Bottom-Up Proteomics of *Escherichia coli* Using Dynamic pH Junction Preconcentration and Capillary Zone Electrophoresis-Electrospray Ionization-Tandem Mass Spectrometry

Guijie Zhu, Liangliang Sun, Xiaojing Yan, and Norman J. Dovichi*

Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States

**Supporting Information**

**ABSTRACT:** We report the use of the dynamic pH junction based capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry (CZE-ESI-MS/MS) for bottom-up proteomics with an electrokinetically pumped sheath-flow nanospray capillary electrophoresis-mass spectrometry (CE-MS) interface and both LTQ-XL and LTQ-Orbitrap-Velos mass spectrometers. Conventional injection of 20 nL of a 1 mg/mL BSA digest identified 37 peptides and produced 66% sequence coverage. In contrast, pH junction injection of 130 nL (or larger) of a 0.05 mg/mL BSA digest identified 40 peptides and produced 70% coverage using a pH 6.5 sample buffer and the LTQ. A 20 nL conventional injection of a 1 mg/mL *Escherichia coli* digest identified 508 peptides and 199 proteins with the LTQ. A 400 nL pH junction injection of a 0.1 mg/mL *E. coli* digest identified 527 peptides and 179 proteins with the LTQ. Triplicate technical replicates of a 0.01 mg/mL BSA digest produced narrower peaks and significant concentration for all but the most acidic components in the sample. Compared with the conventional stacking method, the pH junction method can generate comparable performance for small injection volume (20 nL) and significantly better concentration performance for a large injection volume (200 nL). We also applied the pH junction to three intact standard proteins and observed a >10X increase in peak intensity compared to conventional injection.

Capillary zone electrophoresis is an attractive alternative to reversed-phase liquid chromatography for proteomics research. It provides fast and efficient separations in a very simple flow path. However, application of capillary zone electrophoresis to proteomics research has been discouraged by a fundamental limitation of the technique. When the sample is prepared in the same matrix as the running buffer, the injection volume must be <1% of the capillary volume to minimize band broadening. In this conventional injection mode, the sample volume is often less than 10 nL, which is 2 orders of magnitude smaller than the injection volume commonly used in reversed-phase liquid chromatography. This small injection volume results in poor concentration detection limits and relatively small numbers of protein and peptide identifications.

A number of large injection volume methods have been developed for capillary zone electrophoresis, such as solid-phase microextraction, transient capillary isotachophoresis, and field enhanced sample injection. The dynamic pH junction is a particularly intriguing example of a sample preconcentration technique that facilitates use of large injection volumes. In one mode of the pH junction, amphoretic analyte is prepared in a basic buffer and is negatively charged. The sample is injected by pressure into a capillary filled with an acidic buffer. Application of an electric field will cause the negatively charged analyte to migrate toward the positive electrode at the injection end of the capillary. Analyte is neutralized and is focused to a narrow band when it reaches the acidic separation buffer at the front of the sample plug. Focusing continues until the separation buffer has neutralized the sample buffer. Very large sample volumes can be used with little loss of separation efficiency. Alternatively, the sample can be prepared in an acidic buffer and separated in a basic buffer.

Although the dynamic pH junction technique has a long history, to our knowledge, there have been no reports that employ this technique for bottom-up proteomics analysis. In this manuscript, we coupled the pH junction with capillary zone electrophoresis-tandem mass spectrometry (CZE-MS/MS) for analysis of a tryptic digest of both a standard protein and of the *E. coli* proteome. We also present preliminary results for the concentration of intact standard proteins, mimicking a top-down analysis.

**MATERIALS AND METHOD**

Reagents. Bovine pancreas TPCK-treated trypsin, bovine serum albumin (BSA), cytochrome C (Cyto.c), myoglobin

Received: January 31, 2014
Accepted: May 22, 2014
Published: May 22, 2014

© 2014 American Chemical Society
(Myo), beta casein, urea, ammonium bicarbonate (NH$_4$HCO$_3$), ammonium acetate (NH$_4$C$_2$H$_3$O$_2$), dithiothreitol (DTT), and iodoacetamide (IAA) were purchased from Sigma-Aldrich (St. Louis, MO). Acetonitrile (ACN) and formic acid (FA) were purchased from Fisher Scientific (Pittsburgh, PA). Methanol was purchased from Honeywell Burdick & Jackson (Wicklow, Ireland). Water was deionized by a NanoPure system from Thermo Scientific (Marietta, OH). LPA-coated fused capillary (50 μm i.d./150 μm o.d.) were purchased from Polymicro Technologies (Phoenix, AZ). Complete, mini protease inhibitor cocktail (provided in EASYpacks) was purchased from Roche (Indianapolis, IN).

Preparation of Sample. A 0.5 mg/mL solution of BSA in 100 mM NH$_4$HCO$_3$ (pH 8.0) containing 8 M urea was denatured at 37 °C for 30 min, followed by standard reduction and alkylation with DTT and IAA. Digestion was performed for 12 h at 37 °C with trypsin at a trypsin/protein ratio of 1/30 (w/w).

The E. coli digests were prepared as reported previously. Briefly, after cell culture and cell lysis, 900 μg of E. coli homogenate was purified by acetone precipitation and then dried at room temperature. The dried sample was dissolved in 300 μL of 100 mM NH$_4$HCO$_3$ (pH 8.5) containing 8 M urea. After denaturing at 37 °C for 60 min, the proteins were reduced by 8 mM of DTT (final concentration) at 60 °C for 1 h and alkylated by 20 mM of IAA (final concentration) at room temperature for 30 min in the dark, followed by dilution to 1.5 mL with 100 mM NH$_4$HCO$_3$ (pH 8.5) to reduce the urea concentration less than 2 M. An aliquot of 120 μg of treated proteins was digested by trypsin at a trypsin/protein ratio of 1/30 (w/w) for 4 h at 37 °C.

After trypsin digestion, the BSA and E. coli digests were acidified by FA (0.5% (v/v) final concentration) to terminate the reaction. Then the digests were desalted with a tC18 SepPak column (Waters, Milford, MA), followed by lyophilization with a vacuum concentrator (Thermo Fisher Scientific, Marietta, OH). The dried samples were stored at −20 °C before use.

Two mixtures of standard proteins containing cytochrome c (0.1 mg/mL), myoglobin (0.1 mg/mL), and beta casein (0.5 mg/mL) dissolved in 0.1% (v/v) FA buffer and 10 mM ammonium acetate (pH ~6.5) buffer were prepared for conventional and dynamic pH junction based CZE-MS/MS analysis.

CZE-ESI-MS/MS Analysis. The CZE-ESI-MS/MS system was similar to that reported previously. High voltages were supplied by two Spellman CZE 1000R power supplies. The LPA-coated separation capillary (60 cm, 50 μm i.d./150 μm o.d.) was coupled to an LTQ XL instrument (Thermo Fisher Scientific) with an electrokinetically pumped sheath-flow nanospray interface. The emitter was pulled with a Sutter pipet puller to ~10 μm i.d. Voltage programming was controlled by LabView software. For separation, 300 V/cm was applied for BSA digests and E. coli digests analysis, and 400 V/cm was applied for intact protein analysis. Electrospray was produced at 1.2 kV. Sample injection was performed by pressure. The separation buffer for CZE is 0.1% (v/v) FA, and the electrospray sheath liquid is 10% (v/v) methanol and 0.1% (v/v) FA.

For conventional injection, the sample is prepared in the 0.1% FA separation buffer. For pH junction injection, the sample is prepared in a 10 mM ammonium acetate (pH ~6.5) solution. For conventional stacking injection, the sample is prepared in 50% (v/v) ACN with 0.05% (v/v) FA.

All experiments except the comparison of the pH junction with conventional stacking were performed using an LTQ-XL mass spectrometer (Thermo Fisher Scientific). For protein digests analysis with the LTQ, data dependent acquisition was applied. Full MS scans were acquired over the 395–1900 m/z range. The 10 most intense peaks were selected for fragmentation in the ion trap with normalized collision energy of 35%, activation q = 0.25, activation time of 30 ms, and one microscan. Peaks selected for fragmentation more than once within 45 s were excluded from selection for 60 s. For intact protein analysis, no tandem spectra were acquired, and full MS scans were acquired over the 600–2000 m/z range. Two microscans were applied.

A comparison of the pH junction and conventional stacking was performed on an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific). Full MS scans were acquired in the Orbitrap mass analyzer over the m/z 395−1800 range with resolution of 30 000 (at 400 m/z). The 10 most intense peaks with a charge state ≥2 were fragmented in the higher-energy collisional dissociation (HCD) collision cell and analyzed by the Orbitrap mass analyzer with a resolution of 7 500. One microscan was used. Normalized collision energy was set at 40%. For MS and MS/MS spectra acquisition, the maximum ion inject time was set as 500 and 250 ms, respectively. The precursor isolation width was 2 Da. The target values for MS and MS/MS were set at 1.00 × 10⁵ and 5.00 × 10⁴, respectively. Dynamic exclusion was applied for the experiments. Peaks selected for fragmentation more than once within 25 s were excluded from selection for 25 s.

Data Analysis. Database searching of acquired raw files was performed in Proteome Discoverer 1.3 with the Mascot search engine (version 2.2) against ipi.bovin.v3.68.fasta (for BSA) and the NCBI E. coli DH1 database (for E. coli). Database searching of the reversed database was also performed in order to evaluate the false discovery rate. The database searching parameters included full trypptic digestion and allowed up to two missed cleavages, precursor mass tolerance 2 Da, and fragment mass tolerance 1 Da. Carbamidomethylation (C) was set as fixed modifications. Oxidation (M) was set as variable modification.

Percolator software (version 1.17) integrated in Proteome Discoverer 1.3 was used to evaluate the database search results. Peptide confidence value as high was used to filter the peptide identifications, and the corresponding false discovery rate at the peptide level was less than 1%. On the protein level, the protein grouping was enabled.

Data were imported into Matlab. Selected ion electropherograms were generated with a 10 ppm mass window. An unsupervised least-squares routine (“fit”) was used to fit a Gaussian function to selected ion electropherograms. The peak area was estimated as the product of peak amplitude and width.

RESULTS AND DISCUSSION

Most publications describing the pH junction have used uncoated capillaries and a high concentration separation buffer (i.e., 1 M formic acid) to minimize analyte adsorption onto the inner wall of capillary and to reduce the electroosmotic flow for high-resolution separation. In contrast, we used a linear polyacrylamide coated capillary to eliminate electroosmotic flow and 0.1% formic acid as the separation buffer. For conventional injection, the sample was prepared in
produced a separation window from 10 to 25 min (Figure 1A). We then performed a dynamic pH junction injection of a 0.05 mg/mL BSA digest prepared in 10 mM ammonium acetate (pH ∼6.5). Injection volumes from 40 to 600 nL were used, Figure 1B–F. Larger injection volumes resulted in a modest compression of the separation window and an apparent loss of later migrating components. Nevertheless, good electrophoretic profiles were obtained, even with an injection volume of 600 nL, which is half the capillary volume. A plot of injection volume vs normalization level (NL) was linear up to about 250 nL and then began to saturate, S-Figure 1 in the Supporting Information. Injection of 260 nL of the dilute 0.05 mg/mL BSA digest (Figure 1D) generated a base peak intensity that was comparable to the intensity produced by conventional injection of 20 nL of 1 mg/mL BSA digests (Figure 1A).

We also generated extracted ion electropherograms of three BSA peptides from conventional injection and using the dynamic pH junction, S-Figure 2 in the Supporting Information. The peptide intensity from the dynamic pH junction with injection of 260 nL of 0.05 mg/mL BSA digests is comparable to that from conventional injection of 20 nL of 1 mg/mL BSA digest with superior separation efficiency. Significant enrichment was obtained for the three peptides that have dramatically different isoelectric points (pI 4.5, 6, and 10), which demonstrates that this technique does not have strong bias for either acidic or basic peptides.

We obtained 66% sequence coverage and identified 37 peptides with conventional injection of 20 nL of 1 mg/mL BSA digests. We obtained 70% coverage and identified over 40 peptides using the pH junction with 130 nL or larger injection volume of the 0.05 mg/mL BSA digest. The ion score from the Mascot database search was also analyzed, S-Figure 5 in the Supporting Information. Use of the pH junction to inject a 130 nL or larger volume of 0.05 mg/mL BSA digest generated mean and median ion scores that are comparable with those from conventional injection of 20 nL of a 1 mg/mL BSA digest.

**Application of the Dynamic pH Junction Preconcentration Based CZE-MS/MS for E. coli Digests Analysis.** We compared conventional and dynamic pH junction injections for bottom-up analysis of E. coli digests, Figure 2. For conventional injection, 1 and 0.1 mg/mL E. coli digest samples were prepared in 0.1% (v/v) FA; 20 nL injection volumes were analyzed by the CZE-MS/MS system, Figure 2A,B. As expected, the base peak intensity (normalized level, NL) of the 0.1 mg/mL sample is about 1 order of magnitude lower than that of the 1 mg/mL sample.

For the pH junction, 0.1 and 0.01 mg/mL E. coli digests were prepared in 10 mM ammonium acetate (pH ∼6.5); 400 nL injection volumes (1/3 of the capillary volume) were analyzed by the CZE-MS/MS system (Figure 2C,D). As with the BSA digest, the base peak intensity generated by the dynamic pH junction is about 20 times higher than that of the same sample concentration with conventional injection and the pH junction generated a slightly narrower separation window. Also as with the BSA results, the use of a 400-nL pH junction injection generated more efficient peaks compared to a 20-nL conventional injection, S-Figure 6 in the Supporting Information.

We also analyzed the database search results from conventional and pH junction injection of the E. coli samples. Conventional injection of a 1 mg/mL sample identified 508 peptides and 199 proteins. Dynamic pH junction of a 0.1 mg/mL sample identified 527 peptides and 179 proteins. These single-shot peptide and protein identifications are much smaller than our earlier work, which employed the much more powerful LTQ-Orbitrap Velos or Q-Exactive mass spectrometers.17,22

The pH junction was not quite as successful when applied to a very dilute E. coli digest. Conventional injection of a 0.1 mg/mL sample yielded 366 peptides and 153 proteins. We performed triplicate analysis of 0.01 mg/mL E. coli digest sample with the dynamic pH junction method (injection volume, 400 nL; injection mass, 4 ng), Figure 3. Reasonably reproducible separations were obtained, and the relative standard deviation (RSD) of base peak intensity was about 15%. Because the pH junction employs different sample matrix and running buffers with a long injection length, the capillary
needs to be flushed thoroughly with running buffer between runs to obtain highly reproducible separations.

The pH junction identified 288 ± 9 peptides and 121 ± 5 proteins from this dilute sample, which represents a 20% decrease in the number of identifications compared to conventional injection of the 10X more concentrated sample.

Comparison of Dynamic pH Junction and Conventional Stacking. There are several online preconcentration methods for CE, including conventional stacking, transient isotachophoresis, solid phase extraction, and dynamic pH junction. Of these methods, the pH junction and conventional stacking are the simplest. They require no change in instrument setup, capillary preparation, and CE operation. The only change is using different sample matrix. The pH junction uses a high pH sample buffer while stacking uses a low ionic strength sample buffer.

We compared the preconcentration performance of the dynamic pH junction with the conventional stacking method using a 0.1 mg/mL E. coli digests sample. In this experiment, the LTQ Orbitrap Velos was used for peptide identification, Figure 3 in the Supporting Information.

For a 1 cm injection length (20 nL, 2 ng sample amount), both pH junction and stacking generate a ~2-fold increase in base peak signal compared with normal injection using the same injection length. The number of peptide and protein identifications is comparable for the two methods. Large volume stacking identified 140 proteins and 456 peptides. The pH junction identified 132 proteins and 598 peptides in duplicate runs.

In contrast, the pH junction produced superior results for a larger injection volume. A 10 cm (200 nL, 20 ng) stacking injection identified only 89 proteins and 363 peptides. In contrast, the pH junction with a 10 cm injection length resulted in identification of 185 proteins and 828 peptides, which is comparable with the data from our previous work. However, the E. coli digest concentration used in this work is 10 times lower than the previous work (0.1 mg/mL vs 1 mg/mL).

The results clearly demonstrate that the pH junction is valuable and efficient for low concentration complex proteome digest analysis, which is critically important when only trace amounts of samples are available. In addition, CZE usually has good mass detection limit and worse concentration detection limit due to small sample injection volume. Here, we generated...
enough bottom-up proteomic data to support that the pH junction method is a simple and efficient method to significantly improve the concentration detection limit of CZE for proteomic samples.

**Injection Bias of Dynamic pH Junction.** We compared the performance of the pH junction and normal injection using a 1 cm injection (20 nL) of a 0.1 mg/mL *E. coli* digest, Figure 4.

![Figure 4](image)

**Figure 4.** Base peak electropherograms for a 20 nL (1 cm) injection of a 0.1 mg/mL *E. coli* tryptic digest. The electropherogram generated with a pH junction injection is shown in blue (top), and the electropherogram generated using normal injection is inverted and shown in green (bottom).

Although the injection amounts were the same (2 ng) for both experiments, the pH junction generated more intense base peak electropherograms. We then identified 307 peptides that generated reasonably intense (>1 × 10^5) peaks in both the normal and pH junction data sets. Unsupervised nonlinear least-squares regression was used to fit a Gaussian function to each of the peaks.

The pH junction had little effect on either migration time or peak area. The peptide migration times for the two injection methods was strongly correlated (slope = 1.1 ± 0.2, r = 0.94), S-Figure 8 in the Supporting Information. A plot of the peak area for the pH junction and normal injection was also linear with a slope very near 1 (slope = 0.93 ± 0.03, r = 0.90), S-Figure 9 in the Supporting Information.

The pH junction resulted in a significant decrease in peak width for all but the latest migrating components, Figure 5. The focusing effect was most pronounced for components with a migration time between 12.5 and 20 min. Components with longer migration time, which tend to be strongly acidic, generated wider peaks than the normal injection. As a result, components with migration time longer than ~20 min have lower amplitude in the pH junction experiment. These components are sufficiently acidic to require higher pH sample buffer to undergo pH junction focusing.

**Dynamic pH Junction Preconcentration Based CZE-MS System for Intact Protein Concentration, Separation, and Detection.** We finally applied the dynamic pH junction based CZE-ESI-MS system for separation of intact proteins, Figure 6. Three standard proteins (cytochrome c, myoglobin, and beta casein) with a pI from 5 to 10 were chosen as the sample. When the sample was dissolved in 0.1% (v/v) FA for conventional injection (20 nL), after CZE-MS analysis, cytochrome c and beta casein peaks were observed, but no significant myoglobin peak was produced, Figure 6A. When the sample with same protein concentration was dissolved in 10 mM ammonium acetate (pH ~6.5) and 20 nL of the sample was injected for dynamic pH junction based CZE-MS analysis,
the intensity of peaks corresponding to cytochrome c and casein was 3 times higher than that from conventional injection, and the peak of myoglobin is dramatically increased in amplitude, Figure 6B. When 40 nL of protein sample in 10 mM ammonium acetate (pH ~6.5) was injected for analysis, the protein intensity was around 1 order of magnitude higher than the conventional injection, Figure 6C. These results suggest that the dynamic pH junction based CZE-MS system may be of value in top-down proteomic analysis.

ASSOCIATED CONTENT

* Supporting Information
Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
* E-mail: ndovichi@nd.edu. Phone: +1 574 631 2778.

Notes
The authors declare the following competing financial interest(s): We have just licensed the electrospray interface used in this manuscript to CMP, Inc. and have a financial interest in its commercialization.

ACKNOWLEDGMENTS

We thank Dr. William Boggess in the Notre Dame Mass Spectrometry and Proteomics Facility for his help with this project. This work was funded by the National Institutes of Health (Grant R01GM096767).

REFERENCES