JOURNAL OF
MOLECULAR AND APPLIED BIOANALYSIS

REVIEW ARTICLE
PUBLISHED: 25 May 2017

Capillary zone electrophoresis-electrospray ionization-mass spectrometry is ready for large-scale proteomics

Daoyang Chen and Liangliang Sun*

Department of Chemistry, Michigan State University, East Lansing, Michigan, USA.

Reviewed by:
Dr. Giuliana Bianco
Dr. Kambiz Gilany
Dr. Donatella Nardiello

Correspondence:
Liangliang Sun,
lsun@chemistry.msu.edu

Abstract

Capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry (CZE-ESI-MS/MS) has re-attracted great attentions for large-scale proteomics. Significant progress has been achieved in improving the CE-MS interfaces, the loading capacity and the separation window of CZE in last decade. The state-of-the-art CZE-MS/MS can routinely approach micro-liter scale loading capacity, over 2-hour separation window and 10,000 peptide identifications from complex proteome samples. It is ready for large-scale proteomics. CZE-MS/MS will make significant contributions to at least three fields of proteomics: ultra-deep bottom-up proteomics, large-scale top-down proteomics and native proteomics.

Keywords: CZE-ESI-MS/MS, bottom-up proteomics, top-down proteomics, native proteomics

INTRODUCTION

Proteomics aims to perform large-scale identification and quantification of proteins from complex and various biological samples [1]. It typically employs multi-dimensional liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MS/MS). Originating from the invention of multidimensional protein identification technology (MudPIT) in 2001 by Yates lab [2], multi-dimensional LC-MS/MS has approached great depth of complex proteomes including identification of 10,000 proteins from a human cell line [3-5] and identification of proteins encoded by 84% of the total protein-coding genes in humans [6,7]. However, several issues remain for LC-MS/MS based proteomics.

First, the median protein sequence coverage from LC-MS/MS based large-scale bottom-up proteomics is typically well below 30%, thus leading to challenges for identifying different protein isoforms from same genes. Second, the sensitivity and scale of LC-MS/MS based top-down proteomics is limited. Typically, hundreds of microgram of complex proteome materials are required for identification of thousands of proteoforms with LC-MS/MS [8,9], thus leading to challenges for applying top-down proteomics for characterization of mass-limited samples, e.g., single cells. Third, it is challenging to characterize complex proteomes in cells in their physical conditions ("Native proteomics"). Native mass spectrometry is a valuable technique for protein complex analysis in their physical conditions, providing accurate information about the dynamics of protein complexes [10]. Coupling highly efficient separation technique to native mass spectrometry will be an invaluable tool for native proteomics. Unfortunately, the typically used reversed-phase liquid chromatography (RPLC) for proteomics employs high organic solvents for separation, and organic solvents denature proteins. RPLC-MS cannot be used for native proteomics.

In order to solve those issues, alternative proteomics platforms are imperative. The platforms should significantly improve the protein sequence coverage for bottom-up proteomics, should improve the scale and sensitivity of top-down proteomics and should enable separation and detection of protein complexes in their native conditions for native proteomics.

Capillary zone electrophoresis (CZE) separates analytes based on their size-to-charge ratios. It is typically performed in a fused silica capillary with 10-50 μm inner diameter. It is one simple separation technique with high separation efficiency for different biomolecules, e.g., metabolites, peptides and proteins. It can separate protein complexes in their native conditions. CZE-ESI-MS/MS has a long history for proteomics. In 1996, Figeys *et al.* reported solid-phase microextraction (SPME)-CZE-MS/MS for analysis of yeast proteins separated by two-dimensional gel electrophoresis [11]. In 1999, Tong *et al.* coupled RP based SPME to CZE-MS/MS for two-dimensional peptide separation followed by MS/MS identification [12]. In 1996, Valaskovic *et al.* reported one CZE-MS/MS platform for online separation and identification of intact proteins. The platform only

required attomole amounts of protein molecules and could identify carbonic anhydrase from crude extract of red blood cells through database search [13].

During the 1990s, CZE-MS/MS attracted great attention for analysis of protein samples. However, during the 2000s, the contribution of CZE-MS/MS for large-scale proteomics was ignorable. There are several reasons. First, the interfaces for coupling CZE to MS are not sensitive enough. Second, loading capacity of CZE is typically at low nanoliter level. The low loading capacity limits the capability of CZE-MS/MS for large-scale profiling of complex proteome samples. Third, separation of CZE typically is fast, around 30 minutes or shorter. The fast separation of CZE limits the number of MS/MS spectra that can be acquired during one CZE-MS/MS run.

During the last six years, CZE-MS/MS based proteomics has achieved tremendous improvement [14]. First, two CE-MS interfaces with high sensitivity and good stability have been commercialized. One sheathless interface using porous tip as ESI emitter was invented by Moini group in 2007 [15], was initially commercialized by Beckman Coulter and is one product of Sciex Separations. The other one is the electro-kinetically pumped sheath-flow interface, invented by Dovichi group [16,17] and commercialized by CMP Scientific (Brooklyn, NY), **Figure 1**. The CZE-MS system based on the electro-kinetically pumped sheath-flow interface could approach 1-zmole peptide detection limit (600 molecules) and could identify over 160 proteins from *E.coli* proteome using only 80 pg of protein digests [18]. CZE-MS can produce significantly higher sensitivity than RPLC-MS for peptide detection, and CZE-MS/MS can identify more proteins and peptides than RPLC-MS/MS for mass-limited proteome sample analysis [19,20]. Recently, Yates group compared the sheathless interface based CZE-MS with RPLC-MS for intact protein analysis, demonstrating that CZE-MS has up to 100 times higher sensitivity than RPLC-MS [21]. Besides those two CE-MS interfaces, a flow-through microvial CE-MS interface from Chen group [22] and a tapered-tip CE-MS interface from Nemes group [23] also show their great sensitivity and robustness for characterization of various biomolecules.

Second, the loading capacity and separation window of CZE has been significantly improved during the last decade. In 2010, Busnel *et al.* reported 60-min separation window and over 300 peak capacity for analysis of a complex peptide mixture using sheathless interface based CZE-MS [24]. In 2014, Sun *et al.* reported 90-min separation window, 300 peak capacity and over 10,000 peptide identifications from a HeLa cell proteome digest using the electro-kinetically pumped sheath-flow interface based single-shot CZE-MS/MS [25]. Around 100 nL of peptide mixture was injected for CZE-MS/MS based on field enhanced sample stacking method. In 2014, Zhu *et al.* reported one dynamic pH junction based CZE-MS/MS system with half-a-microliter loading capacity [26]. Very recently, Chen *et al.* systematically optimized the dynamic pH junction based sample stacking method and reported a CZE-MS/MS system with microliter scale loading capacity and 140-min separation window for large-scale bottom-up proteomics [27]. The dynamic pH junction based CZE-MS/MS system significantly reduces the gap between CZE-MS and RPLC-MS in terms of loading capacity and separation window.

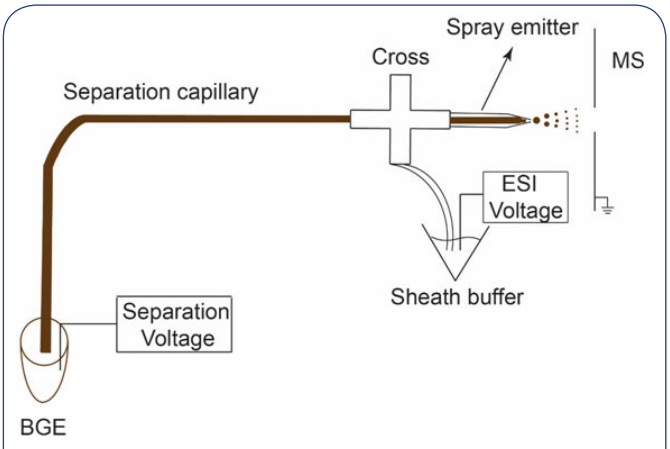

Figure 1. Schematic figure of the electro-kinetically pumped sheath-flow interface. The interface contains a cross and an electrospray emitter pulled from a borosilicate glass tubing (1 mm o.d./0.75 mm i.d.) with a puller. The orifice of the spray emitter is typically 20-40 μm. The separation capillary (typically 50/360 μm i.d./o.d.) was threaded into the spray emitter through the cross. The end of the capillary is typically etched with hydrofluoric acid to reduce the o.d. to around 70 μm. The distance between capillary end and spray emitter orifice is around 200 μm. The optimum distance between spray emitter orifice and mass spectrometer entrance is around 2 mm. Separation voltage was applied in the background electrolyte (BGE) vial for separation and ESI voltage was applied in the sheath buffer vial for ESI.

The state-of-the-art CZE-MS/MS can routinely approach micro-liter scale loading capacity, over 2-hour separation window and 10,000 peptide identifications from complex proteome samples, **Figure 2**. CZE-MS/MS is ready for large-scale proteomics. We believe CZE-MS/MS will make significant contributions to at least three fields of proteomics.

REVIEW
ULTRA-DEEP BOTTOM-UP PROTEOMICS

Multi-dimensional LC-MS/MS based bottom-up proteomics has difficulties to identify different protein isoforms from same genes due to the low protein sequence

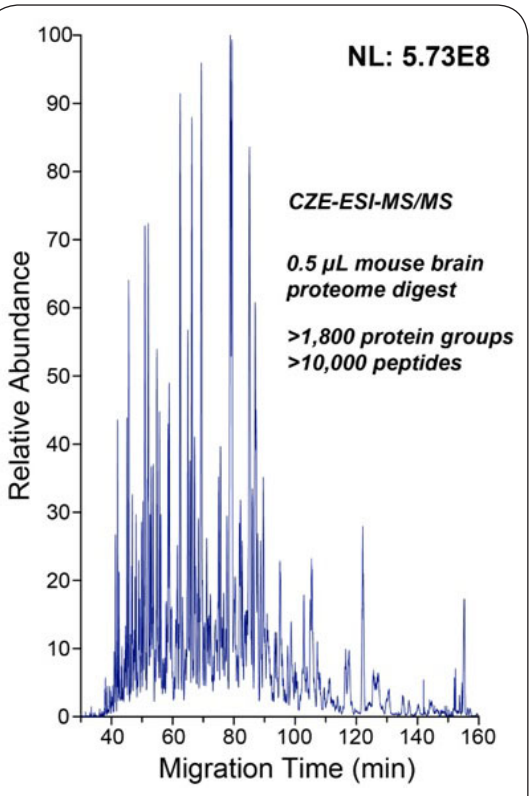

Figure 2. Electropherogram of mouse brain proteome digest analyzed by CZE-ESI-MS/MS. Sample was dissolved in 10 mM ammonium bicarbonate (pH 8). Background electrolyte of CZE was 5% (v/v) acetic acid (pH 2.2). Linear polyacrylamide coated capillary (50/360 µm i.d./o.d., 1 meter long) was used for CZE. Q-Exactive HF mass spectrometer was used for detection. The electro-kinetically pumped sheath-flow interface was used for coupling CZE to MS.

coverage. Better peptide separation is required to better the protein sequence coverage. The state-of-the-art multi-dimensional LC systems only explore the difference of peptides in charge (strong cation/anion exchange, SCX/SAX) and hydrophobicity (RPLC). We need to consider alternative separation techniques with different separation mechanisms to improve the peptide separation dramatically. CZE separates analytes based on their size-to-charge ratios and it is orthogonal to RPLC and SCX/SAX. Coupling SCX/SAX-RPLC to CZE-MS/MS will produce orthogonal three-dimensional peptide separation prior to MS detection, thus leading to significantly better protein sequence coverage, better identification of protein isoforms and deeper proteome coverage.

LARGE-SCALE TOP-DOWN PROTEOMICS

RPLC-MS/MS is typically used for top-down proteomics. The state-of-the-art top-down proteomics can identify 3,000-5,000 proteoforms from human cell lines [8,9]. The scale is much lower than bottom-up proteomics. One of the major reasons relates to the low separation efficiency of RPLC for intact proteins, especially large proteins, due to the low diffusion coefficient of proteins in the pores of the beads. CZE can separate proteins and protein complexes with high efficiency and CZE-MS can produce 100 times higher sensitivity than RPLC-MS for identification of intact proteins [21]. Coupling multi-dimensional LC to CZE-MS/MS will dramatically improve the separation of intact proteins, thus leading to much larger scale of top-down proteomics.

NATIVE PROTEOMICS

Native proteomics requires techniques for highly efficient separation of proteins/protein complexes in close to physical condition. Ion-exchange chromatography has been coupled to native ESI-MS for characterization of proteins in native condition [28]. Kelleher's group developed clear native gel-eluted liquid fraction entrapment electrophoresis (CN-GELFrEE) for separation of protein complexes in native condition [29]. CN-GELFrEE can be further coupled with native MS offline for native proteomics. CZE can separate protein complexes in their native condition, and CZE-MS has been used to separate protein complexes in lysate of red blood cells [30]. Hemoglobin A (tetramer), carbonic anhydrase II-zinc complex and carbonic anhydrase I-zinc complex were well separated and detected with CZE-MS. Coupling ion-exchange chromatography or CN-GELFrEE to CZE-MS will make large-scale native proteomics possible.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

ACKNOWLEDGEMENT

We thank Prof. Xuefei Huang's group at Department of Chemistry, Michigan State University for kindly providing the mouse brain samples. We thank the support from the Michigan State University.

REFERENCES

1. Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR 3rd. Chem Rev. 2013 Apr 10;113(4):2343-94.
2. Washburn MP, Wolters D, Yates JR 3rd. Nat Biotechnol. 2001 Mar;19(3):242-7.
3. Ding C, Jiang J, Wei J, Liu W, Zhang W, Liu M, Fu T, Lu T, Song L, Ying W, Chang C, Zhang Y, Ma J, Wei L, Malovannaya A, Jia L, Zhen B, Wang Y, He F, Qian X, Qin J. Mol Cell Proteomics. 2013 Aug;12(8):2370-80.
4. Geiger T, Wehner A, Schaab C, Cox J, Mann M. Mol Cell Proteomics. 2012 Mar;11(3):M111.014050.

5. Zhao Q, Fang F, Shan Y, Sui Z, Zhao B, Liang Z, Zhang L, Zhang Y. Anal Chem. 2017 Apr 28. doi: 10.1021/acs.analchem.6b04232.
6. Kim MS, *et al* Nature. 2014 May 29;509(7502):575-81.
7. Wilhelm M, *et al* Nature. 2014 May 29;509(7502):582-7.
8. Tran JC, Zamdborg L, Ahlf DR, Lee JE, Catherman AD, Durbin KR, Tipton JD, Vellaichamy A, Kellie JF, Li M, Wu C, Sweet SM, Early BP, Siuti N, LeDuc RD, Compton PD, Thomas PM, Kelleher NL. Nature. 2011 Oct 30;480(7376):254-8.
9. Cai W, Tucholski T, Chen B, Alpert AJ, McIlwain S, Kohmoto T, Jin S, Ge Y. Anal Chem. 2017 May 2. doi: 10.1021/acs.analchem.7b00380.
10. Heck AJ. Nat Methods. 2008 Nov;5(11):927-33.
11. Figeys D, Ducret A, Yates JR 3rd, Aebersold R. Protein identification by solid phase microextraction-capillary zone electrophoresis-microelectrospray-tandem mass spectrometry. Nat Biotechnol. 1996 Nov;14(11):1579-83.
12. Tong W, Link A, Eng JK, Yates JR 3rd. Anal Chem. 1999 Jul 1;71(13):2270-8.
13. Valaskovic GA, Kelleher NL, McLafferty FW. Science. 1996 Aug 30;273(5279):1199-202.
14. Sun L, Zhu G, Yan X, Zhang Z, Wojcik R, Champion MM, Dovichi NJ. Proteomics. 2016 Jan;16(2):188-96.
15. Moini M. Anal Chem. 2007 Jun 1;79(11):4241-6.
16. Wojcik R, Dada OO, Sadilek M, Dovichi NJ. Rapid Commun Mass Spectrom. 2010 Sep 15;24(17):2554-60.
17. Sun L, Zhu G, Zhang Z, Mou S, Dovichi NJ. J Proteome Res. 2015 May 1;14(5):2312-21.
18. Sun L, Zhu G, Zhao Y, Yan X, Mou S, Dovichi NJ. Angew Chem Int Ed. 2013 Dec 16;52(51):13661-4.
19. Faserl K, Sarg B, Kremser L, Lindner H. Anal Chem. 2011 Oct 1;83(19):7297-305.
20. Wang Y, Fonslow BR, Wong CC, Nakorchevsky A, Yates JR 3rd. Anal Chem. 2012 Oct 16;84(20):8505-13.
21. Han X, Wang Y, Aslanian A, Fonslow B, Graczyk B, Davis TN, Yates JR 3rd. J Proteome Res. 2014 Dec 5;13(12):6078-86.
22. Maxwell EJ, Zhong X, Zhang H, van Zeijl N, Chen DD. Electrophoresis. 2010 Apr;31(7):1130-7.
23. Choi SB, Zamarbide M, Manzini MC, Nemes P. J Am Soc Mass Spectrom. 2017 Apr;28(4):597-607.
24. Busnel JM, Schoenmaker B, Ramautar R, Carrasco-Pancorbo A, Ratnayake C, Feitelson JS, Chapman JD, Deelder AM, Mayboroda OA. Anal Chem. 2010 Nov 15;82(22):9476-83.
25. Sun L, Hebert AS, Yan X, Zhao Y, Westphall MS, Rush MJ, Zhu G, Champion MM, Coon JJ, Dovichi NJ. Angew Chem Int Ed. 2014 Dec 8;53(50):13931-3.
26. Zhu G, Sun L, Yan X, Dovichi NJ. Anal Chem. 2014 Jul 1;86(13):6331-6.
27. Chen D, Shen X, Sun L. Analyst. 2017, DOI: 10.1039/C7AN00509A.
28. Muneeruddin K, Nazzaro M, Kaltashov IA. Anal Chem. 2015 Oct 6;87(19):10138-45.
29. Melani RD, Seckler HS, Skinner OS, Do Vale LH, Catherman AD, Havugimana PC, Valle de Sousa M, Domont GB, Kelleher NL, Compton PD. J Vis Exp. 2016 Feb 29;(108):53597.
30. Nguyen A, Moini M. Anal Chem. 2008 Sep 15;80(18):7169-73.