Capillary Electrophoresis Mass Spectrometry for Scalable Single-Cell Proteomics

Bowen Shen†, Leena R. Pade†, Sam B. Choi†, Pablo Muñoz-LLancao‡,3,4, M. Chiara Manzini‡,3 and Peter Nemes†*

1Department of Chemistry and Biochemistry, University of Maryland, College Park, MD, United States, 2Pharmacology and Physiology, School of Medicine and Health Sciences, The George Washington University, Washington, DC, United States, 3Department of Neuroscience and Cell Biology and Child Health Institute of New Jersey, Rutgers Robert Wood Johnson Medical School, The State University of New Jersey, New Brunswick, NJ, United States, 4Department of Cell Biology, Yale University School of Medicine, New Haven, CT, United States

Understanding the biochemistry of the cell requires measurement of all the molecules it produces. Single-cell proteomics recently became possible through advances in microanalytical sample preparation, separation by nano-flow liquid chromatography (nanoLC) and capillary electrophoresis (CE), and detection using electrospray ionization (ESI) high-resolution mass spectrometry (HRMS). Here, we demonstrate capillary microsampling CE-ESI-HRMS to be scalable to proteomics across broad cellular dimensions. This study established proof-of-principle using giant, ~250-µm-diameter cells from embryos of the frog Xenopus laevis and small, ~35-µm-diameter neurons in culture from the mouse hippocampus. From ~18 ng, or ~0.2% of the total cellular proteome, subcellular analysis of the ventral-animal midline (V11) and equatorial (V12) cells identified 1,133 different proteins in a 16-cell embryo. CE-HRMS achieved ~20-times higher sensitivity and doubled the speed of instrumental measurements compared to nanoLC, the closest neighboring single-cell technology of choice. Microanalysis was scalable to 722 proteins groups from ~5 ng of cellular protein digest from identified left dorsal-animal midline cell (D11), supporting sensitivity for smaller cells. Capillary microsampling enabled the isolation and transfer of individual neurons from the culture, identifying 37 proteins between three different cells. A total of 224 proteins were detected from 500 pg of neuronal protein digest, which estimates to a single neuron. Serial dilution returned 157 proteins from sample amounts estimating to about half a cell (250 pg protein) and 70 proteins from ca. a quarter of a neuron (125 pg protein), suggesting sufficient sensitivity for subcellular proteomics. CE-ESI-HRMS complements nanoLC proteomics with scalability, sensitivity, and speed across broad cellular dimensions.

Keywords: capillary electrophoresis, mass spectrometry, single cell, xenopus, mouse, proteomics

INTRODUCTION

Our understanding of cellular biochemistry depends on the ability to measure the constituent biomolecules in the cell. Genes and transcripts are now routinely analyzed with single-cell resolution thanks to automated sample handling using microfluidics, signal amplification using polymerase chain reaction (PCR), and multiplexing barcoding and sequencing using next-generation sequencing.
technologies. These measurements provided novel insights into regulatory mechanisms that control gene transcription as cells differentiate and establish lineages important for the formation of tissues and organs. Due to a complex correlation between transcription and translation, especially in developing biological systems, holistic understanding of cellular biochemistry calls for direct measurement of the proteome. Without a capability to amplify the whole proteome, the challenge has been due to limitations in the sensitivity of HRMS, the contemporary technology of choice for the unbiased and label-free identification and quantification of proteomes.

Single-cell HRMS was recently made possible for proteome studies. The technology was the focus of recent reviews (Rubakhin et al., 2011; Trouillon et al., 2013; Zenobi, 2013; Lombard-Banek et al., 2017; Levy et al., 2018; Zhang and Vertes, 2018; Kelly, 2020; Taylor et al., 2021). Advances in sample processing, separation, and detection enabled the measurement of proteins from minute amounts of materials that are available from cells. These studies usually tailored to specific types of cells in a given biological model. Using capillary microsampling, we demonstrated protein collection from identified cells in frog X. laevis and zebrafish embryos with dexterity, speed, and finesse. The method returned ~744 different proteins from ~1 pg (~15 amol) of sensitivity for model proteins and ~260 zmol (~150,000 copies) for model peptides. These custom-built CE-HRMS platforms identified ~750 proteins from ~5 ng of protein digest, or ~0.05% of the total protein content of a single cell in a 16-cell stage frog embryo. The method returned 428 proteins from ~500 pg of protein digest from cultured neurons, which estimated to single neurons. In theory, CE-HRMS raises a possibility for scalable proteomics across cells of broad dimensions.

In a systematic study here, we assess CE-HRMS for scalable single-cell proteomics from large to small cells. For testing proof-of-principle, we capillary microsample protein content of ~350 µm diameter identified embryonic cells (blastomeres) from 16-cell X. laevis embryos and single ~35 µm diameter neurons cultured from the mouse. To improve sensitivity and analytical speed, we compare protein identifications from diluted protein digests from cultured neurons between nanoLC, the gold standard technology in bottom-up proteomics, and microanalytical CE using HRMS. Encouraged by the identification of hundreds of proteins from the embryonic cells and diluted protein digests estimating to single neurons or less, we also evaluate the platform for proteomics in single neurons in culture. Results from this study demonstrate CE-ESI-HRMS to complement nanoLC proteomics with scalability across broad cellular dimensions and a capability for subcellular sensitivity.

**METHODS**

**Chemicals and Materials**

All solvents including acetic acid (AcOH), acetonitrile (ACN), formic acid (FA), methanol (MeOH), and water were purchased at LC-MS grade from Fisher Scientific. Reagents for bottom-up proteomics were purchased from Sigma-Aldrich (St. Louis, MO), unless specified otherwise. Ethylenediaminetetraacetic acid (EDTA) for neuron cell lysis was from Fisher Scientific (Fair Spring, Inc.).

**METHODS**

**Chemicals and Materials**

All solvents including acetic acid (AcOH), acetonitrile (ACN), formic acid (FA), methanol (MeOH), and water were purchased at LC-MS grade from Fisher Scientific. Reagents for bottom-up proteomics were purchased from Sigma-Aldrich (St. Louis, MO), unless specified otherwise. Ethylenediaminetetraacetic acid (EDTA) for neuron cell lysis was from Fisher Scientific (Fair Spring, Inc.).

**METHODS**

**Chemicals and Materials**

All solvents including acetic acid (AcOH), acetonitrile (ACN), formic acid (FA), methanol (MeOH), and water were purchased at LC-MS grade from Fisher Scientific. Reagents for bottom-up proteomics were purchased from Sigma-Aldrich (St. Louis, MO), unless specified otherwise. Ethylenediaminetetraacetic acid (EDTA) for neuron cell lysis was from Fisher Scientific (Fair Spring, Inc.).

**METHODS**

**Chemicals and Materials**

All solvents including acetic acid (AcOH), acetonitrile (ACN), formic acid (FA), methanol (MeOH), and water were purchased at LC-MS grade from Fisher Scientific. Reagents for bottom-up proteomics were purchased from Sigma-Aldrich (St. Louis, MO), unless specified otherwise. Ethylenediaminetetraacetic acid (EDTA) for neuron cell lysis was from Fisher Scientific (Fair Spring, Inc.).
Lawn, NJ). The CE background electrolyte (BGE) solution was prepared with 25% (v/v) ACN and 1 M FA. The CE-ESI sheath solution consisted of 10% (v/v) ACN containing 0.5% (v/v) AcOH. The capillaries for microsampling Xenopus single cells and performing nano-flow electrospray ionization (nanoESI) were constructed from borosilicate glass capillaries (0.50/1.00 and 0.75/1.00 mm inner/outer diameter, respectively) using a Sutter Instrument P-1000 micropipettte puller (Novato, CA). The extracted proteins were digested with MS-grade TPCK-modified trypsin (part no. 90057, Pierce, Rockford, IL) in 200/500 μL LoBind Eppendorf vials (part no. 9510100602/2431064, Eppendorf, USA), respectively, chosen to reduce adsorptive losses for proteins during sample processing. For CE, fused silica capillaries were purchased from Polymicro Technologies and used without modification (40/105 μm and 20/90 μm inner/outer diameter for Xenopus and mouse neurons, respectively, Phoenix, AZ). Measurements on the Xenopus digests used an electrokinetic pump to supply a low flow of sheath liquid, which was built using the pulled borosilicate capillary (Lombard-Banek et al., 2019). Analyses on the neuronal proteome digest employed a stainless steel tapered-tip metal emitter (100/320 μm inner/outer diameter, New Objective, Woburn, MA) following our recent design (Choi et al., 2018a).

Animal Care

Adult Xenopus laevis frogs (Nasco, Fort Atkinson, WI) and pregnant C57BL/6 dams of mice (Charles River Laboratories, Wilmington, MA) were cared for and humanely used following protocols approved by the Institutional Animal Care and Use Committee of the University of Maryland, College Park (Approval no. #R-FEB-21-07) and the George Washington University (Approval no. A274), respectively.

Protein Preparation From Xenopus laevis Cells

Xenopus embryos were collected during the natural mating of multiple parents. The experiments spanned different days and different parents to account for natural biological variability. The embryos were dejellied with 2% (w/v) cysteine solution and transferred to 100% (w/v) Steinberg’s solution (Lombard-Banek et al., 2019). The 2-cell stage embryos with stereotypical pigmentation were selected, cultured to the 16-cell stage, and the cell types of interest were recognized based on the location, pigmentation, and size. (Moody, 1987). The cell contents were aspirated by capillary microsampling following our recent study, (Lombard-Banek et al., 2019), then transferred to 50 mM ammonium bicarbonate (AmBic) prepared with HPLC grade water.

The extracted proteome was processed via standard bottom-up proteomic workflows. For the CE-MS measurements, the proteome was similarly processed, including reduction and alkylation, and desalted on a ZipTip column (Millipore Sigma) prior to analysis.

Protein Preparation From Cultured Mouse Neurons

Primary cultures of isolated mouse hippocampal neurons were prepared as described elsewhere (Kaech and Baner, 2006). After 14 days of culture, the individual neurons were randomly selected in the culture for analysis. Alternatively, to process the entire cell population, the neurons were rinsed 3 times with 1 mL of sodium dodecyl sulfate (SDS) (10% v/v), then transferred into a LoBind microtube. The collected neurons were pelleted at 800 × g at 4°C for 5 min. The neurons were lysed in SDS with sonication, followed by reduction (5 μL of 1 M dithiothreitol, DTT, 30 min at 60°C), alkylation (10 μL of 1 M iodoacetamide, IAD, 15 min in the dark), and quenching of the reaction (5 μL of DTT). The extracted proteins were purified by acetone precipitation overnight, then reconstituted in 200 μL of 50 mM AmBic, followed by trypsin digestion (1 mg/mL, 0.8 μL) over 12 h at 37°C. The peptides were dried and reconstituted for analysis.

NanoLC-HRMS

The peptides were dried at 60°C in a vacuum concentrator and reconstituted in LC-MS grade water containing 0.1% (v/v) FA. The peptide digest (400 ng) was trapped on a C18 trap column (100 μm i.d., 5 μm particle with 100 Å pores, 2 cm length, Acclaim PepMap 100, Thermo) at 5 μL/min for 5 min, then separated on a C18 analytical column (75 μm i.d., 3 μm particle with 100 Å pores, 25 cm length, Acclaim PepMap 100, Thermo) at 300 nL/min using a 120 min gradient using nanoflow liquid chromatography (Dionex Ultimate 3000 RSLC, Thermo). Elution employed the following gradient of Buffer B (100% ACN, 0.1% v/v FA) as follows (% v/v): 3.2% for 0–5 min, to 35% in 80 min, to 72% in 5 min, held at 72% for 5 min, then ramped to 5% in 5 min and equilibrated at 5% for 20 min. The eluent was electrosprayed through an electriified (+2 kV vs. Earth ground) stainless-steel emitter (part no. ES542, Thermo Fisher).

The generated ions were detected on an orbitrap-quadropole-ion trap tribrid ultrahigh-resolution mass spectrometer (Orbitrap Fusion Lumos, Thermo) executing data-dependent acquisition (DDA) to control tandem MS using the following parameters: MS scan resolution, 120,000 FWHM (m/z 200) in the orbitrap analyzer; cycle time, 3 s; max IT, 50 ms; AGC target, 4 × 10^6 counts; microscans, 1; dynamic exclusion, 60 s. The precursor ions were fragmented in the ion trap with the following settings: isolation window, 1.6 Da; higher-energy collisional dissociation (HCD) using the settings: normalized collision energy (NCE), 30%; AGC target, 1 × 10^4 counts; max IT, 100 ms; microscans, 1.

CE-ESI-MS

Xenopus. The peptide mixture was dried in a vacuum concentrator at 60°C and reconstituted in 2 μL of 75% (v/v) ACN containing 0.05% (v/v) FA (sample solvent). The peptides were separated in a custom-built CE-ESI platform, as described elsewhere. (Lombard-Banek et al., 2019; Lombard-
Banek et al., 2021). Approximately 18 or 5 ng of peptides were hydrodynamically injected into a ~90 cm CE capillary filled with the BGE. The sample was electrophoresed by electrospraying the inlet end of the CE capillary, generating ~5–10 µA current. An electrokinetic pump supplied low sheath-flow at the outlet of the CE capillary. The tip of the ESI emitter was mounted ~1 mm from the mass spectrometer orifice. Approximately 1,000 V ESI voltage was applied to maintain a stable Taylor cone, (Nemes et al., 2007), monitored by a long-working distance objective microscope (Mitutoyo Plan Apo, Edmund Optics, Barrington, NJ) equipped with a CCD camera (EO-2018C, Edmund Optics). For 18 ng of *Xenopus* proteome digest, +23 kV CE voltage was applied, yielding ~1 h separation time. For 5 ng of proteome digest, a +28 kV CE voltage reduced the analysis time to ~30 min.

The separated peptides were measured on a quadrupole orbitrap mass spectrometer (Q Exactive Plus, Thermo Scientific) executing tandem MS using DDA. Dynamic exclusion on the 10 most abundant precursor ions were selected for fragmentation. The following MS (Rubakhin et al., 2011) parameters were used: scan range, m/z 350–900; mass resolution, 70,000 FWHM (m/z 200); AGC target, 1 × 10⁶ ions; C-trap maximum injection time, 50 ms. The MS (Levy et al., 2018) parameters were set as follows: isolation window for precursor ions, 1.5 Da; mass resolution, 17,500 FWHM (m/z 200); AGC target, 5 × 10⁴ ions; C-trap maximum injection time, 100 ms; HCD with nitrogen gas, 28% NCE.

**Neurons**

The proteome digest from the individual neurons and neuron culture was dried in a vacuum concentrator at 60°C and respectively reconstituted in 0.5 and 5 µL of 50% (v/v) ACN containing 0.05% (v/v) AcOH (sample solvent). The CE background electrolyte (BGE) solution was 25% (v/v) ACN with 1 M FA. The CE-ESI sheath solution consisted of 0.1% FA and 50% (v/v) MeOH. The neuron samples were measured on a custom-built microanalytical CE-ESI platform previously developed in our group. (Choi et al., 2017). The peptides were separated in a 90 cm CE capillary and the CE voltage was set at 23 kV on the inlet of the capillary. The sheath solution was supplied at 300 nL/min and a stable Taylor cone was generated after applying +2,700 V ESI voltage. The tapered-tip emitter was set ~2 mm from the MS orifice. The separated peptides were measured on the Q Exactive Plus mass spectrometer as earlier using DDA on the 15 most abundant precursor ions (Top 15) for MS/MS. The MS (Rubakhin et al., 2011) parameters were as follow: scan range, m/z 350–1,800; mass resolution, 35,000 FWHM (m/z 200); isolation window, m/z 0.8; C-trap maximum injection time, 50 ms. For MS (Levy et al., 2018), the parameters were set as follows: mass resolution, 17,500 FWHM (m/z 200), AGC target, 1.5 × 10⁴ ions; maximum injection time, 50 ms; HCD, nitrogen gas at 32% NCE.

**Data Analysis**

The primary MS-MS/MS files were analyzed with Proteome Discoverer version 2.2 (Thermo) with SEQUEST search engine and Minora algorithm for label free quantification (LFQ) analysis or MaxQuant version 1.6.17.0 (Max Planck Institute of Biochemistry) executing the Andromeda search engine (Cox et al., 2014). Using identical search parameters between these bioinformatic platforms, the detected signals were identified against the *Xenopus laevis* proteome (downloaded from UniProt on 12 Sep 2021, containing 43,236 entries) or the mouse proteome (downloaded from SwissProt on 11 August 2021, containing 21,978 entries). The search parameters were as follows: minimum peptide length, 6 amino acids; maximal missed cleavage sites, 2; precursor mass tolerance, 10 ppm; fragment mass tolerance, 0.02 Da (Q Exactive Plus) and 0.6 Da (Lumos, Thermo). The static and dynamic modifications of cysteine carboxymethylation and methionine oxidation were enabled for neuroproteomics by CE-MS and *Xenopus* proteomics by LC-MS. Protein identifications were filtered to <1% false discovery rate (FDR) against a reversed-sequence decoy database, both at the peptide and protein level. Common contaminant proteins were manually curated against the cRAPome (https://www.thegpm.org/crap/) and excluded from the proteins that are reported in this study.

**Safety**

Standard safety protocols were followed when working with chemicals and biological samples. Care was taken when handling microsampling capillaries, electrospray emitters, and metal tapered-tip emitters, which pose needle-stick hazards. All electrically conductive parts of the CE-nanoESI platform were grounded or shielded to prevent electrical shock hazards.
Data Repository
All primary and processed MS-MS/MS data were deposited to the ProteomeExchange Consortium (http://proteomcentral.proteomexchange.org) via the PRIDE (Perez-Riverol et al., 2019) partner repository with the project accession number PXD031955.

RESULTS AND DISCUSSION
CE-HRMS for Scalable Single-Cell Proteomics
The goal of this study was to assess the scalability of CE-HRMS for single-cell proteomics across broad cellular dimensions. Figure 1 presents the microanalytical pillars of our approach. Capillary microsampling was selected to enable the collection of proteins from cells of varying dimensions. We analyzed ~30–300 µm (diameter) blastomeres from popular biological models in cell and developmental and neurobiology in this study. Early X. laevis embryos were used to obtain large, ~250-µm-diameter embryonic cells that are fated to give rise to neural or epidermal tissues in the frog. The proteomic measurements were extended to differentiated cells, specifically ~35 µm neurons in culture from the mouse. For the proteomic characterization of these cells, we benchmarked the performance of microanalytical CE using nanoLC as the closest neighboring reference technology and tested detection to cellular and subcellular amounts of proteins.

Scalability was essential for these proteomic measurements. Figure 1 overviews our approach, starting with the identification of the cells of interest. The cleavage-stage X. laevis embryos provided an ideal biological model, because asymmetric pigmentation and stereotypical cell fate maps enabled accurate cell phenotyping and a large size (~250 µm/cell in the 16-cell embryo) with predictable cell divisions facilitated direct manipulation of the cells (Grant et al., 2013; Pratt and Khakhalin, 2013). Figure 2A demonstrates microscopy-aided capillary microsampling of identified cells in the embryos. The dorsal (called D) cell in the 4-cell embryo formed the dorsal-animal (called D1) cell of the 8-cell embryo, which then divided into the dorsal-animal midline (called D11) in the 16-cell embryo. Capillary microsampling ensured adequate spatial and temporal finesse and scalability to access contents of the cells in the complex embryo, despite synchronous cell divisions occurring every ~15–30 min at this stage of development. An ~10 nL portion was aspirated from the D11 cells into a pulled borosilicate capillary within ~5 s using an in-line connected microinjector delivering negative pressure pulses. The aspirates were ejected into 5 µL of 50 mM of AmBic. This cell stores ~8 µg of total protein within an ~90 nL internal volume (Peuchen et al., 2016). Ca. 90% of the cellular proteome is from yolk, thus yielding an
estimated 1 µg of yolk-free protein amount from each cell (Sun et al., 2016).

Microsampling was potentially applicable to also handling single mammalian neurons. To test the approach, we cultured primary hippocampal neurons on a glass slide. Figure 2B exemplifies readily identifiable neurons and their processes under inspection by an inverted microscope. A borosilicate tubing was fabricated into micropipettes with ~30 µm tip diameter and mounted on a three-axis translation stage under remote control. As earlier, the outlet end of the micropipette was connected to a microinjector delivering pulses of calibrated vacuum. This micropipette was used to extract the whole neuron from the culture, including its soma and the microscopically identifiable portions of its axon and dendrites. Partial cell lysis was observed as the neuron was collected into the micropipette (see inset). We previously estimated that a single neuron yields ~500 pg of total protein content. (Choi et al., 2018a). To account for smaller protein amounts than from the X. laevis cells, the neuron was captured together with ~2 µL of Neurobasal Plus/B27 culture media to reduce protein losses, albeit added salts from the media were anticipated to challenge detection. The procedure was repeated on up to 3 different neurons in the culture.

The collected proteins were processed for bottom-up proteomics. The aspirates were ejected into a microtube containing AmBic and digested to peptides using trypsin following our established protocols (Lombard-Banek et al., 2019) (see Methods). The standard alkylation and reduction steps of bottom of proteomics were eliminated to reduce material losses for the neuron samples. The resulting peptides were separated over ~2 h by gradient C18 reversed-phase chromatography or ~30–60 min electrophoresis using CE. NanoLC analysis required at least ~4 µL of protein digest in the sample loading vial, whence ~2 µL of sample solution containing ~400 ng of protein digest were injected during each technical replicate analysis. In comparison, our custom-built CE platform enabled the loading of ~1–10 nL of the sample, or ~1–10 ng of protein digest, from each ~250 nL aliquot of the sample. These measured protein amounts mark at least an order of magnitude of improvement to help analyze lower protein amounts.

The resulting peptides were detected by ESI-HRMS. Figure 2C shows a representative base peak chromatogram and electropherogram from the same identified D11 X. laevis cell. Separation was completed in ~90 min by LC and ~45 min by CE; electrophoresis essentially halved the analysis time. The resulting peptide data allowed us to identify the source proteins against the X. laevis and mouse proteomes and estimate their relative abundances using established bioinformatic resources (<1% FDR), thus revealing the proteomic profile of the cells.

Benchmarking CE Proteomics Using NanoLC

The proteomic data allowed for the evaluation of CE sensitivity to NanoLC, the closest neighboring single-cell technology. The...
content of a whole single V12 cell was capillary microaspirated and analyzed by both technologies as described earlier. Figure 3A compares the protein identifications from the ~400 ng of protein digest by nanoLC and ~18 ng protein digest by CE. Over ~2 h of chromatography, a total of 797 proteins were identified on average, or 1,156 proteins in cumulative from technical triplicates. The proteins are tabulated in Table 1 in the electronic Supplementary Information (Supplementary Table S1).

We queried the reproducibility of the identifications. Using nanoLC, about 42% of the proteins were detected in all the replicates, and 65% were in at least two. Electrophoresis returned an average of 855, or 1,133 cumulative proteins from technical triplicates (see Supplementary Table S2). About 54% of the proteins were detected in all the replicates, and 73% were in at least two. From 5 ng of protein digest, the approach was scalable to an average of 569, or 722 cumulative proteins (See Supplementary Table S3). Therefore, identifications scaled better than anticipated based on protein amounts. We ascribed these improvements to field-amplified sample stacking allowing to concentrate the sample in the capillary for signal enhancement, higher theoretical plate numbers aiding peptide separation to tailor to the duty cycle of tandem MS, and reduced flow rates during nanoESI facilitating ionization efficiency. CE demonstrated favorable scalability toward trace amounts of protein digests.

The depth of protein identifications matter for biological experiments. Figure 3B compares the types of proteins that were identifiable a single shot measurement of the V12 blastomere using nanoLC (~400 ng) and CE (~18 ng). Despite analyzing a notably lower amount of proteome, ~46% of all the proteins were identified by both technologies. A comparable number of proteins were exclusively identifiable by CE (296 proteins) and nanoLC (292 proteins). These identifications are tabulated in Supplementary Table S4.

As not all identifiable proteins are quantifiable in bottom-up proteomics, Figure 3B also compared protein quantification. Of all the 900 proteins that were quantified by the technologies, only ~44% were commonly quantified. CE was able to quantify about twice as many unique proteins compared to nanoLC. Therefore, while identification was comparable, quantification was found enhanced during electrophoresis.

Orthogonality was compared among the quantifications. Using LFQ, an effective proxy for concentration in both technologies, Figure 3C plots the mean-normalized and log-transformed abundance of proteins that were quantified commonly between nanoLC and CE HRMS. With a calculated product moment ($\rho$) of 0.75, Pearson cross-correlation analysis found the two technologies reproducible at relatively quantifying the proteome. Figure 3D compares the normalized abundances of the proteins. Proteins that were quantifiable by both CE and nanoLC had indistinguishable abundance distributions (Wilcoxon signed test). Surprisingly, the 318 proteins that were quantifiable only by CE were statistically more abundant than the 186 proteins that were exclusively measured by nanoLC. Further, CE-MS achieved a lower number of missing values (~10% missing values) among the quantitative replicates compared to nanoLC (~27% missing values). Therefore, CE improved the sensitivity and robustness of quantification.

**Scaling to Neurons and Subcellular Protein Amounts**

We analyzed the single neurons that were aspirated from the cell culture (recall Figure 2B). CE-ESI-HRMS returned 16, 14, and 7 proteins from single neurons, totaling 37 different proteins among the biological replicates (Supplementary Table S5). These results were encouraging, although revealed notable variability in identification success and proteome coverage. Manual handling of the cells and their processing in microvials likely introduced variations in protein aspiration...
and enhanced adsorptive losses on vial and pipette surfaces. We anticipate that our approach can readily be integrated with recent advances enabling high-efficiency preparation and processing of microscale samples [e.g., NanoPOTS (Zhu et al., 2018), SCOPe2 (Specht et al., 2021)]. Capillary microsampling allowed us to extend CE-HRMS proteomics to single neurons in culture.

These results also suggested scalability to smaller cells and subcellular compartments. To establish proof of principle, we limited variations in sample processing and protein losses by measuring diluted protein digests from neuron cultures in this portion of the study. We previously estimated a single cell to yield ~500 pg of extractable protein amount. (Choi et al., 2018b). To enhance the detectable portion of the neuroproteome, we sought to improve both sensitivity and peptide sequencing. The protein digests were suspended in 75% (v/v) of ACN containing 0.05% (v/v) FA to aid sensitivity by enabling field-amplified sample stacking (Lombard-Banek et al., 2016a; Choi et al., 2017; Baxi et al., 2018; Choi et al., 2018b; Lombard-Banek et al., 2019; Lombard-Banek et al., 2021). We recently found concentration-dependent fragmentation redundancy to be a leading cause of sensitivity loss during CE HRMS employing DDA-driven tandem MS. (Choi et al., 2021). This experience allowed us to define an ion signal abundance-dependent threshold, or a rung of a DDA ladder from our previous work, to seek identifications. To maximize the utilization efficiency of the DDA duty cycle, the 250 peptide-like ions with the most abundant ion signals were excluded from sequencing. The approach was tested for total protein amounts approximating to a single cell (1 nL of 500 pg/nL analyzed), a half cell (1 nL of 250 pg/nL), and a quarter of a cell (1 nL of 125 pg/nL).

Protein identifications and quantification was assessed. Figure 4A tracks the linear dynamic range of the detected proteins using LFQ as a proxy for concentration. A total of 222 out of 224 identified proteins were quantified from ~500 pg of protein digest between technical duplicates, approximating to a single neuron. These proteins were found to span over a 3–4 log-order concentration range based on their LFQ abundances. Technical duplicate of 250 pg protein digest, or ca. half a neuron, returned 157 different proteins with 155 quantified. From approximately a quarter of a cell, or ~125 pg of protein digest, 69 out of 70 identified proteins were quantified. These protein identifications are listed in Supplementary Table S6. The rich information obtained appreciated the molecular state of neuron- and subneuronal equivalent protein amounts.

Smaller protein amounts returned fewer proteins, surprisingly with complementary coverage. Figure 4B annotates the types of proteins that were quantified between the dilutions. Seven proteins were only detected from a quarter portion of the neuronal proteome. Most occupied the lower range of the dynamic concentration span and are known for biological importance. For example, the protein Kcnq5 takes part in forming a potassium channel, which can affect electrical excitability, (Tzingounis et al., 2010), and the protein Gm10358 was linked to carbohydrate degradation during stress response. (Li et al., 2020). If abundant proteins masked the detection of these proteins in the previous measurements, concentration-dilution-caused enhancements in field-amplified sample stacking or DDA stochasticity may explain the observed signal enhancements.

Last, we sought canonical knowledge on the identified proteins. The 224 proteins (500 pg) were mapped to the human genome and their protein-protein associations analyzed. Figure 5 presents the association network based on

FIGURE 5 Reactome analysis of protein-protein interactions from a single-neuron equivalent protein digest (500 pg proteome). More than 700 pathways were represented (Supplementary Table S7). Detected nodes are highlighted in yellow. Pathway key: (A), muscle contraction; (B), digestion and absorption; (C), developmental biology; (D), neuronal system; (E), immune system; (F), signal transduction; (G), hemostasis; (H), gene expression; (I), metabolism of RNA; (J), metabolism; (K), chromatin organization; (L), DNA repair; (M), DNA replication; (N), transport of small molecules; (O), sensory perception; (P), cell cycle; (Q), reproduction; (R), organelle biogenesis and maintenance; (S), metabolism of proteins; (T), programmed cell death; (U), cellular responses to stimuli; (V), autophagy; (W), protein localization; (X), vesicle-mediated transport; (Y), extracellular matrix organization; (Z), cell-cell communication.
Reactome (Jassal et al., 2020). About 750 pathways were represented in the analysis. Multiple cellular functions were accounted for, from cell adhesion, to intracellular signaling, regulation of process outgrowth and synaptic function in neurons. A complete list of the pathways is summarized in Supplementary Table S7. Among the 10 most significantly enriched cellular pathways ($p \leq 10^{-12}$) were Rho GTPase signaling, JAK-STAT signaling, and axon guidance and cell adhesion. Proteins involved in synaptic vesicle release, endocytosis, and synaptic transmission were also enriched. Two processes of interest with a significant $P$ value were also aggrephagy and selective autophagy. Accumulation of misfolded proteins in neurons was associated with several neurodegenerative diseases, such as Parkinson, Alzheimer, and Huntington Disease, (Conway et al., 2020), and the ability to recognize specific targets could identify cellular differences. Detection of a host of important processes with single- and subcellular sensitivity open exciting new directions in neuroscience studies.

**CONCLUSION**

This study found CE to complement nanoLC ESI-HRMS proteomics with sensitivity, speed, and scalability to single cells spanning broad dimensions. From 20- to 100-times less protein digest amounts, CE-ESI was able to identify a similar number of (~1,100 different) proteins from single identified X. laevis cells. The platform was scalable to ~722 proteins from ~5 ng of embryo protein digest, supporting extension of the method to smaller cells. Capillary micromanipulating allowed us to individually aspirate 3 single neurons from culture, identifying 37 different proteins in total. CE-HRMS on neuronal proteome digests estimating to a single neuron (~500 pg proteome) returned 224 different proteins. The method was tested scalable to 157 proteins from approximately half a neuron (~250 pg proteome digest) and 70 proteins from about a quarter of a neuron (~125 pg proteome digest). The quantified proteins were enriched in multiple biological pathways important for cell and nervous system development and cell-to-cell communication. Results from this study demonstrated CE-HRMS-based single-cell proteomics to be scalable for cells measuring ~35–350 µm in diameter, a ~150-fold change in available protein amounts, opening new possibilities in cell biology and neuroscience.

These protein identifications mark only the beginning. Further sensitivity enhancements are possible through better sample processing and smarter data acquisition. Micro- and nano-scale platforms that are capable of minimizing protein or peptide losses during sample handling, such as µPOTS (Xu et al., 2019) and nanoPOTS, (Zhu et al., 2018), hold keys to deepen the detectable coverage of the single-cell proteome. Although this work utilized a custom-built ultrasensitive CE-ESI platform, commercialization of instruments capable of low-volume sample handling recently enabled ultrasensitive detection (Johnson et al., 2022) to support adaptation of our approach in other laboratories. CE-ESI-HRMS proteomics opens new possibilities to help characterize the molecules in the cell, now also including the proteome.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are publicly available. This data can be found here: http://www.proteomexchange.org/, PXD031955.

**ETHICS STATEMENT**

The animal study was reviewed and approved by the IACUC of the George Washington University and the University of Maryland, College Park.

**AUTHOR CONTRIBUTIONS**

PN, BS, and SC designed the study. PM-L and MCM provided the neuron culture. SC aspirated the single neurons from culture. LP (Xenopus) and SC (neurons) prepared the samples. BS (CE-MS, Xenopus), LP (LC-MS, Xenopus), and SC (CE-MS, neurons) measured the samples. BS and PN analyzed the data and interpreted the results. All authors commented on the manuscript.

**FUNDING**

Parts of this work were supported by the National Institute of General Medical Sciences Award No. R35GM124755 (to PN), Arnold and Mabel Beckman Foundation Young Investigator Award (to PN), and a University of Maryland Brain and Behavior Initiative (to PN) and Sampugna/Keeney Fellowship (to BS). The opinions and conclusions expressed in this publication are solely those of the authors and do not necessarily mirror the official views of the sponsors.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2022.863979/full#supplementary-material
Shen et al. Scalable Single-Cell Proteomics

Sweedler, J. V., and Arriaga, E. A. (2007). Single Cell Analysis. *Anal. Bioanal. Chem.* 387 (1), 1–2. doi:10.1007/s00216-006-1609-z

Taylor, M. J., Lukowksi, J. K., and Anderton, C. R. (2021). Spatially Resolved Mass Spectrometry at the Single Cell: Recent Innovations in Proteomics and Metabolomics. *J. Am. Soc. Mass Spectrom.* 32 (4), 872–894. doi:10.1021/jasms.0c00439

Trouillon, R., Passarelli, M. K., Wang, J., Kurczy, M. E., and Ewing, A. G. (2013). Chemical Analysis of Single Cells. *Anal. Chem.* 85 (2), 522–542. doi:10.1021/ac303290s

Tzigonis, A. V., Heidenreich, M., Kharkovets, T., Spitzmaul, G., Jensen, H. S., Nicoll, R. A., et al. (2010). The KCNQ5 Potassium Channel Mediates a Component of the Afterhyperpolarization Current in Mouse hippocampus. *Proc. Natl. Acad. Sci.* 107 (22), 10232–10237. doi:10.1073/pnas.1004644107

Williams, S. M., Liyu, A. V., Tsai, C.-F., Moore, R. J., Orton, D. J., Chrioler, W. B., et al. (2020). Automated Coupling of Nanodroplet Sample Preparation with Liquid Chromatography-Mass Spectrometry for High-Throughput Single-Cell Proteomics. *Anal. Chem.* 92 (15), 10588–10596. doi:10.1021/acs.analchem.0c01551

Wilson, S. R., Vehus, T., Berg, H. S., and Lundanes, E. (2015). Nano-LC in Proteomics: Recent Advances and Approaches. *Bioanalysis* 7 (14), 1799–1815. doi:10.4155/bio.15.92

Xu, K., Liang, Y., Piehowski, P. D., Dou, M., Schwarz, K. C., Zhao, R., et al. (2019). Benchtop-compatible Sample Processing Workflow for Proteome Profiling of. *Anal. Bioanal. Chem.* 411 (19), 4587–4596. doi:10.1007/s00216-018-1493-9

Zenobi, R. (2013). Single-cell Metabolomics: Analytical and Biological Perspectives. *Science* 342 (6163), 1243259. doi:10.1126/science.1243259

Zhang, Z., Dubiak, K. M., Huber, P. W., and Dovichi, N. J. (2020). Miniaturized Filter-Aided Sample Preparation (MICRO-FASP) Method for High Throughput, Ultrasensitive Proteomics Sample Preparation Reveals Proteome Asymmetry in Xenopus laevis Embryos. *Anal. Chem.* 92 (7), 5554–5560. doi:10.1021/acs.analchem.0c00470

Zhang, Z., Sun, L., Zhu, G., Cox, O. F., Huber, P. W., and Dovichi, N. J. (2016). 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. *Anal. Chem.* 88 (1), 877–882. doi:10.1021/acs.analchem.5b03496

Zhu, Y., Piehowski, P. D., Zhao, R., Chen, J., Shen, Y., Moore, R. J., et al. (2018). Nanodroplet Processing Platform for Deep and Quantitative Proteome Profiling of 10–100 Mammalian Cells. *Nat. Commun.* 9, 882. doi:10.1038/s41467-018-03367-w

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Shen, Pade, Choi, Muñoz-LLancao, Manzini and Nemes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*