Ultrasensitive Analysis Hot Paper

DOI: 10.1002/ange.201308139

Ultrasensitive and Fast Bottom-up Analysis of Femtogram Amounts of Complex Proteome Digests**

*Liangliang Sun, Guijie Zhu, Yimeng Zhao, Xiaojing Yan, Si Mou, and Norman J. Dovichi**

Bottom-up proteomics is widely used for qualitative and quantitative characterization of complex biological samples.[1,2] With microgram amounts of material, it is possible to identify more than 10000 proteins from mammalian cell lysates and over 2500 proteins from prokaryote lysates.[3,4] The performance of bottom-up proteomics degrades rapidly for mass-limited samples, such as laser capture microdissected tissues (pure cell populations isolated from a heterogeneous tissue section), circulating tumor cells, single embryos, and single somatic cells. There have been a handful of reports on bottom-up proteomics of nanogram samples by capillary liquid chromatography–electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS). Mann et al. identified 2000 proteins from single pancreatic islets with a protein content of several hundred nanograms.[5] Karger and co-workers identified 566 proteins from 50 ng of *Methanosarcina acetivorans* digest[6] and 163 proteins from approximately 2.5 ng of the tryptic digest of a cervical cancer cell line.[7] Smith et al. detected 870 proteins with an accurate mass and time tags (AMTs) strategy[8] from low nanogram amounts of the digest of *Deinococcus radiodurans*.[9] They also reported the detection of the three most abundant proteins in a 0.5 pg sample with the AMTs method, and reported a detection limit of approximately 10 zmole for one peptide in a digest of bovine serum albumin.[9] Our group used a Q-Exactive mass spectrometer with higher-energy collisional dissociation[10] to identify approximately 100 protein groups from 1 ng of a digest of the RAW264.7 macrophage cell line.[11] All of these analyses required at least one hour of instrument time. Herein, we report an ultrasensitive and fast capillary zone electrophoresis (CZE)-ESI-MS/MS system that is based on an improved electrokinetically pumped sheath–flow interface. We demonstrate that this system is suitable for the rapid bottom-up analysis of femtogram amounts of the *E. coli* protein digest.

CZE-ESI-MS/MS is an attractive method for bottom-up proteomics,[12] and this approach consistently outperforms LC-MS/MS for low nanogram samples.[13–16] The improved performance of CZE for small sample amounts is presumably due to its very simple design, which leads to reduced sample loss on injectors and fittings. Since the pioneering work of Smith and co-workers,[17] electrospray interfaces have been developed for capillary electrophoresis.[18] Two recently developed interfaces should be mentioned: Moini developed a sheathless interface that is based on a very thin porous capillary tip.[19] We, in turn, developed another interface based on an electrokinetically pumped sheath–flow interface (Figure 1 a).[20] Our interface has several advantages, including reduced sample dilution, which is due to a very low sheath flow rate, elimination of mechanical pumps, use of a wide range of separation buffers, and stable operation in the nanospray regime. We recently coupled CZE to a triple-quadrupole mass spectrometer with this interface for quantification of Leu-enkephalin in a complex mixture using multiple reaction monitoring, and we obtained a peptide detection limit of 335 zmole,[21] which suggests that the system should be suitable for ultrasensitive analysis.

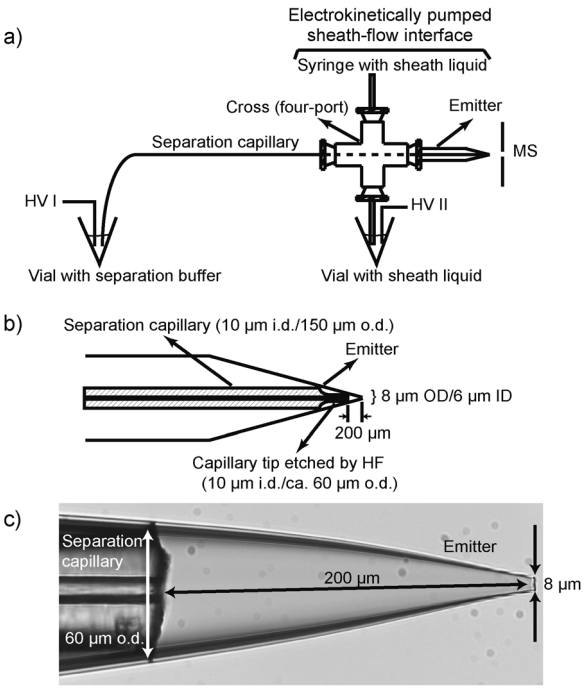

Figure 1. a) CZE-ESI-MS/MS system; b) schematic representation of the etched capillary in the electrospray emitter; c) micrograph of the etched capillary in the emitter. i.d. = inner diameter, o.d. = outer diameter.

[*] Dr. L. Sun, Dr. G. Zhu, Y. Zhao, X. Yan, S. Mou, Prof. Dr. N. J. Dovichi
Department of Chemistry and Biochemistry
University of Notre Dame
Notre Dame, IN 46556 (USA)
E-mail: ndovichi@nd.edu

[**] 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 (R01GM096767).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201308139.

A COMSOL model of the electrokinetically pumped sheath–flow interface predicted, and experiments verified, that sensitivity increases as the distal end of the capillary is brought closer to the emitter orifice.[20] Typical distances between capillary tip and orifice are about 1 mm; they are limited by the outer diameter of the separation capillary that butts against the conical emitter wall. In this work, we etched a few millimeters of the outside of the separation capillary tip with hydrofluoric acid to reduce its outer diameter from approximately 150 μm to approximately 60 μm. This simple step allows us to place the capillary end much closer to the emitter orifice (ca. 200 μm; Figure 1b,c), which results in a dramatic improvement in the sensitivity of the system. We used uncoated fused silica capillaries (32 cm and 40 cm, 10 μm i.d./150 μm o.d.; i.d. = inner diameter, o.d. = outer diameter) for electrophoresis, and a Q-Exactive mass spectrometer for peptide identification (for experimental details, see the Supporting Information).

We first evaluated the effect of the separation voltage for the analysis of 28 pg amounts of *E. coli* digests. Separations were performed at 15 kV (500 V cm^{-1}) and 10 kV (300 V cm^{-1}) in a 32 cm long capillary (Supporting Information, Figure S1). The 10 kV potential produced a wider separation window, which resulted in more protein (129 ± 18 vs. 88 ± 14) and peptide (375 ± 27 vs. 246 ± 19) identifications than at 15 kV. For the following experiments, an electric field of 300 V cm^{-1} was used.

Next, we evaluated the reproducibility of an analysis of 16 pg of *E. coli* protein digest with a 40 cm capillary with our CZE-ESI-MS/MS system. We identified 105 ± 17 proteins and 256 ± 9 peptides by a triplicate bottom-up analysis of tandem mass spectra. The state-of-the-art detection limits for tandem mass spectrometry of complex protein digests is approximately 100 protein identifications at the 1 ng level.[7,9,11,13,16] The use of our system leads to a similar number of protein identifications from a sample that is two orders of magnitude lower in weight.

The separations were reproducible and efficient. The signals from 50 peptides were summed to produce extracted ion electropherograms (Figure 2). The average relative standard deviation of the migration time of 154 peptides was 0.7% (Figure S2). The electrophoretic peaks were quite sharp, with an average width, which is defined as the standard deviation of the Gaussian function used to fit the peaks, of 0.7 s (full width at half height: 1.6 s; Figure S3). We consistently obtained an average of over 300 000 theoretical plates for the peptide separations (Figure S4). Peak intensities were also consistent between runs (Figure S5). Separations were complete in less than 10 min, which is an improvement in analysis time of one order of magnitude compared to state-of-the-art systems for highly sensitive bottom-up proteomics of complex proteomes.

We next determined the relationship between the number of identifications by tandem mass spectrometry and the loaded amounts of *E. coli* digests (Figure 3). In duplicate 400 fg loadings, nine peptides corresponding to 4 ± 1 proteins were confidently identified after manual evaluation of tandem mass spectra (Figure S6). The most abundant protein in *E. coli*, the elongation factor Tu, makes up approximately

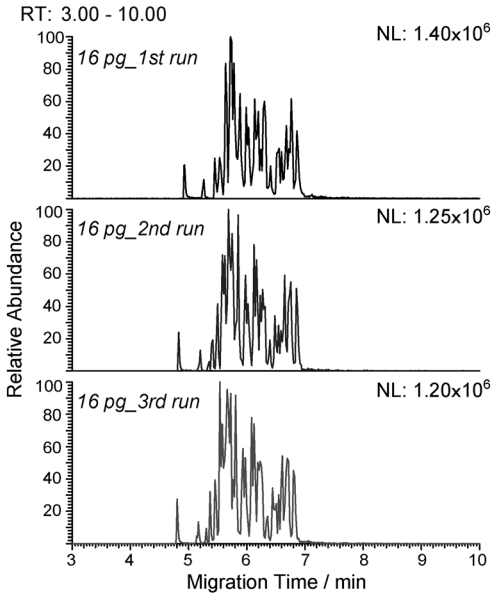

Figure 2. Extracted ion electropherograms of 50 peptides of high intensity based on tandem spectra of *E. coli* digests (16 pg) analyzed by CZE-ESI-MS/MS in triplicate. The mass tolerance for extraction was 2 ppm. NL = normalized level, RT = retention time.

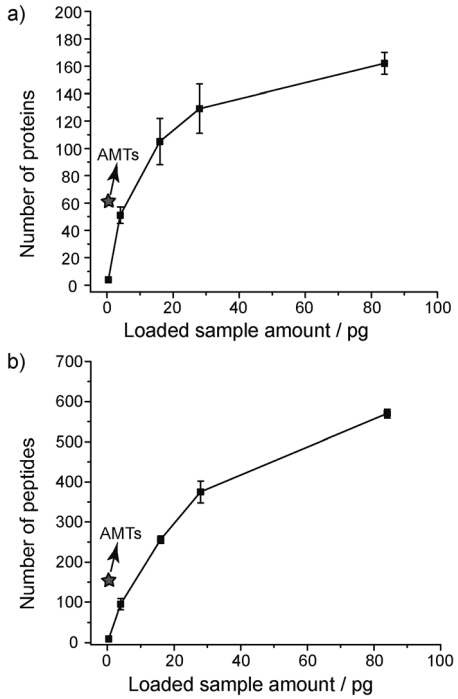

Figure 3. Relationship between the amount of *E. coli* digest and protein (a) and peptide (b) identification by the analysis of tandem spectra (■). Each sample was analyzed in duplicate or triplicate. Identifications based on AMTs of *E. coli* digests (400 fg) are also given (★). The error bars correspond to the standard deviations of the mean.

1% of the total protein mass.[22] An amount of 400 fg of an *E. coli* digest will contain 4 fg of this protein, and the minimum protein amount required for identification by tandem mass spectrometry is less than 4 fg, which is an

improvement of two orders of magnitude compared to the state of art.[9] When the sample amount was increased to 84 pg, the number of protein and peptide identifications increased to 162 ± 8 and 570 ± 11, respectively (Figure 3).

We also applied the AMTs approach developed by Smith and co-workers using the 16 pg *E. coli* data as the database. Over 60 proteins and 150 peptides were identified from the 400 fg *E. coli* digests with a mass tolerance of 3 ppm, a migration-time tolerance of 0.3 min (without alignment), and at least two detected isotopic peaks for each peptide (Figure 3; for the extracted ion electropherograms, see Figure S7). This result corresponds to a 20-fold improvement in the number of protein identifications relative to state-of-the-art AMTs-based sub-picogram proteome analysis.[9]

We finally estimated the peptide detection limit from the 400 fg *E. coli* data. We manually extracted electropherograms for three peptides from the elongation factor Tu, which were identified by MS/MS spectra with a mass tolerance of 1 ppm. The signal-to-noise ratios were obtained with Xcalibur software (Thermo Fisher Scientific), using the noise region from 0.3 min to 2.3 min after the peak (Figure S8). Based on the amount of elongation factor Tu present in the sample (4 fg) and its molecular weight (ca. 43 kDa), approximately 100 zmole of these peptides were taken for analysis. These peptides generated signal-to-noise ratios (S/N) of 270–290; the mass detection limit ($S/N = 3$) is approximately 1 zmole (ca. 600 molecules), which is an improvement of one order of magnitude compared with state-of-the-art MS-based peptide detection.[9]

There are several possible reasons for the high sensitivity of the CZE-MS system. First, the peptides only take around 0.2 s or less to migrate from the end of the capillary to the end of the spray emitter (approximate calculation based on our previous work[20]), which dramatically reduces the sample diffusion in the spray emitter and generates more intense peptide signals, which in turn results in improved peptide detection limits. Second, we can assume that the electroosmotic flow rate in the separation capillary and in the spray emitter is approximately the same because of a similar buffer and applied voltage (ca. $300\ V\,cm^{-1}$); the total flow rate for the spray is about $20\ nL\,min^{-1}$, which should lead to a high ionization efficiency, thus resulting in high sensitivity. Third, we employed capillaries with a quite narrow inner diameter, which reduces sample flow rate and generates very efficient separations.

We also note that the number of protein identifications obtained in this work is approximately 200 proteins because of the relatively short peptide separation window, restricting the number of acquired tandem spectra, which limits the detection of proteins that are of relatively low abundance in biological samples. One way to improve protein identification with the CZE-MS system is to perform online/offline peptide pre-fractionation before CZE-MS analysis or to employ coated capillaries to reduce electroosmosis and to increase the separation window.[16,23]

In summary, an ultrasensitive and high-throughput CZE-ESI-MS/MS system has been developed for femtogram proteomics. The results are an improvement of one- to two-orders of magnitude in the amount of material required for protein identification by tandem mass spectra, in the number of proteins identified by accurate mass and time tags from sub-picogram amounts of a complex protein digest, in peptide mass detection limit, and in analysis time. Faster and more sensitive mass spectrometers will enable further improvements. This technology may, for example, be applied to the analysis of single cells. To date, methods of the highest possible sensitivity for single-cell protein analysis have employed laser-induced fluorescence detection;[24] although it is a sensitive method, fluorescence inherently generates a signal with a low information content that provides only rudimentary information on protein identity. The development of CZE-mass spectrometry systems with detection limits of 1 zmol opens the door to single-cell protein analysis with confident identification of proteins with relatively high abundance.

Received: September 16, 2013
Published online: October 31, 2013

Keywords: electrophoresis · electrospray · mass spectrometry · proteomics · ultrasensitive analysis

[1] M. Gstaiger, R. Aebersold, *Nat. Rev. Genet.* **2009**, *10*, 617 – 627.
[2] Y. Zhang, B. R. Fonslow, B. Shan, M.-C. Baek, J. R. Yates III, *Chem. Rev.* **2013**, *113*, 2343 – 2394.
[3] T. Geiger, A. Wehner, C. Schaab, J. Cox, M. Mann, *Mol. Cell. Proteomics* **2012**, DOI: 10.1074/mcp.M111.014050.
[4] M. Iwasaki, S. Miwa, T. Ikegami, M. Tomita, N. Tanaka, Y. Ishihama, *Anal. Chem.* **2010**, *82*, 2616 – 2620.
[5] L. F. Waanders, K. Chwalek, M. Monetti, C. Kumar, E. Lammert, M. Mann, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18902 – 18907.
[6] G. Yue, Q. Luo, J. Zhang, S.-L. Wu, B. L. Karger, *Anal. Chem.* **2007**, *79*, 938 – 946.
[7] Q. Luo, G. Yue, G. A. Valaskovic, Y. Gu, S.-L. Wu, B. L. Karger, *Anal. Chem.* **2007**, *79*, 6174 – 6181.
[8] R. D. Smith, G. A. Anderson, M. S. Lipton, L. Pasa-Tolic, Y. Shen, T. P. Conrads, T. D. Veenstra, H. R. Udseth, *Proteomics* **2002**, *2*, 513 – 523.
[9] Y. Shen, N. Tolić, C. Masselon, L. Paša-Tolić, D. G. Camp II, K. K. Hixson, R. Zhao, G. A. Anderson, R. D. Smith, *Anal. Chem.* **2004**, *76*, 144 – 154.
[10] J. V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, M. Mann, *Nat. Methods* **2007**, *4*, 709 – 712.
[11] L. Sun, G. Zhu, N. J. Dovichi, *Rapid Commun. Mass Spectrom.* **2013**, *27*, 157 – 162.
[12] L. Sun, G. Zhu, X.Yan, N. J. Dovichi, *Curr. Opin. Chem. Biol.* **2013**, DOI: 10.1016/j.cbpa.2013.07.018.
[13] Y. Wang, B. R. Fonslow, C. C. L. Wong, A. Nakorchevsky, J. R. Yates III, *Anal. Chem.* **2012**, *84*, 8505 – 8513.
[14] L. Sun, G. Zhu, Y. Li, R. Wojcik, P. Yang, N. J. Dovichi, *Proteomics* **2012**, *12*, 3013 – 3019.
[15] K. Faserl, B. Sarg, L. Kremser, H. Lindner, *Anal. Chem.* **2011**, *83*, 7297 – 7305.
[16] G. Zhu, L. Sun, X. Yan, N. J. Dovichi, *Anal. Chem.* **2013**, *85*, 2569 – 2573.
[17] R. D. Smith, C. J. Barinaga, H. R. Udseth, *Anal. Chem.* **1988**, *60*, 1948 – 1952.
[18] E. J. Maxwell, D. D. Y. Chen, *Anal. Chim. Acta* **2008**, *627*, 25 – 33.
[19] M. Moini, *Anal. Chem.* **2007**, *79*, 4241 – 4246.

[20] R. Wojcik, O. O. Dada, M. Sadilek, N. J. Dovichi, *Rapid Commun. Mass Spectrom.* **2010**, *24*, 2554 – 2560.
[21] Y. Li, R. Wojcik, N. J. Dovichi, M. M. Champion, *Anal. Chem.* **2012**, *84*, 6116 – 6121.
[22] M. Wang, M. Weiss, M. Simonovic, G. Haertinger, S. P. Schrimpf, M. O. Hengartner, C. von Mering, *Mol. Cell. Proteomics* **2012**, *11*, 492 – 500.
[23] X. Yan, D. C. Essaka, L. Sun, G. Zhu, N. J. Dovichi, *Proteomics* **2013**, *13*, 2546 – 2551.
[24] D. Cohen, J. A. Dickerson, C. D. Whitmore, E. H. Turner, M. M. Palcic, O. Hindsgaul, N. J. Dovichi, *Annu. Rev. Anal. Chem.* **2008**, *1*, 165 – 190.

