

Video Article

# Large-scale Top-down Proteomics Using Capillary Zone Electrophoresis Tandem Mass Spectrometry

Elijah N. McCool<sup>\*1</sup>, Rachele Lubeckyj<sup>\*1</sup>, Xiaojing Shen<sup>1</sup>, Qiang Kou<sup>2</sup>, Xiaowen Liu<sup>2,3</sup>, Liangliang Sun<sup>1</sup>

<sup>1</sup>Department of Chemistry, Michigan State University

<sup>2</sup>Department of BioHealth Informatics, Indiana University-Purdue University Indianapolis

<sup>3</sup>Center for Computational Biology and Bioinformatics, Indiana University School of Medicine

\*These authors contributed equally

Correspondence to: Liangliang Sun at [lsun@chemistry.msu.edu](mailto:lsun@chemistry.msu.edu)

URL: <https://www.jove.com/video/58644>

DOI: [doi:10.3791/58644](https://doi.org/10.3791/58644)

Keywords: Chemistry, Issue 140, Top-down proteomics, capillary zone electrophoresis, electrospray ionization, tandem mass spectrometry, proteoform, *Escherichia coli*

Date Published: 10/24/2018

Citation: McCool, E.N., Lubeckyj, R., Shen, X., Kou, Q., Liu, X., Sun, L. Large-scale Top-down Proteomics Using Capillary Zone Electrophoresis Tandem Mass Spectrometry. *J. Vis. Exp.* (140), e58644, doi:10.3791/58644 (2018).

## Abstract

Capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry (CZE-ESI-MS/MS) has been recognized as a useful tool for top-down proteomics that aims to characterize proteoforms in complex proteomes. However, the application of CZE-MS/MS for large-scale top-down proteomics has been impeded by the low sample-loading capacity and narrow separation window of CZE. Here, a protocol is described using CZE-MS/MS with a microliter-scale sample-loading volume and a 90-min separation window for large-scale top-down proteomics. The CZE-MS/MS platform is based on a linear polyacrylamide (LPA)-coated separation capillary with extremely low electroosmotic flow, a dynamic pH-junction-based online sample concentration method with a high efficiency for protein stacking, an electro-kinetically pumped sheath flow CE-MS interface with extremely high sensitivity, and an ion trap mass spectrometer with high mass resolution and scan speed. The platform can be used for the high-resolution characterization of simple intact protein samples and the large-scale characterization of proteoforms in various complex proteomes. As an example, a highly efficient separation of a standard protein mixture and a highly sensitive detection of many impurities using the platform is demonstrated. As another example, this platform can produce over 500 proteoform and 190 protein identifications from an *Escherichia coli* proteome in a single CZE-MS/MS run.

## Video Link

The video component of this article can be found at <https://www.jove.com/video/58644/>

## Introduction

Top-down proteomics (TDP) aims for the large-scale characterization of proteoforms within a proteome. TDP relies on the effective liquid-phase separation of intact proteins before electrospray ionization-tandem mass spectrometry (ESI-MS/MS) analysis due to the high complexity and large concentration dynamic range of the proteome<sup>1,2,3,4,5</sup>. Capillary zone electrophoresis (CZE) is a powerful technique for the separation of biomolecules based on their size-to-charge ratios<sup>6</sup>. CZE is relatively simple, requiring only an open tubular-fused silica capillary, a background electrolyte (BGE), and a power supply. A sample of intact proteins can be loaded into the capillary using pressure or voltage, and separation is initiated by immersing both ends of the capillary in the BGE and applying a high voltage. CZE can approach ultra-high separation efficiency (> one million theoretical plates) for the separation of biomolecules<sup>7</sup>. CZE-MS has a drastically higher sensitivity than widely used reversed-phase liquid chromatography (RPLC)-MS for the analysis of intact proteins<sup>8</sup>. Although CZE-MS has a great potential for large-scale top-down proteomics, its wide application in proteomics has been impeded by several issues, including a low sample-loading capacity and narrow separation window. The typical sample loading volume in CZE is about 1% of the total capillary volume, which usually corresponds to less than 100 nL<sup>9,10,11</sup>. The separation window of CZE is usually less than 30 min due to the strong electroosmotic flow (EOF)<sup>9,10</sup>. These issues limit the CZE-MS/MS for the identification of a large number of proteoforms and low abundant proteoforms from a complex proteome.

Much effort has been made to improve the sample loading volume of CZE via online sample concentration methods (e.g., solid-phase microextraction [SPME]<sup>12,13</sup>, field-enhanced sample stacking [FESS]<sup>9,11,14</sup>, and dynamic pH junction<sup>15,16,17,18</sup>). FESS and dynamic pH junction are simpler than SPME, only requiring a significant difference between the sample buffer and the BGE in conductivity and pH. FESS employs a sample buffer with much lower conductivity than the BGE, leading to a stacking of analytes on the boundary between the sample zone and the BGE zone in the capillary. Dynamic pH junction utilizes a basic sample plug (e.g., 50 mM ammonium bicarbonate, pH 8) and an acidic BGE (e.g., 5% [v/v] acetic acid, pH 2.4) on both sides of the sample plug. Upon application of a high positive voltage at the injection end of the capillary, titration of the basic sample plug occurs, focusing the analytes into a tight plug before undergoing a CZE separation. Recently, the Sun group systematically compared FESS and dynamic pH junction for the online stacking of intact proteins, demonstrating that dynamic pH junction could

produce much better performance than FESS for the online concentration of intact proteins when the sample injection volume was 25% of the total capillary volume<sup>19</sup>.

Neutrally coated separation capillaries (e.g., linear polyacrylamide [LPA]) have been employed to reduce the EOF in the capillary, slowing down the CZE separation and widening the separation window<sup>20,21</sup>. Recently, the Dovichi group developed a simple procedure for the preparation of stable LPA coating on the inner wall of capillaries, utilizing ammonium persulfate (APS) as the initiator and temperature (50 °C) for free radical production and polymerization<sup>22</sup>. Very recently, the Sun group employed the LPA-coated separation capillary and the dynamic pH junction method for the CZE separation of intact proteins, reaching a microliter-scale sample loading volume and a 90-min separation window<sup>19</sup>. This CZE system opens the door to using CZE-MS/MS for large-scale top-down proteomics.

CZE-MS requires a highly robust and sensitive interface to couple CZE to MS. Three CE-MS interfaces have been well developed and commercialized in the history of CE-MS, and they are the co-axial sheath-flow interface<sup>23</sup>, the sheathless interface using a porous tip as the ESI emitter<sup>24</sup>, and the electro-kinetically pumped sheath flow interface<sup>25,26</sup>. The electro-kinetically pumped sheath-flow-interface-based CZE-MS/MS has reached a low zeptomole peptide detection limit<sup>9</sup>, over 10,000 peptide identifications (IDs) from the HeLa cell proteome in a single run<sup>14</sup>, a fast characterization of intact proteins<sup>11</sup>, and highly stable and reproducible analyses of biomolecules<sup>26</sup>. Recently, the LPA-coated separation capillary, the dynamic pH junction method, and the electro-kinetically pumped sheath flow interface were used for large-scale top-down proteomics of an *Escherichia coli* (*E. coli*) proteome<sup>19,27</sup>. The CZE-MS/MS platform approached over 500 proteoform IDs in a single run<sup>19</sup> and nearly 6,000 proteoform IDs via coupling with size-exclusion chromatography (SEC)-RPLC fractionation<sup>27</sup>. The results clearly show the capability of CZE-MS/MS for large-scale top-down proteomics.

Herein, a detailed procedure of using CZE-MS/MS for large-scale top-down proteomics is described. The CZE-MS/MS system employs the LPA-coated capillary to reduce the EOF in the capillary, the dynamic pH junction method for the online concentration of proteins, the electro-kinetically pumped sheath flow interface for coupling CZE to MS, an orbitrap mass spectrometer for the collection of MS and MS/MS spectra of proteins, and a TopPIC (TOP-Down Mass Spectrometry-Based Proteoform Identification and Characterization) software for proteoform ID via database search.

## Protocol

### 1. Preparation of LPA Coating on the Inner Wall of the Separation Capillary

#### 1. Pretreatment of the capillary

1. Flush a fused silica capillary (120 cm in length, 50 µm in inner diameter [i.d.], 360 µm in outer diameter [o.d.]) successively with 500 µL of 1 M sodium hydroxide, deionized water, 1 M hydrochloric acid, deionized water, and LC-MS grade methanol using a syringe pump.
2. Dry the capillary with nitrogen gas (10 psi, ≥ 12 h) and fill the capillary with 50% (v/v) 3-(trimethoxysilyl)propyl methacrylate in methanol using a syringe pump. Seal both ends of the capillary with silica rubber and incubate it at room temperature for at least 24 h.  
NOTE: During this step, the inner wall of the capillary is functionalized with C=C. Longer incubation results in better capillary coating due to a more complete reaction.
3. Cut a small portion (~5 mm) of the capillary from both ends with a cleaving stone. Rinse the capillary with methanol (500 µL) using a syringe pump to clean up the unreacted reagents. Dry the capillary with nitrogen gas (10 psi, ≥ 12 h).

#### 2. Preparation of the LPA coating

NOTE: This procedure is based on Zhu *et al.*<sup>22</sup> with minor modifications.

1. Prepare an acrylamide solution (40 mg of acrylamide in 1 mL of water) and ammonium persulfate (APS) solution (5% [w/v] in water).
2. Add 2 - 3 µL of the APS solution to 500 µL of the acrylamide solution, vortex the mixture, and degas it with nitrogen gas for 5 min to remove the oxygen in the solution.
3. Load the mixture into the pretreated capillary using a vacuum, seal both ends of the capillary with silica rubber, and incubate it in a water bath at 50 °C for 40 min.
4. Remove a small portion (~5 mm) of the capillary from both ends with a cleaving stone. Push the unreacted solution out of the capillary with water (200 µL), using the syringe pump.

NOTE: Ensure the polymer pushed out of the capillary is an agarose gel-like consistency. A longer incubation step (up to ~45 - 50 min) can result in a better capillary coating. With longer reaction periods, the capillary may become blocked and high pressure is required to push out the polymer with water. An HPLC pump can be used for this purpose.

### 2. Etching of the Capillary with Hydrofluoric Acid

CAUTION: Use appropriate safety procedures while handling hydrofluoric acid (HF) solutions. All the HF-related operations need to be done in a chemical hood. Before any HF-related operation, make sure that 2.5% calcium gluconate gel is available for use in the case of exposure. Double gloves are required, a typical nitrile glove inside and a heavy neoprene glove outside. Wear a lab coat and chemical safety goggles. After the HF operations, keep liquid and solid hazardous waste separate. The liquid HF waste must be neutralized immediately with a high-concentration sodium hydroxide solution for temporary storage before waste pick-up. The solid HF waste needs to be temporarily stored in a plastic container that is lined with two thick plastic one-gallon Ziploc bags and a lid. Both the solid and liquid waste must be labeled properly.

1. Burn a portion of the capillary using a gentle flame (e.g., pocket lighter) to remove the polyimide outer-coating (1 cm in length) around 4 cm away from one end of the capillary. Gently clean the burnt portion of the capillary by wiping to remove the polyimide coating completely.  
NOTE: Proceed with caution, as the portion of the capillary without the polyimide coating will be a little fragile.

2. Drill a small hole at the end of a 200- $\mu$ L tube around the same size as the capillary outer diameter to sufficiently hold the capillary in place once it is threaded through this hole. Thread the end of the capillary that is close to the burnt portion through the hole until the burnt portion is in the tube.  
NOTE: It is recommended to test the size of the hole with the end of the capillary away from the burnt portion to ensure the correct size, as the burnt portion of the capillary is fragile.
3. Add ~150  $\mu$ L of HF (48% - 51% solution in water) to the 200- $\mu$ L tube so that the HF solution is about halfway up the burnt portion on the capillary. Incubate the capillary in the HF solution at room temperature for 90 - 100 min.  
NOTE: It is recommended that another hole is created in the lid of the tube and some small object (e.g., a small pipette tip) is used to hold up the tube while the incubation is taking place. The pipette tip can be used to puncture foam or some other solid platform from which the reaction chamber can hang from, creating a safe environment. Place paper towels below the tube containing HF and follow proper procedures in case of a spill.
4. Remove the capillary from the tube and wash the exterior with deionized water to remove any residual HF.  
NOTE: Ensure that the outer diameter of the capillary at the narrowest portion of the burnt part is now smaller than 100  $\mu$ m, as it should be.
5. Cut the etched part of the capillary at the middle of the narrowest portion using a cleaving stone to produce a separation capillary with less than 100  $\mu$ m in o.d. at one end. Cut the non-etched end of the capillary to get a 1-m separation capillary.  
NOTE: Dispose of HF waste according to the institutionally prescribed protocol.

### 3. Preparation of the Samples

#### 1. Preparation of a standard protein mixture

1. Prepare a standard protein mixture containing cytochrome c (Cyto.c, 12 kDa, 0.1 mg/mL), lysozyme (14.3 kDa, 0.1 mg/mL),  $\beta$ -casein (24 kDa, 0.4 mg/mL), myoglobin (16.9 kDa, 0.1 mg/mL), carbonic anhydrase (CA, 29 kDa, 0.5 mg/mL), and bovine serum albumin (BSA, 66.5 kDa, 1.0 mg/mL) in LC-MS grade water as a stock solution.  
NOTE: The stock solution can be aliquoted and stored at -80 °C for use.
2. Dilute the stock solution by a factor of 10 with a 50 mM ammonium bicarbonate solution (pH 8.0) in LC-MS grade water for CZE-MS analysis.  
NOTE: Filter the ammonium bicarbonate solution before use with a membrane filter (e.g., 0.2  $\mu$ m of cellulose nitrate membrane and 50 mm in diameter).

#### 2. Preparation of an *E. coli* sample

1. Culture *E. coli* (K-12 MG1655) cells in LB medium at 37 °C while shaking at 225 rpm until the OD600 value approaches 0.7. Collect *E. coli* cells via centrifugation (3,283  $\times g$ , 10 min) and wash them multiple times with phosphate-buffered saline (PBS) to remove the medium.
2. Suspend the *E. coli* cells in the lysis buffer containing 8 M urea, protease inhibitors, and 100 mM ammonium bicarbonate (pH 8.0) and ultrasonicate on ice for 15 min using an ultrasonic cell disruptor with a cup horn for complete cell lysis.
  1. Set the **Duty Cycle (%)** to 50 and set the **Output Control** to 7 for the ultrasonic cell disruptor.
3. Centrifuge the lysate at 18,000  $\times g$  for 10 min. Collect the supernatant and discard the pellet. Use a small aliquot of the supernatant for protein concentration measurement with the bicinchoninic acid (BCA) assay.  
NOTE: The lysate needs to be diluted at least by a factor of three to reduce the urea concentration lower than 3 M in order to become compatible with the BCA assay. The BCA assay is performed based on the manufacturer's procedure.
4. Mix 1 mg of *E. coli* proteins with cold acetone with a volume ratio of 1:4 and keep the mixture at -20 °C overnight to precipitate the proteins.  
NOTE: Perform all experiments using acetone in a chemical hood.
5. Spin down precipitated proteins (12,000  $\times g$ , 5 min) and remove the supernatant. Wash the pellet with cold acetone again (same volume as before) and spin down again.
6. Remove the supernatant and allow the pellet to dry in the chemical hood for a couple of minutes.  
NOTE: Do not overdry the protein pellet.
7. Dissolve the precipitated *E. coli* proteins (1 mg) in 200  $\mu$ L of 8 M urea in 100 mM ammonium bicarbonate (pH 8.0).
8. Denature, reduce, and alkylate the sample.
  1. Incubate the sample at 37 °C for 30 min to denature the proteins.
  2. Reduce with dithiothreitol (DTT) by adding 2  $\mu$ L of 1 M DTT solution (in 100 mM ammonium bicarbonate) into the sample and incubating the sample at 37 °C for 30 min.
  3. Alkylate the proteins with iodoacetamide (IAA) by adding 6  $\mu$ L of 1 M IAA solution (in 100 mM ammonium bicarbonate) to the sample and incubate the sample for 20 min at room temperature in the dark.
  4. Quench the excess IAA with DTT by adding 2  $\mu$ L of 1 M DTT solution and incubate for 5 min at room temperature. Acidify the sample with formic acid (FA) to get a final FA concentration of 1% (v/v).  
NOTE: Reduction and alkylation are performed to aid the unfolding of proteins for downstream fragmentation. Those steps will lose the information of disulfide bonds. If the goal is to study the disulfide bonds on the proteins, avoid those steps. Remember to handle all concentrated acid solutions in a chemical hood.
9. Desalt the sample using a C4 trap column (e.g., 4 mm of i.d., 10 mm in length, packed with 3- $\mu$ m particles having 300-Å pores) with an HPLC system. Use a UV detector at a wavelength of 254 nm for detection.
  1. Set the mobile phase flow rate to 1 mL/min. Activate the column by flushing it with 80% (v/v) acetonitrile (ACN), 0.1% FA in water for 10 min, and equilibrate by flushing it with 2% (v/v) ACN, 0.1% FA in water for 10 min.
  2. Load 500- $\mu$ g proteins onto the column and desalt by flushing them with 2% (v/v) ACN, 0.1% FA in water for 10 min at a 1 mL/min flow rate.
  3. Elute the proteins with 80% (v/v) ACN, 0.1% FA in water for 3 min at a 1 mL/min flow rate. Collect the eluate and lyophilize it with a vacuum concentrator.

NOTE: The C4 trap column can be used to desalt large amounts of proteins but not low micrograms of proteins due to possible sample loss. For limited protein materials, consider using pipette tips with C4 media for protein desalting. Be careful not to overdry the proteins when lyophilizing the sample.

10. Dissolve the 500 µg of proteins (assuming no sample loss during the sample preparation) in 250 µL of 50 mM ammonium bicarbonate (pH 8.0) for CZE-MS experiments.

## 4. Set-up of the CZE-MS/MS System and Analysis of the Samples

### 1. Set-up of the CZE system

1. Prepare a BGE containing 5% or 10% (v/v) acetic acid (pH ~2.4 or ~2.2) in LC-MS grade water. Put ~1.5 mL of the BGE into a BGE vial that matches with the buffer tray of the CZE autosampler.  
NOTE: Make fresh BGE every week and change the BGE solution in the vial every day to avoid any contamination.
2. Add a sample solution (5 - 200 µL) to an insert tube and put the insert tube into a sample vial that matches with the sample tray of the CZE autosampler.
3. Load the sample, the BGE, and the separation capillary into the CZE autosampler.  
NOTE: The language used in this protocol most accurately reflects the use of one particular autosampler (see **Table of Materials**), but the principles can be applied to other autosamplers.
  1. Select **Sample load** in the manual control page of the CZE autosampler. Load the BGE vial into the buffer tray of the autosampler, noting its position in the buffer tray. Load the sample vial into the sample tray of the autosampler, noting its position in the sample tray.  
NOTE: The instrument can be operated in manual control by clicking on the picture of the instrument for **Device Monitor** on the main instrument page and then selecting **Manual** under **Direct Control**.
  2. Load the separation capillary into the CZE autosampler, using the non-etched end of the capillary. Put the autosampler to the **Alignment** position using manual control. Thread the non-etched end of the capillary into a hole that is used to hold the separation capillary until it cannot be physically threaded further.  
NOTE: In this case, the end of the capillary is almost at the same height as the end of the electrode, and at least 50 µL of the sample is required in the sample vial to make sure the end of the capillary is immersed in the sample during injection. If the sample volume is low (*i.e.*, 5 µL), the height of the injection end of the capillary needs to be adjusted as described in step 4.1.3.2.1.
    1. Adjust the height of the injection end of the capillary if the sample volume is lower than 50 µL (*i.e.*, 5 µL). Switch the autosampler to **Standby** using manual control. Type the sample position in the system and move the capillary to the sample. Push the injection end of the capillary further down to reach the bottom of the sample vial.  
NOTE: In this case, the sample injection can be performed with a sample of only 5 µL in the vial.
  3. Continuing to use the manual control, flush the capillary with the BGE for 20 min, using a pressure of 20 psi.

### 2. Set-up of the CZE-MS interface

1. Pull a borosilicate glass capillary (1 mm in o.d., 0.75 mm in i.d., 10 cm long) into two electrospray emitters with an orifice of 20 - 40 µm in o.d.  
NOTE: Keep the length of the electrospray emitter as 4 - 5 cm and cut it with a cleaving stone if necessary. Dispose of all glass waste in a sharps container. The settings to get 20- to 40-µm tips are as follows. Set the heat to 498, the pull to 5, the velocity to 10, the delay to 150, and the pressure to 200. Check the size of the emitters with a microscope before use. Adjust the parameters slightly if necessary.
2. Mount a commercial electro-kinetically pumped sheath flow interface (see **Table of Materials**) to the front of the mass spectrometer. Fill the sheath buffer reservoir with a buffer containing 10% (v/v) methanol and 0.2% (v/v) FA in LC-MS grade water.
3. Flush the **T** in the interface with the sheath buffer *via* applying pressure manually with a syringe. Thread the 1-mm-o.d. end of an electrospray emitter through a sleeve tubing and connect the emitter with one port of the **T** *via* a fitting. Simply flush the **T** manually again with a syringe to fill the emitter with the sheath buffer.
4. Adjust the distance between the orifice of the emitter and the entrance of the mass spectrometer to ~2 mm with the help of the camera that came with the interface. Apply a 2- to 2.2-kV voltage at the sheath buffer vial for electrospray. Adjust the spray voltage to reach a stable electrospray.  
NOTE: If no electrospray or an unstable electrospray is observed, check the interface to make sure there are no large bubbles in the emitter, in the **T**, and in the tubing that used to connect the sheath buffer vial and the **T**. The distance between the orifice of the emitter and the mass spectrometer entrance can be roughly estimated by comparing it with the 1-mm o.d. of the emitter. The spray voltage depends on the size of the emitter. For 20- to 40-µm emitters, 2 - 2.2 kV is good enough. Remember to always be careful when using high voltages.
5. Turn off the spray voltage and gently thread the etched end of the separation capillary through the **T** into the emitter until it cannot be pushed further. Apply a low pressure (*i.e.*, 5 psi) at the injection end of the capillary during this process to make sure there are no bubbles in the capillary.  
NOTE: The capillary is connected to the **T** *via* a sleeve tubing and a fitting. Adjust the position of the etched end of the capillary in the emitter with the help of the camera to within 500 µm. The distance can be roughly estimated by comparing it with the 1-mm o.d. of the emitter.
6. Stop the low pressure at the injection end of the capillary. Flush the emitter a little bit with the sheath buffer. Again, apply 2 - 2.2 kV of spray voltage to test the spray.

### 3. Set-up of a CZE method for sample injection, separation, capillary flushing, and triggering of the mass spectrometer for data acquisition

1. Select **New method** under the **File** drop-down menu on the main instrument screen to initiate a new method.
  2. Select **Inlet** as SV/EV, **Tray** as Sample, **Pressure** as 5 psi, **KV** as 0, and **Duration** as 95 s for a sample injection.  
NOTE: The sample is injected into the capillary via applying pressure. The sample injection volume can be calculated based on the pressure and injection time using Poiseuille's Law. For example, 5 psi for a 95-s sample injection corresponds to about 500 nL of sample-loading volume for a 1-m-long separation capillary (50- $\mu$ m i.d.).
  3. Select parameters for the CZE separation and set up parameters for triggering the data acquisition.
    1. Set **Inlet** as InV, **Tray** as Buffer, **Pressure** as 0 psi, **KV** as 30, and **Duration** as 4,200 s for the separation of the standard protein mixture sample. Set **Inlet** as InV, **Tray** as Buffer, **Pressure** as 0 psi, **KV** as 20, and **Duration** as 6,600 s for the separation of the *E. coli* proteome sample.  
NOTE: The voltage applied for the separation can be adjusted based on the sample complexity. For simple protein samples, 30 kV can be applied to speed up the analysis. For complex samples, 20 kV is applied to slow down the separation and acquire more MS/MS spectra for proteoform IDs.
    2. Set up the parameters for triggering the data acquisition under **Timed Events**. Set **Time (s)** as 0.0 and **Type** as Relay 1 for **Activate** in step 1; Set **Time (s)** as 1.0 and **Type** as Relay 1 for **Deactivate** in step 2.
  4. Select **Inlet** as InV, **Tray** as Buffer, **Pressure** as 10 psi, **KV** as 30, and **Duration** as 600 s for capillary flushing.
  5. Save the method files for the CZE-MS and MS/MS experiments.
4. **Set-up of the MS and MS/MS**
1. Set up the MS and MS/MS parameters for intact protein analysis using a quadrupole-ion trap mass spectrometer (see **Table of Materials**).  
NOTE: The positive ion mode and higher energy collisional dissociation (HCD) for fragmentation are employed.
    1. Adjust the tune file settings: turn on **Intact protein mode** and use a trapping pressure of 0.2. Set the ion transfer capillary temperature to 320 °C and the s-lens RF level to 55.
    2. Build the MS and MS/MS method file.
      1. For full MS, set the number of microscans to 3, the resolution to 240,000 (at  $m/z$  200), the automatic gain control (AGC) target value to 1E6, the maximum injection time to 50 ms, and the scan range to 600 - 2000  $m/z$ .
    3. For MS/MS, use data-dependent acquisition (DDA) for the eight most intense ions in a full MS spectrum. Set the isolation window to 4  $m/z$  and the normalized collisional energy (NCE) to 20%. Set the number of microscans to 1, the resolution to 120,000 (at  $m/z$  200), the AGC target value to 1E5, and the maximum injection time to 200 ms.
    4. Set the intensity threshold for triggering fragmentation to 1E5 and set the ions with a charge state higher than 5 to be isolated for fragmentation. Turn on the **exclude isotopes** and set the dynamic exclusion to 30 s.  
NOTE: The number of microscans for MS/MS can be increased to 3. In this case, a Top3 DDA method can be used in order to reduce the cycle time.
  2. Set up a wired connection between the CE autosampler and the mass spectrometer for the automatic triggering of the data acquisition. Connect one end of a wire to the **Relay 1 Contact A** and **Relay 1 Contact B** on the back of the CE autosampler. Connect the other end of the wire to **Start In -** and **Start In +** on the side of the mass spectrometer.
5. **CZE-MS/MS experiment and analysis of the samples**
1. Apply 30 kV using the CE manual control at the sample injection end and 2 - 2.2 kV at the sheath buffer vial using the external power supply to test the separation capillary and electrospray.  
NOTE: The separation current, in this case, is around 8 - 9  $\mu$ A if 5% (v/v) acetic acid is used as the BGE.
  2. Set up a data acquisition sequence on the mass spectrometer computer by selecting a new sequence.
    1. Select **Unknown** as the sample type and specify a file name and path. Indicate the MS method to be used for each sample run under **Instrument Method**.  
NOTE: Since there is a separate computer for controlling the CE autosampler, the sample position does not need to be specified here.
  3. Set up a CZE separation sequence on the CE computer to indicate the buffer position, sample position, and the CZE method. To start a new sequence, select **Sequence** under the **Analysis** drop-down menu on the main instrument page.
  4. Start the data acquisition method on the mass spectrometer computer first by selecting **Run sequence** (or **Run sample** for single runs). Then, start the CE sequence on the CE autosampler computer.  
NOTE: When the separation voltage is on after the sample injection, a signal is sent from the CE autosampler to the mass spectrometer for triggering the data acquisition. The separation current will decrease at the beginning during the separation due to the dynamic pH junction sample stacking. Then, the current recovers gradually.

## 5. Database Search of the Collected Raw Files with the TopPIC Software

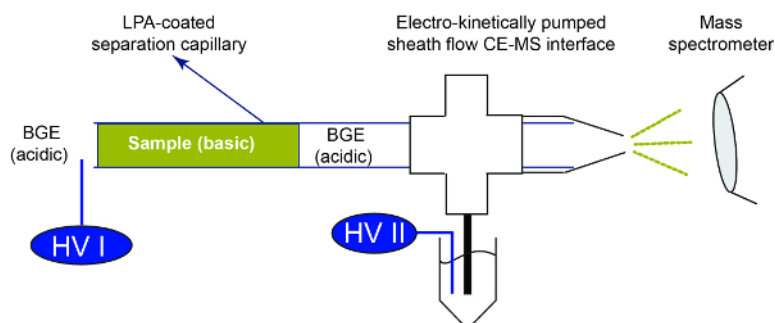
1. Transfer the .raw files, including the MS and MS/MS data, from the mass spectrometer computer to a computer dedicated to database searching.  
NOTE: These .raw files may be large and require a powerful computer in order to reach a fast database search.
2. Download ProteoWizard and TopPIC suite onto the computer dedicated to database searching.  
NOTE: ProteoWizard and TopPIC suite are open-source software and can be found online (ProteoWizard: <http://proteowizard.sourceforge.net>; TopPIC suite: <http://proteomics.informatics.iupui.edu/software/toppic/>). UniProt databases are used for database searches and can be found on the UniProt website.



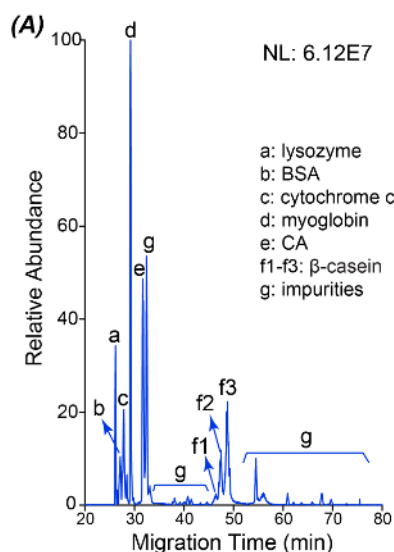
3. Convert the .raw files to .mzML files with msconvert, an MS file format conversion tool in ProteoWizard<sup>28</sup>. In the GUI of msconvert (MSConvertGUI.exe), click the button **Browse** and select the .raw file to be converted. Select mzML as the output format and keep all other options at their default values. Click the button **Start** at the bottom right of the GUI.
4. Convert the .mzML files to .msalign files with the TopFD tool in TopPIC suite<sup>29</sup>.
  1. In the GUI of TopFD (topfd\_gui.exe), click the button **File** and select the .mzML file created in the previous step. Keep the default values of the parameters and, then, click the button **Start** at the bottom right of the GUI.  
NOTE: TopFD converts precursor and fragment isotopic clusters to monoisotopic masses and identifies possible proteoform features in MS1 data by combining precursor isotopic clusters with similar monoisotopic masses and close migration times. Three files will be generated from TopFD, including an ms1.msalign file, an ms2.msalign file, and a .feature file. The ms1.msalign and ms2.msalign files contain deconvoluted monoisotopic masses of MS1 and MS/MS spectra, respectively. The .feature file contains possible proteoform features.
5. Perform a database search with TopPIC (version 1.1.3) in TopPIC suite<sup>30</sup>.
  1. In the TopPIC GUI (toppic\_gui.exe), click on **Database file** and select an appropriate database for searching.
  2. Click on **Spectrum file** and select the ms2.msalign file generated in step 5.4.1 as the input. Select the **MS1 feature file** and choose the .feature file generated in step 5.4.1.
  3. Set cysteine carbamidomethylation (C57) as a fixed modification due to the IAA treatment.  
NOTE: A text file can also be created with various fixed modifications by a text editor. Then, the text file can be selected by using the **File** button next to **Fixed modifications** in the TopPIC GUI.
  4. Select the **Decoy database** feature and, under cutoff settings, select **FDR** (false discovery rate) in the drop-down menu next to **Spectrum level**. Set the FDR to 0.01 at the spectrum level and 0.05 at the proteoform level.  
NOTE: TopPIC provides multiple options for filtering the results: spectrum-level FDR, proteoform-level FDR, or both. The FDR is evaluated based on the target-decoy approach<sup>31</sup>.
  5. Leave **Generating function** unselected and set the **Error tolerance (ppm)** to 15.
  6. Under **Advanced Parameters**, select 2 for the maximum number of mass shifts in the drop-down menu. Set the **Maximum mass shift (Da)** of unknown modifications to 500 Da. Leave all other parameters at their default values and click the button **Start** at the bottom right of the GUI.  
NOTE: The TopPIC software generates two resulting text files. The file with the suffix .OUTPUT\_TABLE contains the list of identified proteoform-spectrum matches (PrSMs), and the one with a suffix .FORM\_OUTPUT\_TABLE contains the list of identified proteoforms. For each PrSM, the software provides general information on the match, the observed fragmentation pattern, matched fragment ions, and detected mass shifts.

## Representative Results

**Figure 1** shows a diagram of the dynamic pH-junction-based CZE-ESI-MS system used in the experiment. A long plug of the sample in a basic buffer is injected into an LPA-coated separation capillary filled with an acidic BGE. After applying high voltages I and II, the analytes in the sample zone will be concentrated via the dynamic pH junction method. To evaluate the performance of the CZE-MS system, a standard protein mixture (cytochrome c, lysozyme,  $\beta$ -casein, myoglobin, CA, and BSA) is typically analyzed. The representative electropherogram for the standard protein mixture is shown in **Figure 2A**. The standard protein mixture is typically run at least in duplicate to evaluate the separation efficiency and the reproducibility of the system. The separation efficiency can be evaluated with the number of theoretical plates of some proteins, as shown in **Figure 2B**. The reproducibility can be evaluated by the relative standard deviations of protein intensity and migration time. **Figure 3A** shows a zoomed-in view of an electropherogram of the *E. coli* protein sample analyzed by the dynamic pH-junction-based CZE-MS/MS. The normalized-level (NL) protein intensity should be on the scale of  $10^8$  if  $1 \mu\text{g}$  of *E. coli* proteins are loaded for the analysis with a quadrupole ion trap mass spectrometer. A zoomed-in view of the electropherogram can be used to assess the separation window of the system. In this case, the separation window is 80 - 90 min. **Figure 3B** shows an example PrSM, including the general corresponding proteoform information, protein sequence, observed fragmentation pattern, and modifications. The very low E-Value ( $2.11\text{E-}48$ ) and Spectral FDR (0) suggest the high confidence of the proteoform ID. The high number of matched fragment ions (60) further indicates the high confidence of the ID. The observed fragmentation pattern shows that the fragmentation of the proteoform is highly efficient covering the termini and middle part of the proteoform. The N-terminal cleavage of three amino acids (MTM) is determined through the database search.



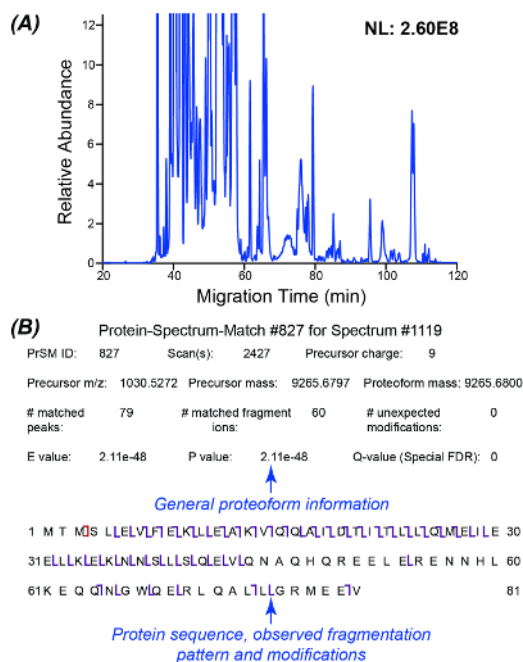
**Figure 1: Diagram of the dynamic pH-junction-based CZE-ESI-MS system.** The sample is dissolved in 50 mM ammonium bicarbonate, pH 8. The BGE is 5% or 10% (v/v) acetic acid, pH ~2.4 or ~2.2. A 1-m LPA-coated capillary is used for separation. An electro-kinetically pumped sheath flow interface is used to couple CZE to MS. A quadrupole ion trap mass spectrometer is used. The high voltage I (HV I) is provided by the power supply integrated into the CE autosampler for separation. High voltage II (HV II) is provided by a separate power supply for electrospray. The mass spectrometer is grounded. The distance between the orifice of the emitter and the entrance of the mass spectrometer is ~2 mm. The distance between the end of the etched capillary in the emitter and the orifice of the emitter is less than 500  $\mu\text{m}$ . The size of the orifice of the electrospray emitter is 20 - 40  $\mu\text{m}$ . [Please click here to view a larger version of this figure.](#)



**(B)**

Peak	Protein	MW (Da)	N
d	myoglobin	17000	90000
e	CA	29000	45000
f3	β-casein	24000	32000

**Figure 2: Data of a standard protein mixture analyzed by the dynamic pH-junction-based CZE-MS.** (A) This panel shows a base peak electropherogram. (B) This panel shows the number of theoretical plates (N) of three proteins. The sample injection volume was 500 nL. HV I was 30 kV for separation, and HV II was 2.2 kV for electrospray. The BGE was 5% (v/v) acetic acid, pH 2.4. The sample was in 50 mM ammonium bicarbonate (pH 8) and contains cytochrome c (0.01 mg/mL), lysozyme (0.01 mg/mL), β-casein (0.04 mg/mL), myoglobin (0.01 mg/mL), carbonic anhydrase (CA, 0.05 mg/mL), and bovine serum albumin (BSA, 0.1 mg/mL). No MS/MS spectra were acquired for the standard protein mixture sample. [Please click here to view a larger version of this figure.](#)



**Figure 3: Data of the *E.coli* proteome analyzed by the dynamic pH-junction-based CZE-MS/MS.** (A) This panel shows a zoomed-in view of the based peak electropherogram of the *E. coli* protein sample. (B) This panel shows an example PrSM identified through the TopPIC database search of the acquired MS/MS spectra. The general corresponding proteoform information, protein sequence, observed fragmentation pattern, and modifications are presented. The matched fragment ions from individual PrSM can also be viewed if needed. HV I was 20 kV for separation, and HV II was 2.2 kV for electrospray. The BGE was 10% (v/v) acetic acid, pH ~2.2. The sample was in 50 mM ammonium bicarbonate (pH 8) and the protein concentration was 2 mg/mL. The sample injection volume was 500 nL. A top8 DDA method was used to acquire the data. [Please click here to view a larger version of this figure.](#)

## Discussion

Here we provide a detailed protocol to use CZE-MS/MS for the high-resolution characterization of proteoforms in simple protein samples and for the large-scale identification of proteoforms in complex proteome samples. A diagram of the CZE-ESI-MS/MS system is shown in **Figure 1**. There are four critical steps in the protocol. First, the preparation of high-quality LPA coating on the inner wall of the separation capillary is extremely important. An LPA-coated separation capillary can reduce the EOF in the capillary, widen the separation window of CZE, and reduce protein adsorption on its inner wall<sup>19</sup>. The polymerization reaction for making the LPA coating is performed via filling the capillary with a mixture of acrylamide (monomer) and ammonium persulfate (polymerization initiator), followed by heating the capillary in a water bath to trigger the reaction<sup>22</sup>. The timing in the water bath is critical. Too short a reaction time will lead to an incomplete reaction and poor coating. We typically allow the reaction to go up to 50 minutes, allowing the reaction to proceed for a longer period for a better coating. With longer reaction periods, the capillary may become blocked and an HPLC pump might need to be used to push out the polymer inside of the capillary with water. One LPA-coated capillary can be continuously used for about one week, based on the literature data and our experience<sup>22</sup>.

The second critical step is coupling CZE to MS with the electro-kinetically pumped sheath flow CE-MS interface<sup>25,26</sup>. The distance between the end of the separation capillary in the ESI emitter and the emitter orifice affects the sensitivity of CZE-MS significantly, and a shorter distance produces a higher sensitivity<sup>25</sup>. The end of the separation capillary needs to be etched with HF, to reduce its outer diameter from 360  $\mu$ m to less than 100  $\mu$ m, allowing the end of the capillary to be pushed close to the emitter orifice. The distance between the emitter orifice and mass spectrometer entrance has been optimized to reach the best sensitivity and good stability<sup>26</sup>. We typically keep the distance around 2 mm.

The third critical step is the CZE-MS and MS/MS analysis with the dynamic pH-junction-based online sample stacking<sup>19</sup>. The pH of the sample buffer and the BGE need to be significantly different in order to ensure efficient sample stacking. The protein concentration in the sample needs to be appropriate. If the protein concentration is too high, the proteins can be precipitated in the capillary during sample stacking. The protein concentrations of the samples used for **Figures 2** and **3** can be considered as typical examples. The last critical step is the database search with TopPIC for proteoform IDs<sup>30</sup>. Multiple steps are required for file conversions before the database search with TopPIC can be performed. A proper FDR filter is necessary to ensure the quality of the identified proteoforms. We typically use a 1% spectrum-level FDR to filter the database search results. We note that the top-down proteomics initiative is a very good resource for top-down proteomics methods and data analysis software<sup>32,33</sup>.

We must note that some parts of the protocol may need to be modified slightly and some troubleshooting may be required. The time of polymerization in the water bath for the preparation of the LPA coating may vary at different labs. The time for HF etching of the separation capillary may need to be slightly modified due to a different temperature in different labs. When setting up the CE-MS interface, make sure the electrospray is stable before threading the separation capillary into the electrospray emitter. If no electrospray or an unstable electrospray is observed, check the interface to make sure there are no big bubbles in the emitter, in the T, or in the tubing that used to connect the sheath buffer vial and the T. In addition, the distance between the orifice of the emitter and the mass spectrometer entrance can affect the electrospray stability and is usually about 2 mm. The electrospray voltage can also influence the spray stability. For 20- to 40- $\mu$ m emitters, 2 - 2.2 kV is



usually good enough. As described in the previous paragraph, the protein concentration in the sample needs to be appropriate. If the protein concentration is too high, the proteins can be precipitated in the capillary during sample stacking. Pay attention to the current profile on the CE autosampler computer. If the current is zero at the very beginning of the CZE run, it means that a plug of air is injected into the capillary during the sample injection step. Make sure the volume of the sample in the vial is large enough. If the current is normal at the beginning and suddenly becomes zero during the run, it usually means that the protein concentration is too high.

CZE-MS/MS based on this protocol enables the high-resolution characterization of simple protein samples and the large-scale top-down proteomics of complex proteomes. As an example, the CZE-MS approached the high-resolution characterization of a standard protein mixture containing six proteins<sup>19</sup>. As shown in **Figure 2**, the CZE-MS clearly separated the six proteins with a high separation efficiency, revealed three forms of  $\beta$ -casein, and detected many impurities. As another example, single-shot CZE-MS/MS reproducibly yielded over 500 proteoform IDs and 190 protein IDs from the *E. coli* proteome, using only 1  $\mu$ g of *E. coli* proteins with a 1% spectrum-level FDR<sup>19</sup>. The CZE-MS/MS improved the number of proteoform IDs from a complex proteome by at least three times, compared with previous single-shot CZE-MS/MS studies. As shown in **Figure 3A**, the CZE-MS/MS simultaneously reached a 500-nL sample-loading volume and a 90-min separation window for the analysis of the *E. coli* proteome<sup>19</sup>. The TopPIC software can provide comprehensive information about the identified proteoforms (**Figure 3B**). The CZE-MS/MS system provides the proteomics community with a useful tool for large-scale and high-resolution top-down proteomics.

CZE-MS/MS still has some limitations for deep top-down proteomics. Although we demonstrated that the CZE-MS/MS can reach over 500 proteoform IDs from the *E. coli* proteome in a single run, the number of proteoform IDs from single-shot CZE-MS/MS is only roughly 50% of that from state-of-the-art RPLC-MS/MS<sup>34,35</sup>. In this protocol, only HCD is used for protein fragmentation, leading to limited fragmentation coverages of identified proteoforms. Several improvements can be made to the CZE-MS/MS protocol to improve both the number and quality of the proteoform IDs from single-shot CZE-MS/MS. First, increasing the length of the separation capillary should provide a wider separation window, leading to more proteoform IDs. Second, using a mass spectrometer with a much higher resolving power, faster scan rate, and multiple fragmentation methods (e.g., HCD,<sup>36</sup> electron transfer dissociation [ETD],<sup>37</sup> and ultraviolet photodissociation [UVPD]<sup>38</sup>) will certainly improve both the number of proteoform IDs and the fragmentation coverages of identified proteoforms.

## Disclosures

The authors have nothing to disclose.

## Acknowledgements

The authors thank Heedeok Hong's group at the Department of Chemistry, Michigan State University, for kindly providing the *Escherichia coli* cells for the experiments. The authors thank the support from the National Institute of General Medical Sciences, the National Institutes of Health (NIH) through Grant R01GM118470 (to X. Liu) and Grant R01GM125991 (to L. Sun and X. Liu).

## References

1. Aebersold, R. *et al.* How many proteoforms are there? *Nature Chemical Biology*. **14** (3), 206-214 (2018).
2. Tran, J. C. *et al.* Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature*, **480** (7376), 254-258 (2011).
3. Catherman, A. D. *et al.* Large-scale Top-down Proteomics of the Human Proteome: Membrane Proteins, Mitochondria, and Senescence. *Molecular and Cellular Proteomics*. **12** (12), 3465-3473 (2013).
4. Cai, W. *et al.* Top-Down Proteomics of Large Proteins up to 223 kDa Enabled by Serial Size Exclusion Chromatography Strategy. *Analytical Chemistry*. **89** (10), 5467-5475 (2017).
5. Toby, T. K., Fornelli, L., Kelleher, N. L. Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annual Review of Analytical Chemistry*. **9** (1), 499-519 (2016).
6. Jorgenson, J. W., Lukacs, K. D. Capillary Zone Electrophoresis. *Science*. **222** (4621), 266-272 (1983).
7. Keithley, R. B. Capillary electrophoresis with three-color fluorescence detection for the analysis of glycosphingolipid metabolism. *Analyst*. **138** (1), 164-70 (2013).
8. Han, X. *et al.* In-Line Separation by Capillary Electrophoresis Prior to Analysis by Top-Down Mass Spectrometry Enables Sensitive Characterization of Protein Complexes. *Journal of Proteome Research*. **13** (12), 6078-6086 (2014).
9. 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 International Edition*. **52** (51), 13661-13664 (2013).
10. Choi, S. B., Lombard-Banek, C., Muñoz-Llanca, P., Manzini, M. C., Nemes, P. Enhanced Peptide Detection Toward Single-Neuron Proteomics by Reversed-Phase Fractionation Capillary Electrophoresis Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*. **29** (5), 913-922 (2018).
11. Sun, L., Knierman, M. D., Zhu, G., Dovichi, N. J. Fast Top-Down Intact Protein Characterization with Capillary Zone Electrophoresis-Electrospray Ionization Tandem Mass Spectrometry. *Analytical Chemistry*. **85** (12), 5989-5995 (2013).
12. Wang, Y., Fonslow, B. R., Wong, C. C., Nakorchevsky, A., Yates 3rd, J. R. Improving the comprehensiveness and sensitivity of sheathless capillary electrophoresis-tandem mass spectrometry for proteomic analysis. *Analytical Chemistry*. **84** (20), 8505-8513 (2012).
13. 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*. **87** (8), 4572-4577 (2015).
14. Sun, L. *et al.* Over 10,000 peptide identifications from the HeLa proteome by using single-shot capillary zone electrophoresis combined with tandem mass spectrometry. *Angewandte Chemie International Edition in English*. **53** (50), 13931-13933 (2014).
15. Aebersold, R., Morrison, H. D. Analysis of dilute peptide samples by capillary zone electrophoresis. *Journal of Chromatography*. **516** (1), 79-88 (1990).

16. Britz-McKibbin, P., Chen, D. D. Y. Selective Focusing of Catecholamines and Weakly Acidic Compounds by Capillary Electrophoresis Using a Dynamic pH Junction. *Analytical Chemistry*. **72** (6), 1242-1252 (2000).
17. Zhao, Y., Sun, L., Zhu, G., Dovichi, N. J. Coupling Capillary Zone Electrophoresis to a Q Exactive HF Mass Spectrometer for Top-down Proteomics: 580 Proteoform Identifications from Yeast. *Journal of Proteome Research*. **15** (10), 3679-3685 (2016).
18. 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*. **86** (13), 6331-6336 (2014).
19. Lubeckyj, R. A., McCool, E. N., Shen, X., Kou, Q., Liu, X., Sun, L. Single-Shot Top-Down Proteomics with Capillary Zone Electrophoresis-Electrospray Ionization-Tandem Mass Spectrometry for Identification of Nearly 600 *Escherichia coli* Proteoforms. *Analytical Chemistry*. **89** (22), 12059-12067 (2017).
20. Busnel, J. M. High capacity capillary electrophoresis-electrospray ionization mass spectrometry: coupling a porous sheathless interface with transient-isotachopheresis. *Analytical Chemistry*. **82** (22), 9476-9483 (2010).
21. Zhang, Z., Peuchen, E. H., Dovichi, N. J. Surface-Confined Aqueous Reversible Addition-Fragmentation Chain Transfer (SCRAFT) Polymerization Method for Preparation of Coated Capillary Leads to over 10 000 Peptides Identified from 25 ng HeLa Digest by Using Capillary Zone Electrophoresis-Tandem Mass Spectrometry. *Analytical Chemistry*. **89** (12), 6774-6780 (2017).
22. Zhu, G., Sun, L., Dovichi, N. J. Thermally-initiated free radical polymerization for reproducible production of stable linear polyacrylamide coated capillaries, and their application to proteomic analysis using capillary zone electrophoresis-mass spectrometry. *Talanta*. **146**, 839-843 (2016).
23. Smith, R. D., Barinaga, C. J., Udseth, H. R. Improved Electrospray Ionization Interface for Capillary Zone Electrophoresis-Mass Spectrometry. *Analytical Chemistry*. **60** (16), 1948-1952 (1988).
24. Moini, M. Simplifying CE-MS Operation. 2. Interfacing Low-Flow Separation Techniques to Mass Spectrometry Using a Porous Tip. *Analytical Chemistry*. **79** (11), 4241-4246 (2007).
25. Wojcik, R., Dada, O. O., Sadilek, M., Dovichi, N. J. Simplified capillary electrophoresis nanospray sheath-flow interface for high efficiency and sensitive peptide analysis. *Rapid Communications in Mass Spectrometry*. **24** (17), 2554-2560 (2010).
26. Sun, L., Zhu, G., Zhang, Z., 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. *Journal of Proteome Research*. **14** (5), 2312-2321 (2015).
27. McCool, E. N. *et al.* Deep Top-Down Proteomics Using Capillary Zone Electrophoresis-Tandem Mass Spectrometry: Identification of 5700 Proteoforms from the *Escherichia coli* Proteome. *Analytical Chemistry*. **90** (9), 5529-5533 (2018).
28. Kessner, D., Chambers, M., Burke, R., Agus, D., Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics*. **24** (21), 2534-2536 (2008).
29. *TopPIC Suite*. <http://proteomics.informatics.iupui.edu/software/topfd/>. June 8, 2017. (2017).
30. Kou, Q., Xun, L., Liu, X. TopPIC : a software tool for top-down mass spectrometry-based proteoform identification and characterization. *Bioinformatics*. **32** (22), 3495-3497 (2016).
31. Elias, J. E., Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods*. **4** (3), 207-214 (2007).
32. *Methods - Consortium for Top Down Proteomics*. <http://www.topdownproteomics.org/resources/methods/> (2018).
33. *Software - Consortium for Top Down Proteomics*. <http://www.topdownproteomics.org/resources/software/> (2018).
34. Ansong, C. *et al.* Top-down proteomics reveals a unique protein S-thiolation switch in *Salmonella* Typhimurium in response to infection-like conditions. *Proceedings of the National Academy of Sciences of the United States of America*. **110** (25), 10153-10158 (2013).
35. Shen, Y. *et al.* High-resolution ultrahigh-pressure long column reversed-phase liquid chromatography for top-down proteomics *Journal of Chromatography A*. **1498**, 99-110 (2017).
36. Olsen, J. V., Macek, B., Lange, O., Makarov, A., Horning, S., Mann, M. Higher-energy C-trap dissociation for peptide modification analysis. *Nature Methods*. **4** (9), 709-712 (2007).
37. Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., Hunt, D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*. **101** (26), 9528-9533 (2004).
38. Shaw, J. B. *et al.* Complete Protein Characterization Using Top-Down Mass Spectrometry and Ultraviolet Photodissociation. *Journal of the American Chemical Society*. **135** (34), 12646-12651 (2013).