UltiMate 3000 RSLCnano system is fully compatible with all the latest Orbitrap mass spectrometers including the Thermo Scientific™ Orbitrap Eclipse™ Tribrid™ mass spectrometer.

thermo scientific

References

1. Boychenko, A.; Pynn, C.; Arrey, T.; Zheng, R.; Decrop, W.; Jehle, P. Tailored high-throughput low-flow LC-MS methods for large scale sample cohort analysis. Thermo Scientific Technical Note 73208.
2. Osorio, D.; Rondón-Villarreal P.; Torres. R. Peptides: Calculate indices and theoretical physicochemical properties of peptides and protein sequences., 2014. URL: http://CRAN.R-project.org/package=Peptides. R Package Version 2.2.
3. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
4. Carr, D.; Lewin-Koh, N.; Maechler, M. Hexagonal Binning Routines., 2019. URL: http://github.com/edzer/hexbin

Find out more at **thermofisher.com/nanoLCMS**

ThermoFisher SCIENTIFIC