

Xiaojing Yan
Liangliang Sun*
Norman J. Dovichi 
Matthew M. Champion 

Department of Chemistry and
Biochemistry, University of Notre
Dame, Notre Dame, IN, USA

Received December 9, 2019
Revised April 29, 2020
Accepted May 16, 2020

Short Communication

Minimal deuterium isotope effects in quantitation of dimethyl-labeled complex proteomes analyzed with capillary zone electrophoresis/mass spectrometry

Stable heavy-isotope labeling is commonly used in quantitative proteomics. Several common techniques incorporate deuterium (^2H) as the heavy isotopic label using reductive amination with formaldehyde. Compared with alternatives, dimethyl labeling reagents are inexpensive and the labeling chemistry is simple and rapid. However, the substitution of hydrogen by deuterium can introduce subtle changes in peptides' polarities, leading to a shift in chromatographic retention times between deuterated and nondeuterated peptides that can lead to quantification deviations. Capillary zone electrophoresis has emerged as a complementary separation for ESI-MS-based proteomics, including targeted and quantitative approaches. The extent to which the deuterium isotope effect impacts CZE-based proteomics, which separates peptides based on their S/N ratios, has not been investigated. To address this issue, CZE was used to analyze dimethyl labeled *E. coli* tryptic digests in 100 min single-shot analyses. The median migration time shift was 0.1 s for light versus heavy labeled peptides, which is 2.5% of the peak width. For comparison, nUHPLC-ESI-MS/MS was used to analyze the same sample. In UPLC, deuterated peptides tended to elute earlier than nondeuterated peptides, with a retention shift of 3 s for light versus heavy labeled peptides, which is roughly half the peak width. This shift in separation time did not have a significant effect on quantitation for either method for equal mixing ratios of the light-intermediate-heavy isotope labeled samples.

Keywords:

Capillary zone electrophoresis / Dimethyl labeling / Isotope effect / Nanospray interface for capillary electrospray / Protein quantification

DOI 10.1002/elps.202000051



Additional supporting information may be found online in the Supporting Information section at the end of the article.

MS-based proteomics is routinely used to investigate proteins and peptides within complex mixtures. Quantitative analysis, which yields information about abundance differences between samples, is of critical importance in most modern MS-based proteomic analyses [1–5]. Compared with label-free methods, which require higher precision mass spectra,

reproducible chromatography, serial injections, and consistent sample preparation [6], stable isotope labeling tends to be more robust, accurate, precise, and typically higher throughput due to multiplexing.

In stable isotope labeling quantitation, proteins or peptides from different samples are labeled with reagents that have characteristic isotopic signatures that produce identifiable mass differences. Labeled samples can be pooled and processed in a single analysis, and relative abundances are determined based on the isotopic signature introduced during sample preparation. Isotope coding approaches are often used as internal standards, in which abundances are determined by comparing to control samples [7]. ^2H , ^{13}C , ^{15}N , and ^{18}O are frequently used heavy isotopes [8].

Correspondence: Dr. Matthew M. Champion, Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556 USA.

E-mail: mchampion@nd.edu

Abbreviations: ABC, ammonium bicarbonate; FA, formic acid; IAA, iodoacetamide; iTRAQ, isobaric tags for relative and absolute quantification; RPLC-MS/MS, reversed-phase liquid chromatography-tandem mass spectrometry; SILAC, stable isotope labeling by amino acids in cell culture; TEAB, tetraethylammonium bicarbonate; TMT, tandem mass tags; UPLC, ultra-performance LC

*Current address: Department of Chemistry Michigan State University East Lansing, East Lansing, MI 48824 USA

Color online: See article online to view Figs. 1–3 in color.

There are two general strategies for incorporating heavy isotopes into samples. Metabolic labeling is accomplished by providing growth medium containing stable heavy isotopes, like ^{13}C and ^{15}N , to an organism or a cell culture, which introduces heavy isotopes exclusively to newly synthesized compounds [8]. Stable isotope labeling by amino acids in cell culture (SILAC) is a popular metabolic labeling method [9]. Cells are fed with growth medium lacking necessary amino acids, these amino acids are supplied in either natural or heavy isotope forms, and the amino acids are incorporated in newly synthesized proteins. Samples from light- and heavy-labeled cultures are mixed, digested with a protease, and subjected to downstream MS-based identification and quantification. The isotopically labeled samples are mixed in the early stages of sample preparation, which diminishes experimental bias [10].

Chemical labeling introduces stable isotopes by reacting proteins or peptides with isotopically labeled reagents or “tags”. This approach is particularly useful for samples derived from animals and plants that are not easily studied with SILAC. Chemical labeling targets primary amines or cysteine residues. Isobaric tagging (isobaric tags for relative and absolute quantification (iTRAQ) and tandem mass tags (TMT)) has gained popularity because labels can be multiplexed for up to 10 samples and because these methods generate identical nominal mass, simplifying isolation of peptides for fragmentation [11]. However, SILAC, iTRAQ, and TMT employ relatively expensive labeling reagents, which can discourage their use in high-throughput and routine-use experiments.

In this paper, we use an inexpensive stable isotope labeling strategy that targets the primary amine groups of peptides through reductive imination [12–14]. Formaldehyde reacts with primary amines of peptides to generate Schiff bases, which are reduced upon addition of cyanoborohydride, attaching a methyl group to the amine. In excess stoichiometry, this process repeats adding a second methyl group. The combination of natural forms of formaldehyde and cyanoborohydride introduces a +28 Da nominal mass increase per primary amine, which is the “light” labeled sample. An “intermediate” labeled sample uses deuterated formaldehyde and cyanoborohydride to generate a +32 Da mass increase per primary amine. Combining deuterated and ^{13}C -formaldehyde and cyanoborodeuteride produces a “heavy” labeled sample, generating a +36 Da mass increase per primary amine [6].

Labeled samples are pooled and typically analyzed using reversed-phase liquid chromatography-tandem mass spectrometry (RPLC–MS/MS). Differentially labeled peptides will ideally elute simultaneously and experience identical ionization conditions. However, deuterated peptides can experience retention time shifts during reversed-phase chromatographic separation [15,16], where deuterated peptides usually elute slightly earlier than nondeuterated [15,17]. Because of this retention time shift, quantitative analysis of control samples and experimental samples occur in dissimilar environments, which may introduce bias in abundance estimates.

CZE is an alternative separation method for bottom-up proteomics. In CZE, analytes are separated under the influence of an electric field within a capillary filled with an electrolyte. With the development of improved CZE–ESI–MS/MS interfaces [18–23], CZE–ESI–MS/MS has attracted increasing attention for MS-based proteomics. The impact of isotopic substitution on the CZE separation, which is based on ions' S/N ratios, should be less than that for RPLC, which separate analytes depending on polarities. However, the effect and potential impact on quantitative proteomic approaches remain untested. Here, we evaluate the effect of deuterium shift on quantitation using both CZE- and UPLC-based analyses.

Bovine pancreas TPCK-treated trypsin, PBS buffer, urea, ammonium bicarbonate (ABC), dithiothreitol (DTT), iodoacetamide (IAA), tetraethylammonium bicarbonate (TEAB), acrylamide, ammonium persulfate, 3-(trimethoxysilyl) propyl methacrylate, formaldehyde, and sodium cyanoborohydride were from Sigma-Aldrich (St. Louis, MO). (CD_2O , $^{13}\text{CD}_2\text{O}$) and sodium cyanoborodeuteride (NaBD_3CN) were from Cambridge Isotopes (Boston, MA). Acetic acid, hydrofluoric acid (HF), formic acid (FA), acetonitrile (ACN) were from Fisher (Pittsburgh, PA). Complete, mini protease inhibitor cocktail (provided in EASYpacks) was from Roche (Indianapolis, IN). Methanol and water were from Honeywell Burdick and Jackson (Wicklow, Ireland). Fused silica capillary (50 μm i.d., 360 μm o.d.) were from Polymicro-Technologies (Phoenix, AZ).

E. coli MG1655 was grown at 37 °C with shaking in LB (Teknova Hollister, CA) from a 1:200 dilution of an overnight culture for 8h to stationary phase. Cells were harvested by centrifugation, washed with cold PBS, and lysed with a bead mill (Biospec Bartlesville, OK) 3x passes 30 s, 100 μm Zr beads, clarified by centrifugation. The lysate was quantified using a BCA assay (Thermo), and extracted using MeOH-CHCl_3 following by vacuum concentration to dryness.

800 μg of the dried protein pellet was dissolved in 250 μL 8 M urea in 50 mM ABC and denatured at 37°C for 1 h, followed by protein reduction in 20 mM DTT at 37°C for 1 h. Proteins were alkylated by addition of iodoacetamide to 50 mM at room temperature for 30 min in the dark. After a 4-fold dilution, trypsin (enzyme to protein ratio 1:30) was added, and the solution was incubated at 37°C overnight. 10 μL FA was added to quench digestion. Digests were desalted by SepPak solid-phase extraction C_{18} (SPE) column (Waters), lyophilized, and stored at -20°C until use.

800 μg tryptic digest was dissolved in 2.4 mL of 100 mM TEAB, and 150 μL (50 μg) of the tryptic digest was aliquoted into three microcentrifuge tubes. Triplex stable isotope dimethyl labeling (light, intermediate, and heavy), 8 μL of 4% (v/v) CH_2O , CD_2O , or $^{13}\text{CD}_2\text{O}$ diluted in 100 mM TEAB was added into each aliquot to label samples with light, intermediate, and heavy dimethyl, respectively. Eight microliters of 0.6 M NaBH_3CN was added to the “light” and “intermediate” label and 8 μL of 0.6 M NaBD_3CN was added to the aliquot “heavy.” After incubating on a benchtop tube mixer at room temperature for 1 h, 24 μL of 5% (v/v) ammonia solution was added to quench the reaction on ice, then 5% FA was added

Table 1. Identification and quantification results of CZE (Upper)/UHPLC–ESI–MS/MS (Lower) for triplex dimethyl labeled *E. coli* tryptic digests mixed in different L/M/H ratios

Experiments L/M/H Ratio	Identified Peptides	Quantified Peptides	Peptide Quantification Rate (%)	Identified Proteins	Quantified Proteins	Protein Quantification Rate (%)
1:1:1	2123	1653	78	498	281	56
1:2:2	1400	1028	73	370	183	50
1:2:4	1745	1075	62	468	208	44
1:4:4	2114	1366	65	500	242	48
1:8:8	2121	1027	48	503	186	37
2:1:2	2442	1754	72	572	310	54
4:1:4	1953	1066	55	505	196	39

Experiments L/M/H Ratio	Identified Peptides	Quantified Peptides	Peptide Quantification Rate (%)	Identified Proteins	Quantified Proteins	Protein Quantification Rate (%)
1:1:1	3320	2772	84	739	501	68
1:2:2	3646	2762	76	770	471	61
1:2:4	3388	2289	68	720	414	58
1:4:4	3232	2116	66	767	383	50
1:8:8	3550	1618	46	748	292	39
2:1:2	2407	1819	76	591	354	60
4:1:4	3655	2167	59	764	388	51

until pH < 4. Light, intermediate, and heavy labeled aliquots were pooled. Labeled samples were mixed at various ratios (see data). After incubating on ice for 30 min, the samples were desalted using SPE, as above, prior to CZE–MS–MS and nUHPLC–MS–MS/MS analysis.

The interior of a fused silica capillary (≈ 100 cm length, 50 μm i.d., 360 μm o.d.) was coated with LPA using a published protocol [24].

The CZE system was similar to a previous description [25]. The electrospray voltage was provided by Spellman CZE 1000R high-voltage power supply (Hauppauge, NY). Electrospray was generated by a third-generation electrokinetically pumped sheath flow interface [26]. The electrospray emitter was a borosilicate glass capillary (1.0 mm o.d., 0.75 mm i.d., 10 cm), pulled by a Sutter instrument P-1000 flaming/brown micropipette puller (Novato, CA). The emitter diameter was ≈ 25 – 35 μm . The sheath buffer was 10% (v/v) methanol with 0.5% (v/v) formic acid, and the nano-electrospray was held at 1.8 kV above ground.

Isotopically labeled *E. coli* tryptic digests were redissolved in 20 μL loading buffer containing 35% (v/v) ACN and 65% (v/v) 0.1% (v/v) FA. A PrinCE auto-sampler (Prince Technologies B.V., Netherlands) was used to sequentially inject samples as well as for voltage control (280 V/cm). Sample was injected into the capillary by applying 200 mbar pressure for 17 s, controlled by the auto-sampler. The separation background electrolyte (BGE) was 5% (v/v) acetic acid.

An LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) was used as the detector. Full mass spectra scans were acquired by the Orbitrap in the m/z range from 350 to 1500 at 60000 mass resolution (at m/z 400). Microscans

was 1 and the maximum injection time was 500 ms. A TOP20 method was used to isolate and fragment the most intense peaks ($z > 2+$) in the collision-induced dissociation (CID) cell with 35% normalized collision energy, and maximum injection time of 100 ms. Dynamic exclusion was set to 60 s.

A nanoACQUITY ultra-performance chromatographic system with a UPLC BEH 120 C_{18} column (Waters, 1.7 μm , 100 $\mu\text{m} \times 100$ mm, column temperature 40°C) was coupled to an LTQ-Orbitrap Velos mass spectrometer. Differentially labeled *E. coli* tryptic digests were redissolved in 10 μL of 0.1% (v/v) FA. Solvent A (0.1% (v/v) FA in water) and solvent B (0.1% (v/v) FA in ACN) served as mobile phases. The gradient profile was: 0–14 min, 5% B; 14–111 min, 5–33% B; 111–115 min, 33–90% B; 115–119, 90% B; 119–120, 90–5% B; 120–135, 5% B; and 165–180, 2% B. The flow rate was 700 nL/min. The sample injection volume was 3.5 μL . The applied electrospray voltage was 1.7 kV, and the MS parameters were identical to those used for CZE–ESI–MS/MS as above.

Tandem Spectra mass matching of RAW files was performed using MaxQuant version 1.5.2.8 with Andromeda search engine [27] against the *E. coli* Uniprot UP000000625 (4391 entries). Searching parameters included full tryptic digestion, up to two missed cleavages, first search peptide tolerance 20 ppm, main search peptide tolerance 6 ppm, and fragment mass tolerance 0.5 Da. Carbamidomethylation (C) was set as a fixed modification, while acetylation (protein N-term), oxidation (M), and deamidation (NQ) were set as variable modifications. False discovery rate (FDR) was set to 0.01 for both protein and peptide identifications. Quantification was set to Dimethyl-light, intermediate, heavy on K, and peptide N termini electropherograms were imported into Matlab

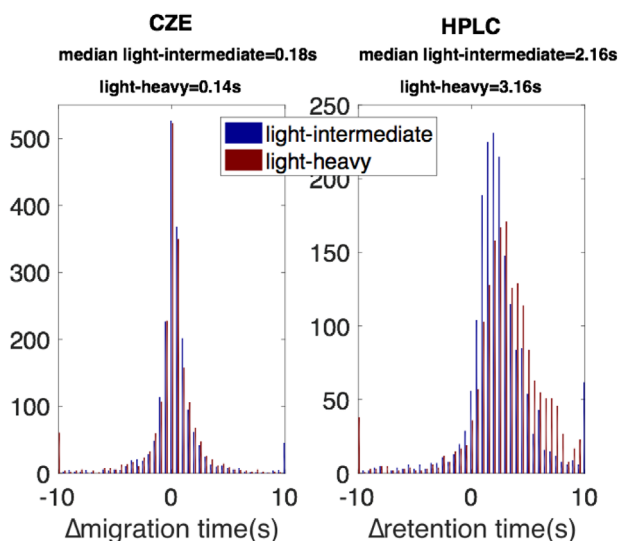


Figure 1. Separation time shift distributions of CZE-ESI-MS/MS and UPLC-ESI-MS/MS for dimethyl labeled *E. coli* tryptic digests mixed in 1:1:1 L/M/H ratio.

for additional processing. RAW data, Matlab, and MaxQuant search processing are available for download from MassIVE at <ftp://massive.ucsd.edu/MSV000084889/>.

We performed CZE-based separation, and a nUPLC-based separation and analysis using the same triplex dimethylated samples, mass spectrometer, and comparable analysis time. To investigate quantitative performance under various peptide abundances, differentially labeled samples were mixed in different light/intermediate/heavy (L/M/H) ratios.

The UPLC analysis employed 1.75 μg sample loading for each 135 min single-shot analysis, while each 100 min single-shot CZE analysis loaded 150 ng per sample. Table 1 compares the identification and quantification results for these methods. CZE-ESI-MS/MS generated 67 (± 14)% of the number of IDs as nUPLC, with similar MS-based analysis time (100 min vs. 135 min), which is consistent with data published elsewhere [26]. The rate at which quantifiable pairs and triplets decreased with increasing L/I/H ratios are in Supporting Information Figure S1a–f. The observed trend is as expected. At a given sample loading amount, a larger abundance difference means at least one label has relatively low abundance, which causes low quantitative yield [13,14].

We next evaluated the migration and retention time shifts from differentially labeled samples, Fig. 1 and Supporting Information Figure S1. Deuterated peptides tend to elute faster than nondeuterated peptides in RPLC; the median retention time shifts are 2.0 s (light vs. intermediate labeled) and 2.9 s (light vs. heavy labeled). In contrast, CZE produced a negligible isotopic shift in migration time; the median migration time shifts are 0.18 s (light vs. intermediate labeled) and 0.12 s (light vs. heavy labeled).

The peak width distributions from both methods are plotted in Fig. 2. The median peak width was ~ 5 s for CZE and ~ 7 s for UPLC. The retention time shift in HPLC produced

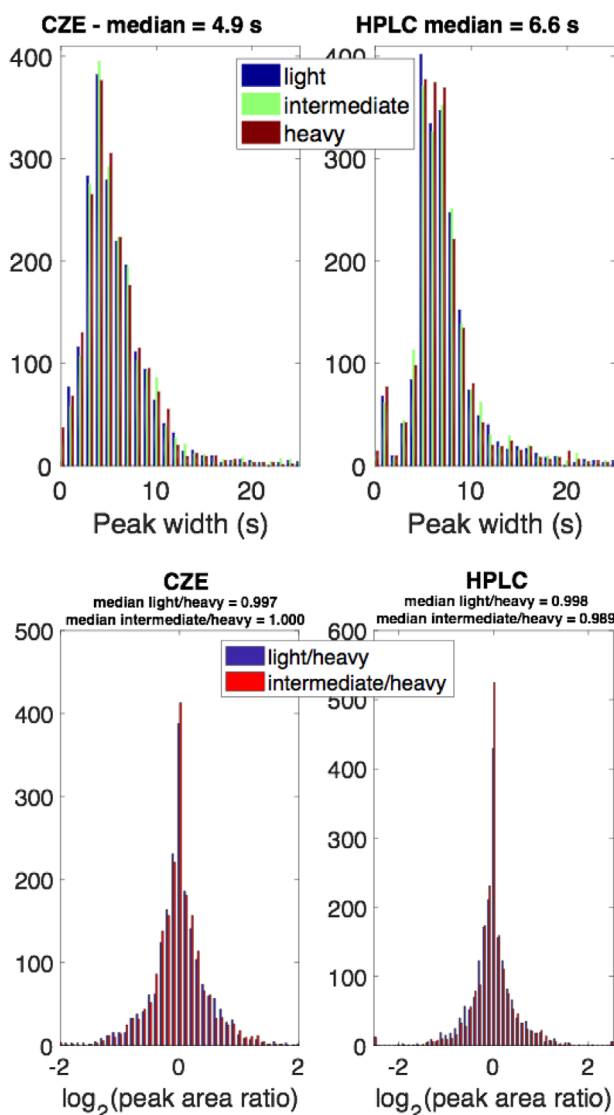


Figure 2. Peak width (top) and peak area (bottom) distributions of CZE-ESI-MS/MS and UPLC-ESI-MS/MS for dimethyl labeled *E. coli* tryptic digests mixed in L/M/H ratio 1:1:1.

by the isotope effect is approximately one-third of the peak width, whereas the migration time shift is only $\approx 3\%$ of the peak width for CZE.

We evaluated the effect of this migration/retention time shift on L/M/H ratios. At a 1:1:1 mixing ratio, the distribution of \log_2 ratio of peak areas in CZE and UPLC was quite similar, Fig. 2, and the median \log_2 ratios are 0.0048 (CZE) and -0.0078 (UPLC), while the medians for light and heavy labeled samples are -0.0174 (CZE) and -0.0259 (UPLC). While UPLC produces much larger retention time shifts than the migration time shift in CZE, this shift is not sufficient to introduce significant ion-suppression for one-to-one mixing. We speculate that unlike in previous reports [16] the improvements of UHPLC, column chemistries, and particle sizes are partially responsible for the mitigation of the observed effects. For those peptides quantified in both methods, Fig. 3

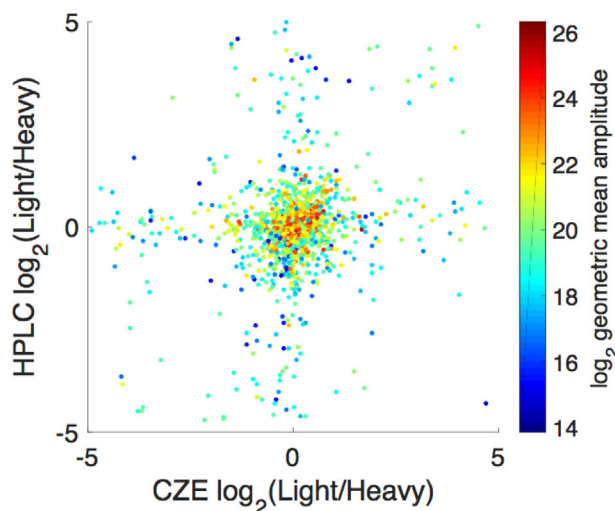


Figure 3. Scatter plot of peptides quantified in both CZE-ESI-MS/MS and UPLC-ESI-MS/MS with their corresponding geometric mean intensities identified in CZE-ESI-MS/MS for dimethyl labeled *E. coli* tryptic digests mixed in L/M/H ratio 1:1:1. Color bar presents the mean of the \log_2 (peptide intensities).

presents a scatter plot of \log_2 intensity ratios generated by the two methods; the color bar presents the \log_2 geometrical average peptide intensity. Peptides with low intensities tend to generate the largest deviations for either method. For peptides with high intensities, the empirical quantification ratios were very close to mixing ratios in both methods.

Recent work re-quantifying these effects on UPLC supports these conclusions that the effects are not particularly pernicious to data acquisition [28]. New methods employing deuterium labels such as pseudoisobaric dimethyl labeling demonstrate that deuterium labels should not preclude accurate proteomics in LC-based approaches [29]. Likewise, we conclude the deuterium isotope effect does not degrade quantitative CZE based proteomics approaches. Dimethyl labeling with CZE-MS is likely useful for small-molecule and metabolome studies, where the use of deuterated internal standards is common.

We thank Dr. William Boggess of the University of Notre Dame Mass Spectrometry and Proteomic Facility for his assistance. This work was supported by the National Institute of Health (R35GM136334 and U54GM105816).

The authors have declared no conflict of interest.

References

- [1] Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., Aebersold, R., *Nat. Biotechnol.* 1999, 17, 994–999.
- [2] Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., *Mol. Cell. Proteomics* 2004, 3, 1154–1169.
- [3] Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., Mann, M., *Mol. Cell. Proteomics* 2002, 1, 376–386.
- [4] Liu, H., Sadygov, R. G., Yates, J. R., *Anal. Chem.* 2004, 76, 4193–4201.
- [5] Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., Mann, M., *Mol. Cell. Proteomics* 2005, 4, 1265–1272.
- [6] Grossmann, J., Roschitzki, B., Panse, C., Fortes, C., Barkow-Oesterreicher, S., Rutishauser, D., Schlapbach, R., *J. Proteomics* 2010, 73, 1740–1746.
- [7] Ong, S. E., Mann, M., *Nat. Chem. Biol.* 2005, 1, 252–262.
- [8] Heck, A. J., Krijgsveld, J., *Expert Rev. Proteomics* 2004, 1, 317–326.
- [9] Ong, S. E., Mann, M., *Nat. Protoc.* 2006, 1, 2650–2660.
- [10] Zhang, G., Fenyö, D., Neubert, T. A., *J. Proteome Res.* 2009, 8, 1285–1292.
- [11] Werner, T., Sweetman, G., Savitski, M. F., Mathieson, T., Bantscheff, M., Savitski, M. M., *Anal. Chem.* 2014, 86, 3594–3601.
- [12] Hsu, J. L., Huang, S. Y., Chow, N. H., Chen, S. H., *Anal. Chem.* 2003, 75, 6843–6852.
- [13] Boersema, P. J., Aye, T. T., van Veen, T. A., Heck, A. J., Mohammed, S., *Proteomics* 2008, 8, 4624.
- [14] Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., Heck, A. J., *Nat. Protoc.* 2009, 4, 484–494.
- [15] Filer, C. N., *J. Labelled Cpd. Radiopharm.* 1999, 42, 169–197.
- [16] Zhang, R., Sioma, C. S., Wang, S., Regnier, F. E., *Anal. Chem.* 2001, 73, 5142–5149.
- [17] Bates, F. S., Fetters, L. J., Wignall, G. D., *Macromolecules* 1988, 21, 1086–1094.
- [18] Olivares, J. A., Nguyen, N. T., Yonker, C. R., Smith, R. D., *Anal. Chem.* 1987, 59, 1230–1232.
- [19] Smith, R. D., Barinaga, C. J., Udseth, H. R., *Anal. Chem.* 1988, 60, 1948–1952.
- [20] Moini, M., *Anal. Chem.* 2007, 79, 4241–4246.
- [21] Maxwell, E. J., Zhong, X., Zhang, H., Zeijl, N., Chen, D. D. Y., *Electrophoresis* 2010, 31, 1130–1137.
- [22] Maxwell, E. J., Chen, D. D. Y., *Anal. Chim. Acta* 2008, 627, 25–33.
- [23] Wojcik, R., Dada, O. O., Sadilek, M., Dovichi, N. J., *Rapid Commun. Mass Spectrom.* 2010, 24, 2554–2560.
- [24] Zhang, Z., Peuchen, E. H., Dovichi, N. J., *Anal. Chem.* 2017, 89, 6774–6780.
- [25] Krylov, S. N., Starke, D. A., Arriaga, E. A., Zhang, Z., Chan, N. W., Palcic, M. M., Dovichi, N. J., *Anal. Chem.* 2000, 72, 872–877.
- [26] Sun, L., Zhu, G., Zhang, Z., Mou, S., Dovichi, N. J., *J. Proteome Res.* 2015, 14, 2312–2321.
- [27] Cox, J., Mann, M., *Nat. Biotechnol.* 2008, 26, 1367–1372.
- [28] Boutilier, J. M., Warden, H., Doucette, A. A., Wentzell, P. D., *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* 2012, 908, 59–66.
- [29] Fang, H., Xiao, H., Li, Y., Yu, F., Xue, B., Tian, Z., *Anal. Chem.* 2016, 14, 7198–205.