A Data-Dependent Acquisition Ladder for Ultrasensitive (Neuro)Proteomics

Sam B. Choi^{1,2}, Pablo Muñoz-LLancao^{3,5,6}, M. Chiara Manzini^{3,5}, and Peter Nemes^{1,2,4*}

¹Department of Chemistry & Biochemistry, University of Maryland, College Park, MD 20742; Departments of ²Chemistry, ³Pharmacology & Physiology, and ⁴Anatomy & Cell Biology, School of Medicine and Health Sciences, The George Washington University, 2300 I (Eye) Street, NW, Washington, DC 20037; ⁵Department of Neuroscience & Cell Biology, Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ, 08901; ⁶Department of Cell Biology, Yale University School of Medicine, New Haven, CT 06510

*Correspondence to: Department of Chemistry & Biochemistry, University of Maryland, 8051 Regents Drive, College Park, MD 20742, USA. Phone: (1) 301-405-0373. Fax: (1) 301-314-9121. E-mail: nemes@umd.edu.

Running title (max 100 characters): DDA Ladder for Ultrasensitive Proteomics

Keywords (max 10 keywords): neuroscience, neuron, proteomics, mass spectrometry, data dependent acquisition, capillary electrophoresis, mouse

Author Contributions. P.N. and S.B.C. designed the research. P.M-L. and M.C.M. prepared the neuron culture. S.B.C. processed the neuron culture, prepared the sample, and performed the measurements. S.B.C. and P.N. analyzed the data and interpreted the results. P.N. and S.B.C. wrote the manuscript. All authors commented on the manuscript.

ABSTRACT

Measurement of broad types of proteins from a small number of cells to single cells would help to better understand the nervous system but requires significant leaps in high-resolution mass spectrometry (HRMS) sensitivity. Microanalytical capillary electrophoresis electrospray ionization (μ CE-ESI) offers a path to ultrasensitive proteomics by integrating scalability with sensitivity. We report here a data acquisition strategy that expands the detectable and quantifiable proteome in trace amounts of digests using μ CE-ESI-HRMS. Data-dependent acquisition (DDA) was programmed to progressively exclude high-intensity peptide signals during repeated measurements. These nested experiments formed rungs of our “DDA ladder.” The method was tested for replicates analyzing $\sim$ 500 pg of protein digest from cultured hippocampal (primary) neurons (mouse), which estimates to the total amount of protein from a single neuron. Analysis of net amounts approximating to $\sim$ 10 neurons identified 428 nonredundant proteins (415 quantified), an $\sim$ 35% increase over traditional DDA. The identified proteins were enriched in neuronal marker genes and molecular pathways of neurobiological importance. The DDA ladder deepens the detectable proteome from trace amounts of proteins, expanding the analytical toolbox of neuroscience.

INTRODUCTION

Ultrasensitive high-resolution mass spectrometry (HRMS) presents a window to the molecular composition of cells, raising a potential to better understand neurobiological processes at the level of neurons and neural circuits. Using high-performance nano-flow liquid chromatography (nanoLC) HRMS, it is now possible to quantify ~13,000 different proteins from neuron cultures.¹⁻² To deepen the detectable coverage of the proteome, shot gun proteomics currently processes sizable populations, usually millions, of neurons to analyze ~100 ng–1 µg of protein. Averaging over an entire cell population, however, potentially loses information critical to particular neurons or neural circuits; it also elevates chemical noise during measurements, further limiting insights into the biochemistry of the neuron.³ A single mammalian neuron yields only ~500 pg of protein extract,⁴ which is ~1000–100,000-fold less than usually detected in nanoLC HRMS. Identification and quantification of the single-neuron proteome hinges on the development of HRMS approaches capable of trace-level sensitivity.

Ultrasensitive HRMS bridges over this technological gap for the proteome (reviewed in references⁵⁻⁷). Instruments automating volume-limited liquid handling extended nanoLC tandem HRMS to improving sensitivity. Up to 4,002 proteins were identified from yeast using LC-HRMS.⁸ Recent nanoPOTS (N2) arrays analyzing 100 pg of tryptic protein digest identified ~1,300 proteins from single murine cells.⁹ The automated single cell proteomics (SCoPE2) workflow enhanced throughput to ~200 single cells in 24 h using a refined nanoLC system, while reporting on ~1,000 proteins per cell.¹⁰ These experiments deepened the detectable coverage of the single cell proteome, albeit at the expense of long experimental times (~2–5 h per analysis).

Capillary electrophoresis (CE) HRMS delivers speed and sensitivity for single-cell analyses. We and others developed robust methodologies and custom-built CE-ESI instruments capable of zeptomole–femtomole sensitivity (reviewed in references ¹¹⁻¹²). For example, ~200 protein groups were identified from 5 ng of *Pyrococcus furiosus* digest¹³ and up to 1,209 proteins groups from single embryonic cells dissected from¹⁴⁻¹⁵ or analyzed *in situ/vivo*¹⁶⁻¹⁷ in chordate embryos of important biological models (*Xenopus laevis*, zebrafish). A self-built CE platform and a sheath-flow tapered-tip CE-ESI design equipped HRMS with ~260 zmol sensitivity (156,000 copies) and robust operation on a quadrupole time-of-flight (TOF) mass spectrometer, allowing us to identify ~217 protein groups from ~1 ng protein (approximating ca. two neurons).¹⁸ An electrokinetically pumped low-flow sheath-flow CE-ESI interface achieved 330 zmol sensitivity, identifying 1,249 protein groups from 300 ng protein digest from *Xenopus laevis* eggs,¹⁹ ~4,400 protein groups from 220 ng of K562-cell digest via 2 h of separation,²⁰ and ~100 proteins from 16 pg of *E. coli* protein digest.²¹ Proteins in these experiments were identified in a short amount of time, usually within an hour.

We recently showed that compressed peptide separation during electrophoresis challenges proteomics sensitivity.^{4, 18} Peptides in these experiments separated over a ~20–30 min effective separation window in total analysis times that were ~2–10-times faster than those typical in nanoLC. The ion flux in CE was chemically complex yet transient, lasting a few seconds only. These conditions challenged the duty cycle of peptide isolation, activation, and fragmentation due to limits in the duty cycle of modern tandem HRMS. In our recent work analyzing protein amounts estimating to ~2 mammalian neurons, only ~43% of peptide-like ion signals with unique m/z and separation time (called molecular features, MFs) were identifiable as peptide spectral matches (PSMs) from ~217 proteins.⁴ Supplementing electrophoresis with an orthogonal

mode of separation, such as high-pH fractionation, allowed us to reduce the complexity of the chemical space in the temporal dimension, improving HRMS sensitivity to ~225 proteins from amounts approximating to single neurons.⁴ Even with orthogonal separation, however, the resulting MS/MS transitions exhausted the bandwidth of data-dependent acquisition (DDA) targeting the most abundant peptides for fragmentation. In mass spectrometers enhancing resolution in trade of speed, such as the orbitrap vs. TOF, we anticipate compact separation to exacerbate identifications in CE.

Iterative DDA presents a potential a way around the challenge. In nanoLC HRMS, the same sample is typically analyzed multiple times using the same DDA method (technical replicates) to leverage stochastic selection of complementary sets of peptides for identification.²²⁻²⁴ Triplicates detected ~20% more proteins, with this gain quickly diminishing after the 5th replicate, reaching ~95% saturation by the 10th.²² Other DDA methods employed advanced iterative strategies to boost identifications by exclusion or inclusion of ions of interest (mass-to-charge, m/z values).²⁵⁻²⁸ Precursor ion exclusion (PIE), for example, progressively excluded peptides by dynamically updating the ion exclusion list for each replicate analysis. The approach helped nanoLC-HRMS to 533 proteins from 5 replicates of 1 µg of yeast whole cell lysate digest, equating to an ~51% improvement compared to the standard approach.²⁵ Modified versions of PIE enhanced peptide identification by 52% with 4 measurements on 16 µg of IgG1 monoclonal antibody digest.²⁹ Alternatively, post analysis data acquisition (PAnDA) iteratively prioritized unselected peptide features for targeted analyses with each replicate.³⁰ The inclusion ion list measured 1,059 protein groups from 6 replicates of 4 µg of protein digest (*C. elegans*), resulting in an ~18% increase in identifications compared to standard DDA. By anticipating peptides in real-time based on a known elution order, 80% more proteins were identifiable from 4 measurements in mouse.³¹

These DDA approaches from studies employing nanoLC raise promising potentials to deepen proteomic sensitivity using CE-HRMS.

Here, we build on iterative DDA to advance ultrasensitive CE-HRMS to neuroproteomics on an orbitrap tandem mass spectrometer. DDA methods executing a nested series of ion exclusion lists formed rungs of our “DDA ladder,” allowing us to probe deeper into the proteome of cultured primary neurons. The method combined high spectral resolution from an orbitrap mass spectrometer (~120,000) with simplicity. Unlike earlier variants of DDA, we established a static list of ions for exclusion based on abundance tiers. In a proof of principle, the method achieved detection of 428 proteins (415 quantified), including many important in neurobiology, from sample amounts estimating to ~10 neurons. The DDA ladder expands the family of tandem MS approaches and emerging technologies, such as μ CE-ESI, supporting ultrasensitive neuroscience research.

METHODS

Materials. Unless otherwise stated, all materials were purchased at reagent grade from Sigma-Aldrich (St. Louis, MO). Standard lysozyme, myoglobin, and papain dissociation system were obtained from Worthington Biochemical Corp (Lakewood, NJ). Sodium dodecyl sulfate (SDS, 10% v/v) was from Amresco (Solon, OH). Ethylenediamine tetraacetic acid (EDTA), LC-MS grade acetonitrile (ACN), formic acid (FA), acetic acid (AcOH), methanol (MeOH), MS-grade trypsin protease, and water (Optima) were supplied by Fisher Scientific (Fair Lawn, NJ). Ammonium bicarbonate was from Avantor (Center Valley, PA). Hank’s balanced salt solution, fetal bovine serum, penicillin-streptomycin, minimum essential medium, neurobasal medium, glutamine, and poly-L-ornithine were obtained from Gibco (Grand Island, NY). B27, N-2,

glutamax, and pyruvate supplements were from Thermo Fisher Scientific (Waltham, MA). Fused silica capillaries (20/90 μm inner/outer diameter and 75/350 μm inner/outer diameter) were from Polymicro Technologies (Phoenix, AZ) and used as received. The stainless steel tapered-tip metal emitter (100/320 μm inner/outer diameter) was manufactured by New Objective (Woburn, MA). To minimize peptide losses due to nonspecific adsorption, the shot-gun proteomics workflow was conducted in LoBind protein microtubes from Eppendorf (Hauppage, NY).

Buffers and Standard Solutions. The *cell lysis buffer* was prepared to contain: 5 mM EDTA, 20 mM Tris-HCl, 1% (v/v) SDS, and 35 mM NaCl. The *neuronal plating medium* was prepared as described elsewhere³² to contain: 0.6% (v/v) D-glucose, 10% (v/v) horse serum, 1% glutamine, and 1% penicillin-streptomycin in $1\times$ MEM. The *neuro-maintain medium* was prepared to contain: 1% penicillin-streptomycin, 1% glutamax, 1% N-2, 2% B27, and 1% pyruvate in neurobasal medium. Prior to usage, all media were filtered through a 0.2 μm porous mesh.

Neuron Culture. All procedures regarding the maintenance and humane treatment of mice were authorized by the Institutional Animal Care and Use Committee of the George Washington University (Approval Number A274). Adult pregnant C57BL/6 dams of mice (*Mus musculus*) were purchased from Charles River Laboratories (Wilmington, MA). Primary cultures of isolated mouse hippocampal neurons were prepared as described elsewhere.³² After 14 days of culture, each plated well was washed 3 times with 1 mL of SDS (10% v/v), and the cultured neurons were gently scraped to be transferred to a 1 mL LoBind microtube using micropipette with a LoBind tip. The collected neurons were centrifuged at $800\times g$ for 5 min at 4 °C and stored at -80 °C until processing for measurement.

Bottom-up Proteomics. The collected cultured neurons were processed using a standard bottom-up proteomic workflow (see reference ²⁴). 200 μ L of lysis buffer was added to the collected neurons with 15 min of sonication in cold water bath ($\sim 4^\circ\text{C}$). The lysate was reduced (5 μ L of 1 M dithiothreitol, DTT, 30 min at 60°C), alkylated (10 μ L of 1 M iodoacetamide, IAD, 15 min in the dark) and quenched (5 μ L of DTT). The lysate was centrifuged at $11,200 \times g$ for 10 min at 4°C , and the supernatant was transferred into a new 2 mL LoBind microtube. Proteins in the aliquot were purified by precipitation in 1 mL of chilled acetone (-20°C , $5\times$ volume of sample aliquot) over 12 h, followed by centrifugation at $11,200 \times g$ for 10 min at 4°C . The protein pellet was rinsed with chilled acetone, vacuum-dried at room temperature, and reconstituted in 200 μ L of ammonium bicarbonate (50 mM). The final concentration of protein content was 0.5 $\mu\text{g}/\mu\text{L}$ based on the standard bicinchoninic acid assay (Thermo Scientific, Waltham, MA). Aliquots of 40 μ L were transferred into 5 separate 500 μ L LoBind microtubes to serve as technical replicates. Each aliquot was digested with 0.8 μ L of trypsin (1 mg/mL) over 12 h at 37°C . The resulting peptides were vacuum-dried and reconstituted in 40 μ L of 50% (v/v) ACN containing 0.05% (v/v) AcOH, chosen to help field-amplified sample stacking during electrophoresis.¹⁸ The peptide concentration of each aliquot was quantified using the Pierce colorimetric peptide assay (Thermo).

Microanalytical CE-ESI-HRMS. This study utilized a home-built co-axial sheath-flow μ CE-nanoESI platform¹⁸ with the following settings: CE, 90 cm capillary at 23 kV vs. Earth ground (applied to the inlet); CE-ESI sheath solution, 50% MeOH (0.1% FA) supplied at 300 nL/min; ESI, +2,700 V at 2 mm from the MS inlet, operated in the cone-jet spraying regime for efficient ion generation.³³ Peptide ions were detected using a quadrupole orbitrap mass spectrometer (Q Exactive Plus, Thermo Scientific, Waltham, MA) between m/z 350–1,800 at 35,000 FWHM

resolution (128 ms transient, 110 ms fill time). Higher-energy collisional dissociation (HCD) was triggered by DDA with the following instrumental settings: peak width (FWHM), 13 s; mass exclusion mass tolerance, 10 ppm; isolation window, 0.8 m/z; peptide matches, on; apex trigger, turned off; ion signals excluded below +2 charge state; ion signal intensity threshold, 1.5×10^3 counts; maximum ion trap time, 50 ms for MS¹. HCD employed the following conditions: m/z range, 200–2,000; resolution, 17,500 FWHM (64 ms transient, 50 ms fill time) for MS²; normalized collision energy, 32; dynamic exclusion, 5.0 s; maximum IT, 50 ms for MS²; TopN, 15; loop count, 15.

Data Analysis. Using *MaxQuant*³⁴ (version 1.6.17.0), the MS–MS/MS data were searched against the mouse proteome database containing 16,915 entries (downloaded from SwissProt on September 13th, 2017). The search parameters were: digestion, tryptic; missed cleavages, maximum 2; minimum peptide length, 5; minimum number of unique peptides, 1; fixed modification, carbamidomethylation at cysteine; variable modification, oxidation at methionine; match tolerance for main search peptide tolerance and MS/MS, ± 4.5 ppm and ± 20 ppm, respectively; isotope match tolerance, 2 ppm; decoy mode, revert; label-free quantitation, enabled; match between runs, turned off. Proteins were identified and filtered to <1% false discovery rate (FDR), computed against a reversed-sequence decoy database. The resulting proteins were grouped based on the closest parsimony principle. *Common contaminant proteins* were manually removed from the list of proteins that are reported in this work. MFs were surveyed in MzMine 2.0³⁵ with the following parameters; mass detector, exact mass; noise level, 5,000; chromatogram builder, on; scans, MS¹; min time span (seconds), 12 s; minimum height, 5,000; m/z tolerance, 0.05 m/z or 20 ppm. Gene ontology annotation of biological and molecular

function were conducted in the PANTHER Classification System.³⁶ Brain anatomy was modeled using Brain Explorer 2.³⁷

Safety Considerations. Metal tapered-tip emitters and fused silica capillaries pose needle-stick hazard and should be handled with care. Handling of chemicals followed standard safety protocols. Electrically conductive components of the μ CE-ESI-HRMS platform were grounded (Earth) or isolated to eliminate electrical shock hazard.

RESULTS AND DISCUSSION

Addressing A Technological Gap. The goal of this study was to enhance CE-ESI-HRMS sensitivity for ultrasensitive neuroproteomics. CE-HRMS enabled the analysis of hundreds to thousands of different molecules in single cells and limited protein digests in exceptional sensitivity, in cases ~200–1,000-times higher than nanoLC.^{4, 14-18} Biomolecules in these experiments migrated in a considerably short time frame, typically with an effective separation window lasting ~15–30 min. For complex ‘omes, particularly an expanded peptidome in bottom-up proteomics, compressed separation taxes identification due to limitations in HRMS-MS/MS sensitivity and speed.⁴ To enhance detection, electrophoresis employing long separation times³⁸ or orthogonal dimensions, such as ion exchange³⁹⁻⁴⁰, size-exclusion⁴¹, or high-pH C18 chromatography¹⁸ reduce molecular complexity over time by tailoring the transient ion flux to the duty cycle of MS/MS. Technological implementation, however, becomes increasingly challenging for trace amounts. Peak broadening due to longer separations is anticipated to reduce sensitivity, as modeled by the Van Deemter equation. Analyte transfer between separation dimensions risks peptide losses. These technologies also require skilled expertise and access to advanced, often custom-built, instrumentation, further constraining ultrasensitive CE-HRMS to specialized laboratories.

To advance CE sensitivity, we proposed to improve the success of data acquisition in HRMS–MS/MS. Our experimental rationale is shown in **Figure 1**. We used a custom-built (one-dimensional) μ CE-ESI system¹⁸ to analyze ~1 nL of protein digest from 250 nL of sample in each measurement.^{4, 18} These samples contained protein amounts approximating to ~2 neurons. Other custom-made or commercial CE systems with compatibility to handling limited sample

volumes serve as alternative platforms in other laboratories. We reasoned that a microanalysis by CE allows for deepening proteomic coverage via repeated measurements if each replicate reports progressively deeper into the proteome, all the while still consuming trace amounts of total proteins. Our goal in this work was to analyze total protein amounts that are commensurate with limited populations of cells, such as less than 10 neurons.

Method development began with control experiments optimizing the mass spectrometer for sensitivity. A protein digest was prepared from a culture of primary hippocampal neurons (mouse). The digest was diluted so ~1 ng of material was analyzed per replicate for method development, estimating to ~2 neurons.⁴ The peptides were electrophoresed, ionized, and detected in a quadrupole orbitrap tandem mass spectrometer (Q Exactive Plus, Thermo, see **Methods**). We chose DDA to govern MS/MS, because this method is robust, well-established, and available on most mass spectrometers. We anticipate these considerations to facilitate adoption of the method in other laboratories. A single-stage MS scan surveyed signals (m/z values), followed by fragmentation of the most abundant ions with a peptide-like isotope pattern. DDA parameters were optimized for identification based on technical duplicates. A dynamic exclusion of 5 s yielded 380 proteins (vs. 334 proteins at 2 s, 286 at 10 s, and 330 at 15 s), an MS¹ automatic gain control (AGC) of 10^6 counts returned 432 proteins (vs. 364 proteins at 5×10^5 counts), and selection of the top 15 most abundant ions for MS/MS (TopN) gave 518 proteins (vs. 359 proteins at Top12, 474 at Top20). A list of proteins identified from the triplicates consuming a total of 3 ng of digest is tabulated in **Supplementary Table 1A (Table S1A)**.

These measurements were tested for scalability to lesser protein amounts. Analysis of 500 pg of digest reported 190 proteins (~1 cell). The number of identifications increased with each technical replicate. Technical duplicates gave 213 proteins (~2 cells) and triplicates returned 244 proteins cumulatively (~3 neurons, see proteins in **Table S1B**). Detection of proteins with important roles in molecular pathways agreed with our earlier studies, demonstrating robust performance from CE-ESI-HRMS.^{4, 18} **Table S1C** uses the PANTHER Classification System to compare these pathway functions between 1 ng and 500 pg of measurements. Reduction of sample amounts resulted in missed detection for 339 proteins, or ~58% of all proteins identified. These proteins were implicated in several pathways important for health, including signaling. Neurobiological research would benefit from enhanced CE-HRMS sensitivity capable of deeper coverage of molecular pathways in tissue biopsies and small populations of cells.

To tackle this challenge, we began with understanding the current sources of limitations hindering DDA. In shot-gun proteomics, protein identification requires fragmentation of MFs that can be assigned to a peptide sequence with low error (<1% FDR), producing peptide spectral matches (PSMs). In our optimized method (see earlier), each DDA cycle targeted maximum 15 MFs. **Figure 2** compares the temporal evolution of MFs and PSMs to reflect on the utilization efficiency of the MS/MS duty cycle. During the ~60 min experiment, 935 peptides were detected (<1% FDR) migrating through the capillary within ~15 min, between ~30 min to ~45 min (**Fig. 2A**). This short effective separation window witnessed a rapidly increasing number of MFs. Their success of assignment as a PSM substantially varied. At the most compact zone of separation, only ~20–30% of MFs were identified. The success of identification increased to ~35–50% as molecular complexity relaxed toward the end of electrophoresis. These results from a quadrupole orbitrap mass spectrometer agreed with limited identification success (~45%) that

we reported from a quadrupole time-of-flight instrument,⁴ highlighting an analyzer-independent need for better data acquisition during CE-HRMS. **Figure 2B** evaluates MS/MS transitions underpinning the PSMs on the orbitrap. In response to increasing MFs, DDA boosted the rate of MS/MS transitions, albeit only up to the minimal duty cycle. Notably, ~95.5% of the MS/MS transitions exhausted the maximal C-trap fill time. Therefore, DDA curtailed MS/MS rate to trade identification for sensitivity.

But conducting many MS/MS scans does not necessarily yield complementary PSMs. In a representative dataset, 2,196 different PSMs were acquired from 1,203 different peptides; thus, the same peptide was fragmented ~1.8-times on average (MS² redundancy). For example, the peptide HFFTVDPR (from the protein Sideroflexin-3) was fragmented once, whereas the ~100-times more concentrated NLDIERPTYTNLNR (Tubulin alpha-1B chain) was fragmented 8 times. **Figure 2C** systematically compares MS² redundancy with DDA efficiency as a function of peptide abundance using label-free quantification (LFQ).¹⁵ More abundant peptides were more likely to yield PSMs. This redundancy ca. tripled with each decade of ion signal abundance across the ~3 log-order observed dynamic range. In parallel, the efficiency of cycle utilization during DDA, approximated here as the PSMs/peptide ratio, took a steep decline. The datasets were strongly anticorrelated ($\rho = -0.899$, Pearson). Between the extremes, an ~9-fold increase in MS² redundancy translated to a 15-fold reduction in cycle utilization (see **Fig. 2C**). Therefore, more abundant peptides were not only increasingly more likely to be redundantly fragmented, but they also required progressively more time from DDA, exacerbating a decline in identification success.

A Guided Design for the DDA Ladder. These limitations, we reasoned, could be (partially) addressed by reducing PSM redundancy. To develop this method, we sought to better understand

DDA economy. **Figure 3A** traces the cumulative for ion signal abundance and PSMs as a function of concentration (calculated LFQ intensities). Ca. 65%, ~75%, and ~85% of total signal abundance was produced by the 150, 250, and 350 most abundant (top) peptides. They respectively accounted for ~25%, ~37%, and ~50% of all the PSMs performed. This information supported our working hypothesis: DDA redundancy hindered protein identifications.

After validating this hypothesis, we developed an improved version of DDA in this study. Based on our results, we proposed that exclusion of these redundant features during replicate analyses would free up DDA bandwidth to lower-abundance peptides in subsequent measurements. Rather than executing the conventional strategy, in which the most abundant peptide ions are excluded at the risk of missing corresponding protein identifications, our approach was to nest DDA methods based on signal abundance. In theory, this approach afforded detection for the abundant proteins based the preceding set of measurements (see triplicate later), before filtering them out in subsequent analyses. It followed that protein identifications from each abundance level of exclusion would mutually complement each other, with each exclusion level deepening quantitative depth. As proof of principle, we supplemented the control DDA (without ion exclusion) with three nested 3 DDA methods that progressively excluded ions based on ion signal abundance (recall **Fig. 1**). These DDA methods offered vantage points at different depths into the proteome, akin to rungs of a ladder; therefore, we termed the method a “DDA ladder.”

The DDA ladder was tested for performance. Protein amounts of 500 pg, which approximates to a single neuron (cell),⁴ were measured in triplicate as the control (without exclusion of ions). The resulting proteins were ranked by LFQ abundance (recall **Table S1B**). The triplicates were repeated while excluding the most abundant 150, 250, and 350 (“top”) ions, identifying 308 (**Table S1D**), 358 (**Table S1E**), and 339 (**Table S1F**) proteins, respectively. We inspected

utilization of the DDA cycles by calculating the time that was spent between the survey scan and the last tandem MS event. **Figure 3B** exemplifies the analysis for conventional DDA and expands the analysis the experimental conditions (see **inset**). Ca. 42% of the cycles were completed without fragmentation (see 0 MS²); therefore, these scans did not identify any protein. Of the cycles performing tandem MS, ~10–30% targeted 1–5 peptide ions (see MS²/Cycle in the **inset**). Only less than ~2% of the cycles were filled with 15 MS² scans. Disregarding the abundant ions provided improvement. By excluding the top 250 ions, the fraction of DDA cycles completing 1 tandem MS was tripled and those conducting the maximal 15 MS/MS scans was improved by 6-fold. Cycles completing transitions between these boundaries became less likely (see 2–14 MS² scans). Therefore, exclusion of the most abundant ions enhanced utilization of the limited MS/MS duty cycle.

Better MS/MS utilization raised the possibility of identifying complementary proteins. **Figure 3C** compares identified proteins from excluding the top 150, 250, and 350 peptide ions. While ~48% of proteins were commonly identifiable, each of the tested rungs reported on a different set of proteins. Each tested exclusion level provided ~4.5–13% proteins that were only identifiable at each specific DDA rung. **Table S1G** shows pathway enrichment analysis for these proteins (PANTHER). Cytoskeletal regulation, Huntington disease, gonadotropin-releasing hormone receptor pathways, Parkinson disease, glycolysis, and EGF and FGF signaling, Wnt signaling, and the TCA cycle were top pathways represented by proteins that were commonly identified at each rung of the ladder. These pathways were complemented with additional proteins from exclusion of the top 150 ions. The second rung of exclusion expanded pathway coverage with angiogenesis, endothelin signaling, and PDGF signaling. Exclusion of the top 350 ion signals

provided additional proteins in the pathways covered by the lower rungs and may be beneficial if additional protein identifications are desired in a particular study.

The DDA ladder was benchmarked (**Fig. 4**). As the reference, standard DDA (without ion exclusion) was chosen, which is the closest neighboring technology to the DDA ladder presented here. **Figure 4A** predicts the cumulative number of proteins on the basis of up to quintuplet technical replicates (Control). As also found by others,²² rapid saturation in identification provided diminishing returns from the replicates. While triplicates identified 244 proteins in our study, the projections were 287 proteins at 6, 319 at 9, and 338 at 12 replicates. Each additional replicate after triplicate measurement returned only ~5% more proteins. Triplicates were a satisfactory trade between identification number, sample consumption, and analysis for our study. Each additional rung of the DDA ladder identified ~21%, ~34%, and ~42% more proteins compared to the control, improving cumulative identifications by ~42%, 24%, and 12%, respectively. With a diminishing return in cumulative identification, we chose the first two rungs (top 150 and 250 ions excluded) sufficient to supplement the control experiments for this study. Consuming ~500 pg proteins per analysis, the finalized DDA ladder for this study therefore required the analysis of a total of ~4.5 ng of protein between 9 measurements, estimating to proteins from ~10 neurons.

We inquired about the resulting depth of proteome coverage. The DDA ladder identified 428 proteins, 415 of which were also quantified (see **Table S1G**). To compare the abundance of the quantified proteins, we minimized technical variability in LFQ between replicates. The LFQ values were scaled using linear regression on the basis of the commonly quantified proteins (control vs. top 150 excluded: $R^2 = 0.93$; control vs. top 250 excluded: $R^2 = 0.90$). **Figure 4B** evaluates the quantitative dynamic range of the proteins. The most abundant proteins were

quantified without ion exclusion (see control with $\mu = 7.1$, Gaussian mean). Exclusion of the top 150 and 250 ions expanded the linear dynamic range of quantification. Therefore, each additional rung of the DDA ladder helped us quantify proteins deeper in the concentration range occupied by the neuroproteome, supporting our working hypothesis.

Last, we asked whether deeper quantification also translated into better coverage of molecular pathways. We compared pathways that were enriched by proteins detected by standard DDA consuming ~ 3 ng (see proteins in **Table S1A**) and ~ 1.5 ng (**Table S1B**) as well as the DDA ladder using ~ 4.5 ng (**Table S1H**) of protein digest. The analysis results are shown in **Figure 5**. A list of the respective pathways is tabulated in **Tables S1I (Eye)**. The DDA ladder reported on all the pathways that were represented by analysis of 1.5 ng of digest using conventional DDA. Our method also recovered proteins from ~ 9 of 21 different pathways that were only detected by 3 ng of analysis previously. These pathways included angiogenesis, Alzheimer disease-amyloid secretase pathway, and several signaling pathways including alpha adrenergic receptor, toll receptor, PDGF, interleukin, and endothelin signaling. Further, the DDA ladder also uncovered previously unrepresented pathways, including transcription regulation by bZIP transcription factor, vitamin D metabolism and pathway, Hedgehog signaling pathway, sulfate assimilation, and cholesterol biosynthesis. Overall, our findings show that protein detection in a single cell-equivalent volume can be extended to identify proteins beyond the most abundant cytoskeletal and synaptic components to further probe signaling and metabolism in healthy and diseased cells.

CONCLUSIONS

This study advanced CE-HRMS to trace-sensitive neuroproteomics. A DDA ladder encompassing two tiers of filters reduced a measurement bias toward high-abundance peptide ions, thus expanding the quantifiable linear dynamic range of the proteome. The method improved utilization of the limited DDA duty cycle, reduced redundant MS/MS events, and promoted the selection of a larger number of ions for targeted for fragmentation. Compared to standard DDA, the approach was better prepared to handle high peptidome complexity transiently unfolding during compressed electrophoresis. CE-HRMS employing the DDA ladder complements classical nanoLC-based proteomics with fast analysis (~30–45-min separation), ultrahigh sensitivity, and robust operation. Recent commercialization of CE instruments capable of handling limited amounts of samples, such as less than ~1–5 μ L, may substitute the custom-built microanalytical CE platform in this study. With well-established usage and availability on commercial mass spectrometers, we anticipate the DDA ladder to be readily adaptable in other laboratories to other types or limited amounts of proteins via shot-gun proteomics.

Continuous advances at multiple stages of the proteomic workflow raise the possibility of further sensitivity improvements. Technologies enabling the isolation and handling of miniscule amounts of proteins with reduced loss, for example, by *in vivo* microsampling¹⁶, nanoPOTS⁹, and ScoPE¹⁰ offer viable solutions for biopsies and limited populations of cells, including single cells. Detection of 428 proteins from protein amounts estimating ~10 neurons in this study, including many with important biological functions in homeostasis and disease, marks another

technological leap in ultrasensitive proteomics, expanding the bioanalytical toolbox of neuroscience.

ACKNOWLEDGEMENT. This work was supported by the Arnold and Mabel Beckman Foundation Beckman Young Investigator Award (to P.N.) and a COSMOS Club Award (to S.B.C).

FIGURES

Figure 1. Experimental strategy of the nested DDA ladder for ultrasensitive CE-ESI-HRMS. The 150, 250, and 350 most abundant (top) peptides were excluded during replicate measurement of protein digests, with each replicate approximating to a single neuron (~500 pg/cell). Scale bar, 50 μm .

Figure 2. Challenges in trace-sensitive neuroproteomics using CE-HRMS with standard DDA. **(A)** Comparison of molecular features (MFs) and peptide spectral matches (PSMs), uncovering rapidly diminishing success at identifying putative peptide signals during the most compressed portion of electrophoresis. **(B)** Evaluation of tandem MS events underlying peptide identifications, showing optimal success at minimal MS/MS duty cycle. **(C)** More abundant peptides were fragmented at increasing redundancy (black curve), causing a steep decline in the

efficiency of the DDA cycles identifying PSMs as peptides (10-point rolling average shown, white curve). Key: ρ , Coefficient of Pearson product-moment correlation.

Figure 3. Assessment of MS–MS/MS economy to guide DDA ladder design. **(A)** The 150, 250, and 350 most abundant peptides accounted for the majority of signal abundance (LFQ intensity) and peptide spectral matches (PSMs) in control experiments. **(B)** Evaluation of the DDA cycle

time, revealing underfilling of the top-15 peptide ions targeted for tandem MS. **(Inset)** Most DDA cycles were completed after fragmenting only 1–5 peptide ions. Exclusion up to the 250 most abundant ions (–250 Top) improved full utilization of the DDA cycle. **(C)** Each additional rung of the DDA ladder identified a complementary set of proteins in addition to the control (without ion exclusion).

Figure 4. Depth of proteome coverage. **(A)** Cumulative protein identifications from technical replicates by standard DDA (Control) and a nested DDA ladder excluding the top 150, 250, and 350 ions. The first two rungs of the ladder were chosen sufficient for this study. **(B)** Dynamic range of quantification by the DDA ladder, revealing quantification of lower-abundance proteins with each tested rung. Proteins only quantifiable in the control (grey) and after exclusion of the 150 (orange) and 250 (green) most abundant peptides are overlaid on the cumulative (black).

Figure 5. Gene anthology annotation of pathway coverage between the conventional DDA (control) and the DDA ladder. The numbers count different pathways enriched by the detected proteins. The DDA ladder uncovered several important pathways for neurobiology that would have been lost to limited utilization efficiency of the DDA cycle. Pathway annotation is provided in **Table S11**.

REFERENCES

1. Hebert, A. S.; Richards, A. L.; Bailey, D. J.; Ulbrich, A.; Coughlin, E. E.; Westphall, M. S.; Coon, J. J., The one hour yeast proteome. *Mol. Cell. Proteomics* **2014**, *13* (1), 339-347.
2. Picotti, P.; Clement-Ziza, M.; Lam, H.; Campbell, D. S.; Schmidt, A.; Deutsch, E. W.; Rost, H.; Sun, Z.; Rinner, O.; Reiter, L.; Shen, Q.; Michaelson, J. J.; Frei, A.; Alberti, S.; Kusebauch, U.; Wollscheid, B.; Moritz, R. L.; Beyer, A.; Aebersold, R., A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. *Nature* **2013**, *494* (7436), 266-270.
3. Altschuler, S. J.; Wu, L. F., Cellular heterogeneity: Do differences make a difference? *Cell* **2010**, *141* (4), 559-563.
4. Choi, S. B.; Lombard-Banek, C.; Munoz-Llancao, P.; Manzini, M. C.; Nemes, P., Enhanced peptide detection toward single-neuron proteomics by reversed-phase fractionation capillary electrophoresis mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (5), 913-922.
5. Nisar, M. S.; Zhao, X., High resolution mass spectrometry for single cell analysis. *International Journal of Mass Spectrometry* **2020**, *450*, 116302.
6. Moshkovskii, S. A.; Lobas, A. A.; Gorshkov, M. V., Single cell proteogenomics — Immediate prospects. *Biochemistry (Moscow)* **2020**, *85* (2), 140-146.
7. Evers, T. M. J.; Hochane, M.; Tans, S. J.; Heeren, R. M. A.; Semrau, S.; Nemes, P.; Mashaghi, A., Deciphering metabolic heterogeneity by single-cell analysis. *Anal. Chem.* **2019**, *91* (21), 13314-13323.
8. Richards, A. L.; Hebert, A. S.; Ulbrich, A.; Bailey, D. J.; Coughlin, E. E.; Westphall, M. S.; Coon, J. J., One-hour proteome analysis in yeast. *Nat. Protoc.* **2015**, *10* (5), 701-714.
9. Woo, J.; Williams, S. M.; Aguilera-Vazquez, V.; Sontag, R. L.; Moore, R. J.; Markillie, L. M.; Mehta, H. S.; Cantlon, J.; Adkins, J. N.; Smith, R. D.; Clair, G. C.; Pasa-Tolic, L.; Zhu, Y., High-throughput and high-efficiency sample preparation for single-cell proteomics using a nested nanowell chip. *bioRxiv* **2021**, 2021.02.17.431689.
10. Petelski, A. A.; Emmott, E.; Leduc, A.; Huffman, R. G.; Specht, H.; Perlman, D. H.; Slavov, N., Multiplexed single-cell proteomics using SCoPE2. *bioRxiv* **2021**, 2021.03.12.435034.
11. Gomes, F. P.; Yates, III, Recent trends of capillary electrophoresis mass spectrometry in proteomics research. *Mass Spectrom. Rev.* **2019**, *38* (6), 445-460.
12. Drouin, N.; van Mever, M.; Zhang, W.; Tobolkina, E.; Ferre, S.; Servais, A. C.; Gou, M. J.; Nyssen, L.; Fillet, M.; Lageveen-Kammeijer, G. S. M.; Nouta, J.; Chetwynd, A. J.; Lynch, I.; Thorn, J. A.; Meixner, J.; Lossner, C.; Taverna, M.; Liu, S.; Tran, N. T.; Francois, Y.; Lechner, A.; Nehm, R.; Banni, G. A. D.; Nasreddine, R.; Colas, C.; Lindner, H. H.; Faserl, K.; Neususs, C.; Nelke, M.; Lammerer, S.; Perrin, C.; Bich-Muracciale, C.; Barbas, C.; Gonzalez, A. L.; Guttman, A.; Szigeti, M.; Britz-McKibbin, P.; Kroezen, Z.; Shanmuganathan, M.; Nemes, P.; Portero, E. P.; Hankemeier, T.; Codesido, S.; Gonzalez-Ruiz, V.; Rudaz, S.; Ramautar, R., Capillary electrophoresis mass spectrometry at trial by metabo-ring: Effective electrophoretic mobility for reproducible and robust compound annotation. *Anal. Chem.* **2020**, *92* (20), 14103-14112.
13. Wang, Y.; Fonslow, B. R.; Wong, C. C. L.; Nakorchevsky, A.; Yates, J. R., Improving the comprehensiveness and sensitivity of sheathless capillary electrophoresis–tandem mass spectrometry for proteomic analysis. *Anal. Chem.* **2012**, *84* (20), 8505-8513.
14. Lombard-Banek, C.; Moody, S. A.; Nemes, P., Single-cell mass spectrometry for discovery proteomics: quantifying translational cell heterogeneity in the 16-cell frog (*Xenopus*) embryo. *Angew. Chem.-Int. Edit.* **2016**, *55* (7), 2454-2458.

15. Lombard-Banek, C.; Reddy, S.; Moody, S. A.; Nemes, P., Label-free quantification of proteins in single embryonic cells with neural fate in the cleavage-stage frog (*Xenopus laevis*) embryo using capillary electrophoresis electrospray ionization high-resolution mass spectrometry (CE-ESI-HRMS). *Mol. Cell. Proteomics* **2016**, *15* (8), 2756-2768.
16. Lombard-Banek, C.; Li, J.; Portero, E. P.; Onjiko, R. M.; Singer, C. D.; Plotnick, D.; Al Shabeeb, R. Q.; Nemes, P., *In vivo* subcellular mass spectrometry enables proteo-metabolomic single-cell systems biology in a chordate embryo developing to a normally behaving tadpole (*X. laevis*). *Angew. Chem.-Int. Edit.* **2021**, *60* (23), 12852-12858.
17. Lombard-Banek, C.; Moody, S. A.; Manzin, 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. *Anal. Chem.* **2019**, *91* (7), 4797-4805.
18. Choi, S. B.; Zamarbide, M.; Manzini, M. C.; Nemes, P., Tapered-tip capillary electrophoresis nano-electrospray ionization mass spectrometry for ultrasensitive proteomics: the mouse cortex. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 597-607.
19. Sun, L. L.; Zhu, G. J.; Zhang, Z. B.; Mou, S.; Dovichi, N. J., Third-generation electrokinetically pumped sheath-flow nanospray interface with improved stability and sensitivity for automated capillary zone electrophoresis mass spectrometry analysis of complex proteome digests. *J. Proteome Res.* **2015**, *14* (5), 2312-2321.
20. Zhang, Z. B.; Hebert, A. S.; Westphall, M. S.; Qu, Y. Y.; Coon, J. J.; Dovichi, N. J., Production of over 27 000 peptide and nearly 4400 protein identifications by single-shot capillary zone electrophoresis mass spectrometry via combination of a very-low electroosmosis coated capillary, a third-generation electrokinetically-pumped sheath-flow nanospray interface, an orbitrap fusion lumos tribrid mass spectrometer, and an advanced-peak determination algorithm. *Anal. Chem.* **2018**, *90* (20), 12090-12093.
21. Sun, L. L.; Zhu, G. J.; Zhao, Y. M.; Yan, X. J.; Mou, S.; Dovichi, N. J., Ultrasensitive and fast bottom-up analysis of femtogram amounts of complex proteome digests. *Angew. Chem.-Int. Edit.* **2013**, *52* (51), 13661-13664.
22. Liu, H. B.; Sadygov, R. G.; Yates, J. R., A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **2004**, *76* (14), 4193-4201.
23. Tabb, D. L.; Vega-Montoto, L.; Rudnick, P. A.; Variyath, A. M.; Ham, A. J. L.; Bunk, D. M.; Kilpatrick, L. E.; Billheimer, D. D.; Blackman, R. K.; Cardasis, H. L.; Carr, S. A.; Clauser, K. R.; Jaffe, J. D.; Kowalski, K. A.; Neubert, T. A.; Regnier, F. E.; Schilling, B.; Tegeler, T. J.; Wang, M.; Wang, P.; Whiteaker, J. R.; Zimmerman, L. J.; Fisher, S. J.; Gibson, B. W.; Kinsinger, C. R.; Mesri, M.; Rodriguez, H.; Stein, S. E.; Tempst, P.; Paulovich, A. G.; Liebler, D. C.; Spiegelman, C., Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2010**, *9* (2), 761-776.
24. Zhang, Y. Y.; Fonslow, B. R.; Shan, B.; Baek, M. C.; Yates, J. R., Protein analysis by shotgun/bottom-up proteomics. *Chem. Rev.* **2013**, *113* (4), 2343-2394.
25. Wang, N.; Li, L., Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Anal. Chem.* **2008**, *80* (12), 4696-4710.
26. Kreimer, S.; Belov, M. E.; Danielson, W. F.; Levitsky, L. I.; Gorshkov, M. V.; Karger, B. L.; Ivanov, A. R., Advanced precursor ion selection algorithms for increased depth of bottom-up proteomic profiling. *J. Proteome Res.* **2016**, *15* (10), 3563-3573.

27. Wu, W. W.; Shen, R. F.; Park, S. S.; Martin, B.; Maudsley, S., Precursor ion exclusion for enhanced identification of plasma biomarkers. *Proteom. Clin. Appl.* **2012**, *6* (5-6), 304-308.
28. Voisin, S. N.; Krakovska, O.; Matta, A.; DeSouza, L. V.; Romaschin, A. D.; Colgan, T. J.; Siu, K. W. M., Identification of novel molecular targets for endometrial cancer using a drill-down LC-MS/MS approach with iTRAQ. *PLoS One* **2011**, *6* (1), e16352.
29. Zhang, Z. Q., Automated precursor ion exclusion during LC-MS/MS data acquisition for optimal ion identification. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (8), 1400-1407.
30. Hoopmann, M. R.; Merrihew, G. E.; von Haller, P. D.; MacCoss, M. J., Post analysis data acquisition for the iterative MS/MS sampling of proteomics mixtures. *J. Proteome Res.* **2009**, *8* (4), 1870-1875.
31. Bailey, D. J.; McDevitt, M. T.; Westphall, M. S.; Pagliarini, D. J.; Coon, J. J., Intelligent data acquisition blends targeted and discovery methods. *J. Proteome Res.* **2014**, *13* (4), 2152-2161.
32. Kaech, S.; Banker, G., Culturing hippocampal neurons. *Nat. Protoc.* **2006**, *1* (5), 2406-2415.
33. Nemes, P.; Marginean, I.; Vertes, A., Spraying mode effect on droplet formation and ion chemistry in electrosprays. *Anal. Chem.* **2007**, *79* (8), 3105-3116.
34. Tyanova, S.; Temu, T.; Cox, J., The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **2016**, *11* (12), 2301-2319.
35. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Oresic, M., MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **2010**, *11*, 11.
36. Mi, H. Y.; Muruganujan, A.; Thomas, P. D., PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* **2013**, *41* (D1), D377-D386.
37. Lau, C.; Ng, L.; Thompson, C.; Pathak, S.; Kuan, L.; Jones, A.; Hawrylycz, M., Exploration and visualization of gene expression with neuroanatomy in the adult mouse brain. *BMC Bioinformatics* **2008**, *9*, 11.
38. Chen, D. Y.; Shen, X. J.; Sun, L. L., Capillary zone electrophoresis-mass spectrometry with microliter-scale loading capacity, 140 min separation window and high peak capacity for bottom-up proteomics. *Analyst* **2017**, *142* (12), 2118-2127.
39. Chen, D.; Shen, X.; Sun, L., Strong cation exchange-reversed phase liquid chromatography-capillary zone electrophoresis-tandem mass spectrometry platform with high peak capacity for deep bottom-up proteomics. *Analytica Chimica Acta* **2018**, *1012* (5), 1-9.
40. Zhang, Z. B.; Yan, X. J.; Sun, L. L.; Zhu, G. J.; 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. *Anal. Chem.* **2015**, *87* (8), 4572-4577.
41. Shen, X. J.; Kou, Q.; Guo, R. Q.; Yang, Z. C.; Chen, D. Y.; Liu, X. W.; Hong, H.; Sun, L. L., Native proteomics in discovery mode using size-exclusion chromatography capillary zone electrophoresis tandem mass spectrometry. *Anal. Chem.* **2018**, *90* (17), 10095-10099.