

# Intact Proteoform Analysis by Capillary Electrophoresis–Mass Spectrometry. Are We There Yet?

Noah Gould<sup>+</sup>, Qianjie Wang<sup>+</sup>, Jeffrey N. Agar, Jennifer S. Brodbelt, Daoyang Chen, Kellye A. Cupp-Sutton, Elena Domínguez-Vega, Fei Fang, Marianne Fillet, Matthew S Fischer, Attila Gaspar, Ying Ge, Marie-Jia Gou, Michal Greguš, Christoph Gstöttner, Narmin Hamidli, Amanda Helms, Md Amin Hossain, Kyle J. Juetten, Neil L. Kelleher, Tobias Kraus, Eli J. Larson, J. Scott Mellors, Cynthia Nagy, Christian Neusüß, Erin A. Redman, Jasmin Schairer, Si Wu, Tian Xu, Zhitao Zhao, Guijie Zhu, Alexander R. Ivanov<sup>#\*</sup>, Kevin Jooß<sup>#\*</sup>, and Liangliang Sun<sup>#\*</sup>

**Abstract:** Mass spectrometry (MS)-based top-down proteomics (TDP) has emerged as a powerful tool for characterizing proteoforms to advance both fundamental and translational research. TDP requires high-efficiency liquid-phase separation, high-resolution MS, and tandem MS. Capillary zone electrophoresis (CZE)-MS has been proposed as a promising analytical technique for protein analysis decades ago because of its unique and valuable features, including high separation efficiency and high detection sensitivity. However, CZE-MS has not been widely adopted by the proteomics community, mainly due to concerns with its robustness and reproducibility. Here, we hypothesized that CZE-MS is sufficiently robust and reproducible for broad adoption due to the continued efforts of the community over the last three decades. In this work, for the first time, research teams from around the world validated the robustness, repeatability, and reproducibility of CZE-MS for TDP in both simple and complex model proteoform mixtures employing a full spectrum of commercially available capillary electrophoresis (CE)-MS interfaces, instrumentation, and compared CZE-MS performance with state-of-the-art liquid chromatography (LC)-MS methods. This study offers the research community an informative resource of ready-to-use experimental CE-MS techniques and a better understanding of the CZE-MS approach and its potential in TDP, accelerating the broad adoption of CZE-MS in proteoform research.

[\*] N. Gould<sup>+</sup>, J. N. Agar, M. Greguš, M. A. Hossain, A. R. Ivanov<sup>#</sup>

Barnett Institute of Chemical and Biological Analysis and  
Department of Chemistry & Chemical Biology, Northeastern  
University, Boston, MA, USA

E-mail: [a.ivanov@northeastern.edu](mailto:a.ivanov@northeastern.edu)

Q. Wang<sup>+</sup>, F. Fang, G. Zhu, L. Sun<sup>#</sup>

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

E-mail: [lsun@chemistry.msu.edu](mailto:lsun@chemistry.msu.edu)

J. S. Brodbelt, A. Helms, K. J. Juetten

Department of Chemistry, University of Texas, Austin, TX, USA

D. Chen

Discovery Biologics, Merck Research Laboratory, USA

K. A. Cupp-Sutton, S. Wu, Z. Zhao

Department of Chemistry and Biochemistry, University of Alabama,  
Tuscaloosa, AL, USA

E. Domínguez-Vega, C. Gstöttner

Center for Proteomics and Metabolomics, Leiden University Medical  
Center, Leiden, The Netherlands

M. Fillet, M.-J. Gou

Laboratory for the Analysis of Medicines, Center for Interdisciplinary  
Research on Medicines (CIRM), University of Liege, Quartier  
Hospital, Liege, Belgium

M. S Fischer, Y. Ge, E. J. Larson

Department of Chemistry, University of Wisconsin-Madison,  
Madison, Wisconsin 53706, USA

Y. Ge

Department of Cell and Regenerative Biology and Human  
Proteomics Program, University of Wisconsin-Madison, Madison,  
Wisconsin 53705, USA

A. Gaspar, N. Hamidli, C. Nagy

Department of Inorganic and Analytical Chemistry, University of  
Debrecen, Debrecen, Hungary

M. A. Hossain

Department of Neurosurgery, Brigham and Women's Hospital,  
Harvard Medical School, Boston, MA, USA

N. L. Kelleher, T. Xu

Departments of Chemistry and Molecular Biosciences, the  
Chemistry of Life Processes Institute, and the Proteomics Center of  
Excellence, Northwestern University, Evanston, IL, USA

T. Kraus, C. Neusüß, J. Schairer

Department of Chemistry, Aalen University, Aalen, Germany

[+] Both authors contributed equally to this work.

[#] The corresponding authors contributed equally to this work and are  
listed in alphabetical order.

Additional supporting information can be found online in the  
Supporting Information section

© 2025 The Author(s). *Angewandte Chemie International Edition*  
published by Wiley-VCH GmbH. This is an open access article under  
the terms of the [Creative Commons Attribution](#) License, which  
permits use, distribution and reproduction in any medium, provided  
the original work is properly cited.

## Introduction

Proteoforms are protein products derived from a single gene that can become functionally distinct due to mutations, polymorphisms, alternative splicing, and post-translational modifications (PTMs).<sup>[1]</sup> Understanding and characterizing these proteoforms is fundamental in unraveling their roles in biological processes and, thus, advancing our comprehension of human health and disease.<sup>[2–4]</sup> While the experimental intact mass of a protein obtained through single-stage mass spectrometry (MS) can often suggest a specific proteoform,<sup>[5]</sup> particularly with high mass accuracy,<sup>[6]</sup> deeper analysis is typically essential to characterize and identify proteoforms. Identification of the precise modification site, often acting as the definite marker for a proteoform, usually requires tandem mass spectrometry (MS/MS).<sup>[7–9]</sup> The analysis of intact proteoforms followed by fragmentation, referred to as “top-down” MS, has been significantly improved over the last two decades, primarily due to advancements in instrumentation as well as putting high emphasis on software development.<sup>[10,11]</sup> Protocols in top-down proteomics (TDP) distinguish themselves from conventional bottom-up proteomics approaches by omitting steps that include endoproteinase digestion prior to the analysis.<sup>[12]</sup> This allows the unequivocal characterization of individual intact proteoforms and serves as a complementary method to other MS-based analyses to enable the analysis of complex proteomes of biological samples.

Primarily, reversed-phase liquid chromatography (RPLC) has been extensively utilized in TDP as an upfront separation technique due to its well-established column and stationary phase technologies, compatibility with ESI, and robustness across various HPLC instrumentation and research laboratories.<sup>[10–15]</sup> Nonetheless, RPLC encounters challenges in efficiently separating larger intact proteoforms (>30 kDa) due to their strong interactions with the stationary phase, which is typically composed of alkyl-functionalized porous silica-based beads. Moreover, mass transfer processes within the stationary phase and from the mobile phase into the stationary phase are particularly slow for large molecular species, leading to substantial band broadening. Recent studies have demonstrated the effective application of ion exchange chromatography as an alternative LC mode of separation in characterizing proteoforms of intact biopharmaceutical proteins.<sup>[16,17]</sup> Complex proteomes, such as those found in human cell lysates, present a serious challenge, estimated to encompass over a million different proteoforms.<sup>[18,19]</sup> To over-

come the limitations of conventional techniques (e.g., RPLC) and address current challenges of the proteome complexity, the development of alternative separation techniques with increased efficiency and capacity for proteoform analysis is urgently needed to propel TDP and expedite the progress of the Human Proteoform Project.<sup>[18–21]</sup>

Capillary zone electrophoresis (CZE) emerged in the 1980s as an invaluable technique for separating small and large biomolecules, particularly proteoforms.<sup>[22,23]</sup> CZE is carried out in an open tubular fused-silica capillary or a channel in a microfluidic device filled with a background electrolyte (BGE) solution, without the need for a stationary phase, under the influence of an electric field. The separation of proteoforms is driven by differences in their electrophoretic mobility ( $\mu_{\text{ep}}$ ), which is proportional to the ratio of a proteoform's net charge to its hydrodynamic radius. While capillary electrophoresis (CE) is an umbrella term encompassing various capillary-based electrophoretic techniques, in this paper, we will use the terms CE and CZE interchangeably for the sake of simplicity. Generally, proteoforms exhibit much lower diffusion coefficients compared to peptides and metabolites, leading to exceptionally high separation efficiency of proteoforms (nearly one million theoretical plates) with subtle structural differences, especially if charge-inducing PTMs are involved.<sup>[23,24]</sup> Furthermore, minimal sample consumption (typically few nL per injection, ng to pg in protein mass), low flow rates (nL min<sup>-1</sup>), and high peak capacity in CE are advantageous for intact proteoform analysis.<sup>[24–27]</sup>

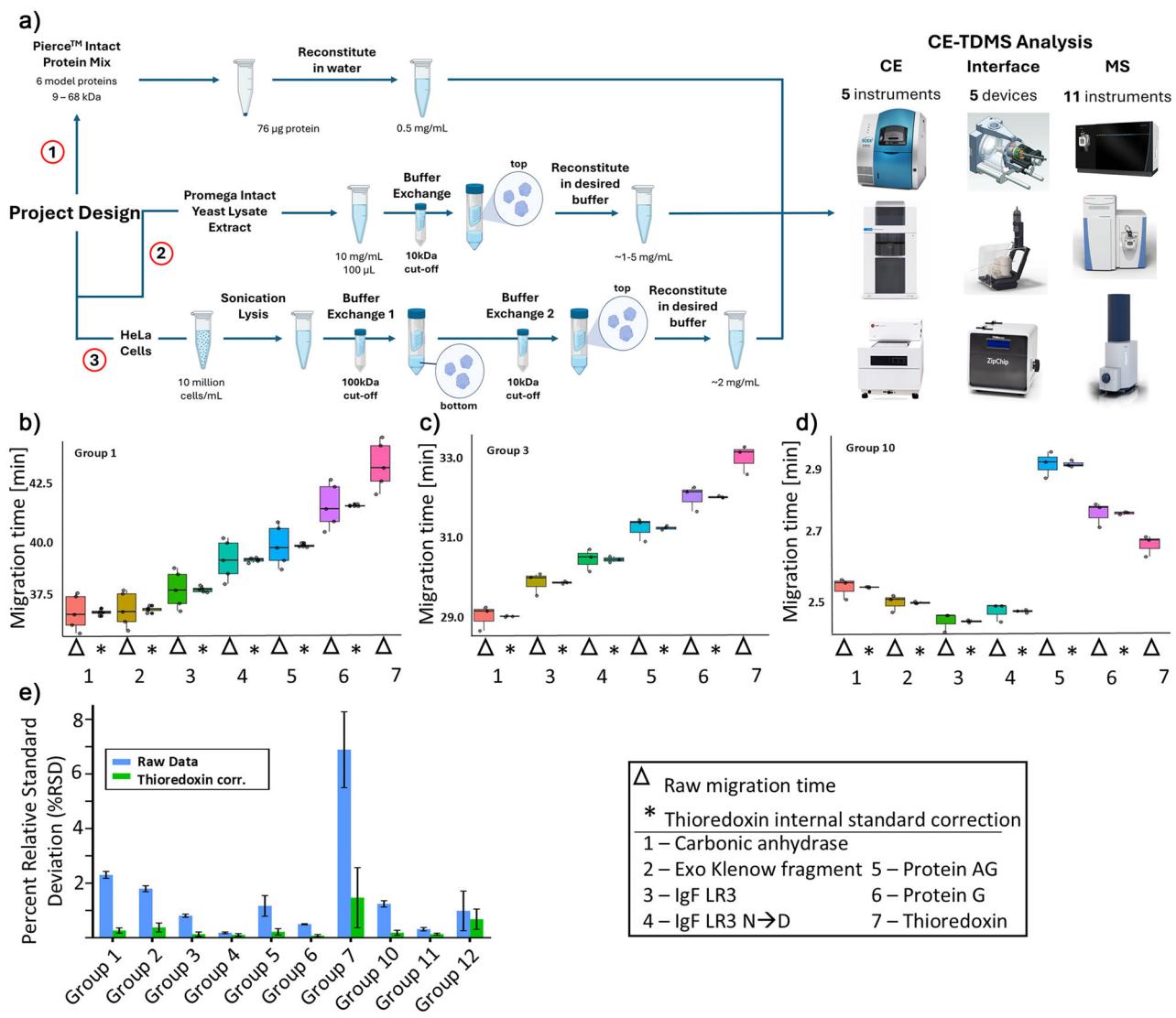
The progress and advantages of CE-based analysis raise the question of why CZE-MS has not been more widely deployed in TDP until recently. There is a general notion in parts of the proteomics community that CZE-MS is neither robust/reproducible nor sensitive enough for efficient proteoform characterization. In addition, CE-MS interfacing is considered technically challenging compared to LC-MS, i.e., due to the necessity to apply an electric field across the separation capillary. These concerns have certainly been true historically but have been the focus of industrial and academic development and technological advancements over the last decades.<sup>[22]</sup> For instance, CE-MS interfacing has been improved dramatically in regard to sensitivity and robustness. Multiple distinct interface iterations, including sheath-flow and sheathless designs, have been commercialized, making CE-MS substantially more user-friendly and accessible.<sup>[28–35]</sup> In addition, different types of CE platforms, including capillary- and microfluidic chip-based, have been established over the past years.<sup>[29–31]</sup> These more advanced setups have been successfully employed for characterizing biopharmaceuticals (e.g., monoclonal antibodies, bi-specific therapeutics, etc.),<sup>[13,36–39]</sup> profiling proteoforms in cells or tissues globally,<sup>[40–43]</sup> analyzing single cells,<sup>[26,30,44–48]</sup> understanding critical proteins implicated in disease<sup>[7,49,50]</sup> in a proteoform-specific manner, deciphering nanoparticle protein corona in nanomedicine,<sup>[51,52]</sup> and enabling native proteomics of endogenous protein complexes in cells.<sup>[53–57]</sup> The separation mechanisms of CZE and conventional RPLC can be viewed as largely complementary (charge-to-size ratio versus hydrophobicity). Therefore, CZE-MS analysis

---

J. S. Mellors, E. A. Redman  
908 Devices Inc., Boston, MA, USA

S. Wu, Z. Zhao  
Department of Chemistry and Biochemistry, University of  
Oklahoma, Norman, OK, USA

K. Joos<sup>#</sup>  
Division of Bioanalytical Chemistry, Department of Chemistry and  
Pharmaceutical Sciences, Amsterdam Institute of Molecular and Life  
Sciences, Vrije Universiteit Amsterdam, Centre for Analytical  
Sciences Amsterdam (CASA), Amsterdam, The Netherlands  
E-mail: [k.jooss@vu.nl](mailto:k.jooss@vu.nl)



**Figure 1.** a) General overview of the workflow of the project, with three main experimental arms: the analysis of the Pierce™ intact protein mix, the Promega intact yeast lysate extract, and the analysis of HeLa cell lysate. Representative boxplots of migration times generated by Groups 1, 3, and 10 (Figure 1b-d, respectively) without migration time normalization (Δ) and with normalization using thioredoxin (Trx) as an internal standard (\*). (e) Bar chart comparing %RSD values before and after internal standard correction ( $n = 3$ ). The error bars in (e) represent the standard deviation of RSD values for different proteins in the sample.

offers additional insights into the proteoform landscape. This was recently demonstrated for the characterization of three human tissues, namely heart, small intestine, and spleen, where samples were analyzed by both RPLC- and CZE-MS-based TDP workflows.<sup>[41]</sup> In this study, 28% of proteins and 56% of proteoforms were uniquely identified by CZE-MS, highlighting the gain and complementarity of CZE-MS- and LC-MS-based workflows offer to the proteomics community.

Given recent advances, it is the right time to reevaluate the performance and potential of CZE-MS for TDP and, in this way, attract more attention and interest to the CE-MS-based TDP techniques by presenting the results of this worldwide cross-laboratory study (12 research groups). The experimental design is shown in Figure 1a. The location distribution of the 12 research groups is shown in Figure S1. Our goal is to provide a repository of information generated in multiple

laboratories using a variety of CZE-MS platforms handled by novices and experts in the field, including capillary- and microfluidic chip-based CE setups (Figure 1a). Additionally, the diversity of CE-MS interfaces ( $n = 5$ ) and MS platforms ( $n = 11$ ) employed by the participants in this study covers a wide range (i.e., nearly all) of commercially available CE systems and the majority of mass spectrometers suitable for TDP applications, which represents a major novum for such studies. The details of the CE-MS systems used by the 12 research teams are listed in Table 1. The goal of this study is not to replicate a standardized CE-MS setup across multiple sites, but rather to showcase the diversity of CE-MS systems accessible to the TDP community and to demonstrate that reproducible and sensitive results can be achieved with any of these configurations. We also aimed to test the hypothesis that through the continued efforts of the CE-MS community

**Table 1:** The details of CE-MS systems (CE-MS interfaces and mass spectrometers) used by the 12 research groups in this study.<sup>a)</sup>

Group	Interface Type	Interface Model	Interface Manufacturer	MS Type	MS Model
1	Sheathless	Sciex OptiMS	Sciex	Orbitrap	Eclipse - Tribrid
2	Nano Sheath-flow	EMASS-II	CMP Scientific	Orbitrap	Q Exactive HF
3	Sheathless	Sciex OptiMS	Sciex	Orbitrap	Q Exactive UHMR
4	Sheathless	Sciex OptiMS	Sciex	QTOF	Bruker Impact
5	Sheath-flow	Agilent Triple-Tube	Agilent	QTOF	Bruker maXis II
6	Nano Sheath-flow	nanoCEeasy		Orbitrap	Fusion Lumos - Tribrid
7	Nano Sheath-flow	EMASS-II	CMP Scientific	TOF	Agilent TOF 6230
8	Nano Sheath-flow	EMASS-II	CMP Scientific	Orbitrap	Q Exactive HFX
9	Sheath-flow	Agilent Triple-Tube	Agilent	IM-QTOF	Agilent 6560 IM Q-TOF
10	Microchip	ZipChip	908 Devices	Orbitrap	Exploris 240 BioPharma
11	Microchip	ZipChip	908 Devices	TOF	Bruker timsTOF Pro
12	Sheathless	Custom		Orbitrap	Exploris 240

<sup>a)</sup> The Sciex OptiMS interface (Sciex) was developed by the Moini group.<sup>[34]</sup> The EMASS-II interface (CMP Scientific) was developed by the Dovichi group, and its original name is the electrokinetically pumped sheath-flow nanospray interface.<sup>[29]</sup> The nanoCEeasy is the modified version of the electrokinetically pumped sheath-flow nanospray interface.<sup>[31]</sup>

over the past three decades, CZE-MS has achieved levels of robustness and reproducibility for broad adoption in the field of TDP, and that it can offer unique, complementary information compared to traditional LC-based approaches. Optimized methods and benchmark data at this breadth of the evaluated instrumentation will enable users of all skill levels to obtain high-quality data for the analysis of intact proteoforms with their respective CZE-MS platforms. Here, we present methods, data analysis approaches, and results acquired for identical samples by nearly a dozen research laboratories around the globe, demonstrating that CZE-MS is a flexible, powerful, and reproducible technique for intact proteoform analysis in various laboratory settings. In this study, we provide CZE-MS practitioners with a guide to the best practices at each step of the intact CZE-MS proteoform analysis workflows and review corresponding benchmark data acquired using representative experimental approaches. Finally, we explore potential future directions for subsequent studies, along with addressing the remaining challenges that still impact CE-MS in TDP analysis. We will also highlight strategies to overcome these obstacles, aiming to encourage the growth of CE-MS-based applications in TDP and potentially other research fields. All method details are described in Supporting Information I.

## Results and Discussion

### CE-MS Analysis of Intact Protein Standard Mixture

The selected commercially available standard mixture contains six proteins—*Bos taurus* carbonic anhydrase II (CA-II), insulin-like growth factor 1 long R3 (IgF-1 LR3), *Staphylococcus aureus* protein AG (chimeric), *Streptococcus dysgalactiae* protein G, thioredoxin (Trx), and *Escherichia coli* exonuclease (Exo) Klenow fragment (EKFr)—ranging from approximately 9 to 68 kDa (Table S1). According to the manufacturer, IgF-1 LR3 includes a prominent deamidation variant, bringing the total to seven major analyte species in the sample. This protein standard mix is widely used in protein characterization as a model system for method development,

serving as a standard protein ladder in gel electrophoresis. Consequently, it was selected for initial benchmarking and showcase of CE-MS performance. Moreover, the deamidated variant of IgF-1 LR3 is particularly valuable for demonstrating the ability of CZE to efficiently separate proteoforms carrying net charge-modifying PTMs. Twelve research groups analyzed this sample using the CE-MS instrumentation, protocols, and settings outlined in Protocol 1a and Protocol 1b in Supporting Information I. Detailed experimental conditions are listed in Tables S2 and S3.

Most participants detected at least six of the seven protein species (Table S4). Importantly, ten groups detected and achieved baseline separation of the deamidated IgF-1 LR3 proteoform, highlighting the ability of CE to resolve and analyze charge variant species—a critical feature in both biological research and biopharmaceutical applications for the regulation and characterization of biological therapeutics. However, only six groups detected EKFr, likely due to its large size (~68 kDa), and associated challenges for some MS platforms, along with its co-migration with CA-II, potentially causing substantial ionization suppression.

The migration order for the observed proteins shows a high level of consistency between research groups using similar CE conditions (Table S4). Variations in BGE composition, particularly increased acid strength or concentration, can impact migration order by shifting pH-dependent protein charge states and potentially affecting conformation, both key factors in CZE's charge- and hydrodynamic volume (i.e., size)-based separation mechanism. Groups that used relatively similar CE conditions, i.e., Groups 1, 2, 3, and 9, show an extremely consistent migration order: CA-II migrating first, followed by EKFr (if observed), IgF-1 LR3, deamidated IgF-1 LR3, protein AG, protein G, and finally Trx. The one major deviation in terms of migration order was observed in one group using a microchip-based CE platform with 50% acetonitrile in the BGE, which is expected to modify the migration order relative to aqueous systems. Additionally, the capillary/channel surface chemistry and charge under the conditions of the BGE are important factors that contribute to the observed migration patterns. The average relative standard deviations (RSDs) of migration times for each

observed protein across triplicate measurements are less than 3% for eleven labs and about 7% for one lab, indicating reasonable reproducibility of various CE-MS systems (Table S5). Migration time correction using one of the proteins, Trx, as a proxy internal standard further improved the migration time RSDs to below 1% (except Group 7), (Table S6 and Figure 1b–e), indicating excellent reproducibility of CZE-MS for protein analysis. Unsupervised hierarchical clustering analysis of Trx-corrected migration times confirmed that groups using comparable CE-MS instrumentation and conditions produce similar results. Clustering patterns corresponded with capillary coating chemistries (e.g., linear polyacrylamide used by Groups 1–4; Table S2) and MS optimization, with groups detecting EKFr clustering together (Figure S2). These findings collectively demonstrate strong consistency and reproducibility among CE-MS systems across laboratories.

Notably, the microchip-based CE system achieved efficient separation and detection of all seven proteoform species in the standard protein mixture in under three minutes, using a field strength of  $500 \text{ V cm}^{-1}$  (Figure 1d). This is substantially higher than the field strengths used by groups with traditional CE systems (Groups 1–9,  $165\text{--}385 \text{ V cm}^{-1}$ ). Such microchip CE systems are well-suited for high-throughput applications, including quality control or research and development environments in the biopharmaceutical industry. Overall, all groups achieved separation and detection of observed species in under 45 min, which is at least comparable, if not faster than, the analysis times obtained in traditional reversed-phase and ion-exchange LC. Additionally, capillary length, BGE composition, capillary surface coating, and other conditions can be further optimized to meet specific analytical needs, such as an increase in throughput for industrial, academic, or clinical applications. In general, CE systems offer arguably more flexibility than conventional LC approaches as they do not require a specified stationary phase, allowing for more straightforward adjustments and optimization of separation conditions.

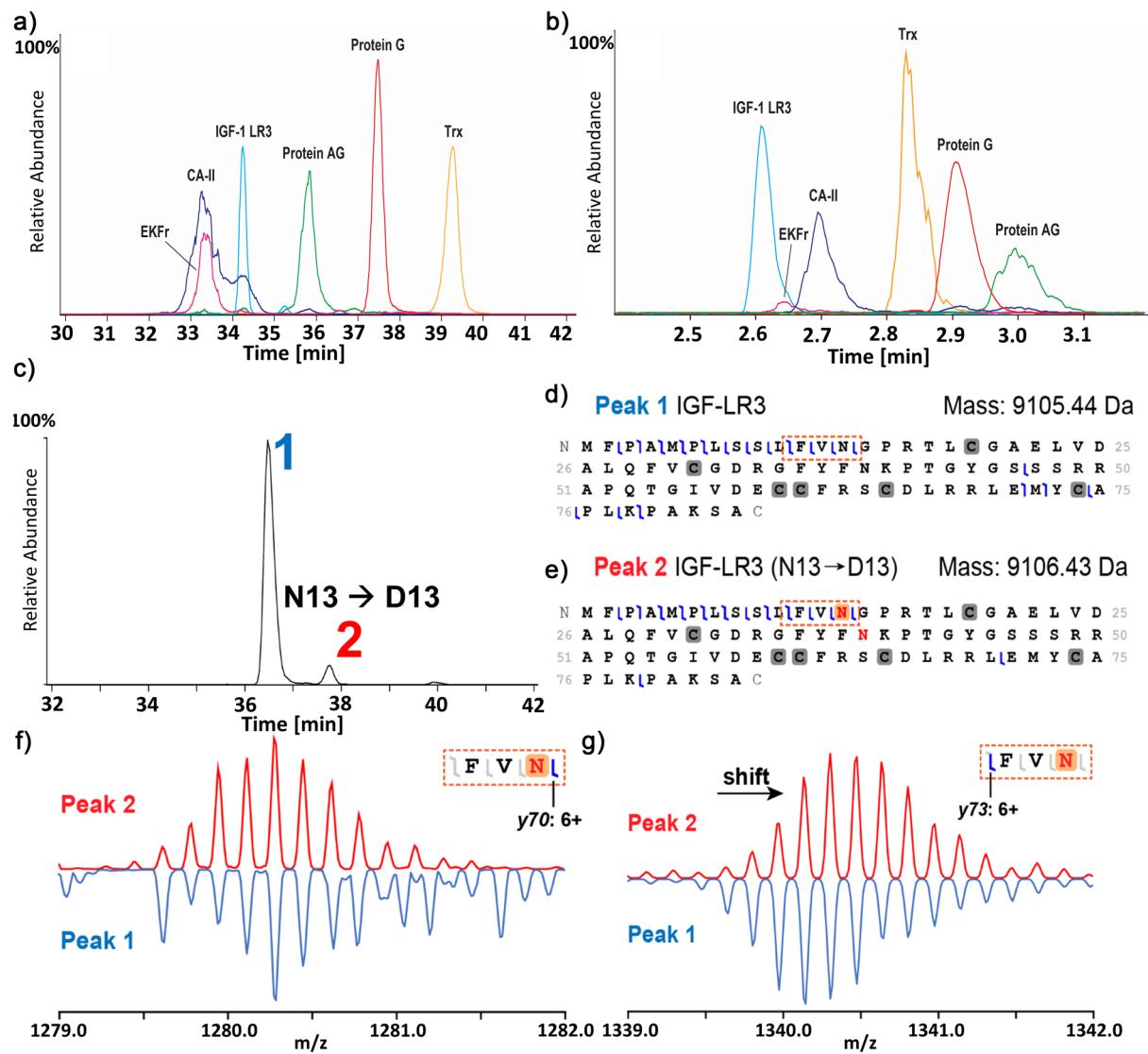
MS/MS fragmentation was performed by several groups to gather sequence-level information and PTMs of the proteins (Figures 2 and S3). As mentioned above, one of the distinct features of CE is the ability to efficiently resolve proteoforms containing PTMs that induce a change in the net charge of the protein; for instance, deamidation is a known critical PTM (which may also be a degradation product during biomanufacturing and storage) that adds an additional negative charge to the protein through the conversion of glutamine or asparagine into their acidic analogs. This modification is highly relevant in the production of biotherapeutics as a critical quality attribute (CQA) while also playing an essential role in human health and disease. Figure 2 illustrates the capability to characterize deamidated proteoforms of IgF-1 LR3 via CE-MS. Baseline separation between peak 1 and peak 2 was achieved (Resolution = 3.6,  $n = 5$ ; Figure 2c). These peaks exhibit a distinct mass shift ( $\Delta M = 0.99 \text{ Da}$ ), indicative of a deamidation event. Fragment maps for peaks 1 and 2 are depicted in Figure 2d,e. The sequence coverage of the fragment maps was compromised due to the presence of three intact disulfide bridges prior to fragmentation, resulting

in limited coverage in the middle part of the amino acid sequence. Nevertheless, clear indicators for the localization of the deamidation sites were observed. The corresponding +1 Da mass shift was detected in all y-ions toward the N-terminus from the amino acid residue N13 for peak 2, as exemplified by y73 (6+) in Figure 2g. In contrast, fragment ion y70 (6+, between N13 and G14) no longer exhibits the same mass shift (Figure 2f), pinpointing the deamidation site at N13. This precise localization underscores CE's capability to resolve and identify proteoforms (e.g., deamidated and non-deamidated species) that would co-elute under reversed-phase LC conditions, highlighting its exceptional sensitivity to charge-based PTMs in top-down proteomics.

### CE-MS Analysis of Yeast Cell Lysate

The intact yeast (*Saccharomyces cerevisiae*) protein extract was analyzed using different settings outlined in Protocol 3, Tables S7 and S8. Figure 3 shows base peak electropherograms for the yeast cell lysate sample obtained by eight participants using various CE-MS platforms. The different capillary coatings [LPA (Groups 1 and 2), bare fused silica without coating (Groups 3 and 5), PEO (Group 6), PVA (Group 7), PEI (Group 12)] and CE separation channel format [capillary (Groups 1–3, 5–7, and 12) and microchip (Group 11)] can have substantial effect on the CZE separation profile due to changes in electroosmotic flow and interactions of proteins with the capillary surface (Figure 3a–h). Despite these methodological differences, the CZE-MS separation profiles are highly consistent across the triplicate measurements within each laboratory (Figure 3i–k). Notably, the microchip-based CZE-MS system employed a much shorter separation channel (22 cm versus  $\geq 70 \text{ cm}$  for capillary systems), allowing high-throughput analysis of complex samples (Figure 3g,k).

After database search of raw files containing MS/MS data using the TopPIC Suite pipeline,<sup>[58,59]</sup> between 120–1700 proteoforms corresponding to 50–275 proteoform families were identified from five research groups (Table S9). The higher proteoform counts observed for some groups can be attributed to greater sample injection amounts (18 versus 160 ng), wider CZE separation window (20 versus 60 min), or faster data acquisition duty cycle and sensitivity from more advanced mass spectrometers (e.g., Eclipse Tribrid versus Q-Exactive HF) (Table S9). Importantly, the number of proteoform identifications was consistent across the triplicate analyses within most laboratories, as reflected by standard deviations below 10% (Table S9). These results demonstrate that reproducible CZE-MS analysis of complex proteomes can be achieved using a range of CE-MS interfaces, instruments, and mass spectrometers. In addition, the molecular mass distributions of identified proteoforms were broadly consistent across research groups (Figure 4a). Group 3 also analyzed the same yeast lysate sample using both CZE-MS and LC-MS coupled to the same MS platform under nearly identical conditions. Across groups, the pI distributions of observed proteoforms were similar regardless of capillary coating chemistry (Figure S4). An exception was Group 6,

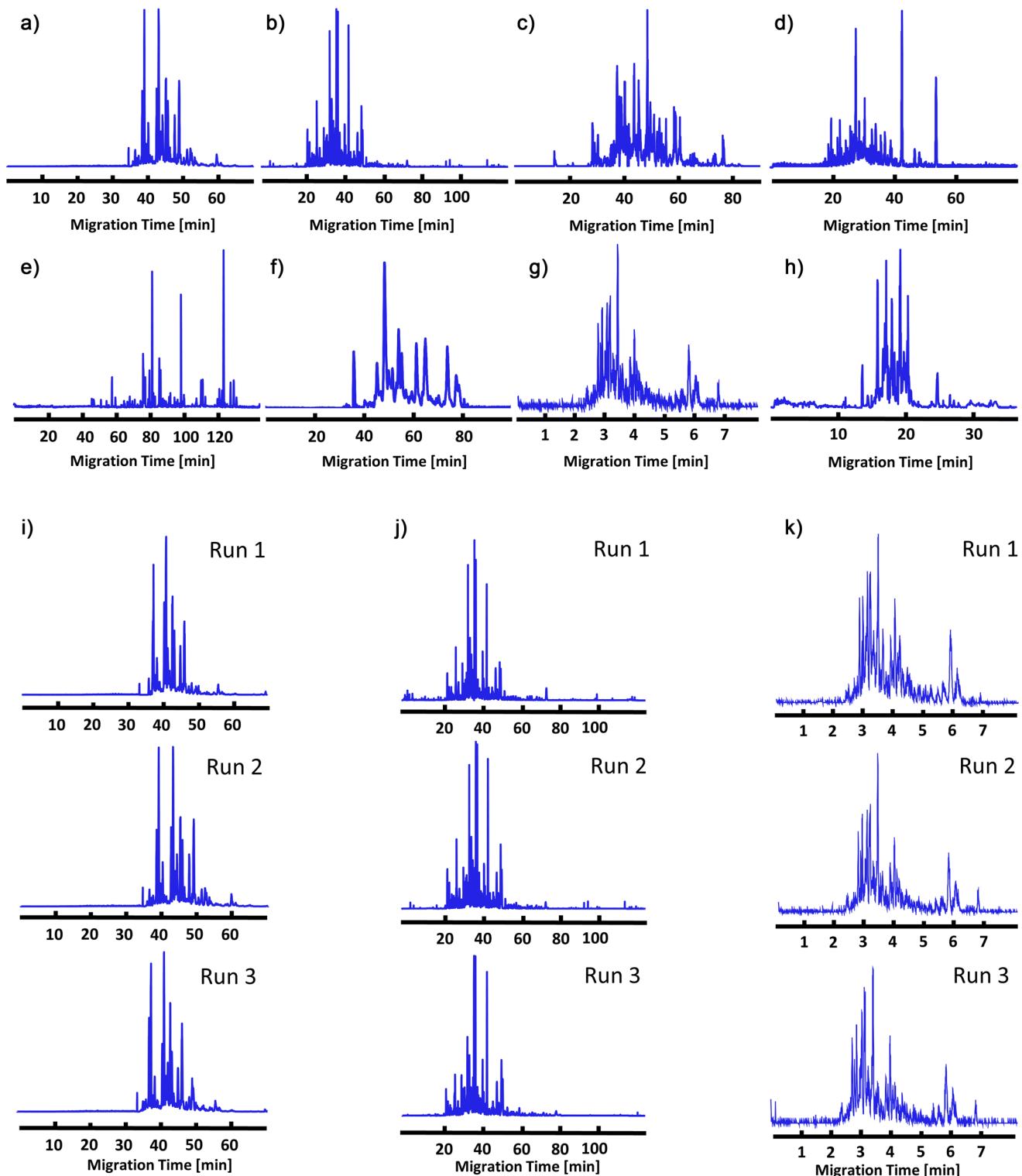


**Figure 2.** Comparison of representative electropherograms obtained for standard protein mix under denaturing conditions via traditional standard CE instrumentation from Group 1 a) and microchip-based from Group 10 b). c) Extracted ion electropherogram (EIE) of IGF-1 LR3 displaying two distinct peaks. EIE trace was generated using the three most abundant charge states. d) and e) Illustrated fragment maps produced by peaks 1 and 2, respectively. Regions of interest for localizing deamidation are outlined by boxes with orange dashed lines. Deconvoluted mass of peak 2 exhibits a distinct mass shift associated with deamidation. f) and g) Comparison of isotopic distributions of diagnostic fragment ions  $y_{70}$  and  $y_{73}$  for peak 1 (blue, upside down) and peak 2 (red). A mass shift between peaks 1 and 2 is evident for fragment ion  $y_{73}$  but not for  $y_{70}$ , indicating the presence of deamidation at amino acid residue N13.

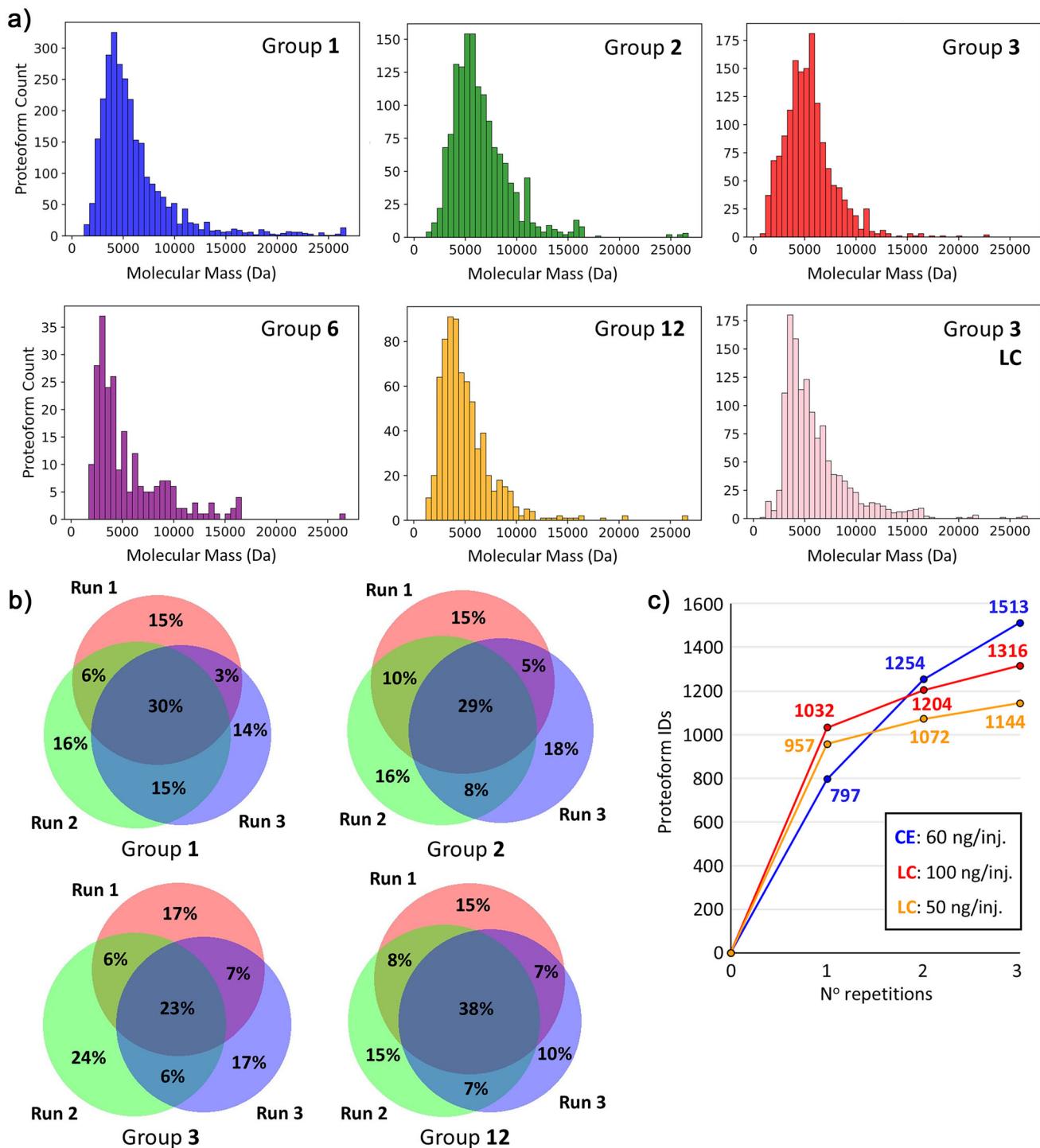
which injected the smallest sample amount ( $\sim 20$  ng) and detected the fewest number of proteins; the low sample injection amount and number of protein IDs are a probable cause of the slightly altered pI distribution. In Group 3's comparison, LC-MS yielded a higher proportion of high-pI proteins than CZE-MS. In the CZE-MS experiment, a 60- $\mu$ g injection and 90-min separation were used, while in LC-MS, a 100- $\mu$ g injection and 60-min effective separation gradient were employed. Despite these differences, both approaches produced comparable proteoform and protein identifications and exhibited similar molecular mass distributions (Figure 4a).

To further assess the reproducibility of proteoform identification across the evaluated experimental conditions, Venn

diagrams were generated to compare triplicate yeast lysate analyses from four study participants (Figure 4b). Despite differences in capillary coating, instrumentation, and methods, an average of 30% of proteoforms were consistently identified across all three runs within each group, while only 17% of proteoforms were uniquely identified in a single run, underscoring the method's reliability. Furthermore, Group 3 directly compared LC-MS (50- $\mu$ g injection) and CE-MS (60- $\mu$ g injection) analyses performed on the same mass spectrometer using identical data acquisition parameters (Figure S5). Combined, the two methods yielded 2169 identified proteoforms, with CE-MS identifying 369 (32%) more proteoforms than LC-MS at comparable sample loads. Even when the LC-MS injection amount was doubled to



**Figure 3.** Electropherograms of yeast proteins from eight participants, Groups 1–3, 5–7, 11, and 12, respectively. a)–f) and h) Base peak electropherograms for participants using traditional capillary electrophoresis. g) Base peak electropherogram data using the microchip-based system. i) Triplicate electropherograms from group 1. j) Triplicate electropherograms from group 2. k) Triplicate electropherograms from group 11.



**Figure 4.** a) Molecular mass distributions of the identified proteoforms from five groups of CE participants and LC data provided by Group 3. b) Venn diagrams for the overlap of identified proteoforms across triplicate runs for the analysis of the yeast cell lysate sample from four study participants. Participants (Groups 1 & 2) used traditional CE instrumentation with near-neutral linear polyacrylamide-coated capillaries. Study participants (Group 3) used traditional CE instrumentation with an uncoated, bare-fused silica capillary. Participants (Group 12) used an in-house built CE instrument with a positively coated, polyethylenimine, capillary. In all Venn diagrams the red circle, green circle, and purple circle correspond to the first, second, and third runs, respectively. c) Comparison of the observed gains in unique proteoform identifications with increasing replicates for CE (blue) and LC (red and orange).

100 ng, CE-MS still identified 197 (15%) more proteoforms (Figure 4c). These results highlight the orthogonality and complementarity of the two methods: 48% of the proteoforms were unique to CE-MS, while only 21% overlapped between the two approaches. In single-injection experiments, LC-MS identified more proteoforms than CE-MS, likely due to longer LC gradients and broader peaks, which increase the number of proteoforms sampled per run. In contrast, CE separations typically produce narrower peaks, which—although generally advantageous—can challenge the MS and MS/MS acquisition rates and duty cycles, potentially limiting identifications per injection. However, the number of unique identifications with CE-MS increased more rapidly with repeated injections, ultimately surpassing LC-MS in replicate injection workflows (Figure 4c). Accordingly, the proteoform identification overlaps across triplicate RPLC-MS/MS analyses (100 ng injection) was higher than for CZE-MS/MS (60 ng injection) (Figure S5). Interestingly, RPLC-MS/MS with 50- $\mu$ g injections produced comparable proteoform overlap across triplicate runs to CZE-MS/MS, likely because the lower LC sample load reduced proteoform signal intensity, leading to more stochastic data acquisition. We also compared yeast protein identifications across CE-MS datasets from different groups. Among the groups with comparable injection amounts (groups 1, 2, and 3), a total of 359 proteins were observed, with 109 (~30%) shared across all three groups (Figure S6A). A large portion of the variability in identified proteoforms between groups can likely be attributed to differences in MS methods and instrumentation. When all groups were included, mapping their identified proteins onto a protein copy number per cell S-curve of the *S. cerevisiae* proteome (Figure S6B) revealed a high degree of overlap among highly abundant proteins.

We further evaluated the quantitative reproducibility of CZE-MS/MS under various instrument conditions using the feature intensity of the overlapped proteoforms between replicates. The correlated intensities of identified proteoforms across triplicates for individual groups are shown in Figure 5a. The Pearson's correlation coefficient values were  $0.93 \pm 0.05$  for all pairwise comparisons, indicating high quantitative reproducibility. Separation consistency was evaluated by comparing migration times of proteoforms observed across runs (Figure S7). For proteoforms detected in multiple runs, migration times plotted against each other showed strong linear correlations ( $R^2 > 0.98$ ), demonstrating excellent run-to-run reproducibility. Additionally, migration time correction was performed for the runs of Group 12 (Figure S8), which resulted in an improvement in the linearity of the data (i.e., the slope closer to 1). Comparison between CE (Figures S7A and S7B) and traditional nanoflow LC using a packed C4 column (Figure S7C) revealed that CE can achieve separation time reproducibility comparable to LC. Nevertheless, further studies are needed to confirm the long-term reproducibility of CZE-MS over extended timeframes. High migration time reproducibility is critical for implementing time-based peak migration alignment and targeted qualitative and quantitative MS data acquisition, which can enhance identification coverage, data consistency, quantitative accuracy, and sensitivity.

One key advantage of CZE is the predictability of migration time based on the assessed electrophoretic mobility,  $\mu_{ep}$ . The separation in CZE is driven mainly by the charge-to-size ratio of the analytes, and we could employ accurate prediction to cross-validate the presence of charge-modifying PTMs. Based on the definition, the experimental electrophoretic mobility ( $\mu_{ep}$ ) can be determined by dividing the experimental electrophoretic velocity of the solute by the applied electrical field (E), as

$$\mu_{ep} = \frac{v_{ep}}{E} = \frac{\frac{L}{t_m}}{\frac{V}{L}} = \frac{L^2}{V * t_m} \quad (1)$$

where L is the capillary length in cm,  $t_m$  is the migration time in seconds, and V is the applied potential in kV.

The theoretical  $\mu_{ep, pred}$  is calculated based on the semi-empirical model.<sup>[60]</sup> In brief, the charge  $Q$  of the analyte of the proteinaceous nature is determined by the number of positively charged amino acid residues (Ks, R<sub>s</sub>, H<sub>s</sub>, and the N-terminal amino group) in the acidic background electrolyte. A mass-to-charge ratio for the precursor ion is measured by the mass spectrometer and deconvoluted to the experimental molecular mass ( $M$ ).

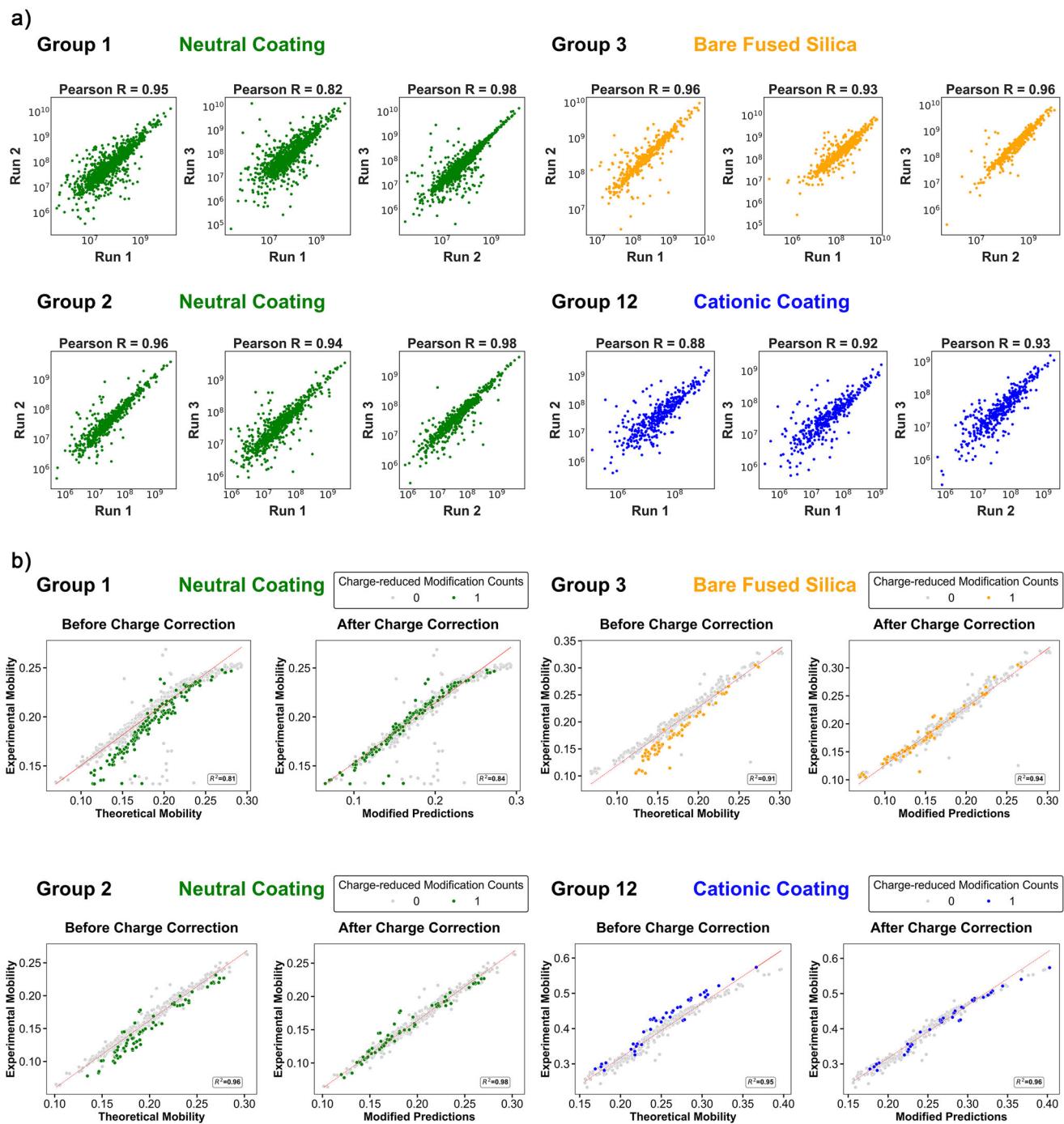
$$Predicted \mu_{ep, pred} = 5 * \frac{\ln(1 + 0.35Q)}{M^{0.411}} \quad (2)$$

As observed in Figure 5b, non-modified proteoforms aligned closely with the predicted trendline, while proteoforms carrying charge-reducing PTMs (i.e., acetylation) deviated from it. Accounting for the loss of a positive charge through acetylation improved correlation with the predicted model, confirming the ability of CZE to detect and validate charge-modifying PTMs. This additional dimension of proteoforms could not only cross-validate charge-reducing PTMs but also help resolve unknown mass shifts influenced by charge effects.

Notably, the variation of the theoretical and experimental  $\mu_{ep}$  values can be attributed to the different separation conditions, such as internal pressure (e.g., 0.5 psi) during the separation, while the linear relationship remains unchanged. Adjustments for capillary length or separation voltage could be incorporated through scaling coefficients. For cationic coatings, the equation is modified to account for the reversed electroosmotic flow (EOF), by incorporating  $\mu_{EOF}$  and a linear coefficient. Importantly, the linearity is consistent across all conditions, including neutral coating, cationic coating, and uncoated negatively charged bare fused silica. Overall, the excellent correlation between experimental and predicted  $\mu_{ep}$  of proteoforms using straightforward calculations makes CZE-MS/MS an attractive analytical method for TDP of complex proteomes with high confidence.

### CE-MS of HeLa Cell Lysate

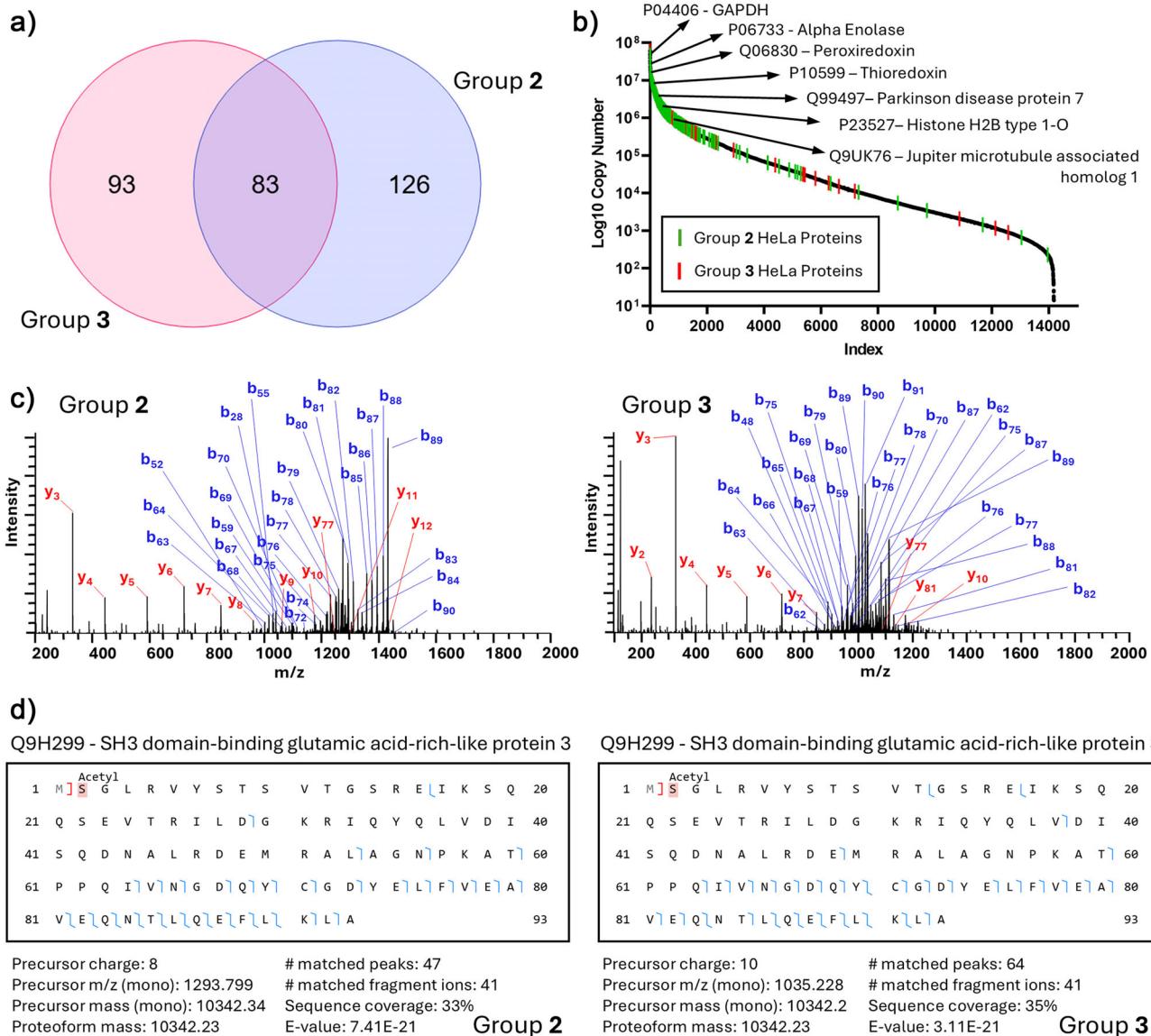
To further evaluate the CZE-MS/MS technique for complex proteomic samples, we analyzed HeLa cell lysates using two independent CZE-MS/MS systems (Groups 2 and 3). Samples



**Figure 5.** a) The log–log plots of proteoform intensity correlations between triplicates within groups. The Pearson's R score of the linear correlation and capillary condition are listed above each plot. b) The linear correlations between theoretical and experimental electrophoretic mobility ( $\mu_{ep}$ ) of proteoforms before and after charge correction (see Equation 1). Proteoforms with unidentified mass shifts and proteoforms with migration time less than 12 min were removed for better confidence. Unmodified proteoforms were labeled grey, and proteoforms with varying numbers of charge-reduced PTMs (acetylation) were labeled in colors specified in inserts. The theoretical electrophoretic mobility is calculated based on the number of positive residues (Ks, Rs, Hs, and the N-terminal amino group). The charge correction subtracts 1 from the calculated net charge value due to the presence of each charge reducing PTM.

were prepared according to protocol 4 and analyzed using settings outlined in Table S10. Group 2 identified a total of 209 HeLa proteins and 714 proteoforms across three runs with an injection amount of  $\sim$ 100 ng, while group 3 observed 176 proteins and 441 proteoforms across three runs with an

injection of  $\sim$ 20 ng of protein. Group 2 detected an average of 171 proteins and 495 proteoforms per run with RSDs of 2% and 5%, respectively. Group 3 detected an average of 108 proteins and 219 proteoforms per run with RSDs of 2% and 13%, respectively. 83 proteins (~27%) were identified



**Figure 6.** Comparison of CE-MS top-down proteomic analysis of HeLa lysate between groups 2 and 3. a) Venn diagram of the overlap in observed HeLa proteins between groups 2 and 3. b) S-curve of protein copy numbers per cell from previously published deep proteomic profiling of HeLa cells, with the observed HeLa proteins from groups 2 and 3 marked with green and red lines, respectively. Select proteins observed across both groups are highlighted. c) Exemplary MS/MS spectrum for SH3 domain-binding glutamic acid-rich-like protein 3 (Q9H299) with identified b- and y-fragment ions observed by group 2 (left) and group 3 (right). d) Fragmentation map, identification results, and fragmentation coverage for Q9H299 from group 2 (left) and group 3 (right).

by both groups (Figure 6a). Similar to the yeast analysis, we mapped the overlapping HeLa proteins to the reported protein copy numbers and onto a resulting HeLa proteome S-curve, highlighting seven exemplary proteins observed by both groups that are known to play key roles in cellular function and pathologies (Figure 6b). Additionally, we examined the MS/MS spectra for proteoforms characterized by both groups 2 and 3 and observed comparable levels of fragmentation and sequence coverage. An example is shown in Figure 6c,d, showing the MS/MS spectra and sequence coverage of one SH3 domain-binding glutamic acid-rich-like protein 3 proteoform from both groups. Although slight variations were observed in the distribution of b- and y-

ions between the two groups due to some differences in MS/MS settings (Table S10), the agreement on the overall fragmentation patterns and PTM across two independent CZE-MS systems demonstrate the consistency of CZE-MS/MS for proteoform characterization in complex biological samples under well-controlled conditions.

## Conclusion

In this work, we systematically assessed the performance and consistency of various CZE approaches, CE-MS interfaces, and MS platforms across multiple laboratories worldwide

for TDP characterization of both model protein standard mixtures and complex cell lysates. Using identical samples, we assessed both capillary-based and microfluidic CE-MS systems. Our results demonstrated a high degree of migration time reproducibility for intact protein standards, with relative standard deviation (RSD) values below 1%. The diversity of CE instrumentation, ranging from commercial to home-built systems with various capillary geometries and chemistries, supports industry-standard throughputs while preserving CE's intrinsic size- and charge-based separation capabilities for TDP analysis. We highlighted CE-MS's strength to efficiently resolve proteoforms carrying charge-altering PTMs, such as achieving baseline separation of a deamidated IgF-LR3 proteoform from its unmodified counterpart. CE-MS also proved highly effective for TDP of complex proteomes, yielding more proteoform identifications than traditional nanoLC-MS-based TDP using bead-packed C4 columns with comparable sample loads. The orthogonality between CE-MS and RPLC-MS was demonstrated, with almost 50% of total proteoforms being uniquely identified by CE-MS, underscoring the potential gain CE-MS provides for the TDP community. Furthermore, analyses of cell lysates demonstrated CE-MS's reproducibility and predictive capability (yeast lysate), as electrophoretic mobilities and migration times closely matched semi-empirical predictions. We also showed that CE-MS can be used for TDP of more biologically relevant complex samples, such as immortalized human cell lines (HeLa), and achieve similar levels of identifications and overlapping identifications across different instrumentation and groups as was observed in the analysis of the yeast lysate sample. Collectively, these findings provide strong evidence that CZE-MS/MS is ready for broad adoption for TDP of complex biological samples.

This study involves the evaluation of both capillary- and microfluidic chip-based CZE separations. Microfluidic CZE-MS provides clear advantages for high-throughput analysis, as demonstrated by our data, owing to the short separation channel (i.e., 10 and 22 cm) and strong electric field achievable in the commercial microfluidic CE system. In contrast, capillary-based CZE-MS typically requires a substantially longer separation capillary (i.e., ~60–100 cm) due to the design of commercial CE systems, which limits throughput. However, the use of longer capillaries enables higher peak and sample loading capacity separations of complex proteoform mixtures, thereby enhancing proteome coverage, as supported by our data. Furthermore, as we showcased here, various CE-MS interfaces can be employed in TDP to achieve reproducible proteoform separations and detection. Each CE-MS interface has distinct operational characteristics that have been comprehensively summarized in recent reviews.<sup>[22]</sup> We recommend that new users consider these features, along with the specific project goals (i.e., high-throughput versus deep coverage), when selecting a CZE-MS platform. Additionally, we encourage participation in practical training opportunities, such as training sessions provided by CE-MS instrument vendors and the annual CE-MS summer school organized at Michigan State University.

The participation of multiple CE-MS experts from laboratories around the globe in this study prompted discussions

on current challenges and future directions for advancing CE-MS in TDP. A major challenge in TDP of complex mixtures, such as total lysates, remains the reduced proteome coverage compared to traditional bottom-up analysis. Difficulties in measuring large proteoforms (i.e., >30 kDa), membrane proteoforms, and heavily modified proteoforms stem from issues related to the analyte solubility, frontend separation performance, ionization efficiency, desolvation/declustering efficiency, resolution of closely spaced ion species, gas-phase fragmentation efficiency, detection sensitivity, and data analysis. Another potential challenge shared by all nanoflow liquid phase separation-based techniques (e.g., CE and LC), but especially pronounced in the techniques utilizing pulled nano-electrospray emitters, is the suboptimal electrospray stability when analyzing samples of increased complexity. Sample constituents, such as lipids, detergents, and cell debris, can negatively impact spray stability and even block the capillary. CE-MS typically employs capillaries with uniform inner diameters or microfluidic channel cross-sections and unrestricted tips, making them less prone to clogging, provided that effective high-pressure rinsing is implemented between analyses. While CZE-MS offers several advantages, it may suffer from its low sample loading capacity for complex samples. Several additional improvements in enhancing sample loading capacity and capillary inner surface coating chemistries will further advance CZE-MS/MS for TDP of complex biological samples. Recently, significant efforts have been made to boost the sample loading capacity, e.g., dynamic pH junction and solid phase microextraction to better tackle "real-world" sample volumes, especially in single-cell and limited sample analysis applications.<sup>[27,61,62]</sup> Additionally, further investigations into long-term reproducibility across multiple replicates and separation capillaries and microchips, which was beyond the scope of this project, will be instrumental for the research community to fully address concerns about robustness and reproducibility. Recent work has already been reported<sup>[63]</sup> that shows the long-term reproducibility of CZE-MS for TDP, and it serves as a helpful template for future studies using different sample types. Finally, the chemistry of the capillary inner wall coating is central for achieving highly efficient proteoform separations, especially for the analysis of large proteoforms. Both neutral and cationic capillary coatings have been broadly employed for CZE-MS/MS-based protein and proteoform studies.<sup>[22,37,64–67]</sup> Continued research into coating reproducibility, stability, and applicability to diverse proteoform classes, e.g., large proteoforms, membrane proteins, and protein complexes, will further enhance the performance and utility of CE-MS for comprehensive proteoform characterization.

Based on the current state of the field and recent technological advances, we identify several promising and impactful application areas for CE-MS in TDP: I) CE-MS-based TDP appears to be an attractive approach to discovering novel proteoform biomarkers of diseases (e.g., cancer and Alzheimer's disease).<sup>[4,42,68]</sup> II) CE-MS can be instrumental in the top-down MS-based characterization of biopharmaceuticals (e.g., monoclonal antibodies, bi-/tri-specific nanobodies, among others) due to its high separation resolution (especially for charge variants and variably modified species) and superb

detection sensitivity.<sup>[13,36,37,69,70]</sup> III) CE-MS also shows great potential for native TDP, which aims to achieve proteome-wide characterization of protein complexes in cells and other biological systems.<sup>[53,57,71,72]</sup> IV) Targeted native CE-MS applications are poised to become valuable tools in structural biology, providing detailed insights into protein complex stoichiometry, conformational states, and interactions.<sup>[54–56]</sup> V) Due to its minimal sample consumption and high detection sensitivity, CE-MS offers unique advantages for the analysis of amount-limited biomedical samples, including single cells and sub- $\mu$ L aspirates from scarce biological and clinical sources.<sup>[26,44–46,48]</sup> VI) Finally, CE-MS-based TDP holds transformative potential for applications in nanomedicine, early-stage disease diagnostics, and monitoring of therapy treatment outcomes. The proteoform level resolution can be leveraged to characterize nanoparticle coronas, single cells, extracellular vesicles and particles, and proteoform alterations in tissues and biofluids,<sup>[51,52,73]</sup> potentially reshaping our understanding of biology manifested in such sample types and advancing nanomedicine.

A central objective of this study was to demonstrate that CE-MS possesses the reproducibility and sensitivity required for effective proteoform characterization and that it is suitable for broad adoption in TDP analysis. Across multiple platforms and laboratories, CE-MS showed strong reproducibility, achieving relative standard deviations (RSDs) below 1% for migration times after internal standard correction. When directly compared with LC-MS data generated in the same laboratory under identical MS conditions and data analysis parameters, CE-MS showed comparable sensitivity and produced similar numbers of proteoform identifications across triplicate runs. It is important to note that LC-MS-based TDP continues to advance rapidly.<sup>[74–76]</sup> Although the LC-MS performance in this study aligns with previously published results under comparable conditions,<sup>[74–76]</sup> it may not fully represent the current, rapidly advancing state-of-the-art in LC-MS-based TDP. Nonetheless, the advantage of CE-MS was pronounced in the large number of uniquely identified proteoforms, nearly 50%, which were undetected by the LC-MS platform, consistent with previous findings.<sup>[41]</sup> This comparison underscores the complementary nature of these two separation techniques and highlights the value of incorporating CE-MS into the suite of tools consistently used in proteomics research. Finally, due to the narrow peak shapes produced in CE, some proteoforms may still remain undetected as MS/MS acquisition may struggle to keep pace with the efficiency of CE-based separation. However, this limitation also represents a significant opportunity. As advances in MS acquisition speed and parallelization continue, CE-MS is poised to capitalize on these developments, enabling even deeper and more comprehensive proteoform coverage in future TDP studies.

## Supporting Information

Experimental methods, data analysis description, and additional supporting tables and figures can be found in the supporting information file. The mass spectrometry pro-

teomics data have been deposited to the ProteomeXchange Consortium via the PRIDE<sup>[77]</sup> partner repository with the dataset identifier PXD059108.

## Author Contributions

N.G. and Q.W. contributed equally by preparing all figures and writing the manuscript with A.R.I., K.J., and L.S. A.R.I., K.J., and L.S. (names of the corresponding authors are listed in alphabetical order) conceived the study design and contributed equally to this project. Experiments were performed by N.G., Q.W., D.C., F.F., M.S.F, M.J.G., M.G., C.G., N.H., A.H., M.A.H., K.J.J., T.K., E.J.L., E.A.R., J.S., Z.Z., and G.Z., who all contributed data that was incorporated into the final results. All authors contributed to the discussions as well as the editing and formatting of the manuscript.

## Acknowledgements

The authors are grateful to all members of the Board of Directors of the Consortium for Top-Down Proteomics for initiating the study and for valuable discussions. A.R.I. was in part supported by the National Institutes of Health, and specifically, the National Institute of General Medical Sciences under award number NIH NIGMS R35GM136421 (A.R.I.), 5P41GM128577 (A.R.I., with Prof. V. Wysocki as PI), R41GM156145 (A.R.I.), and the National Cancer Institute under award number NCI R01CA218500 (A.R.I.). L.S. was supported by the National Institute of General Medical Sciences through grant R35GM153479 (L.S.) and the National Cancer Institute (NCI) through the grant R01CA247863 (L.S.). J.B. would like to acknowledge the National Institutes of Health under award number NIH R35GM139658 (J.B.). E.D-V. would like to acknowledge support from the LUMC fellowship (2020). A.G. would like to thank the National Research, Development and Innovation Office, Hungary (K124134). Y.G. would like to acknowledge National Institute of Health (NIH) R01 GM117058 and HL109810 (Y.G.). S.W. would like to acknowledge the National Institute of Health (NIH) under award numbers NIAID R01AI141625 and NIAID 2U19AI062629, the Oklahoma Center for Advancement of Science and Technology (OCAST) under award number HR23-169, and S.W. would like to thank the University of Alabama for the support received from the startup fund.

## Materials and Correspondence

Correspondence and material requests should be addressed to Alexander R. Ivanov ([a.ivanov@northeastern.edu](mailto:a.ivanov@northeastern.edu)), Kevin Jooß ([k.jooss@vu.nl](mailto:k.jooss@vu.nl)), and Liangliang Sun ([lsun@chemistry.msu.edu](mailto:lsun@chemistry.msu.edu)).

## Conflict of Interests

Authors with company affiliations are as follows: D.C. (Merck and Co., Inc., South San Francisco, CA, USA), J.S.M (908

Devices Inc.), and E.A.R (908 Devices Inc.). Otherwise, the authors declare there are no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

### Data Availability Statement

The data that support the findings of this study are openly available in PRIDE<sup>[77]</sup> at [www.ebi.ac.uk/pride/](http://www.ebi.ac.uk/pride/), under the dataset identifier PXD 059108.

**Keywords:** Analytical methods • Electrophoresis • Mass spectrometry • Proteomics

- [1] L. M. Smith, N. L. Kelleher, *Nat. Methods* **2013**, *10*, 186–187, <https://doi.org/10.1038/nmeth.2369>.
- [2] S. Haider, R. Pal, *Curr. Genomics* **2013**, *14*, 91–110, <https://doi.org/10.2174/1389202911314020003>.
- [3] Q. Shi, L. Qin, W. Wei, F. Geng, R. Fan, Y. S. Shin, D. Guo, L. Hood, P. S. Mischel, J. R. Heath, *Proc Natl Acad Sci U S A* **2012**, *109*, 419–424, <https://doi.org/10.1073/pnas.1110865109>.
- [4] J. S. Brodbelt, *Curr. Opin. Chem. Biol.* **2022**, *70*, 102180, <https://doi.org/10.1016/j.cbpa.2022.102180>.
- [5] M. R. Shortreed, B. L. Frey, M. Scalf, R. A. Knoener, A. J. Cesnik, L. M. Smith, *J. Proteome Res.* **2016**, *15*, 1213–1221, <https://doi.org/10.1021/acs.jproteome.5b01090>.
- [6] M. Mann, N. L. Kelleher, *Proc Natl Acad Sci U S A* **2008**, *105*, 18132–18138, <https://doi.org/10.1073/pnas.0800788105>.
- [7] N. Siuti, N. L. Kelleher, *Nat. Methods* **2007**, *4*, 817–821, <https://doi.org/10.1038/nmeth1097>.
- [8] J. B. Shaw, W. Li, D. D. Holden, Y. Zhang, J. Griep-Raming, R. T. Fellers, B. P. Early, P. M. Thomas, N. L. Kelleher, J. S. Brodbelt, *J. Am. Chem. Soc.* **2013**, *135*, 12646–12651, <https://doi.org/10.1021/ja4029654>.
- [9] Y. Ge, B. G. Lawhorn, M. ElNaggar, E. Strauss, J.-H. Park, T. P. Begley, F. W. McLafferty, *J. Am. Chem. Soc.* **2002**, *124*, 672–678, <https://doi.org/10.1021/ja011335z>.
- [10] D. P. Donnelly, C. M. Rawlins, C. J. DeHart, L. Fornelli, L. F. Schachner, Z. Lin, J. L. Lippens, K. C. Aluri, R. Sarin, B. Chen, C. Lantz, W. Jung, K. R. Johnson, A. Koller, J. J. Wolff, I. D. G. Campuzano, J. R. Auclair, A. R. Ivanov, J. P. Whitelegge, L. Pasa-Tolic, J. Chamot-Rooke, P. O. Danis, L. M. Smith, Y. O. Tsybin, J. A. Loo, Y. Ge, N. L. Kelleher, J. N. Agar, *Nat. Methods* **2019**, *16*, 587–594, <https://doi.org/10.1038/s41592-019-0457-0>.
- [11] D. S. Roberts, J. A. Loo, Y. O. Tsybin, X. Liu, S. Wu, J. Chamot-Rooke, J. N. Agar, L. Pasa-Tolic, L. M. Smith, Y. Ge, *Nat Rev Methods Primers* **2024**, *4*.
- [12] T. K. Toby, L. Fornelli, N. L. Kelleher, *Annu. Rev. Anal. Chem.* **2016**, *9*, 499–519, <https://doi.org/10.1146/annurev-anchem-071015-041550>.
- [13] T. Xu, Q. Wang, Q. Wang, L. Sun, *Anal. Methods* **2024**, *16*, 4664–4682, <https://doi.org/10.1039/D4AY00651H>.
- [14] T. Xu, F. Zhang, D. Chen, L. Sun, D. Tomazela, L. Fayadat-Dilman, *mAbs* **2023**, *15*, 2229102, <https://doi.org/10.1080/19420862.2023.2229102>.
- [15] L. Fornelli, K. Srzentic, T. K. Toby, P. F. Doubleday, R. Huguet, C. Mullen, R. D. Melani, H. D. Santos Seckler, C. J. DeHart, C. R. Weisbrod, K. R. Durbin, J. B. Greer, B. P. Early, R. T. Fellers, V. Zabrouskov, P. M. Thomas, P. D. Compton, N. L. Kelleher, *Mol. Cell. Proteomics* **2020**, *19*, 405–420, <https://doi.org/10.1074/mcp.TIR119.001638>.
- [16] F. Füssl, A. Criscuolo, K. Cook, K. Scheffler, J. Bones, *J. Proteome Res.* **2019**, *18*, 3689–3702.
- [17] F. Füssl, A. Trappe, S. Carillo, C. Jakes, J. Bones, *Anal. Chem.* **2020**, *92*, 5431–5438.
- [18] L. M. Smith, N. L. Kelleher, *Science* **2018**, *359*, 1106–1107, <https://doi.org/10.1126/science.aat1884>.
- [19] R. Aebersold, J. N. Agar, I. J. Amster, M. S. Baker, C. R. Bertozzi, E. S. Boja, C. E. Costello, B. F. Cravatt, C. Fenselau, B. A. Garcia, Y. Ge, J. Gunawardena, R. C. Hendrickson, P. J. Hergenrother, C. G. Huber, A. R. Ivanov, O. N. Jensen, M. C. Jewett, N. L. Kelleher, L. L. Kiessling, N. J. Krogan, M. R. Larsen, J. A. Loo, R. R. Ogorzalek Loo, E. Lundberg, M. J. MacCoss, P. Mallick, V. K. Mootha, M. Mrksich, T. W. Muir, et al., *Nat. Chem. Biol.* **2018**, *14*, 206–214, <https://doi.org/10.1038/nchembio.2576>.
- [20] L. M. Smith, J. N. Agar, J. Chamot-Rooke, P. O. Danis, Y. Ge, J. A. Loo, L. Paša-Tolić, Y. O. Tsybin, N. L. Kelleher, T. C. FT-D Proteomics, *Sci. Adv.* **7**, eabk0734.
- [21] C. Seydel, *Nat. Methods* **2022**, *19*, 1036–1040, <https://doi.org/10.1038/s41592-022-01599-9>.
- [22] D. Chen, E. N. McCool, Z. Yang, X. Shen, R. A. Lubeckyj, T. Xu, Q. Wang, L. Sun, *Mass Spectrom Rev* **2023**, *42*, 617–642, <https://doi.org/10.1002/mas.21714>.
- [23] J. W. Jorgenson, K. D. Lukacs, *Science* **1983**, *222*, 266–272, <https://doi.org/10.1126/science.6623076>.
- [24] R. A. Lubeckyj, A. R. Basharat, X. Shen, X. Liu, L. Sun, *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1435–1445, <https://doi.org/10.1007/s13361-019-02167-w>.
- [25] Z. Zhao, Y. Guo, T. Chowdhury, S. Anjum, J. Li, L. Huang, K. A. Cupp-Sutton, A. Burgett, D. Shi, S. Wu, *Anal. Chem.* **2024**, *96*, 8763–8771, <https://doi.org/10.1021/acs.analchem.4c01119>.
- [26] K. R. Johnson, Y. Gao, M. Greguš, A. R. Ivanov, *Anal. Chem.* **2022**, *94*, 14358–14367, <https://doi.org/10.1021/acs.analchem.2c03045>.
- [27] R. A. Lubeckyj, E. N. McCool, X. Shen, Q. Kou, X. Liu, L. Sun, *Anal. Chem.* **2017**, *89*, 12059–12067, <https://doi.org/10.1021/acs.analchem.7b02532>.
- [28] C. C. Liu, J. Zhang, N. J. Dovichi, *Rapid Commun. Mass Spectrom.* **2005**, *19*, 187–192, <https://doi.org/10.1002/rcm.1769>.
- [29] L. Sun, G. Zhu, Z. Zhang, S. Mou, N. J. Dovichi, *J. Proteome Res.* **2015**, *14*, 2312–2321, <https://doi.org/10.1021/acs.jproteome.5b00100>.
- [30] S. B. Choi, A. M. Polter, P. Nemes, *Anal. Chem.* **2022**, *94*, 1637–1644.
- [31] J. Schlecht, A. Stoltz, A. Hofmann, L. Gerstung, C. Neususs, *Anal. Chem.* **2021**, *93*, 14593–14598.
- [32] S. Carillo, C. Jakes, J. Bones, *J. Pharm. Biomed. Anal.* **2020**, *185*, 113218, <https://doi.org/10.1016/j.jpba.2020.113218>.
- [33] E. J. Maxwell, D. D. Chen, *Anal. Chim. Acta* **2008**, *627*, 25–33, <https://doi.org/10.1016/j.aca.2008.06.034>.
- [34] M. Moini, *Anal. Chem.* **2007**, *79*, 4241–4246, <https://doi.org/10.1021/ac0704560>.
- [35] X. Zhong, E. J. Maxwell, C. Ratnayake, S. Mack, D. D. Chen, *Anal. Chem.* **2011**, *83*, 8748–8755, <https://doi.org/10.1021/ac202130f>.
- [36] A. M. Belov, L. Zang, R. Sebastianiano, M. R. Santos, D. R. Bush, B. L. Karger, A. R. Ivanov, *Electrophoresis* **2018**, *39*, 2069–2082, <https://doi.org/10.1002/elps.201800067>.
- [37] D. R. Bush, L. Zang, A. M. Belov, A. R. Ivanov, B. L. Karger, *Anal. Chem.* **2016**, *88*, 1138–1146, <https://doi.org/10.1021/acs.analchem.5b03218>.
- [38] L. A. Khawli, S. Goswami, R. Hutchinson, Z. W. Kwong, J. Yang, X. Wang, Z. Yao, A. Sreedhara, T. Cano, D. Tesar, I. Nijem, D. E. Allison, P. Y. Wong, Y. H. Kao, C. Quan, A. Joshi, R. J. Harris, P. Motchnik, *mAbs* **2010**, *2*, 613–624, <https://doi.org/10.4161/mabs.2.6.13333>.

[39] A. Shah, R. Desai, W. Cui, J. J. Harrahy, A. R. Ivanov, *Analyst* **2023**, *148*, 665–674, <https://doi.org/10.1039/D2AN01724E>.

[40] E. N. McCool, T. Xu, W. Chen, N. C. Beller, S. M. Nolan, A. B. Hummon, X. Liu, L. Sun, *Sci. Adv.* **2022**, *8*, eabq6348, <https://doi.org/10.1126/sciadv.abq6348>.

[41] B. S. Drown, K. Jooß, R. D. Melani, C. Lloyd-Jones, J. M. Camarillo, N. L. Kelleher, *J. Proteome Res.* **2022**, *21*, 1299–1310, <https://doi.org/10.1021/acs.jproteome.2c00034>.

[42] E. N. McCool, R. A. Lubeckyj, X. Shen, D. Chen, Q. Kou, X. Liu, L. Sun, *Anal. Chem.* **2018**, *90*, 5529–5533, <https://doi.org/10.1021/acs.analchem.8b00693>.

[43] F. P. Gomes, J. K. Diedrich, A. J. Saviola, E. Memili, A. A. Moura, J. R. Yates, 3rd, *Anal. Chem.* **2020**, *92*, 2979–2987, <https://doi.org/10.1021/acs.analchem.9b03856>.

[44] K. Bagwe, N. Gould, K. R. Johnson, A. R. Ivanov, *Trends Analys. Chem.* **2023**, *165*, 117117.

[45] C. Lombard-Banek, S. A. Moody, P. Nemes, *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 2454–2458, <https://doi.org/10.1002/anie.201510411>.

[46] J. A. Melby, K. A. Brown, Z. R. Gregorich, D. S. Roberts, E. A. Chapman, L. E. Ehlers, Z. Gao, E. J. Larson, Y. Jin, J. R. Lopez, J. Hartung, Y. Zhu, S. J. McIlwain, D. Wang, W. Guo, G. M. Diffee, Y. Ge, *Proc Natl Acad Sci U S A* **2023**, *120*, e2222081120, <https://doi.org/10.1073/pnas.2222081120>.

[47] J. S. Mellors, K. Jorabchi, L. M. Smith, J. M. Ramsey, *Anal. Chem.* **2010**, *82*, 967–973, <https://doi.org/10.1021/ac902218y>.

[48] T. Vaclavek, F. Foret, *Electrophoresis* **2022**, *44*, 313–322.

[49] P. Mao, D. Wang, *J. Proteome Res.* **2014**, *13*, 1560–1569, <https://doi.org/10.1021/pr401074t>.

[50] L. Wei, Z. R. Gregorich, Z. Lin, W. Cai, Y. Jin, S. H. McKiernan, S. McIlwain, J. M. Aiken, R. L. Moss, G. M. Diffee, Y. Ge, *Mol. Cell. Proteomics* **2018**, *17*, 134–145, <https://doi.org/10.1074/mcp.RA117.000124>.

[51] S. A. Sadeghi, A. A. Ashkarran, Q. Wang, G. Zhu, M. Mahmoudi, L. Sun, *ACS Nano* **2024**, *38*, 26024–26036, <https://doi.org/10.1021/acsnano.4c04675>.

[52] G. Zhu, S. A. Sadeghi, M. Mahmoudi, L. Sun, *Chem Commun* **2024**, *60*, 11528–11531, <https://doi.org/10.1039/D4CC02666G>.

[53] Q. Wang, Q. Wang, Z. Qi, W. Moeller, V. H. Wysocki, L. Sun, *Angew. Chem. Int. Ed.* **2024**, *63*, e202408370.

[54] A. L. Marie, F. Georgescu, K. R. Johnson, S. Ray, J. R. Engen, A. R. Ivanov, *Adv. Sci.* **2024**, *11*, e2306824, <https://doi.org/10.1002/advs.202306824>.

[55] K. Jooß, L. F. Schachner, R. Watson, Z. B. Gillespie, S. A. Howard, M. A. Cheek, M. J. Meiners, A. Sobh, J. D. Licht, M. C. Keogh, N. L. Kelleher, *Anal. Chem.* **2021**, *93*, 5151–5160.

[56] M. R. Mehaffey, Q. Xia, J. S. Brodbelt, *Anal. Chem.* **2020**, *92*, 15202–15211, <https://doi.org/10.1021/acs.analchem.0c03784>.

[57] X. Shen, Q. Kou, R. Guo, Z. Yang, D. Chen, X. Liu, H. Hong, L. Sun, *Anal. Chem.* **2018**, *90*, 10095–10099, <https://doi.org/10.1021/acs.analchem.8b02725>.

[58] Q. Kou, L. Xun, X. Liu, *Bioinformatics* **2016**, *32*, 3495–3497, <https://doi.org/10.1093/bioinformatics/btw398>.

[59] A. R. Basharat, Y. Zang, L. Sun, X. Liu, *Anal. Chem.* **2023**, *95*, 8189–8196, <https://doi.org/10.1021/acs.analchem.2c05244>.

[60] D. Chen, R. A. Lubeckyj, Z. Yang, E. N. McCool, X. Shen, Q. Wang, T. Xu, L. Sun, *Anal. Chem.* **2020**, *92*, 3503–3507, <https://doi.org/10.1021/acs.analchem.9b05578>.

[61] J. A. Colon Rosado, L. Sun, *J. Am. Soc. Mass Spectrom.* **2024**, *35*, 1120–1127, <https://doi.org/10.1021/jasms.3c00429>.

[62] Y. Wang, B. R. Fonslow, C. C. Wong, A. Nakorchevsky, J. R. Yates, 3rd, *Anal. Chem.* **2012**, *84*, 8505–8513, <https://doi.org/10.1021/ac301091m>.

[63] S. A. Sadeghi, W. Chen, Q. Wang, Q. Wang, F. Fang, X. Liu, L. Sun, *J. Proteome Res.* **2024**, *23*, 1399–1407, <https://doi.org/10.1021/acs.jproteome.3c00872>.

[64] Q. Wang, G. Gao, F. Fang, Q. Wang, P. K. Lundquist, L. Sun, *Anal. Chim. Acta* **2024**, *1328*, 343162, <https://doi.org/10.1016/j.aca.2024.343162>.

[65] L. Leclercq, M. Morvan, J. Koch, C. Neususs, H. Cottet, *Anal. Chim. Acta* **2019**, *1057*, 152–161, <https://doi.org/10.1016/j.aca.2019.01.008>.

[66] Y. Li, P. D. Compton, J. C. Tran, I. Ntai, N. L. Kelleher, *Proteomics* **2014**, *14*, 1158–1164, <https://doi.org/10.1002/pmic.201300381>.

[67] R. Haselberg, C. K. Ratnayake, G. J. de Jong, G. W. Somsen, *J. Chromatogr. A* **2010**, *1217*, 7605–7611, <https://doi.org/10.1016/j.chroma.2010.10.006>.

[68] F. Fang, T. Xu, H. T. Chien Hagar, S. Hovde, M. H. Kuo, L. Sun, *J. Proteome Res.* **2024**, *23*, 5085–5095, <https://doi.org/10.1021/acs.jproteome.4c00587>.

[69] Z. Wu, H. Wang, J. Wu, Y. Huang, X. Zhao, J. B. Nguyen, M. P. Rosconi, E. A. Pyles, H. Qiu, N. Li, *J. Pharm. Biomed. Anal.* **2023**, *223*, 115147, <https://doi.org/10.1016/j.jpba.2022.115147>.

[70] T. Xu, L. Han, A. M. George Thompson, L. Sun, *Anal. Methods* **2022**, *14*, 383–393, <https://doi.org/10.1039/D1AY01556G>.

[71] A. M. Belov, R. Viner, M. R. Santos, D. M. Horn, M. Bern, B. L. Karger, A. R. Ivanov, *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 2614–2634, <https://doi.org/10.1007/s13361-017-1781-1>.

[72] A. K. Schwenzer, L. Kruse, K. Jooß, C. Neusüß, *Proteomics* **2023**, *24*, e2300135.

[73] S. A. Sadeghi, F. Fang, R. Tabatabaeian Nimavard, Q. Wang, G. Zhu, A. A. Saei, L. Sun, M. Mahmoudi, *Nat. Protoc.* **2025**, <https://doi.org/10.1038/s41596-025-01229-6>.

[74] Y. Shen, N. Tolić, P. D. Piehowski, A. K. Shukla, S. Kim, R. Zhao, Y. Qu, E. Robinson, R. D. Smith, L. Paša-Tolić, *J. Chromatogr. A* **2017**, *1498*, 99–110.

[75] M. Greguš, Y. Gao, K. R. Johnson, S. Ray, A. L. Marie, A. R. Ivanov, *Anal. Chem.* **2025**, *97*, 14189–14199.

[76] P. T. Kaulich, K. Jeong, O. Kohlbacher, A. Tholey, *Nat. Methods* **2024**, *21*, 2397–2407, <https://doi.org/10.1038/s41592-024-02481-6>.

[77] Y. Perez-Riverol, J. Bai, C. Bandla, D. García-Seisdedos, S. Hewapathirana, S. Kamatchinathan, D. J. Kundu, A. Prakash, A. Frericks-Zipper, M. Eisenacher, M. Walzer, S. Wang, A. Brazma, J. A. Vizcaño, *Nucleic Acids Res.* **2022**, *50*, D543–D552, <https://doi.org/10.1093/nar/gkab1038>.

Manuscript received: August 20, 2025

Revised manuscript received: October 21, 2025

Manuscript accepted: November 07, 2025

Version of record online: December 18, 2025