analytical chemistry

pubs.acs.org/ac Article

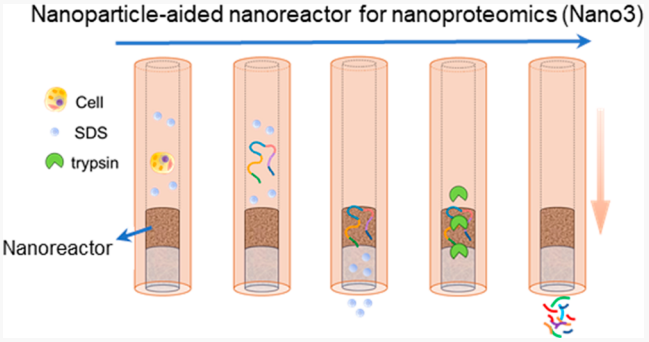
Nanoparticle-Aided Nanoreactor for Nanoproteomics

Zhichang Yang, Zhaoran Zhang, Daoyang Chen, Tian Xu, Yuan Wang, and Liangliang Sun*

Cite This: *Anal. Chem.* 2021, 93, 10568−10576 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Large-scale bottom-up proteomics of few even single cells is crucial for a better understanding of the roles played by cell-to-cell heterogeneity in disease and development. Novel proteomic methodologies with extremely high sensitivity are required for few even single-cell proteomics. Sample processing with high recovery and no contaminants is one key step. Here we developed a nanoparticle-aided nanoreactor for nanoproteomics (Nano3) technique for processing low-nanograms of mammalian cell proteins for proteome profiling. The Nano3 technique employed nanoparticles packed in a capillary channel to form a nanoreactor (≤30 nL) for concentrating, cleaning, and digesting proteins originally in a lysis buffer containing sodium dodecyl sulfate (SDS), followed by nanoRPLC-MS/MS analysis. The Nano3 method identified a 40-times higher number of proteins based on MS/MS from 2-ng mouse brain protein samples compared to the SP3 (single-pot solid-phase-enhanced sample preparation) method, which performed the sample processing using the nanoparticles in a 10 μL solution in an Eppendorf tube. The data indicates a drastically higher sample recovery of the Nano3 compared to the SP3 method for processing mass-limited proteome samples. In this pilot study, the Nano3 method was further applied in processing 10−1000 HeLa cells for bottom-up proteomics, producing 441 ± 263 (n = 4) (MS/MS) and 983 ± 292 (n = 4) [match between runs (MBR)+MS/MS] protein identifications from only 10 HeLa cells using a Q-Exactive HF mass spectrometer. The preliminary results render the Nano3 method a useful approach for processing few mammalian cells for proteome profiling.



The advances in separation techniques and mass spectrometers have enabled comprehensive proteome profiling using bottom-up proteomics, and the identification of over 10 000 gene products from human cells was reported by multiple laboratories.[1−5] To reach the "dark corner" of the human proteome, that is, proteins with low copy numbers per cell, it usually requires hundreds of micrograms to milligrams of material and extensive fractionation.[6] It was reported that, with a charge state of 5−10, single ion detection was possible using orbitrap instrumentation.[7] The state-of-the-art MS platforms have achieved zmol level limit of detections for peptides in complex samples.[8−14] If a single mammalian cell contains 12 000−15 000 gene products with a dynamic range spanning from one copy per cell to millions of copies per cell, more than 4000 gene products could be identified from a single mammalian cell if protein recoveries from sample preparation and separation are 100%.[9,15] However, only few laboratories reported the identification of hundreds to more than 1000 proteins from a single mammalian cell.[14,16,17] The challenge of proteomic analysis of single cells or few cells lies significantly in the sample preparation step. A mammalian somatic cell is only 10−20 μm in size and contains only hundreds of picogram of proteins in mass.[9,18] A full recovery of the protein material at such trace amounts is extremely challenging with traditional sample processing techniques due to significant sample loss caused by adsorption of proteins/peptides on surfaces, such as processing containers and pipet tips. Researchers have been making great efforts in exploring novel sample processing techniques that are suitable for mass-limited proteome samples (e.g., single or small numbers of cells).[14,16,19−25]

The basic idea of the new sample preparation methods for mass-limited samples is to decrease sample processing volume and eliminate sample transfer. The NanoPOTS (nanodroplet processing in one pot for trace samples) method developed by Zhu et al. is a nice example.[14] The method performed all the bottom-up sample preparation steps in a nanowell with a total volume of only 200 nL. It has achieved nearly 400 to over 1000 protein identifications (IDs) from a single HeLa cell, corresponding to 1000−4000 peptide IDs based on MS/MS.[16,26] Due to an extremely small volume of sample handling,

Received: April 21, 2021
Accepted: July 13, 2021
Published: July 23, 2021

https://doi.org/10.1021/acs.analchem.1c01704
Anal. Chem. 2021, 93, 10568−10576

NanoPOTs required careful operations in a humid chamber. MicroFASP is another example developed by Zhang et al.,[25] and it is a modified version of the FASP (filter-aided sample preparation) method.[27] For the microFASP, a miniature filter membrane of 0.1 mm^2 was installed into a 20 μL pipet tip for sample processing, drastically reducing sample loss during sample preparation. The sample processing volume was maintained at a low microliter level. Over 1800 proteins and 13 000 peptides were identified based on MS/MS when a sample containing 100 MCF-7 cells was processed by the microFASP. More importantly, microFASP does not require special instrumentation.

The SP3 (single-pot solid-phase-enhanced sample preparation) is also one of sample processing approaches suitable for mass-limited samples.[22,23] All sample preparation steps were performed in a single Eppendorf tube. Under high organic content (>70% acetonitrile (ACN)), proteins were effectively captured on paramagnetic beads through hydrophilic interaction while salts and detergents were effectively removed. After that, proteins on beads were digested by trypsin, followed by MS analysis. The SP3 outperformed FASP in terms of proteome coverage when low μg of protein materials were processed, and high quantitative reproducibility was also documented for the SP3 method.[28] Over 15 000 unique peptides and nearly 3000 proteins were identified based on MS/MS when a sample containing only 1000 HeLa cells was processed by SP3.[22] Recently, an automated SP3 (auto-SP3) method was developed for high-throughput processing of mass-limited samples for bottom-up proteomics.[29] However, SP3 is mainly operated in Eppendorf tubes and requires a microliter-level solution for sample processing, limiting its performance for processing low nanograms of complex proteome samples.

Unlike NanoPOTs, microFASP, and SP3, where samples were processed in an open environment, some microreactors with small volumes in fused silica capillaries have been developed to process proteins in a closed environment and are potential alternatives for trace material processing.[30−32] The microreactors in fused silica capillaries can be easily sealed and provide a closed environment for all sample preparation steps.

Inspired by the microreactors and the SP3 method, in this work, we present a new sample processing technique, nanoparticle-aided nanoreactor for nanoproteomics (Nano3), for bottom-up proteomics of mass-limited samples. The Nano3 method employs the same nanoparticles and principle as the SP3 method but carries out the sample processing in a nanoreactor with a total volume of 30 nL or smaller. Paramagnetic beads were packed into a fused silica capillary to form a nanoreactor for capturing proteins from cells lysed by a lysis buffer containing sodium dodecyl sulfate (SDS) and ACN. After flushing the nanoreactor with ACN to remove SDS, proteins captured on nanoparticles were digested into peptides by a plug of trypsin solution, followed by peptide collection from the nanoreactor via flushing the reactor with a buffer containing 2% (v/v) ACN and 0.1% (v/v) formic acid and nanoRPLC-MS/MS analysis. We compared the performance of Nano3 and SP3 for processing 50, 10, and 2 ng of mouse brain proteome samples. Nano3 outperformed SP3 regarding the number of protein IDs and intensity, indicating better overall sample recovery. We further validated the Nano3 method for processing 1000, 100, and 10 HeLa cells, corresponding to 100, 10, and 1 ng proteins in mass with the assumption of 100 pg proteins per HeLa cell.

EXPERIMENTAL SECTION

Details of materials and reagents are listed in Supporting Information (SI) I. All the capillaries involved in Nano3 and protein transfer were pretreated with BSA (2 mg/mL) solution to reduce sample loss due to adsorption based on our recent study.[33] Briefly, 2 mg/mL BSA solution was injected into the capillary and stored in the capillary for 10 min at room temperature. The BSA solution was then flushed out with water. Flushing continued for 30 min. The capillary was then flushed with methanol and air-dried before use.

Fabrication of the Nanoreactor for the Nano3. A capillary (200 μm i.d., 360 μm o.d., 10 cm long) was installed with a polymer frit first for packing hydrophilic paramagnetic beads. The polymer solution for frit was made according to the manufactural protocol. Briefly, 15 μL of Kasil-1624 and 5 μL Kasil-1 were mixed. After that, 5 μL of formamide was added into the mixture and vortexed for a few seconds. Frit material was then introduced into the capillary through capillary action. Both ends of the capillary were sealed with rubber. The capillary was then incubated in an 80 °C water bath overnight. Prior to paramagnetic beads packing, the capillary was rinsed with methanol. Both kinds of paramagnetic beads (hydrophilic and hydrophobic) with about 60 nm in size (SI Figure S1) were mixed with a 1:1 ratio, rinsed with water, and were resuspended in 70% (v/v) ACN. The beads solution was introduced into the capillary using a syringe and manual pump. The length of the beads in the capillary was controlled to be about 1 mm, corresponding to about 30 nL total volume. The whole capillary was filled with 70% (v/v) ACN before use. The total volume of the capillary containing the nanoreactor was 3−4 μL.

Mouse Brain Protein Preparation. Mouse brain proteins were extracted in a lysis buffer containing 2% SDS, 100 mM NH_4HCO_3 (pH 8.0), complete protease inhibitors, and phosphatase inhibitors. After reduction and alkylation, 200-μg proteins were used for validating the sensitivity of our LC-MS platform and comparing the SP3 and Nano3 methods. The detailed sample preparation procedure is described in Supporting Information I.

HeLa Cell Preparation. The cells (originally from ATCC) were cultured at 37 °C under 5% CO_2 in Dulbecco's Modified Eagle's Medium, supplemented with 10% fetal bovine serum, 1× glutamine, and 1000 units/mL penicillin/streptomycin (all of reagents above from Thermo Fisher Scientific). After cell culture, cells were harvested, and cell concentration was measured using a hemocytometer. After careful PBS rinsing, cells suspended in PBS were diluted into concentrations of 1000 cells/μL, 100 cells/μL, and 10 cells/μL using the PBS buffer based on the hemocytometer result. To minimize cell lysis in PBS, the entire process was performed within 20 min, and diluted cell samples were immediately processed by the Nano3 method.

Another batch of cultured HeLa cells (2×10^6 cells) was lysed in 200-μL lysis buffer (2% SDS, 100 mM NH_4HCO_3, pH 8.0, complete protease inhibitor, and phosphatase inhibitor) with ultrasonication for 10 min on ice and 95 °C for 5 min. The lysate was centrifuged at 14 000g for 5 min. The supernatant was collected and subjected to a BCA assay for protein concentration measurement. The sample was also used to validate the Nano3 method.

Mouse Brain Sample Processing Using the SP3 and Nano3 Methods. For the SP3 method, 20 μL of each of the

https://doi.org/10.1021/acs.analchem.1c01704
Anal. Chem. 2021, 93, 10568−10576

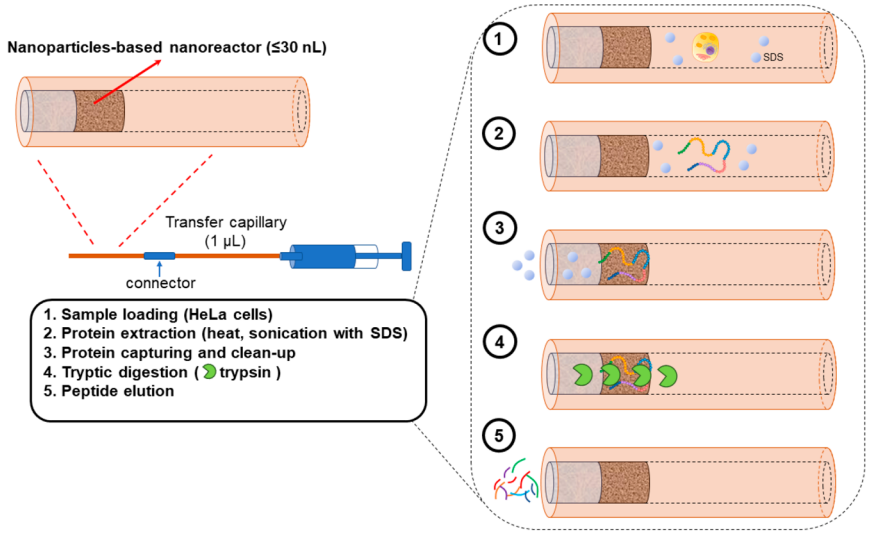


Figure 1. Schematic of the general workflow of sample processing with the Nano3 method.

paramagnetic beads stock solution (50 μg/μL) was combined and was rinsed with water for a few times. Twenty μL water was then used to resuspend the paramagnetic beads. Reduced and alkylated mouse brain proteins in 2% SDS lysis buffer was diluted with 70% (v/v) ACN into 50 ng/μL, 10 ng/μL, and 2 ng/μL. One μL of protein solution from each dilution was combined with 1 μL of beads solution. Then 5 μL of ACN was added into the tube, so ACN concentration was higher than 70% (v/v) in the SP3 system. The mixture was incubated for over 18 min for thorough protein binding. A magnet was placed under the tube for 2 min to separate beads from the solution. Supernatant was carefully removed with the magnet on, and beads were rinsed with 10 μL of 70% ethanol two times and 10 μL 100% ACN one time. Eight μL of 100 mM NH_4HCO_3 (pH 8) was then added into the tube to resuspend the beads. One μL of trypsin solution in 100 mM NH_4HCO_3 (pH 8) with different concentrations (10 ng/μL, 2.5 ng/μL, and 1 ng/μL) was added to 50 ng, 10 ng, and 2 ng protein samples, respectively, for protein digestion (37 °C, overnight). After digestion, the supernatant (about 6 μL) was directly deposited into a nanoLC insert tube for nanoRPLC-MS/MS analysis. The experiment was performed twice at each sample amount.

For the Nano3 method, the basic workflow for sample processing is shown in Figure 1. A detailed workflow is shown in SI Figure S2. Sample transfer was performed through an empty capillary (100 μm i.d. × 12.5 cm length, 1 μL total volume). For the mouse brain samples, the proteins were loaded into the nanoreactor capillary not the intact cells. Reduced and alkylated mouse brain protein solution in the 2% SDS lysis buffer was diluted with 70% ACN into 50 ng/μL, 10 ng/μL, and 2 ng/μL. Each protein solution was pushed into the 1 μL transfer capillary, and the protein solutions in the transfer capillary containing 50 ng, 10 ng, and 2 ng proteins were pushed onto the nanoreactor through a syringe filled with 70% ACN for the Nano3 sample processing. After protein loading, beads were continuously rinsed with 70% ACN for an additional 8 μL to remove the SDS. One μL of trypsin solution (in 100 mM NH_4HCO_3) with the same concentration settings as the SP3 method for different amounts of mouse brain proteins was pushed into the nanoreactor through the transfer capillary to cover the nanoreactor. Both ends of the nanoreactor capillary were sealed with rubber, and the nanoreactor was incubated in a 37 °C water bath overnight for tryptic digestion. After digestion, the nanoreactor was rinsed with 5−6 μL of a buffer containing 2% ACN and 0.1% formic acid to elute the peptides into a nanoLC insert tube for nanoRPLC-MS/MS analysis. The experiment was performed twice at each sample amount.

Sample Processing of Few HeLa Cells With the Nano3 Method. The HeLa cell processing with the Nano3 method was similar to the mouse brain protein processing with some modifications, Figure 1. A detailed workflow is shown in SI Figure S2. The nanoreactor capillary was first filled with a buffer containing 2% SDS and 80% ACN. The syringe used for cell loading onto the nanoreactor was also filled with the same buffer. The cell solution was first transferred into the 1 μL transfer capillary to control the number of cells for processing. For the cell solutions with concentrations of 1000 cells/μL, 100 cells/μL, and 10 cells/μL, the number of cells in the transfer capillary was approximately 1000 cells, 100 cells, and 10 cells. The transfer capillary was connected to the nanoreactor capillary at one end and to the syringe filled with a buffer containing 2% SDS and 80% ACN at the other end. The 1 μL cell solution surrounded by the buffer containing SDS and ACN was pushed into the nanoreactor capillary. After that, both ends of the nanoreactor capillary was sealed with a rubber and the capillary was sonicated for 10 min, followed by incubation in a 95 °C water bath for 10 min for cell lysis and protein denaturation. Then the protein solution was pushed onto the nanoreactor for protein capturing, and the nanoreactor was flushed with at least 10 μL 70% ACN to remove SDS. The rest of the sample processing steps were the same as the mouse brain protein processing. Trypsin amount for digestion was 10 ng, 5 ng, and 2 ng for 1000 cells, 100 cells, and 10 cells, respectively. The experiment was performed twice for the 1000, 100, and 10 cells. We repeated the whole experiment one more time (batch 2), starting from cell culture, with duplicate preparation of 10, 100, and 1000 cell samples. In total, four biological replicates were performed for the 10−1000 cell samples. The sample insert tube for nanoRPLC-MS/MS analysis was pretreated with the BSA solution as described in our previous work.[33]

https://doi.org/10.1021/acs.analchem.1c01704
Anal. Chem. 2021, 93, 10568−10576

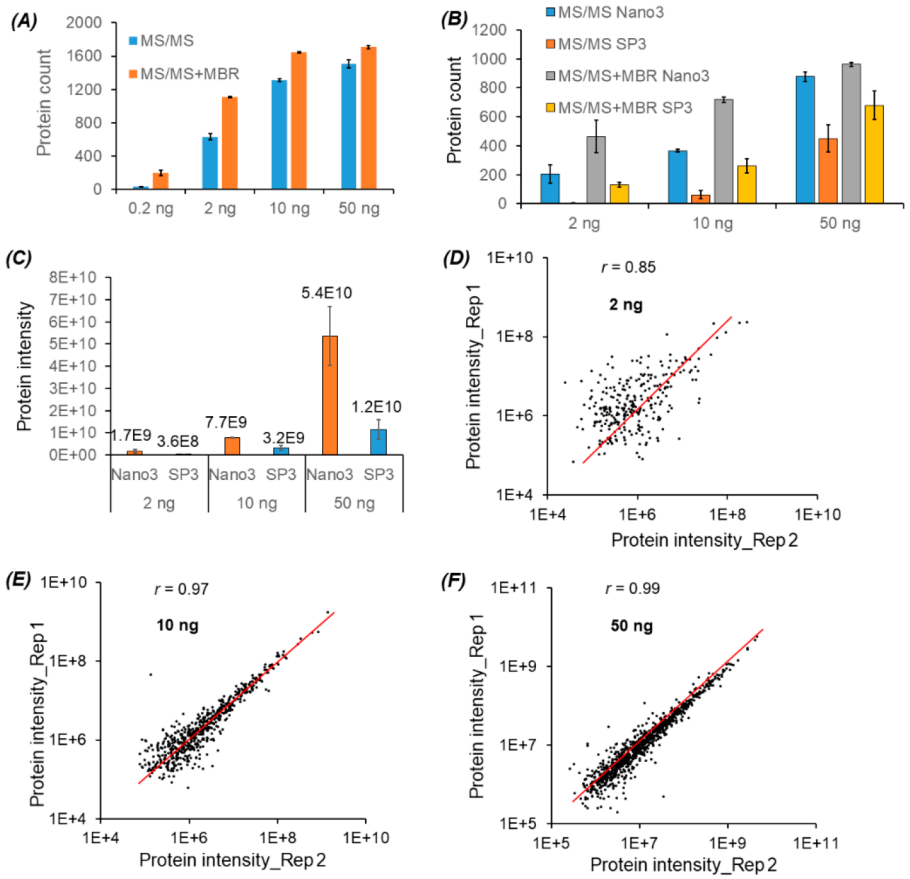


Figure 2. Summary of the data of low-nanograms of mouse brain samples. (A) The number of protein IDs as a function of the loaded peptide amount. The error bars represent the standard deviations of the number of protein IDs from triplicate LC-MS analyses. (B) The number of protein IDs from 2 to 50 ng of mouse brain protein samples processed by the Nano3 and SP3 methods. The error bars represent the standard deviations of the number of protein IDs from duplicate sample preparations. (C) Average protein intensity (MS/MS+MBR) from the 2−50 ng of mouse brain protein samples processed by the Nano3 and SP3 methods. The error bars represent the standard deviations of the average protein intensity from duplicate sample preparations. (D)−(F) Log−log plots of protein intensity (MS/MS+MBR) correlations between duplicate preparations of 2, 10, and 50 ng of mouse brain protein samples using the Nano3 method.

We also examined the number of HeLa cells in 1 μL of 10 cells/μL cell solutionundera microscope by depositing 1 μL of the cell solution onto a glass slide. The cell count varied from 3 cells to 20 cells in 1 μL of the solution across six different examinations with a median of 10 cells.

We also performed a blank experiment to confirm that the protein IDs from HeLa cells were not from contaminates in the cell culture medium. After cultured HeLa cells were washed with PBS, they were suspended in the PBS buffer for further analysis. After gentle centrifugation, the cells were pelleted, and a small aliquot of the supernatant was collected for processing with the Nano3 method. The steps for processing the blank sample are the same as the 10−1000 cell samples.

Direct Sampling of Mass-Limited HeLa Cell Lysates in the Eppendorf Tube for Processing with the Nano3 Method. A HeLa cell lysate was serially diluted to a concentration of 1 ng/μL with a buffer containing 2% SDS and 70% ACN. One μL of diluted protein solution containing only 1 ng of protein in mass was deposited into a 0.6 mL low retention Eppendorf tube for sample processing with the Nano3 method. A capillary (200 μm i.d.) containing a 100-μm-long nanoreactor was used. The capillary was first filled with 70% ACN. Because the capillary only contained a very small amount of paramagnetic beads, the backpressure was low enough for withdrawing the HeLa protein sample in the Eppendorf tube directly using a syringe pump (KD Scientific, Holliston, MA) operated in the withdraw mode, SI Figure S2. The protein solution was continuously drawn into the nanoreactor capillary until no solution was left in the tube. The protein solution in the nanoreactor capillary was then pushed through the nanoreactor for protein loading with a buffer containing 70% ACN and was further processed using a similar procedure to the mouse brain samples. One μL of 70% ACN was added into the Eppendorf tube for protein rinsing after the protein solution was completely drawn. The 1 μL of rinsing solution was also loaded onto the nanoreactor for sample processing. Triplicate sample preparations were performed. For digestion, 2 ng of trypsin was used.

NanoRPLC-MS/MS and Data Analysis. A Q-Exactive HF (Thermo Fisher Scientific) mass spectrometer was used. All MS raw files were processed with MaxQuant 1.5.5.1.[34] The match between runs (MBR) function was used.[35] The details were described in SI Supporting Information I.

RESULTS AND DISCUSSION

Comparisons of the SP3 and Nano3 Methods for Processing Low-Nanograms of a Complex Proteome Sample. Before we compared the SP3 and Nano3 methods

https://doi.org/10.1021/acs.analchem.1c01704
Anal. Chem. 2021, 93, 10568−10576

for mass-limited sample processing for bottom-up proteomics, we first evaluated the sensitivity of our nanoRPLC-MS/MS (Q-Exactive HF mass spectrometer) platform for the analysis of trace amounts of a complex proteome digest. A 100 μg aliquot of reduced and alkylated mouse brain protein sample was digested with trypsin following the SP3 procedure.[22,23] The digest was used for the system evaluation. A self-packed nanoRPLC capillary column (75 μm i.d. × 50 cm long, ReproSil-Pur, 120 Å, C18-AQ, 1.9 μm beads) was used for peptide separation.

Different amounts of mouse brain peptides ranging from 0.2 ng to 50 ng were loaded onto the nanoRPLC-MS/MS system for analysis in triplicate, and each LC-MS run used a 75 min gradient. As shown in Figure 2A, by MS/MS only, fewer than 30 proteins were identified when only a 0.2-ng protein digest was analyzed. That is about the amount of proteins from 1 to 2 single mammalian cells with a size of 10−20 μm.[18] The number of identified proteins increased to 627 when 2-ng of peptides were loaded. The numbers of protein IDs from 10-ng and 50-ng peptides were comparable (1313 vs. 1505). By performing the database search of all the raw files from 0.2 to 50 ng peptides together with the MaxQuant software and turning on the match between runs (MBR) function, the number of protein IDs from 0.2-ng peptides was boosted to 190, which is over 6-folds higher than that from MS/MS only. The number of protein IDs from the 2-ng peptide sample was also improved by nearly 80% compared to the data of MS/MS alone (1106 vs. 627). The data suggests that the MBR function is extremely useful for mass-limited samples. Also, the nanoRPLC-MS/MS system had nice reproducibility for analysis of the 0.2−50-ng peptide samples regarding the number of protein IDs from triplicate analyses.

After evaluating the sensitivity of our nanoRPLC-MS/MS system, we compared the performance of the SP3 and Nano3 methods for processing 2−50 ng of mouse brain proteins, Figure 2B. The Nano3 method clearly outperformed the SP3 method for processing 2−50-ng mouse brain proteins regarding the number of protein IDs. For example, the Nano3 method identified 40-times and 6-times higher number of proteins based on MS/MS than the SP3 method starting from 2-ng (206 vs. Five proteins) and 10-ng proteins (367 vs. 62), respectively. If we consider the protein IDs from both MS/MS and MBR, the Nano3 method still produced about 260% and 170% more protein IDs than the SP3 method from the 2-ng and 10-ng protein samples. For the 50-ng protein sample, the Nano3 generated 40% more protein IDs than the SP3 (963 vs. 679).

The drastically better overall performance of the Nano3 method compared to the SP3 method is most likely due to much higher sample recovery from the Nano3 approach, demonstrated by the substantially higher protein intensity produced by the Nano3 method, Figure 2C and SI Figure S3. The much better protein recovery is due to the substantially smaller sample processing volume of the Nano3 approach, leading to less sample loss because of adsorption to surfaces and higher protein concentration for more efficient tryptic digestion. The SP3 method is a good option for the preparation of submicrograms of proteome samples. It requires at least a 10-μL solution for processing samples to make sure that the paramagnetic beads are freely suspended.[23] In our experiment, the sample processing volume using the SP3 method was about 10 μL. The protein concentration in the 10 μL solution was 0.2−5 ng/μL for the 2−50-ng protein samples. For the Nano3 method, the proteins were digested in the nanoreactor, which had a 30 nL volume. The protein concentration in the nanoreactor was over 2 orders of magnitude higher than that during the SP3 processing. Minimized sample processing volume is a key point for the preparation of mass-limited proteome samples for large-scale proteome profiling.[36]

We need to highlight several advantages of the Nano3 method. First, like the SP3 method, the Nano3 method is compatible with various detergents and chaotropic reagents because proteins can be captured by the hydrophilic nanoparticles with high efficiency under a high ACN concentration environment, and detergents and chaotropic reagents can be removed via flushing with ACN. Second, the total volume of nanoreactor is only 30 nL in the experiment. The extremely small volume not only reduces protein loss during processing due to adsorption onto surfaces, but it also affords a relatively high protein concentration for digestion, ensuring high enzymatic digestion efficiency. We need to note that the volume of the nanoreactor can be further reduced easily via packing lower amounts of nanoparticles in the capillary or using capillaries with a smaller inner diameter. Third, the Nano3 technique employs a relatively closed environment for sample processing, making the protein and peptide storage easy before MS analysis.

The Nano3 method was quantitatively reproducible for processing low ng of mouse brain proteome samples regarding the intensity of quantified proteins from two replicates, Figures 2D-2F. For the 10-ng and 50-ng samples, the linear correlation coefficients of protein intensity were 0.97 and 0.99, respectively. For the 2-ng sample, a reasonable correlation coefficient of 0.85 was still achieved.

We further examined the peptide length and missed cleavage from the Nano3 method and compared them with those from the regular in-solution digestion and SP3 method. For the Nano3 and SP3, we selected the 10 ng sample for this purpose. For the normal in-solution digestion, over 100 μg mouse brain proteins were processed with a protein to trypsin mass ratio of 30:1, and 10 ng peptides were analyzed by the same nanoRPLC-MS/MS method. The Nano3 method performed as well as the SP3 and regular in-solution digestion regarding the peptide length and the number of missed cleavages of peptides, SI Figure S4.

Proteomics analysis of mass-limited samples requires both a highly efficient sample preparation method and a highly sensitive LC-MS/MS platform for peptide measurements. Our RPLC-MS/MS system with a 75-μm-i.d. capillary column only identified 26 ± 4 (N = 3) protein IDs based on MS/MS when analyzing 0.2-ng mouse brain peptides, Figure 2A. To further improve the sensitivity of the platform, we tested another capillary column with a smaller inner diameter of 50 μm and a length of 50 cm. A lower flow rate of 90 nL/min or 80 nL/min was employed for the 50-μm-i.d. column compared to the 75-μm-i.d. column. We increased the length of the gradient to 105 min because of the delay of the chromatography caused by the low flow rate. We analyzed 0.2 ng and 2 ng of the mouse brain digest using the new RPLC-MS/MS system with the 50-μm-i.d. column. As shown in SI Figure S5A, we identified 224 ± 24 (N = 2) proteins based on MS/MS from the 0.2-ng sample, which is about 10-times higher than that from the 75-μm-i.d. column. After employing the MBR function and using the 2-ng data as the library, we identified 640 ± 41 (N = 2) proteins from the 0.2-ng sample. The new RPLC-MS/MS system

https://doi.org/10.1021/acs.analchem.1c01704
Anal. Chem. 2021, 93, 10568−10576

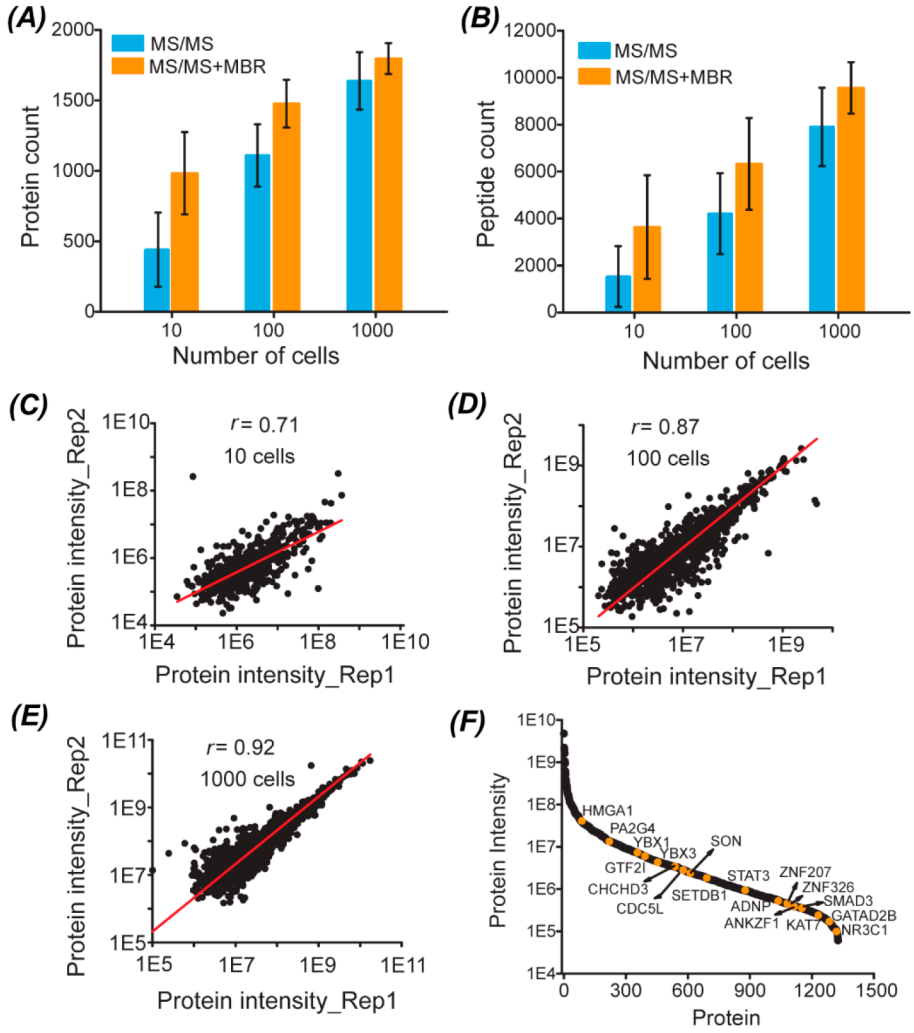


Figure 3. Summary of the data of 10−1000 HeLa cells processed by the Nano3 method. The number of identified proteins (A) and peptides (B) from 10, 100, 1000 HeLa cells. The error bars represent the standard deviations of the number of protein and peptide IDs from quadruplicate sample preparations (two batches and duplicate preparations per batch). (C)−(E) Log−log plots of protein intensity (MS/MS+MBR) correlations between duplicate preparations of 10, 100, and 1000 HeLa cells using the Nano3 method. (F) Distribution of protein intensity (MS/MS+MBR) from one 10-cells sample showing the protein intensity dynamic range. The highlighted ones are 18 transcription factors identified in the run.

showed a drastically better sensitivity for protein ID from mass-limited samples, and it was used for the following experiments. SI Figure S5B shows an example chromatogram of the 2-ng mouse brain peptide sample after analyzed by the new nanoRPLC-MS/MS system.

Application of the Nano3 Method in Processing 10−1000 HeLa Cells. The HeLa cell solution was serially diluted with the PBS buffer into three different concentrations (1000 cells/μL, 100 cells/μL, 10 cells/μL). Then 1 μL of the HeLa cell solutions with different concentrations were transferred into the nanoreactor capillary for sample processing with the Nano3 method, Figure 1. The approximate number of cells processed by the Nano3 method was 1000, 100, and 10 for the 1000 cells/μL, 100 cells/μL, and 10 cells/μL samples. We performed two replicates of the sample processing and repeated the whole experiment once starting from the cell culture. As shown in Figure 3A, we achieved 1638 ± 204 (n = 4), 1110 ± 221 (n = 4), and 441 ± 263 protein IDs by MS/MS from the 1000, 100, and 10 cells, respectively. Using the MBR function, the number of protein IDs was boosted by 123%, 33%, and 10% for the 10, 100, and 1000 cell samples compared to the MS/MS only data. 983 ± 292 (n = 4) proteins were identified from the 10 HeLa cells with MS/MS and MBR, corresponding to, on average, 3635 unique peptides. The numbers of identified peptides from the 10−1000 cells are shown in Figure 3B. We also processed one blank sample, and only 16 proteins were identified. The identified proteins from the 10−1000 cells are listed in Supporting Information II. The MS raw files have been deposited to the ProteomeXchange Consortium via the PRIDE[37] partner repository with the data set identifier PXD027135.

We further evaluated the quantitative reproducibility of the workflow for analyzing few HeLa cells with the protein intensity data, Figure 3C−E. Analyses of 100 and 1000 HeLa cells produced reasonably good reproducibility regarding the protein intensity with linear correlation coefficients (r) of 0.87 and 0.92. We observed significant variations of protein intensity across replicate processing of 10 HeLa cells, Figure 3C. We also determined higher variations of the number of protein and peptide IDs from the 10 cells compared to the 100 and 1000 cell samples, Figure 3A,B. We speculate that the higher variations of protein ID and intensity from the 10-cell

https://doi.org/10.1021/acs.analchem.1c01704
Anal. Chem. 2021, 93, 10568−10576

samples are due to two reasons. One is the much lower peptide/protein intensity from the 10-cell samples compared to 100−1000 cell samples due to the much smaller amounts of protein materials analyzed. The other reason relates to the significant variations of the number of cells in the samples. We injected 1 μL of 10 cells/μL solutions into the nanoreactor for processing, and we assumed that the cells were uniformly distributed in the solution, corresponding to 10 cells injected for sample preparation. To mimic the actual numbers of cells injected, we examined the cell count under a microscope by depositing 1 μL of 10 cells/μL solutions onto a glass slide. The cell count varied from 3 cells to 20 cells across six different examinations with a median of about 10 cells.

We need to highlight that the intensity of identified proteins from the 10-cell sample spanned across 5 orders of magnitude, and 18 transcription factors were confidently identified from only 10 HeLa cells, Figure 3F. The identified transcription factors are labeled in the figure. The transcription factors were determined by comparing the identified proteins from the 10-cell sample with a transcription factor database reported in the literature containing over 1600 transcription factors.[38] It has been estimated that in mammalian cells, transcription factors have a copy number of 10 000 to 300 000 per cell.[39] Peptides from 10 HeLa cells in our study were dissolved in about 5 μL prior of LC-MS analysis. From there, we estimated the concentration of transcription factors in the 5 μL solution was in the range of 33 fM to 1 pM, indicating a high sensitivity of the overall workflow.

We need to point out that the experiments of 10−1000 cells were carried out by employing a 1 μL transfer capillary for controlling the number of cells injected for the Nano3 processing. This approach is appropriate to evaluate the performance of the Nano3 method for the preparation of few mammalian cells. However, it may not be straightforward for processing mass-limited samples from laser capture microdissection (LCM) of tissues and cells isolated by fluorescence activated cell sorting (FACS) using the Nano3, because under those situations, the samples are usually transferred into Eppendorf tubes or wells. To demonstrate the potential of the Nano3 method for processing trace protein samples placed in Eppendorf tubes, we put a 1 μL aliquot of a HeLa cell lysate containing 1 ng proteins in a 0.6 mL low retention Eppendorf tube and processed the sample with the Nano3 method with the assistance of a syringe pump operated under a withdraw mode. The 1 ng HeLa cell proteins were dissolved in a buffer containing 2% SDS and 70% ACN.

In this experiment, we used a nanoreactor with a total length of only 100 μm with a total volume of 3 nL to reduce the backpressure of the nanoreactor for withdrawing the solution into the nanoreactor capillary directly using a syringe pump. After the 1-ng HeLa cell proteins were withdrawn into the nanoreactor capillary, the sample was processed using a similar approach as the mouse brain sample, SI Figure S2. After nanoRPLC-MS/MS analyses, 692 ± 182 proteins and 1997 ± 536 peptides (N = 3) were identified from the 1 ng HeLa cell protein sample based on MS/MS. 990 ± 149 proteins and 3494 ± 705 peptides (N = 3) were identified considering both MS/MS and MBR with the protein and peptide IDs from one 1000 HeLa cell sample as the library. The overall workflow has reasonable reproducibility across triplicate preparations of the 1 ng HeLa cell lysates regarding the number of identified proteins and peptides. The data here demonstrate the potential of using the Nano3 method for processing mass-limited samples from LCM or FACS.

CONCLUSIONS

In this pilot study, we developed a novel sample preparation method (Nano3) for processing trace complex proteome samples with high efficacy for bottom-up proteomics. The Nano3 method employed the basic concept of minimizing the sample processing volume. Proteins extracted from cells in a lysis buffer containing high-concentration SDS were concentrated and digested into peptides in the nanoreactor (≤30 nL total volume). The sample processing was performed in a relatively closed environment, facilitating the sample storage before LC-MS analysis. The Nano3 method identified a 40-times higher number of proteins based on MS/MS than the SP3 method starting from 2-ng mouse brain proteins, most likely due to its more than 100 times smaller sample processing volume than the SP3 method, reducing sample loss and improving the tryptic digestion. About 1000 proteins, including 18 transcription factors, were identified from only 10 HeLa cells processed by the Nano3 method, demonstrating the potential of the Nano3 method for advancing bottom-up proteomics of few mammalian cells.

We expect the number of protein IDs and quantitative reproducibility from few mammalian cells (e.g., 10 HeLa cells) processed by the Nano3 method can be boosted obviously through several improvements. First, in this proof-of-principle study, we did not systematically optimize the Nano3 method for processing few cells. We expect the sample recovery can be improved significantly after optimizing the volume of the nanoreactor, the trypsin concentration for digestion, the tryptic digestion time, and the procedure for peptide elution from the nanoreactor. Second, the current Nano3 method employs manual operations, which will introduce variations during sample preparation unavoidably. We will integrate the Nano3 technique with an automated sample preparation system to boost the quantitative reproducibility and sample processing throughput. Third, the number of protein ID and peptide signal can be increased by employing a liquid-phase separation-MS/MS system with much better sensitivity than the system used in the current study. For example, a nanoRPLC-MS/MS system with a 20-μm-i.d. or 30-μm-i.d. RPLC column and one of the most advanced mass spectrometers (e.g., Orbitrap Fusion Lumos) will be certainly helpful for pursuing a better proteome coverage with improved reproducibility from trace samples.[14,16] Capillary zone electrophoresis (CZE)-MS/MS could be a useful alternative for analyzing the trace samples processed by the Nano3 method because it outperformed nanoRPLC-MS/MS for the characterization of mass-limited samples regarding the number of protein ID and it has shown low zmole even ymole level limit of detections for peptides.[8,40−44] We believe that the optimized and automated Nano3 method coupled with an advanced liquid-phase separation-MS/MS system will be a useful tool for large-scale proteome profiling of few mammalian cells.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.1c01704.

The experimental details; TEM image of the magnetic beads used; Schematic workflow of the Nano3 sample

https://doi.org/10.1021/acs.analchem.1c01704
Anal. Chem. 2021, 93, 10568−10576

processing; Protein intensity (MS/MS only) from the 2−50 ng of mouse brain protein samples processed by the Nano3 and SP3 methods; Comparisons of length and missed cleavages of peptides from mouse brain proteins processed by Nano3, SP3 and regular in-solution method; Identification results from 2 ng and 0.2 ng of mouse brain peptides using optimized LC-MS platform (PDF)

Identified proteins from 10 to 1000 HeLa cells processed by the Nano3 method (XLSX)

AUTHOR INFORMATION

Corresponding Author

Liangliang Sun − *Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, United States;* orcid.org/0000-0001-8939-5042; Phone: 1-517-353-0498; Email: lsun@chemistry.msu.edu

Authors

Zhichang Yang − *Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, United States*

Zhaoran Zhang − *Department of Animal Sciences, Michigan State University, East Lansing, Michigan 48824, United States*

Daoyang Chen − *Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, United States*

Tian Xu − *Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, United States*

Yuan Wang − *Department of Animal Sciences, Michigan State University, East Lansing, Michigan 48824, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.1c01704

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Chen Chen's research group in the Department of Animal Science, Michigan State University for kindly providing us the mouse brains for the project. We thank the support from the National Institute of General Medical Sciences (NIGMS) through Grant R01GM125991, the National Science Foundation through Grant DBI1846913 (CAREER Award), and National Institute of Food and Agriculture (MICL02540) for the research project.

REFERENCES

(1) Nagaraj, N.; Wisniewski, J. R.; Geiger, T.; Cox, J.; Kircher, M.; Kelso, J.; Paabo, S.; Mann, M. *Mol. Syst. Biol.* **2011**, *7*, 548.

(2) Geiger, T.; Wehner, A.; Schaab, C.; Cox, J.; Mann, M. *Mol. Cell Proteomics* **2012**, *11*, M111 014050.

(3) Bekker-Jensen, D. B.; Kelstrup, C. D.; Batth, T. S.; Larsen, S. C.; Haldrup, C.; Bramsen, J. B.; Sorensen, K. D.; Hoyer, S.; Orntoft, T. F.; Andersen, C. L.; Nielsen, M. L.; Olsen, J. V. *Cell Syst.* **2017**, *4*, 587−599 e4.

(4) Bache, N.; Geyer, P. E.; Bekker-Jensen, D. B.; Hoerning, O.; Falkenby, L.; Treit, P. V.; Doll, S.; Paron, I.; Muller, J. B.; Meier, F.; Olsen, J. V.; Vorm, O.; Mann, M. *Mol. Cell Proteomics* **2018**, *17*, 2284−2296.

(5) Orre, L. M.; Vesterlund, M.; Pan, Y.; Arslan, T.; Zhu, Y.; Fernandez Woodbridge, A.; Frings, O.; Fredlund, E.; Lehtio, J. *Mol. Cell* **2019**, *73*, 166−182 e7.

(6) Zubarev, R. A. *Proteomics* **2013**, *13*, 723−6.

(7) Zubarev, R. A.; Makarov, A. *Anal. Chem.* **2013**, *85*, 5288−96.

(8) Sun, L.; Zhu, G.; Zhao, Y.; Yan, X.; Mou, S.; Dovichi, N. J. *Angew. Chem., Int. Ed.* **2013**, *52*, 13661−13664.

(9) Zhang, P.; Gaffrey, M. J.; Zhu, Y.; Chrisler, W. B.; Fillmore, T. L.; Yi, L.; Nicora, C. D.; Zhang, T.; Wu, H.; Jacobs, J.; Tang, K.; Kagan, J.; Srivastava, S.; Rodland, K. D.; Qian, W. J.; Smith, R. D.; Liu, T.; Wiley, H. S.; Shi, T. *Anal. Chem.* **2019**, *91*, 1441−1451.

(10) Shen, Y.; Tolic, N.; Masselon, C.; Pasa-Tolic, L.; Camp, D. G., 2nd; Hixson, K. K.; Zhao, R.; Anderson, G. A.; Smith, R. D. *Anal. Chem.* **2004**, *76*, 144−154.

(11) Li, S.; Plouffe, B. D.; Belov, A. M.; Ray, S.; Wang, X.; Murthy, S. K.; Karger, B. L.; Ivanov, A. R. *Mol. Cell Proteomics* **2015**, *14*, 1672−1683.

(12) Smith, R. D.; Shen, Y.; Tang, K. *Acc. Chem. Res.* **2004**, *37*, 269−278.

(13) Sun, X.; Kelly, R. T.; Tang, K.; Smith, R. D. *Analyst* **2010**, *135*, 2296−2302.

(14) Zhu, Y.; Piehowski, P. D.; Zhao, R.; Chen, J.; Shen, Y.; Moore, R. J.; Shukla, A. K.; Petyuk, V. A.; Campbell-Thompson, M.; Mathews, C. E.; Smith, R. D.; Qian, W. J.; Kelly, R. T. *Nat. Commun.* **2018**, *9*, 882.

(15) Schwanhausser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M. *Nature* **2011**, *473*, 337−342.

(16) Cong, Y.; Liang, Y.; Motamedchaboki, K.; Huguet, R.; Truong, T.; Zhao, R.; Shen, Y.; Lopez-Ferrer, D.; Zhu, Y.; Kelly, R. T. *Anal. Chem.* **2020**, *92*, 2665−2671.

(17) Brunner, A.-D.; Thielert, M.; Vasilopoulou, C.; Ammar, C.; Coscia, F.; Mund, A.; Horning, O. B.; Bache, N.; Apalategui, A.; Lubeck, M.; Raether, O.; Park, M. A.; Richter, S.; Fischer, D. S.; Theis, F. J.; Meier, F.; Mann, M., Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *bioRxiv* **2020**, 2020.12.22.423933.

(18) Cohen, D.; Dickerson, J. A.; Whitmore, C. D.; Turner, E. H.; Palcic, M. M.; Hindsgaul, O.; Dovichi, N. J. *Annu. Rev. Anal. Chem.* **2008**, *1*, 165−190.

(19) Li, Z. Y.; Huang, M.; Wang, X. K.; Zhu, Y.; Li, J. S.; Wong, C. C. L.; Fang, Q. *Anal. Chem.* **2018**, *90*, 5430−5438.

(20) Shao, X.; Wang, X.; Guan, S.; Lin, H.; Yan, G.; Gao, M.; Deng, C.; Zhang, X. *Anal. Chem.* **2018**, *90*, 14003−14010.

(21) Budnik, B.; Levy, E.; Harmange, G.; Slavov, N. *Genome Biol.* **2018**, *19*, 161.

(22) Hughes, C. S.; Foehr, S.; Garfield, D. A.; Furlong, E. E.; Steinmetz, L. M.; Krijgsveld, J. *Mol. Syst. Biol.* **2014**, *10*, 757.

(23) Hughes, C. S.; Moggridge, S.; Muller, T.; Sorensen, P. H.; Morin, G. B.; Krijgsveld, J. *Nat. Protoc.* **2019**, *14*, 68−85.

(24) Williams, S. M.; Liyu, A. V.; Tsai, C. F.; Moore, R. J.; Orton, D. J.; Chrisler, W. B.; Gaffrey, M. J.; Liu, T.; Smith, R. D.; Kelly, R. T.; Pasa-Tolic, L.; Zhu, Y. *Anal. Chem.* **2020**, *92*, 10588−10596.

(25) Zhang, Z.; Dubiak, K. M.; Huber, P. W.; Dovichi, N. J. *Anal. Chem.* **2020**, *92*, 5554−5560.

(26) Cong, Y.; Motamedchaboki, K.; Misal, S. A.; Liang, Y.; Guise, A. J.; Truong, T.; Huguet, R.; Plowey, E. D.; Zhu, Y.; Lopez-Ferrer, D.; Kelly, R. T. *Chem. Sci.* **2021**, *12*, 1001−1006.

(27) Wisniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. *Nat. Methods* **2009**, *6*, 359−362.

(28) Sielaff, M.; Kuharev, J.; Bohn, T.; Hahlbrock, J.; Bopp, T.; Tenzer, S.; Distler, U. *J. Proteome Res.* **2017**, *16*, 4060−4072.

(29) Müller, T.; Kalxdorf, M.; Longuespée, R.; Kazdal, D. N.; Stenzinger, A.; Krijgsveld, J. *Mol. Syst. Biol.* **2020**, *16*, No. e9111.

(30) Tian, R.; Wang, S.; Elisma, F.; Li, L.; Zhou, H.; Wang, L.; Figeys, D. *Mol. Cell Proteomics* **2011**, *10*, M110 000679.

(31) Zhao, Q.; Liang, Y.; Yuan, H.; Sui, Z.; Wu, Q.; Liang, Z.; Zhang, L.; Zhang, Y. *Anal. Chem.* **2013**, *85*, 8507−8512.

(32) Wang, F.; Wei, X.; Zhou, H.; Liu, J.; Figeys, D.; Zou, H. *Proteomics* **2012**, *12*, 3129−3137.

(33) Yang, Z.; Shen, X.; Chen, D.; Sun, L. *J. Proteome Res.* **2019**, *18*, 4046−4054.

(34) Cox, J.; Mann, M. *Nat. Biotechnol.* **2008**, *26*, 1367−1372.

(35) Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M. *Mol. Cell Proteomics* **2014**, *13*, 2513−2526.

(36) Yang, Z.; Sun, L. *Anal. Methods* **2021**, *13*, 1214−1225.
(37) Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Pérez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yilmaz, S.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaíno, J. A. *Nucleic Acids Res.* **2019**, *47* (D1), D442−D450.
(38) Lambert, S. A.; Jolma, A.; Campitelli, L. F.; Das, P. K.; Yin, Y.; Albu, M.; Chen, X.; Taipale, J.; Hughes, T. R.; Weirauch, M. T. *Cell* **2018**, *172*, 650−665.
(39) Biggin, M. D. *Dev. Cell* **2011**, *21*, 611−626.
(40) Zhu, G.; Sun, L.; Yan, X.; Dovichi, N. J. *Anal. Chem.* **2013**, *85*, 2569−2573.
(41) Ludwig, K. R.; Sun, L.; Zhu, G.; Dovichi, N. J.; Hummon, A. B. *Anal. Chem.* **2015**, *87*, 9532−9537.
(42) Wang, Y.; Fonslow, B. R.; Wong, C. C.; Nakorchevsky, A.; Yates, J. R., 3rd *Anal. Chem.* **2012**, *84*, 8505−8513.
(43) Lombard-Banek, C.; Moody, S. A.; Nemes, P. *Angew. Chem., Int. Ed.* **2016**, *55*, 2454−2458.
(44) Amenson-Lamar, E. A.; Sun, L.; Zhang, Z.; Bohn, P. W.; Dovichi, N. J. *Talanta* **2019**, *204*, 70−73.

https://doi.org/10.1021/acs.analchem.1c01704
Anal. Chem. 2021, 93, 10568−10576