Analytical Methods

Recent technical progress in sample preparation and liquid-phase separation-mass spectrometry for proteomic analysis of mass-limited samples

Journal:	*Analytical Methods*
Manuscript ID	AY-CRV-01-2021-000171.R1
Article Type:	Critical Review
Date Submitted by the Author:	14-Feb-2021
Complete List of Authors:	Yang, Zhichang; Michigan State University, Chemistry Sun, Liangliang; Michigan State University, Chemistry;

Recent technical progress in sample preparation and liquid-phase separation-mass spectrometry for proteomic analysis of mass-limited samples

Zhichang Yang and Liangliang Sun*

Department of Chemistry, Michigan State University, 578 S Shaw Lane, East Lansing, MI 48824 USA

* Corresponding author. Email: lsun@chemistry.msu.edu ; Phone: 517-353-0498

Abstract

Mass spectrometry (MS)-based proteomics has enabled the identification and quantification of thousands of proteins from complex proteomes in a single experiment. However, its performance for mass-limited proteome samples (*e.g.,* single cells and tissue samples from laser capture microdissection) is still not satisfying. The development of novel proteomic methodologies with better overall sensitivity is vital. During the last several years, substantial technical progress has been achieved for the preparation and liquid-phase separation-MS characterization of mass-limited proteome samples. In this review, we summarize recent technological progress of sample preparation, liquid chromatography (LC)-MS, capillary zone electrophoresis (CZE)-MS and MS instrumentation for bottom-up proteomics of trace biological samples, highlight some exciting applications of the novel techniques for single-cell proteomics, and provide a very brief perspective about the field at the end.

Mass spectrometry (MS)-based proteomics has become one of the major tools to measure protein molecules in a biological system globally. Studies on protein identity, quantity, post-translational modifications (PTMs), localization, temporal and spatial dynamics, and even structural assembly have been implemented using MS-based proteomics. One of the ultimate goals of proteomic research is comprehensive profiling of whole proteomes of cells. Typically, a comprehensive profiling study requires at least tens of micrograms of starting protein material. The requirement of a large amount of protein material is mainly due to two reasons: first, protein abundance in cells can have a very high dynamic range, from one copy per cell to millions of copies per cell, and second, the sensitivity of modern MS-based platform is not high enough.[1-3] The protein mass in a single human somatic cell is on the order of sub-nanogram,[4, 5] and tens of micrograms correspond to hundreds of thousands of human cells. In this kind of bulk studies, the proteome heterogeneity of individual cells is lost.

Proteomic measurement of individual cells is vital for understanding the roles played by cell-to-cell heterogeneity in fundamental biological processes and disease development.[6-8] Even in one cell type, various cell populations exist and have distinct gene expression profiles.[9] Significant cell-to-cell heterogeneity exists in cellular responses to dynamic environmental stimulus.[10] Even under a seemingly stable and identical environment, cells can display heterogeneous behaviors.[11] During the early embryogenesis, blastomeres in one embryo gradually differentiate from each other at the molecular level (*e.g.,* protein), establishing the foundation for organogenesis.[12, 13] Strong protein-level heterogeneity across cancer cells in one tumor makes drug development for cancer treatment challenging.[6, 14, 15] Technical breakthrough in developing extremely sensitive proteomic methodologies will enable global characterization of proteins in a small number of cells and even single cells, leading to substantial impact on the understanding of various biological questions in cancer biology, developmental biology, neuroscience, etc.

The state-of-the-art proteomic platforms have achieved the sensitivity of 1-100 zmol for proteins.[4, 16-21] Mammalian cells can express 12,000 to 15,000 different proteins spanning seven orders of magnitude in concentration. With the modern proteomic

techniques, over 4,000 proteins could be identified in theory from a single mammalian cell.[22] However, only few studies have reported the identification of hundreds of proteins from single human cells.[21, 23] One of the major challenges of comprehensive characterization of the mass-limited samples is the sample loss during the sample preparation and liquid-phase separation caused by dead adsorption of proteins/peptides on surfaces, such as processing containers, pipette tips, the stationary phase of liquid chromatography (LC) columns, and sample loading valves in LC, leading to low sample recovery from sample preparation and limited sensitivity of LC-MS for proteomics.

Researchers have been dedicating to explore novel sample processing and liquid-phase separation-MS methods to decrease the sample loss and boost the overall sensitivity of proteomics. Several review papers have been published and summarized the recent advances in proteomic studies of mass-limited samples, including the development in sample processing, fractionation, separation, and MS instrumentation.[24-29] In this review, we highlight the most recent technical progress of sample preparation, LC-MS, and capillary electrophoresis (CE)-MS, and MS instrumentation for proteomic analyses of mass-limited samples and some exciting single-cell proteomics studies of various biological questions.

1. Development of novel sample preparation methods with better sample recovery for proteomics

Reducing sample loss due to non-specific interactions between proteins/peptides and surfaces is critical in processing trace proteome samples. It can be achieved by minimizing sample processing volume, eliminating sample transferring, and blocking the non-specific interactions.

1.1 Minimizing sample processing volume

Controlling the sample processing in a minimal volume (*i.e.,* nanoliters) is one efficient approach for preparing mass-limited samples. In 2018, Zhu *et al.* developed a strategy called NanoPOTS (nanodroplet processing in one pot for trace samples) platform for proteomic analyses of small numbers of human cells.[21] Nanowells were manufactured on a glass chip with a hydrophilic surface and with diameters of 1 mm, **Figure 1 A**. All

steps of sample processing, including protein extraction, reduction and alkylation, and digestion, took place in the nanowells with a total volume of 200 nL. The NanoPOTS platform significantly reduced the reaction volume by 99.5% compared to regular sample processing in a 0.5-mL Eppendorf tube. In addition to the nanowell, the hydrophilic surface can reduce hydrophobic adsorption. Rapigest was used as the surfactant to facilitate cell lysis while did not affect trypsin digestion and MS signal.[30] The NanoPOTS has been applied to few-cells and even single-cell proteomics,[31] laser capture microdissected (LCM) tissues with a diameter low to 50 μm,[32] and circulating tumor cells.[33] Hundreds or over one thousand of proteins were identified from single HeLa cells with or without the match-between-runs algorithm.[34] The algorithm is integrated with the MaxQuant software.[35] The fact that the match-between-runs algorithm boosted the number of protein identifications drastically indicated the value of high mass accuracy and reproducible retention time of parent peptides on protein identification from trace proteome samples.

A nanoliter-scale oil-air-droplet (OAD) chip was introduced in 2018 by *Li et al.* for sample preparation of a low population of cells.[36] In a sandwiched device, a droplet containing cells was deposited on to a low retention millimeter-sized chip. It was isolated from the outside world by an oil layer to prevent evaporation of liquid, **Figure 1B**. Surface tension prevented the direct contact of the oil and sample droplet. Followed by the deposition of the cell droplet, a series of reagents were added into the droplet to implement cell lysis, protein reduction, alkylation, and digestion. The total droplet volume was estimated at around 550 nL. A C18 prepacked capillary was then inserted into the droplet. Peptides were directly loaded onto the capillary column from the droplet through pressurization, and online desalting was performed. Using this approach, 51 to 1,360 protein groups were identified from 1 to 100 HeLa cells. When comparing the OAD and NanoPOTS methods, we noted that the number of protein identifications from single HeLa cells using the OAD chip was much lower than that using the NanoPOTS platform (tens of proteins *vs.* hundreds of proteins). However, it doesn't necessarily indicate lower sensitivity or more sample loss of the OAD method compared to the NanoPOTS method because different LC-MS/MS systems were used for the two

methods (50-i.d. LC column plus Orbitrap Elite *vs.* 30-i.d. column plus Orbitrap Fusion Lumos or Orbitrap Eclipse Tribrid).

Recently, another on-chip based microfluidic device was developed by the Vinh group.[37] The device integrated an ultrafiltration membrane in a micro-reaction chamber with 1.2 μL volume in total. The membrane divided the chamber in half with one side (0.6 μL) for protein extraction and digestion. The membrane played the role of filtration membrane applied in the filter aided sample reparation (FASP) technique[38] for protein clean up and allowed clean peptides to be eluted for the downstream LC-MS/MS. A multi-reagent pump system and multi-way valve were integrated with the chipfilter device to deliver sample and reagent, **Figure 1C**. The chipfilter method achieved ten times higher sensitivity compared with the traditional FASP strategy when the starting protein material was 1 μg. Although the chipfilter was not applied on low numbers of human cells as NanoPOTS and the OAD chip, the chipfilter has great potential on preparing low numbers of cells. First, it is directly connected to the LC-MS/MS platform with no need for material transfer. Second, all the sample processing steps occur in one side of the reaction chamber upstream of the membrane, which is only 600 nL, leading to a comparable volume level as the NanoPOTS and OAD chip. Third, the excellent sealing of the chamber prevents liquid evaporation. Further sensitivity improvement is highly possible with a smaller chamber volume and a special surface treatment device to reduce dead adsorption.

An integrated proteome analysis device (iPAD), **Figure 1F**, was developed by the Zhang group in 2015 for ultrasensitive proteome profiling of only 100 living cancer cells.[39] Cells were suspended in a cold solution (4 °C) containing salts (NH_4HCO_3, guanidine hydrochloride, and EDTA) and trypsin at a known cell concentration. An exact proportion of cell solution containing 100 cells was directly drawn into a fused silica capillary loop (100 μm i.d.× 40 cm, 3.2 μL in volume). One hour of heat treatment (50 °C) facilitated the cell lysis and digestion in the capillary loop. The digested peptides were then directly loaded onto a trap column for further LC-MS analysis. 813 proteins were identified from the 100 cells. In 2018, Shao *et. al.* developed an optimized version of iPAD technique, iPAD-1, for proteomics analysis of single HeLa cells.[40] In iPAD-1, a

narrower capillary and simpler valve system were applied, **Figure 1G**. Single HeLa cell was aspirated into the capillary under microscope monitoring. The total volume of the processing capillary in iPAD-1 is only 20 nL (22 μm i.d. × 5 cm). In addition to heat treatment, an ultrasonication probe was placed close to the capillary to facilitate cell lysis and protein digestion. The processing capillary was then connected to an LC column through a union with zero dead volume for LC-MS analysis. Over 180 proteins were confidently identified with MS/MS from one HeLa cell. The ultrasensitive performance of the iPAD-1 device can be attributed to the tiny sample processing volume (20 nL) and the direct connection between sample processing capillary with LC-MS without sample transfer.

All the methods we discussed above handle liquid at the nanoliter level and require special liquid handling systems. For example, NanoPOTS requires 70% humidity to reduce liquid evaporation so the entire sample processing is implemented in a closed humid chamber. The OAD chip method requires an installation of a self-alignment monolithic device for droplet deposition, pressurization, and sample loading for LC-MS/MS. The chipfilter technique requires specialization of a multi-reagent pump and multi-way valve for reagent delivering and a good sealing reaction chamber for sample processing. The iPAD method also requires a multi-port valve to integrate the pumping system, processing capillary, and LC system together. The requirement of specific instrumentation limits the adaptability of the methods across different labs.

To overcome the instrumentation availability issue of the above-mentioned methods, a micro-FASP technique was published by Zhang *et al.* recently.[41] The strategy adapted the idea of conventional FASP but reduced the surface area of the filtration membrane to 0.1 mm^2. The membrane was integrated within a 20 μl pipette tip with bottom and top support for immobilization, **Figure 1 D**. The sample loading, buffer washing, and elution volume were all controlled at microliter level. Peptides were directly eluted into sample vials for LC-MS/MS analysis to avoid additional liquid transferring. The micro-FASP identified a comparable number of proteins and 20% more peptides compared to the conventional FASP when 10 times lower of protein amount (1 μg *vs* 10 μg) was consumed. Over 3,000 proteins were identified, starting with 1000 MCF-7 cells,

indicating the great potential of the micro-FASP method for proteomic analysis of few human cells. More importantly, it does not require special instrumentation for sample preparation.

1.2 Eliminating sample transferring

Another crucial point on treating trace amount of protein materials is to reduce liquid transferring as much as possible during sample processing. Proteomic sample preparation normally employs detergents (*e.g.,* SDS) or chaotropic reagent (*e.g.,* urea) to facilitate cell lysis and protein extraction. Most of the detergents and chaotropic reagents are incompatible with downstream enzymatic protein digestion and LC-MS analysis. Detergents usually need to be removed through ultrafiltration or precipitation, and chaotropic reagents need to be removed through desalting to ensure compatibility with follow-up LC-MS experiments. Those steps lead to limited sample loss when hundreds of micrograms of proteins are available but result in serious sample loss for trace protein materials.

Researchers have been searching for ways to avoid further sample clean up to reduce sample loss. Budnik *et al.* used water as lysis buffer, lysed single cells through mechanical sonication, and denatured proteins through high temperature.[2] Since chemicals were obviated, no further clean-up was applied. Detergent Rapigest was applied in the preparation of mass-limited proteome samples because Rapigest is compatible with enzymatic digestion and is degraded into non-interfering products under an acidic condition.[21] Organic solvent trifluoroethanol was also employed for the preparation of mass-limited samples.[4, 42] Trifluoroethanol can assist the cell lysis and protein denaturation. More importantly, it can be removed easily by lyophilization.

In 2014, Hughes *et al.* introduced a single-pot solid-phase-enhanced sample preparation (SP3) method.[43] The method utilizes strong detergents (*e.g.,* SDS) for protein extraction from cells. The cell lysates are incubated with carboxyl-coated paramagnetic nanoparticles under a high concentration of acetonitrile (>70%). Proteins are captured on the beads through hydrophilic interaction, and detergents can be removed efficiently via washing multiple times with organic solvents. Then the captured proteins are digested by enzymes on beads, followed by peptide elution from beads

using an aqueous solution for LC-MS/MS analysis, **Figure 1E**. The paramagnetic nanoparticles have a large surface area, ensuring strong interaction with proteins and supporting extensive washing. The SP3 method achieves sample preparation in a single Eppendorf tube, reduces sample loss during sample transferring, and enables the preparation of mass-limited samples for proteomic analysis. Multiple studies have used the SP3 method for processing the mass- limited samples and achieved good sensitivity on proteome profiling. Hughes *et al*. processed single *Drosophila* embryos containing only 200 ng of proteins and identified almost 3,000 proteins when using SP3 for sample processing.[43] Griesser *et al.* applied SP3 to process proteins extracted from formalin-fixed paraffin-embedded (FFPE) sample. From tissue containing about 3,000 cells, over 5,600 proteins were confidently quantified.[44] Yang *et al.* identified 3,600 proteins from a protein amount equal to 1,000 HEK-293t cells using CE-MS/MS, when SP3 was used for sample processing.[45] We need to point out that the SP3 method is operated in Eppendorf tubes and requires microliter-level solutions for sample preparation, resulting in a limited performance for the preparation of low nanograms of proteome samples.

1.3 Blocking the non-specific interactions

Blocking the non-specific adsorption of proteins/peptides on the surfaces of Eppendorf tubes, sample vials, and even beads packed in the LC columns has also been implemented in proteomic analyses of mass-limited samples. Yang *et al.* proved that treatment of the sample vials with bovine serum albumin (BSA) improved the number of protein identifications (IDs) by 40% and peptide intensity by 4-fold compared to that without treatment when only nanograms of peptides were in the sample vials for MS analysis.[45] BSA peptides were also used to block the dead adsorption of beads packed in reversed-phase LC columns before the columns were used for LC-MS/MS analysis of trace amounts of peptides.[46] Similar treatment of LC columns with an *E. coli* digest was also reported when mammalian proteome samples were analyzed.[36] Dou *et al.* applied *n*-Dodecyl *β*-D-maltoside (DDM) (0.01%) as a collection buffer additive in the process of fraction collection and achieved significantly higher numbers of protein IDs compared to that without the DDM additive.[47] This outcome is because of the reduced interaction between peptides and the collection device surface in the presence of low-concentration

DDM.[48] Another effective approach for reducing non-specific binding between peptides to the potentially interacting surface is to introduce carrier peptides into the peptide samples. The carrier peptides have a much higher concentration than the peptides in the samples and are responsible for blocking the dead adsorption. One of the perfect examples is the Single Cell ProtEomics by Mass Spectrometry (SCoPE-MS) method developed by Budnik *et al.*[2] In the SCoPE-MS method, one of the tandem mass tag (TMT) channels is used to label the carrier sample, and other TMT channels are used to label the target mass-limited samples (*i.e.,* single mammalian cells). The carrier sample derives from the same source of protein materials as the single-cell samples but with a significantly higher peptide amount. The use of a high concentration of TMT-labeled carrier sample not only reduces the peptide loss of mass-limited samples but also boosts the peptide signal in mass spectra, facilitating the identification and quantification of peptides from the mass-limited samples. In addition, the involvement of TMT labeling also improves the throughput of proteomic analysis of trace proteome samples, such as single cells.[49, 50]

2. Development of novel LC-MS and CE-MS methods for highly sensitive proteomics

Coupling liquid-phase separations (LC and CE) to MS and MS/MS plays vital roles in comprehensive and sensitive profiling of complex proteomes with high concentration dynamic ranges. Millions of different peptide molecules exist in a typical proteome digest. Highly efficient separation of peptides before MS and MS/MS reduces ionization suppression and boosts the sensitivity of peptide measurement.

RPLC-MS is routinely used for proteomics. The flow rate of mobile phase can have a significant effect on the electrospray ionization efficiency of peptides and usually, a low flow rate yields better ionization efficiency.[51, 52] Researchers have deployed RPLC columns with narrow inner diameter (i.d.) to accommodate lower flow rate of LC separation and to increase the sensitivity of LC-MS. It has been pointed out that the sensitivity of LC-MS can be boosted by decreasing the inner diameter (d) of the RPLC column from d_1 to d_2, and the improvement factor (f) is equal to d_1^2/d_2^2.[53] The i.d. of RPLC columns has been decreased from 75 μm as the most routinely used for RPLC-

MS to 30 μm or smaller. [17, 18, 21, 23, 54, 55] The flow rate is on the order of tens of nanoliter per min to even picoliter per min. Extremely high sensitivity has been achieved for bottom-up proteomics by applying such small i.d. RPLC columns. In 2004, a trace proteomics study was performed by coupling a 15-μm-i.d. RPLC column with an 11.4 T FT-ICR mass spectrometer. The flow rate of the system was controlled at 20 nL/min, and a peak capacity of 103 was achieved.[17] About 14% of the total proteome of *Deinococcus radiodurans* were confidently identified when only 2.5 ng of peptides were consumed. The sensitivity was proved to be at the low zmol level, corresponding to thousands of peptide molecules. Recently, 20-μm-i.d. RPLC columns have been applied to the characterization of peptides of single human cells processed by the NanoPOTS method, enabling the identification of over 300 proteins from single cells by MS/MS.[23] Porous layer open tube (PLOT) columns with small i.d. have also been employed for proteomics of mass-limited samples.[18, 54] In PLOT, a poly(styrene-divinylbenzene) in-situ polymerization was implemented in a 10-μm-i.d. capillary column with polymer thickness of 1-2 μm, **Figure 2A**. The thin wall of the polymer as stationary phase significantly reduces the resistance to mass transfer and boosts the separation efficiency.[54] By applying a 10-μm i.d. PLOT column, 4,000 proteins were identified with the consumption of peptides corresponding to 100-200 cells.[18] Recently, Xiang *et al.* introduced an extremely narrow open tubular (NOT) column with only 2-μm i.d. for trace proteomics analysis.[55] The 2-μm-i.d. NOT column was operated at a flow rate lower than 1 nL/min and achieved over 1,000 protein IDs when only 75 pg of tryptic peptides were loaded onto the system, demonstrating the tremendous potential for single-cell proteomics. The PLOT/NOT columns still have some limitations. For example, they could be clogged easily due to the extremely small i.d. and further improvement on robustness is essential before their widespread adoption.

Besides RPLC-MS, capillary zone electrophoresis (CZE)-MS has also shown great potential for proteomic analysis of mass-limited samples. Multiple studies have compared the performance of CZE-MS and RPLC-MS for discovery proteomics in terms of sensitivity.[56-60] It has been found that CZE-MS outperforms RPLC-MS regarding the number of peptide and protein IDs when the sample size is smaller than 10 ng. The higher sensitivity of CZE-MS can attribute to several reasons. First, CZE can achieve

very high separation efficiency for large biomolecules. One million of theoretical plates have been achieved by CZE for separation of proteins.[61] Second, the flow rate in CZE separation is on the order of low nL/min when capillaries with neutral coatings are employed, ensuring the high electrospray ionization efficiency of peptides and proteins. Third, the absence of stationary phase and direct sample injection from a sample vial without a valve and transferring tubing reduce sample loss. Low zmole peptide detection limits have been reported using CZE-MS.[16, 62]

CZE-MS typically has an excellent mass detection limit but a poor concentration detection limit because of the low sample loading capacity of CZE (low nL). The low sample loading capacity had impeded CZE-MS/MS for large-scale proteomics of mass-limited samples. During recent years, various online concentration approaches have been evaluated for boosting the sample loading capacity of CZE-MS for large-scale proteomics of nanograms of proteome samples. Sun *et al.* applied CZE-MS/MS for large-scale proteomic analysis of a HeLa cell proteome digest with a filed enhanced sample stacking method for increasing the sample loading volume (100 nL), identifying 10,000 peptides and 2,000 proteins from only 400 ng of Hela digest.[63] Chen *et al.* optimized the dynamic pH junction-based sample stacking method for CZE-MS/MS according to work reported by the Dovichi group[64] and reported a CZE-MS/MS system with a microliter-scale sample loading volume and an over 2-hour separation window, establishing the foundation of using CZE-MS/MS for large-scale proteomics.[65] More recently, the dynamic pH junction-based CZE-MS/MS has shown great potential for large-scale proteomics of mass-limited samples.[45, 66] The CZE-MS/MS system identified on average 100 proteins via consuming only 250 pg of an MCF7 proteome digest, corresponding to the protein content of roughly one MCF7 cell in mass.[66] Yang *et al.* reported the identification of over 6,500 proteins from an MCF7 cell lysate starting with only 500-ng peptides via coupling the nanoflow RPLC (nanoRPLC) fractionation with the dynamic pH junction-based CZE-MS/MS.[45] The well orthogonal separations of nanoRPLC and CZE for peptides guarantee the high peak capacity of the nanoRPLC-CZE-MS/MS platform for bottom-up proteomics, **Figure 2B**. Coupling the SP3 sample preparation method and nanoRPLC-CZE-MS/MS enabled the identification of nearly 4,000 proteins from 5,000 HEK293T cells with the consumption of a peptide amount

that corresponded to only roughly 1,000 cells. One technical challenge still exists for using CZE-MS/MS for mass-limited samples and it is how to fully use all the available mass-limited materials for the measurements.

Coupling SPME (solid phase microextraction) to CZE-MS/MS could be an efficient approach for mass-limited proteome samples because it theoretically can eliminate the limitation of sample loading volume of CZE as peptides can be preconcentrated by the SPME material before CZE-MS. In addition, the eluted sample plug from SPME can be further online concentrated with transient isotachophoresis (tITP)[58] and dynamic pH junction via carefully choosing the elution buffer and separation buffer.[67, 68] The SPME-CZE-MS has shown over 3-folds more protein IDs than nanoRPLC-MS when only 5-ng peptides were used, clearly demonstrating the advantage of SPME-CZE-MS for trace proteome samples.[58] Zhang *et al.* reported the identification of over 1,000 proteins from 50-ng of *Xenopus laevis* proteome digest with the SPME-CZE-MS/MS.[68] However, the robustness of the SPME-CZE-MS/MS needs to be improved, and SPME-CZE-MS/MS needs to be coupled with the sample preparation methods mentioned above for validating its performance for analyses of mass-limited proteome samples.

Both CZE-MS and nanoRPLC-MS have demonstrated the power of large-scale proteomics of mass-limited samples. We believe that better separation of peptides will lead to higher overall sensitivity of bottom-up proteomics, benefiting the characterization of mass-limited samples. Our recent successes in coupling nanoRPLC to the dynamic pH junction-based CZE-MS/MS for highly sensitive proteomics further prove this point.[45] We expect that improved versions of the nanoRPLC-CZE-MS/MS system via employing smaller i.d. RPLC columns and the SPME-CZE-MS/MS will certainly advance the proteomic analysis of mass-limited samples substantially.

3. Advances of MS instrumentation for highly sensitive proteomics

Advances of MS instrumentation in the last decade have played a crucial role in enabling ultrasensitive characterization of proteins from complex but mass-limited samples.

Ion mobility MS (IMS) separates gas-phase ions by their mobility within gas media. It provides an additional dimension of separation of analytes besides liquid-phase separations (*i.e.,* LC/CE) and MS. High field asymmetric waveform ion mobility spectrometry (FAIMS) separates gas-phase ions based on their characteristic difference in mobility in a high and low electric field.[69] FAIMS provides some advantages for mass-limited yet complex samples such as single cells, including reducing mass spectrum background noise and providing reproducible online fractionations of ions.[70] FAIMS was not widely applied in MS instrumentation until recently because it attenuated ion signal up to one order of magnitude, and because the long ion transmission time of FAIMS resulted in delay switching of compensation voltage (CV). In 2018, Pfammatter *et al.* introduced a novel FAIMS device with shorter CV switch time and improved sensitivity, and integrated it with the Orbitrap Tribrid mass spectrometer.[71] The novel FAIMS-Orbitrap platform achieved better detection of low-abundance peptides that were underrepresented in the platform without FAIMS and effectively reduced ratio compression effect in TMT quantification.[72] FAIMS-Orbitrap has shown superior performance on proteomic analysis of trace samples. For example, more than 1,000 proteins were identified from 5 ng of a HeLa digest with only a 5-min gradient.[3] In another example, more than 1,000 proteins were identified by MS/MS from a single HeLa cell using FAIMS-Orbitrap, which is 2.3 times more protein IDs than the platform without FAIMS.[73]

Another IMS, trapped ion mobility MS (TIMS), has also been integrated with Time-of-Flight (TOF) MS for proteomic analysis of mass-limited samples. Unlike FAIMS, TIMS separates gas-phase ions based on their collision cross sections. The separation is typically performed within 10s or 100s milliseconds.[74] TOF-MS is famous for its extremely high scan speed, making it perfect to be coupled with TIMS since ions that emerge from TIMS can be efficiently sampled by TOF analyzer. Bruker's timsTOF pro is a representative of interfacing TIMS with TOF-MS. By applying Parallel Accumulation - SErial Fragmentation (PASEF)[75] in timsTOF pro, peptide sequencing speed is significantly improved without loss of sensitivity. By using timsTOF pro, more than 2,500 proteins were identified from 10 ng of a HeLa digest within a 30-min acquisition time.[76] Very recently, timsTOF pro was applied for a single-cell proteomics study. Over 800

proteins were identified from single HeLa cells, and over 420 single HeLa cells were analyzed, representing one of the largest data sets of single-cell proteomics.[77]

Orbitrap-based instruments have been widely used by the proteomics community. Since the first commercial instrument incorporated with Orbitrap analyzer was introduced in 2005, successive generations of Orbitrap-based instruments have been developed with improved scan speed, sensitivity, and resolution.[78] Up to 40-Hz acquisition rate (at resolution setting of 7,500 at m/z 200) and half million resolution (m/z 200) were achieved in the latest model of Orbitrap instrumentation. The superior resolving power, scan rate, and sensitivity make Orbitrap instrumentation the most widely used MS in proteomic analysis of trace materials, especially single cells.[23,36,73]

Not only MS instrumentation, the data acquisition strategies also play vital roles for sensitive and reproducible characterizations of trace protein materials. Compared to conventional data-dependent acquisition (DDA), the data-independent acquisition (DIA) strategy has unique advantages, especially for proteomics analysis of trace protein materials. Saha-Shah *et al.* systematically compared DDA and DIA for proteomics analyses of single blastomeres from *Xenopus laevis* embryos (1-cell to 128-cell stages), and observed that DIA had better sensitivity when tiny quantities of proteome samples were analyzed and exhibited great quantitative reproducibility.[79]

4. Applications of the techniques to single-cell proteomics

The advance of techniques regarding sample preparation, separation and MS instrumentation enables researchers to explore proteome difference at single-cell resolution. Lombard-Banek *et al.* used a capillary microprobe to sample cellular proteins from single blastomeres in *Xenopus* early-stage embryos, followed by CZE-MS/MS-based label-free quantitative bottom-up proteomics.[80] Significant proteome differences were observed at the single-cell level between blastomeres collected from the animal pole and vegetal pole, **Figure 3A**. Brunner *et al.* processed FACS isolated single cells in 384-well plate with very small processing volume (1-2 μL) and applied an improved LC-MS platform to study proteome difference in cells at different cell-cycle stages.[77] Cell-cycle of HeLa cells was arrested by drug treatment to produce four cell populations enriched in specific cell-cycle stages. Although all cells were HeLa cells, significant

proteome differences were observed between cells at different cell-cycle stages, **Figure 3B**. This single-cell proteomics result also reflected different protein amounts in cells at different proliferation states by summarizing identified peptides intensity, **Figure 3C**. Significant cell-to-cell heterogeneity at the proteome level was also demonstrated in a single cell proteomics study of progenitor cell and descendant cell.[81] In the study, Zhu *et al*. processed individual hair cells and its progenitor, supporting cells, with NanoPOTS, and performed single-cell proteomics of these two kinds of cells with LC-MS equipped with an ultranarrow bore separation column (30 μm i.d.). By using FM1-43 as a labeling reagent (a membrane probe for identifying actively firing neurons), hair cells and supporting cells were distinguished through fluorescence-activated cell sorting, based on the fact that hair cells can be labeled more strongly by FM1-43 than supporting cells. The NanoPOTS-assisted single-cell proteomics study identified 60 proteins from a single hair cell and 600 proteins from a pool of 20 hair cells. By checking the identified proteins from a pool of 20 cells in each population (hair and supporting cells), different proteins were significantly enriched in different cell types, **Figure 3D**. Specht *et al.* applied SCoPE2, an optimized version of SCoPE method, for proteomics analyses of 1,490 single cells (monocytes and macrophage cells) and quantified over 3,000 proteins from those single cells.[82] Principle component analysis of the large single-cell-proteomics data set clearly separated the two cell types and revealed cellular heterogeneity, **Figure 3E.**

Single-cell proteomics can shed light on the molecular mechanisms of cell differentiation. In the Lombard-Banek's work, cellular proteins were sampled from a blastomere called midline animal-dorsal cell (at 16-cell stage) and its descendant cells at different cell-division stages (at 32, 64, and 128-cell stages) prior to mid-blastula transition (MBT).[80] Before MBT, there is no *de novo* transcription. Since all descendant cells arise from the same founder cell, it is surprising to see the proteome change over embryo development as shown in the hierarchical cluster analysis result, **Figure 3F**. In the Zhu's work, although less than 100 proteins were identified from a single cell due to the extremely small cell size, with sufficient sample size, developmental trajectories from various protein expression patterns at the single-cell level can be established, **Figure**

3G.[81] These single-cell proteomics studies revealed the protein expression dynamics during cell differentiation.

Conclusions

Proteomics of mass-limited samples has been advanced aggressively in recent years because of significant technical progress in sample preparation, liquid-phase separation, and MS instrumentation. It took over 20 years from the detection of hemoglobin from single erythrocytes[83] to the identification of hundreds of proteins from a single HeLa cell[23] using MS. We expect that with further advancement of sample preparation methods and nanoRPLC/CE-MS platforms, the sensitivity and throughput of proteomics will be improved drastically, enabling routine proteomic characterization of mass-limited samples (*e.g.,* single cells) with high proteome coverage.

We also need to note that in almost all the proteomic studies of mass-limited samples, bottom-up proteomics is employed due to the fact that it has drastically better sensitivity than the top-down proteomics that directly measures intact proteins (proteoforms) instead of peptides. However, the most recent top-down proteomic studies of mass-limited proteome samples demonstrated the identification and quantification of thousands of proteoforms from nanograms of proteome samples using CZE-MS/MS[61] as well as the identification of over 600 proteoforms from only ~770 HeLa cells using nanoRPLC-MS/MS and NanoPOTS-based sample preparation.[48] The data suggest the potential of using top-down proteomics for the characterization of mass-limited proteome samples, even single cells.

Acknowledgments

We thank the support from the National Science Foundation (CAREER Award, Grant DBI1846913) and the National Institutes of Health (Grant R01GM125991).

Conflict of Interest Statement

The authors have no conflict of interest related to this work.

Reference

1. Zubarev, R. A., The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics* **2013,** *13* (5), 723-6.

2. Budnik, B.; Levy, E.; Harmange, G.; Slavov, N., SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome biology* **2018,** *19* (1), 161.

3. Bekker-Jensen, D. B.; Martinez-Val, A.; Steigerwald, S.; Ruther, P.; Fort, K. L.; Arrey, T. N.; Harder, A.; Makarov, A.; Olsen, J. V., A Compact Quadrupole-Orbitrap Mass Spectrometer with FAIMS Interface Improves Proteome Coverage in Short LC Gradients. *Molecular & cellular proteomics : MCP* **2020,** *19* (4), 716-729.

4. 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., Carrier-Assisted Single-Tube Processing Approach for Targeted Proteomics Analysis of Low Numbers of Mammalian Cells. *Analytical chemistry* **2019,** *91* (2), 1441-1451.

5. Wisniewski, J. R.; Hein, M. Y.; Cox, J.; Mann, M., A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. *Molecular & cellular proteomics : MCP* **2014,** *13* (12), 3497-506.

6. Cohen, A. A.; Geva-Zatorsky, N.; Eden, E.; Frenkel-Morgenstern, M.; Issaeva, I.; Sigal, A.; Milo, R.; Cohen-Saidon, C.; Liron, Y.; Kam, Z.; Cohen, L.; Danon, T.; Perzov,

N.; Alon, U., Dynamic proteomics of individual cancer cells in response to a drug. *Science* **2008,** *322* (5907), 1511-6.

7. Semrau, S.; van Oudenaarden, A., Studying lineage decision-making in vitro: emerging concepts and novel tools. *Annual review of cell and developmental biology* **2015,** *31*, 317-45.

8. Symmons, O.; Raj, A., What's Luck Got to Do with It: Single Cells, Multiple Fates, and Biological Nondeterminism. *Molecular cell* **2016,** *62* (5), 788-802.

9. Georgiev, H.; Ravens, I.; Benarafa, C.; Forster, R.; Bernhardt, G., Distinct gene expression patterns correlate with developmental and functional traits of iNKT subsets. *Nature communications* **2016,** *7*, 13116.

10. Kalisky, T.; Quake, S. R., Single-cell genomics. *Nature methods* **2011,** *8* (4), 311-4.

11. Luo, T.; Fan, L.; Zhu, R.; Sun, D., Microfluidic Single-Cell Manipulation and Analysis: Methods and Applications. *Micromachines* **2019,** *10* (2).

12. Kimmel, C. B.; Ballard, W. W.; Kimmel, S. R.; Ullmann, B.; Schilling, T. F., Stages of embryonic development of the zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists* **1995,** *203* (3), 253-310.

13. Wagner, D. E.; Weinreb, C.; Collins, Z. M.; Briggs, J. A.; Megason, S. G.; Klein, A. M., Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **2018,** *360* (6392), 981-987.

14. Marusyk, A.; Almendro, V.; Polyak, K., Intra-tumour heterogeneity: a looking glass for cancer? *Nature reviews. Cancer* **2012,** *12* (5), 323-34.

15. Dagogo-Jack, I.; Shaw, A. T., Tumour heterogeneity and resistance to cancer therapies. *Nature reviews. Clinical oncology* **2018,** *15* (2), 81-94.

16. Sun, L.; Zhu, G.; Zhao, Y.; Yan, X.; Mou, S.; Dovichi, N. J., Ultrasensitive and fast bottom-up analysis of femtogram amounts of complex proteome digests. *Angewandte Chemie* **2013,** *52* (51), 13661-4.

17. Shen, Y.; Tolic, N.; Masselon, C.; Pasa-Tolic, L.; Camp, D. G., 2nd; Hixson, K. K.; Zhao, R.; Anderson, G. A.; Smith, R. D., Ultrasensitive proteomics using high-efficiency on-line micro-SPE-nanoLC-nanoESI MS and MS/MS. *Analytical chemistry* **2004,** *76* (1), 144-54.

18. Li, S.; Plouffe, B. D.; Belov, A. M.; Ray, S.; Wang, X.; Murthy, S. K.; Karger, B. L.; Ivanov, A. R., An Integrated Platform for Isolation, Processing, and Mass Spectrometry-based Proteomic Profiling of Rare Cells in Whole Blood. *Molecular & cellular proteomics : MCP* **2015,** *14* (6), 1672-83.

19. Smith, R. D.; Shen, Y.; Tang, K., Ultrasensitive and quantitative analyses from combined separations-mass spectrometry for the characterization of proteomes. *Accounts of chemical research* **2004,** *37* (4), 269-78.

20. Sun, X.; Kelly, R. T.; Tang, K.; Smith, R. D., Ultrasensitive nanoelectrospray ionization-mass spectrometry using poly(dimethylsiloxane) microchips with monolithically integrated emitters. *The Analyst* **2010,** *135* (9), 2296-302.

21. 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., Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. *Nature communications* **2018,** *9* (1), 882.

22. Schwanhausser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M., Global quantification of mammalian gene expression control. *Nature* **2011,** *473* (7347), 337-42.

23. Cong, Y.; Liang, Y.; Motamedchaboki, K.; Huguet, R.; Truong, T.; Zhao, R.; Shen, Y.; Lopez-Ferrer, D.; Zhu, Y.; Kelly, R. T., Improved Single-Cell Proteome Coverage Using Narrow-Bore Packed NanoLC Columns and Ultrasensitive Mass Spectrometry. *Analytical chemistry* **2020,** *92* (3), 2665-2671.

24. Altelaar, A. F.; Heck, A. J., Trends in ultrasensitive proteomics. *Current opinion in chemical biology* **2012,** *16* (1-2), 206-13.

25. Kelly, R. T., Single-cell Proteomics: Progress and Prospects. *Molecular & cellular proteomics : MCP* **2020,** *19* (11), 1739-1748.

26. Feist, P.; Hummon, A. B., Proteomic challenges: sample preparation techniques for microgram-quantity protein analysis from biological samples. *International journal of molecular sciences* **2015,** *16* (2), 3537-63.

27. Levy, E.; Slavov, N., Single cell protein analysis for systems biology. *Essays in biochemistry* **2018,** *62* (4), 595-605.

28. Zhu, Y.; Piehowski, P. D.; Kelly, R. T.; Qian, W. J., Nanoproteomics comes of age. *Expert review of proteomics* **2018,** *15* (11), 865-871.

29. Yi, L.; Piehowski, P. D.; Shi, T.; Smith, R. D.; Qian, W. J., Advances in microscale separations towards nanoproteomics applications. *Journal of chromatography. A* **2017,** *1523*, 40-48.

30. Pop, C.; Mogosan, C.; Loghin, F., Evaluation of Rapigest Efficacy for the Digestion of Proteins from Cell Cultures and Heart Tissue. *Clujul medical* **2014,** *87* (4), 258-62.

31. Zhu, Y.; Clair, G.; Chrisler, W. B.; Shen, Y.; Zhao, R.; Shukla, A. K.; Moore, R. J.; Misra, R. S.; Pryhuber, G. S.; Smith, R. D.; Ansong, C.; Kelly, R. T., Proteomic Analysis of Single Mammalian Cells Enabled by Microfluidic Nanodroplet Sample Preparation and Ultrasensitive NanoLC-MS. *Angewandte Chemie* **2018,** *57* (38), 12370-12374.

32. Zhu, Y.; Dou, M.; Piehowski, P. D.; Liang, Y.; Wang, F.; Chu, R. K.; Chrisler, W. B.; Smith, J. N.; Schwarz, K. C.; Shen, Y.; Shukla, A. K.; Moore, R. J.; Smith, R. D.; Qian, W. J.; Kelly, R. T., Spatially Resolved Proteome Mapping of Laser Capture Microdissected Tissue with Automated Sample Transfer to Nanodroplets. *Molecular & cellular proteomics : MCP* **2018,** *17* (9), 1864-1874.

33. Zhu, Y.; Podolak, J.; Zhao, R.; Shukla, A. K.; Moore, R. J.; Thomas, G. V.; Kelly, R. T., Proteome Profiling of 1 to 5 Spiked Circulating Tumor Cells Isolated from Whole Blood Using Immunodensity Enrichment, Laser Capture Microdissection, Nanodroplet Sample Processing, and Ultrasensitive nanoLC-MS. *Analytical chemistry* **2018,** *90* (20), 11756-11759.

34. Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M., Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & cellular proteomics : MCP* **2014,** *13* (9), 2513-26.

35. Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M., Andromeda: a peptide search engine integrated into the MaxQuant environment. *Journal of proteome research* **2011,** *10* (4), 1794-805.

36. Li, Z. Y.; Huang, M.; Wang, X. K.; Zhu, Y.; Li, J. S.; Wong, C. C. L.; Fang, Q., Nanoliter-Scale Oil-Air-Droplet Chip-Based Single Cell Proteomic Analysis. *Analytical chemistry* **2018,** *90* (8), 5430-5438.

37. Ndiaye, M. M.; Ta, H. P.; Chiappetta, G.; Vinh, J., On-Chip Sample Preparation Using a ChipFilter Coupled to NanoLC-MS/MS for Bottom-Up Proteomics. *Journal of proteome research* **2020,** *19* (7), 2654-2663.

38. Wisniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M., Universal sample preparation method for proteome analysis. *Nature methods* **2009,** *6* (5), 359-62.

39. Chen, Q.; Yan, G.; Gao, M.; Zhang, X., Ultrasensitive Proteome Profiling for 100 Living Cells by Direct Cell Injection, Online Digestion and Nano-LC-MS/MS Analysis. *Analytical chemistry* **2015**, 87 (13), 6674-80.

40. Shao, X.; Wang, X.; Guan, S.; Lin, H.; Yan, G.; Gao, M.; Deng, C.; Zhang, X., Integrated Proteome Analysis Device for Fast Single-Cell Protein Profiling. *Analytical chemistry* **2018**, 90 (23), 14003-14010

41. Zhang, Z.; Dubiak, K. M.; Huber, P. W.; Dovichi, N. J., Miniaturized Filter-Aided Sample Preparation (MICRO-FASP) Method for High Throughput, Ultrasensitive Proteomics Sample Preparation Reveals Proteome Asymmetry in Xenopus laevis Embryos. *Analytical chemistry* **2020,** *92* (7), 5554-5560.

42. Wang, H.; Qian, W. J.; Mottaz, H. M.; Clauss, T. R.; Anderson, D. J.; Moore, R. J.; Camp, D. G., 2nd; Khan, A. H.; Sforza, D. M.; Pallavicini, M.; Smith, D. J.; Smith, R. D., Development and evaluation of a micro- and nanoscale proteomic sample preparation method. *Journal of proteome research* **2005,** *4* (6), 2397-403.

43. Hughes, C. S.; Foehr, S.; Garfield, D. A.; Furlong, E. E.; Steinmetz, L. M.; Krijgsveld, J., Ultrasensitive proteome analysis using paramagnetic bead technology. *Molecular systems biology* **2014,** *10*, 757.

44. Griesser, E.; Wyatt, H.; Ten Have, S.; Stierstorfer, B.; Lenter, M.; Lamond, A. I., Quantitative Profiling of the Human Substantia Nigra Proteome from Laser-capture Microdissected FFPE Tissue. *Molecular & cellular proteomics : MCP* **2020,** *19* (5), 839-851.

45. Yang, Z.; Shen, X.; Chen, D.; Sun, L., Improved Nanoflow RPLC-CZE-MS/MS System with High Peak Capacity and Sensitivity for Nanogram Bottom-up Proteomics. *Journal of proteome research* **2019,** *18* (11), 4046-4054.

46. Kulak, N. A.; Geyer, P. E.; Mann, M., Loss-less Nano-fractionator for High Sensitivity, High Coverage Proteomics. *Molecular & cellular proteomics : MCP* **2017,** *16* (4), 694-705.

47. Dou, M.; Tsai, C. F.; Piehowski, P. D.; Wang, Y.; Fillmore, T. L.; Zhao, R.; Moore, R. J.; Zhang, P.; Qian, W. J.; Smith, R. D.; Liu, T.; Kelly, R. T.; Shi, T.; Zhu, Y., Automated Nanoflow Two-Dimensional Reversed-Phase Liquid Chromatography System Enables In-Depth Proteome and Phosphoproteome Profiling of Nanoscale Samples. *Analytical chemistry* **2019,** *91* (15), 9707-9715.

48. Zhou, M.; Uwugiaren, N.; Williams, S. M.; Moore, R. J.; Zhao, R.; Goodlett, D.; Dapic, I.; Pasa-Tolic, L.; Zhu, Y., Sensitive Top-Down Proteomics Analysis of a Low Number of Mammalian Cells Using a Nanodroplet Sample Processing Platform. *Analytical chemistry* **2020,** *92* (10), 7087-7095.

49. Tsai, C. F.; Zhao, R.; Williams, S. M.; Moore, R. J.; Schultz, K.; Chrisler, W. B.; Pasa-Tolic, L.; Rodland, K. D.; Smith, R. D.; Shi, T.; Zhu, Y.; Liu, T., An Improved Boosting to Amplify Signal with Isobaric Labeling (iBASIL) Strategy for Precise Quantitative Single-cell Proteomics. *Molecular & cellular proteomics : MCP* **2020,** *19* (5), 828-838.

50. Dou, M.; Clair, G.; Tsai, C. F.; Xu, K.; Chrisler, W. B.; Sontag, R. L.; Zhao, R.; Moore, R. J.; Liu, T.; Pasa-Tolic, L.; Smith, R. D.; Shi, T.; Adkins, J. N.; Qian, W. J.; Kelly, R. T.; Ansong, C.; Zhu, Y., High-Throughput Single Cell Proteomics Enabled by Multiplex Isobaric Labeling in a Nanodroplet Sample Preparation Platform. *Analytical chemistry* **2019,** *91* (20), 13119-13127.

51. Wilm, M.; Mann, M., Analytical properties of the nanoelectrospray ion source. *Analytical chemistry* **1996,** *68* (1), 1-8.

52. Marginean, I.; Tang, K.; Smith, R. D.; Kelly, R. T., Picoelectrospray ionization mass spectrometry using narrow-bore chemically etched emitters. *Journal of the American Society for Mass Spectrometry* **2014,** *25* (1), 30-6.

53. Meiring, H. D.; Heeft, E.; Hove, G. J.; Jong, A., Nanoscale LC-MS(n): Technical design and applications to peptide and protein analysis. *Journal of Separation Science* **2002,** *25*, 557-568.

54. Yue, G.; Luo, Q.; Zhang, J.; Wu, S. L.; Karger, B. L., Ultratrace LC/MS proteomic analysis using 10-microm-i.d. Porous layer open tubular poly(styrene-divinylbenzene) capillary columns. *Analytical chemistry* **2007,** *79* (3), 938-46.

55. Xiang, P.; Zhu, Y.; Yang, Y.; Zhao, Z.; Williams, S. M.; Moore, R. J.; Kelly, R. T.; Smith, R. D.; Liu, S., Picoflow Liquid Chromatography-Mass Spectrometry for Ultrasensitive Bottom-Up Proteomics Using 2-mum-i.d. Open Tubular Columns. *Analytical chemistry* **2020,** *92* (7), 4711-4715.

56. Faserl, K.; Sarg, B.; Kremser, L.; Lindner, H., Optimization and evaluation of a sheathless capillary electrophoresis-electrospray ionization mass spectrometry platform for peptide analysis: comparison to liquid chromatography-electrospray ionization mass spectrometry. *Analytical chemistry* **2011,** *83* (19), 7297-305.

57. Li, Y.; Champion, M. M.; Sun, L.; Champion, P. A.; Wojcik, R.; Dovichi, N. J., Capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry as an alternative proteomics platform to ultraperformance liquid chromatography-electrospray ionization-tandem mass spectrometry for samples of intermediate complexity. *Analytical chemistry* **2012,** *84* (3), 1617-22.

58. Wang, Y.; Fonslow, B. R.; Wong, C. C.; Nakorchevsky, A.; Yates, J. R., 3rd, Improving the comprehensiveness and sensitivity of sheathless capillary electrophoresis-tandem mass spectrometry for proteomic analysis. *Analytical chemistry* **2012,** *84* (20), 8505-13.

59. Zhu, G.; Sun, L.; Yan, X.; Dovichi, N. J., Single-shot proteomics using capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry with production of more than 1250 Escherichia coli peptide identifications in a 50 min separation. *Analytical chemistry* **2013,** *85* (5), 2569-73.

60. Sun, L.; Li, Y.; Champion, M. M.; Zhu, G.; Wojcik, R.; Dovichi, N. J., Capillary zone electrophoresis-multiple reaction monitoring from 100 pg of RAW 264.7 cell lysate digest. *The Analyst* **2013,** *138* (11), 3181-8.

61. Lubeckyj, R. A.; Basharat, A. R.; Shen, X.; Liu, X.; Sun, L., Large-Scale Qualitative and Quantitative Top-Down Proteomics Using Capillary Zone Electrophoresis-Electrospray Ionization-Tandem Mass Spectrometry with Nanograms of Proteome Samples. *Journal of the American Society for Mass Spectrometry* **2019,** *30* (8), 1435-1445.

62. Amenson-Lamar, E. A.; Sun, L.; Zhang, Z.; Bohn, P. W.; Dovichi, N. J., Detection of 1zmol injection of angiotensin using capillary zone electrophoresis coupled to a Q-Exactive HF mass spectrometer with an electrokinetically pumped sheath-flow electrospray interface. *Talanta* **2019,** *204*, 70-73.

63. Sun, L.; Hebert, A. S.; Yan, X.; Zhao, Y.; Westphall, M. S.; Rush, M. J.; Zhu, G.; Champion, M. M.; Coon, J. J.; Dovichi, N. J., Over 10,000 peptide identifications from the HeLa proteome by using single-shot capillary zone electrophoresis combined with tandem mass spectrometry. *Angewandte Chemie* **2014,** *53* (50), 13931-3.

64. Zhu, G.; Sun, L.; Yan, X.; Dovichi, N. J., Bottom-up proteomics of Escherichia coli using dynamic pH junction preconcentration and capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry. *Analytical chemistry* **2014,** *86* (13), 6331-6.

65. Chen, D.; Shen, X.; Sun, L., Capillary zone electrophoresis-mass spectrometry with microliter-scale loading capacity, 140 min separation window and high peak capacity for bottom-up proteomics. *The Analyst* **2017,** *142* (12), 2118-2127.

66. Yang, Z.; Shen, X.; Chen, D.; Sun, L., Microscale Reversed-Phase Liquid Chromatography/Capillary Zone Electrophoresis-Tandem Mass Spectrometry for Deep and Highly Sensitive Bottom-Up Proteomics: Identification of 7500 Proteins with Five Micrograms of an MCF7 Proteome Digest. *Analytical chemistry* **2018,** *90* (17), 10479-10486.

67. Zhang, Z.; Yan, X.; Sun, L.; Zhu, G.; Dovichi, N. J., Detachable strong cation exchange monolith, integrated with capillary zone electrophoresis and coupled with pH gradient elution, produces improved sensitivity and numbers of peptide identifications during bottom-up analysis of complex proteomes. *Analytical chemistry* **2015,** *87* (8), 4572-7.

68. Zhang, Z.; Sun, L.; Zhu, G.; Cox, O. F.; Huber, P. W.; Dovichi, N. J., Nearly 1000 Protein Identifications from 50 ng of Xenopus laevis Zygote Homogenate Using Online Sample Preparation on a Strong Cation Exchange Monolith Based Microreactor Coupled with Capillary Zone Electrophoresis. *Analytical chemistry* **2016,** *88* (1), 877-82.
69. Kolakowski, B. M.; Mester, Z., Review of applications of high-field asymmetric waveform ion mobility spectrometry (FAIMS) and differential mobility spectrometry (DMS). *Analyst* **2007**, *132* (9), 842-64.
70. Swearingen, K. E.; Moritz, R. L., High-field asymmetric waveform ion mobility spectrometry for mass spectrometry-based proteomics. *Expert review of proteomics* **2012**, *9* (5), 505-17.
71. Pfammatter, S.; Bonneil, E.; McManus, F. P.; Prasad, S.; Bailey, D. J.; Belford, M.; Dunyach, J. J.; Thibault, P., A Novel Differential Ion Mobility Device Expands the Depth of Proteome Coverage and the Sensitivity of Multiplex Proteomic Measurements. *Molecular & cellular proteomics* **2018**, *17* (10), 2051-2067.
72. Pfammatter, S.; Bonneil, E.; Thibault, P., Improvement of Quantitative Measurements in Multiplex Proteomics Using High-Field Asymmetric Waveform Spectrometry. *Journal of proteome research* **2016**, *15* (12), 4653-4665.
73. 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., Ultrasensitive single-cell proteomics workflow identifies >1000 protein groups per mammalian cell. *Chemical Science* **2021**,*12*, 1001-1006.
74. Eiceman, G. A.; Karpas, Z.; Hill, H. H., Ion mobility spectrometry. CRC press: Boca Raton, Fla. [etc.], 2016.
75. Meier, F.; Beck, S.; Grassl, N.; Lubeck, M.; Park, M. A.; Raether, O.; Mann, M., Parallel Accumulation–Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device. *Journal of proteome research* **2015**, *14* (12), 5378-5387.
76. Meier, F.; Brunner, A. D.; Koch, S.; Koch, H.; Lubeck, M.; Krause, M.; Goedecke, N.; Decker, J.; Kosinski, T.; Park, M. A.; Bache, N.; Hoerning, O.; Cox, J.; Rather, O.; Mann, M., Online Parallel Accumulation-Serial Fragmentation (PASEF) with a Novel

Trapped Ion Mobility Mass Spectrometer. *Molecular & cellular proteomics* **2018**, *17* (12), 2534-2545.

77. 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.

78. Eliuk, S.; Makarov, A., Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annual review of analytical chemistry* **2015**, *8*, 61-80.

79. Saha-Shah, A.; Esmaeili, M.; Sidoli, S.; Hwang, H.; Yang, J.; Klein, P. S.; Garcia, B. A., Single Cell Proteomics by Data-Independent Acquisition To Study Embryonic Asymmetry in Xenopus laevis. *Analytical chemistry* **2019**, *91* (14), 8891-8899.

80. Lombard-Banek, C.; Moody, S. A.; Manzini, M. C.; Nemes, P., Microsampling Capillary Electrophoresis Mass Spectrometry Enables Single-Cell Proteomics in Complex Tissues: Developing Cell Clones in Live Xenopus laevis and Zebrafish Embryos. *Analytical chemistry* **2019,** *91* (7), 4797-4805.

81. Zhu, Y.; Scheibinger, M.; Ellwanger, D. C.; Krey, J. F.; Choi, D.; Kelly, R. T.; Heller, S.; Barr-Gillespie, P. G., Single-cell proteomics reveals changes in expression during hair-cell development. *eLife* **2019,** *8*.

82. Specht, H.; Emmott, E.; Petelski, A. A.; Huffman, R. G.; Perlman, D. H.; Serra, M.; Kharchenko, P.; Koller, A.; Slavov, N., Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome biology* **2021**, *22* (1), 50.

83. Hofstadler, S. A.; Swanek, F. D.; Gale, D. C.; Ewing, A. G.; Smith, R. D., Capillary electrophoresis-electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry for direct analysis of cellular proteins. *Analytical chemistry* **1995,** *67* (8), 1477-80.

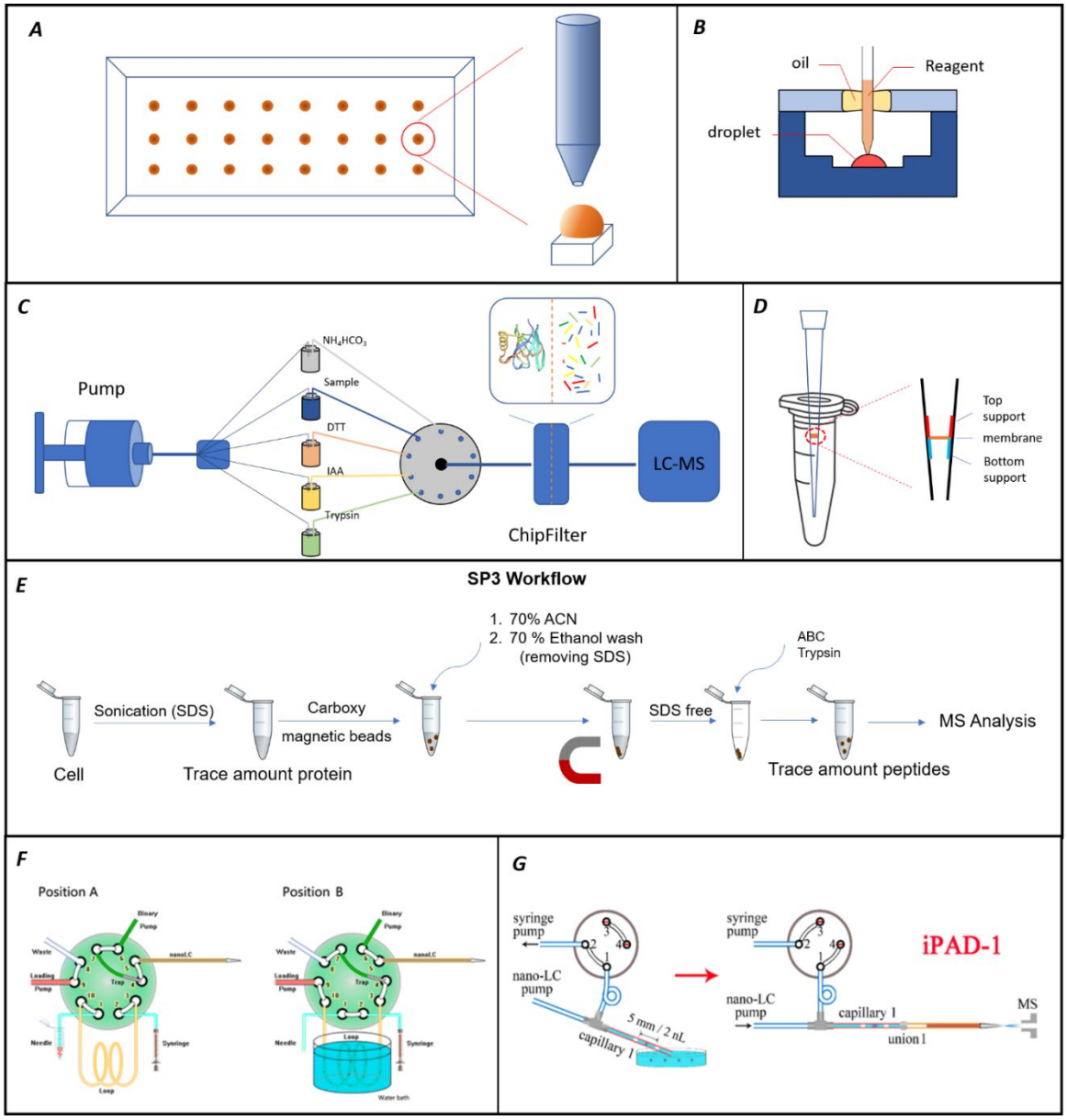

Figure 1. Schematic diagrams of different sample preparation methods. (A) NanoPOTS. Reproduced from ref. 21 with permission from Springer Nature, copyright (2018); (B) OAD. Reproduced from ref. 36 with permission from American Chemical Society, copyright (2018); (C) ChipFilter. Reproduced from ref. 37 with permission from American Chemical Society, copyright (2020); (D) MicroFASP. Reproduced from ref. 41 with permission from American Chemical Society, copyright (2020); (E) SP3; (F) iPAD. Reproduced from ref. 39 with permission from American Chemical Society, copyright (2015); (G) iPAD-1. Reproduced from ref. 40 with permission from American Chemical Society, copyright (2018).

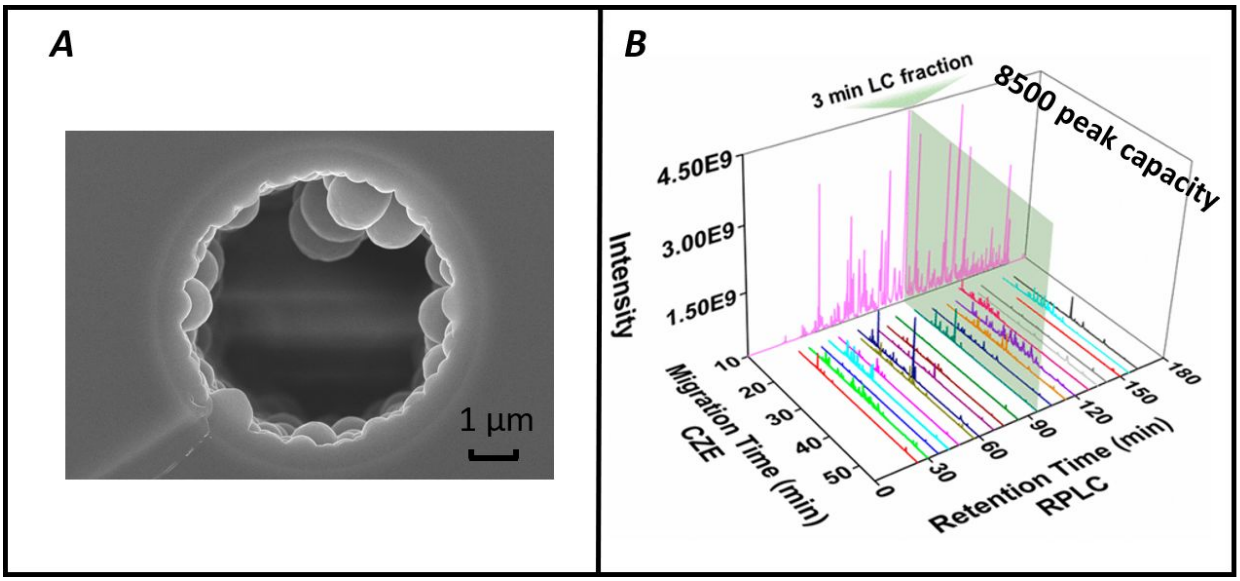

Figure 2. (A) SEM image of the cross section of a PLOT column. (B) The orthogonal separations of peptides using nanoRPLC and CZE. Reproduced from ref. 45 with permission from American Chemical Society, copyright (2019).

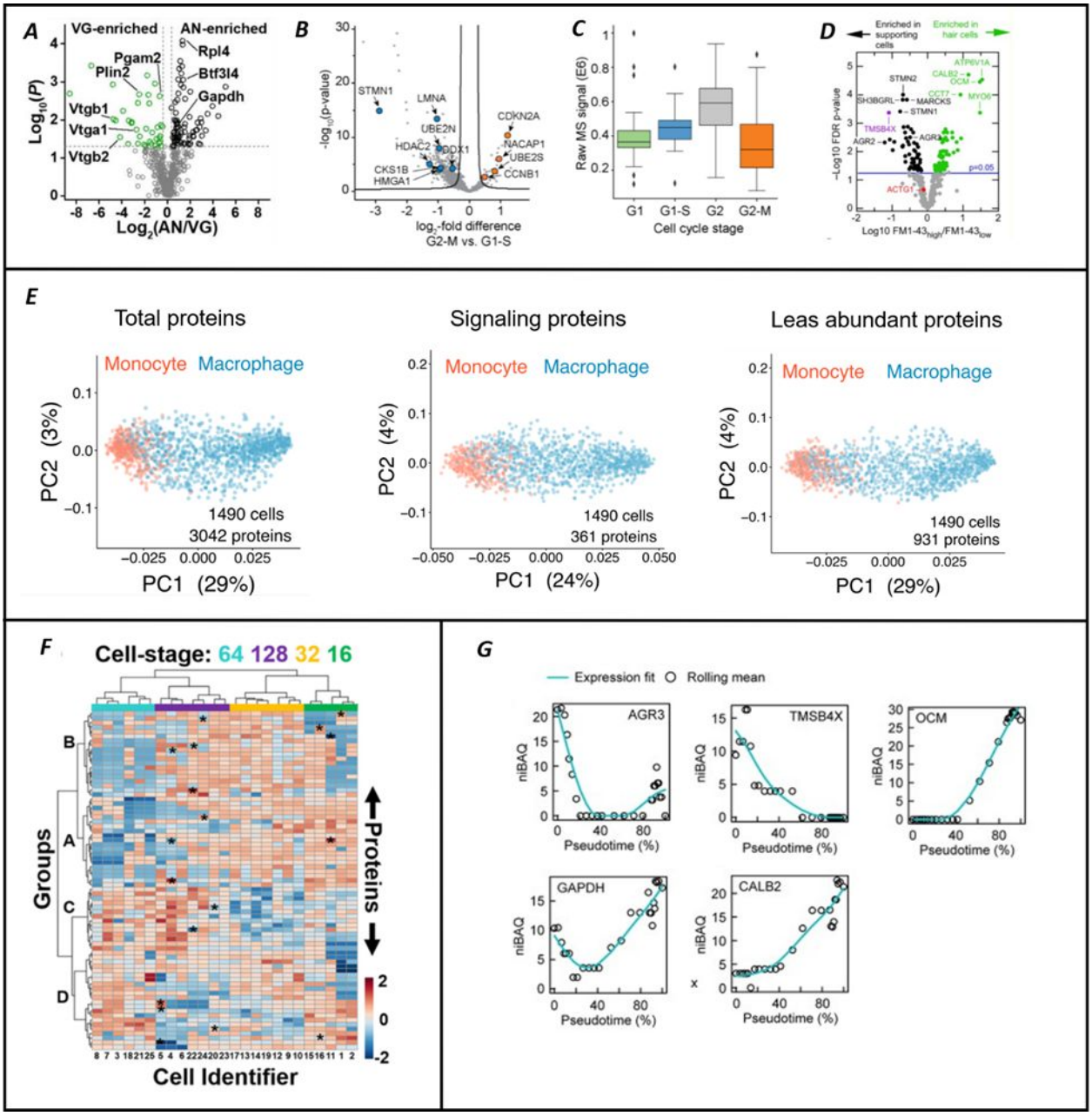

Figure 3. (A) Volcano plot revealing significant proteomic differences between blastomeres from animal and vegetal poles. Reproduced from ref. 80 with permission from American Chemical Society, copyright (2019). (B) Volcano plot of the quantified proteins from single cells in the two drug arrested states. Reproduced from ref. 77 with permission from bioRxiv, copyright (2020). (C) Violin plot of total protein signals of the analyzed single cells in the indicated cell cycle stages as enriched by the drug

treatments. Reproduced from ref. 77 with permission from bioRxiv, copyright (2020). (D) Volcano plot showing significant proteomic differences between FM1-43$_{high}$/FM1-43$_{low}$ cells (hair cell/supporting cell). Reproduced from ref. 81 with permission from eLife, copyright (2019). (E) Weighted principal component analysis (PCA) of 1490 single cells using all 3,042 proteins, signaling proteins and least abundant proteins quantified across single cells. Cells are colored by cell type. The more spread-out of macrophage cells indicates significant cellular heterogeneity after differentiation from homogeneous monocytes cells. Reproduced from ref. 82 with permission from BioMed Central, copyright (2021). (F) Hierarchical cluster analysis-heat map of quantified proteins from single *Xenopus* blastomeres isolated from various developmental stages. Reproduced from ref. 80 with permission from American Chemical Society, copyright (2019). (G) Absolute expression dynamics (log2 niBAQ) of 5 proteins as a function of pseudotime. Reproduced from ref. 81 with permission from eLife, copyright (2019).