Decoding Disease: Insights from Proteomics in Cancer and Viral Pathogenesis

ARTICLE COLLECTION

Sponsored by:

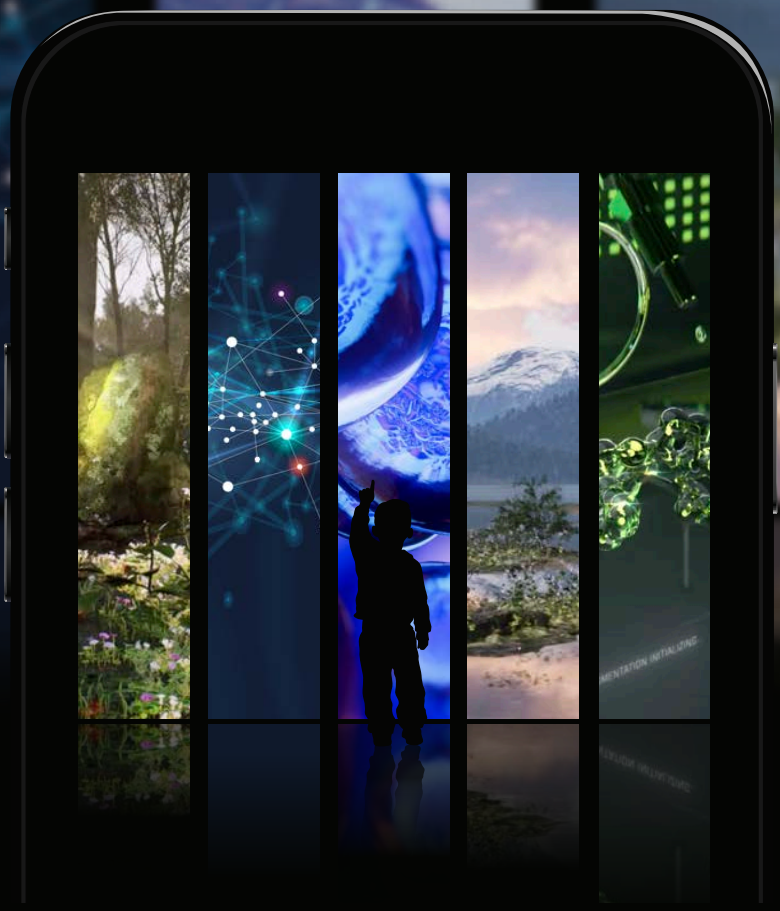
WILEY Analytical Science

SCIEX
The Power of Precision

SCIEX

The Power of Precision

Extraordinary science



Coined by philosopher and science historian Thomas Kuhn,

“extraordinary science”

refers to discoveries that go beyond current knowledge, challenging understandings and paving the way for scientific revolutions.

WATCH NOW

Be inspired by the extraordinary science from proteomics researchers using SCIEX ZenoTOF 7600 system.

Visit:

sciex.com/extraordinary-science

Contents

FURTHER READING AND RESOURCES (CLICK BELOW TO LEARN MORE)

sciex.com/extraordinary-science

Introduction

Proteomics, the study of proteins and their functions within biological systems, plays a crucial role in understanding diseases such as cancer and viral infections. By analyzing the proteome, researchers can gain insights into the molecular mechanisms underlying these diseases and identify potential diagnostic and therapeutic targets.

In cancer research, proteomics has contributed significantly to the identification and characterization of specific proteins associated with different types and stages of the disease. By comparing the proteomes of cancer cells to healthy cells, scientists can discover biomarkers that indicate the presence of cancer, predict patient outcomes, and guide personalized treatment strategies. Proteomic technologies allow for the analysis of protein expression levels, post-translational modifications, and protein interactions, providing a comprehensive understanding of the complex biological pathways involved in cancer development and progression. Moreover, proteomics has proven valuable in identifying potential targets for therapeutic interventions. By analyzing the proteomic profiles of cancer cells, researchers can identify vulnerabilities and specific proteins that are crucial for cancer cell survival. These findings can inform the development of targeted therapies that selectively inhibit these proteins, leading to improved treatment outcomes and reduced side effects compared to traditional chemotherapy approaches.

In the context of viral infections, proteomics plays a vital role in understanding the interaction between viruses and host cells. By studying the viral proteome, researchers can identify viral proteins involved in infection, viral replication, and evasion of host immune responses. Additionally, proteomics can reveal host cell proteins that are dysregulated during viral infection, providing insights into the molecular mechanisms underlying virus-host interactions. Proteomic studies of viral infections have led to significant discoveries, including the identification of viral proteins as targets for antiviral drugs and the characterization of host cellular pathways hijacked by viruses. Furthermore, proteomic analyses have been instrumental in elucidating host immune responses to viral infections and identifying potential biomarkers for diagnostic and prognostic purposes.

Overall, proteomics has revolutionized the field of disease research, particularly in the areas of cancer and viral infections. By uncovering the intricate details of protein expression, modifications, and interactions, proteomic studies provide a deeper understanding of the molecular basis of diseases and pave the way for the development of targeted therapies and improved diagnostic approaches.

This article collection begins with a study by Dutt, M. *et al.* [1] that aimed to identify serum glycoprotein biomarkers for early detection of high-grade serous ovarian cancer (HGSOC). The researchers used a glycoproteomics methodology called lectin magnetic bead array (LeMBA)-mass spectrometry (MS) to analyze serum samples from HGSOC patients. They identified several candidate proteins and lectins that showed potential as biomarkers for HGSOC. Validation analysis confirmed the altered glycoforms of these proteins in HGSOC, and a multimarker signature showed promise in distinguishing HGSOC from benign and healthy groups. The study also found that some glycoforms were altered in preclinical samples collected before HGSOC diagnosis, suggesting the potential for early detection. These findings provide a basis for further research in larger cohorts.

In response to the increasing cases of monkeypox (MPX) outside of previously endemic areas, Wang, Z. *et al.* [2] conducted a study to better understand the disease. They analyzed the plasma proteome of a group of MPX patients with similar infection histories and clinical manifestations. The study found that MPX is associated with a strong plasma proteomic response, particularly in nutritional and acute phase response proteins. There was also a correlation between plasma proteins and disease severity. Comparing the host response in MPX to that of COVID-19, similarities and differences were observed. For example, CFHR1 was induced in COVID-19 but suppressed in MPX, highlighting variations in the role of the complement system between the two infectious diseases. Interestingly, the study found that a COVID-19 biomarker panel assay could potentially be repurposed for MPX, indicating some overlap in response proteins. Using a targeted protein panel assay, encouraging results were obtained, as the assay was able to distinguish MPX cases from healthy controls. These findings provide an initial understanding of the MPX human host response at a proteomic level and suggest further exploration of protein-panel assays in emerging infectious diseases.

Finally, Maráková, K. *et al.* [3] report a study that focuses on the possibilities and limitations of capillary electrophoresis with mass spectrometry (CE-MS) for protein analysis at the intact level. CE is a separation technique valued for its efficiency, low sample consumption, reproducibility, and compatibility with liquid chromatography. While CE experiments typically use optical detection methods, the combination of CE with mass spectrometry has been developed to provide structural information and overcome the limitations of optical detection. CE-MS is increasingly used in protein analysis, particularly in biopharmaceutical and biomedical research. It

allows for the determination of physicochemical and biochemical parameters of proteins, offering excellent performance in characterizing biopharmaceuticals and biomarker discovery. This review also discusses various CE modes, CE-MS interfaces, and approaches to prevent protein adsorption and enhance sample loading capacity. It also summarizes recent developments and applications in the field of biopharmaceutical and biomedical analysis from 2018 to March 2023.

Through the methods and applications presented in this article collection, we hope to educate researchers on new technologies and methodologies in the field of proteomics. To gain a deeper understanding of available options for improving your research, we encourage you to visit SCIEX where you'll find a collection of 'extraordinary science' from proteomics researchers.

Róisín Murtagh

Editor at *Wiley Analytical Science*

References

[1] Dutt, M. *et al.* (2023). Discovery and validation of serum glycoprotein biomarkers for high grade serous ovarian cancer. *Proteomics - Clinical Applications*. DOI: 10.1002/prca.202200114.

[2] Wang, Z. *et al.* (2022). The human host response to monkeypox infection: a proteomic case series study. *EMBO Molecular Medicine*. DOI: 10.15252/emmm.202216643.

[3] Maráková, K. *et al.* (2023). Capillary electrophoresis-mass spectrometry for intact protein analysis: Pharmaceutical and biomedical applications (2018–March 2023). *Journal of Separation Science*. DOI: 10.1002/jssc.202300244.

Received: 16 December 2022 | Revised: 6 April 2023 | Accepted: 27 April 2023

DOI: 10.1002/prca.202200114

Proteomics Clinical Applications

RESEARCH ARTICLE

Discovery and validation of serum glycoprotein biomarkers for high grade serous ovarian cancer

Mriga Dutt[1] | Gunter Hartel[1] | Renee S. Richards[1] | Alok K. Shah[1] | Ahmed Mohamed[1] | Sophia Apostolidou[2] | Aleksandra Gentry-Maharaj[2] | Australian Ovarian Cancer Study Group | John D. Hooper[3] | Lewis C. Perrin[3,4] | Usha Menon[2] | Michelle M. Hill[1,5]

[1]QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

[2]MRC Clinical Trials Unit, Institute of Clinical Trials and Methodology, University College London, London, UK

[3]Mater Research Institute – The University of Queensland, Translational Research Institute, Woolloongabba, QLD, Australia

[4]Mater Adult Hospital, South Brisbane, QLD, Australia

[5]UQ Centre for Clinical Research, Faculty of Medicine, The University of Queensland, Brisbane, Australia

Correspondence
Usha Menon, MRC Clinical Trials Unit at UCL, 90 High Holborn, 2nd Floor, London WC1V 6LJ, UK.
Email: u.menon@ucl.ac.uk

Michelle Hill, QIMR Berghofer Medical Research Institute, Herston, Brisbane, QLD 4006, Australia.
Email: m.hill2@uq.edu.au

Funding information
Ovarian Cancer Research Foundation (Australia), Grant/Award Number: GA-2018-04; Medical Research Council, Grant/Award Numbers: G9901012, G0801228; Peter MacCallum Foundation, Grant/Award Number: N/A; MRC core funding for MRC CTU, Grant/Award Number: MC_UU_00004/01; UK Department of Health and Social Care, Grant/Award Number: N/A; Cancer Research UK, Grant/Award Number: C1479/A2884; Ovarian Cancer Australia; The Eve Appeal; The Cancer Council Victoria, Queensland Cancer Fund, The Cancer Council New South Wales, The Cancer Council South Australia, The Cancer Council Tasmania and

Abstract

Purpose: This study aimed to identify serum glycoprotein biomarkers for early detection of high-grade serous ovarian cancer (HGSOC), the most common and aggressive histotype of ovarian cancer.

Experimental design: The glycoproteomics pipeline lectin magnetic bead array (LeMBA)-mass spectrometry (MS) was used in age-matched case-control serum samples. Clinical samples collected at diagnosis were divided into discovery ($n = 30$) and validation ($n = 98$) sets. We also analysed a set of preclinical sera ($n = 30$) collected prior to HGSOC diagnosis in the UK Collaborative Trial of Ovarian Cancer Screening.

Results: A 7-lectin LeMBA-MS/MS discovery screen shortlisted 59 candidate proteins and three lectins. Validation analysis using 3-lectin LeMBA-multiple reaction monitoring (MRM) confirmed elevated A1AT, AACT, CO9, HPT and ITIH3 and reduced A2MG, ALS, IBP3 and PON1 glycoforms in HGSOC. The best performing multimarker signature had 87.7% area under the receiver operating curve, 90.7% specificity and 70.4% sensitivity for distinguishing HGSOC from benign and healthy groups. In the preclinical set, CO9, ITIH3 and A2MG glycoforms were altered in samples collected 11.1 ± 5.1 months prior to HGSOC diagnosis, suggesting potential for early detection.

Abbreviations: AAL, Aleuria Aurantia Lectin; AOCS, Australian Ovarian Cancer Study; Con-A, Concanavalin-A; ECA, Erythrina Cristagalli Lectin; HGSOC, High grade serous ovarian cancer; LeMBA, Lectin magnetic bead array; PHA-L, Phaseolus Vulgaris Leucoagglutinin; SNA, Sambucus Nigra Lectin; STL, Solanum Tuberosum Lectin; UKCTOCS, United Kingdom Collaborative Trial of Ovarian Cancer Screening; UKOPS, United Kingdom Ovarian Population Study; WFA, Wisteria Floribunda Lectin.

The Cancer Foundation of Western Australia, Grant/Award Numbers: 191, 211, 182; NIHR University College London Hospitals Biomedical Research Centre, Grant/Award Number: N/A; QIMR Berghofer Medical Research Institute, Grant/Award Number: N/A; Medical Research and Materiel Command, Grant/Award Number: DAMD17-01-1-0729; National Health and Medical Research Council, Grant/Award Numbers: ID199600, ID400281, ID400413

Conclusions and clinical relevance: Our findings provide evidence of candidate early HGSOC serum glycoprotein biomarkers, laying the foundation for further study in larger cohorts.

KEYWORDS
high grade serous ovarian cancer, lectin magnetic bead array (LeMBA), mass spectrometry, ovarian cancer screening, serum glycoprotein biomarker

1 | INTRODUCTION

While ovarian cancer is only the third most common gynaecological cancer worldwide, it is the leading cause of gynaecological cancer mortality. One of the main factors for this high mortality rate is diagnosis at an advanced stage due to non-specific symptoms and the lack of effective predictive or diagnostic blood biomarkers. Late diagnosis is associated with high mortality. Only 29% of women with distant metastases survive 5 years, compared to 92% with localized disease [1]. However, the existing ovarian cancer diagnostic tests in clinical use, transvaginal ultrasound and serum CA125, do not have the sensitivity required for detecting the disease in early stage. Indeed, two large ovarian cancer screening trials using a combination of these modalities found no evidence of a reduction in disease-specific mortality on long-term follow-up [2, 3]. Furthermore, neither test is specific to cancer and both trials reported unnecessary surgery in women without cancer [4]. This has led to a concerted effort to discover biomarkers for early detection of ovarian cancer [5].

Cancer is associated with alterations in the glycosylation machinery and glycan structures on circulating proteins [6]. Several studies have shown that specific glycoforms of cancer biomarkers can improve specificity. For ovarian cancer, glycosylated forms of CA125 measured by microarray [7], lectin immunoassay [8] or glycosylation-specific antibodies [9] can significantly improve differential diagnosis. This suggests that a glycoform-specific glycoprotein biomarker panel may achieve the high specificity and sensitivity required for ovarian cancer screening. Glycomic and glycoproteomics studies on ovarian cancer serum and tissues have revealed differential abundance of several types of N-glycans in ovarian cancer, including fucose, sialic acid, high mannose types [10–15]. However, these potential biomarkers are yet to be clinically validated.

Here, we report on a study using lectin magnetic bead array (LeMBA)-coupled mass spectrometry (MS) platform [16] for ovarian cancer serum glycoprotein biomarker discovery and validation. LeMBA is a one-pot, high throughput glycoproteomics method with no need for abundant serum protein depletion, potentially increasing robustness of the biomarker development process as we previously reported for oesophageal adenocarcinoma [17, 18] and canine haemangiosarcoma [19]. It has not been previously used for glycoprotein biomarker studies in ovarian cancer. We have focused on the most common and aggressive histotype, high grade serous ovarian cancers (HGSOC), which accounts for ~70% of ovarian cancers and most of the disease-specific mortality [5]. Furthermore, for discovery of biomarkers with the potential for early detection, in addition to using samples collected at clinical diagnosis as is the norm, we also evaluated samples collected prior to ovarian cancer diagnosis from the multicentre randomised controlled trial, the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS).

2 | MATERIALS AND METHODS

2.1 | Study design

Case control studies were undertaken using serum samples from women with HGSOC patients (cases) and two age-matched groups (controls)—women with a benign ovarian neoplasm and healthy women using sample sets from three independent cohorts - (1) a clinical set of 30 serum samples from the United Kingdom Ovarian Population Study (UKOPS) collected from women at diagnosis of HGSOC ($n = 10$), benign ovarian neoplasms ($n = 10$) and healthy controls ($n = 10$). (Table S1); (2) a pre-clinical set of 30 serum samples from the UKCTOCS trial [20] collected from women at a mean interval of 11.1 ± 5.1 months prior to diagnosis of HGSOC ($n = 10$), benign ovarian neoplasms ($n = 10$) and healthy controls ($n = 10$). (Table S1); and (3) a clinical set of 95 serum samples from the Australian Ovarian Cancer Study (AOCS) collected from women at diagnosis of HGSOC ($n = 39$), benign ovarian neoplasms ($n = 28$) and healthy controls ($n = 28$) (Table S2).

The discovery phase included the UKOPS clinical set and the UKCTOCS pre-clinical set. A shortlist of candidate proteins and lectins from the discovery phase was then validated using the independent clinical set from the AOCS (Figure 1). This study was approved by the QIMR Berghofer Medical Research Institute Research Ethics committee, and the East Midlands—Derby Research Ethics Committee in the UK. AOCS was approved by the Human Research Ethics Committees at the Peter MacCallum Cancer Centre, QIMR Berghofer Medical Research Institute, University of Melbourne and all participating hospitals.

2.2 | Biomarker discovery phase

LeMBA-MS was used for both discovery and validation phases (Figure 1). For discovery, seven lectins were selected from the literature [10, 11, 13, 15]: *Aleuria aurantia* (AAL), Concanavalin-A (Con-A),

Erythrina cristagalli (ECA), *Phaseolus vulgaris* Leucoagglutinin (L-PHA), *Sambucus nigra* (SNA), *Solanum tuberosum* (STL) and *Wisteria floribunda* (WFA), which preferentially target glycoproteins with fucose (α1-3, α1-4, α1-6 linked), mannose (oligomannose and hybrid-type), galactose (β1-4-linked terminal), 2,6-branched tri-, tetraantennary complex-type N-glycan, sialic acid (α2-6-linked and Tn antigen), N-acetylglucosamine ((GlcNAcβ1-4)n (Chitin), oligosaccharide containing GlcNAc and MurNAc), and N-acetylgalactosamine (terminal), respectively. All chemicals and reagents were purchased from Sigma-Aldrich, USA unless stated otherwise.

Statement of Clinical Relevance

Ovarian cancer continues to be associated with high disease mortality. Much of this is due to the diagnosis at an advanced stage of the most common and aggressive histotype—high grade serous ovarian cancer (HGSOC). This has led to significant efforts to detect the disease earlier when treatment is more effective. However, to date there is no effective screening test. We describe discovery and validation of a novel blood glycoprotein signature using lectin magnetic bead array (LeMBA)-coupled mass spectrometry for HGSOC using a case control study of clinical samples collected at diagnosis. Three biomarker candidates were altered 4–18 months prior to cancer diagnosis in a pilot case control study using preclinical samples from the UK Collaborative Trial of Ovarian Cancer Screening. Our findings suggest that serum glycoproteins could be novel biomarkers for earlier detection of HGSOC and lay the foundation for further study in larger cohorts.

2.2.1 | Serum denaturation

Thawed serum samples were centrifuged at 16,000 g and 4°C for 15 min to remove cellular debris and the supernatant protein concentration was determined by BCA protein assay (Pierce, Thermo Fischer). To minimise batch effects, bulk serum denaturation was performed for the entire project. An aliquot of each serum sample containing 800 µg of protein was diluted to 10 µg/µL in denaturation buffer (20 mM Tris-HCl pH 7.4, 1% v/v sodium dodecyl sulphate (SDS) and 5% v/v Triton X-100. The internal standard protein chicken ovalbumin was added to each serum sample at 10 pmol. Protein disulphide bonds were reduced by adding 20 mM dithiothreitol (Thermo Fisher, USA) to the samples and incubating at 37°C for 30 min. Following this, 100 mM iodoacetamide (Thermo Fisher, USA) was added to each sample and incubated at room temperature for 30 min in the dark to alkylate free thiol groups. The denatured serum samples were diluted 20 times in LeMBA binding buffer (20 mM Tris-HCl pH 7.4, 300 mM NaCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$, 1% Triton, 1 unit protease inhibitor cocktail) to yield a final protein concentration of 0.5 µg/µL. Aliquots of 100 µL were transferred to microplates in a randomized layout in preparation for LeMBA. Prepared plates were sealed and frozen at −80°C until use.

2.2.2 | Lectin magnetic bead pulldown and on-bead trypsin digest

LeMBA-MS with the selected seven lectins was performed as previously described [16, 18]. First, individual lectins were conjugated to MyOne tosyl-activated Dynabeads (Invitrogen, Australia) by incubating 50 µg of selected lectin (Vector Laboratories, USA) with 100 µL of Dynabeads at 37°C for 24 h. The resulting lectin-bead conjugate was treated with 2% w/v glycine solution to reduce nonspecific binding, and further incubated at 37°C for 16 h. The blocked beads were washed and diluted in lectin storage buffer (20 mM Tris pH 7.4, 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$, 0.5% Triton-X 100, 1 unit protease inhibitor cocktail).

Pulldown using the prepared lectin magnetic beads was performed on the AssayMAP Bravo liquid handler workstation (Agilent Technologies, USA) using one lectin per microplate. Briefly, 50 µL conjugated beads and 100 µL denatured serum was added to each well of a 96-well microtiter plate (Greiner, USA), and glycoprotein capture was performed at 4°C for 1 h under gentle shaking. Post incubation, the conjugated beads with the captured glycoproteins were washed seven times in 50 mM ammonium bicarbonate buffer with two microplate changes to minimise trace detergent, reducing and alkylating agent concentrations. The captured glycoproteins were digested with trypsin at 37°C for 18 h after adding 50 mM ammonium bicarbonate buffer and 1 µg of sequencing grade porcine trypsin (Promega, Australia) to each well. The plate was sealed for enzyme digestion. The trypsin was inactivated with 1% v/v formic acid (FA; Merck, USA) and the digested peptides were collected, dried down in a vacuum concentrator, sealed and stored at −80°C until use.

2.2.3 | Data dependent acquisition mass spectrometry

Shotgun proteomics using data-dependent acquisition was performed on a SCIEX 5600 TripleTOF 5600+ mass spectrometer (SCIEX, USA) coupled to a Shimadzu LC-20AD Prominence nano liquid chromatography system (Shimadzu, Japan). All solvents and reagents were of MS grade (Thermo Fisher, USA). The mass spectrometer was controlled using Analyst 1.7 software (SCIEX, USA). Digested peptides were resuspended in 0.1% v/v FA and injected onto a Protecol C18 analytical column (200 Å, 3 µm, 150 mm × 150 µm, Trajan Scientific, Australia) connected to a Protecol guard column (Polar 120 Å, 3 µm, 10 mm × 300 µm, Trojan Scientific, Australia) and the sample injection order was randomised in the worklist. Column compartment was maintained at 45°C. The peptides were eluted using mobile phase A (0.1% v/v FA) over the specified gradient of mobile phase B (95%

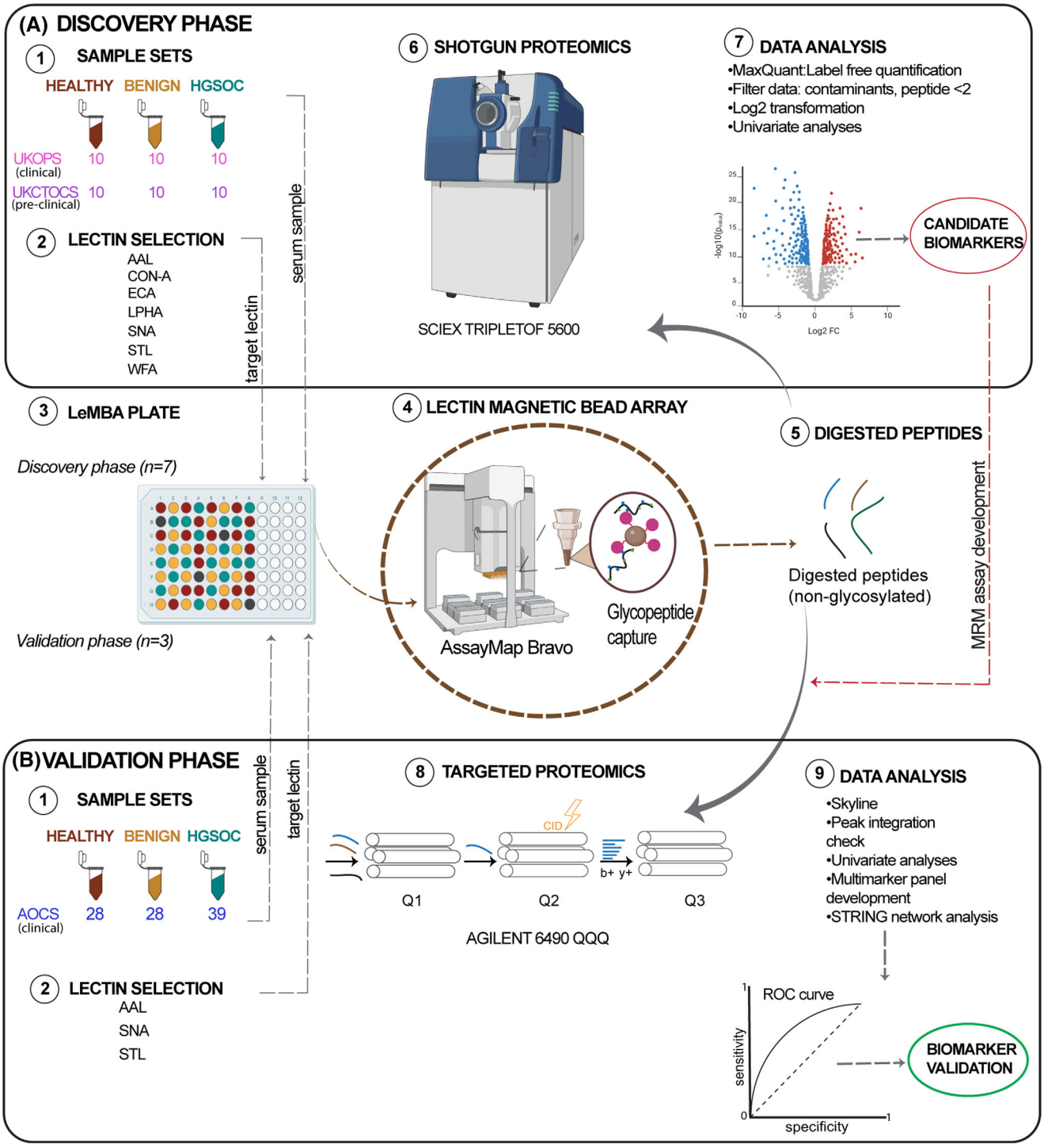


FIGURE 1 Biomarker study design. Discovery and validation of HGSOC biomarkers was conducted in two phases, starting from separate clinical cohorts (1) and sera collection (2). Lectin selection (3) was based on literature for discovery phase and the discovery results for the validation phase. Both phases use LeMBA (4), liquid handler-assisted pulldown (5) and on-bead digestion (6). Shotgun mass spectrometry was conducted for discovery phase (7) followed by discovery of candidates (8) for development of a targeted mass spectrometry assay (9) for validation phase. Both univariate and multivariate analyses were conducted for biomarker validation (10).

acetonitrile, 5% v/v water, 0.1% v/v FA) for 60 min at a flow rate of 1.2 µL/min (5% B at 3 min; 30% B at 37 min; 50% B at 45 min; 100% B at 47 min; 100% B at 51 min; 5% B at 53 min until end of run). The nanospray ion source was set as follows: ion source gas 1 = 35 psi, curtain gas = 30 psi, ion spray floating voltage = 2400 V and interface heater temperature = 180°C. The ion optics parameters were set as declustering potential = 100 V and collision energy (survey scan) = 10 V. Data acquisition was performed using the information dependent acquisition (IDA) and the top 30 precursors from each survey scan were selected for fragmentation. The MS1 spectra was acquired in positive polarity within the mass range = *m/z* 350–1250 Da, with the accumulation time of 250 ms. The precursor selection mass window in the quadrupole was set to unit resolution (*m/z* 0.7 window). The MS/MS spectra were acquired using collision induced dissociation (CID) within the mass range = *m/z* 100–1500 Da with the following parameters: charge states +2 to +5, accumulation time = 100 ms, fragmentation threshold = 150 cps, dynamic exclusion = 15 s and collision energy voltage was set as rolling collision energy with a collision energy spread of 3.

The acquired raw ion spectra for each lectin batch were searched against the reviewed UniProt human proteome database (20,365 proteins, accession date 1 January 2020) using MaxQuant software, v.1.6.6.0 [21]. The MaxQuant contaminant database (247 entries) was also searched to identify contaminants such as keratin. MaxQuant parameters were set as follows: Digestion = trypsin, with two missed cleavages; fixed modification was set to cysteine carbamidomethylation; variable modifications were set as methionine oxidation and N-terminal acetylation; label free quantification (LFQ) was enabled with minimum ratio count set to 2; unique and razor peptides were used for protein identification; match between runs was set as TRUE; and false discovery rate (FDR) for protein and peptide identification was set at 0.01. AB Sciex Q-TOF was set as instrument type using default settings. First search peptide tolerance was set to 0.07 Da and 0.06 Da for the main search. MS/MS tolerance was set at 40 ppm. The search results were imported into R software v1.4.1103 (www.R-project.org) for further data processing and statistical analyses.

2.2.4 | Mass spectrometry data processing and statistical analysis

The generated protein list for each lectin batch was filtered to remove contaminants, reverse identified protein IDs, proteins with < 2 peptide IDs and score < 5. Proteins which were missing in < 25% of all samples were considered missing at random and imputed using localized least square regression (llsimpute) [22]. Proteins missing in > 25% were imputed with the minimum detected value (values drawn randomly from a normal distribution centred at sample minimum and with SD estimated from non-missing proteins). Quantitative analysis was conducted at the protein level using the summed intensity of all peptides mapped to each protein. Log2 transformed data were analysed using the R limma package [23] and Student's T-test. As candidates from the discovery phase will be further confirmed by targeted MS, and application of false discovery rate to the dataset yielded very few significant differences, we shortlisted candidates based on non-adjusted *p*-values. Differentially abundant proteins were visualised by volcano plots, using the criteria *p*-value < 0.05 and Log_2 fold change > 1, and the nomenclature 'lectin-UniProt entry name'. All graphical output has been generated using R or GraphPad Prism v9 (San Diego, USA) and figures prepared using Illustrator v26.3.1 (Adobe Inc, USA) and Biorender (www.biorender.com).

2.3 | Biomarker validation phase

2.3.1 | Multiple reaction monitoring assay development

A custom multiple reaction monitoring (MRM) assay was developed for the list of protein biomarker candidates discovered from UKCTOCS and UKOPS clinical sets, after manually removing immunoglobulins from the list. An initial transition list was selected by matching to a spectral library generated from *in silico* trypsin digest of the discovery phase raw spectral files ($n = 420$) in Skyline v 21.1.0.278 (http://skyline.maccosslab.org/), using human proteome as background proteome. For each candidate biomarker protein, a minimum of 10 peptides were selected, each consisting of at least six transitions (b and y ions). Additionally, the 'Unique Peptides' parameter in Skyline was applied to check for peptide uniqueness to a single protein.

For MRM method optimisation, digested peptide samples from all the lectin batches and across both discovery sets were pooled into a single sample. To monitor retention time across all runs, three stable isotope standard (SIS) peptides (VTSIQDWVQK, NLAVSQVVHK, LSPIYNLVPVK) were added. The final dynamic MRM method consisted of 60 proteins (59 candidate biomarker proteins + one chicken ovalbumin protein), 176 peptides (170 candidate peptides + three SIS peptides + three chicken ovalbumin peptides) and 860 transitions with a delta retention time of 1 min.

2.3.2 | LeMBA-MRM-MS

Candidate biomarker validation was performed on the independent clinical sample set from AOCS. Based on the discovery phase results, AAL, SNA and STL lectins were selected for the validation phase and the serum samples were subjected to the LeMBA workflow as described in the discovery phase. Prior to MS injection, the above three SIS peptides were spiked-in to the samples for monitoring retention time stability across runs. MRM-MS was performed on an Agilent 6490 triple quadrupole mass spectrometer coupled to an Agilent 1290 Infinity UHPLC system, equipped with an Agilent jet stream + ESI source. The mass spectrometer was controlled by MassHunter software (Agilent Technologies, USA). Digested peptides (10 µL, 20 µg) were injected onto a reverse phase AdvanceBio Peptide Mapping analytical column (150 × 2.1 mm i.d., 2.7 µm, part number 653750–902, Agilent Technologies, USA) connected to a 5 mm long guard column and the sample

injection order was randomised in the worklist. The column compartment was maintained at 50°C. The peptides were eluted using mobile phase A (0.1% v/v FA) over the specified gradient of mobile phase B (100% acetonitrile, 0.1% v/v FA) for 35 min at a flow rate of 0.4 mL/min (3% B at 0 min; 30% B at 20 min; 40% B at 24 min; 95% B at 24.5 min; 95% B at 28.5 min and 3% B at 29 min until end of run). The mass spectrometer operated in positive ion mode and the source parameters were set as 150 V high pressure RF, 60 V low pressure RF, 4000 V capillary voltage, 300 V nozzle voltage, 30 psi nebulizer gas flow, 15 L/min drying gas flow at a temperature of 150°C, 11 L/min sheath gas flow at a temperature of 250°C and 200 V delta EMV. The quadrupole was set at unit resolution [0.7 Da full width at half maximum in the first quadrupole (Q1) and the third quadrupole (Q3)], fragmentor at 380 V and cell accelerator voltage at 4 V.

Due to a mass spectrometer software failure, data for the first batch of AAL-pulldown samples were not saved. The entire plate had to be re-run using remaining sample volume, and 25 samples were noted to have lower remaining volume. In addition, an injection problem was noted for one SNA-pulldown sample.

2.3.3 | Data analysis

The data analysis for each lectin MRM-MS was performed independently with no comparisons performed across the three lectins. MRM-MS raw data for the three lectins were exported to Skyline v 21.1.0.278 (downloaded January 2022, http://skyline.maccosslab.org/) to manually check for correct peak integration and the peak area of each measured transition was exported and further analysed in R (v1.4.1103). For each measured peptide in each sample, the transition peak areas were summed and then Log_2 transformed.

Data quality control was conducted using peak area for the three internal standard chicken ovalbumin peptides (Figure S1) and the three spiked-in SIS peptides (Figure S2). While most chicken ovalbumin peptides were consistent, the 25 AAL and one SNA samples with noted aberrations at the MS step showed larger variability, and were removed as outliers (Figure S1, S2). After outlier removal, the calculated %CV for all peptide standards was less than 10% (except peptide AAL-VASMASEK, 11.3%) (Table S3) and the peak area distribution of all samples remaining in the analysis exhibited a normal data distribution (Figure S3). As the dataset has low %CV, we decided normalisation is not required. Student's *t*-test was performed between the case control groups and false discovery rate using Benjamini–Hochberg method was applied to identify significantly differing proteins at *p*-adjusted-value < 0.05. All graphical output has been generated using R or GraphPad Prism and figures prepared using Adobe Illustrator and Biorender.

2.3.4 | Multi-marker panel development

Generalized regression with binomial distribution and lasso estimation was used to develop multi-marker panels using JMP Pro version 16.2.0 (JMP Pro Inc., Carey, NC, USA). Performance of the multi-marker panels were assessed using leave-one-out cross-validation, where area under the receiver operating curve, specificity, sensitivity was calculated on the left-out observations. The number of times each parameter (peptide) was chosen in the cross-validation models is presented as an indication of the relative importance of the markers for prediction. All models were also run using peptides standardised by subtraction of the mean of the three internal standard chicken ovalbumin peptides, but are not reported as they yielded comparable results. Inclusion of the 25 AAL outliers also was tested and led to similar albeit slightly worse prediction models.

3 | RESULTS

3.1 | Discovery of candidate serum glycoprotein biomarkers

In the UKOPS discovery set, we found 15 and 16 differentially abundant proteins when HGSOC samples were compared to benign and healthy samples, respectively (Figure 2, Table 1, Table S4). Compared to benign samples, UKOPS HGSOC samples exhibited an increase in AAL-A1AG1 (alpha-1-acid-glycoprotein 1), AAL-A1AG2 (alpha-1-acid-glycoprotein-2), AAL-FA5 (coagulation factor V), AAL-HGF (hepatocyte growth factor), STL-FIBA (fibrinogen alpha chain), AAL-FIBB (fibrinogen beta chain), LPHA-ITIH3 (inter-alpha-trypsin-inhibitor heavy chain 3), as well as decreased SNA-CO8G (complement component C8 gamma chain), SNA-FA10 (coagulation factor X), LPHA-PON1 serum paraoxonase/arylesterase 1) and STL-PON1 (Figure 2A). When compared to healthy samples, the UKOPS HGSOC samples exhibited an increase in ECA-CO9 (complement component C9) and ECA-LG3BP (galectin-3-binding protein), and a reduction of ECA-C8B (complement component C8B), ECA-CO8G, ECA-LUM (lumican) and ECA-THBG (thyroxine-binding globulin), STL-ACTG (actin), STL-CBG (corticosteroid-binding globulin, STL-CHLE (cholinesterase) and STL-PON1 (Figure 2B).

In the UKCTOCS discovery set, we found 14 and 12 differentially abundant proteins in HGSOC samples when compared to benign and healthy samples, respectively (Figure 2, Table 1, Table S5). Compared to benign samples, UKCTOCS HGSOC samples showed an increase in WFA-ALDOA (fructose bisphosphates aldolase A), WFA-CO9 and WFA-HV307 (immunoglobulin heavy-variable 3–7), SNA-F13B (coagulation factor XIII B chain) and SNA-ZA2G (zinc-alpha-2-glycoprotein), as well as reduction in AAL-MASP1 (mannan-binding lectin serine protease 1), ECA-APOH (beta-2 glycoprotein-1), ECA-CO7 (complement component C7), ECA-CO8G, ECA-FBLN1 (fibulin-1) and ECA-PON1 (Figure 2C). On the other hand, when compared to healthy samples, HGSOC samples showed increased STL-A1AT (alpha-1-antitrypsin), STL-HEP2 (heparin cofactor 2), STL-ITIH3, STL-IGLL5 (immunoglobulin lambda like growth factor), STL-LG3BP, and reduced AAL-MASP1, ECA-CO8G, ECA-FBLN1 and ECA-PGRP2 (peptidoglycan recognition protein-2), WFA-CO7 and WFA-FA12 (coagulation factor XII) (Figure 2D).

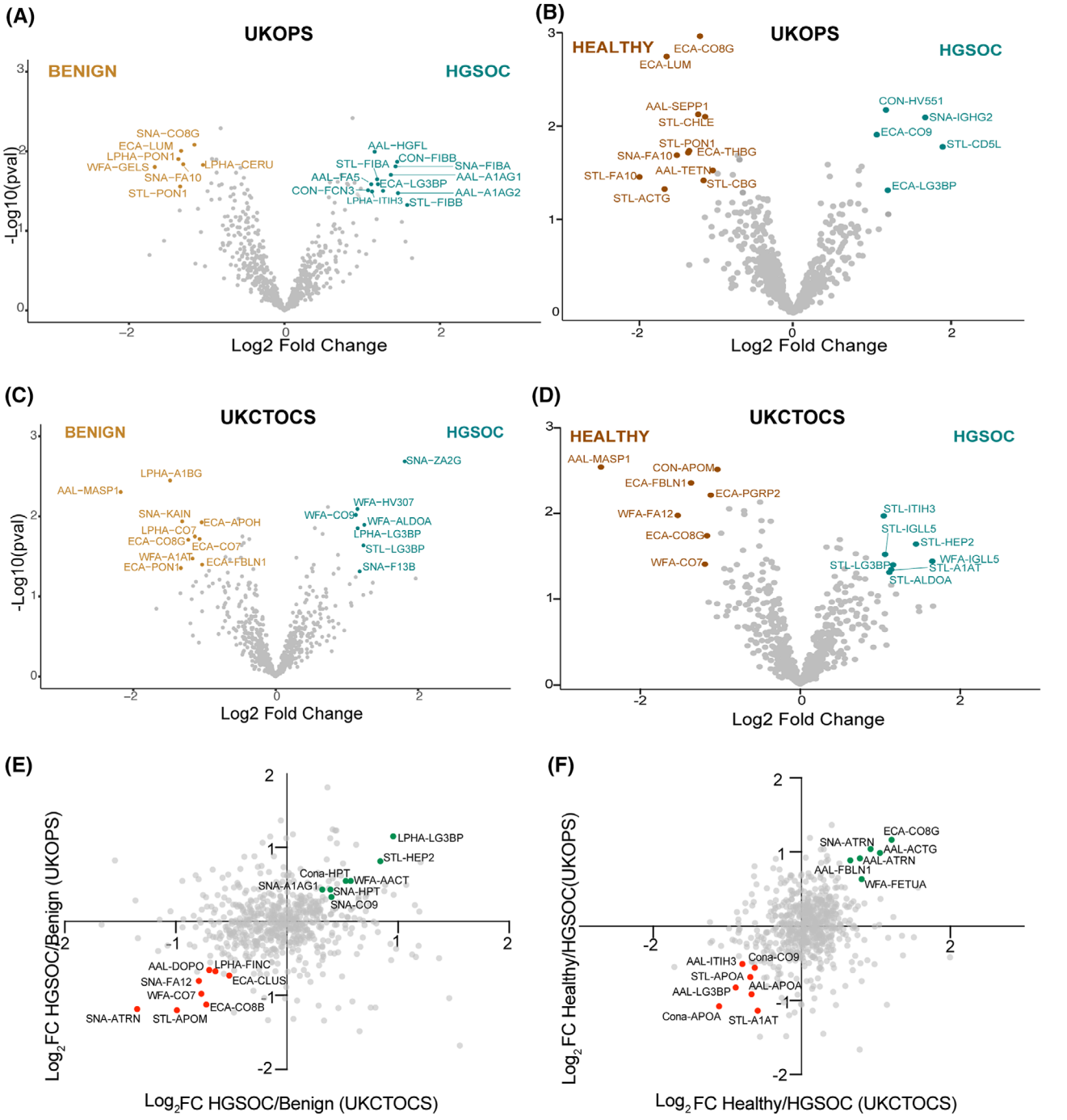


FIGURE 2 Biomarker discovery data. Volcano and two-way scatter plots visualising the differentially abundant proteins and correlated proteins, respectively, between the benign and HGSOC (A, C, E) and healthy and HGSOC (B, D, F) clinical comparisons for UKOPS and UKCTOCS sample sets. The volcano plots highlight all differential glycoproteins according to the criteria $p < 0.05$, Log_2 Fold Change > 1. The scatter plots highlight select glycoproteins (Log_2 Fold Change > 0.5) that are upregulated (green dots) and downregulated (red dots) in both sample sets. All candidates are indicated using the nomenclature 'lectin-Uniprot entry name'.

Additionally, there was an overlap of a subset of candidates that displayed the same expression trend in both the sample sets such as LPHA-LG3BP, STL-HEP2 and SNA-CO9 in the HGSOC versus benign comparison (Figure 2E) and ECA-C08G, SNA-ATRN and STL-A1AT in the healthy versus HGSOC comparison (Figure 2F).

3.2 | Validation of candidate biomarkers

Three lectins (AAL, SNA and STL) with the largest number of candidate proteins discovered in both UKOPS and UKCTOCS sets (Table 1) were selected for validation in the independent clinical AOCS cohort.

TABLE 1 Discovery phase biomarker candidates.

	HGSOC versus Benign		HGSOC versus Healthy		Total number
LECTIN	UKOPS	UKCTOCS	UKOPS	UKCTOCS	
AAL	A1AG1, A1AG2, FA5, HGFL	MASP1	SEPP1	MASP1	6
CON-A	FIBB, FCN3	–	HV551	APOM	4
ECA	LUM, LG3BP	APOH, CO7, CO8G, FBLN1, PON1	CO9, LG3BP, CO8B, CO8G, LUM, THBG	CO8G, FBLN1, PGRP2	11
L-PHA	PON1, CERU, ITIH3	-	–	–	6
SNA	FA10, CO8G, FIBA	KAIN, F13B, ZA2G	IGHG2, CBPN, FA10	–	9
STL	PON1, FIBA, FIBB	LG3BP	CD5L, CHLE, PON1, FA10, ACTG, CBG	A1AT, ITIH3, HEP2, IGLL5, LG3BP	13
WFA	GELS	A1AT, CO9, ALDOA, HV307	-	IGLL5, CO7, FA12	8
Total number	15	15	14	12	

Note: For each lectin, significant proteins with p-value < 0.05 and $\log_2$FC > 1 identified in each clinical cohort have been detailed out below. The total number of candidate proteins across each lectin and clinical cohort accounts for overlaps. The proteins are labelled by their UniProt entry name.

The list of protein candidates discovered from UKOPS and UKCTOCS were combined to generate a list of 44 proteins, which fell short of the target number of ~60 candidate proteins that we previously used for biomarker validation [17, 18]. In order to assess additional candidates which may be just outside of the $p < 0.05$ cut-off, we expanded the selection threshold to p-value < 0.1 and removed Log_2 fold change filtering. This resulted in an initial list of 102 proteins which was filtered down to a final MRM target list of 59 candidates based on suitability of protein tryptic peptides for MRM. The developed custom MRM assay measured 170 peptides from the 59 candidate proteins, with at least three peptides per protein and 4–5 transitions per peptide. The full MRM-MS data are provided in Table S6.

Univariate analysis for HGSOC versus benign and HGSOC versus healthy samples was conducted at the peptide level on each lectin dataset, with significance cut-offs set at adjusted p-value < 0.05 and $\log_2$ fold change > 0.5 (Table S7). Overall, for the HGSOC versus benign comparison, we found 53 (AAL), 80 (SNA) and 49 (STL) peptides with an overlap of 15 common peptides that were mapped to seven proteins (Figure 3A, C, E, G). Likewise, for the HGSOC versus healthy comparison, we found 58 (AAL), 88 (SNA) and 74 (STL) peptides, with an overlap of 38 peptides that were mapped to 18 proteins (Figure 3B, D, F, H). To summarise the peptide differential expression data into HGSOC glycoprotein biomarkers, we next looked for consistency in the peptide differential expression. Each glycoprotein candidate was measured by three or more non-glycosylated peptides, but not all peptides showed significance or consistent direction of change, possibly related to proteoforms or protein cleavage. Filtering for consistent direction of change (up/down) across all measured peptides revealed 21%–31% of proteins had all consistent peptides, with five proteins elevated in HGSOC (A1AT, AACT, CO9, HPT and ITIH3), and four down-regulated proteins (A2MG, ALS, IBP3 and PON1) across the three lectins (Table 2). The validated glycoproteins had diverse lectin-binding affinities, from binding to all three lectins (CO9, A2MG) to a single lectin (AAL-HPT, STL-PON1) (Table 2).

We used STRING v 11.5 to investigate the interactions between the nine validated biomarker proteins. The developed protein-protein interaction (PPI) network had 16 edges (expected 0), and significantly more interactions than expected (PPI enrichment p-value of $<1 \times 10^{-16}$) (Figure S4A). Functional enrichment analyses revealed significant enrichments in Gene Ontology Cellular Component terms Blood microparticle (five out of 115 genes, FDR 1.67×10^{-6}), extracellular exosome (eight out of 2099 genes, FDR 8.21 1.67×10^{-5}), insulin-like growth factor ternary complex (two out of four genes, FDR 0.0048), amongst others, as well as KEGG pathway, Complement and coagulation cascades (three out of 82 genes, FDR 0.0022) (full analysis in Table S8).

3.3 | Development of multi-marker signature for HGSOC

The receiver operating curve (AUC), specificity and sensitivity of the developed multi-peptide models are detailed in Table 3, Table S9. All four models performed similarly with the AAL signature having the highest AUC (87.5%), sensitivity (70.4%) and specificity (90.7%). To further inspect the stable peptides for each of the models, we filtered peptides chosen in at least 50% of the cross-validation runs (Table 3). This analysis revealed several interesting observations. The IBP3 peptide ALAQCAPPPAVCAELVR was always selected for each lectin signature, indicating strong predictive value for HGSOC. Two peptides were highly stable for SNA, STL and the combined signatures, namely, A2MG_NEDSLVFVQTDK and CHLE_NIAAFGGNPK. For the combined signature, both SNA and STL binding IBP3_ALAQCAPPPAVCAELVR were selected with high stability (100% and 91.6%, respectively).

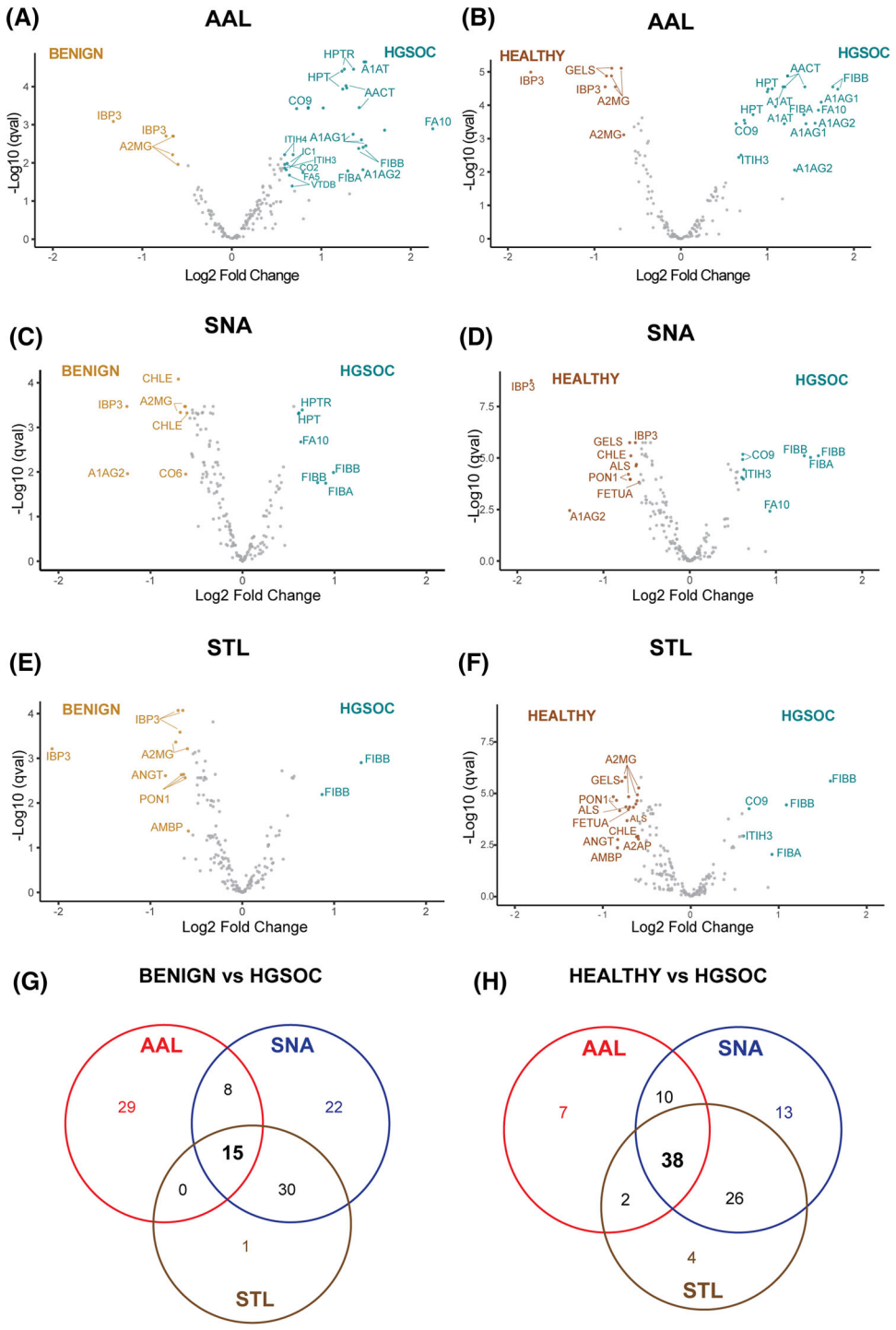


FIGURE 3 Biomarker validation data. LeMBA-MRM data were analysed for differentially abundant peptides between benign and HGSOC (A, C, E, G), and healthy and HGSOC (B, D, F, H) for AAL (A, B), SNA (C, D), and STL (E, F) lectins, respectively. Each dot in the volcano plot indicates a peptide, labelled only by the corresponding UniProt ID for the protein for visualization. The overlap between candidates for each lectin is shown in G, and H.

Strikingly, the stable peptides in the combined signature comprised two SNA and five STL peptides, with no AAL peptides. PPI analysis again revealed significant interactions with enrichment value of 1.29×10^{-7} (Figure S4B), and enrichment of the GO Cellular Component term Blood microparticle (four out of 115 genes, FDR 3.28×10^{-5}), as well as the KEGG pathway Complement and coagulation cascades (four out of 82 genes, FDR 1.74×10^{-6}). Additionally, the GO biological process Blood coagulation, Fibrin clot formation was also highly enriched (four out of 26 genes, FDR 8.62×10^{-7}, Table S8).

TABLE 2 Validated biomarker glycoproteins.

	AAL	SNA	STL
Number of significant peptides	47	41	38
Protein numbers with any significant peptide	23	20	19
Protein numbers with all peptides consistent	5 (21.7%)	6 (30%)	6 (31.6%)
Increased in HGSOC	A1AT, AACT, CO9, HPT	AACT, CO9, ITIH3	CO9, ITIH3
Reduced in HGSOC	A2MG	A2MG, ALS, IBP3	A2MG, ALS, IBP3, PON1

Note: Table shows the number of significant peptides for either HGSOC versus benign or HGSOC versus healthy comparison (q-value < 0.05 and $\log_2$FC > 0.5), the number of proteins with any significant peptide, and the number of proteins with all measured peptides significant and consistent in direction. Protein with all peptides consistent are arranged by alphabetical order of their Uniprot entry name, according to the direction of change in HGSOC.

3.4 | Evaluation of validated biomarker candidates for early HGSOC detections

To determine if any of the nine validated univariate protein biomarkers were altered in the pre-clinical samples, we re-examined the UKCTOCS discovery LeMBA-DDA-MS data set for the nine proteins (Table 2). Three proteins, namely, CO9, ITIH3 and A2MG were significantly altered in the UKCTOCS samples. Figure 4 illustrates the comparative data from discovery UKCTOCS (protein level) and validation (peptide level) phases. AAL-CO9 (Figure 4A) and STL-ITIH3 (Figure 4B) were significantly higher in HGSOC group compared to benign and healthy groups in the UKCTOCS set, while SNA-A2MG was lower in HGSOC group (Figure 4C). STL-CO9 was also elevated in HGSOC versus other groups in the UKCTOCS set, albeit not statistically significant (Figure 4A), while STL-A2MG was not detected in the UKCTOCS or UKOPS datasets likely due to the lower sensitivity of DDA-MS compared to MRM-MS. The three early detection biomarkers CO9, ITIH3 and A2MG interacted in a tight network (enrichment p-value 3.91e−06) (Figure S4C), which was functionally enriched in the KEGG pathway Complement and coagulation cascades (FDR 0.00183) and in the Uniprot annotated keyword of Serine protease inhibitors (FDR 0.0351, Table S8).

4 | DISCUSSION

This study, involving multiple independent clinical and pre-clinical sample sets, provides evidence of glycoproteins as serum biomarkers for HGSOC and lays the foundation for further research on larger patient cohorts. Excitingly, three glycoproteins (CO9, ITIH3 and A2MG) were altered in pre-clinical serum samples collected 11.1 ± 5.1 months prior to HGSOC diagnosis, suggesting promise in early detection. The observed changes in proteins pulled down by the three lectins (AAL, SNA and STL) suggest alterations in α-fucose, sialic acid and N-acetylglucosamine during HGSOC development that require further characterization.

To increase likelihood of successful biomarker development, our glycoprotein-focused biomarker pipeline addressed the issue of technical and biological variations by using LeMBA as a common platform across all phases, and developing multi-marker panels, respectively. The LeMBA-MRM platform also has the advantage of being able to be deployed as a clinical assay [24] reducing the time it takes for the findings to be translated for patient use. Alternatively, lectin-immunoassays can be developed for the discovered biomarkers [25].

We employed a phased biomarker study design to operate within budget [26] where the discovery phase screen uses a relatively small sample size to generate a shortlist of lectins and protein candidates for validation in a larger cohort. In view of the small discovery samples size, our choice of low-stringency statistics on the discovery data, and experimental design to analyse all candidate proteins against the three selected lectins (using the single MRM assay) was ultimately critical for successful biomarker validation. Notably, only one specific lectin-protein discovery phase candidate (STL-PON1, Table 1) was ultimately confirmed in the validation cohort (Table 2). The validated biomarkers were comprised mostly of lectin-protein combinations that were just outside the initial cut-off for the discovery analysis but were analysed in the validation cohort as the custom MRM assay was used on all three selected lectins. This outcome highlights the need for data-specific statistical approaches and considered (non-)use of multiple-testing adjustment in discovery science. Aside from statistics and sample size, biological variation related to evolution of the cancer between the discovery clinical (UKOPS) and validation clinical (AOCS) samples is likely to have contributed to the observed differences.

Strikingly, the validated biomarker proteins physically interact and are functionally enriched in the Complement and coagulation cascades, two components of the innate immune system. This finding is in line with the higher risk of venous thromboembolism (VTE) in cancer patients [27, 28] and the emerging concept of a tumour 'coagulome', a cancer-driven network of molecular effectors favouring thrombosis or bleeding [29, 30]. The analysis of the cancer coagulome found expression of genes encoding six pro-coagulant and fibrinolytic factors (*F3*, *PLAU*, *PLAT*, *PLAUR*, *SERPINB2*, and *SERPINE1)* in The Cancer Genome Atlas [29], and correlated with VTE incidence in 32 cancer types in a previously reported Dutch study [27]. While a moderate correlation was found between VTE risk and expression of Tissue Factor (F3), a major pro-coagulant factor [29], the authors noted heterogeneity

TABLE 3 Signature peptides from lectin-pulldowns for distinguishing HGSOC from benign/healthy and the stability of each peptide assessed by number of times selected during cross-validations.

Lectin (sample size)	AAL ($n = 70$)	SNA ($n = 95$)	STL ($n = 95$)	Combined ($n = 95$)
Model performance (range)	AUC = 87.6% (78.4%, 96.8%)	AUC = 84.6% (76.1%, 93.0%)	AUC = 85.0% (76.6%, 93.4%)	AUC = 84.8% (76.4%, 93.3%)
	Sens = 70.4% (51.5%, 84.1%)	Sens = 64.1% (48.4%, 77.3%)	Sens = 64.1% (48.4%, 77.3%)	Sens = 64.1% (48.4%, 77.3%)
	Spec = 90.7% (78.4%, 96.3%)	Spec = 87.5% (76.4%, 93.8%)	Spec = 85.7% (74.3%, 92.6%)	Spec = 89.3% (78.5%, 95.0%)
UniProt Name_Peptide	**Times selected**	**Times selected**	**Times selected**	**Times selected**
IBP3_ALAQCAPPPAVCAELVR	70 (100%)	95 (100%)	95 (100%)	SNA 95 (100%) STL 87 (91.6%)
A1AT_QINDYVEK	70 (100%)			
HPT_VTSIQDWVQK	69 (98.5%)		83 (87.4%)	
IBP3_FLNVLSPR	69 (98.5%)			
GELS_QTQVSVLPEGGETPLFK	68 (97.1%)			
FETUA_HTLNQIDEVK	64 (91.4%)			
FIBB_YQISVNK	46 (65.7%)			
FA10_ETYDFDIAVLR		94 (98.9%)		
A2MG_NEDSLVFVQTDK		93 (97.9%)	95 (100%)	STL 95 (100%)
CHLE_NIAAFGGNPK		93 (97.9%)	67 (70.5%)	SNA 92 (96.8%)
KNG1_ENFLFLTPDCK		93 (97.9%)		
ITIH2_VQSTITSR		90 (94.7%)		
ITIH3_EHLVQATPENLQEAR		86 (90.5%)	91 (95.8%)	
C1QC_QTHQPPAPNSLIR		85 (89.5%)		
HPT_DYAEVGR		84 (88.4%)		
THRB_SGIECQLWR		79 (83.2%)		
KNG1_EGDCPVQSGK			95 (100%)	STL 95 (100%)
FIBB_QDGSVDFGR			78 (82.1%)	
FA10_IVGGQECK				STL 66 (69.5%)
FIBB_SILENLR				STL 60 (63.2%)

Abbreviations: AUC, area under the receiver operating curve; Sens, sensitivity; Spec, specificity.

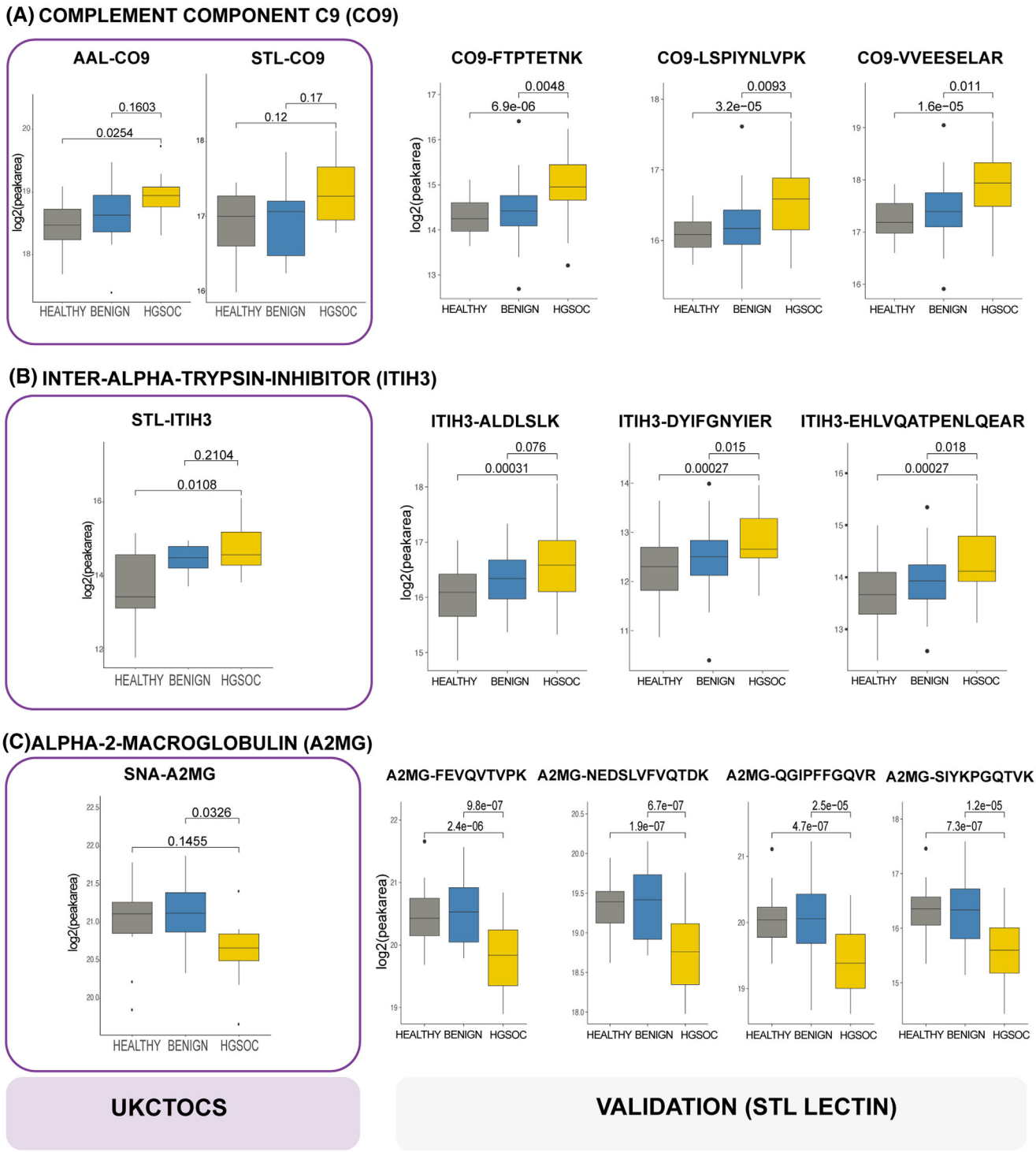


FIGURE 4 Potential HGSOC early detection biomarkers. Comparison of glycoprotein biomarker data between UKCTOCS discovery case control set (left panel, LeMBA-DDA-MS protein level data), and STL-pulldown validation case control set (right panel, LeMBA-MRM-MS, peptide level data) for (A) CO9, (B) ITIH3, and (C) A2MG. For comparison, the unadjusted *p*-values are shown for both data sets.

across tumour types. Intriguingly, while ovarian cancer had one of the highest VTE incidences in this study, the six examined genes showed moderate expression levels in this cancer [29], indicating important roles of other mechanisms and procoagulant factors. Indeed, our findings identified altered glycosylation of two additional proteins not yet considered in the cancer coagulome, namely, thrombin (THRB) and coagulation factor X (FA10, Table 3), which should be evaluated for inclusion in the coagulome.

Proteins of the Complement and coagulation cascades interact at several levels, share common regulator proteins and both systems act on immune and endothelial cells [31, 32]. Interestingly, both cascades are activated or regulated by extracellular vesicles [30, 32], small membrane-enclosed vesicles released by cells for inter-cellular communication. Our biomarkers were enriched in the Gene Ontology term 'Blood microparticles', defined as a type of extracellular vesicle (EV) devoid of nucleic acids, released from several cell types including platelets, blood cells and endothelial cells. This finding is supported by two previous independent cohort studies reporting elevated serum/plasma EV procoagulant activity in ovarian cancer patients [33, 34]. While a recent meta-analysis reports VTE to be higher in advanced serous ovarian cancer, clear cell histology and ascites at diagnosis [35], our study has found evidence of perturbations in the glycosylation of Complement and coagulation cascade several months prior to ovarian cancer diagnosis.

We were highly encouraged to find three validated biomarkers to be altered several months prior to cancer diagnosis in the UKCTOCS set. Levels of two candidates were elevated (glycoforms of C9 and ITIH3) and one was lower (SNA-A2MG) compared to healthy and benign samples. All three proteins are associated with inflammation, with C9 and A2MG having known roles in the Complement and coagulation pathways. Of the three potential early detection biomarkers, only A2MG has previously been reported to be altered in ovarian cancer. Similar to our findings, Miyamoto et al. reported decreased A2MG protein in ovarian cancer patients compared to healthy controls [36]. A2MG is a broad-spectrum protease inhibitor which inhibits thrombin and the complement pathway [37]. Reduced A2MG levels may indicate elevated Coagulation and complement pathway activity, although the role of the SNA-A2MG glycoform is currently unknown.

C9 is the terminal Component of the complement cascade, which has also been found to be elevated in serum of gastric [38], lung [39], colorectal [40, 41] and esophageal cancers [17, 18] through proteomics or glycoproteomic approaches. In our esophageal adenocarcinoma biomarker study, C9 glycoforms binding to each of the six shortlisted lectins (AAL, EPHA, JAC, NPL, PSA, WGA) was also significantly elevated in esophageal adenocarcinoma compared to the precursor benign condition, Barrett's esophagus but with some variability in healthy groups [18]. The current study shortlisted three lectins (AAL, SNA, STL) for the validation phase, and all three C9 glycoforms measured were significantly elevated in HGSOC, suggesting that C9 is subjected to increased aberrant glycosylation that may contribute to disease progression. Interestingly, AAL-C9 glycoform levels have been reported to be elevated in lung and stomach cancers, intermediate in hepatocellular carcinoma and low in breast cancer [39]. Recently, we reported the release of $C9^{+}$ EVs by esophageal adenocarcinoma cells as a potential mechanism of the elevated serum C9 glycoform in esophageal cancer [42]. However, the specific glycosylation differences in cancer serum and EVs, as well as the molecular mechanisms underpinning the altered C9 glycosylation in different cancer types remains to be determined.

Inter-alpha-trypsin inhibitor family members, including ITIH3, have been implicated in inflammation and carcinogenesis [43]. In addition to its protease inhibitor activity, ITIH3 is thought to stabilize the extracellular matrix through binding to hyaluronic acid [44]. Interestingly, previous reports suggest blood ITIH3 to be elevated for a similar range of cancers as C9, namely, lung [45], gastric [46], pancreatic [47] and colorectal [48] cancers.

In conclusion, we report the discovery and validation of serum glycoprotein markers for HGSOC using a lectin-assisted proteomics workflow that is directly translatable to blood tests. The validated markers show high specificity when bench-marked against the existing ovarian cancer biomarker, CA125. Their utility in ovarian cancer diagnosis and monitoring will need to be evaluated in additional cohorts, such as at diagnosis and following surgery/chemotherapy. Furthermore, several markers were elevated months prior to cancer diagnosis, and should be further evaluated as for ovarian cancer screening. Functional enrichment of the validated markers highlights blood microparticle (EV)-mediated complement and coagulation activity ahead of clinical diagnosis of HGSOC. Further investigation on the contribution of EV-mediated complement and coagulation in ovarian cancer development may provide mechanisms for prevention.

ACKNOWLEDGEMENTS

We are grateful for the expert technical assistance from UQ Centre for Clinical Research Mass Spectrometry Facility, particularly Mr. Buddhika Jayakody Arachchige. We also thank Dr. Thomas Stoll (QIMR Berghofer Medical Research Institute) for assistance in uploading proteomics data to repository. The AOCS acknowledges the cooperation of the participating institutions in Australia and acknowledges the contribution of the study nurses, research assistants and all clinical and scientific collaborators to the study. The complete AOCS Study Group can be found at www.aocstudy.org. The UCL team acknowledges the contribution of the volunteers without whom the trial would not have been possible, and everyone involved in conduct and oversight of UKCTOCS/UKOPS. We would like to thank all of the women who participated in these research programs. The project was funded by the Ovarian Cancer Research Foundation (Australia) to MMH (GA-2018-04). MMH's salary was supported by the QIMR Berghofer Medical Research Institute. The Australian Ovarian Cancer Study Group was supported by the U.S. Army Medical Research and Materiel Command under DAMD17-01-1-0729, The Cancer Council Victoria, Queensland Cancer Fund, The Cancer Council New South Wales, The Cancer Council South Australia, The Cancer Council Tasmania and The Cancer Foundation of Western Australia (Multi-State Applications 191, 211 and 182) and the National Health and Medical Research Council of Australia (NHMRC; ID199600; ID400413 and ID400281). The Australian Ovarian Cancer Study gratefully acknowledges additional support from Ovarian Cancer Australia and the Peter MacCallum Foundation. The UKOPS study was funded by The Eve Appeal (The Oak Foundation). UKCTOCS was funded by Medical Research Council (G9901012 and G0801228), CRUK (C1479/A2884), and the Department of Health, with additional support from The Eve Appeal. Usha

Menon has salary support from the NIHR University College London Hospitals Biomedical Research Centre and Usha Menon, Sophia Apostolidou, Aleksandra Gentry-Maharaj are supported by MRC CTU at UCL core funding (MC_UU_00004/01).

CONFLICT OF INTEREST STATEMENT

Usha Menon had shares (2011–2020) awarded by UCL in Abcodia which had an interest in early detection of cancer. SA had salary support from Abcodia between 2011 and 2021. Usha Menon, Sophia Apostolidou and Aleksandra Gentry-Maharaj are involved in institutional research collaborations with iLOF, Micronoma, RNA Guardian, Mercy Bioanalytics and Synteny Biotechnology. All other authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The mass spectrometry proteomics data for the discovery phase have been deposited to the ProteomeXchange Consortium via the PRIDE [49] partner repository with the dataset identifier PXD032299. The targeted mass spectrometry data for the validation phase have been deposited to PRIDE via Panorama Public database with the dataset identifier PXD033108.

REFERENCES

1. Siegel, R. L., Miller, K. D., & Jemal, A. (2020). Cancer statistics, 2020. *CA: A Cancer Journal for Clinicians*, *70*, 7–30. https://doi.org/10.3322/caac.21590
2. Menon, U., Gentry-Maharaj, A., Burnell, M., Singh, N., Ryan, A., Karpinskyj, C., Carlino, G., Taylor, J., Massingham, S. K., Raikou, M., Kalsi, J. K., Woolas, R., Manchanda, R., Arora, R., Casey, L., Dawnay, A., Dobbs, S., Leeson, S., Mould, T., … Parmar, M. (2021). Ovarian cancer population screening and mortality after long-term follow-up in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet*, *397*, 2182–2193. https://doi.org/10.1016/S0140-6736(21)00731-5
3. Pinsky, P. F., Yu, K., Kramer, B. S., Black, A., Buys, S. S., Partridge, E., Gohagan, J., Berg, C. D., & Prorok, P. C. (2016). Extended mortality results for ovarian cancer screening in the PLCO trial with median 15years follow-up. *Gynecologic Oncology*, *143*, 270–275. https://doi.org/10.1016/j.ygyno.2016.08.334
4. Henderson, J. T., Webber, E. M., & Sawaya, G. F. (2018). Screening for ovarian cancer. *Jama*, *319*, 595–606. https://doi.org/10.1001/jama.2017.21421
5. Nash, Z., & Menon, U. (2020). Ovarian cancer screening: Current status and future directions. *Best Practice & Research. Clinical Obstetrics & Gynaecology*, *65*, 32–45. https://doi.org/10.1016/j.bpobgyn.2020.02.010
6. Pinho, S. S., & Reis, C. A. (2015). Glycosylation in cancer: mechanisms and clinical implications. *Nature Reviews Cancer*, *15*, 540–555. https://doi.org/10.1038/nrc3982
7. Chen, K., Gentry-Maharaj, A., Burnell, M., Steentoft, C., Marcos-Silva, L., Mandel, U., Jacobs, I., Dawnay, A., Menon, U., & Blixt, O. (2013). Microarray glycoprofiling of CA125 improves differential diagnosis of ovarian cancer. *Journal of Proteome Research*, *12*, 1408–1418. https://doi.org/10.1021/pr3010474
8. Gidwani, K., Huhtinen, K., Kekki, H., Van Vliet, S., Hynninen, J., Koivuviita, N., Perheentupa, A., Poutanen, M., Auranen, A., Grenman, S., Lamminmäki, U., Carpen, O., Van Kooyk, Y., & Pettersson, K. (2016). A Nanoparticle-Lectin Immunoassay Improves Discrimination of Serum CA125 from Malignant and Benign Sources. *Clinical Chemistry*, *62*, 1390–1400. https://doi.org/10.1373/clinchem.2016.257691
9. Bayoumy, S., Hyytiä, H., Leivo, J., Talha, S. M., Huhtinen, K., Poutanen, M., Hynninen, J., Perheentupa, A., Lamminmäki, U., Gidwani, K., & Pettersson, K. (2020). Glycovariant-based lateral flow immunoassay to detect ovarian cancer–associated serum CA125. *Communications Biology*, *3*, 460. https://doi.org/10.1038/s42003-020-01191-x
10. Biskup, K., Braicu, E. I., Sehouli, J., Fotopoulou, C., Tauber, R., Berger, M., & Blanchard, V. (2013). Serum glycome profiling: a biomarker for diagnosis of ovarian cancer. *Journal of Proteome Research*, *12*, 4056–4063. https://doi.org/10.1021/pr400405x
11. Dědová, T., Braicu, E. I., Sehouli, J., & Blanchard, V. (2019). Sialic acid linkage analysis refines the diagnosis of ovarian cancer. *Frontiers in Oncology*, *9*, 261. https://doi.org/10.3389/fonc.2019.00261
12. Hu, Y., Pan, J., Shah, P., Ao, M., Thomas, S. N., Liu, Y., Chen, L., Schnaubelt, M., Clark, D. J., Rodriguez, H., Boja, E. S., Hiltke, T., Kinsinger, C. R., Rodland, K. D., Li, Q. K., Qian, J., Zhang, Z., Chan, D. W., Zhang, H., … Shi, Z. (2020). Integrated proteomic and glycoproteomic characterization of human high-grade serous ovarian carcinoma. *Cell Reports*, *33*, 108276.
13. Shetty, V., Hafner, J., Shah, P., Nickens, Z., & Philip, R. (2012). Investigation of ovarian cancer associated sialylation changes in N-linked glycopeptides by quantitative proteomics. *Clinical Proteomics*, *9*, 10. https://doi.org/10.1186/1559-0275-9-10
14. Sinha, A., Hussain, A., Ignatchenko, V., Ignatchenko, A., Tang, K. H., Ho, V. W. H., Neel, B. G., Clarke, B., Bernardini, M. Q., & Ailles, L., Kislinger, T. (2019). N-glycoproteomics of patient-derived xenografts: A strategy to discover tumor-associated proteins in high-grade serous ovarian cancer. *Cell Systems*, *8*, 345–351.e4 e344.
15. Wu, J., Xie, X., Liu, Y., He, J., Benitez, R., Buckanovich, R. J., & Lubman, D. M. (2012). Identification and confirmation of differentially expressed fucosylated glycoproteins in the serum of ovarian cancer patients using a lectin array and LC–MS/MS. *Journal of Proteome Research*, *11*, 4541–4552. https://doi.org/10.1021/pr300330z
16. Choi, E., Loo, D., Dennis, J. W., O'leary, C. A., & Hill, M. M. (2011). High-throughput lectin magnetic bead array-coupled tandem mass spectrometry for glycoprotein biomarker discovery. *Electrophoresis*, *32*, 3564–3575. https://doi.org/10.1002/elps.201100341
17. Shah, A. K., Hartel, G., Brown, I., Winterford, C., Na, R., Cao, K.-A. L., Spicer, B. A., Dunstone, M. A., Phillips, W. A., Lord, R. V., Barbour, A. P., Watson, D. I., Joshi, V., Whiteman, D. C., & Hill, M. M. (2018). Evaluation of serum glycoprotein biomarker candidates for detection of esophageal adenocarcinoma and surveillance of Barrett's esophagus. *Molecular & Cellular Proteomics*, *17*, 2324–2334. https://doi.org/10.1074/mcp.RA118.000734
18. Shah, A. K., Cao, K.-A. L., Choi, E., Chen, D., Gautier, B., Nancarrow, D., Whiteman, D. C., Saunders, N. A., Barbour, A. P., Joshi, V., & Hill, M. M. (2015). Serum glycoprotein biomarker discovery and qualification pipeline reveals novel diagnostic biomarker candidates for esophageal adenocarcinoma. *Molecular & Cellular Proteomics*, *14*, 3023–3039. https://doi.org/10.1074/mcp.M115.050922
19. Oungsakul, P., Choi, E., Shah, A. K., Mohamed, A., O'leary, C., Duffy, D., Hill, M. M., & Bielefeldt-Ohmann, H. (2021). Candidate glycoprotein biomarkers for canine visceral hemangiosarcoma and validation using semi-quantitative lectin/immunohistochemical assays. *Veterinary Sciences*, *8*, 38. https://doi.org/10.3390/vetsci8030038
20. Menon, U., Gentry-Maharaj, A., Hallett, R., Ryan, A., Burnell, M., Sharma, A., Lewis, S., Davies, S., Philpott, S., Lopes, A., Godfrey, K., Oram, D., Herod, J., Williamson, K., Seif, M. W., Scott, I., Mould, T., Woolas, R., Murdoch, J., … Jacobs, I. (2009). Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: Results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS).

The Lancet Oncology, *10*, 327–340. https://doi.org/10.1016/S1470-2045(09)70026-9

21. Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*, 1367–1372. https://doi.org/10.1038/nbt.1511
22. Kim, H., Golub, G. H., & Park, H. (2005). Missing value estimation for DNA microarray gene expression data: Local least squares imputation. *Bioinformatics*, *21*, 187–198. https://doi.org/10.1093/bioinformatics/bth499
23. Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*, e47. https://doi.org/10.1093/nar/gkv007
24. Bringans, S., Duong, M., Laming, C., Di Prinzio, P., Hill, M., & Lipscombe, R. (2022). Establishing a mass spectrometry based diagnostic test for oesophageal cancer. *Diseases of the Esophagus*, *35*, doac051203. https://doi.org/10.1093/dote/doac051.203
25. Webster, J. A., Wuethrich, A., Shanmugasundaram, K. B., Richards, R. S., Zelek, W. M., Shah, A. K., Gordon, L. G., Kendall, B. J., Hartel, G., Morgan, B. P, Trau, M., & Hill, M. M. (2021). Development of endo-screen chip, a microfluidic pre-endoscopy triage test for esophageal adenocarcinoma. *Cancers (Basel)*, *13*, 2865.
26. Rifai, N., Gillette, M. A., & Carr, S. A. (2006). Protein biomarker discovery and validation: The long and uncertain path to clinical utility. *Nature Biotechnology*, *24*, 971–983. https://doi.org/10.1038/nbt1235
27. Blom, J. W., Vanderschoot, J. P. M., Oostindiër, M. J., Osanto, S., Van Der Meer, F. J. M., & Rosendaal, F. R. (2006). Incidence of venous thrombosis in a large cohort of 66 329 cancer patients: Results of a record linkage study. *Journal of Thrombosis and Haemostasis*, *4*, 529–535. https://doi.org/10.1111/j.1538-7836.2006.01804.x
28. Stein, P. D., Beemath, A., Meyers, F. A., Skaf, E., Sanchez, J., & Olson, R. E. (2006). Incidence of venous thromboembolism in patients hospitalized with cancer. *American Journal of Medicine*, *119*, 60–68. https://doi.org/10.1016/j.amjmed.2005.06.058
29. Saidak, Z., Soudet, S., Lottin, M., Salle, V., Sevestre, M.-A., Clatot, F., & Galmiche, A. (2021). A pan-cancer analysis of the human tumor coagulome and its link to the tumor immune microenvironment. *Cancer Immunology, Immunotherapy*, *70*, 923–933. https://doi.org/10.1007/s00262-020-02739-w
30. Galmiche, A., Rak, J., Roumenina, L. T., & Saidak, Z. (2022). Coagulome and the tumor microenvironment: an actionable interplay. *Trends in Cancer*, *8*, 369–383. https://doi.org/10.1016/j.trecan.2021.12.008
31. Oncul, S., & Afshar-Kharghan, V. (2020). The interaction between the complement system and hemostatic factors. *Current Opinion in Hematology*, *27*, 341–352. https://doi.org/10.1097/MOH.0000000000000605
32. Ratajczak, M. Z., & Ratajczak, J. (2021). Innate immunity communicates using the language of extracellular microvesicles. *Stem Cell Reviews and Reports*, *17*, 502–510. https://doi.org/10.1007/s12015-021-10138-6
33. Claussen, C., Rausch, A.-V., Lezius, S., Amirkhosravi, A., Davila, M., Francis, J. L., Hisada, Y. M., Mackman, N., Bokemeyer, C., Schmalfeldt, B., Mahner, S., & Langer, F. (2016). Microvesicle-associated tissue factor procoagulant activity for the preoperative diagnosis of ovarian cancer. *Thrombosis Research*, *141*, 39–48. https://doi.org/10.1016/j.thromres.2016.03.002
34. Zhang, W., Peng, P., Ou, X., Shen, K., & Wu, X. (2019). Ovarian cancer circulating extracelluar vesicles promote coagulation and have a potential in diagnosis: An iTRAQ based proteomic analysis. *BMC Cancer*, *19*, 1095. https://doi.org/10.1186/s12885-019-6176-1
35. Weeks, K. S., Herbach, E., Mcdonald, M., Charlton, M., & Schweizer, M. L. (2020). Ovarian cancer patients by stage, histology, cytoreduction, and ascites at diagnosis. *Obstetrics and Gynecology International*, *2020*, 1.
36. Miyamoto, S., Stroble, C. D., Taylor, S., Hong, Q., Lebrilla, C. B., Leiserowitz, G. S., Kim, K., & Ruhaak, L. R (2018). Multiple reaction monitoring for the quantitation of serum protein glycosylation profiles: application to ovarian cancer. *Journal of Proteome Research*, *17*, 222–233. https://doi.org/10.1021/acs.jproteome.7b00541
37. Vandooren, J., & Itoh, Y. (2021). Alpha-2-macroglobulin in inflammation, immunity and infections. *Frontiers in Immunology*, *12*, 803244.
38. Chong, P.-K., Lee, H., Loh, M. C. S., Choong, L.-Y., Lin, Q., So, J. B. Y., Lim, K. H., Soo, R. A., Yong, W. P., Chan, S. P., Smoot, D. T., Ashktorab, H., Yeoh, K. G., & Lim, Y. P. (2010). Upregulation of plasma C9 protein in gastric cancer patients. *Proteomics*, *10*, 3210–3221. https://doi.org/10.1002/pmic.201000127
39. Narayanasamy, A., Ahn, J.-M., Sung, H.-J., Kong, D.-H., Ha, K.-S., Lee, S.-Y., & Cho, J.-Y. (2011). Fucosylated glycoproteomic approach to identify a complement component 9 associated with squamous cell lung cancer (SQLC). *Journal of Proteomics*, *74*, 2948–2958. https://doi.org/10.1016/j.jprot.2011.07.019
40. Chantaraamporn, J., Champattanachai, V., Khongmanee, A., Verathamjamras, C., Prasongsook, N., Mingkwan, K., Luevisadpibul, V., Chutipongtanate, S., & Svasti, J. (2020). Glycoproteomic analysis reveals aberrant expression of complement C9 and fibronectin in the plasma of patients with colorectal cancer. *Proteomes*, *8*, 26.
41. Murakoshi, Y., Honda, K., Sasazuki, S., Ono, M., Negishi, A., Matsubara, J., Sakuma, T., Kuwabara, H., Nakamori, S., Sata, N., Nagai, H., Ioka, T., Okusaka, T., Kosuge, T., Shimahara, M., Yasunami, Y., Ino, Y., Tsuchida, A., Aoki, T., … Yamada, T. (2011). Plasma biomarker discovery and validation for colorectal cancer by quantitative shotgun mass spectrometry and protein microarray. *Cancer Science*, *102*, 630–638. https://doi.org/10.1111/j.1349-7006.2010.01818.x
42. Kolka, C. M., Webster, J., Lepletier, A., Winterford, C., Brown, I., Richards, R. S., Zelek, W. M., Cao, Y., Khamis, R., Shanmugasundaram, K. B., Wuethrich, A., Trau, M., Brosda, S., Barbour, A., Shah, A. K., Eslick, G. D., Clemons, N. J., Morgan, B. P, & Hill, M. M. (2022). C5b-9 membrane attack complex formation and extracellular vesicle shedding in barrett's esophagus and esophageal adenocarcinoma. *Frontiers in Immunology*, *13*, 842023.
43. Hamm, A., Veeck, J., Bektas, N., Wild, P. J., Hartmann, A., Heindrichs, U., Kristiansen, G., Werbowetski-Ogilvie, T., Del Maestro, R., Knuechel, R., & Dahl, E. (2008). Frequent expression loss of Inter-alpha-trypsin inhibitor heavy chain (ITIH) genes in multiple human solid tumors: a systematic expression analysis. *BMC Cancer*, *8*, 25. https://doi.org/10.1186/1471-2407-8-25
44. Huang, L., Yoneda, M., & Kimata, K. (1993). A serum-derived hyaluronan-associated protein (SHAP) is the heavy chain of the inter alpha-trypsin inhibitor. *Journal of Biological Chemistry*, *268*, 26725–26730. https://doi.org/10.1016/S0021-9258(19)74373-7
45. Heo, S.-H., Lee, S.-J., Ryoo, H.-M., Park, J.-Y., & Cho, J.-Y. (2007). Identification of putative serum glycoprotein biomarkers for human lung adenocarcinoma by multilectin affinity chromatography and LC-MS/MS. *Proteomics*, *7*, 4292–4302. https://doi.org/10.1002/pmic.200700433
46. Chong, P. K., Lee, H., Zhou, J., Liu, S.-C., Loh, M. C. S., Wang, T. T., Chan, S. P., Smoot, D. T., Ashktorab, H., So, J. B. Y., Lim, K. H., Yeoh, K. G., & Lim, Y. P. (2010). ITIH3 is a potential biomarker for early detection of gastric cancer. *Journal of Proteome Research*, *9*, 3671–3679. https://doi.org/10.1021/pr100192h
47. Liu, X., Zheng, W., Wang, W., Shen, H., Liu, L., Lou, W., Wang, X., & Yang, P. (2017). A new panel of pancreatic cancer biomarkers discovered using a mass spectrometry-based pipeline. *British Journal of Cancer*, *117*, 1846–1854. https://doi.org/10.1038/bjc.2017.365
48. Kopylov, A. T., Stepanov, A. A., Malsagova, K. A., Soni, D., Kushlinsky, N. E., Enikeev, D. V., Potoldykova, N. V., Lisitsa, A. V., & Kaysheva, A. L. (2020). Revelation of proteomic indicators for colorectal cancer in initial stages of development. *Molecules (Basel, Switzerland)*, *25*, 619.

49. Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D. J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., Walzer, M., Wang, S., Brazma, A., & Vizcaíno, J. A. (2022). The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Research*, *50*, D543–D552. https://doi.org/10.1093/nar/gkab1038

SUPPORTING INFORMATION

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

Report

TRANSPARENT PROCESS OPEN ACCESS

EMBO Molecular Medicine

The human host response to monkeypox infection: a proteomic case series study

Ziyue Wang[1,†], Pinkus Tober-Lau[2,†], Vadim Farztdinov[3,†], Oliver Lemke[1], Torsten Schwecke[1], Sarah Steinbrecher[2], Julia Muenzner[1], Helene Kriedemann[2], Leif Erik Sander[2,4], Johannes Hartl[1], Michael Mülleder[3,†], Markus Ralser[1,4,5,*,†] & Florian Kurth[2,†]

Abstract

The rapid rise of monkeypox (MPX) cases outside previously endemic areas prompts for a better understanding of the disease. We studied the plasma proteome of a group of MPX patients with a similar infection history and clinical manifestation typical for the current outbreak. We report that MPX in this case series is associated with a strong plasma proteomic response among nutritional and acute phase response proteins. Moreover, we report a correlation between plasma proteins and disease severity. Contrasting the MPX host response with that of COVID-19, we find a range of similarities, but also important differences. For instance, CFHR1 is induced in COVID-19, but suppressed in MPX, reflecting the different roles of the complement system in the two infectious diseases. Of note, the spatial overlap in response proteins suggested that a COVID-19 biomarker panel assay could be repurposed for MPX. Applying a targeted protein panel assay provided encouraging results and distinguished MPX cases from healthy controls. Hence, our results provide a first proteomic characterization of the MPX human host response and encourage further research on protein-panel assays in emerging infectious diseases.

Keywords host response; monkeypox virus; proteomics; viral infection
Subject Categories Microbiology, Virology & Host Pathogen Interaction; Proteomics
DOI 10.15252/emmm.202216643 | Received 25 July 2022 | Revised 9 September 2022 | Accepted 9 September 2022 | Published online 28 September 2022
EMBO Mol Med (2022) 14: e16643

Introduction

The outbreak of monkeypox (MPX) with currently more than 40,000 confirmed infections worldwide, is exceptional in scale and spread (Kraemer *et al*, 2022), and has been declared a global emergency by the WHO (World Health Organisation, 2022a). MPX is caused by the zoonotic monkeypox virus (MPXV), a member of the genus *Orthopoxvirus* (World Health Organisation, 2022b). The first human MPX case was reported in 1970 in the Democratic Republic of the Congo (DRC), which is still the region with the highest level of endemicity in Africa (Bunge *et al*, 2022). Several outbreaks have been reported from African countries during the past decades, but research on MPX has largely been neglected. The clinical presentation often includes typical skin lesions, fever, and swollen lymph nodes. MPX is usually self-limiting, but severe cases can occur and a case fatality rate of 1–10% has been reported from Africa, with generally higher case fatality associated with infections from the Central African viral clade compared to the West African virus clade (Bunge *et al*, 2022).

The molecular epidemiology of the current MPX outbreak suggests that the current strain is closely related to that of a 2018–2019 outbreak in the United Kingdom and may have been circulating in the human population for some time, possibly with adaptation to the human host (Isidro *et al*, 2022; World Health Organisation, 2022c). In the current outbreak, there is a clear predominance of infections among men who have sex with men (MSM), and several large public events have been associated with the rapid emergence of cases in different parts of the world. Currently, transmission via close skin and mucosal contact, possibly including sexual transmission, seems likely (Dye & Kraemer, 2022; European Centre for Disease Prevention and Control, 2022; Pfäfflin *et al*, 2022; Thornhill *et al*, 2022). Even though the current outbreak is still in its early stages, a self-limiting course cannot be assumed; rather, it is a longer-term public-health problem that will hopefully bring diagnostic and therapeutic benefits to endemic African countries.

The COVID-19 pandemic has reminded us of the need to create infrastructure and methodologies to respond rapidly to emerging pathogens. Mass spectrometry-based proteomics is one of the emerging technologies in this regard, which due to the technical and analytical advances during the last years is increasingly moving into clinical applications (Liotta *et al*, 2001; Messner *et al*, 2020; Struwe *et al*, 2020; He *et al*, 2022). In the early phase of the COVID-19 pandemic, proteomic analyses provided rapid insights into the nature of

1 Department of Biochemistry, Charité – Universitätsmedizin Berlin, Berlin, Germany
2 Department of Infectious Diseases and Respiratory Medicine, Charité – Universitätsmedizin Berlin, Berlin, Germany
3 Core Facility High Throughput Mass Spectrometry, Charité – Universitätsmedizin Berlin, Berlin, Germany
4 Berlin Institute of Health, Berlin, Germany
5 The Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK
*Corresponding author. Tel: +49 30 45052814; E-mail: markus.ralser@charite.de
†These authors contributed equally to this work

the human response to SARS-CoV-2 and captured hallmarks of its immune evasion strategies and pathophysiology, including its impact on the complement system, coagulation cascade, and inflammatory and nutritional response machinery (D'Alessandro *et al*, 2020; Messner *et al*, 2020; Shen *et al*, 2020; Demichev *et al*, 2021; Overmyer *et al*, 2021; Nuñez *et al*, 2022). Furthermore, proteomic signatures turned out to classify disease severity in COVID-19 and allow for outcome prediction weeks in advance (Völlmy *et al*, 2021; Demichev *et al*, 2022; Nuñez *et al*, 2022). Recently, we were able to show the strength of mass spectrometry-based proteomics for rapid translation to medical care by generating a routine-applicable proteomic biomarker panel which predicted COVID-19 severity and outcome in a multicohort study (Wang *et al*, 2022a). While such proteomic assays are currently primarily used to monitor clinical trials, they are increasingly being considered for their potential to optimize treatment and resource allocation, as well as to aid navigation of difficult triaging situations in the event of a pandemic.

Here, we describe the proteomic changes in a case series, a small but characteristic group of patients hospitalized due to MPXV infection that share a similar disease and infection history. We detect significant and consistent proteomic changes caused by MPXV infection, enabling us to characterize the MPX host response at the proteomic level despite the moderate cohort size of a case series, in a timely manner. We report several protein markers that correlate with disease severity in the tested cases, that classify the disease proteome, and that contrast the human host response of MPXV to that of SARS-CoV-2 infection. Because we detected a partial overlap between the MPX and COVID-19 host response proteome, we also used a targeted proteomic panel developed for COVID-19 (Wang *et al*, 2022a) to explore the possibility of repurposing existing biomarker panels for classifying newly emerging infections. Although our results are derived from a small number of cases, they nonetheless suggest that repurposing of multiplex panel assays might be a viable strategy to improve pandemic preparedness. Our case series study provides a biochemical characterization of the MPX host response and reveals correlation of host proteins with MPX disease severity, and expands knowledge on protein panel testing for emerging infections.

Results

MPX patient case series and clinical presentation

A group of five patients were hospitalized at Charité University Hospital between 26th and 31st May 2022 for treatment of MPX, detected by PCR from cutaneous blisters. Interestingly, all patients had attended the same social event 10–14 days before developing symptoms, three of whom considered it most likely to have been infected on that occasion. We then included a 6th patient with an unrelated infection history who was hospitalized in mid-June 2022, but that otherwise had a related disease history. All six patients were of European descent, and all self-identified as men having sex with men (MSM) having practiced receptive anal sexual intercourse within 14 days prior to hospitalization. The group of patients was therefore notably homogeneous regarding history and time course of infection, triggering our interest in a case series study.

Overall, MPX patients exhibited mild to moderate symptoms, and no severe systemic affections such as encephalitis, myocarditis, or kidney failure were observed. Prodromes included fever, myalgia, and fatigue, and had already subsided in all patients by the time of admission to the hospital. The number of MPX skin lesions ranged from 5 to 36 and there were no clinical or laboratory signs of organ dysfunction. In all patients, the chief complaint and cause of hospitalization was severe anal or perianal pain requiring systemic analgesics in addition to topical treatment. Samples for proteome measurements were taken at a median of 8 days after symptom onset. Comorbidities included HIV (n = 2, both well controlled on antiretroviral therapy), other STIs (n = 1), and hepatitis C (n = 1). Patients were discharged with alleviated symptoms after 3–6 days. A summary of clinical characteristics is given in Table 1.

The partial sequence of the genome of the MPXV isolate obtained from one of the patients was determined and is available on GenBank (ON813251.2).

To gain maximum information from the case series cohort, we assembled two control cohorts. The first consisted of 15 age- and sex-matched healthy volunteers (Table EV2). Ten patients with SARS-CoV-2 infection, hospitalized due to moderate COVID-19 (grade 3 on the 8-point WHO ordinal scale, i.e., without the need for supplemental oxygen therapy), constituted the second control group. Their proteomes were measured within the same batch on our MS platforms, but had also been analyzed by us as part of a previous study (Demichev *et al*, 2021).

Table 1. Patient characteristics.

	MPX cases (n = 6)	
Male, *n* (%)	6	100%
Age, years	31	IQR: 27–41; range: 26–49
BMI, kg/m^2	22.0	IQR: 19.6–23.4; range: 17.6–25.1
Comorbidities, *n* (%)	3	50%
HIV, *n* (%)	2	33%
Hepatitis C, *n* (%)	1	17%
Other STIs[a], *n* (%)	1	17%
Δ symptom onset to sample, days	8	IQR: 5–14; range: 5–17
Δ PCR to sample, days	3.5	IQR 1.5–5; range: 0–5
Fever, *n* (%)	6	100%
Number of lesions	9	IQR: 5–20; range: 5–36
Duration of hospital stay, days	3.5	IQR: 3–5; range: 3–6
C-reactive protein at admission, mg/l	20.0	IQR: 10.4–57.9; range: 8.7–120.8
Leukocytes at admission, per nl	9.7	IQR: 8.3–11.7; range: 8.1–12.9
Lymphocytes at admission, per nl	3.1	IQR: 1.6–3.7; range: 1.4–3.8
Lactate dehydrogenase, U/l	214	IQR: 203–273; range: 181–381

BMI, body mass index; HIV, human immunodeficiency virus infection; STI, sexually transmitted infection.
Data are presented as median and IQR; range, unless otherwise specified.
[a]Other STIs: co-infection with Neisseria gonorrhoeae, Ureaplasma, and Mycoplasma hominis.

A plasma proteomic signature of MPXV infection

Because of the moderate size of the case series study, we focused on obtaining maximally precise proteomic measurements and contrasted against both control groups. For obtaining proteomic measurements, we prepared tryptic digests from the MPX cases, matched healthy controls, and patients with moderate COVID-19, and included a broad panel of stable-isotope-labeled internal standards (PQ500, Biognosys). The tryptic digests obtained were then recorded using an online coupling of microflow chromatography and Zeno SWATH DIA, a latest generation of DIA proteomic technology (preprint: Wang *et al*, 2022b). Indeed, to our knowledge, the present study represents the first biomedical application of Zeno SWATH MS. After data were recorded as a single batch, raw data were processed with DIA-NN (Demichev *et al*, 2020), and data were post-processed to detect differentially concentrated proteins as well as the enrichment of pathway terms using pathway definitions from REACTOME (Croft *et al*, 2014). A workflow diagram of the procedures is provided (Fig 1A).

Considering the relatively mild severity of clinical symptoms and skin manifestation, the data revealed a substantial proteomic response to MPXV infection within the abundant "functional fraction" of the plasma proteome. This proteome fraction constitutes more than 99% of the plasma proteomic mass and is composed of around 300 proteins, most of which directly function in the plasma (Anderson & Anderson, 2002). As 200–300 of them are consistently quantified using high-throughput proteomics in neat plasma (Messner *et al*, 2020), and because this fraction contains more than 50 typical protein biomarkers (Demichev *et al*, 2021) that capture host physiological parameters (Vernardis *et al*, 2022), this functional fraction of the plasma proteome is of special interest for the development of clinical assays (Wang *et al*, 2022a). After pre-processing,

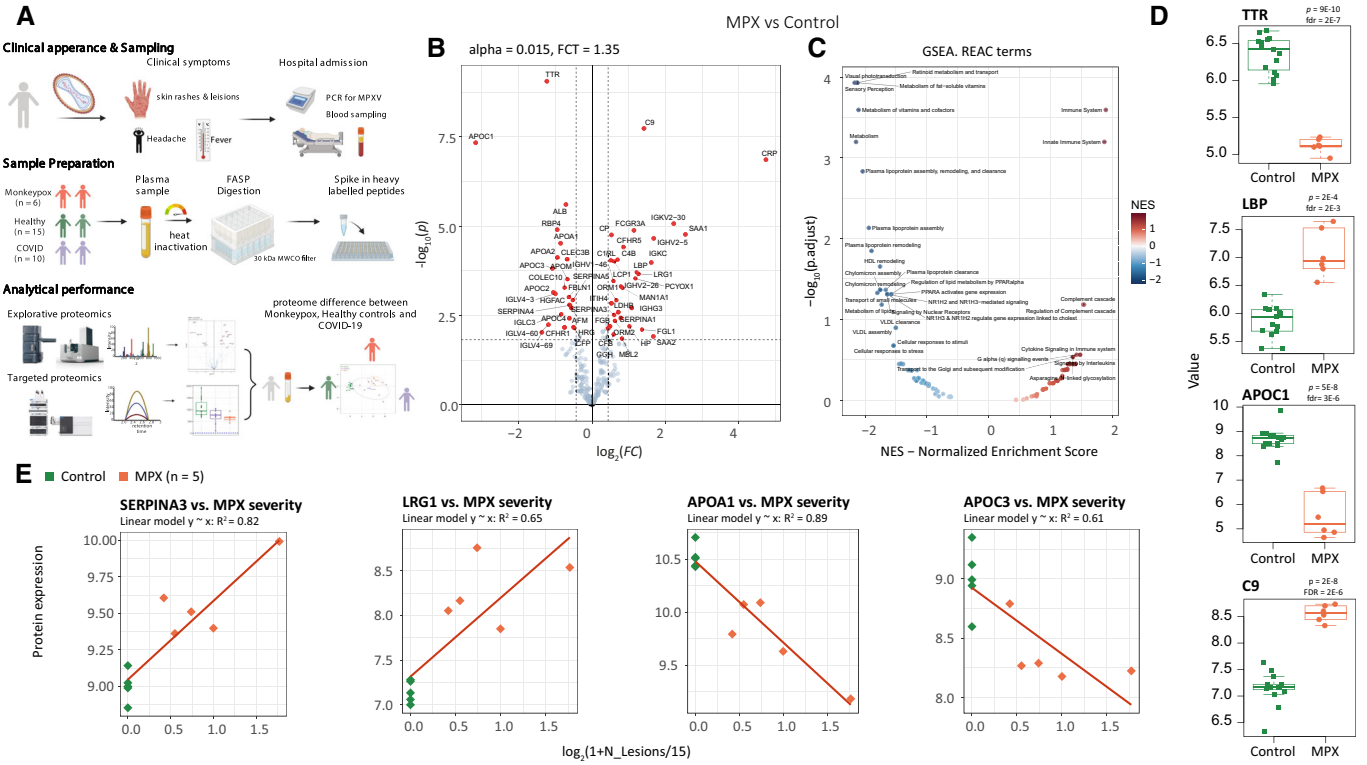


Figure 1. The human host response to monkeypox virus infection determined at the level of the plasma proteome.

A Schematic overview of the workflow using discovery proteomics (Zeno SWATH MS (Wang *et al*, 2022b)) in parallel to a targeted proteomic assay that quantifies COVID-19 severity biomarkers (Wang *et al*, 2022a) to characterize the plasma proteome in an MPX case series, and compare the proteomes to those of healthy volunteers and COVID-19 patients.

B Volcano plot of contrast MPX vs healthy controls; $\alpha <= 0.015$ and $|logFC| >= 1.35$ were used for selection of regulated proteins.

C Gene set analysis (GSEA) of REACTOME (Croft *et al*, 2014) terms enrichment for contrast MPX vs control. Y-axis shows $-\log_{10}$ of adjusted *P*-value (fdr) for Normalized Enrichment Score (x-axis) for each term. Terms with fdr <= 0.3 are labeled.

D Boxplots illustrating key proteins that differ between patients with MPX and controls (*P*-values and fdr for corresponding contrast MPX vs Control are provided in brackets): TTR (*P*-value = 9E-10, fdr = 2E-7), LBP (*P*-value = 2E-4, fdr = 2E-3), APOC1 (*P*-value = 5E-8, fdr = 3E-6), and C9 (*P*-value = 2E-8, fdr = 2E-6). Here, as usual, the central bar marks the median (second quartile), the bottom edge of the box marks the first quartile, the top edge of the box marks the third quartile, and the bottom and top whiskers mark the minimum and maximum values that are not outliers. The specific values of the protein expressions are also shown. Provided *P*-values are obtained from moderated statistics implemented in limma, dfrs were calculated according to Benjamini-Hochberg.

E Correlation between MPX severity ($N_{Skin\ lesions}$) and protein expression (y-axis). One MPX patient had an unclear additional skin condition (not a pure case of MPX) and therefore was excluded from the regression analysis that compares the number of skin lesions with the proteome; however, the proteome of this patient was largely in agreement with those of the other MPX cases (Fig EV3). As a measure of MPX severity, the $\log_2(1 + N_{Lesions} / 15)$ was used. Here $N_{Lesions}$ is the number of lesions. R^2 shows squared correlation coefficient. MPX patients are colored orange, control patients green.

226 of the highly abundant proteins were found consistently quantified in the neat plasma sample. We detected low within-group coefficients of variation, below 25% for MPX and control, and about 34% for COVID-19 cases, indicating a high quantitative precision of the measurements, but also the presence of a biological signal (Fig EV1C). Indeed, we found 56 of the major plasma proteins to be differentially abundant in MPX patients compared to healthy controls. Twenty-four of these were lower concentrated in MPX, and 32 detected at a higher concentration (Fig 1B). The nature of the affected proteins indicated the molecular processes affected by MPX, as revealed by an enrichment analysis. For example, we see "immune system" and "regulation of complement cascade" mostly enriched among upregulated pathways. Among downregulated pathways, "plasma lipoprotein assembly" and "metabolism of fat-soluble vitamins" are enriched (Fig 1C).

At the level of individual proteins, the greatest differences between cases and controls were found in proteins associated with the acute phase response. These included significantly lower levels of the negative acute phase proteins TTR, ALB, and RBP4, as well as higher levels of acute phase proteins CRP, SAA1, SERPINA3, LBP, CP, and LRG1. Of note, various proteins involved in hepatic lipid metabolism and nutrient transport (APOA1, APOA2, APOC1, APOC2, APOC3) were lower in MPX patients than in controls, a known but not fully understood phenomenon also observed in other infections (Hardardóttir *et al*, 1995) (Fig 1D). Compared to controls, MPX patients exhibited a significantly higher level of complement component 9, the main element of the channel part of the membrane attack complex. Also, TTR in combination with the differentially expressed apolipoproteins is noteworthy, as it is a marker for malnutrition (Dellière *et al*, 2018), and we recently found it as a rapid responder in a caloric-restriction experiment conducted with healthy volunteers (Vernardis *et al*, 2022). We first speculated that acute MPX could result in a reduced caloric intake in affected patients. However, this picture was not confirmed by the clinical records of our patients, indicating that TTR is also part of the host response. We did not observe a significant influence of the concomitant conditions such as HIV or hepatitis C on the plasma proteomes. Results of the plasma proteomic response in patients with and without concomitant HIV infection are shown in Fig EV2. Both patients with HIV had immunologically well-controlled infections with suppressed viral load. Nevertheless, these patients can exhibit signs of ongoing immune activation, but if this response to HIV infection was present, it was masked by the acute response of the plasma proteome to the acute MPXV infection.

Next, we tested whether there is a relationship between the proteomic response and the number of skin lesions observed in our patients, determined as a proxy of disease severity. Several peptides showed a statistically robust correlation with the number of lesions, including the upregulated acute phase proteins SERPINA3, SAA1, and LRG1, as well as the downregulated apolipoproteins APOA1, APOA2, and APOC3 (Fig 1E). In particular, LRG1, an upstream modifier of TGF-beta signaling, is being increasingly recognized as an important contributor to disease pathogenesis and hence as a potential therapeutic target in a range of inflammatory conditions (Camilli *et al*, 2022). Despite the moderate size of the case series, our data suggests a consistent proteomic response in MPX cases that reflects the extent of skin manifestation and disease severity in MPX. Our case series did not contain severe cases. We can hence not predict the proteomic profile expected in severe cases, but our results suggest that with a more severe disease, stronger proteomic changes might become prevalent.

Relationship and intersection of the acute phase proteomic responses of MPX and COVID-19

The plasma proteome has similarly been shown to distinguish between different degrees of disease severity in other viral infections, including Ebola (Viodé *et al*, 2022) and COVID-19 (D'Alessandro *et al*, 2020; Shen *et al*, 2020; Demichev *et al*, 2021, 2022; Nuñez *et al*, 2022). To investigate to which degree this classification is due to a similar or divergent set of protein markers, we compared the MPX proteome response to that of an age- and sex-matched group of patients with moderately symptomatic COVID-19 (hospitalized, but without need of supplemental oxygen). The proteome obtained for these two patient groups revealed both an overlap in some response proteins and differences between the host responses against the two viral pathogens in other proteins. A simple hierarchical clustering based on Ward's agglomeration of Euclidean distances clearly separated healthy controls from MPX and COVID-19 cases (Fig 2A), and a protein expression analysis revealed differentially expressed proteins that are common between both diseases, but also those that differentiate the two infections from each other (Fig 2B (central part of the cloud), full-scale figure in Fig EV4A). Consistently, a principal component analysis (PCA) separated both patient groups (and controls), indicating that despite an overlap in several factors, the proteomes are discriminatory between MPX and COVID-19 (Fig 2C).

Contrasting the signatures at the protein level revealed that of the 56 proteins differentially expressed in MPX cases compared to healthy controls, 37 are also differentially expressed in COVID-19 patients with the same direction of regulation (Fig EV4A, Venn diagram). These include 12 proteins of the acute phase response such as SAA1 and LBP, and 12 proteins involved in coagulation, including FGB and SERPINA4, all of which have been found to be differentially expressed depending on COVID-19 disease severity.

Furthermore, we found 19 proteins that were differentially abundant in MPX but not in COVID-19. For instance, LCP1 (logFC(MPX-Control) = 0.6 ± 0.1, logFC(COVID-19-Control) = 0 ± 0.1), and LDHB (logFC(MPX-Control) = 0.7 ± 0.2, logFC(COVID-19-Control) = -0.1 ± 0.2) were found to be only upregulated in MPX (Fig 2D and E). LCP1 is interesting, because as L-plastin, it has been associated with membrane dynamics and the cytoskeleton and is an early tumor marker in kidney cancer (Ralser *et al*, 2005; Su Kim *et al*, 2013). Another protein that triggered our attention was CFHR1 (logFC(MPX-Control) = -0.8 ± 0.3, logFC(COVID-19-Control) = 0.4 ± 0.2), an inhibitor of the terminal pathway of the complement cascade, which was downregulated in MPX but was upregulated in COVID-19, where it is a marker of disease severity (D'Alessandro *et al*, 2020; Shen *et al*, 2020; Demichev *et al*, 2021). Indeed, hyperactivation of the complement system has been shown as a key feature for the pathophysiology of COVID-19 (Georg *et al*, 2022), but according to our proteome data, it is less important in MPX. Of note, it is plausible that additional differentially abundant proteins can be identified by proteomics in other or larger cohorts.

We deemed our case series too small to construct a robust classifier that identifies MPX cases on the basis of their proteome.

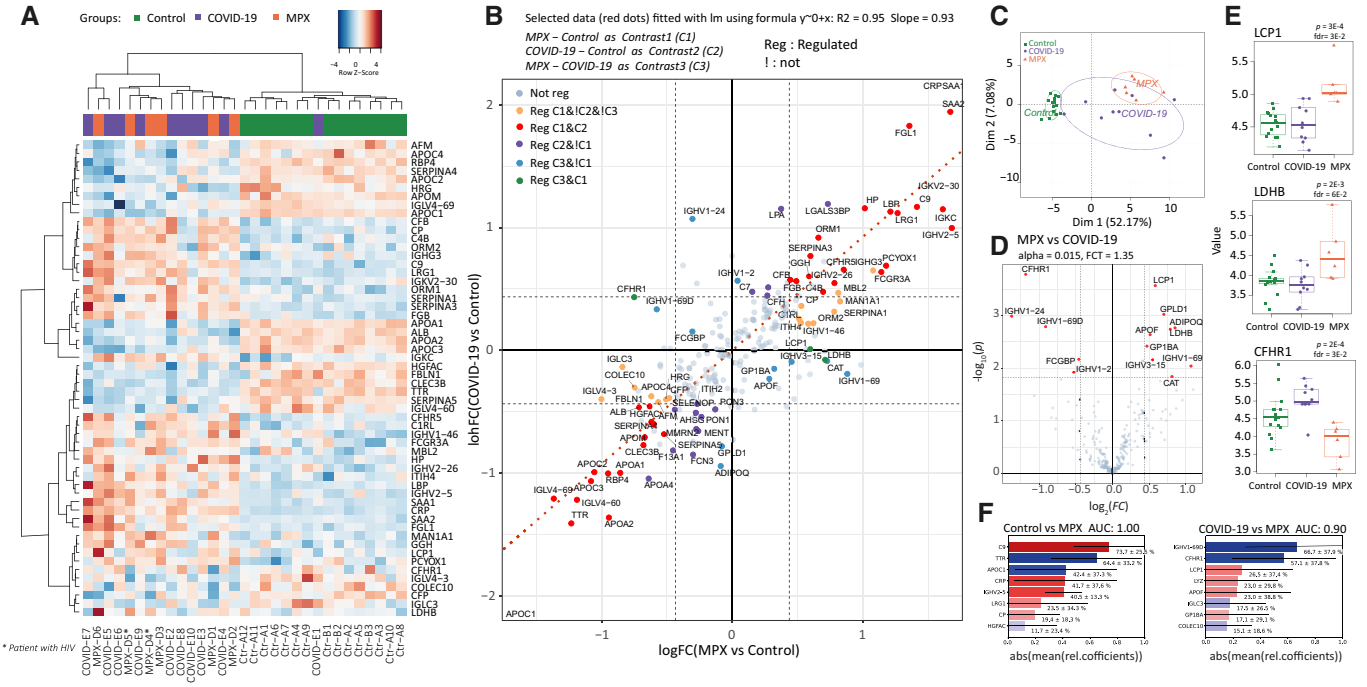


Figure 2. Differences and similarities between the plasma proteome upon infection with MPXV and SARS-CoV-2.

A Heatmap displaying hierarchical clustering using differentially regulated proteins between patients with MPX, COVID-19, and controls.

B Scatterplot of log fold-change (logFC) for contrast MPX vs control (C1, x-axis) and logFC for contrast COVID-19 vs control (C2, y-axis). Only the central part of the cloud is shown here. Three truncated dots (APOC1, CRP, and SAA1) are shown in the lower left and upper right corner. A full-scale figure is presented in Fig EV4A. Differentially abundant (Regulated; "Reg") proteins are color coded, with the red color corresponding to 37 proteins differentially abundant in both MPX vs control (C1) and COVID-19 vs control (C2), the orange color corresponding to proteins specifically changed in MPX vs control (C1) only (16 proteins), and the green color corresponding to proteins responding to both MPX vs control (C1) and MPX vs COVID-19 (C3) (3 proteins). There are no intersections between COVID-19 vs control (C2) and MPX vs COVID-19 (C3). The blue color corresponds to proteins responding to MPX vs COVID-19 (C3), but not in MPX vs control (C1) (11 proteins), and the pink color to proteins responding to COVID-19 vs control (C2) only (19 proteins). The red dotted line shows a linear regression through the red dots, i.e., proteins differentially abundant in MPX vs control (C1) and COVID-19 vs control (C2). Note that orange and pink points have the same direction of regulation in both MPXV vs control (C1) and COVID-19 vs control (C2). Only green and blue dots (except three proteins: ADIPOQ, GPLD1, and IGHV1-2) have opposite directions in C1 and in C2.

C *Post hoc* PCA score plot using proteins shown in (a).

D Differentially regulated proteins of patients with MPX and COVID-19 (Volcano Plot); $\alpha <= 0.015$ and $|logFC| >= 1.35$ were used for selection of regulated proteins. The chosen significance level ensured that fdr for this contrast was below 22%.

E Key proteins that differ between patients with MPX and COVID-19 (Boxplots) (*P*-values and fdr for corresponding contrast MPX vs COVID-19 are provided in brackets): LCP1 (*P*-value = 3E-4, fdr = 3E-2), LDHB (*P*-value = 2E-3, fdr = 6E-2), CFHR1 (*P*-value = 2E-4, fdr = 3E-2). Here the central bar marks the median (second quartile), the bottom edge of the box marks the first quartile, the top edge of the box marks the third quartile, and the bottom and top whiskers mark the minimum and maximum values that are not outliers. The specific values of the protein expressions are also shown. Provided *P*-values are obtained from moderated statistics implemented in limma, dfrs were calculated according to Benjamini-Hochberg.

F Top 8 proteins of an SVM-trained model discriminating between healthy controls ($n = 15$) and MPX cases ($n = 6$) (left) or COVID-19 ($n = 10$) and MPX ($n = 6$) cases (right). Means of the relative coefficients over a 5-fold cross-validation are shown. Error bars denote the standard deviations. Red denotes positive, blue denotes negative coefficients. The AUC was calculated based on withheld samples that were not used for training the model.

However, we explored our data to see whether such an approach should be encouraged, and used a machine learning classifier to complement PCA and differential protein expression analysis, in the characterization of the MPX host response. Therefore, we tested a strictly cross-validated classifier to distinguish between MPX cases and healthy controls, as well as between MPX and COVID-19 cases on the basis of their proteomes. Within our data set, we achieved for both cases a differentiation with a high AUC on the test data that were withheld during training (Fig 2F). Encouragingly, the top-ranked differentiators identified by the machine-learning algorithm were also among the most differentially expressed proteins, like C9 (logFC(MPX-Control) = 1.4 ± 0.2) and TTR (logFC(MPX-Control) = −1.2 ± 0.1) for differentiating MPX cases and healthy individuals, or CHFR1 (logFC(MPX-COVID-19) = −1.2 ± 0.3) or LCP1 (logFC (MPX-COVID-19) = 0.6 ± 0.1) for differentiating MPX from COVID-19, respectively (Fig 2F). We note that due to the limited number of MPX cases in the case series, we can currently not validate the transferability of the model to other data and cohorts. Our data suggests however that the construction of such models appears feasible, once larger cohort data is available. It is further noteworthy that longitudinal proteome analysis in COVID-19 revealed a spike in proteomic response in the early disease phase, triggered by the inflammatory response, and that this early response signature was most predictive of outcome (Demichev *et al*, 2021). These results suggest that future studies should also follow the proteomic response to MPX in a longitudinal fashion.

Hence, our data provide a differentiated picture of the acute proteomic response that follows the two viral infections. On the one hand, we describe various acute phase proteins responding to both COVID-19 and MPX; on the other hand, both viral infections exhibit distinct proteomic response patterns, for instance, concerning the activation of the complement system. Hence, proteomics was effective in obtaining valuable insights even from a case series study.

Potential to repurpose proteomic assays to rapidly respond to emerging viral infections

Due to the partial overlap between the COVID-19 and MPX host responses, we speculated that there might be a potential to repurpose COVID-19 biomarker panel tests to MPX. We recently demonstrated the translational potential of plasma proteomics for applicability in clinical practice through the transfer of protein marker candidates which had been identified by discovery proteomics in COVID-19 into a routinely applicable targeted protein panel assay. The assay absolutely quantifies up to 50 peptides derived from 30 COVID-19-related plasma biomarker proteins and captured hallmarks of COVID-19 in a multi-cohort observational study conducted using routine-lab-compatible high-flow chromatography and LC-MRM acquisition (Wang *et al*, 2022a). The LC–MRM assay consistently quantified 32 of the peptides in plasma samples from MPX cases, controls, and in COVID-19 patient samples. Despite the assay being developed to quantify COVID-19 severity, a PCA on the peptides quantified also separated MPX patient samples from controls (Fig 3A). Moreover, a hierarchical clustering of the protein quantities that differed between healthy controls and MPX cases classified the disease samples (Fig 3B). This separation was driven by differential plasma levels of several proteins involved in the inflammatory and immune-mediated host response, e.g., increased levels of SERPINA3 and LYZ, or decreased levels of TF, TTR, HRG, PGLYRP2, and APOA1 (Fig 3C). Based on this proteomics data, we tested a classifier that distinguished between MPX cases and healthy controls within our data set to obtain a feature importance (Fig 3D). The most important features of the classifier (e.g., SERPINA3, AFM, PGLYRP2, and TF) overlapped with the differentially concentrated proteins in the COVID-19 plasma samples. Hence, within the limitations that our study is based on a small case series, our data suggest there is sufficient overlap among host response proteins among different viral diseases so that upon the disease specific adaptation of the statistical models, the biomarker panel assay could be adapted and repurposed to different viral infections, for instance, to improve pandemic preparedness.

Discussion

As case numbers rise, the current knowledge gap on the molecular etiology of MPX—a disease that has been known in central Africa

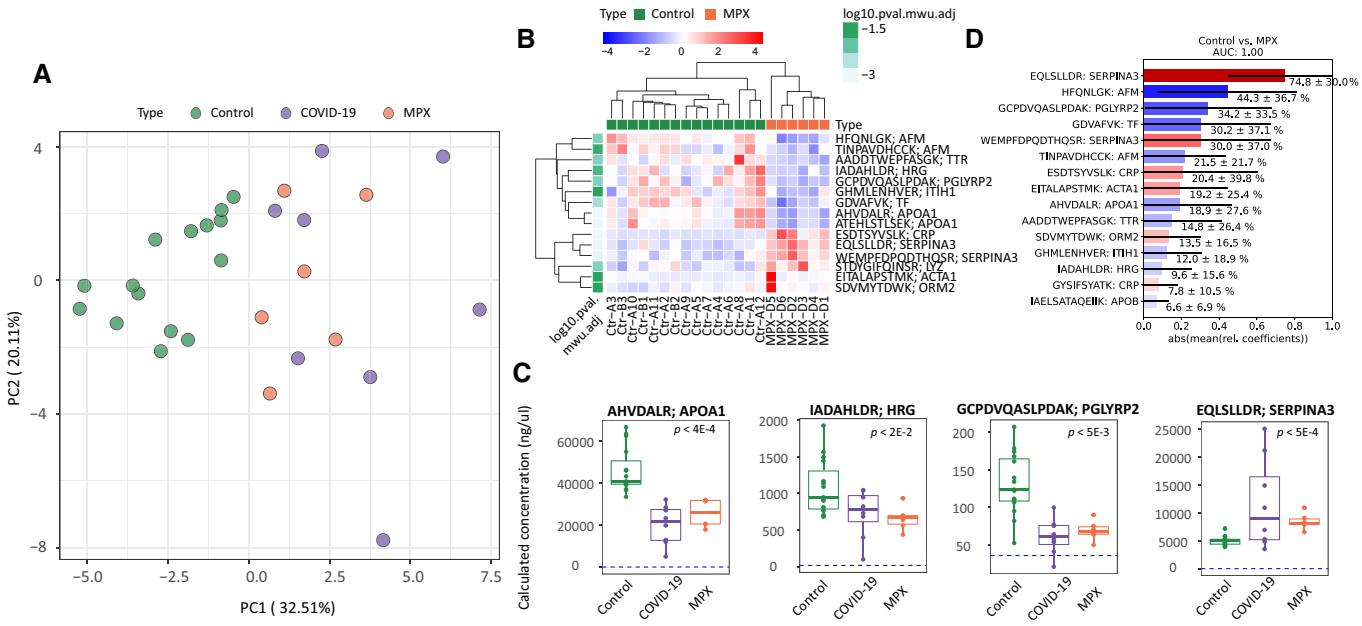


Figure 3. A targeted, multi-protein panel assay developed for COVID-19 infection discriminates patients with MPX from controls.

A Principal component analysis (PCA) of controls, patients with MPX, and COVID-19 with 32 peptides absolutely quantified in all samples using liquid chromatography selective reaction monitoring (LC-SRM).

B Hierarchical clustering using differentially regulated proteins between patients with MPX and controls (Heatmap); $P < 0.05$ with Mann–Whitney *U* test with FDR-based multiple testing correction.

C Key proteins that differ between patients with MPX and controls, and COVID-19 (Boxplots). Dashed blue lines indicate the lowest detected peptide concentration from calibration curves. Statistics based on Mann–Whitney *U* test with multiple testing corrections. For boxplots, median is indicated by a solid line, hinges show the 25th and 75th percentiles, whiskers show values that, at maximum, are within 1.5 times the interquartile range.

D Top 15 peptides of an SVM-trained model discriminating between healthy controls (n = 15) and MPX cases (n = 6). Means of the relative coefficients over a 5-fold cross-validation are shown. Error bars denote the standard deviations. Red denotes positive, blue denotes negative coefficients. The AUC was calculated based on withheld samples that were not used for training the model.

for more than 50 years—becomes ever more apparent and calls urgently for a better understanding of this disease. In this context, a case series of individuals with similar demographics, timing, and course of disease who likely contracted the infection at the same social event caught our attention. Usually, the host response to a viral pathogen would be investigated in larger cohorts. However, considering urgently needed data and the parallel disease history of our case series, we speculated that because of the homogeneous and representative nature of cohort considering the current outbreak, even a low number of individuals may provide a clear proteomic signal, allowing us to provide a timely assessment of the host response to MPXV infection.

Indeed, analyzing the host response of the MPX patients at the proteome level provided a surprisingly clear picture, even in this small cohort, especially when comparing the proteomes to age- and sex-matched healthy individuals or patients with severity-matched, moderate COVID-19. Our dataset showed increased levels of specific acute phase proteins and overall lower nutritional response proteins such as TTR and apolipoproteins in MPX when compared to healthy controls. However, key pathways altered in COVID-19, including the complement and coagulation systems, were affected to a much lesser extent. The proteomic response described in our study therefore reflects the different pathophysiology connected with MPXV and SARS-CoV-2 as well as the mild to moderate disease severity in MPX observed in the current outbreak so far (Pfäfflin *et al*, 2022; Thornhill *et al*, 2022). Additional cohort studies will be required to validate our results in the broader context. Reassuringly however, most proteins identified by our non-targeted proteomic technique to be differentially abundant in MPX, have a known biological role in the acute phase response to viral infections. The correlation of numerous of the inflammatory proteins with disease severity gives additional and orthogonal confidence in our results.

We identified several peptides that showed a statistically robust correlation with disease severity as determined by the number of skin lesions. Organ dysfunction and severe disease have so far only sporadically been reported in the current outbreak in Europe and the US (Thornhill *et al*, 2022). MPX is however known to cause severe and lethal disease in endemic regions in Africa, with reported case fatality of up to 10% (Bunge *et al*, 2022).

Emerging pathogens with pandemic potential require fast responses, and an attractive possibility to achieve that is in the repurposing of existing procedures, diagnostics, and therapies, whenever possible. Prognostic biomarker panel assays were discussed during the COVID-19 pandemic for the monitoring of clinical trials, for supporting clinical decisions, and for their potential to support the navigation through difficult triaging situations (Struwe *et al*, 2020; Papadopoulou *et al*, 2022; Wang *et al*, 2022a). Drawing from our previous experience and based on the signature of the MPX human host response in discovery proteomics in the present study, we explored the application of a biomarker panel designed to classify patients with COVID-19 in routine laboratories (Wang *et al*, 2022a) on this very different viral disease. Within the limitations of a small case series study, the biomarker panel captured hallmarks of MPXV infection and facilitated a classification of patients with MPX and healthy controls in our sample set using an SVM model. The attractiveness of MRM panel assays is that they can be implemented in clinical workflows and are of low cost per sample. A panel of severity markers could be of help in endemic regions and possibly help to elucidate the pathophysiological differences between the Central African and West African clade of MPXV in the future. Our case series was too small to determine if the biomarker panel can be used to predict disease features, e.g., time to recovery, or to discriminate the effectiveness of therapeutic options. However, the correlation of the proteomic response with the number of skin lesions suggests that a predictive application of proteomics is possible for MPX and suggests conducting respective cohort studies in the near future. Indeed, we hope that the clear proteomic signature revealed by our case series justifies larger studies involving different cohorts and longitudinal sample in in the near future.

We believe our study demonstrates two essential aspects which are important for pandemic preparedness. First, our study exemplifies that when time is of the essence, proteomics can deliver valuable information on the molecular disease etiology of a moderate number of affected individuals, at least when their disease history is homogeneous and/or representative as in our case series study. Our results therefore imply that plasma proteomics might be particularly valuable for rare and neglected diseases, where proteomics may become an increasingly attractive toolkit for systemic analyses, despite limited case numbers. Indeed, given that symptoms were relatively mild, the proteomic host response to MPXV was distinct, with about one quarter of the highly abundant functional fraction of the plasma proteome changing. Second, our data suggested that there could be an untapped potential in the repurposing of biomarker panel assays across viral disease: although the overall proteomic signature clearly distinguished MPX from COVID-19, there was a sufficient overlap in the host response signature, so that we could distinguish MPX patients from healthy controls on the basis of a COVID-19 proteomic panel assay. Indeed, although our data hence shows that the individual biomarkers are not specific to a particular infection, the pattern in which they respond seems highly discriminatory. Although these results are to be regarded preliminary due to the moderate size of our cohort, our data suggests that one could generate a proteomic panel assay that is applicable across different viral diseases; in case of a new viral agent, one could hence measure the same panel of biomarkers, and only would need to adapt the data analysis and ML models to the novel agent. Future studies are needed to substantiate the viability of this possibility.

Materials and Methods

Patient cohort, biosamples, and clinical data

Patients with PCR-confirmed MPXV infection were recruited in a prospective observational study on the clinical and molecular characteristics of MPX. Written informed consent for collection of clinical data and blood was obtained from all patients before inclusion. Biosampling for proteomic measurements was performed on day 1–3 after admission to the hospital. Clinical data were captured in a purpose-built database. The study was approved by the ethics committee of Charité—Universitätsmedizin Berlin (EA2/139/22) and conducted in accordance with the Declaration of Helsinki and guidelines of Good Clinical Practice (EMA, 1996/2018). Biosamples for the cohort of patients with COVID-19 were obtained from the PaCOVID-19 study, a prospective observational cohort study on the

pathophysiology of COVID-19 conducted at Charité—Universitätsmedizin Berlin (Kurth *et al*, 2020; Thibeault *et al*, 2021). Biosamples for the cohort of healthy controls were obtained from a clinical study including healthy volunteers (Hillus *et al*, 2021).

Reagents and consumables

Water was from Merck (LiChrosolv LC–MS grade; Cat# 115333), acetonitrile was from Biosolve (LC–MS grade; Cat# 012078), trypsin (sequence grade; Cat# V511X) and trypsin/LysC mix (mass-spec grade; Cat# V5072) were from Promega, 1,4-dithiothreitol (DTT; Cat# 6908.2) was from Carl Roth, urea (puriss. P.a., reag. Ph. Eur.; Cat# 33247), Tris(2-carboxyethyl) phosphine hydrochloride (TCEP; Cat# 646547), and RIPA buffer (Cat# R0278) were from Merck, ammonium bicarbonate (ABC; eluent additive for LC–MS; Cat# 40867), 2-chloroacetamide (Cat# 22788), and dimethyl sulfoxide (DMSO; Cat# 41648) were from Fluka, formic acid (LC–MS grade; eluent additive for LC–MS; Cat# 85178), PCR sealing foil sheets (Cat# AB-0626), and Pierce quantitative fluorometric peptide assays (Cat# 23290) were from Thermo Fisher Scientific, bovine serum albumin (BSA; albumin Bovine Fraction V, Very Low Endotoxin, Fatty Acid-free; Cat# 47299) was from Serva, 96-well ultrafiltration plates (AcroPrep) Advance Filter Plates for Ultrafiltration, 1 ml, Omega 30 K MWCO (Cat# 8165) were from PALL, 96-well LoBind plates (Cat# ER0030129512-25EA) were from Merck, stable isotopic labeled (SIL) reference peptides for discovery proteomics (PQ500 Reference Peptides) were from Biognosys.

Sample preparation

Plasma samples were diluted 1:10 in RIPA buffer and heated at 95°C for 10 min. After cooling to room temperature (RT), 15 µl (~ 100 µg protein) were processed by FASP as previously described with minor modifications (Fossati *et al*, 2021) and transferred to a 96-well ultrafiltration plate mounted onto a collection plate (96-well LoBind plate). Liquid was removed by centrifugation (30 min, 1,800 × rcf, 20°C). Samples were denatured and reduced in 50 µl TUA buffer (8 M urea, 20 mM ammonium bicarbonate, 5 mM TCEP) for 30 min at room temperature without shaking. Following thiol alkylation (addition of 10 µl CA buffer (50 mM 2-chloroacetamide, 20 mM ABC) and incubation in the dark at RT for 30 min), the plate was centrifuged (30 min, 1,800 × rcf, 20°C). Samples were washed twice (30 min, 1,800 × rcf, 20°C) with 100 µl 20 mM ABC. Following an additional centrifugation to remove residual liquid (60 min, 1,800 × rcf, 20°C), the filter plate was moved to a fresh collection plate. To each well 50 µl 20 mM ABC containing 1 µg of trypsin/LysC mix was added, the plate was sealed with an adhesive PCR sealing foil sheet, and incubated at 37°C for 15 h. Peptides were collected by centrifugation (30 min, 1,800 × rcf, 20°C). Following the addition of 70 µl of HPLC-grade water to each well, the plate was centrifuged once more. The collection plate was then placed in a SpeedVac and samples were evaporated to complete dryness. Peptides were reconstituted in formic acid (30 µl, 0.1% v/v). Peptide concentration was determined using the Pierce quantitative fluorometric peptide assay.

For discovery proteomics, all samples (QCs, monkeypox, COVID-19, and healthy controls) were diluted to 200 ng/µl. The stable isotopic labeled reference peptides (PQ500 Reference Peptides) stock was prepared as described in the vendor's protocol (PQ500™ Reference Peptides Kit for Human Samples MANUAL), and diluted 1:10 in 50/50 v/v ACN:H2O. 2 µl of diluted PQ500 stock solution were spiked into 18 µl of the 200 ng/µl sample before transfer to vials for injection. For targeted proteomics, 15 µl of pre-digested heavy labeled standards (details in Wang *et al*, 2022a) were spiked into 10 µl samples (QCs, monkeypox, COVID-19, and healthy controls) and 20 µl were injected into the LC–MS system.

Mass spectrometry

Discovery proteomics using Zeno SWATH MS (preprint: Wang *et al*, 2022b) Tryptic digests were analyzed on a 7600 ZenoTOF mass spectrometer system (SCIEX), coupled to an ACQUITY UPLC M-Class system (Waters). 2 µl of each sample (360 ng sample + 0.02 µl PQ500, Biognosys) were loaded on a HSS T3 column (300 µm × 150 mm, 1.8 µm, Waters) heated to 35°C, then chromatographically separated with a 20-min gradient using a flow rate of 5 µl/min (Zelezniak *et al*, 2018). A Zeno SWATH acquisition scheme with 85 variable-size windows and 11-ms accumulation time with 1.4 s cycle time was used (preprint: Wang *et al*, 2022b) which allows for MS detection for average 7 points per chromatographic peak with the chosen chromatography.

Targeted proteomics by multiple reaction monitoring (plasma biomarker panel; Wang et al, 2022a)

Tryptic digests were analyzed on a 6495C triple quadrupole mass spectrometer (Agilent) coupled to a 1290 Infinity II UHPLC system (Agilent). Prior to MS analysis, samples were chromatographically separated on an InfinityLab Poroshell 120 EC-C18 column (2.1 × 50 mm, 1.9 µm, Agilent) heated to 45°C with a flow rate of 800 µl/min. The 6495C mass spectrometer was controlled by MassHunter Workstation software (LC–MS/MS Data Acquisition for 6,400 series Triple Quadrupole, Version 10.1 (Agilent)) and was operated in positive electrospray ionization mode. Samples were analyzed in dynamic multiple reaction monitoring (MRM) mode with both quadrupoles operating at unit resolution (Wang *et al*, 2022a).

Data processing

Discovery proteomics

The Zeno SWATH raw proteomics data was processed using DIA-NN (Demichev *et al*, 2020), 1.8.1 beta 20, availabe on github (DIA-NN github repository). The MS2 and MS1 mass accuracies were set to 20 and 12 ppm, and the scan window to 7. For the discovery approach, we used a publicly available spectral library for human plasma (Bruderer *et al*, 2019) and replaced spectra and RT information with DIA-NN in silico prediction. Protein inference was switched off and the match-between-runs (MBR) option was enabled. The processing pipeline is available in Supplementary Materials.

Targeted proteomics

LC–MRM data were processed using MassHunter Quantitative Analysis, v10.1 (Agilent). No blinding was done during peak integration. Peptide absolute concentration (expressed in ng/ml) was determined from calibration curves, constructed with native and SIL peptide standards in surrogate matrix (40 mg/ml BSA), and manually

validated. Linear regression analysis of each calibration curve was performed using custom R code (with 1/x weighting). Detailed information on transitions and matching of native peptides and internal standards can be found in Wang *et al* (2022a). Peptides with > 40% of values below the lowest or above the highest detected calibrant concentration across all samples were removed from analysis.

Data analysis

Clinical data analysis

Pseudonymized clinical data were processed using JMP Pro 16 (SAS Institute).

Discovery proteomics data analysis

Peptide expressions were first normalized within each clinical group. No blinding was done in normalization. To deal with a higher number of missing values in plasma proteomics compared to those obtained from cellular proteomics, we adopted the following approach: Peptides with excessive missing values (> 40% per group) were excluded from our analysis. This group-based thresholding delivered approximately the same number of peptides as the 26% presence threshold applied to the total set. The missing values of remaining peptides were imputed group-based using the PCA method (Josse & Husson, 2016). The group-based imputation allowed to avoid admixing of information from other groups. After imputation, an additional step of normalization was applied to the total set without using group information. In both cases, normalization was performed with LIMMA (Ritchie *et al*, 2015) implementation of cyclic loess method (Bolstad *et al*, 2003) with option "fast" (Ballman *et al*, 2004). To obtain a quantitative protein data matrix, the log2-intensities of peptides were filtered, only peptides belonging to one protein group were kept, and then summarized into protein log intensity using the PLM method (Bolstad, 2008, 41–59) implemented in the preprocessCore R package (Bolstad, 2021).

Statistical analysis of proteomics data was carried out in R using publicly available packages. Linear modeling was based on the R package LIMMA (Ritchie *et al*, 2015). The following model was applied to each tissue dataset (log2(p) is the log2-transformed expression of a protein): log2(p) ~ 0 + Class. The categorical factor Class had three levels: MPX, COVID-19, and control; reference level: control. For correlation between MPX severity (NSkin lesions) and protein expression, log2(1 + NLesions / 15) was used for linear regression. Log base 2 transformation was applied to bring the number of lesions to the same scale as protein expressions, 1 was added to guarantee that N lesions = 0 is transformed to 0, and division by Nmean = 15 (average number of lesions) was applied to map the average lesions number to 1. Also note that for 0 < N/Nmean < 1.5, deviations of f(N) from linearity are less than 12%.

For finding regulated features, the following criteria were applied for all contrasts: Significance level alpha was set to 0.015, which guaranteed the Benjamini–Hochberg (Benjamini & Hochberg, 1995) false discovery rate below 5% for contrast MPX vs Control, below 4% for contrast COVID-19 vs Control, and below 22% for contrast MPX vs COVID-19. The log fold-change threshold was applied to all contrasts to guarantee that the measured signal is above the average noise level. As such we took the median residual standard deviation of linear model: log2(T) = median residual SD of linear modeling (= log2(1.35)). Functional GSEA analysis was carried out using the clusterProfiler R package (Yu *et al*, 2012). For selecting the most (de)regulated pathway terms, we applied filter: 3 ≤ term size ≤ 300. The data matrix and description are provided in Dataset EV1.

The paper explained

Problem

Until the recent outbreak, monkeypox was mainly confined to endemic areas in West and Central Africa, gaining little research interest. Aiming to breach the knowledge gap, we applied state-of-the-art plasma proteomics to a group of six patients with similar disease history and severity.

Results

Applying a recent proteomic method, ZenoSWATH-MS, on plasma samples obtained from a small but characteristic case series, we report distinct changes in proteins involved in the acute phase and nutritional response. Several proteins correlated with the number of skin lesions, indicating a potential use as disease severity markers. Comparing the proteomes to those of matched patients with COVID-19, we found numerous similarities. Moreover, we explored the usefulness of applying a proteomic COVID-19 biomarker panel assay to monkeypox cases and obtained a classification of the different disease groups.

Impact

This study is the first characterization of the human host response to monkeypox infection, offering insights into the pathophysiology. Moreover, we speculate that there is a thus far untapped potential for accelerating the response to disease outbreaks through the repurposing of biomarker assays.

Classifier construction and protein/peptide ranking

To complement the principal component and differential protein expression analysis, we constructed classifiers using a linear support vector machine (sklearn.svm.LinearSVC()) as implemented in scikit-learn 1.0.2 (Pedregosa *et al*, 2011) with an L1-penalty and balanced class-weights. The maximum number of iterations was increased to 10,000 to ensure convergence. As input, the log2-transformed quantities of the discovery proteomics and the 32 quantified peptides of the MRM panel were used, respectively.

The models were constructed and tested using a 5-fold shuffled and stratified cross-validation as implemented in sklearn.model_selection.StratifiedKFold(). For each iteration, 4 folds were used for training, 1 fold was used for testing the model. The data were scaled using sklearn.preprocessing.StandardScaler() fitted on the training data.

The AUC was calculated for the test data that were not used for training the model after all 5 iterations, resulting in one predicted value for every sample. For each iteration, the coefficients of the trained model were extracted and normalized by the maximum absolute coefficient of this iteration. For the plots, the mean and the standard deviation (error bars) of all 5 coefficients per protein/peptide were calculated and sorted according to the absolute mean. For reproducibility, the seed was fixed to 42.

Targeted proteomics

Significance testing of the absolute peptide concentrations and the sample type (control, MPX) was performed using Mann–Whitney *U*

test with multiple testing correction (where indicated). Test results are provided in Table EV1, P-values < 0.05 were considered significant.

Data availability

Internal patient IDs were changed at random within groups. Data matrix for discovery proteomics is available in (Dataset EV1). Proteomic raw data are deposited on PRIDE (https://www.ebi.ac.uk/pride/) under the project accession: PXD036074.

Expanded View for this article is available online.

Acknowledgments

Figure 1A and graphical abstract were made using BioRender (Biorender.com). We thank Vadim Demichev, Eva Tranter, Gisèle Godzick-Njomgang, Daniel Wendisch, Linda Jürgens, and Frieder Pfäfflin for their valuable input and help in collecting biosamples, and Hezi Tenenboim for proofreading our manuscript. We thank Christof von Kalle and the Charité-BIH Clinical Study Centre for their support in setting up our translational research platform. This research was funded in part by the Berlin Institute of Health (BIH), Wellcome Trust (IA 200829/Z/16/Z to MR), the European Research Council (ERC) under grant agreement ERC-SyG-2020 951475 (to MR), by the Ministry of Education and Research (BMBF), as part of the National Research Node Mass spectrometry in Systems Medicine (MSCoresys), under grant agreement 031L0220 (to MR), the Deutsche Forschungsgemeinschaft (DFG) (INST 335/797-1) (to MM) and the Berlin University Alliance (501_Massenspektrometrie, 501_Linklab). We further acknowledge funding by Deutsche Forschungsgemeinschaft (DFG) supporting ZW's PhD studies as part of the SFB/TRR 186. Open access funding enabled and organized by Projekt DEAL.

Author contributions

Ziyue Wang: Formal analysis; visualization; methodology; writing—original draft; writing—review and editing. **Pinkus Tober-Lau:** Data curation; writing—original draft; writing—review and editing. **Vadim Farztdinov:** Formal analysis; visualization; writing—original draft; writing—review and editing. **Oliver Lemke:** Formal analysis; visualization; writing—original draft; writing—review and editing. **Torsten Schwecke:** Methodology; writing—original draft. **Sarah Steinbrecher:** Data curation. **Julia Muenzer:** Formal analysis; writing—original draft. **Helene Kriedemann:** Data curation. **Leif Erik Sander:** Project administration. **Johannes Hartl:** Data curation; formal analysis; methodology; writing—original draft; writing—review and editing. **Michael Mülleder:** Conceptualization; supervision; funding acquisition; methodology; writing—original draft; project administration; writing—review and editing. **Markus Ralser:** Conceptualization; supervision; funding acquisition; writing—original draft; project administration; writing—review and editing. **Florian Kurth:** Conceptualization; data curation; supervision; writing—original draft; project administration; writing—review and editing.

Disclosure and competing interests statement

The authors declare that they have no conflict of interest.

References

Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 1: 845 – 867

Ballman KV, Grill DE, Oberg AL, Therneau TM (2004) Faster cyclic loess: normalizing RNA arrays via linear models. *Bioinformatics* 20: 2778 – 2786

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R I State Dent Soc* 57: 289 – 300

Bolstad B (2008) Preprocessing and Normalization for Affymetrix GeneChip Expression repoMicroarrays. In *Methods in Microarray Normalization*, P Stafford (ed), pp 41 – 59. Boca Raton, FL: Taylor & Francis

Bolstad B (2021) preprocessCore: a collection of pre-processing functions

Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19: 185 – 193

Bruderer R, Muntel J, Müller S, Bernhardt OM, Gandhi T, Cominetti O, Macron C, Carayol J, Rinner O, Astrup A *et al* (2019) Analysis of 1508 plasma samples by capillary-flow data-independent acquisition profiles proteomics of weight loss and maintenance. *Mol Cell Proteomics* 18: 1242 – 1254

Bunge EM, Hoet B, Chen L, Lienert F, Weidenthaler H, Baer LR, Steffen R (2022) The changing epidemiology of human monkeypox-A potential threat? A systematic review. *PLoS Negl Trop Dis* 16: e0010141

Camilli C, Hoeh AE, De Rossi G, Moss SE, Greenwood J (2022) LRG1: an emerging player in disease pathogenesis. *J Biomed Sci* 29: 6

Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, Caudy M, Garapati P, Gillespie M, Kamdar MR *et al* (2014) The Reactome pathway knowledgebase. *Nucleic Acids Res* 42: D472 – D477

D'Alessandro A, Thomas T, Dzieciatkowska M, Hill RC, Francis RO, Hudson KE, Zimring JC, Hod EA, Spitalnik SL, Hansen KC (2020) Serum proteomics in COVID-19 patients: altered coagulation and complement status as a function of IL-6 level. *J Proteome Res* 19: 4417 – 4427

Dellière S, Neveux N, De Bandt J-P, Cynober L (2018) Transthyretin for the routine assessment of malnutrition: A clinical dilemma highlighted by an international survey of experts in the field. *Clin Nutr* 37: 2226 – 2229

Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M (2020) DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* 17: 41 – 44

Demichev V, Tober-Lau P, Lemke O, Nazarenko T, Thibeault C, Whitwell H, Röhl A, Freiwald A, Szyrwiel L, Ludwig D *et al* (2021) A time-resolved proteomic and prognostic map of COVID-19. *Cell Syst* 12: 780 – 794

Demichev V, Tober-Lau P, Nazarenko T, Lemke O, Kaur Aulakh S, Whitwell HJ, Röhl A, Freiwald A, Mittermaier M, Szyrwiel L *et al* (2022) A proteomic survival predictor for COVID-19 patients in intensive care. *PLOS Digit Health* 1: e0000007

Dye C, Kraemer MUG (2022) Investigating the monkeypox outbreak. *BMJ* 377: o1314

EMA (1996/2018) ICH E6 (R2) Good clinical practice. In *European Medicines Agency* [Internet]. 17 Sep 2018. https://www.ema.europa.eu/en/ich-e6-r2-good-clinical-practice

European Centre for Disease Prevention and Control (2022) Joint ECDC-WHO Regional Office for Europe Monkeypox Surveillance Bulletin

Fossati A, Frommelt F, Uliana F, Martelli C, Vizovisek M, Gillet L, Collins B, Gstaiger M, Aebersold R (2021) System-wide profiling of protein complexes via size exclusion chromatography-mass spectrometry (SEC-MS). In *Methods in Molecular Biology*, JM Walker (ed), pp 269 – 294. Cham: Springer Nature

Georg P, Astaburuaga-García R, Bonaguro L, Brumhard S, Michalick L, Lippert LJ, Kostevc T, Gäbel C, Schneider M, Streitz M *et al* (2022) Complement activation induces excessive T cell cytotoxicity in severe COVID-19. *Cell* 185: 493 – 512

Hardardóttir I, Grunfeld C, Feingold KR (1995) Effects of endotoxin on lipid metabolism. *Biochem Soc Trans* 23: 1013 – 1018

He B, Huang Z, Huang C, Nice EC (2022) Clinical applications of plasma proteomics and peptidomics: towards precision medicine. *Proteomics Clin Appl* e2100097

Hillus D, Schwarz T, Tober-Lau P, Vanshylla K, Hastor H, Thibeault C, Jentzsch S, Helbig ET, Lippert LJ, Tscheak P *et al* (2021) Safety, reactogenicity, and immunogenicity of homologous and heterologous prime-boost immunisation with ChAdOx1 nCoV-19 and BNT162b2: a prospective cohort study. *Lancet* 9: 1255 – 1265

Isidro J, Borges V, Pinto M, Ferreira R, Sobral D, Nunes A, Dourado Santos J, José Borrego M, Núnci S, Pelerito A, *et al* (2022) First draft genome sequence of Monkeypox virus associated with the suspected multi-country outbreak, May 2022 (confirmed case in Portugal). *Virological*

Josse J, Husson F (2016) missMDA: a package for handling missing values in multivariate data analysis. *J Stat Softw* 70: 1 – 31

Kraemer MUG, Tegally H, Pigott DM, Dasgupta A, Sheldon J, Wilkinson E, Schultheiss M, Han A, Oglia M, Marks S *et al* (2022) Tracking the 2022 monkeypox outbreak with epidemiological data in real-time. *Lancet Infect Dis* 22: 941 – 942

Kurth F, Roennefarth M, Thibeault C, Corman VM, Müller-Redetzky H, Mittermaier M, Ruwwe-Glösenkamp C, Heim KM, Krannich A, Zvorc S *et al* (2020) Studying the pathophysiology of coronavirus disease 2019: a protocol for the Berlin prospective COVID-19 patient cohort (Pa-COVID-19). *Infection* 48: 619 – 626

Liotta LA, Kohn EC, Petricoin EF (2001) Clinical proteomics: personalized molecular medicine. *JAMA* 286: 2211 – 2214

Messner CB, Demichev V, Wendisch D, Michalick L, White M, Freiwald A, Textoris-Taube K, Vernardis SI, Egger A-S, Kreidl M *et al* (2020) Ultra-high-throughput clinical proteomics reveals classifiers of COVID-19 infection. *Cell Syst* 11: 11 – 24

Nuñez E, Orera I, Carmona-Rodríguez L, Paño JR, Vázquez J, Corrales FJ (2022) Mapping the Serum Proteome of COVID-19 Patients; Guidance for Severity Assessment. *Biomedicine* 10: 1690

Overmyer KA, Shishkova E, Miller IJ, Balnis J, Bernstein MN, Peters-Clarke TM, Meyer JG, Quan Q, Muehlbauer LK, Trujillo EA *et al* (2021) Large-scale multi-omic analysis of COVID-19 severity. *Cell Syst* 12: 23 – 40

Papadopoulou G, Manoloudi E, Repousi N, Skoura L, Hurst T, Karamitros T (2022) Molecular and clinical prognostic biomarkers of COVID-19 severity and persistence. *Pathogens* 11: 311

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V *et al* (2011) Scikit-learn: Machine learning in Python. *J Mach Learn Res* 12: 2825 – 2830

Pfäfflin F, Wendisch D, Scherer R, Jürgens L, Godzick-Njomgang G, Tranter E, Tober-Lau P, Stegemann MS, Corman VM, Kurth F *et al* (2022) Monkeypox in-patients with severe anal pain. *Infection* https://doi.org/10.1007/s15010-022-01896-7

Ralser M, Nonhoff U, Albrecht M, Lengauer T, Wanker EE, Lehrach H, Krobitsch S (2005) Ataxin-2 and huntingtin interact with endophilin-A complexes to function in plastin-associated pathways. *Hum Mol Genet* 14: 2893 – 2909

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43: e47

Shen B, Yi X, Sun Y, Bi X, Du J, Zhang C, Quan S, Zhang F, Sun R, Qian L *et al* (2020) Proteomic and metabolomic characterization of COVID-19 patient sera. *Cell* 182: 59 – 72

Struwe W, Emmott E, Bailey M, Sharon M, Sinz A, Corrales FJ, Thalassinos K, Braybrook J, Mills C, Barran P (2020) The COVID-19 MS Coalition—accelerating diagnostics, prognostics, and treatment. *Lancet* 395: 1761 – 1762

Su Kim D, Choi YD, Moon M, Kang S, Lim J-B, Kim KM, Park KM, Cho NH (2013) Composite three-marker assay for early detection of kidney cancer. *Cancer Epidemiol Biomarkers Prev* 22: 390 – 398

Thibeault C, Mühlemann B, Helbig ET, Mittermaier M, Lingscheid T, Tober-Lau P, Meyer-Arndt LA, Meiners L, Stubbemann P, Haenel SS *et al* (2021) Clinical and virological characteristics of hospitalised COVID-19 patients in a German tertiary care centre during the first wave of the SARS-CoV-2 pandemic: a prospective observational study. *Infection* 49: 703 – 714

Thornhill JP, Barkati S, Walmsley S, Rockstroh J, Antinori A, Harrison LB, Palich R, Nori A, Reeves I, Habibi MS *et al* (2022) Monkeypox virus infection in humans across 16 countries – April-June 2022. *N Engl J Med* 387: 679 – 691

Vernardis S, Demichev V, Lemke O, Grüning N-M, Messner C, White M, Pietzner M, Peluso A, Collet T-H, Henning E *et al* (2022) Acute caloric restriction acts on the plasma proteome and reveals Apolipoprotein C1 as a signal of nutritional state and metabolic disease. *Res Square*

Viodé A, Smolen KK, Fatou B, Wurie Z, Van Zalm P, Konde MK, Keita BM, Ablam RA, Fish EN, Steen H (2022) Plasma proteomic analysis distinguishes severity outcomes of human Ebola virus disease. *MBio* e0056722

Völlmy F, van den Toorn H, Zenezini Chiozzi R, Zucchetti O, Papi A, Volta CA, Marracino L, Vieceli Dalla Sega F, Fortini F, Demichev V *et al* (2021) A serum proteome signature to predict mortality in severe COVID-19 patients. *Life Sci Alliance* 4: e202101099

Wang Z, Cryar A, Lemke O, Tober-Lau P, Ludwig D, Helbig ET, Hippenstiel S, Sander L-E, Blake D, Lane CS *et al* (2022a) A multiplex protein panel assay for severity prediction and outcome prognosis in patients with COVID-19: an observational multi-cohort study. *eClinicalMedicine* 49: 101495

Wang Z, Mülleder M, Batruch I, Chelur A, Textoris-Taube K, Schwecke T, Hartl J, Causon J, Castro-Perez J, Demichev V, *et al* (2022b) High-throughput proteomics of nanogram-scale samples with Zeno SWATH DIA. *medRxiv* https://doi.org/10.1101/2022.04.14.488299 [PREPRINT]

World Health Organisation (2022a) Second meeting of the International Health Regulations (IHR) Emergency Committee regarding the multi-country outbreak of monkeypox

World Health Organisation (2022b) WHO Fact Sheet on Monkeypox

World Health Organisation (2022c) WHO Director-General's opening remarks at the COVID-19 media briefing– 1 June 2022

Yu G, Wang L-G, Han Y, He Q-Y (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16: 284 – 287

Zelezniak A, Vowinckel J, Capuano F, Messner CB, Demichev V, Polowsky N, Mülleder M, Kamrad S, Klaus B, Keller MA *et al* (2018) Machine learning predicts the yeast metabolome from the quantitative proteome of kinase knockouts. *Cell Syst* 7: 269 – 283

Received: 12 April 2023 | Revised: 7 May 2023 | Accepted: 7 May 2023

DOI: 10.1002/jssc.202300244

JOURNAL OF SEPARATION SCIENCE

REVIEW ARTICLE

Capillary electrophoresis-mass spectrometry for intact protein analysis: Pharmaceutical and biomedical applications (2018–March 2023)

Katarína Maráková[1,2] | **Martina Opetová**[1,2] | **Radovan Tomašovský**[1,2]

[1]Faculty of Pharmacy, Department of Pharmaceutical Analysis and Nuclear Pharmacy, Comenius University Bratislava, Bratislava, Slovakia

[2]Faculty of Pharmacy, Toxicological and Antidoping Center, Comenius University Bratislava, Bratislava, Slovakia

Correspondence
Katarína Maráková, Comenius University Bratislava, Faculty of Pharmacy, Department of Pharmaceutical Analysis and Nuclear Pharmacy, Odbojarov 10, SK-832 32 Bratislava, Slovakia.
Email: marakova@fpharm.uniba.sk

Funding information
Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic and the Slovak Academy of Sciences under the projects, Grant/Award Numbers: VEGA 1/0483/20, VEGA 1/0514/22

Capillary electrophoresis is recognized as a valued separation technique for its high separation efficiency, low sample consumption, good economic and ecological aspects, reproducibility, and complementarity to traditional liquid chromatography techniques. Capillary electrophoresis experiments are generally performed utilizing optical detection, such as ultraviolet or fluorescence detectors. However, in order to provide structural information, capillary electrophoresis hyphenated to highly sensitive and selective mass spectrometry has been developed to overcome the limitations of optical detections. Capillary electrophoresis-mass spectrometry is increasingly popular in protein analysis, including biopharmaceutical and biomedical research. It is frequently applied for the determination of physicochemical and biochemical parameters of proteins, offers excellent performance for in-depth characterizations of biopharmaceuticals at various levels of analysis, and has been also already proven as a promising tool in biomarker discovery. In this review, we focus on the possibilities and limitations of capillary electrophoresis-mass spectrometry for protein analysis at their intact level. Various capillary electrophoresis modes and capillary electrophoresis-mass spectrometry interfaces, as well as approaches to prevent protein adsorption and to enhance sample loading capacity, are discussed and the recent (2018–March 2023) developments and applications in the field of biopharmaceutical and biomedical analysis are summarized.

KEYWORDS
biological fluids, biomarkers, biopharmaceuticals, critical quality attributes, top-down proteomics

Article Related Abbreviations: ACE, affinity capillary electrophoresis; BsAb, bispecific antibody; CEX, cation exchange chromatography; CQAs, critical quality attributes; CSE, capillary sieving electrophoresis; DPJ, dynamic pH junction; EKS, electrokinetic supercharging; EMA, European Medicines Agency; FDA, Food and Drug Administration; FESI, field-enhanced sample injection; hCG, human chorionic gonadotropin; HCP, host cell protein; iCIEF, imaged capillary isoelectric focusing; LE, leading electrolyte; LLE, liquid-liquid extraction; mAb, monoclonal antibody; mCE, microchip capillary electrophoresis; MCE, mobility capillary electrophoresis; M_r, molecular mass; NACE, nonaqueous capillary electrophoresis; PS, polarity switching; PSA, prostate-specific antigen; PTM, post-translational modification; RCAD, renal cysts and diabetes syndrome; SDS, sodium dodecyl sulfate; TE, terminating electrolyte; tITP, transient isotachophoresis.

1 | INTRODUCTION

Proteins are important biomacromolecules composed of amino acids playing crucial roles in all living organisms. It is now becoming clear that to fully understand (patho)physiology, an understanding of the different protein functions and interactions is crucial. Proteins act as enzymes, hormones, receptors, antigens, antibodies, transporters, and other functional and structural elements. They ensure various vital important functions, such as the storage and transfer of genetic information, immunity, and overall control of the organism. Post-translational modifications (PTMs) are an integral part of protein structure and are important for the stability and proper biological functioning of proteins. Therefore, qualitative and quantitative evaluation of PTMs is important for protein biopharmaceuticals as well as in analyses of biomarkers [1–3].

Many proteins, especially monoclonal antibodies, have become important biopharmaceuticals [4]. The molecular complexity of these therapeutic agents requires a large variety of analytical methods to identify and monitor harmful modifications and degradation products. Various separation methods, including CE, can be applied for the testing of biopharmaceuticals at all stages of the manufacturing process, their physicochemical characterization, impurities analysis, stability determination, and biosimilarity assessment [5, 6]. The evaluation of size, charge, and glycan heterogeneity is often required by regulatory authorities and CE methods are becoming increasingly popular even with the potential of implementation of good manufacturing practice.

Even though many diseases are currently still dependent on diagnosis using invasive methods, in recent decades more and more diagnostic and monitoring actions can be performed using biomarkers. These molecules can provide a great tool for improving diagnostics, monitoring the state of the disease, determining its prognosis, and can predict the treatment effectiveness and success. Ideally, biomarkers should be obtained from biological fluids given their noninvasive nature. Single biomarkers in many cases lack specificity, sensitivity, and predictive value, therefore also proteomic profiling in this area is of particular interest [7]. However, to achieve a reliable outcome or clinically valuable biomarkers, analytical advances must be coupled with adequate sample preparation methods, bioinformatics tools, and appropriate data analysis [8–10].

Since the completion of the Human Genome Project, which finished well ahead of schedule thanks also to the massive application of capillary electrophoresis, CE is now a minority technique when compared to the widespread use of LC. LC is a widely recognized work-horse technique also for proteomic applications in both laboratory and clinical settings mainly due to its high separation efficiency, robustness, reproducibility, large sample loading capacity, wide separation window, and good solvent compatibility with MS detection [11, 12]. However, CE emerged as a mature and robust separation technique increasingly popular in peptidomic and proteomic analysis [13–19]. CE methods can be found in all major pharmacopeias and are successfully utilized in biopharmaceutical research and industry [20–27]. CE, as an electric field-mediated separation technique first introduced in 1981 [28], offers different selectivity and therefore complementarity to LC techniques. It provides fast highly efficient separations and requires ultralow sample amounts and volumes per single analysis. Moreover, utilizing also very low amounts of mostly aqueous reagents and solvents is much more eco-friendly. Continuous advancement of CE has been focused on overcoming key obstacles, such as low concentration sensitivity and strong capillary wall adsorption of proteins, which could be particularly problematic when employed to analyze complex biological samples. Microchip CE (mCE) is also a trending technology thanks to its short analysis time, high throughput and sensitivity, and very low sample consumption [14, 29–31].

Although the development was initially put into LC-MS systems due to its straightforward coupling and better stability of operation, the hyphenation of CE with highly selective and sensitive MS is emerging as a promising bioanalytical tool [32–36]. The hyphenation of MS detector to CE has advantages such as a lower detection limit than UV, no need for chromophore and fluorophore groups in the analyte, subsequent resolution of comigrating analytes, and elucidation of the structure of unknown molecules. A special interface and volatile buffers are needed for the on-line hyphenation of CE with MS [37–40]. The high salt content and complexity of the biological samples are often problematic for MS detection.

There are two main strategies for the MS, LC-MS, and CE-MS analyses of proteins. The more traditionally used "bottom-up" proteomics, where proteins are first enzymatically digested into their constituent peptides, and "top-down" proteomics, where proteins are analyzed in their intact form (native or denatured). The progress of top-down proteomics requires analytical methods with high peak capacity for proteoform separation and high sensitivity for proteoform detection as proteome samples are usually extremely complex. It has been estimated that the human proteome contains over 1 million proteoforms [41] and the advanced CE-MS was able to identify almost 6000 proteoforms from a complex proteome sample, which represents one of the largest top-down proteomic datasets so far [42]. However, the separation window and sample loading capacity of CE-MS is at least 10 times narrower and 100 times lower than that of LC-MS, respectively, which

hindered the adoption of CE-MS in large-scale top-down proteomics.

Several general review papers have been recently published on the instrumental developments of CE-MS [33], the application of CE or CE-MS in proteomics [13–15, 29, 34], as well as more specific reviews covering recent advances and applications of CE methods for the analysis of biopharmaceuticals [20–25, 43] and clinical proteomic applications [10, 44–46]. A comprehensive review of the publications reported between 1987 and 2007 on CE-MS for the analysis of intact proteins was published by Haselberg et al. [47] with updates covering the years 2007–2010 [48] and 2010–2012 [49] and book chapters focused on four practical CE-MS methodologies [50] and a general chapter covering various technical parameters of CE-MS for intact protein analysis [51]. Therefore, for a more comprehensive overview of the general technical developments and applications of top-down proteomics, the reader is referred to those reviews. This review aims to present developments solely in the field of intact protein analysis by CE-MS with a focus on pharmaceutical and biomedical applications. The present review will cover (i) the basic information about the biopharmaceuticals and proteomic biomarkers, (ii) important technical aspects of CE-MS; its advantages and drawbacks, CE modes and MS interfaces with a specific focus on intact protein analysis as well as prevention of protein adsorption and enhancement of sample loading capacity, (iii) CE-MS applications for the analysis of biopharmaceuticals and biomarkers based on peptides and intact proteins in the recent years 2018–March 2023, and (iv) our conclusions and future perspectives in the field.

2 | PEPTIDE- AND PROTEIN-BASED THERAPEUTICS

Biopharmaceuticals represent a very rapidly growing broad class of therapeutic agents for the advanced treatment of more than thirty life-threatening and severe diseases including cancer, autoimmune disorders, diabetes, cardiovascular, inflammatory, and infectious diseases. Biopharmaceuticals include diverse peptide- and protein-based modalities, such as monoclonal antibodies (mAbs), bi-and multi-specific antibodies, antibody-drug conjugates, fusion proteins, growth factors, cytokines, hormones, therapeutic enzymes, blood factors, anticoagulants, but also non-proteomic agents, such as nucleic acids and viral vectors [4, 52]. Hundreds of biopharmaceuticals have been developed and approved by the Food and Drug Administration (FDA) and the European Medicines Agency (EMA), and came to the market since the approval of the very first recombinant therapeutic biologic, i.e., insulin, in 1982 [53]. mAbs represent the most important class of biopharmaceuticals and the most successful category of pharmaceutical products occupying the top selling ranks for drugs since the last decade [54, 55]. As the patents of many biopharmaceuticals have already expired or are close to the expiration date, we can expect further expansion and more of these therapeutics coming to the market in the near future thanks to the biosimilar industry. Biosimilars are biopharmaceuticals that have demonstrated similarity in their structure, function, quality, safety, and efficacy to the innovator product [56, 57].

Biopharmaceuticals are, unlike small therapeutic molecules, primarily produced in living organisms, therefore a major concern with these products is their heterogeneity in terms of size, charge, and PTMs. The contributions of the host cell line, culture media, and conditions during the bioprocessing manifest into the mutations during translation, uncontrolled PTMs, aggregation, 3D conformation alterations, and degradation during the storage and transportation and result in a population of heterogenous related protein variants that may affect the safety and efficacy of the final therapeutic products. Therefore, the established manufacturing process as well as the final therapeutic product are subjected to quality control to assure that the acceptance criteria are within defined specifications of pharmacopeias, the International Council for Harmonization, FDA, and/or EMA guidelines, and biopharmaceuticals are produced in consistent quality. Monitoring of the so-called critical quality attributes (CQAs), such as primary sequence, charge, and size heterogeneity, glycosylation pattern as well as other PTMs potentially involved in the alteration of the pharmacokinetic and pharmacodynamic properties of the proteins and the presence of the host cell proteins (HCPs), is often required by regulatory authorities as outlined by the International Council for Harmonization guidelines [58–60]. Since biosimilars are also produced by living cells, even minor alterations during the manufacturing process may lead to significant consequences on the quality and safety of the final product. Thus, extensive analytical characterization is mandatory also to prove their similarity to the innovator's product [57, 61].

Because of the natural complexity of biopharmaceuticals, there is a crucial need for superior analytical methods to provide a comprehensive in-depth characterization of these molecules, their variants, and degradation products. Besides traditionally established separation techniques based on liquid chromatography, such as SEC, RP-HPLC, IEC, and slab gel electrophoresis, such as sodium dodecyl sulfate (SDS)-PAGE and 2D-differential gel electrophoresis [57, 62], CE offers numerous benefits and unique features of the separation mechanisms and therefore represents one of the most powerful tools for the characterization of

biopharmaceuticals. CE can be increasingly found in all major pharmacopeias and techniques as CZE, CGE, and CIEF are recommended for the analysis of biopharmaceutical products during process development, characterization, and quality control. Often, a combination of multiple analytical approaches is necessary to provide a detailed characterization of a complex biotherapeutic. In the past decade, several hundred papers have been published on the analytical and structural characterization of biotherapeutics, mainly mAbs, and the trend will certainly continue to expand in the future. The application of different CE techniques for the separation of biopharmaceuticals has been discussed in several recent excellent reviews [20–25, 43] and the reader is referred to those for more in-depth information. Just a brief information about the critical quality attributes and the position of CE in their monitoring will be presented in **Chapter 2.1** and applications of CE-MS for peptide- and protein-based therapeutics will be discussed in **Chapter 5.1**.

2.1 | Critical quality attributes

CQAs are physical, chemical, biological, or microbiological characteristics that should be within an appropriate limit, range, or distribution to ensure the desired quality of a product. CQAs of biopharmaceuticals include the determination of charge and size heterogeneity, glycosylation pattern as well as other PTMs, and the presence of the HCPs.

2.1.1 | Size heterogeneity

Biopharmaceuticals may contain multiple species of various sizes, ranging from a few nanometers to micrometers. Product-related aggregates (dimers, tetramers, and even multimers), also known as high molecular weight species, and fragments of low molecular weight relative to the parent molecule require close monitoring during manufacturing, release, stability, and storage and must meet the acceptance criteria as they may reduce the potency and increase the immunogenicity of the product [63]. SDS-CGE by using borate cross-linked dextran gels is the most used method for protein size heterogeneity assessment, aggregation, and fragment impurities characterization, which provides orthogonal information to size exclusion chromatography [64–66].

2.1.2 | Charge heterogeneity

The presence of product-related species, differing from the main product in terms of charge, can arise due to the presence of PTMs and degradation reactions such as deamidation, C-terminal lysine processing, and glycation. Their presence in the biopharmaceuticals may alter the product efficacy and pharmacokinetics, cause complete product inactivation, or increase immunogenicity [24, 67]. The most common charge variant monitoring methods in the biotechnology industry include CIEF or imaged CIEF (iCIEF). CIEF and iCIEF allow the determination of the apparent pI of a biotherapeutic, levels of acidic, main, and basic variants as well as impurities and they have become reference methods for the characterization of charge heterogeneity for biopharmaceutical companies [68, 69].

2.1.3 | Glycosylation

Additional heterogeneity may be introduced into biopharmaceuticals due to PTMs and biochemical changes, such as deamidation and oxidation. One of the major PTMs occurred in biopharmaceuticals is glycosylation, an addition of glycans (sugar moieties) to the molecule. Glycosylation is a highly heterogeneous process resulting in protein molecules with various glycosylation patterns in terms of glycosylation sites (macro-heterogeneity), and glycan structures (micro-heterogeneity). N- and O-linked glycans are the most occurring forms [70]. The glycosylation complexity varies by the host cell and bioprocessing parameters, such as growth phase, nutrition, oxygen level, pH, and temperature. Glycan content testing at drug product release is required to ensure profile consistency [71]. N-glycan analysis can occur at four levels: intact, subunit, peptide, and released glycan. The number and relative abundance of glycoforms are better investigated at the intact level as it does not introduce the sample preparation bias (no enzymatic pre-treatment is needed). CGE can separate glycoforms, as glycans modify the mobility of the protein in CGE with their mass but also with their charge since the SDS molecules do not interact with them. CZE and CIEF methods can differentiate glycan-altering modifications such as sialylation, acetylation, sulfation, or phosphorylation, due to their impact on the overall charge density of the protein [72]. With continuous advances in the CZE-MS instrumentation, this method could represent an excellent way to conduct intact analysis of glycoproteins and has the potential to be the main technique for routine analyses at least for mildly glycosylated proteins [73].

2.1.4 | Host cell proteins

HCPs are other than the desired recombinant product produced by the host cell line during its life cycle. The accepted level of HCPs is below 100 ppm in the final

biotherapeutic product. A higher level of HCPs may lead to unwanted immune responses in patients and may affect the stability and potency of the product. Due to its high resolving power, CE is gaining attention for its potential in HCP analysis and detection [74].

3 | PEPTIDE- AND PROTEIN-BASED BIOMARKERS

A biological marker can be defined as a characteristic that is objectively measured and evaluated as an indicator of physiological and pathological processes or pharmacological responses to therapeutic intervention [75]. Although the term biomarker became more widely known around the 1980s, biomarkers were around long before that. The term "honey urine" was first used by the ancient Hindus, approximately 1000 years before Europeans noticed the sweetness of urine in diabetics. The Hindus of this period first described conditions such as polyuria and glycosuria when they noticed that the urine of certain sick people attracted insects [76].

So far, more than 22 000 human diseases have been identified, but only a fraction of them has a sufficiently accurate, sensitive, and specific diagnostic method [77]. Despite decades of intensive research and over 700 000 scientific publications referring to biomarkers [78], we can only find dozens of clinically relevant biomarkers in oncology [79]. The failure of the biomarker transformation from the research laboratory to clinical use may be responsible for several factors, such as the inadequate experiment design, the lack of standardized procedures and quality control for sampling procedure and analysis, or the failure of the validation process [77]. Although a newly discovered promising biomarker has only a small chance and many obstacles on the way to routine use in clinical practice, the potential of the industry is enormous, as evidenced by the projected increase of the clinical biomarkers market size from the current 23 billion USD (in 2022) to over 54 billion USD in 2032 [80]. This growth is driven by the increasing prevalence of chronic non-communicable diseases, the growing pursuit of personalized medicine, the preference for non-invasive or minimally invasive methods, and the emphasis on early diagnosis to reduce mortality and treatment costs.

Peptides and proteins are attractive choices in the search for new biomarkers because of their many physiological and pathological functions. Many pathological events are manifested as changes in the peptides and proteins through their various PTMs, such as glycosylation. These macromolecules are also naturally present in the intercellular space and body fluids, such as serum, plasma, urine, saliva, sweat, or even exhaled breath condensate. Thanks to their occurrence in easily accessible compartments, these changes can be observed using non-invasive or minimally invasive sampling and subsequent proteomic analysis. These sampling techniques also require minimal expertise and are more affordable compared to, for example, biopsy samples. In this context, proteomic analyses of biological fluids can be a promising approach for discovering new biomarkers and further understanding pathophysiology, diagnostics, therapy strategies, or prevention for many diseases.

Endogenous proteins are mostly present in very complex mixtures and many of them occur at very low concentration levels. Complex biological matrices involve a wide variety of interfering compounds, exogenous and endogenous, which makes it impossible to directly analyze these samples. Traditionally, immunoaffinity methods, such as western blot or ELISA, have been preferred for the quantification of proteins in biological matrices [81]. Immunoaffinity methods are characterized by excellent sensitivity, but their effectivity varies due to the quality of used antibodies and cross-reactivity [82]. MS-based methods are newer techniques for protein analysis compared to immunoaffinity methods, although the invention of both dates back to the beginning of the 20th century. Due to its independence from antibodies, MS-based proteomics provides high analytical specificity, distinguishing protein isoforms, and is an ideal technique for the discovery, verification, and validation of novel proteomic biomarkers. When using MS detection, matrix components can alter the ionization efficiency of a target analyte (ion suppression or enhancement) by interfering with its ionization and causing the matrix effect. Therefore, an efficient separation technique before MS detection is necessary. Although the majority of methods utilized for proteomic biomarker analysis in complex biological samples represent LC-MS methods [11], the incorporation of electromigration separation methods is still growing as manifested in several review papers [35, 44–46, 10, 83–85].

Protein analysis in biological matrices often requires a pre-separation and/or preconcentration of proteins before their analysis by separation methods. Proteins can be extracted from biological matrices employing various off-line and on-line techniques such as homogenization, centrifugation, precipitation, liquid-liquid extraction, SPE, SPME, (micro)dialysis, and ultrafiltration [86, 87]. Even though, there is an effort to develop non-immunoaffinity sample preparation methods for the extraction of multiple intact proteins from biological matrices [88], sample enrichment methods based on immunoaffinity are still predominantly applied for intact protein analysis. Sample preparation techniques for extraction and fractionation of intact proteins from biological matrices for MS analysis were extensively reviewed in our recent paper [89]. The current advances in the analysis of peptide and intact protein biomarkers by CE-MS are discussed in **Chapter 5.2**.

3.1 | Biological fluids

Biological fluids are the source of many proteins for proteomic analysis and are extremely complex matrices due to the amount and variety of chemicals they contain. Biological fluids consist of water, dissolved or dispersed salts, lipids, metabolites, and various proteins with countless chemical structures, molecular weights, and properties. Biological fluids can be classified, according to the degree of invasiveness during their collection, to biological fluids obtained by minimally invasive methods (urine, saliva, sweat, or tears), more invasive methods (blood, blood plasma, serum, or amniotic fluid) and highly invasive methods (cerebrospinal fluid, mammary duct fluid, pleural effusion, or ascitic fluid) [90]. The most frequently collected biological fluids suitable for diagnosis and the search for potential biomarkers are blood and urine, which contain proteins secreted by various damaged organ tissues.

Blood is a complex mixture of red and white blood cells and platelets dispersed in blood plasma. It has a pH ranging from 7.35 to 7.45. Blood plasma is a liquid component that makes up approximately 55% of blood and contains fibrinogen. Blood plasma can be obtained when blood is collected in the presence of an anticoagulant (e.g., EDTA or heparin) and the blood cells are removed by centrifugation. Plasma consists of 90% water along with lipids and several water-soluble substances, nutrients, regulatory proteins, antibodies, hormones, and electrolytes. Blood serum is also a liquid component of blood, but it is acquired after blood has naturally clotted. Unlike blood plasma, blood serum does not contain fibrinogen or other coagulation factors, but also contains proteins, antibodies, antigens, hormones, and electrolytes [91].

Urine is formed in the kidneys by ultrafiltration of blood plasma. The main part of urine (approximately 94%) consists of water and hormones, metabolites, and electrolytes. Urine pH ranges from pH 4.6 to 8. Serum proteins are filtered in the glomeruli based on their size, charge, and shape [92]. After passing through the glomeruli, the highly abundant serum proteins are reabsorbed in the proximal renal tubules. The concentration of proteins under physiological conditions in the urine of a healthy person is very low (less than 100 mg/L), which is about 1000 times less than in other biological fluids, such as plasma [93]. Urine is a biological fluid that can be obtained non-invasively in large volumes, making it an excellent candidate for studying biomarkers [94]. A recent review paper describes the latest applications of CZE-MS for the analysis of biologically significant substances in urine [95].

Cerebrospinal fluid is a colorless liquid found in the brain and spinal cord. It contains lower concentrations of proteins than the plasma and has a different representation of electrolytes. The collection of cerebrospinal fluid is extremely invasive, but it can be used, for example, in the study of biomarkers involved in neurological diseases [96].

Saliva is a valuable resource for the search for known and potential biomarkers because its collection is very simple and much less invasive than blood or cerebrospinal fluid collection. Quantification of the salivary proteome has been used to search for potential biomarkers that could contribute to the early diagnosis of diseases [97].

Other biological fluids such as tears and sweat are relatively easy to obtain by non-invasive methods but are not routinely used in proteomic biomarker analysis because they have a rather variable composition and proteins are found in only negligible amounts in these biological fluids.

4 | CE-MASS SPECTROMETRY

In the late 80s, CE emerged as a powerful technique to analyze biomolecules due to its high-resolution power, fast analysis, minimal sample (nl, pg), buffer consumption, and low operating costs [98]. CE analysis of large biomolecules, having low diffusion coefficients in the liquid solution, leads to outstanding separation efficiency, up to one million theoretical plates [99]. On-capillary detection avoids dead volumes, and CE capillaries with a lack of stationary phase cause less risk of nonspecific interactions and injection-to-injection carryover. Aqueous CE conditions represent an eco-friendlier alternative to LC separation, are generally more compatible with biological samples, and preserve the high-order structure of proteins. Contrarily, the presence of high concentrations of organic solvents in LC could manifest relatively easily into conformational changes and protein denaturation. Ultralow sample volumes are especially useful for the analysis of samples of biological origin such as body fluids and biopsies. CE has also the capability for high-resolution separation of protein complexes under native conditions [100] and the analytical throughput can be increased also by multisegment injection [101]. Moreover, CE represents a complementary approach with a different separation mechanism to more traditional and commonly applied chromatographic approaches.

CE is typically performed in fused silica capillaries with inner diameter in a range of 10–75 μm and the typical length of a capillary is from 20 to 100 cm. Separation is based on the different migration velocities of the analytes in the applied electric field according to the ratio of the charge to the size of the molecule. Proteins and their proteoforms differ by charge, size, mass, and steric conformation. Some PTMs are relatively bulky and increase the ionic radius, reducing the electrophoretic velocity of the protein. Other PTMs alter both protein mass

and charge causing changes in the electrophoretic mobility. CE methods can be employed for the determination of several important physicochemical parameters of proteins, for example, effective mobility, effective charge, pI, molecular mass (M_r), diffusion coefficient, the binding or dissociation constants of their complexes, and kinetic parameter of their reactions and interactions with other molecules.

In CE analysis of proteins, UV-absorption spectrophotometry, and MS are the most frequently used detection schemes, followed by fluorescence detection (mostly in laser-induced fluorescence mode). MS as both universal and selective, highly sensitive and information-rich analytical method has become the most powerful detection mode in proteomic analysis. In proteomic research, MS detection is mostly combined with LC separations [11, 12], however, CE-MS is often employed in diverse proteomic applications as an alternative and/or complementary technique to LC-MS [15, 34, 35, 45, 51, 102, 103]. For the analysis of peptides and proteins, CE is mostly hyphenated with MS detection on-line via ESI and much less off-line via MALDI [33]. CE-MS interfaces are discussed more in **Chapter 4.4**. The composition and concentration of BGE are important parameters that should be carefully selected when developing a CE-ESI-MS method. Non-volatile constituents, surfactants, and high buffer concentrations lead to analyte signal suppression, high background signals, and source contamination. To prevent the entrance of ESI-interfering compounds to the ESI-MS interface; partial filling techniques, alternative ion sources for ESI, or on-line/off-line multidimensional separation systems are utilized [104]. CE separation is typically fast with a separation window in a range of 1–30 min. However, the narrow separation window limits the number of MS or MS/MS spectra that can be acquired during one run, which can be problematic for complex top-down proteomic analysis. Currently, almost all known types of MS analyzers can be coupled to a CE system. Low-resolution mass analyzers, quadrupole, and ion traps can cover a mass range of up to m/z 2000–4000, allowing the detection of multiple charged proteins obtained from ESI. High-resolution TOF is the most applied mass analyzer for the analysis of intact proteins followed by Fourier-transform ion cyclotron resonance and Orbitrap mass analyzers. Intact protein fragmentation has long been a challenge. Except for the most frequently used collision-induced dissociation, techniques that enable efficient intact protein fragmentation—like electron transfer dissociation and higher energy collision dissociation —have been also introduced for CE-MS analysis [99, 105]. Bioinformatic tools for intact proteins, for example, ProSight [106], TopPIC [107], Proteoform Suite [108], MASHSuite [109], and pTop [110] are also a necessary part of complex top-down proteomic analysis.

To further improve the performance of CE, mCE devices were developed, for example, the ZipChip approach (908 devices) with the chips inserted directly in front of the MS orifice and automatically positioned to create the optimal electrospray. The application of mCE dramatically increased the throughput of the analysis. It enables the separation of proteins in a few seconds to a few minutes thanks to the short capillary length and high separation efficiency. mCE is on-line coupled to ESI-MS through on-chip emitters, such as the sharp corner of the microchip or the monolithically fabricated nanospray tip. The off-line approach with MALDI-MS is utilized rarely. For more information about mCE coupled with MS, the reader is referred to the recent reviews [14, 29–31].

Besides multiple advantages, CE has also several shortcomings [98]. One major issue during CE analysis of proteins is proteins adsorption on the fused-silica capillary wall via binding to the free silanol groups, thus influencing the separation efficiency. Another major drawback is the relatively low concentration sensitivity of CE, even when hyphenated with more sensitive MS detection, due to the low injected sample volume (nanoliter scale).

4.1 | Prevention of protein adsorption

Adsorption of biomolecules (especially basic peptides and proteins) on bare-fused silica capillary walls is an unwanted phenomenon as it strongly negatively influences separation efficiency and precision, as well as EOF consistency [111]. The adsorption process may even be irreversible, making proper analyte quantitation and further use of the capillary almost impossible. The protein isoelectric point (pI) and its relation to the pH of the BGE determines the adhesion of a protein onto the capillary surface. If the pH = pI, protein attachment is most pronounced due to the minimized lateral repulsion between adhered proteins [112]. The new method was recently developed by Leclercq et al. [113] for quantification of the protein adsorption onto the coated or uncoated capillary based on the determination of the retention factor and separation efficiency of individual proteins at different separation voltages. Various strategies were reported to reduce protein adsorption by either modifying the density of charge on the protein or the capillary wall.

4.1.1 | Appropriate rinsing protocol between consecutive CE runs

The proper rinsing of the capillary is essential to minimize protein adsorption and to promote the desorption of protein residues prior to the next CE run. Some

post-adsorption conformational changes of proteins can occur and turn protein adhesion into an irreversible event within 24 h and change the capillary surface [114, 115].

4.1.2 | BGE with an extremely low or high pH

When pH > pI of the protein, both the protein and silica surface are fully negatively charged, and an electrostatic repulsion occurs, preventing protein adsorption. In the case of pH < pI of the protein, the silica wall surface and a protein are counter-charged and electrostatic adsorption occurs. However, once pH ≤ 2.0, protein adsorption weakens by reducing the silica net charge close to zero (pKa of silica surface is estimated to be 4.5). Low pH will also suppress the EOF resulting in slower CE separation of proteins with low pIs and enlarging the separation window (necessary for large-scale proteomics). On the other side, one must keep in mind that working at an extreme pH reduces charge differences between proteins, may cause protein unfolding, and increase the risk for aggregation and precipitation (especially under highly acidic conditions) [116, 117]. This approach is also not suitable for native proteomics.

4.1.3 | Higher ionic strength of BGE

Increased BGE concentration promotes the competition of electrolyte ions with proteins for available adhesion sites on the capillary surface. However, it also increases Joule heating resulting in peak broadening and worse resolution [118]. This approach is also problematic with MS detection, as concentrated BGEs decrease the ionization efficiency.

4.1.4 | Addition of organic solvents into the BGE

Hydrophobic interactions and/or protein conformational changes can also be involved in the adsorption process, therefore the addition of organic solvents into BGEs can prevent protein adsorption. The influence of organic solvents on protein adsorption can be different and must be tested case by case [119]. Generally, this approach is very convenient, especially for on-line CE-ESI-MS analysis. The increased volatility of solvents lowers ESI currents and increases ionization efficiency. The reduced number of side electrochemical reactions at the ESI tip stabilizes the ESI current and decreases background noise. Organic solvent addition also alters separation selectivity and/or efficiency, reduces Joule heating, and, in certain cases, enhances the analyte solubility.

4.1.5 | Capillary coatings

Capillary coatings prevent either electrostatic and/or hydrophobic interactions or create electrostatic repulsion of proteins with a capillary wall. Coatings differ in the mode of attachment of the coating material (dynamic and permanent) and the charge of the coating. Charged coating materials (positively or negatively charged) suppress ionic interactions. Neutral coatings cover the silanol groups. A good capillary coating should be highly stable over a broad pH range and affect minimally or not at all the separation and/or detection system. Capillary coatings also enable an adjustment of EOF making it possible to increase the differences in the migration times of proteins and thus improving selectivity and enhancing the separation window. Suppressed EOF is necessary to perform several CE modes, such as CGE, CIEF, ITP, and ACE. However, when a CE-MS combination is performed a constant and high EOF is needed to maintain a stable electrical circuit. The analysis of complex samples can be problematic when the separation window is small due to the too-high EOF. Cationic coatings lead to the reversed EOF, and the improvement of the separation window is usually modest due to the strong EOF inside of the capillary. In the case of neutral coatings, the absence of a constant flow rate can lead to problems with ESI stability and longer separation times. This problem can be overcome by positive pressure at the capillary inlet. Principles and molecular mechanisms of protein adsorption as well as prevention strategies have been discussed in a couple of reviews and the reader is referred there for more details [114, 120, 121]. Coatings used for intact protein analysis by CE-MS are listed in Table 1.

Dynamic coatings represent small reagents adhered onto the capillary surface and thus masking the negative capillary charges. Their attachment is weak and only temporary, thus regular regeneration is mandatory. These coating reagents are structurally mostly polymers (polyamines or polysaccharides). Neutral coatings represent poly(ethylene oxide), cellulose derivatives, dextrans, and methyl chitosan. Cationic coatings include polybrene, polyarginine, and ionic liquids. Hydroxypropyl methylcellulose-coated capillary is considered a reference coating for mAb charge variants analysis [21]. The advantage of dynamic coating is that it is simple, but its handicap is that coating agents are in BGEs and can interfere with the separation or detection, therefore these coatings are not recommended for CE-MS [120].

Permanent coatings are either physically adsorbed or covalently bonded onto the capillary inner wall creating stable surface coverage and enabling to perform separations without the addition of the coating materials into the BGE. Physically adsorbed coatings have simpler

TABLE 1 Capillary coatings used for intact protein analysis by on-line CE-MS.

Coating	Type	Surface charge	References
polyarginine	dynamic	cationic	[254]
linear polyacrylamide (LPA)	permanent (covalent)	neutral	[122, 231]
polyvinyl alcohol (PVA)	permanent (covalent)	neutral	[235–237]
hydroxypropyl methylcellulose (HPMC)	permanent (covalent)	neutral	[236]
N-acryloylamidoethoxyethanol	permanent (covalent)	neutral	[255]
PS1 neutral coating	permanent (covalent)	neutral	[239, 253, 238]
linear carbohydrate polymer (LCP)	permanent (covalent)	neutral	[248]
polyethyleneimine (PEI)	permanent (covalent)	cationic	[229, 230, 233, 234]
(3-aminopropyl) diisopropylethoxysilane	permanent (covalent)	cationic	[256]
omega-iodoalkylammonium salts (M7C4I)	permanent (covalent)	cationic	[232]
polybrene (PB)	permanent (adsorbed)	cationic	[122]
PB-dextran sulfate-PB	permanent (adsorbed)	cationic	[122, 144]
successive multiple ionic polymer layer (SMIL)	permanent (adsorbed)	cationic	[122, 125, 126]
anionic polymer	permanent (adsorbed)	anionic	[257]

coating procedures, easy coating regeneration, and minimal dependence on surface chemistry. The preparation of covalent coatings is usually more laborious and time-consuming. These coatings have long-term stability, and good compatibility with MS detection but have low stability at extreme pH and are impossible to regenerate if the coating deteriorates. Permanently coated capillaries are commercially available, but the id is usually limited to 50 μm or greater and the price is usually high. Successive multiple ionic polymer layers coating, which combines alternating polycationic and polyanionic polymer layers boosting the durability of the coating, was successfully used in several recent publications [122–126], however, more studies are needed to investigate the performance of successive multiple ionic polymer layers for CE-MS top-down proteomics. An interesting approach for on-line preconcentration of proteins through switchable protein adsorption/desorption properties of mixed polymer brushes coatings was applied for on-line preconcentration of lysozyme in hen egg white [127, 128], bovine serum albumin [129] and pepsin [130]. The achieved LOD of lysozyme was 4.5 pg/ml, which represents 100 000 times lower LOD compared to an uncoated capillary.

Despite a strong effort and research in innovative types of capillary coatings, the effective suppression of protein adsorption to the inner capillary wall still remains one of the major challenges and the entire prevention of protein adsorption in CE is so far a myth. Numerous reports showing high plate numbers and sharp peaks are reported only for a set of selected highly purified standards, however, when complex biological matrices are analyzed, the results would most likely change, as also matrix components will be adsorbed on the capillary walls.

4.2 | Enhancement of sample loading capacity

Small injection volumes in CE are responsible for relatively poor mass concentration detection sensitivity even when sensitive MS detectors are used. Trace analyte analysis in complex biological matrices requires proper preconcentration methods. Two different approaches can be utilized for analyte pre-concentration: 1) traditional extraction methods that are used for sample pre-treatment and sample clean-up, and 2) on-line electrodriven sample stacking procedures that are unique features of CE.

The main microextraction methods that were coupled to CE-MS are SPME, hollow-fiber liquid-phase microextraction, electro-membrane extraction, and dispersive liquid-phase microextraction methods. These methods usually yield enrichment factors of 10 or 100, are time-consuming, tedious, and commonly operated in an off-line scheme [32], however on-line SPE in combination with CE methods was also developed [131, 132]. SPE offers a significant sample complexity reduction, however, the eluted sample volume is often larger than 1%–2% of the conventional capillary volume used for the sample injection. Additional stacking techniques are therefore still required for appropriate sensitivity enhancement. Kuzyk et al. [51] provided a decision flowchart to ease the choice of the proper technique to use (Figure 1).

On-line preconcentration techniques, such as field-amplified sample stacking, large-volume sample stacking, field-enhanced sample injection (FESI), dynamic pH junction (DPJ), transient ITP (tITP), electrokinetic supercharging (EKS), and sweeping employ various electrophoretic effects using different electrolyte systems

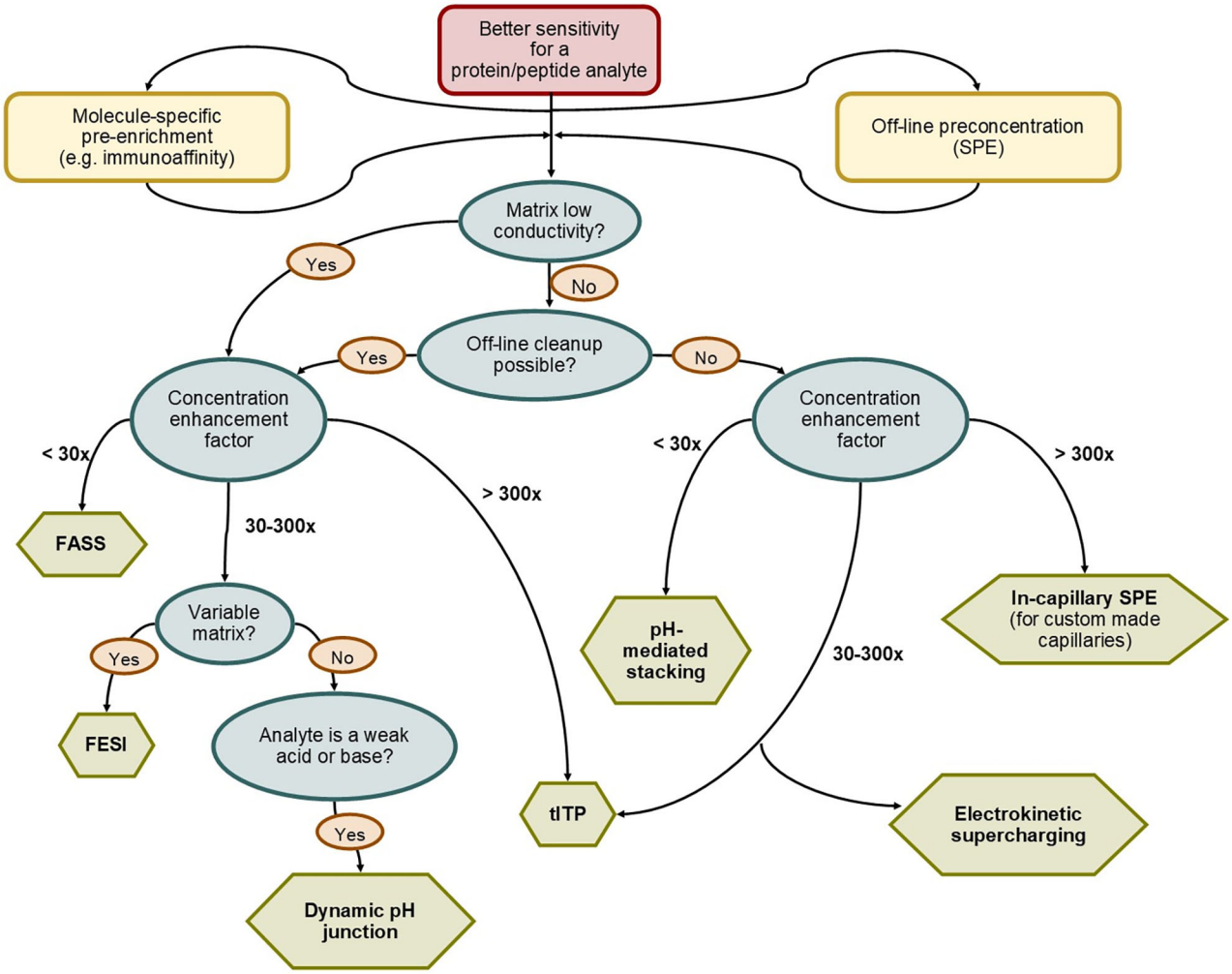


FIGURE 1 Decision flowchart to select the most suitable CE sample preconcentration technique. Modified and reproduced from Kuzyk et al. [51].

and can be performed right inside the CE capillary. Their main advantage is that the loading sample volume can be increased while resolution remains unaffected leading to increased detection sensitivity of the method. Several recent reviews devoted to on-line CE preconcentration were published [133, 134]. Despite multiple advantages, the number of these in-capillary focusing methods, especially methods using non-volatile micelles such as sweeping, are incompatible with CE-MS. Since the CE-MS technique is incompatible with samples with high content of salts, in-capillary preconcentration methods such as field-amplified sample stacking, FESI, and large volume sample stacking, which are suitable for low ionic strength samples, are widely used in CE-MS. However, biological fluids are natural sources of higher content of salts and for those samples are these in-capillary preconcentration methods very challenging. pH-driven focusing techniques such as DPJ, and tITP are suitable for amphoteric compounds including proteins, peptides, and amino acids and for samples with a high concentration of salts, therefore those are the methods mostly applied for CE-MS analysis of intact proteins [32]. The DPJ mechanism lies in the differential pH value at the interface between the sample zone and the electrolyte zone, where changes in charge and electrophoretic mobility of the amphiphilic analytes occur, resulting in a focus of the analyte zone at that interface [135]. tITP uses the isotachophoretic principle and can be applied to concentrate analytes using suitable electrolytes, leading (LE) and terminating (TE), containing ions with higher and lower electrophoretic mobility than the analytes, respectively. A plug of a sample dissolved in LE and a plug of TE are sequentially introduced into the capillary. Applying voltage to the separation capillary first results in the separation, concentration, and focus of underrepresented analytes depending on the concentration of the leading ions by the principle of isotachophoresis, and then the electrophoretic separation occurs. The schemes of tITP and DPJ principles are shown in Figure 2 and an overview

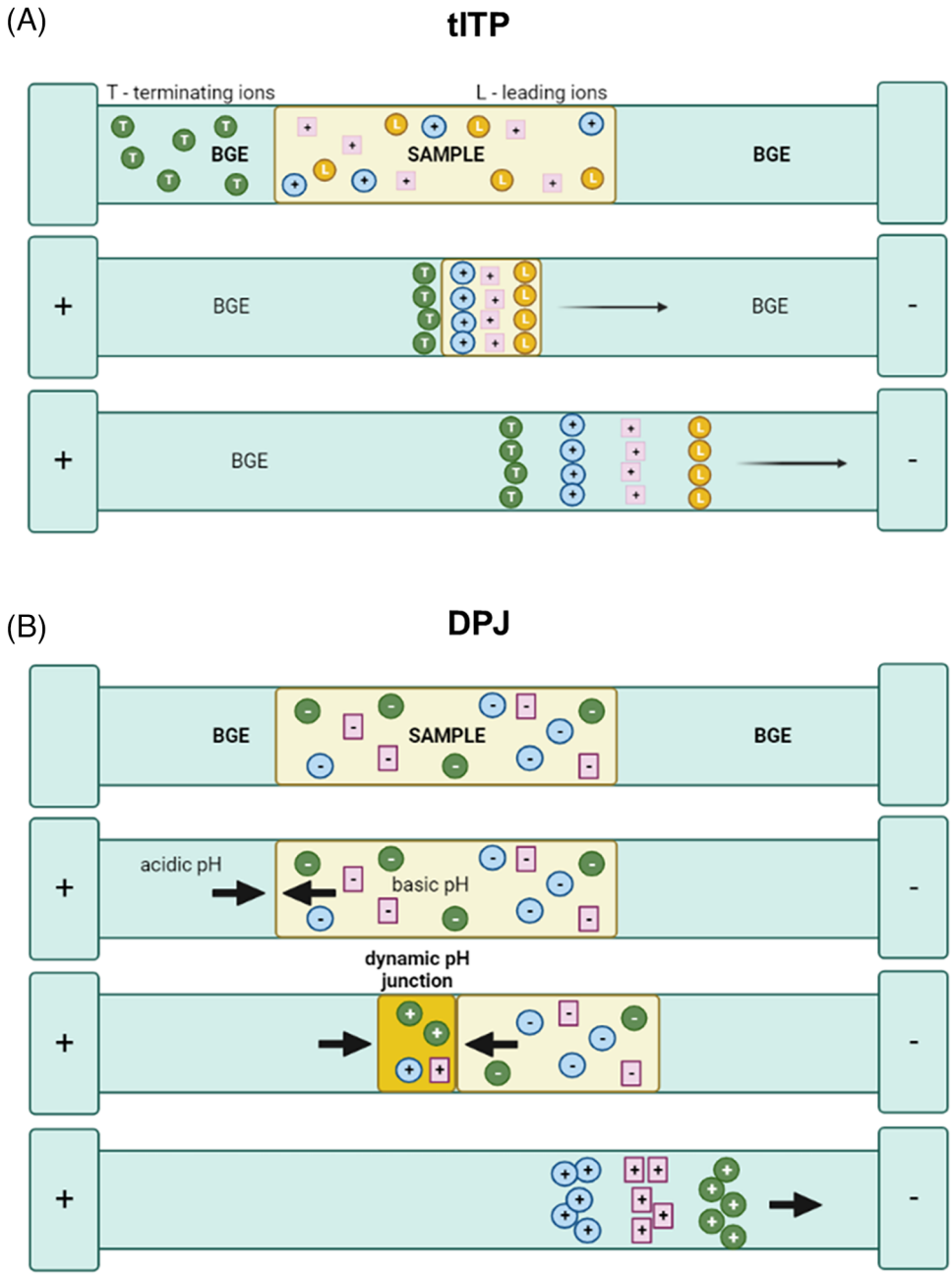


FIGURE 2 The mechanism of stacking by (A) transient ITP and (B) dynamic pH junction.

of their applications on intact protein and peptide analysis in different matrices is presented in Table 2.

Recently, a novel approach using voltage polarity switching transient ITP (PS-tITP) was introduced, where the sample ions move back and forth in a separation capillary during their initial tITP focusing stage by switching the separation voltage polarity, in order to increase the sample loading volume and improve the separation efficiency. Different modes of PS-tITP were investigated [136, 137] and an increased sample loading volume to about 70% of the separation capillary without compromising the resolution can be achieved.

Various preconcentration techniques can be also combined to achieve multiple stacking which can lead to selective interference removal, analyte enrichment, and in final a clearer and more informative electropherogram. EKS is a two-step stacking technique that employs tITP to preconcentrate analytes after a significantly long FESI [87]. This combination can provide significantly higher enrichment factors than tITP or FESI can achieve alone. EKS has limited applicability in highly conductive sample matrices since electrokinetic sample injection is used in the FESI step. This shortage can generally be overcome by significant dilution of the sample or applying a sample cleanup

TABLE 2 Applications of on-line sample preconcentration (transient ITP [tITP] and dynamic pH junction [DPJ]) in CE-MS for intact proteins and peptides.

Analyte	Matrix	CE	Detection	Sample pretreatment	Ref.
lysozyme, cytochrome c, β-lactoglobulin, ribonuclease A, myoglobin, carbonic anhydrase	water	LPA coated capillary (75 μm id) BGE (0.02 M 6-aminohexanoic acid, pH 4.4)	QQQ (sheath-liquid)	tITP LE: 0,01 M AmAc, pH 4.4 TE: 0.001 M HAc	[258]
model protein mixture (M_r 8–29 kDa); P. aeruginosa proteome (M_r 30–80 kDa)	water, PA01 whole cell lysate	polyarginine coated capillary (50 μm id)	Q-Orbitrap (electrokinetically pumped sheath-flow nanospray)	ACN precipitation, GELFrEE fractionation transient pseudo-ITP (60% IPA, 0.8% HFo)	[254]
cytochrome c, myoglobin, β-casein	water	LPA coated capillary (50 μm id) BGE (0.1% HFo)	ion trap (electrokinetically pumped sheath-flow nanospray)	DPJ (10 mM AmAc, pH 6.5)	[259]
model protein mixture (M_r 12–66 kDa, pI 4.5–11), *E. coli* proteome	water, cell lysate	LPA coated capillary (50 μm id) BGE (5% or 10% HAc)	ion trap, Q-Orbitrap (electrokinetically pumped sheath-liquid)	DPJ (50 mM ammonium bicarbonate, pH 8.0)	[260]
model protein mixture (M_r 16.9–66.5 kDa), E. coli proteome	water, cell lysate	LPA coated capillary (50 μm id) BGE (10% HAc, pH 2.2)	Q-Orbitrap (electrokinetically pumped sheath-liquid)	DPJ (50 mM ammonium bicarbonate, pH 8.0)	[99]
model protein mixture (M_r 5.8–29 kDa), yeast proteome	water, cell lysate	LPA coated capillary (50 μm id) BGE (5% HAc)	Q-Orbitrap (electrokinetically pumped sheath-liquid)	DPJ	[261]

(Continues)

TABLE 2 (Continued)

Analyte	Matrix	CE	Detection	Sample pretreatment	Ref.
peptidic hormones	rat hypothalamus tissue	BFS capillary (50 μm id)	Q (sheath-liquid)	tITP LE: 50 mM AmAc, pH 4.8 TE: 50 mM HAc	[262]
Pyrococcus furiosus proteome	cell lysate	PEI coated capillary (30 μm id) BGE (0.1% Hac, 20% IPA)	Orbitrap (sheathless)	tITP LE: 50 mM AmAc, pH 4.5 TE: BGE	[263]
model peptide mixture	cell lysate	BFS capillary (50 μm id) BGE (1 M HFo, pH 2.0)	UV (200 nm); Ion trap (sheath-liquid)	DPJ (100 mM borate buffer, pH 10.0; 50% ACN)	[264]
E. coli proteome	cell lysate	LPA coated capillary (50 μm id) BGE (10% HAc, pH 2.2)	Q-Orbitrap (sheath-liquid)	SEC, RPLC prefractionation DPJ (50 mM ammonium bicarbonate, pH 8.0)	[42]
rRGD-hirudin	human urine	BFS capillary (70 μm id)	Q (sheath-liquid)	DPJ (100 mM ammonium bicarbonate)	[265]
model peptide mixture	human urine	BFS capillary (50 μm id) BGE (0.5 M HFo, pH 2.15)	Q (sheath-liquid)	DPJ (50 mM AmAc, pH 7.5)	[266]
opioid peptides	plasma	BFS capillary (75 μm id)	DAD (195 nm); TOF (sheath-liquid)	ACN precipitation, MWCO (3 kDa), online SPE, tITP (LE: 0.25% ammonium, TE: 0.1% HAc)	[267]

Abbreviations: ACN, acetonitrile; AmAc, ammonium acetate; BGE, background electrolyte; CE, capillary electrophoresis; DAD, diode-array detector; DPJ, dynamic pH junction; BFS, bare fused silica; GELFrEE, gel-eluted liquid fraction entrapment electrophoresis; HAc, acetic acid; HFo, formic acid; IPA, isopropyl alcohol; LE, leading electrolyte; LPA, linear polyacrylamide; MS, mass spectrometer; MWCO, molecular weight cut-off; PEI, polyethylenimine; Q, quadrupole; QQQ, triple quadrupole; RPLC, reversed-phase liquid chromatography; SEC, size-exclusion chromatography; SPE, solid-phase extraction; TE, terminating electrolyte; tITP, transient isotachophoresis; TOF, time-of-flight.

step to minimize the harmful effects of the salt. Nyssen et al. [138] reported a sheathless CE-ESI-MS/MS method for the determination of parathyroid hormone and its variants and using the EKS preconcentration step the high pg/ml LOQ range was reached.

A new approach named velocity gap was suggested for the selective enrichment of low-abundance lysozyme (8 μg/ml) from the model mixture comprising high-abundance BSA (36 mg/ml) [139]. The fractionation and enrichment were accomplished by modification of electric field strengths in different sections of the capillary. Only the front sample plug containing lysozyme was enriched by stacking and about 99% of highly abundant BSA was removed.

4.3 | CE modes for intact protein analysis

The most popular CE modes in the intact protein analysis include SDS-CGE, capillary isoelectric focusing (regular, CIEF, and imaged, iCIEF, modes), and CZE. Other CE modes, such as ITP, EKC, CEC, and affinity electrophoresis (ACE) have also been applied for various proteomic applications.

When analysis of complex protein samples, such as body fluids, tissue extracts, and cell lysates is required, the application of multidimensional systems is often necessary. The combination of LC and CE techniques is even being referred to as a "Swiss knife" for proteomics investigations due to their complementarity [140] and, for example, heart-cut nano-LC-CZE-MS for intact proteins mixture [141] and SDS-CGE-RPLC-MS for top-down identification of mAb fragments [142] were developed. The main advantage of 2D CE-MS setup is the possibility to perform a high resolving CE separation in the first dimension in any CE mode (CZE, CIEF, and CGE) with non-volatile BGEs followed by a second dimension able to fully remove the ESI-interfering components from the analytes prior to MS detection to provide interference-free mass spectra [104].

4.3.1 | Capillary zone electrophoresis

The simplest and most widely used capillary electromigration technique, also for intact protein analysis, is CZE due to its simplicity, high throughput, and flexibility regarding analysis conditions (denaturing or native conditions, MS-compatible or not). CZE can be used to confirm protein identity, detect impurities and PTMs, and characterize charge heterogeneity. CZE has grown in popularity and is being added to biologics release testing panels.

Generally, CZE separates proteins based on their different migration velocities (i.e., charge-to-size ratio) contrary to RPLC which separates analytes based mainly on their hydrophobicity. Typically, the sample is injected as a plug at one end of a capillary filled with a BGE buffered at a pH providing the desired mobilities and capable of maintaining a constant field strength. Choice of BGE conditions (composition, pH, and ionic strength), and capillary surface are key elements for successful CZE analyses of intact proteins. The choice of pH should consider the protein solubility together with the capillary coating stability as well as the detection type. The CZE–ESI–MS system requires the use of volatile solutions and is incompatible with the presence of salts that would cause adduct formation and ion suppression. Therefore, the choice of the sample matrix and BGE is limited for CZE–ESI–MS, typically to acetic and formic acid, ammonium acetate, and ammonium formate. Acetic acid is often used instead of formic acid for CZE separations due to its very low conductivity which avoids Joule heating. Ammonium acetate (pH 5.0–7.0) is the main sample buffer and BGE for native or near-native CZE-ESI-MS characterization of mAbs stability and aggregate formation [143, 144]. It preserves non-covalent interactions and conformational heterogeneity of a protein. Protein samples can also be prepared in denaturing solutions, mainly aqueous solutions of acid and organic solvent, to reduce sample zone conductivity and enhance the resolution of separation. Such sample matrix causes protein unfolding and dissociation of non-covalent interactions, thus enhancing the ESI efficiency by increasing a charge state. Additives such as organic solvents (ACN, methanol, and isopropyl alcohol) sometimes enhance protein solubility and decrease hydrophobic interactions between different protein variants.

Besides aqueous BGEs, a broad range of organic solvents (alcohols, ACN, etc.) can be seen in a nonaqueous CE-MS (NACE-MS) [145]. Improved solubility and reduced aggregation of hydrophobic macromolecules in the presence of organic solvents lead to the stable and robust analysis of biomolecules [146, 147]. NACE could be a potential method for the quality control of biopharmaceuticals. However, the use of NACE in biological samples is focused on special applications since biological matrices are water-based [148].

4.3.2 | Capillary gel electrophoresis

CGE or capillary sieving electrophoresis (CSE) is now a recognized analytical tool for the characterization of size heterogeneities during the development and quality control process of biopharmaceuticals, where it complements

SEC [149]. It allows the investigation of protein stability and purity (product-related impurities, fragmentation, and aggregation), accurate and precise quantification of proteins, and N-glycan analysis [64, 66].

In CSE (CGE), distinct sieving matrices, such as replaceable and water-soluble linear or slightly branched polymers, for example, linear or cross-linked polyacrylamide, PEG, poly(ethylene oxide), and dextran, are used as a separation medium and act also as dynamic capillary coatings. The most common (and commercial) approach uses a capillary consisting of an inner surface composed of bare-fused silica and filled with a sieving matrix of dextran cross-linked with borate [64]. CGE analysis is performed under denaturing conditions with surfactants, mostly sodium dodecyl sulfate (SDS), or sodium hexadecyl sulfate [149, 150]. Additives such as glycerol serve to enhance the resolution of separation.

When sensitivity is a concern, the signal can be enhanced through the use of CE with laser-induced fluorescence detection either using native fluorescence or after proper labeling [66]. Online hyphenation of MS to SDS-CGE has not yet been achieved at the commercial stage and requires more indirect approaches. The main bottleneck is high ion suppression caused by the sieving matrix components, such as SDS and high concentrations of co-ions. Different in-capillary strategies, tITP, injections of preplugs and post-plugs of organic solvents, stripping agents (cyclodextrins), or cationic surfactants, have been tested to remove SDS before MS detection [151, 152]. Another option is the use of multidimensional systems, heart-cut SDS-CGE-CZE-MS [153] and SDS-CE-RPLC-MS [142]. Microchip CGE was utilized for the quality control and purity of mAbs [154, 155] and the detection of C-reactive protein in serum samples [156].

A new method referred to as packed CE, where capillary is packed with nonporous colloidal silica creating arbitrarily wide pores, thus, functioning as CGE/CSE, was applied for separation of protein standards (11–155 kDa) and stressed pharmaceutical-grade IgG4 protein sample, giving baseline resolution of monomer, dimer, trimer, and tetramer in less than 10 min [157].

4.3.3 | CIEF and imaged CIEF

CIEF is a high-resolution electrophoretic method with a unique capability to separate proteins or other (poly)ampholytes according to their pIs. A mixture of ampholytes and a sample is used to establish the pH gradient inside a neutral-coated capillary (e.g., fluorocarbon, polyvinyl alcohol, and dynamic coating). After the application of an electric field, a pH gradient is established and proteins are focused on the position with pH equal to their pI value, and then forced to move toward the detector using chemical, electrophoretic, or hydrodynamic mobilization. When the proteins are focused into highly concentrated bands at their pI precipitation and aggregation can occur, causing clogging of capillary or irreproducible results. To avoid such issues and enhance protein solubility, additives like surfactants and denaturants (e.g., urea, sucrose, glycerol, sulfobetaine, and taurine or their mixture) must be added to the sample [158]. Urea is the most used one, however, the pI shifts may be induced by its denaturing effect [159] and protein carbamylation by isocyanic acid derived from urea [160]. Since the mobilization step tends to broaden peaks and increase analysis time, imaged CIEF (iCIEF) has been established to achieve shorter run time, higher throughput (the use of very short capillary only 5 cm length), better resolution, and reproducibility thanks to the whole column imaging technology within a transparent capillary [68].

In comparison with other CE modes, CIEF and especially iCIEF have the best resolution for the analysis of proteins and peptides. Resolving power depends on field strength and the slope of a linear pH gradient, with the shallower gradients improving the resolution. CIEF and iCIEF with UV detection (at 280 nm) have become the method of choice for the characterization of charge variants and have been implemented in quality control laboratories as reference methods for charge heterogeneity analysis and play an important role in the development of biosimilars [24, 159, 161, 162]. CIEF allows the determination of levels of acidic, main, and basic protein variants as well as degradation impurities or PTMs that impact pI value. Compared to IEC, CIEF separates analytes based on an overall charge, thus, a particular proteoform could not be separated by IEC but can be by CE and thereby yielding complementary information as orthogonal techniques.

CIEF has also been coupled to MS, nevertheless, hyphenation with MS remains problematic due to the high concentrations and non-volatility of ampholytes and additives and the transfer of the very narrow-focused zones into the MS without peak-broadening. Coupling strategies and applications of CIEF-MS have been reviewed by Hühner et al. [163]. Glycerol with its low conductivity and compatibility with MS could serve as an alternative additive to stabilize the pH gradient and interferences with MS detection could be avoided by using ACN in the sheath-liquid to break hydrogen bonds of glycerol with proteins [164]. Furthermore, CIEF is often used as the first step in the 2D-separations of complex protein mixtures for preconcentration of diluted proteins and the introduction of a second dimension, like CZE or RPLC providing a tool to separate the ampholytes from the analytes [165, 166]. iCIEF-MS analysis can be achieved by off-line fractionation where fractions are first collected, processed with

MS-compatible components if needed, and subsequently introduced into MS for characterization. Although more time-consuming and labor-intensive, off-line fractionation is not limited to the separation of analytes requiring only MS-compatible components and may be necessary to fully characterize some highly complex proteins at the intact level. The new preparative iCIEF instrument allows simultaneous isolation and characterization of individual protein charge variants and can be directly connected to MS to deliver the fractionated protein samples for characterization [167–170]. During pressure mobilization, an electric field keeps the samples in the separation capillary focused. The separation capillary has a larger inner diameter, and the transfer capillary has a smaller inner diameter. Hence, this difference greatly minimizes the remixing of separated protein isomers.

The conventional CIEF was successfully miniaturized into a microfluidic CIEF system, which provides comparable resolution and precision and ultrafast charge heterogeneity assessment [171]. Commercially available microchip iCIEF-MS is a promising new combination and show great potential for applications such as early developmental screening and monitoring of specific protein quality attributes [172–174].

4.3.4 | Capillary ITP

ITP is a separation mode of CE, where analytes are separated as adjoining successive zones in the order of their decreasing electrophoretic mobilities. ITP uses a discontinuous system of two electrolytes, the LE and the TE containing ions with electrophoretic mobility higher and lower than that of the sample, respectively. The sample plug is injected into the interface between LE and TE. The main characteristic of ITP is the so-called self-sharpening effect, therefore, there is no dispersion of individual zones, even though neighboring zones have a common interface. The principle of this mechanism is the increasing intensity of the electric field from the leading to the TE based on the Kohlrausch control function. The concentrations of analyte ions in individual zones are adapted to the concentration of LE ions and the result of this process is that although the more abundant components of the mixture may be diluted, the analytes at the low concentration level will be concentrated and the concentration range of the complex mixture will be significantly reduced [175]. Malá and Gebauer [176, 177] recently published a nice overview of the development of ITP from 1967 to 2022.

Challenges associated with working with proteins include that their solubility is sometimes limited and the fact that they exhibit a wide range of pH and buffer-dependent electrophoretic mobilities due to their different pIs and sizes. This diversity makes it difficult to develop a generally applicable ITP assay design [178]. Also, the low separation efficiency and difficulties in finding an appropriate spacer limit its applicability. Therefore, ITP is an effective part of multidimensional separations when combined on-line with other separation techniques, for example, Gysler et al. [179] and Bergmann et al. [180] applied an ITP-CZE-MS method for the rapid and efficient analysis of recombinant human interleukin-6 and interleukin-3. ITP can be also used in a preparative format [181–184], where single fractions can be manually transferred to other separation devices. Another interesting format of ITP is its on-chip use in microfluidic devices [178, 185]. ITP-MS is still rarely used because of a low number of MS-compatible electrolytes [186], although the work of Crosnier de Lassichere showed its applicability to analyze amyloid beta peptides in cerebrospinal fluid [187].

Applications of ITP also involve the use of ITP to aid chemical reactions involving proteins (e.g., for immunoassays) [188]. ITP was used to pre-concentrate and deliver target proteins in innovative surface-based immunoassays [189], the effective utilization of the focused sample was achieved by stop-and-diffuse ITP mode or by applying a counter-flow ITP mode. Using the green fluorescent protein as a model protein, a 1300-fold improvement in LOD compared to a standard immunoassay was achieved.

4.3.5 | Other CE modes for intact protein analysis

EKC is a hybrid electrokinetic and chromatographic method. Its major advantage is that it can separate both charged and uncharged compounds. The separation is based on different interactions of the analyte with the pseudostationary phase and if they are charged also on their different electrophoretic mobilities. In the major EKC method, MEKC, many different types of surfactants are used, such as anionic alkyl sulfates (e.g., SDS), cationic alkylammonium salts (e.g., cetyltrimethylammonium bromide), ampholytic surfactant (e.g., 3-[(3-cholamidopropyl) dimethylammonio] propane sulfonate), non-ionic surfactant (e.g., Brij 35) and other special pseudophase constituents, such as polymer surfactants and ionic liquids. EKC analyses of proteins were recently reviewed by Gao et al. [190]. MEKC-UV was evaluated for a mAbs charge heterogeneity study [191, 192]. Micelle-forming molecules like SDS are among the strongest inhibitors of ESI efficiency, therefore applying counter-migrating micelles or partial-filling techniques is needed to prevent the micelles (and their monomers) to enter the MS or volatile surfactants, for example, perfluorooctanoic acid, can be applied for direct MEKC-MS coupling [193].

CEC is a hybrid of chromatography with an electrically driven system (EOF-driven) combining the advantages of the CE (high separation efficiency) and LC techniques (versatility and higher loading capacity) [194]. Despite this fact, CEC did not become a widely used analytical method. The open-tubular mode of CEC with a stationary phase immobilized on the inner fused silica capillary wall has several advantages, such as easy preparation, avoidance of back-pressure problems, bubble formation, and frit fabrication [195]. A poly(norepinephrine)-coated open-tubular column for the separation of various basic and acidic proteins and glycoforms of recombinant human erythropoietin [196] and a fibrin coating for mAb variants characterization [197] were developed by Xiao et al. Advances in CEC for the analysis of proteins were recently reviewed by Hajba and Guttman [198] and general CEC-MS hyphenation in the paper by Della Posta et al. [36].

ACE utilizes the specific interactions between ligands, receptors, and/or antibodies and analytes as a separation mechanism and is used for selective analysis of proteins in complex mixtures, and especially for studies of protein interactions with variable ligands under native conditions via determination of the binding or dissociation constants of the formed complexes [199, 200]. ACE was successfully applied for charge heterogeneity profiling of biopharmaceuticals [201] and binding of mAb proteoforms to Fc receptors also in combination with MS detection [202, 203].

A new type of CE method, mobility CE (MCE), in which a constant liquid flow is used to replace EOF, has been applied for protein structure analysis and effective charge characterization [204, 205]. MCE separates the analytes based on their charge states, hydrodynamic sizes, and geometries. MCE has the capability of not only separating complex samples but also acquiring the sizes of biomolecules. MCE with native MS detection was used to determine the geometric structure of different globular proteins [206].

4.4 | CE-MS interfaces

A recent overview of CE-MS interfacing techniques for proteomics is published by Mikšík [34]. For the analysis of peptides and proteins, CE-MS is mostly hyphenated on-line via ESI and much less off-line via MALDI. The off-line CE-MALDI-MS arrangement has several advantages: (i) better tolerance to salts, (ii) possibility to store samples/analytes, (iii) independent optimization of CE separation and MS detection, and (iv) less multiply charged ions that simplify the interpretation of mass spectra. However, the most important problem with CE-MALDI-MS is maintaining electrical continuity when collecting a CE effluent at the capillary outlet. Fractions are transported onto the MALDI plate employing EOF, pressure, or sheath-liquid. Specific CE-MALDI-MS proteomic applications were described in the literature [207–209].

ESI is the most common MS ion source for large molecule analysis. Its integration with CE requires an interface to maintain stable electric contact at the CE outlet electrode and steady spray formation by ESI. Significant amounts of effort have been devoted to CE-ESI interface design and several different interfaces to hyphenate CE and ESI-MS have been described [210]. CE-ESI interfaces can be classified into three main categories: (i) sheathless (flow rate at range tens of nl/min), (ii) sheath-liquid (flow rate in the range of μl/min), and (iii) liquid junction interfaces (flow rate at sub-μl/min). Commercially available interfaces represent coaxial sheath-liquid interface G1607B (Agilent Technologies, USA), the sheathless porous tip interface CESI 8000 (AB SCIEX, USA), and nanoflow sheath-liquid interface EMASS-II (CMP Scientific, USA). The specific class is represented by microfabricated CE devices (microchips) with integrated spray emitters. A detailed discussion of most of the microchip CE-MS interfacing is summarized in a recent review [31].

4.4.1 | Sheath-liquid interface

The coaxial sheath-liquid interface based on the triple tube design is a "traditional" interface for coupling CE-MS developed by Smith et al. [211] in 1988. In this interface, the CE capillary is surrounded by two metal tubes. The inner steel tube delivers the sheath-liquid and the outer one delivers the nebulizing gas. Conductivity and volatility of sheath-liquid are important parameters in obtaining a stable MS signal. Sheath-liquid flow rates are around 1–10 μl/min (typically 4 μl/min). Sheath-liquid interfaces provide better optimization of the ionization process due to the BGE component flexibility. Moreover, certain reagents could be added to the sheath-liquid for promoting a particular chemical reaction upon mixing with the capillary effluent within the CE-MS interface. This "sheath-flow chemistry" may be used to enhance ionization rates for intact proteins by adding so-called supercharging reagents (e.g., 3-nitrobenzyl alcohol and sulfolane) [212, 213]. However, compromised sensitivity may occur due to the dilution effects of the sheath-liquid flow. Other drawbacks are a suction effect caused by a nebulizing gas and resulting in parabolic flow inside the capillary and decreased separation efficiency [34] and the requirement to use a relatively long separation capillary leading to longer separation times. The electrochemical reactions and ionization processes occurring during the ESI ionization can cause clogging of the separation capillary resulting in decreased detection sensitivity [214]. Sauer

et al. [215] introduced a new sheath-flow interface, based on the gold-coated stainless-steel tube serving as an emitter, to couple CE with Orbitrap-MS for peptide analysis. The separation capillary passes through the emitter coaxially and an isocratic syringe pump was used to deliver a sheath-liquid. However, the developed interface lacks adequate robustness and sensitivity.

4.4.2 | Liquid junction interface

A liquid junction interface (nanoflow spray liquid interface) was developed to overcome the high dilution of CE effluent by the traditional coaxial sheath-liquid interface. A pressurized version of the liquid junction interface was capable to work with a nanospray needle (10 μm id) at flow rates of tens of nL/min [216]. In some liquid-junction geometries, the electrical connection is made through a small gap between the CE capillary and ESI spray needle. This gap allows for the addition of low volumes of a spray liquid, thereby improving the CE effluent's MS compatibility [217]. The Dovichi group reported the electrokinetically pumped sheath flow interface in 2010 [218], which produced very stable spray liquid flow by EOF generated by the electrospray potential right at the inner wall of the 2–10 μm id glass emitter. The capillary, the electrospray emitter, and the spray liquid tubing were connected via a polyether ether ketone cross. The spray liquid flows over the end of the separation capillary, closing the circuit and mixing with the capillary effluent inside the tip. This interface has been later commercialized by CMP Scientific. ESI voltage was applied via a platinum electrode placed in the spray liquid reservoir. A new easy-to-use, sensitive, and robust EOF-driven nano-ESI source, nanoCEasy, based on 3D printed parts was proposed by Schlecht et al. [219]. A novel hybrid self-aligning liquid junction-type interface was microfabricated from polyimide (Figure 3) and connected a commercial CE analyzer to the fused silica capillary without any need for adjustment [220].

4.4.3 | Sheathless interface

The sheathless interfacing for CE-MS coupling has been pioneered by Smith's group in the late 1980s [221]. The main advantage of this interface is that it does not dilute eluent/sample and reduce ion suppression therefore should improve ionization efficiencies, increase sensitivity, and allow the use of highly aqueous solutions [34]. The required conductivity may be established by either metal coating of the capillary, by making the capillary tip porous, or by adding a microelectrode in the CE system. A flow inside the separation capillary must be established by pressure or electroosmosis to deliver the liquid into the electrospray and to maintain stable electrical contact. The flow rate at 10 nl/min allows low ion suppression and high sensitivity. As there is only one BGE for separation and ionization, all separation conditions have a direct influence on the spray performance [222]. The commercially available sheathless interface (CESI 8000 by Sciex) with the OptiMS sprayer is based on the porous tip design, which was introduced by Moini in 2007 [223]. In this sheathless interface, the 3–4 cm of the distal end of the separation capillary (30 μm id) is etched with hydrofluoric acid. The electrical contact for the CE is achieved through the ESI needle, which is filled with a conductive liquid and by the porous capillary protruding from the needle.

5 | APPLICATIONS

5.1 | CE-MS analysis of peptide- and protein-based therapeutics

A majority of CE-MS separations of intact proteins are focused on the analysis of biopharmaceutical products. Characterization of recombinant protein biopharmaceuticals represents a key step during the drug development or quality control process. The recent applications of CE methods with various detection systems for the analysis of biopharmaceuticals on multiple levels were reviewed elsewhere [20, 22–25, 43]. The recent applications of the CE-MS method for the characterization of biopharmaceuticals based on peptides and intact proteins in the recent 5 years (2018–March 2023) are discussed in this chapter and summarized in Table 3.

CZE-MS hyphenation has been mainly applied for the analysis of related process impurities, degradation products, and various PTMs, including glycosylation. Piešťanský et al. recently developed a CZE-ESI-QQQ system for the analysis of immunogenic synthetic peptide in a conjugate with BSA as a carrier [224] and for quantitation of therapeutic peptide, triptorelin [225]. In the first work, an effective non-enzymatic release step of the peptide from the final peptide conjugate based on acid hydrolysis (2% formic acid) was successfully tested and implemented for simple, rapid, and robust quantification of immunogens in modern immunotherapeutics [224]. In the second work, an FDA-validated CZE-ESI-QQQ system with a multisegment injection and in-capillary FESI preconcentration for determination of triptorelin improved the limit of detection 50 times (5 ng/ml) and increased the sample throughput three times in comparison to a conventional CE approach [225]. To better characterize the heterogeneity of human chorionic gonadotropin (hCG), its analysis at the intact level by different electrophoretic techniques was undertaken

TABLE 3 An overview of CE-MS methods for the analysis of intact protein-based therapeutics and therapeutic peptides.

Analyte	CE mode	Separation conditions	Detection	Sample pretreatment	Aim of the analysis	Reference
triptorelin	CZE	BFS (50 μm id) BGE: 1 M HFo, pH 1.88	QQQ (sheath-liquid)	FESI (10 kV for 20 s)	quantitative analysis	[225]
immunogenic synthetic peptide in a conjugate with BSA	CZE	BFS (50 μm id) BGE: 1 M HFo, pH 1.88	QQQ (sheath-liquid)	acidic hydrolysis (2% HFo)DPJ (12.5% AmOH)	quantitative analysis	[224]
r-hCG	CZE	BFS (50 μm id) BGE: 800 mM HFo + 800 mM HAc, pH 2.2 20 % MeOH	QQQ(sheath-liquid)	desalting (MWCO, 10 kDa)	PTMs	[226]
insulin insulin lispro	CZE	BFS (50 μm id) BGE: 50 mM AmAc, pH 9.0	QTOF(sheath-liquid)	CZE separation of deamidation products	degradation products	[227]
insulin	CZE	BFS (50 μm id) BGE: 50 mM AmAc + 20% IPA, pH 9.0	QTOF(sheath-liquid)		degradation products PTMs	[228]
infliximab	CZE	PB-DS-PB coated capillary (50 μm id) BGE: 40 mM AmAc, pH 6	QTOF(sheath-liquid)	desalting (MWCO, 10 kDa)	degradation products	[144]
BsAb	CZE	cationic PEI coated (30 μm id) BGE: 3% HAc	QTOF FT-ICR (sheathless)	desalting (MWCO, 10 kDa)	macro and microheterogeneity assessment	[229]
BsAb	CZE	cationic PEI coated (30 μm id) BGE: 10% HAc	QTOF (sheathless)	desalting (MWCO, 10 kDa)	degradation products	[230]
trastuzumab infliximab ustekinumab monovalent and bivalent nanobodies	CZE	Neutral coating based on PA (30 μm id) BGE: 50 mM AmAc (pH 3.0) and 50 mM HAc (nanobodies) 10% HAc (mAb)	QTOF (sheathless)	desalting (MWCO, 10 kDa)	glycosylation (mAb) stability and degradation products (nanobodies)	[231]
IgG mAb	CZE	cationic M7C4I coating BGE: 0.2% HFo and 10% IPAnative analysis:neutral PA coated BGE: 40 mM AmAc, pH 7.5	Q-Orbitrap (sheathless)	Bio-Spin™ P-6 gel columns (SEC)	glycosylation PTMs	[232]

(Continues)

TABLE 3 (Continued)

Analyte	CE mode	Separation conditions	Detection	Sample pretreatment	Aim of the analysis	Reference
trastuzumab rituximab palivizumab	CZE	cationic PEI coated (30 μm id) BGE: 3% HAc	Q-TOF (sheathless)	desalting (MWCO, 10 kDa)	glycosylation charge variants	[233]
seven mAbs	CZE	cationic PEI coated (30 μm id) BGE: 3% HAc	IM-QTOF (sheathless)	desalting (MWCO, 10 kDa)	charge variants PTMs glycosylation	[234]
two mAbs	CZE	neutral PVA coated (50 μm id)BGE: 2 M HAc	Orbitrap (nanoCEasy) QTOF (sheath-liquid)		glycosylation	[235]
infliximab	CIEF	neutral PVA and HPC coated (50 μm id) anolyte – 1% HAc catholyte– 1% AmOH 20% glycerol	Q-TOF Q-Orbitrap-IT (flow-through microvial interface)		charge variants	[236]
four mAbs	CIEF	neutral PVA coated (50 μm id) anolyte – 1% HAc catholyte– 1% AmOH 20% glycerol	Q-TOF (flow-through microvial interface)	desalting (MWCO, 30 kDa)	charge variants structural heterogeneity	[237]
trastuzumab bevacizumab infliximab cetuximab	CIEF	neutral PS1 coated (50 μm id)catholyte: 0.2 M AmOHanolyte: 1% HFo15% glycerol	TOF (EMASS-II)	desalting (MWCO, 30 kDa)	charge variants	[238]
bevacizumab	CZE CIEF	CIEF: Neutral PS1 coated (50 μm id)Pharmalyte® 3–10 and Pharmalyte® 8–10.5 Buffer S35CZE: neutral PS2 coated (50 μm id)	TOF (EMASS-II)	desalting (MWCO, 10 kDa)	size heterogeneity	[239]
pembrolizumab	iCIEF	acrylamide derivative coated (200 μm id)100 μm id inlet capillary HR AESlyte 6–8 (carrier ampholytes)	Q-Orbitrap (sheath-liquid micro-tee connection)		charge variants	[169]

(Continues)

TABLE 3 (Continued)

Analyte	CE mode	Separation conditions	Detection	Sample pretreatment	Aim of the analysis	Reference
mAbs (bevacizumab pembrolizumab) ADC (T-DM1)	iCIEF	acrylamide derivative coated (200 μm id)100 μm id transfer capillaryvarious HR AESlytes (carrier ampholytes)10% formamide	Q-Orbitrap (sheath-liquid micro-tee connection)		charge variants	[170]
nine mAbs	iCIEF	acrylamide derivative coated (200 μm id)100 μm id transfer capillary	Q-Orbitrap (sheath-liquid, micro-tee connection)		charge variants	[168]
fusion mAb trastuzumab	iCIEF	acrylamide derivative or fluorocarbon (FC) coated (200 μm id) 50 μm id transfer capillary AESlyte HR 6–8 (fusion mAb) Pharmalyte 3–10 and Pharmalyte 8–10.5 (trastuzumab)	QTOF and Orbitrap (nanoCEasy)		charge variants	[167]
mAb Fc receptors	ACE	OptiMS neutrally coated capillaries BGE: 50 mM AmAc pH 6.8	Orbitrap (sheathless)	desalting (MWCO, 10 kDa)	mAb-receptor interactions	[203]
mAb Fc receptors	ACE	OptiMS neutrally coated capillaries BGE: 50 mM AmAc pH 6.0	Orbitrap (sheathless)	desalting (MWCO, 10 kDa)	mAb-receptor interactions	[202]
recombinant humanized IgG1 mAb	mCZE	ZipChip BGE: 0.2% HAc + 10% IPA (pH 3.17)	Q-Orbitrap (sheathless)		charge variants PTMs	[240]
fusion proteins	mCZE	ZipChip BGE: 0.2% HAc + 10% IPA (pH 3.17)	Q-Orbitrap (sheathless)		glycosylation	[241]
cetuximab	mCZE	ZipChip ^{HRN} Native Antibodies Kit (proprietary formulation based on AmAc at pH 5.5)	Q-Orbitrap (sheathless)	desalting (MWCO, 10 kDa)	glycosylation charge variants	[242]
rituximab trastuzumab bevacizumab	mCZE	ZipChip ^{HRN} Native antibody kit (proprietary formulation based on AmAc at pH 5.5) + 4% DMSO	Q-Orbitrap (sheathless)	desalting (MWCO, 10 kDa)	charge variants	[243]

(Continues)

TABLE 3 (Continued)

Analyte	CE mode	Separation conditions	Detection	Sample pretreatment	Aim of the analysis	Reference
canonical IgG1	mCZE	ZipChip HRN Native Antibodies Kit (proprietary formulation based on AmAc at pH 5.5)	Q-Orbitrap (sheathless)		charge variants degradation products	[244]
IgG4 mAbIgG1 mAb	mCZE	ZipChip HRN Native Antibodies Kit (proprietary formulation based on AmAc at pH 5.5)	Q-Orbitrap (sheathless)		charge variants	[245]
various mAbs and BsAbs	mCZE	ZipChip HRN Native Antibodies Kit (proprietary formulation based on AmAc at pH 5.5) + 3.8% DMSO	Orbitrap (sheathless)		charge variants	[268]
trastuzumab biosimilar	microchip iCIEF	hydrophilic-coated borosilicate microchipscatholyte: 1% diethylamineanolyte: 1% HFo	Q-TOF (nanoelectrospray capillary cap)	ZebaTM7K MWCO spin desalting column	charge variants	[172]
Bs Ab	microchipiCIEF	microchip type not specified.catholyte: 1% diethylamineanolyte: 1% HFo	Q-TOF (nanoelectrospray capillary cap)	ZebaTM7K MWCO spin desalting column	charge variants	[173]
IgG1 and IgG2 mAb	microchipiCI	coated microfluidic chipanolyte: 1% HFocatholyte: 1% diethylamine	Q-Orbitrap		charge variants	[174]
trastuzumab	iCIEF-CZE	iCIEF: catholyte—100 mM NaOH anolyte—80 mM phosphoric acid 0.1% methylcellulose CZE: transfer neutral PVA coated (50 μm id)	QQQ (sheath-liquid)		charge variants	[247]
NIST mAb	CIEF-CZE	neutral LCP coated (50 μm id)BGE: 25 mM AmAc (pH6.8)catholyte:50 mM AmAc (pH 9.0)	Q-TOF (EMASS-II)	Bio-Spin P-6 gel columns (SEC)	glycosylation protein complexes	[269]

Abbreviations: ACE, affinity capillary electrophoresis; AmAc, ammonium acetate; AmOH, ammonium hydroxide; BFS, bare fused silica; BGE, background electrolyte; BSA, bovine serum albumin; bsAb, bispecific antibody; CIEF, capillary isoelectric focusing; CZE, capillary zone electrophoresis; DMSO, dimethyl sulfoxide; FESI, field-enhanced sample injection; FT-ICR, Fourier-transform ion cyclotron resonance; HAc, acetic acid; HFo, formic acid; HPC, hydroxypropyl methylcellulose; iCIEF, imaging capillary isoelectric focusing; IM, ion mobility; IPA, isopropyl alcohol; LCP, linear carbohydrate polymer; LPA, linear polyacrylamide; mAb, monoclonal antibody; mCZE, microchip capillary zone electrophoresis ; MeOH, methanol; MWCO, molecular weight cut-off; PA, polyacrylamide; PB-DS-PB, polybrene dextran sulfate; PEI, polyethyleneimine; PTM, post-translational modification; PVA, polyvinyl alcohol; QQQ, triple quadrupole; Q-TOF, quadrupole time-of-flight; r-hCG, recombinant human chorionic gonadotropin; SEC, size-exclusion chromatography.

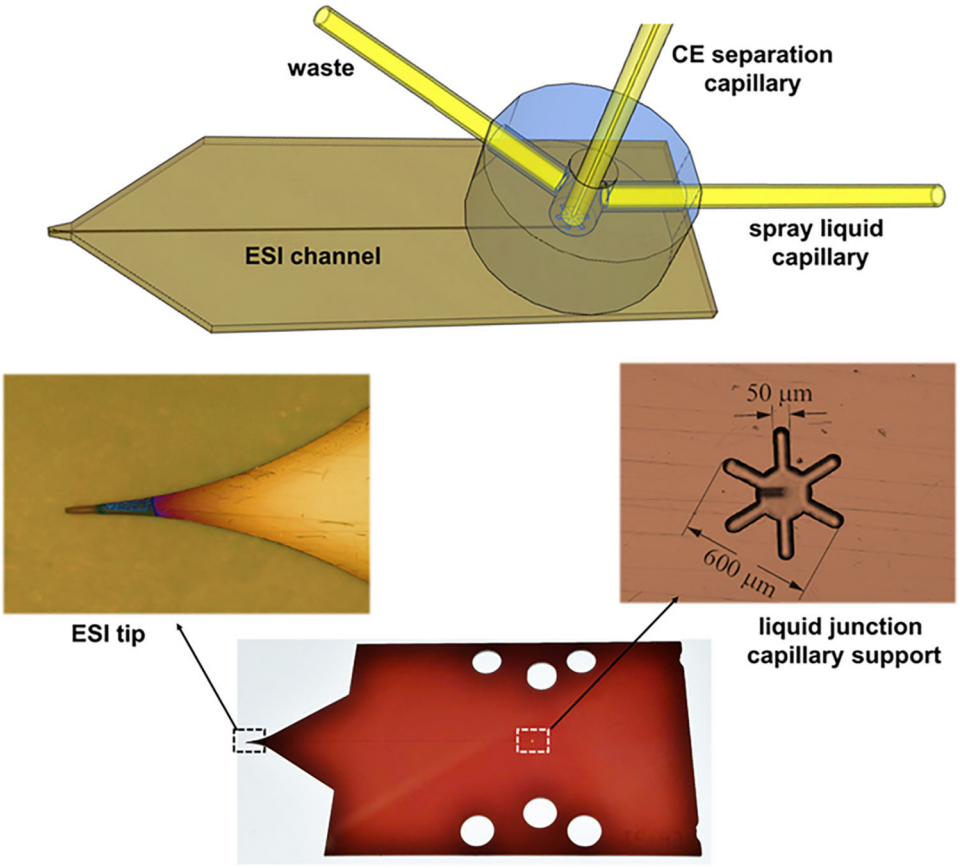


FIGURE 3 Scheme of the interface (top) and photographs of the liquid junction chip details. Reprinted with permission [220].

for the first time by Camperi et al. [226]. This study confirmed the very high heterogeneity of hCG isoforms and CZE-ESI-QQQ analysis demonstrated its potential for a fingerprinting approach as it allowed to differentiate the two hCG-based drugs.

High-resolution MS is the most utilized detection system for the characterization of biopharmaceutical products. CZE-ESI-QTOF was employed for analyzing deamidation and other degradation products of insulin and its analog, lispro [227, 228]. The developed method was efficiently applied for monitoring the degradation rate of insulin and the formation of different deamidation isoforms. However, identification of the exact position of deamidation sites in the insulin molecule remains a challenge. CZE-ESI-QTOF with a sheath-liquid interface was used also for studying the unfolding and aggregation of fresh and storage stressed (6 months at 4°C) mAb (infliximab) in a non-denaturing environment, using 40 mM ammonium acetate, pH 6.0 as BGE [144]. In this work, infliximab dimers formation was investigated and attributed to the interaction between unfolded infliximab molecules via F(ab) regions.

Gstöttner et al. [229] applied CZE-ESI hyphenated through the sheathless interface to QTOF or Fourier-transform ion cyclotron resonance for intact analysis of two homologous bispecific antibodies (BsAbs) for assessing their macro- and microheterogeneity. For intact BsAbs, the separation permitted the characterization of free light chains, homo- and heterodimers as well as incomplete assemblies. The potential of the developed sheathless CZE-ESI-QTOF to monitor the exchange efficiency and stability of in-house produced IgG4 monovalent BsAbs was also inspected [230]. BsAbs were separated from the parent monospecific antibody versions permitting reliable determination of the exchange efficiency even for BsAb2 where masses of the different species overlapped. The method also allowed monitoring of degradation products after prolonged storage and in vivo experiments.

Haselberg et al. [231] reported a sheathless CZE-ESI-QTOF for heterogeneity assessment of three nanobody preparations and three mAb therapeutics. Baseline resolution with narrow peaks and high plate numbers of 100,000 with overall migration time RSDs less than 2.2% was obtained. For all mAbs, separation of Fc/2 charge variants, including sialylated glycoforms and other PTMs, was achieved. Different glycosylated forms of IgG1 mAb were investigated by Belov et al. [232] under denaturing conditions by CZE-ESI-Orbitrap coupled with the sheathless interface. In addition, intact IgG1 mAb was also

studied under native conditions, and dimers of mAb were detected.

The applicability of CZE-ESI-QTOF for intact trastuzumab, rituximab, and palivizumab glycosylated isoforms analysis under denaturing conditions [233] and CZE-ESI-IM-QTOF in glycoforms and other PTMs analysis of seven commercial therapeutic antibodies [234] was demonstrated by Giorgetti et al. Cationic polyethyleneimine-coated capillaries were applied in both works. Concerning the charge variant analysis, potential aspartic acid isomerization modification and asparagine deamidation have been observed.

Naumann et al. [235] recently presented CZE methods hyphenated through the sheath-liquid interface to QTOF and liquid junction nanoCEasy interface to Orbitrap for the high throughput non-targeted screening of the glycosylation pattern of intact mAb at the low μg/ml level directly from cell supernatants without any sample preparation. Both MS instruments were sufficient to determine the glycosylation pattern of the five major glycoforms, however, the Orbitrap achieved higher sensitivity.

Recently published works suggested automated CIEF-MS and iCIEF-MS as promising tools for quality control of therapeutic biopharmaceuticals regarding charge variants and PTMs by providing high-resolution separation and accurate mass determination. In 2018, Wang et al. [236] introduced two instrumental platforms based on CIEF-QTOF and CIEF-Orbitrap for the separation and identification of charge variants of the commercial mAb formulation (infliximab). The on-line hyphenation was facilitated by a flow-through microbial interface made of stainless steel with high chemical resistance and mechanical robustness. With only 30 ng of mAb sample consumed in a single injection, four charge variants (the main intact mAbs along with two basic and one acidic) with 0.05–0.2 pI differences and 13 glycoforms were detected. The same sandwich CIEF-QTOF approach, where anolyte, sample, and catholyte segments were sequentially injected into a neutrally (polyvinyl alcohol) coated capillary, was implemented for the charge and structural heterogeneity analysis of four intact mAbs [237]. A single run required less than 400 ng mAb and the main intact mAbs along with one acidic and two basic variants was observed in less than 100 min.

Dai et al. [238] developed an automated CIEF-ESI-TOF method with neutrally coated capillary and electrokinetically pumped sheath-liquid nanospray (EMASS-II) for the separation and characterization of mAb charge variants. Sufficient robustness and accuracy assessments of sheath-liquid have been shown by studies characterizing degradation variants of mAbs and their fragments via screening and monitoring of specific modifications. Using this method, the charge variants of various mAbs, including trastuzumab, bevacizumab, infliximab, and cetuximab, were well-resolved. Li et al. recently presented [239] two separation approaches, CZE-TOF and CIEF-TOF, for the identification of bevacizumab's clipping variant providing a powerful addition to the traditional CE-SDS analysis workflow. The intact masses of all species in the bevacizumab were identified in the CZE-MS analysis. Furthermore, CIEF-MS of the intact bevacizumab confirmed the existence of the clipping variant. These two complementary approaches represent orthogonal verification for size heterogeneity characterization.

Very recently, Zhang et al. [169] published a work, where a cutting-edge preparative iCIEF platform was employed for analytical profiling, MS coupling, and fraction collection for charge variant analysis of biopharmaceuticals. This multiple-operation mode system can be rapidly and flexibly switched just by changing customized capillary separation cartridges without further configurations. The whole workflow provides a comprehensive and revolutionary technology for protein drug quality control monitoring, MS coupling for fingerprinting intact protein, and LC-MS peptide mapping in depth within 45–60 min. The innovative microliter interface improved the sensitivity of identifying protein charge variants. MS-compatible amphoteric electrolytes and both polymer-free and urea-free cartridges in iCIEF analysis have been innovatively developed. The established methodology was employed for the charge heterogeneity characterization of pembrolizumab [169] and for commercial monoclonal antibodies (mAbs) and antibody-drug-conjugates [170]. Wu et al [168] recently compared this innovative iCIEF-MS approach for charge variant analysis of diverse mAbs to strong cation exchange-MS. It was found that iCIEF-MS outperformed strong cation exchange-MS in this study by demonstrating its outstanding sensitivity, low carryover, accurate protein identification, and higher separation resolution. Schlecht et al. [167] presented straightforward direct coupling of iCIEF–MS via the novel nanoCEasy interface to two different MS instruments providing separation and subsequent on-line deep characterization of charge variants of two intact mAb (Figure 4). The presented setup provided a large potential for mAb charge heterogeneity characterization in biopharmaceutical applications.

Gstöttner et al. [202, 203] developed an approach based on mobility shift-ACE–MS to determine the binding of coexisting mAb proteoforms to Fc receptors. The approach required only low microgram amounts of antibody and receptor and coupling through a sheathless interface allowed functional characterization of mAbs with a high sensitivity and dynamic range. Hyphenation with native MS provided unique capabilities for simultaneous heterogeneity assessment for mAbs, Fc receptors, and formed complexes.

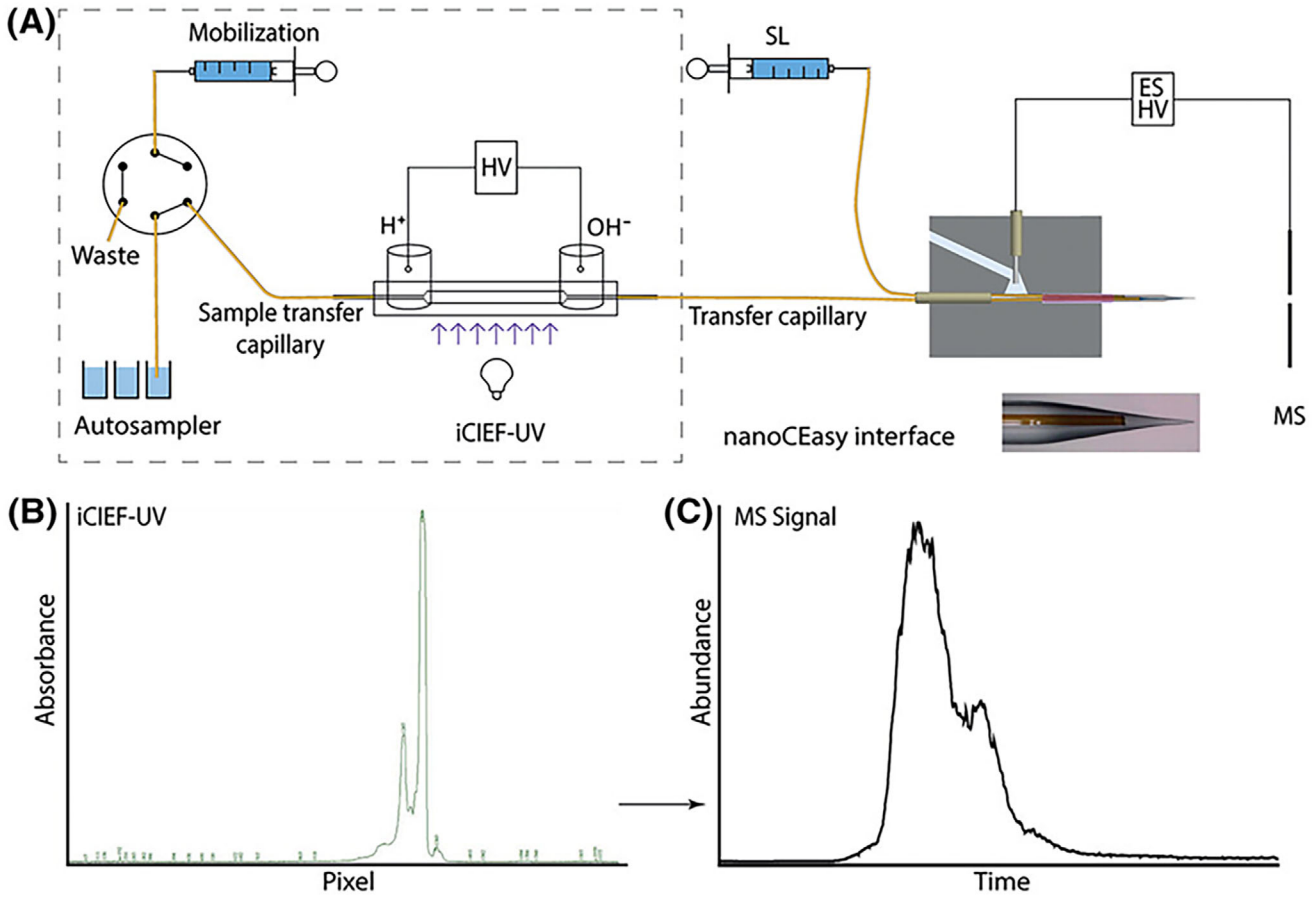


FIGURE 4 (A) Setup of CEInfinite coupled to mass spectrometry. Schematic view of the CEInfinite with autosampler, syringe pump, and imaged capillary isoelectric focusing (iCIEF) cartridge connected with the nanoCEasy electrospray ionization–mass spectrometry (ESI–MS) interface; (B) iCIEF–UV profile of trastuzumab (2 mg/ml). Focusing voltage: 1500 V (1 min) and 3000 V (10 min). UV detection at 280 nm; and (C) base peak electropherogram (BPE) m/z 1500−4000 of trastuzumab (2 mg/ml, mobilization speed 30 nl/min) with Orbitrap Fusion Lumos. Reprinted [167].

Moreover, techniques such as microfluidic systems and 2D CE-MS are now emerging to offer new possibilities beyond actual limits. Chen et al. [240] presented a fast and robust ZipChip mCZE-MS workflow for intact mAb characterization under denaturing conditions. This method is a useful tool for the rapid screening of therapeutic mAbs, including the characterization of glycoforms and other PTMs. mCZE-MS analysis of intact mAb showed all major glycoforms, as well as all three possible lysine variants, and enabled their relative quantification. A similar ZipChip mCZE-MS approach was implemented by Deyanova et al. [241] for monitoring the glycosylation profile of a fusion protein with complex sialylated glycoprofile within 6 min. The method can be used to estimate the relative abundance of each glycoform, generating signature fingerprints for the molecule.

Comparative glycoform heterogeneity and charge variant characterization of chimeric mAb, cetuximab, by pH gradient cation exchange chromatography (CEX)-MS and mCZE-MS using the ZipChip platform was performed by Fussl et al. [242]. Both techniques have demonstrated the capability for the separation of eight major peaks of cetuximab charge variants. Besides, the resolution of distinct glycoform complexity within unique charge variant peaks was obtained using native mCZE-MS based on hydrodynamic radius-dependent separation (Figure 5). Totally, 109 isoforms were found via CEX-MS while for CE-MS it was 218 isoforms. The CEX-MS method seemed to be less prone to interferences rendered by drug formulations. However, CE-MS did not require individual method optimization for each specific mAb because the parameters used are suitable for charge variants analysis of conventional mAbs. Such an analytical approach was also successfully applied to investigate the charge heterogeneity of rituximab, trastuzumab, and bevacizumab [243]. The method allowed achieving the confident identification of up to 52 proteoforms in trastuzumab, as well as mAb fragments in rituximab. The detected proteoforms were present at levels as low as 0.01% using just 1 ng of sample and showed improved separation resolution and good

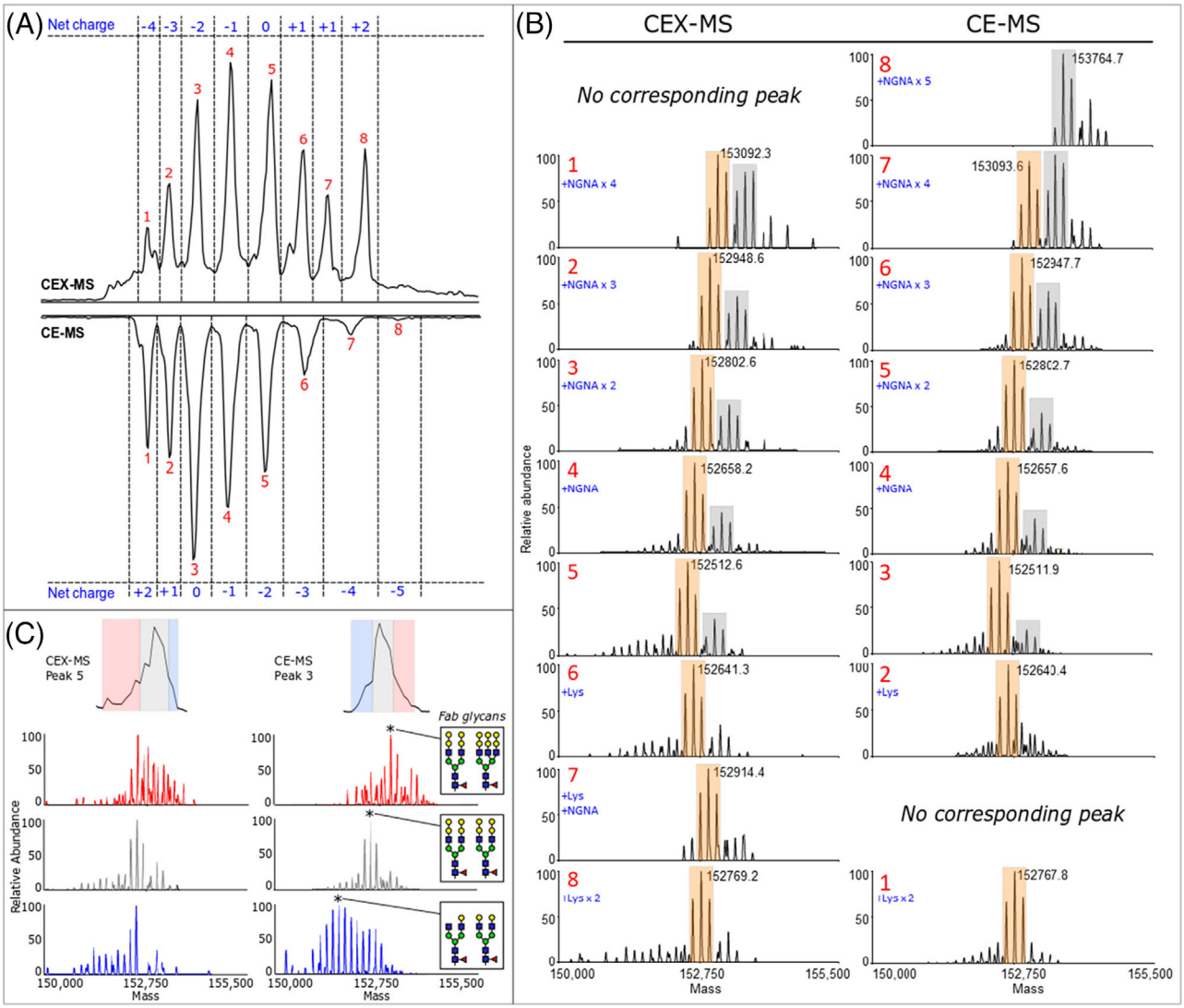


FIGURE 5 (A) Cetuximab chromatogram and electropherogram. Peaks are labeled in red in order of increasing retention/migration time. The net charge is represented in blue. (B) Deconvoluted averaged mass spectra corresponding to each peak. Spectra from the two separation approaches containing the same charge variant species are aligned horizontally. Peak assignment is based on the corresponding red labels in panel A. Modifications that are causing differences in the net charge are indicated in blue, whereas only the most probable combination of modifications is given. The most abundant species corresponding to biantennary Fab glycan pairs are shown in orange, whereas the most abundant forms carrying triantennary Fab glycans are labeled in gray. (C) Deconvoluted spectra of peak front, center, and tail of peaks 5 and 3 of the CEX and CE separation, respectively. Associated spectra are aligned horizontally and are shown in the same color. Reprinted with permission [242].

comparability with previously reported profiles obtained using IEC chromatography.

Analysis of intact mAb stressed at high temperature under native conditions by mCZE-MS by Sun et al. [244] resulted in the characterization of basic and acidic variants of normal and stressed mAb. The authors concluded that the basic variants in the unstressed sample were produced by C-terminal amidation, while the acidic variants were produced by deamidation. In stressed samples, a change in the acidic and main peaks was caused by deamidation, and changes in the basic peaks were caused by both deamidation and oxidation. In the work of Cao et al. [245] charge variants' identities were assigned based on characteristic mass shifts, knowledge learned from peptide mapping, and changes in electrophoretic mobility. Major mAb glycoforms of each charge variant were resolved and identified in the deconvoluted mass spectra. Wu et al. [246] performed mCZE-MS analysis under native conditions to evaluate the charge variants and impurities in therapeutic antibodies including IgG mAbs, BsAbs, and alternative formats such as therapeutic antibodies with the addition or removal of the antigen-binding domain. The high

sensitivity analysis by the developed method enabled the detection of fragments with very low abundance in the antibody sample and monospecific mAb impurities in the BsAb sample.

A recent microchip-based iCIEF-MS system with MS-compatible conditions was developed by Mack et al. [172] to characterize charge variants of intact mAb. This innovative enhanced microfluidic chip-based integrated iCIEF-MS technology is capable to assess multi-attribute characterization (characterization and identification of protein charge and mass isoforms) of intact antibodies in a single run within 15 min. In total, 33 separate molecular features were characterized by the iCIEF-MS system representing a dramatic increase in the ability to monitor multiple intact mAb critical quality attributes in a single comprehensive assay. Ostrowski et al. [173] very recently applied this technology as a powerful approach for assessing product quality and provided charge variant analysis with the direct identification of all charge isoforms of the BsAb. This system provided the orthogonal separation required to resolve and identify acidic PTMs including difficult-to-detect deamidation and glycation events at the intact protein level. Six charge variant peaks were resolved and directly identified by on-line iCIEF-MS. In addition to acidic charge variants, basic variants were identified as C-terminal lysine, *N*-terminal cyclization, proline amidation, and the combination of modifications, including lysine and one or two hexose additions.

He et al. [174] demonstrated the comparability of the pI value measurement and relative charge species distributions between the new microfluidic iCIEF-MS system and the control data from a frequently utilized methodology in the biopharmaceutical industry for several blinded development-phase biopharmaceutical mAb across a wide pI range of 7.3–9.0. The results showed acidic and basic shifts caused by sialic acid additions and the presence of unprocessed lysine residues.

Besides a direct coupling of CIEF and ESI-MS, CIEF has also been coupled to CZE-MS using a mechanical valve or a nanoliter valve for high-resolution characterization of intact proteins and mAb charge variants, to clean up the non-volatile additives and improve mass spectra quality. iCIEF in combination with the nanoliter valve as an interface to the second CZE-MS dimension efficiently provided MS characterization of the main charge variants from trastuzumab [247]. CIEF also complemented native CZE-MS for the analysis of standard mAbs [248]. An online sample stacking based on CIEF in a narrow pH range was developed to expand the loading capacity as well as improve the CZE separation in the native conditions for the first time.

5.2 | CE-MS analysis of peptide- and protein-based biomarkers

CE-MS is a powerful tool in the context of proteomic clinical application, as over 380 manuscripts focused on peptidomics and/or proteomics were published within the last 20 years [45]. However, only a few scientific papers deal with the targeted proteomic analysis of intact proteins or peptides by CE in biological samples, and even fewer publications used a validated method. In this chapter, we summarize and discuss the recent (2018–March 2023) applications of CE-MS for the analysis of peptides and intact proteins as potential biomarkers in biological fluids and these applications are also listed in Table 4.

Pont et al. [132] reported the use of polymeric monoliths with gold nanoparticles (AuNP@monolith) as a solid sorbent in on-line SPE-CE-ESI-(IT)MS for the analysis of human transthyretin. Two microcartridge designs of on-line SPE-CE are shown in Figure 6. This type of sorbent due to the highly active surface gives appropriate recoveries and enrichment factors (50-fold). However, incorporating AuNPs into polymer monoliths makes method selectivity problematic when using complex biological samples.

Ricci et al. by the CE-ESI-TOF method investigated the urinary proteome of pediatric renal cysts and diabetes syndrome (RCAD) patients and different controls to identify peptide biomarkers and gain further knowledge about the pathophysiology of this disorder [249]. The protocol included sample preparation consisting of ultrafiltration with 20 kDa molecular weight cut-off filters, desalting based on the SEC principle using pharmacodynamic-10 desalting columns, and lyophilization. The authors found 146 peptides associated with RCAD, which were further tested as a classifier on an independent cohort. The study demonstrates the difference in urinary proteome between pediatric RCAD patients and many other renal diseases, suggesting the pathophysiological differences behind these disorders. Another application of this method was made by Mavrogeorgis et al. [250]. Their work aimed to evaluate the reproducibility, variability, and efficiency of urinary peptide detection by CE-MS. The performance of the method was evaluated by 72 measurements of a standard urine sample. The study showed a high correlation between runs, low variation of the highest signal intensities (coefficient of variation <10%) and very low variation of biomarker panels applied (coefficient of variation close to 1%), which proves the value of CE-MS in clinical applications.

Moran et al. [251] profiled prostate-specific antigen (PSA) proteoforms from urine by CE coupled with HRMS using a sheathless porous tip interface, with a pretreatment based on immunoaffinity SPE using anti-PSA

TABLE 4 An overview of published methods for the analysis of intact proteins and peptides in biological matrices by CE-MS.

Analyte	Matrix	Sample preparation	Separation parameters	Analytical method	LOD [μg/ml]	Reference
transthyretin	serum	on-line SPE (polymeric monoliths with gold nanoparticles)	72 cm × 75 μm fused-silica; BGE: 10 mM AmAc (pH 5.0)	CZE-ESI-IT (sheath-liquid)	0.005 (standard)	[132]
urinary peptidome and proteome	urine	ultrafiltration (20 kDa MWCO filter), SEC (PD-10 column)	90 cm × 50 μm fused-silica;BGE: 0.94% HFo + 20% ACN	CZE-ESI-TOF(sheath-liquid)	–	[249, 250]
PSA proteoforms	urine	IA SPE (anti-PSA nanobodies on Sepharose beads)	91 cm × 30 μm coated with PEI; BGE: 20% HAc	CZE-nanoESI-QTOF sheathless (porous tip)	–	[251]
Aβ1-38, Aβ1-40, Aβ1-42,	CSF	three-step ultrafiltration (10 kDa, and 2 × 3 kDa MWCO filtration)	100 cm × 100 μm fused-silica; LE: 80 mM acetate; TE: 0.175% AmOH	ITP-ESI-TOF(sheath-liquid)	0.13 – 4.5 $\times 10^{-3}$	[187]
α-synuclein	RBC lysate	on-line aptamer affinity SPE (magnetic beads)	72 cm × 75 μm fused-silica; BGE: 100 mM HAc	CZE-ESI-TOF (sheath-liquid)	0.2	[131]
hemoglobin isoforms	RBC lysate	–	100 cm × 50 μm PS1 neutral coating; BGE: 2% HFo + 20% ACN	CZE-ESI-Orbitrap (sheath-liquid)	–	[253]
angiotensin peptides	mice SFO, PVN	solvent extraction	85 cm × 40 μm fused-silica; BGE: 1 M HFo + 25% ACN	CZE-nanoESI-LTQ-Orbitrap	0.5 – 5 $\times 10^{-3}$	[252]
hemoglobin isomers	DBS	solvent extraction	70 cm × 50 μm SMIL coated; BGE: 2 M HAc	CZE-ESI-Q-TOF-Orbitrap (sheath-liquid)	–	[125]

Abbreviations: ACN, acetonitrile; AmAc, ammonium acetate; AmOH, ammonium hydroxide; BGE, background electrolyte; CE, capillary electrophoresis; CSF, cerebrospinal fluid; DBS, dried blood spot; ESI, electrospray ionization; HAc, acetic acid; HFo, formic acid; IA, immunoaffinity; IT, ion trap; ITP, isotachophoresis; LTQ, linear trap quadrupole; MWCO, molecular weight cut-off; PEI, polyethyleneimine; PSA, prostate-specific antigen; PVN, paraventricular nucleus; RBC, red blood cells; SEC, size-exclusion chromatography; SFO, subfornical organ; SMIL, successive multiple ionic-polymer layer; SPE, solid phase extraction; TOF, time-of-flight.

nanobodies bound to Sepharose beads. Six proteolytic cleavage variants of PSA were identified with an overall 77 PSA N-glycans determined (Figure 7). Furthermore, the authors performed a bottom-up analysis, and both approaches identified a similar relative abundance of PSA glycoforms. According to the authors, these results represent an important basis for future characterization and biomarker studies.

Crosnier de Lassichére et al. [187] coupled less conventional capillary electrophoretic method—ITP with MS detection for the determination of amyloid β (Aβ) peptides in cerebrospinal fluid samples from Alzheimer's disease patients. By coupling this powerful preconcentration technique, the group reached sensitivity at the sub-nM levels. By determining the Aβ 1–42/Aβ 1–40 ratios, they successfully discriminate patients with Alzheimer's disease from a healthy control.

Lombard-Banek et al. developed a microanalytical laboratory-build CE platform coupled to a high-resolution MS equipped with nano-ESI for analysis of angiotensin peptides in mice brain tissues sampled from subfornical organ and paraventricular nucleus of the hypothalamus [252]. This method revealed different angiotensin peptide levels in the two locations of the brain and with parallel reaction monitoring MS configuration achieved excellent sensitivity at the level of approximately 0.5–5 ng/ml.

Peró-Gascón et al. [131] described an on-line aptamer affinity-SPE CE-MS method for the analysis of intact α-synuclein in the blood. This complex method along with separation and detection also includes purification and

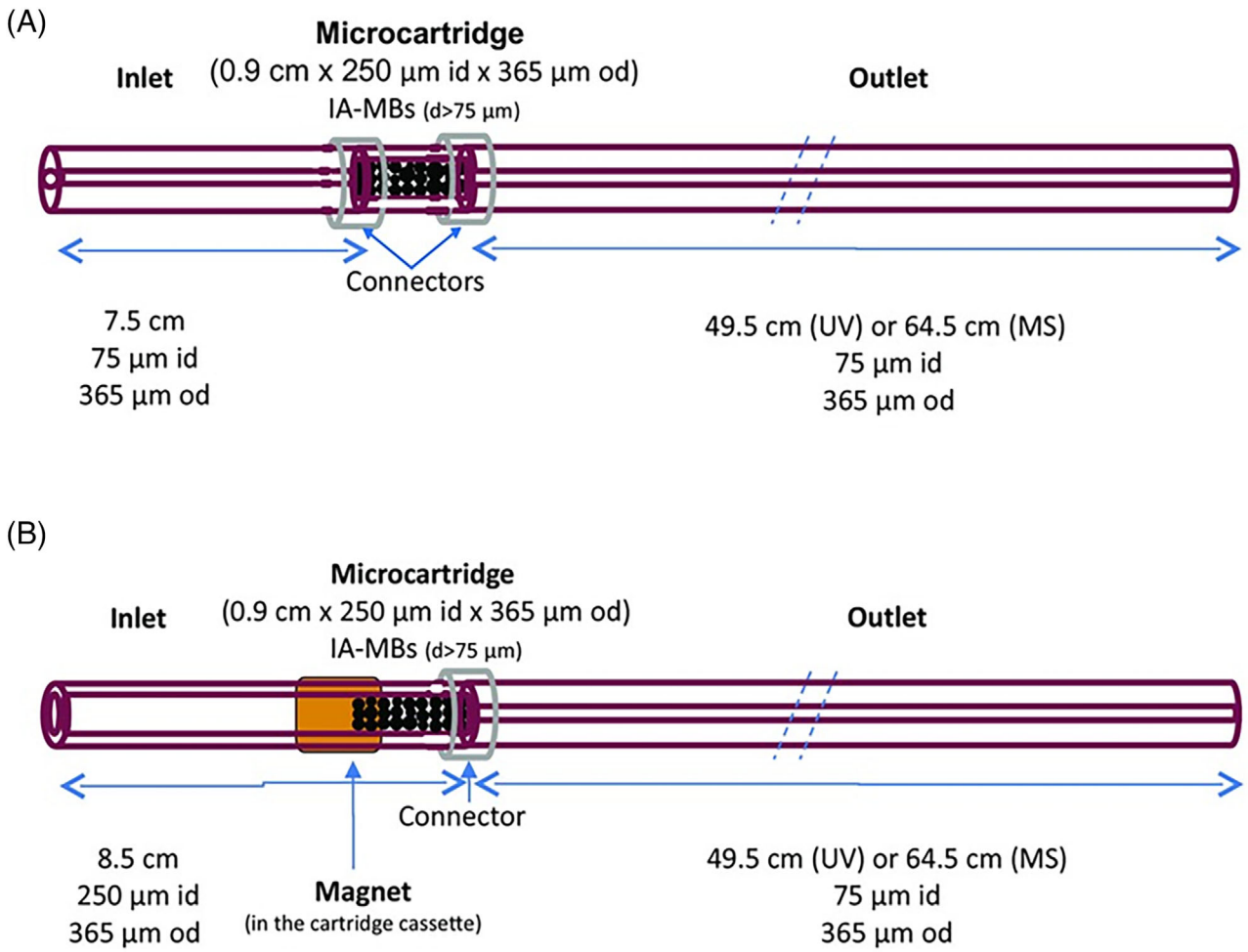


FIGURE 6 Representations of the microcartridge designs (A) Protein A Ultrarapid Agarose or AffiAmino Ultrarapid Agarose™ magnetic beads are trapped in a microcartridge body of 250 µm id due to their particle size and (B) Protein A Ultrarapid Agarose™ or AffiAmino Ultrarapid Agarose magnetic beads are retained in one of the ends of a piece of 250 µm id capillary and a magnet prevents the shift and loss of the magnetic beads. (The first design could not be applied with SiMAG-Protein A and Dynabeads Protein A magnetic beads because both are very small. Similarly, in the second case the magnet should cover the whole microcartridge body). Reprinted with permission [132].

preconcentration of the red blood cell lysate samples with a LOD 100-fold lower than conventional CE-MS. The novel method was applied to the detection of α-synuclein proteoforms in Parkinson's disease patients. However, the authors report that *N*-acetylated α-synuclein is the main and the only proteoform detected in Parkinson's disease patients and healthy controls.

Luo et al. [253] recently analyzed red blood cell lysate for the identification of intact hemoglobin variants using HRMS coupled with CE without any sample purification method. With neutral capillary coating, they were able to baseline separate a total of 15 hemoglobin subunits from 18 patient samples. According to the authors, the ability to characterize the primary structures of hemoglobin makes this method a suitable complement or even partial replacement of the conventional methods used to identify hemoglobin variants.

Stolz et al. [125] described a flexible and versatile CE-MS method for intact hemoglobin proteoforms screening from a dried blood spot sample with simple sample preparation based on solvent extraction. Thanks to the 5-layer multiple ionic polymer capillary coating were achieved exceptional separation power enabled the separation of positional isomers of glycated α- and β-hemoglobin chains. The method showed a good correlation with the results obtained in a clinical routine method and, according to the authors, represents a valuable tool for the deeper characterization of clinical and veterinary samples.

6 | CONCLUDING REMARKS

Capillary electrophoresis is a powerful, cost-effective, and eco-friendly technique. It offers attractive features for

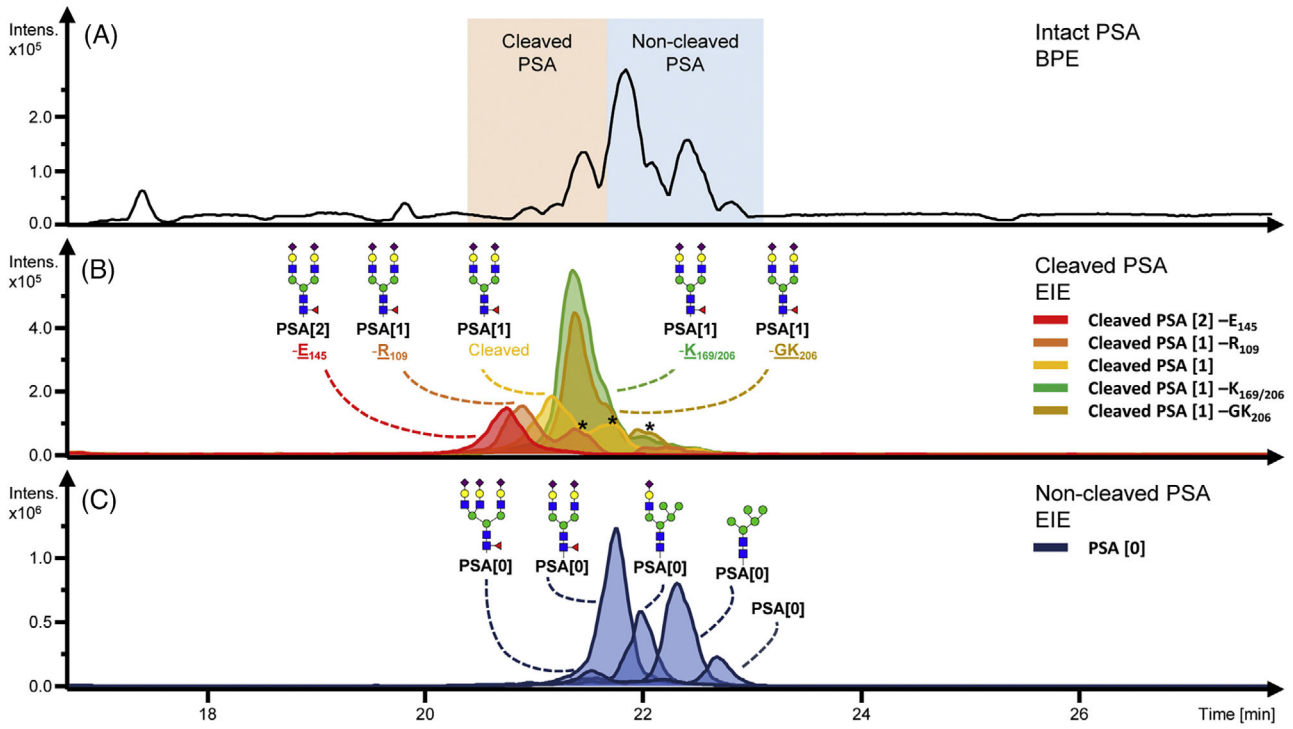


FIGURE 7 Intact protein analysis using CE-ESI-MS of PSA captured from a patient urinary pool. (A) Base peak electropherogram (BPE) of intact urinary PSA. A variation in PSA proteoforms can be observed with and without internal cleavages. (B) The EIEs illustrate the most abundant glycoprotein present per proteoform. Proteoforms with different numbers of internal cleavages can be observed with the number of internal cleavages indicated in the square brackets: PSA[2] -E145, double cleaved PSA at E145 and one other unidentified position; PSA[1] -R109, cleaved PSA at R109; PSA[1] -Cleaved, cleaved PSA at an unidentified position; PSA[1] -K169/206, cleaved PSA at either K169 or K206; PSA[1] -GK206, cleaved PSA at GK206 (C) For non-cleaved PSA the most abundant tri-, di- and mono-sialylated form is illustrated as well as the most abundant high mannose type and the non-glycosylated form of the protein. Asterisk (*) indicates overlapping m/z values in different electrophoretic peaks. Blue square: *N*-acetylglucosamine, green circle: mannose, yellow circle: galactose, red triangle: fucose, pink diamond: *N*-acetylneuraminic acid. Reprinted [251].

miniaturization, sample preparation automation, and integration with the separation step. Moreover, CE methods are orthogonal to LC techniques, and in many cases, they can serve as a complementary technique to enrich the information outcome. The coupling of CE to highly sensitive and selective MS detection has matured into a powerful analytical separation method that offers fast and highly efficient protein separations on multiple levels of complexity, including intact protein analysis. Research performed during the last years demonstrates that CE-MS is rapidly growing in the field of the biopharmaceutical industry. However, limitations of CE-MS still appear when this method is utilized for the analysis of intact proteins in complex biological fluids.

Nevertheless, CE-MS continues to evolve, the fundamental technical challenges are being addressed by analytical scientists and more commercial solutions are being offered. Currently, various commercially coated capillaries to prevent protein adsorption on the inner wall of the fused silica capillaries are available. However, the development of novel more efficient capillary coatings and additives to reduce non-specific interactions and increase protein sample solubility and stability remains a challenge. Therefore, more research devoted to the development of novel coating materials can be expected. The novel, more convenient, and effective sample clean-up and enrichment approaches are still in good demand to increase the sample loading capacity of CE-MS for proteomic analysis. Technical advances, hardware, and software development will further allow us to get more valuable information from the intact level analysis and advance also multimarker-based clinical applications.

In summary, CE-MS will continue to grow into a more advanced analytical tool to help solve problems in diverse areas. It can be expected, that CE-MS will be also in the future broadly involved in biopharmaceutical and biomedical analysis to help understand the roles played by proteoforms and protein complexes in various diseases. CE-MS analysis of intact proteins under the native conditions, the coupling of originally non-MS-compatible modes of CE to MS, and multidimensional set-ups to get more reliable information is the future next steps for proteomic

analysis. Ongoing implementation and further development of simplified, miniaturized, and portable CE devices, to lower analysis costs and improve user experience, will also continue with the advancement in 3D printing and microfabrication. However, apart from the instrumental advances we also need to have more well-trained CE scientists willing to learn new methodologies to fully utilize all advantages CE methods can offer us. We hope to see and possibly contribute to the development and use of CE methods for intact protein analysis in real biological samples.

ACKNOWLEDGMENTS

This work was supported by the Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic and the Slovak Academy of Sciences under the projects VEGA 1/0483/20 and VEGA 1/0514/22. The authors would like to gratefully acknowledge Assoc. Prof. Jozef Marák for his valuable comments and suggestions during the manuscript writing.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID

Katarína Maráková https://orcid.org/0000-0003-2170-6582

REFERENCES

1. Wu Z, Jankowski J. Impact of post-translational modification on the genesis and progression of diseases. Mol Aspects Med. 2022;86:101105.
2. Smith LM, Kelleher NL. Proteoform: a single term describing protein complexity. Nat Methods. 2013;10:186–7.
3. Virág D, Dalmadi-Kiss B, Vékey K, Drahos L, Klebovich I, Antal I, et al. Current trends in the analysis of post-translational modifications. Chromatographia 2020;83:1–10.
4. Mullard A. FDA approves 100th monoclonal antibody product. Nat Rev Drug Discov. 2021;20:491–5.
5. Ramachandra B. Development of impurity profiling methods using modern analytical techniques. Crit Rev Anal Chem. 2017;47:24–36.
6. Pandey S, Pandey P, Tiwari G, Tiwari R. Bioanalysis in drug discovery and development. Pharm Methods. 2010;1:14.
7. Mischak H, Apweiler R, Banks RE, Conaway M, Coon J, Dominiczak A, et al. Clinical proteomics: A need to define the field and to begin to set adequate standards. Proteomics Clin Appl. 2007;1:148–56.
8. Patrie SM, Cline EN. Proteomic and metabolomic approaches to biomarker discovery. Amsterdam, NL: Elsevier; 2020.
9. Van Gool A, Corrales F, Čolović M, Krstić D, Oliver-Martos B, Martínez-Cáceres E, et al. Analytical techniques for multiplex analysis of protein biomarkers. Expert Rev Proteomics. 2020;17:257–73.
10. Albalat A, Husi H, Stalmach A, Schanstra JP, Mischak H. Classical MALDI-MS versus CE-based ESI-MS proteomic profiling in urine for clinical applications. Bioanalysis. 2014;6:247–66.
11. Kang L, Weng N, Jian W. LC–MS bioanalysis of intact proteins and peptides. Biomed Chromatogr. 2020;34:e4633. https://doi.org/10.1002/bmc.4633
12. Bults P, Spanov B, Olaleye O, van de Merbel NC, Bischoff R. Intact protein bioanalysis by liquid chromatography – High-resolution mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci. 2019;1110–1:155–67.
13. Štěpánová S, Kašička V. Applications of capillary electromigration methods for separation and analysis of proteins (2017–mid 2021) – a review. Anal Chim Acta. 2022;1209:339447. https://doi.org/10.1016/j.aca.2022.339447
14. Štěpánová S, Kašička V. Recent developments and applications of capillary and microchip electrophoresis in proteomics and peptidomics (2015-mid 2018). J Sep Sci. 2019;42:398–414.
15. Shen X, Yang Z, McCool EN, Lubeckyj RA, Chen D, Sun L. Capillary zone electrophoresis-mass spectrometry for top-down proteomics. TrAC Trends Analyt Chem. 2019;120:115644.
16. Meyer S, Clases D, Gonzalez de Vega R, Padula MP, Doble PA. Separation of intact proteins by capillary electrophoresis. Analyst 2022;147:2988–96.
17. Kašička V. Recent developments in capillary and microchip electroseparations of peptides (2019–mid 2021). Electrophoresis 2022;43:82–108.
18. Kašička V. Peptide mapping of proteins by capillary electromigration methods. J Sep Sci. 2022;45:4245–79.
19. Bhimwal R, Rustandi RR, Payne A, Dawod M. Recent advances in capillary gel electrophoresis for the analysis of proteins. J Chromatogr A. 2022;1682:463453.
20. Lechner A, Giorgetti J, Gahoual R, Beck A, Leize-Wagner E, François YN. Insights from capillary electrophoresis approaches for characterization of monoclonal antibodies and antibody drug conjugates in the period 2016–2018. J Chromatogr B Analyt Technol Biomed Life Sci. 2019;1122–3:1–17.
21. Gahoual R, Giorgetti J, Beck A, Giorgetti E, François YN. Capillary electromigration separation methods. Amsterdam, NL: Elsevier; 2018.
22. Stutz H. Advances and applications of capillary electromigration methods in the analysis of therapeutic and diagnostic recombinant proteins – A Review. J Pharm Biomed Anal. 2023;222:115089.
23. Kumar R, Guttman A, Rathore AS. Applications of capillary electrophoresis for biopharmaceutical product characterization. Electrophoresis 2022;43:143–66.
24. Kaur H, Beckman J, Zhang Y, Li ZJ, Szigeti M, Guttman A. Capillary electrophoresis and the biopharmaceutical industry: Therapeutic protein analysis and characterization. TrAC Trends Analyt Chem. 2021;144:116407.
25. Dadouch M, Ladner Y, Perrin C. Analysis of monoclonal antibodies by capillary electrophoresis: Sample preparation, separation, and detection. Separations 2021;8:1–30.
26. Puerta A, Garcia-Lopez D, Tejedor-Matellanes P, Gomez-Ruiz L, de la Cruz-Rodriguez R, de Frutos M. Capillary gel electrophoresis of very high molecular weight glycoproteins. Commercial and tailor-made gels for analysis of human

monomeric and secretory immunoglobulin A. J Chromatogr A. 2023;1688:463689.
27. Kwok T, Chan SL, Zhou M, Schaefer A, Bo T, Huang T, et al. High-efficient characterization of complex protein drugs by imaged capillary isoelectric focusing with high-resolution ampholytes. Sep Sci Plus. 2023;6:2200142.
28. Jorgenson JW, Lukacs K. DeArman., Zone electrophoresis in open-tubular glass capillaries. Anal Chem. 1981;53:1298–302.
29. Štěpánová S, Kašička V. Recent developments and applications of capillary and microchip electrophoresis in proteomics and peptidomics (mid-2018–2022). J Sep Sci. 2023, 2300043.
30. Wuethrich A, Quirino JP. A decade of microchip electrophoresis for clinical diagnostics – a review of 2008–2017. Anal Chim Acta. 2019;1045:42–66.
31. Naghdi E, Moran GE, Reinau ME, De Malsche W, Neusüß C. Concepts and recent advances in microchip electrophoresis coupled to mass spectrometry: Technologies and applications. Electrophoresis 2023;44:246–67.
32. Seyfinejad B, Jouyban A. Capillary electrophoresis-mass spectrometry in pharmaceutical and biomedical analyses. J Pharm Biomed Anal. 2022;221:115059.
33. Stolz A, Jooß K, Höcker O, Römer J, Schlecht J, Neusüß C. Recent advances in capillary electrophoresis-mass spectrometry: Instrumentation, methodology and applications. Electrophoresis 2019;40:79–112.
34. Mikšík I. Coupling of CE-MS for protein and peptide analysis. J Sep Sci. 2019;42:385–97.
35. Chen D, McCool EN, Yang Z, Shen X, Lubeckyj RA, Xu T, et al. Recent advances (2019–2021) of capillary electrophoresis-mass spectrometry for multilevel proteomics. Mass Spectrom Rev. 2023;42:617–42.
36. Della Posta S, Fanali C, Gallo V, Fanali S. Recent advances in the hyphenation of electromigration techniques with mass spectrometry. TrAC Trends Analyt Chem. 2022;157:116800.
37. Elshamy YS, Strein TG, Holland LA, Li C, DeBastiani A, Valentine SJ, et al. Nanoflow Sheath Voltage-Free Interfacing of Capillary Electrophoresis and Mass Spectrometry for the Detection of Small Molecules. Anal Chem. 2022;94:11329–36.
38. Ferré S, Drouin N, González-Ruiz V, Rudaz S. Evaluation of a nanoflow interface based on the triple-tube coaxial sheath-flow sprayer for capillary electrophoresis-mass spectrometry coupling in metabolomics. J Chromatogr A. 2021;1641:461982.
39. Konášová R, Koval D, Hošek J, Kašička V. Investigating the position of the separation capillary and emitter tube tips in a nanoflow sheath-liquid CE-ESI-MS interface to decouple the ESI potential. Talanta. 2021;228:122212.
40. Konášová R, Koval D, Dytrtová JJ, Kašička V. Comparison of two low flow interfaces for measurement of mobilities and stability constants by affinity capillary electrophoresis–mass spectrometry. J Chromatogr A. 2018;1568:197–204.
41. Aebersold R, Agar JN, Amster IJ, Baker MS, Bertozzi CR, Boja ES, et al. How many human proteoforms are there? Nat Chem Biol. 2018;14:206–14.
42. McCool EN, Lubeckyj RA, Shen X, Chen D, Kou Q, Liu X, et al. Deep top-down proteomics using capillary zone electrophoresis-tandem mass spectrometry: identification of 5700 proteoforms from the escherichia coli proteome. Anal Chem. 2018;90:5529–33.
43. van der Burg D, Josefsson L, Emmer Å, Sänger-van de Griend CE. Recent capillary electrophoresis applications for upstream and downstream biopharmaceutical process monitoring. TrAC Trends Analyt Chem. 2023;160. https://doi.org/10.1016/j.trac.2023.116975
44. Pontillo C, Filip S, Borràs DM, Mullen W, Vlahou A, Mischak H. CE-MS-based proteomics in biomarker discovery and clinical application. Proteomics Clin Appl. 2015;9:322–34.
45. Latosinska A, Siwy J, Mischak H, Frantzi M. Peptidomics and proteomics based on CE-MS as a robust tool in clinical application: The past, the present, and the future. Electrophoresis 2019;40:2294–300.
46. Stalmach A, Albalat A, Mullen W, Mischak H. Recent advances in capillary electrophoresis coupled to mass spectrometry for clinical proteomic applications. Electrophoresis 2013;34:1452–64.
47. Haselberg R, de Jong GJ, Somsen GW. Capillary electrophoresis–mass spectrometry for the analysis of intact proteins. J Chromatogr A. 2007;1159:81–109.
48. Haselberg R, de Jong GJ, Somsen GW. Capillary electrophoresis-mass spectrometry for the analysis of intact proteins 2007–2010. Electrophoresis. 2011;32:66–82.
49. Haselberg R, de Jong GJ, Somsen GW. CE-MS for the analysis of intact proteins 2010–2012. Electrophoresis 2013;34:99–112.
50. Domínguez-Vega E, Haselberg R, Somsen GW. Capillary Zone Electrophoresis–Mass Spectrometry of Intact Proteins. In: Tran N, Taverna M. (eds) Capillary Electrophoresis of Proteins and Peptides. Methods in Molecular Biology, New York, NY: Humana Press; 2016.
51. Kuzyk VO, Somsen GW, Haselberg R. CE-MS for proteomics and intact protein analysis. In: Colnaghi Simionato AV, editor. Separation techniques applied to omics sciences. advances in experimental medicine and biology. Cham: Springer International Publishing; 2021. pp. 51–86.
52. Lu R-M, Hwang Y-C, Liu I-J, Lee C-C, Tsai H-Z, Li H-J, et al. Development of therapeutic antibodies for the treatment of diseases. J Biomed Sci. 2020;27:1.
53. Johnson IS. Human insulin from recombinant DNA technology. Science. 1983;219:632–37.
54. Grilo AL, Mantalaris A. The increasingly human and profitable monoclonal antibody market. Trends Biotechnol. 2019;37:9–16.
55. Brown DG, Wobst HJ. A decade of FDA-approved drugs (2010–2019): trends and future directions. J Med Chem. 2021;64:2312–38.
56. Agbogbo FK, Ecker DM, Farrand A, Han K, Khoury A, Martin A, et al. Current perspectives on biosimilars. J Ind Microbiol Biotechnol. 2019;46:1297–311.
57. Ratih R, Asmari M, Abdel-Megied AM, Elbarbry F, el Deeb S. Biosimilars: review of regulatory, manufacturing, analytical aspects and beyond. Microchem J. 2021;165:106143.
58. ICH. ICH-guideline Q6B: Specifications: test procedures and acceptance criteria for biotechnological/biological products. 1999:1–16.
59. ICH. ICH Guideline Q5E: Comparability of biotechnological/biological products subject to changes in their manufacturing process. 2004:1–12.
60. ICH. ICH guideline Q11: Development and manufacture of drug substances (chemical entities and biotechnological/biological entities). 2012:1–26.
61. Rahalkar H, Cetintas HC, Salek S. Quality, Non-clinical and Clinical Considerations for Biosimilar Monoclonal Antibody Development: EU, WHO, USA, Canada, and BRICS-TM Reg-

ulatory Guidelines. Front Pharmacol. 2018;9:1079. https://doi.org/10.3389/fphar.2018.01079
62. Tartaglia A, Locatelli M, Samanidou V. Trends in the analysis of biopharmaceuticals by HPLC. Curr Anal Chem. 2020;16:52–8.
63. Kaur H. Stability testing in monoclonal antibodies. Crit Rev Biotechnol. 2021;41:692–714.
64. Guttman A, Filep C, Karger BL. Fundamentals of capillary electrophoretic migration and separation of SDS proteins in borate cross-linked dextran gels. Anal Chem. 2021;93:9267–76.
65. Filep C, Szigeti M, Farsang R, Haberger M, Reusch D, Guttman A. Multilevel capillary gel electrophoresis characterization of new antibody modalities. Anal Chim Acta. 2021;1166:338492.
66. Zhu Z, Lies M, Silzel J. Sodium dodecyl sulfate-capillary gel electrophoresis with native fluorescence detection for analysis of therapeutic proteins. J Pharm Biomed Anal. 2022;213:114689.
67. Chung S, Tian J, Tan Z, Chen J, Lee J, Borys M, et al. Industrial bioprocessing perspectives on managing therapeutic protein charge variant profiles. Biotechnol Bioeng. 2018;115:1646–65.
68. Kahle J, Wätzig H. Determination of protein charge variants with (imaged) capillary isoelectric focusing and capillary zone electrophoresis. Electrophoresis 2018;39:2492–511.
69. Xu T, Sun L. A mini review on capillary isoelectric focusing-mass spectrometry for top-down proteomics. Front Chem. 2021;9:651757. https://doi.org/10.3389/fchem.2021.651757
70. Schjoldager KT, Narimatsu Y, Joshi HJ, Clausen H. Global view of human protein glycosylation pathways and functions. Nat Rev Mol Cell Biol. 2020;21:729–49.
71. Jefferis R. Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action. Trends Pharmacol Sci. 2009;30:356–62.
72. Camperi J, Pichon V, Delaunay N. Separation methods hyphenated to mass spectrometry for the characterization of the protein glycosylation at the intact level. J Pharm Biomed Anal. 2020;178:112921. https://doi.org/10.1016/j.jpba.2019.112921
73. De Leoz MLA, Duewer DL, Fung A, Liu L, Yau HK, Potter O, et al. NIST interlaboratory study on glycosylation analysis of monoclonal antibodies: Comparison of results from diverse analytical methods. Mol Cell Proteomics. 2020;19:11–30.
74. Kumar R, Shah RL, Ahmad S, Rathore AS. Harnessing the power of electrophoresis and chromatography: offline coupling of reverse phase liquid chromatography-capillary zone electrophoresis-tandem mass spectrometry for analysis of host cell proteins in monoclonal antibody producing CHO cell line. Electrophoresis 2021;42:735–41.
75. Atkinson AJ, Colburn WA, DeGruttola VG, DeMets DL, Downing GJ, Hoth DF, et al. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69:89–95.
76. Sanders LJ. From Thebes to Toronto and the 21st century: an incredible journey. Diabetes Spectrum. 2002, 15.
77. Nakayasu ES, Gritsenko M, Piehowski PD, Gao Y, Orton DJ, Schepmoes AA, et al. Tutorial: best practices and considerations for mass-spectrometry-based protein biomarker discovery and validation. Nature Protocols. 2021;16:3737–60.
78. Burke HB. Predicting clinical outcomes using molecular biomarkers. Biomark Cancer. 2016;8:89.
79. Selleck MJ, Senthil M, Wall NR. Making meaningful clinical use of biomarkers. Biomark Insights. 2017;12:1177271917715236. https://doi.org/10.1177/1177271917715236
80. Research and Markets. Global Clinical Biomarkers Market Research Report 2022: Rising Key Player Initiatives & Increasing Demand for Clinical Biomarker Product, https://www.globenewswire.com/en/news-release/2022/09/08/2512260/28124/en/Global-Clinical-Biomarkers-Market-Research-Report-2022-Rising-Key-Player-Initiatives-Increasing-Demand-for-Clinical-Biomarker-Products.html (last time accessed: February 13, 2023)
81. Badve S, Kumar GL. Predictive biomarkers in oncology: applications in precision medicine. Cham: Springer; 2018.
82. Voskuil JLA, Banham AH, Lund-Johansen F, Weller MG, Glerup S. The challenges with the validation of research antibodies. F1000Res. 2017;6:161. https://doi.org/10.12688/f1000research.10851.1
83. Sastre Toraño J, Ramautar R, de Jong G. Advances in capillary electrophoresis for the life sciences. J Chromatogr B Analyt Technol Biomed Life Sci. 2019;1118–9:116–36.
84. Kristoff CJ, Bwanali L, Veltri LM, Gautam GP, Rutto PK, Newton EO, et al. Challenging bioanalyses with capillary electrophoresis. Anal Chem. 2020;92:49–66.
85. Aitekenov S, Gaipov A, Bukasov R. Review: Detection and quantification of proteins in human urine. Talanta. 2021;223:121718. https://doi.org/10.1016/j.talanta.2020.121718
86. Jarvas G, Guttman A, Miękus N, Bączek T, Jeong S, Chung DS, et al. Practical sample pretreatment techniques coupled with capillary electrophoresis for real samples in complex matrices. TrAC Trends Analyt Chem. 2020;122. https://doi.org/10.1016/j.trac.2019.115702
87. Breadmore MC, Grochocki W, Kalsoom U, Alves MN, Phung SC, Rokh MT, et al. Recent advances in enhancing the sensitivity of electrophoresis and electrochromatography in capillaries and microchips (2016–2018). Electrophoresis 2019;40:17–39.
88. Maráková K, Renner BJ, Thomas SL, Opetová M, Tomašovský R, Rai AJ, et al. Solid phase extraction as sample pretreatment method for top-down quantitative analysis of low molecular weight proteins from biological samples using liquid chromatography – triple quadrupole mass spectrometry. Anal Chim Acta. 2023;1243:340801. https://doi.org/10.1016/j.aca.2023.340801
89. Thomas SL, Thacker JB, Schug KA, Maráková K. Sample preparation and fractionation techniques for intact proteins for mass spectrometric analysis. J Sep Sci. 2021;44:211–46.
90. Hanash SM, Pitteri SJ, Faca VM. Mining the plasma proteome for cancer biomarkers. Nature 2008;452:571–9.
91. Luque-Garcia JL, Neubert TA. Sample preparation for serum/plasma profiling and biomarker identification by mass spectrometry. J Chromatogr A. 2007;1153:259–76.
92. Haraldsson B, Sörensson J. Why do we not all have proteinuria? An update of our current understanding of the glomerular barrier. News Physiol Sci. 2004;19:7–10.
93. Adachi J, Kumar C, Zhang Y, Olsen JV, Mann M. The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. Genome Biol. 2006;7:R80.
94. Mischak H, Julian BA, Novak J. High-resolution proteome/peptidome analysis of peptides and low-molecular-weight proteins in urine. Proteomics Clin Appl. 2007;1:792–804.
95. Hrušková H, Voráčová I, Řemínek R, Foret F. Current applications of capillary electrophoresis-mass spectrometry for the

analysis of biologically important analytes in urine (2017 to mid-2021): A review. J Sep Sci. 2022;45:305–24.
96. Macron C, Lane L, Núnez Galindo A, Dayon L. Deep dive on the proteome of human cerebrospinal fluid: a valuable data resource for biomarker discovery and missing protein identification. J Proteome Res. 2018;17:4113–26.
97. Wu S, Brown JN, Tolić N, Meng D, Liu X, Zhang H, et al. Quantitative analysis of human salivary gland-derived intact proteome using top-down mass spectrometry. Proteomics. 2014;14:1211–22.
98. Voeten RLC, Ventouri IK, Haselberg R, Somsen GW. Capillary electrophoresis: trends and recent advances. Anal Chem. 2018;90:1464–81.
99. Lubeckyj RA, Basharat AR, Shen X, Liu X, Sun L. Large-scale qualitative and quantitative top-down proteomics using capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry with nanograms of proteome samples. J Am Soc Mass Spectrom. 2019;30:1435–45.
100. Jooß K, McGee JP, Melani RD, Kelleher NL. Standard procedures for native CZE-MS of proteins and protein complexes up to 800 kDa. Electrophoresis 2021;42:1050–9.
101. Boley DA, Zhang Z, Dovichi NJ. Multisegment injections improve peptide identification rates in capillary zone electrophoresis-based bottom-up proteomics. J Chromatogr A. 2017;1523:123–6.
102. Shah U, Patel A, Patel R, Patel M, Patel A, Kavad M, et al. Overview on capillary electrophoresis with mass spectrometry: application in peptide analysis and proteomics. Int J Pharm Chem Anal. 2021;8:6–15.
103. Gomes FP, Yates JR. Recent trends of capillary electrophoresis-mass spectrometry in proteomics research. Mass Spectrom Rev. 2019;38:445–60.
104. Schlecht J, Jooß K, Neusüß C. Two-dimensional capillary electrophoresis-mass spectrometry (CE-CE-MS): coupling MS-interfering capillary electromigration methods with mass spectrometry. Anal Bioanal Chem. 2018;410:6353–9.
105. McCool EN, Lodge JM, Basharat AR, Liu X, Coon JJ, Sun L. Capillary zone electrophoresis-tandem mass spectrometry with activated ion electron transfer dissociation for large-scale top-down proteomics. J Am Soc Mass Spectrom. 2019;30:2470–9.
106. Zamdborg L, LeDuc RD, Glowacz KJ, Kim Y-B, Viswanathan V, Spaulding IT, et al. ProSight PTM 2.0: improved protein identification and characterization for top down mass spectrometry. Nucleic Acids Res. 2007;35:W701–6.
107. Kou Q, Xun L, Liu X. TopPIC: a software tool for top-down mass spectrometry-based proteoform identification and characterization. Bioinformatics 2016;32:3495–7.
108. Cesnik AJ, Shortreed MR, Schaffer LV, Knoener RA, Frey BL, Scalf M, et al. Proteoform suite: software for constructing, quantifying, and visualizing proteoform families. J Proteome Res. 2018;17:568–78.
109. Cai W, Guner H, Gregorich ZR, Chen AJ, Ayaz-Guner S, Peng Y, et al. MASH suite pro: a comprehensive software tool for top-down proteomics. Mol. Cell Proteomics. 2016;15:703–14.
110. Sun R-X, Luo L, Wu L, Wang R-M, Zeng W-F, Chi H, et al. pTop 1.0: a high-accuracy and high-efficiency search engine for intact protein identification. Anal Chem. 2016;88:3082–90.
111. Wätzig H, Kaupp S, Graf M. Inner surface properties of capillaries for electrophoresis. TrAC Trends Analyt Chem. 2003;22:588–604.
112. Verzola B, Gelfi C, Righetti PG. Protein adsorption to the bare silica wall in capillary electrophoresis. J Chromatogr A. 2000;868:85–99.
113. Leclercq L, Renard C, Martin M, Cottet H. Quantification of adsorption and optimization of separation of proteins in capillary electrophoresis. Anal Chem. 2020;92:10743–50.
114. Stutz H. Protein attachment onto silica surfaces - a survey of molecular fundamentals, resulting effects and novel preventive strategies in CE. Electrophoresis 2009;30:2032–61.
115. Suratman A, Wätzig H. Reproducible protein analysis by CE using linear polyacrylamide-coated capillaries and hydrochloric acid rinsing. Electrophoresis 2007;28:2324–8.
116. Zhu M, Rodriguez R, Hansen D, Wehr T. Capillary electrophoresis of proteins under alkaline conditions. J Chromatogr A. 1990;516:123–31.
117. McCormick RM. Capillary zone electrophoretic separation of peptides and proteins using low pH buffers in modified silica capillaries. Anal Chem. 1988;60:2322–8.
118. Xuan X, Li D. Analytical study of Joule heating effects on electrokinetic transportation in capillary electrophoresis. J Chromatogr A. 2005;1064:227–37.
119. Staub A, Comte S, Rudaz S, Veuthey J-L, Schappler J. Use of organic solvent to prevent protein adsorption in CE-MS experiments. Electrophoresis 2010;31:3326–33.
120. Hajba L, Guttman A. Recent advances in column coatings for capillary electrophoresis of proteins. TrAC Trends Analyt Chem.. 2017;90:38–44.
121. Huhn C, Ramautar R, Wuhrer M, Somsen GW. Relevance and use of capillary coatings in capillary electrophoresis–mass spectrometry. Anal Bioanal Chem. 2010;396:297–314.
122. Hamidli N, Andrasi M, Nagy C, Gaspar A. Analysis of intact proteins with capillary zone electrophoresis coupled to mass spectrometry using uncoated and coated capillaries. J Chromatogr A. 2021;1654:462448. https://doi.org/10.1016/j.chroma.2021.462448
123. Leclercq L, Morvan M, Koch J, Neusüß C, Cottet H. Modulation of the electroosmotic mobility using polyelectrolyte multilayer coatings for protein analysis by capillary electrophoresis. Anal Chim Acta. 2019;1057:152–61.
124. Roca S, Leclercq L, Gonzalez P, Dhellemmes L, Boiteau L, Rydzek G, et al. Modifying last layer in polyelectrolyte multilayer coatings for capillary electrophoresis of proteins. J Chromatogr A. 2023;1692:463837.
125. Stolz A, Hedeland Y, Salzer L, Römer J, Heiene R, Leclercq L, et al. Capillary zone electrophoresis-top-down tandem mass spectrometry for in-depth characterization of hemoglobin proteoforms in clinical and veterinary samples. Anal Chem. 2020;92:10531–9.
126. Goyon A, Francois YN, Colas O, Beck A, Veuthey JL, Guillarme D. High-resolution separation of monoclonal antibodies mixtures and their charge variants by an alternative and generic CZE method. Electrophoresis. 2018;39:2083–90.
127. Hou M, Zhang M, Chen L, Gong K, Pan C, Wang Y. Amplification of lysozyme signal detected in capillary electrophoresis using mixed polymer brushes coating with switchable properties. Talanta. 2019;202:426–35.

128. Wang Y, Li M, Hu F, Wang Y. Online preconcentration of lysozyme in hen egg white using responsive polymer coating in CE. J Sep Sci. 2021;44:3477–88.
129. Liu X, Pan C, Wang Y. PMOXA/PAA brushes toward on-line preconcentration for BSA in capillary electrophoresis. Chin J Chem Phys. 2019;32:497–507.
130. Li M, Wang Y, He K, Wang Y. Determination of pepsin by capillary electrophoresis using mixed polymer coated capillary with switchable properties toward protein adsorption/desorption. J Sep Sci. 2022;45:1960–70.
131. Pero-Gascon R, Benavente F, Minic Z, Berezovski MV, Sanz-Nebot V. On-line aptamer affinity solid-phase extraction capillary electrophoresis-mass spectrometry for the analysis of blood α-synuclein. Anal Chem. 2020;92:1525–33.
132. Pont L, Marin G, Vergara-Barberán M, Gagliardi LG, Sanz-Nebot V, Herrero-Martínez JM, et al. Polymeric monolithic microcartridges with gold nanoparticles for the analysis of protein biomarkers by on-line solid-phase extraction capillary electrophoresis-mass spectrometry. J Chromatogr A. 2020;1622:461097. https://doi.org/10.1016/j.chroma.2020.461097
133. Kawai T. Recent studies on online sample preconcentration methods in capillary electrophoresis coupled with mass spectrometry. Chromatography 2017;38:1–8.
134. Šlampová A, Malá Z, Gebauer P. Recent progress of sample stacking in capillary electrophoresis (2016–2018). Electrophoresis 2019;40:40–54.
135. Zhu G, Sun L, Dovichi NJ. Dynamic pH junction preconcentration in capillary electrophoresis- electrospray ionization-mass spectrometry for proteomics analysis. Analyst 2016;141:5216–20.
136. Wu H, Yi L, Wojcik R, Shi T, Tang K. A separation voltage polarity switching method for higher sample loading capacity and better separation resolution in transient capillary isotachophoresis separation. Analyst 2019;144:454–62.
137. Tang J, Wu H, Hu JJ, Yu J, Zhang J, Wang C, et al. On a separation voltage polarity switching transient capillary isotachophoresis method for higher sample loading capacity and better separation performance. Analyst 2021;146:124–31.
138. Nyssen L, Fillet M, Cavalier E, Servais AC. Highly sensitive and selective separation of intact parathyroid hormone and variants by sheathless CE-ESI-MS/MS. Electrophoresis 2019;40:1550–7.
139. Zhang Y, Fu X, Ping G. Selective enrichment of low-abundance compounds in a mixture by capillary electrophoresis. J Chromatogr A. 2021;1635:461737. https://doi.org/10.1016/j.chroma.2020.461737
140. Di Venere M, Viglio S, Sassera D, Fumagalli M, Bardoni A, Salvini R, et al. Do the complementarities of electrokinetic and chromatographic procedures represent the "Swiss knife" in proteomic investigation? An overview of the literature in the past decade. Electrophoresis 2017;38:1538–50.
141. Jooß K, Scholz N, Meixner J, Neusüß C. Heart-cut nano-LC–CZE–MS for the characterization of proteins on the intact level. Electrophoresis 2019;40:1061–5.
142. Wang WH, Cheung-Lau J, Chen Y, Lewis M, Tang QM. Specific and high-resolution identification of monoclonal antibody fragments detected by capillary electrophoresis–sodium dodecyl sulfate using reversed-phase HPLC with top-down mass spectrometry analysis. MAbs. 2019;11:1233–44.
143. Belov AM, Viner R, Santos MR, Horn DM, Bern M, Karger BL, et al. Analysis of proteins, protein complexes, and organellar proteomes using sheathless capillary zone electrophoresis - native mass spectrometry. J Am Soc Mass Spectrom. 2017;28:2614–34.
144. Le-Minh V, Tran NT, Makky A, Rosilio V, Taverna M, Smadja C. Capillary zone electrophoresis-native mass spectrometry for the quality control of intact therapeutic monoclonal antibodies. J Chromatogr A. 2019;1601:375–84.
145. Jussila M, Sinervo K, Porras SP, Riekkola M-L. Modified liquid junction interface for nonaqueous capillary electrophoresis-mass spectrometry. Electrophoresis 2000;21:3311–7.
146. Assunção NA, Grange Deziderio LA, Paulino LG, Lupetti KO, Carrilho E. Nonaqueous capillary electrophoresis in coated capillaries: An interesting alternative for proteomic applications. Electrophoresis 2005;26:3292–9.
147. Cheng J, Chen DDY. Nonaqueous capillary electrophoresis mass spectrometry method for determining highly hydrophobic peptides. Electrophoresis 2018;39:1216–21.
148. Duarte LM, Moreira RC, Coltro WKT. Nonaqueous electrophoresis on microchips: a review. Electrophoresis 2020;41:434–48.
149. Sänger-van de Griend CE. CE-SDS method development, validation, and best practice—an overview. Electrophoresis 2019;40:2361–74.
150. Beckman J, Song Y, Gu Y, Voronov S, Chennamsetty N, Krystek S, et al. Purity determination by capillary electrophoresis sodium hexadecyl sulfate (CE-SHS): a novel application for therapeutic protein characterization. Anal Chem. 2018;90:2542–7.
151. Sánchez-Hernández L, Montealegre C, Kiessig S, Moritz B, Neusüß C. In-capillary approach to eliminate SDS interferences in antibody analysis by capillary electrophoresis coupled to mass spectrometry. Electrophoresis. 2017;38:1044–52.
152. Quirino JP. Sodium dodecyl sulfate removal during electrospray ionization using cyclodextrins as simple sample solution additive for improved mass spectrometric detection of peptides. Anal Chim Acta. 2018;1005:54–60.
153. Römer J, Montealegre C, Schlecht J, Kiessig S, Moritz B, Neusüß C. Online mass spectrometry of CE (SDS)-separated proteins by two-dimensional capillary electrophoresis. Anal Bioanal Chem. 2019;411:7197–206.
154. Wenz C, Marchetti-Deschmann M, Herwig E, Schröttner E, Allmaier G, Trojer L, et al. A fluorescent derivatization method of proteins for the detection of low-level impurities by microchip capillary gel electrophoresis. Electrophoresis 2010;31:611–7.
155. Seyfried BK, Marchetti-Deschmann M, Siekmann J, Bossard MJ, Scheiflinger F, Turecek PL, et al. Microchip capillary gel electrophoresis of multiply PEGylated high-molecular-mass glycoproteins. Biotechnol J. 2012;7:635–41.
156. Herwig E, Marchetti-Deschmann M, Wenz C, Rüfer A, Redl H, Bahrami S, et al. Sensitive detection of C-reactive protein in serum by immunoprecipitation–microchip capillary gel electrophoresis. Anal Biochem. 2015;478:102–6.
157. Ragland TS, Gossage MD, Furtaw MD, Anderson JP, Steffens DL, Wirth MJ. Electrophoresis of megaDalton proteins inside colloidal silica. Electrophoresis 2019;40:817–23.

158. Farmerie L, Rustandi RR, Loughney JW, Dawod M. Recent advances in isoelectric focusing of proteins and peptides. J Chromatogr A. 2021;1651:462274.
159. Schmailzl J, Vorage MW, Stutz H. Intact and middle-down CIEF of commercial therapeutic monoclonal antibody products under non-denaturing conditions. Electrophoresis 2020;41:1109–17.
160. Sun S, Zhou J-Y, Yang W, Zhang H. Inhibition of protein carbamylation in urea solution using ammonium-containing buffers. Anal Biochem. 2014;446:76–81.
161. Li X, Shi X, Qin X, Yu L, Zhou Y, Rao C. Interlaboratory method validation of imaged capillary isoelectric focusing methodology for analysis of recombinant human erythropoietin. Analytical Methods. 2020;12:3836–43.
162. Wu G, Yu C, Wang W, Zhang R, Li M, Wang L. A platform method for charge heterogeneity characterization of fusion proteins by icIEF. Anal Biochem. 2022;638:114505.
163. Hühner J, Lämmerhofer M, Neusüß C. Capillary isoelectric focusing-mass spectrometry: Coupling strategies and applications. Electrophoresis 2015;36:2670–86.
164. Mokaddem M, d'Orlyé F, Varenne A. Online capillary isoelectric focusing-electrospray ionization mass spectrometry (CIEF-ESI MS) in glycerol–water media for the separation and characterization of hydrophilic and hydrophobic proteins. In: Tran N, Taverna M, editors. Capillary electrophoresis of proteins and peptides. methods in molecular biology. New York, NY: vol 1466. Humana Press; 2018.pp. 57–66.
165. Hühner J, Neusüß C. CIEF-CZE-MS applying a mechanical valve. Anal Bioanal Chem. 2016;408:4055–61.
166. Fan W, Li X, Long Z, Pei D, Shi X, Wang G, et al. Integrating ultra-high-performance liquid chromatography tandem mass spectrometry and imaged capillary isoelectric focusing for in-depth characterization of complex fusion proteins. Rapid Commun. Mass Spectrom. 2023;37:e9484. https://doi.org/10.1002/rcm.9484
167. Schlecht J, Moritz B, Kiessig S, Neusüß C. Characterization of therapeutic mAb charge heterogeneity by iCIEF coupled to mass spectrometry (iCIEF–MS). Electrophoresis. 2023;44:540–8.
168. Wu G, Yu C, Wang W, Du J, Fu Z, Xu G, et al. Mass spectrometry-based charge heterogeneity characterization of therapeutic mAbs with imaged capillary isoelectric focusing and ion-exchange chromatography as separation techniques. Anal Chem. 2023;95:2548–60.
169. Zhang X, Chen T, Li V, Bo T, Du M, Huang T. Cutting-edge mass spectrometry strategy based on imaged capillary isoelectric focusing (iCIEF) technology for characterizing charge heterogeneity of monoclonal antibody. Anal Biochem. 2023;660:114961. https://doi.org/10.1016/j.ab.2022.114961
170. Zhang X, Kwok T, Zhou M, Du M, Li V, Bo T, et al. Imaged capillary isoelectric focusing (iCIEF) tandem high resolution mass spectrometry for charged heterogeneity of protein drugs in biopharmaceutical discovery. J Pharm Biomed Anal. 2023;224:115178. https://doi.org/10.1016/j.jpba.2022.115178
171. Kinoshita M, Nakatsuji Y, Suzuki S, Hayakawa T, Kakehi K. Quality assurance of monoclonal antibody pharmaceuticals based on their charge variants using microchip isoelectric focusing method. J Chromatogr A. 2013;1309:76–83.
172. Mack S, Arnold D, Bogdan G, Bousse L, Danan L, Dolnik V, et al. A novel microchip-based imaged CIEF-MS system for comprehensive characterization and identification of biopharmaceutical charge variants. Electrophoresis 2019;40:3084–91.
173. Ostrowski MA, Mack S, Ninonuevo M, Yan J, ElNaggar M, Gentalen E, et al. Rapid multi-attribute characterization of intact bispecific antibodies by a microfluidic chip-based integrated icIEF-MS technology. Electrophoresis 2023;44:378–86.
174. He X, ElNaggar M, Ostrowski MA, Guttman A, Gentalen E, Sperry J. Evaluation of an icIEF-MS system for comparable charge variant analysis of biotherapeutics with rapid peak identification by mass spectrometry. Electrophoresis 2022;43:1215–22.
175. Everaerts FM, Mikkers FEP, Verheggen TPEM. Isotachophoresis. Sep Purif Methods. 1977;6:287–351.
176. Malá Z, Gebauer P. Analytical isotachophoresis 1967–2022: from standard analytical technique to universal on-line concentration tool. TrAC Trends Analyt Chem. 2023; 158. https://doi.org/10.1016/j.trac.2022.116837
177. Malá Z, Gebauer P. Recent progress in analytical capillary isotachophoresis (2018 - March 2022). J Chromatogr A. 2022;1677:463337.
178. Ramachandran A, Santiago JG. Isotachophoresis: theory and microfluidic applications. Chem Rev. 2022;122:12904–76.
179. Gysler J, Mazereeuw M, Helk B, Heitzmann M, Jaehde U, Schunack W, et al. Utility of isotachophoresis–capillary zone electrophoresis, mass spectrometry and high-performance size-exclusion chromatography for monitoring of interleukin-6 dimer formation. J Chromatogr A. 1999;841:63–73.
180. Bergmann J, Jaehde U, Schunack W. Quantitative trace analysis of interleukin-3, interleukin-6, and basic model proteins using isotachophoresis-capillary zone electrophoresis with hydrodynamic counterflow. Electrophoresis 1998;19:305–10.
181. Madajová V, Šimuničová E, Kaniansky D, Marák J, Zelenská V. Fractionation of glycoforms of recombinant human erythropoietin by preparative capillary isotachophoresis. Electrophoresis 2005;26:2664–73.
182. Kondeková M, Staňová A, Marák J. Methodological aspects of an off-line combination of preparative isotachophoresis and high-performance liquid chromatography with mass spectrometry in the analysis of biological matrices. Electrophoresis 2014;35:1173–80.
183. Marák J, Staňová A, Gajdoštínová S, Škultéty L, Kaniansky D. Some possibilities of an analysis of complex samples by a mass spectrometry with a sample pretreatment by an offline coupled preparative capillary isotachophoresis. Electrophoresis 2011;32:1273–81.
184. Staňová A, Marák J, Rezeli M, Páger C, Kilár F, Kaniansky D. Analysis of therapeutic peptides in human urine by combination of capillary zone electrophoresis-electrospray mass spectrometry with preparative capillary isotachophoresis sample pretreatment. J Chromatogr A. 2011;1218:8701–7.
185. Jacroux T, Bottenus D, Rieck B, Ivory CF, Dong W. Cationic isotachophoresis separation of the biomarker cardiac troponin I from a high-abundance contaminant, serum albumin. Electrophoresis 2014;35:2029–38.
186. Melzer T, Wimmer B, Bock S, Posch TN, Huhn C. Challenges and applications of isotachophoresis coupled to mass spectrometry: a review. Electrophoresis 2020;41:1045–59.
187. Crosnier de Lassichère C, Duc Mai T, Taverna M. Antibody-free detection of amyloid beta peptides biomarkers in cerebrospinal

fluid using capillary isotachophoresis coupled with mass spectrometry. J Chromatogr A. 2019;1601:350–6.
188. Khnouf R, Han C. Isotachophoresis-enhanced immunoassays: challenges and opportunities. IEEE Nanotechnol Mag. 2020;14:6–17.
189. Paratore F, Zeidman Kalman T, Rosenfeld T, Kaigala GV, Bercovici M. Isotachophoresis-based surface immunoassay. Anal Chem. 2017;89:7373–81.
190. Gao F, Wang X-F, Zhang B. Research and application progress of micellar electrokinetic chromatography in separation of proteins. Chin J Anal Chem. 2019;47:805–13.
191. Kahle J, Zagst H, Wiesner R, Wätzig H. Comparative charge-based separation study with various capillary electrophoresis (CE) modes and cation exchange chromatography (CEX) for the analysis of monoclonal antibodies. J Pharm Biomed Anal. 2019;174:460–70.
192. Zayed S, Belal F. Determination of the monoclonal antibody tocilizumab by a validated micellar electrokinetic chromatography method. Chromatographia 2022;85:481–8.
193. Moreno-González D, Haselberg R, Gámiz-Gracia L, García-Campaña AM, de Jong GJ, Somsen GW. Fully compatible and ultra-sensitive micellar electrokinetic chromatography-tandem mass spectrometry using sheathless porous-tip interfacing. J Chromatogr A. 2017;1524:283–9.
194. Mao Z, Chen Z. Advances in capillary electro-chromatography. J Pharm Anal. 2019;9:227–37.
195. Hu L, Yin S, Zhang H, Yang F. Recent developments of monolithic and open-tubular capillary electrochromatography (2017–2019). J Sep Sci. 2020;43:1942–66.
196. Xiao X, Zhang Y, Wu J, Jia L. Poly(norepinephrine)-coated open tubular column for the separation of proteins and recombination human erythropoietin by capillary electrochromatography. J Sep Sci. 2017;40:4636–44.
197. Xiao X, Wang W, Zhang Y, Jia L. Facile preparation of fibrin coated open tubular column for characterization of monoclonal antibody variants by capillary electrochromatography. J Pharm Biomed Anal. 2017;140:377–83.
198. Hajba L, Guttman A. Recent advances in capillary electrochromatography of proteins and carbohydrates in the biopharmaceutical and biomedical field. Crit Rev Anal Chem. 2021;51:289–98.
199. Olabi M, Stein M, Wätzig H. Affinity capillary electrophoresis for studying interactions in life sciences. Methods. 2018;146:76–92.
200. Zhang C, Woolfork AG, Suh K, Ovbude S, Bi C, Elzoeiry M, et al. Clinical and pharmaceutical applications of affinity ligands in capillary electrophoresis: A review. J Pharm Biomed Anal. 2020;177:112882.
201. Hutanu A, Kiessig S, Bathke A, Ketterer R, Riner S, Olaf Stracke J, et al. Application of affinity capillary electrophoresis for charge heterogeneity profiling of biopharmaceuticals. Electrophoresis 2019;40:3014–22.
202. Gstöttner C, Hook M, Christopeit T, Knaupp A, Schlothauer T, Reusch D, et al. Affinity Capillary Electrophoresis-Mass Spectrometry as a Tool to Unravel Proteoform-Specific Antibody-Receptor Interactions. Anal Chem. 2021;93:15133–41.
203. Gstöttner C, Knaupp A, Vidarsson G, Reusch D, Schlothauer T, Wuhrer M, et al. Affinity capillary electrophoresis – mass spectrometry permits direct binding assessment of IgG and FcγRIIa in a glycoform-resolved manner. Front Immunol. 2022;13:980291. https://doi.org/10.3389/fimmu.2022.980291
204. Zhang W, Wu H, Zhang R, Fang X, Xu W. Structure and effective charge characterization of proteins by a mobility capillary electrophoresis based method. Chem Sci. 2019;10:7779–87.
205. He M, Luo P, Hong J, Wang X, Wu H, Zhang R, et al. Structural analysis of biomolecules through a combination of mobility capillary electrophoresis and mass spectrometry. ACS Omega. 2019;4:2377–86.
206. Wu H, Zhang R, Zhang W, Hong J, Xiang Y, Xu W. Rapid 3-dimensional shape determination of globular proteins by mobility capillary electrophoresis and native mass spectrometry. Chem Sci. 2020;11:4758–65.
207. Ruel C, Morani M, Bruneel A, Junot C, Taverna M, Fenaille F, et al. A capillary zone electrophoresis method for detection of Apolipoprotein C-III glycoforms and other related artifactually modified species. J Chromatogr A. 2018;1532:238–45.
208. Biacchi M, Bhajun R, Saïd N, Beck A, François YN, Leize-Wagner E. Analysis of monoclonal antibody by a novel CE-UV/MALDI-MS interface. Electrophoresis 2014;35:2986–95.
209. Zuberovic A, Ullsten S, Hellman U, Markides KE, Bergquist J. Capillary electrophoresis off-line matrix-assisted laser desorption/ionisation mass spectrometry of intact and digested proteins using cationic-coated capillaries. Rapid Commun Mass Spectrom. 2004;18:2946–52.
210. Wu H, Tang K. Highly sensitive and robust capillary electrophoresis-electrospray ionization-mass spectrometry: interfaces, preconcentration techniques and applications. Rev Anal Chem. 2020;39:45–55.
211. Smith RD, Barinaga CJ, Udseth HR. Improved electrospray ionization interface for capillary zone electrophoresis-mass spectrometry. Anal Chem. 1988;60:1948–52.
212. Bonvin G, Rudaz S, Schappler J. In-spray supercharging of intact proteins by capillary electrophoresis–electrospray ionization–mass spectrometry using sheath liquid interface. Anal Chim Acta. 2014;813:97–105.
213. Sundberg BN, Lagalante AF. Coaxial electrospray ionization for the study of rapid in-source chemistry. J Am Soc Mass Spectrom. 2018;29:2023–9.
214. Girault H, Liu B, Qiao L, Bi H, Prudent M, Lion N, et al. Electrochemical reactions and ionization processes. Eur J Mass Spectrom. 2010;16:341–9.
215. Sauer F, Sydow C, Trapp O. A robust sheath-flow CE-MS interface for hyphenation with Orbitrap MS. Electrophoresis 2020;41:1280–6.
216. Kusý P, Klepárník K, Aturki Z, Fanali S, Foret F. Optimization of a pressurized liquid junction nanoelectrospray interface between CE and MS for reliable proteomic analysis. Electrophoresis 2007;28:1964–9.
217. Krenkova J, Foret F. On-line CE/ESI/MS interfacing: Recent developments and applications in proteomics. Proteomics 2012;12:2978–90.
218. Wojcik R, Dada OO, Sadilek M, Dovichi NJ. Simplified capillary electrophoresis nanospray sheath-flow interface for high efficiency and sensitive peptide analysis. Rapid Commun Mass Spectrom. 2010;24:2554–60.
219. Schlecht J, Stolz A, Hofmann A, Gerstung L, Neusüß C. nanoCEasy: an easy, flexible, and robust nanoflow sheath

liquid capillary electrophoresis-mass spectrometry interface based on 3D printed parts. Anal Chem. 2021;93:14593–8.
220. Krenkova J, Kleparnik K, Luksch J, Foret F. Microfabricated liquid junction hybrid capillary electrophoresis-mass spectrometry interface for fully automated operation. Electrophoresis 2019;40(18-19):2263–70. https://doi.org/10.1002/elps.201900049
221. Smith RD, Olivares JA, Nguyen NT, Udseth HR. Capillary zone electrophoresis-mass spectrometry using an electrospray ionization interface. Anal Chem. 1988;60:436–41.
222. Jarvas G, Fonslow B, Yates JR, Foret F, Guttman A. Characterization of a porous nano-electrospray capillary emitter at ultra-low flow rates. J Chromatogr Sci. 2017;55:47–51.
223. Moini M. Simplifying CE–MS operation. 2. interfacing low-flow separation techniques to mass spectrometry using a porous tip. Anal Chem. 2007;79:4241–6.
224. Piestansky J, Barath P, Majerova P, Galba J, Mikus P, Kovacech B, et al. A simple and rapid LC-MS/MS and CE-MS/MS analytical strategy for the determination of therapeutic peptides in modern immunotherapeutics and biopharmaceutics. J Pharm Biomed Anal. 2020, 189, https://doi.org/10.1016/j.jpba.2020.113449
225. Piešťanský J, Čižmárová I, Štefánik O, Matušková M, Horniaková A, Majerová P, et al. Capillary electrophoresis–mass spectrometry with multisegment injection and in-capillary preconcentration for high-throughput and sensitive determination of therapeutic decapeptide triptorelin in pharmaceutical and biological matrices. Biomedicines 2021;9:1488. https://doi.org/10.3390/biomedicines9101488
226. Camperi J, De Cock B, Pichon V, Combes A, Guibourdenche J, Fournier T, et al. First characterizations by capillary electrophoresis of human Chorionic Gonadotropin at the intact level. Talanta 2019;193:77–86.
227. Pajaziti B, Petkovska R, Andrási M, Nebija D. Application of the capillary zone electrophoresis (CZE) and capillary gel electrophoresis (CGE) for the separation of human insulin, insulin lispro and their degradation products. Pharmazie 2020;75:167–71.
228. Andrasi M, Pajaziti B, Sipos B, Nagy C, Hamidli N, Gaspar A. Determination of deamidated isoforms of human insulin using capillary electrophoresis. J Chromatogr A. 2020;1626:461344. https://doi.org/10.1016/j.chroma.2020.461344
229. Gstöttner C, Nicolardi S, Haberger M, Reusch D, Wuhrer M, Domínguez-Vega E. Intact and subunit-specific analysis of bispecific antibodies by sheathless CE-MS. Anal Chim Acta. 2020;1134:18–27.
230. Gstöttner C, Vergoossen DLE, Wuhrer M, Huijbers MGM, Domínguez-Vega E. Sheathless CE-MS as a tool for monitoring exchange efficiency and stability of bispecific antibodies. Electrophoresis 2021;42:171–6.
231. Haselberg R, De Vijlder T, Heukers R, Smit MJ, Romijn EP, Somsen GW, et al. Heterogeneity assessment of antibody-derived therapeutics at the intact and middle-up level by low-flow sheathless capillary electrophoresis-mass spectrometry. Anal Chim Acta. 2018;1044:181–90.
232. Belov AM, Zang L, Sebastiano R, Santos MR, Bush DR, Karger BL, et al. Complementary middle-down and intact monoclonal antibody proteoform characterization by capillary zone electrophoresis – mass spectrometry. Electrophoresis 2018;39:2069–82.
233. Giorgetti J, Lechner A, Del Nero E, Beck A, François Y-N, Leize-Wagner E. Intact monoclonal antibodies separation and analysis by sheathless capillary electrophoresis-mass spectrometry. Eur J Mass Spectrom. 2019;25:324–32.
234. Giorgetti J, Beck A, Leize-Wagner E, François YN. Combination of intact, middle-up and bottom-up levels to characterize 7 therapeutic monoclonal antibodies by capillary electrophoresis – Mass spectrometry. J Pharm Biomed Anal. 2020;182:113107. https://doi.org/10.1016/j.jpba.2020.113107
235. Naumann L, Schlossbauer P, Klingler F, Hesse F, Otte K, Neusüß C. High-throughput glycosylation analysis of intact monoclonal antibodies by mass spectrometry coupled with capillary electrophoresis and liquid chromatography. J Sep Sci. 2022;45:2034–44.
236. Wang L, Bo T, Zhang Z, Wang G, Tong W, Da Yong Chen D. High Resolution capillary isoelectric focusing mass spectrometry analysis of peptides, proteins, and monoclonal antibodies with a flow-through microvial interface. Anal Chem. 2018;90:9495–503.
237. Wang L, Chen DDY. Analysis of four therapeutic monoclonal antibodies by online capillary isoelectric focusing directly coupled to quadrupole time-of-flight mass spectrometry. Electrophoresis 2019;40:2899–907.
238. Dai J, Lamp J, Xia Q, Zhang Y. Capillary Isoelectric Focusing-Mass Spectrometry Method for the Separation and Online Characterization of Intact Monoclonal Antibody Charge Variants. Anal Chem. 2018;90:2246–54.
239. Li M, Zhao X, Shen D, Wu G, Wang W, Yu C, et al. Identification of a monoclonal antibody clipping variant by cross-validation using capillary electrophoresis – sodium dodecyl sulfate, capillary zone electrophoresis – mass spectrometry and capillary isoelectric focusing – mass spectrometry. J Chromatogr A. 2022;1684:463560. https://doi.org/10.1016/j.chroma.2022.463560
240. Chen C-H, Feng H, Guo R, Li P, Laserna AKC, Ji Y, et al. Intact NIST monoclonal antibody characterization—proteoforms, glycoforms—using CE-MS and CE-LIF. Cogent Chem. 2018;4:1480455.
241. Deyanova EG, Huang RYC, Madia PA, Nandi P, Gudmundsson O, Chen G. Rapid fingerprinting of a highly glycosylated fusion protein by microfluidic chip-based capillary electrophoresis–mass spectrometry. Electrophoresis 2021;42:460–4.
242. Fussl F, Trappe A, Carillo S, Jakes C, Bones J. Comparative elucidation of cetuximab heterogeneity on the intact protein level by cation exchange chromatography and capillary electrophoresis coupled to mass spectrometry. Anal Chem. 2020;92:5431–8.
243. Carillo S, Jakes C, Bones J. In-depth analysis of monoclonal antibodies using microfluidic capillary electrophoresis and native mass spectrometry. J Pharm Biomed Anal. 2020;185:113218.
244. Sun Q, Wang L, Li N, Shi L. Characterization and monitoring of charge variants of a recombinant monoclonal antibody using microfluidic capillary electrophoresis-mass spectrometry. Anal Biochem. 2021;625:114214. https://doi.org/10.1016/j.ab.2021.114214
245. Cao L, Fabry D, Lan K. Rapid and comprehensive monoclonal antibody characterization using microfluidic CE-MS. J Pharm Biomed Anal. 2021;204. https://doi.org/10.1016/j.jpba.2021.114251

246. Wu Z, Wang H, Wu J, Huang Y, Zhao X, Nguyen JB, et al. High-sensitivity and high-resolution therapeutic antibody charge variant and impurity characterization by microfluidic native capillary electrophoresis-mass spectrometry. J Pharm Biomed Anal. 2023;223:115147. https://doi.org/10.1016/j.jpba.2022.115147
247. Montealegre C, Neusüß C. Coupling imaged capillary isoelectric focusing with mass spectrometry using a nanoliter valve. Electrophoresis. 2018;39:1151–4.
248. Shen X, Liang Z, Xu T, Yang Z, Wang Q, Chen D, et al. Investigating native capillary zone electrophoresis-mass spectrometry on a high-end quadrupole-time-of-flight mass spectrometer for the characterization of monoclonal antibodies. Int J Mass Spectrom. 2021;462:116541.
249. Ricci P, Magalhães P, Krochmal M, Pejchinovski M, Daina E, Caruso MR, et al. Urinary proteome signature of Renal Cysts and Diabetes syndrome in children. Sci Rep. 2019;9:1–10.
250. Mavrogeorgis E, Mischak H, Latosinska A, Siwy J, Jankowski V, Jankowski J. Reproducibility evaluation of urinary peptide detection using CE-MS. Molecules. 2021;26. https://doi.org/10.3390/molecules26237260
251. Moran AB, Domínguez-Vega E, Nouta J, Pongracz T, de Reijke TM, Wuhrer M, et al. Profiling the proteoforms of urinary prostate-specific antigen by capillary electrophoresis – mass spectrometry. J Proteomics. 2021;238:104148.
252. Lombard-Banek C, Yu Z, Swiercz AP, Marvar PJ, Nemes P. A microanalytical capillary electrophoresis mass spectrometry assay for quantifying angiotensin peptides in the brain. Anal Bioanal Chem. 2019;411:4661–71.
253. Luo RY, Wong C, Xia JQ, Glader BE, Shi RZ, Zehnder JL. Neutral-coating capillary electrophoresis coupled with high-resolution mass spectrometry for top-down identification of hemoglobin variants. Clin Chem. 2023;69:56–67.
254. Li Y, Compton PD, Tran JC, Ntai I, Kelleher NL. Optimizing capillary electrophoresis for top-down proteomics of 30–80 kDa proteins. Proteomics. 2014;14:1158–64.
255. Meixner M, Pattky M, Huhn C. Novel approach for the synthesis of a neutral and covalently bound capillary coating for capillary electrophoresis-mass spectrometry made from highly polar and pH-persistent N-acryloylamido ethoxyethanol. Anal Bioanal Chem. 2020;412:561–75.
256. Batz NG, Mellors JS, Alarie JP, Ramsey JM. Chemical vapor deposition of aminopropyl silanes in microfluidic channels for highly efficient microchip capillary electrophoresis-electrospray ionization-mass spectrometry. Anal Chem. 2014;86:3493–500.
257. Bendahl L, Hansen SH, Gammelgaard B. Capillaries modified by noncovalent anionic polymer adsorption for capillary zone electrophoresis, micellar electrokinetic capillary chromatography and capillary electrophoresis mass spectrometry. Electrophoresis. 2001;22:2565–73.
258. Thompson TJ, Foret F, Vouros P, Karger BL. Capillary electrophoresis/electrospray ionization mass spectrometry: improvement of protein detection limits using on-column transient isotachophoretic sample preconcentration. Anal Chem. 1993;65:900–6.
259. Zhu G, Sun L, Yan X, Dovichi NJ. Stable, reproducible, and automated capillary zone electrophoresis–tandem mass spectrometry system with an electrokinetically pumped sheath–flow nanospray interface. Anal Chim Acta. 2014;810:94–8.
260. Lubeckyj RA, McCool EN, Shen X, Kou Q, Liu X, Sun L. Single-shot top-down proteomics with capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry for identification of nearly 600 Escherichia coli proteoforms. Anal Chem. 2017;89:12059–67.
261. Zhao Y, Sun L, Zhu G, Dovichi NJ. Coupling capillary zone electrophoresis to a Q Exactive HF mass spectrometer for top-down proteomics: 580 proteoform identifications from yeast. J Proteome Res. 2016;15:3679–85.
262. Xia S, Zhan L, Qiu B, Lu M, Chi Y, Chen G. On-line preconcentration and quantitative analysis of peptide hormone of brain and intestine using on-column transient isotachophoresis coupled with capillary electrophoresis/electrospray ionization mass spectrometry. Rapid Commun Mass Spectrom. 2008;22:3719–26.
263. Han X, Wang Y, Aslanian A, Bern M, Lavallée-Adam M, Yates JR. Sheathless capillary electrophoresis-tandem mass spectrometry for top-down characterization of pyrococcus furiosus proteins on a proteome scale. Anal Chem. 2014;86:11006–12.
264. Hasan MN, Park SH, Oh E, Song EJ, Ban E, Yoo YS. Sensitivity enhancement of CE and CE-MS for the analysis of peptides by a dynamic pH junction. J Sep Sci. 2010;33:3701–9.
265. Su S, Yu Y. Online preconcentration of recombinant Arg-Gly-Asp-hirudin using dynamic pH junction for analysis in human urine samples by capillary electrophoresis-mass spectrometry. J Chromatogr A. 2009;1216:1490–5.
266. Ye H, Xia S, Lin W, Yu L, Xu X, Zheng C, et al. CE-ESI-MS coupled with dynamic pH junction online concentration for analysis of peptides in human urine samples. Electrophoresis 2010;31:3400–6.
267. Medina-Casanellas S, Benavente F, Barbosa J, Sanz-Nebot V. Transient isotachophoresis in on-line solid phase extraction capillary electrophoresis time-of-flight-mass spectrometry for peptide analysis in human plasma. Electrophoresis 2011;32:1750–9.
268. Wu Z, Wang H, Wu J, Huang Y, Zhao X, Nguyen JB, et al. High-sensitivity and high-resolution therapeutic antibody charge variant and impurity characterization by microfluidic native capillary electrophoresis-mass spectrometry. J Pharm Biomed Anal. 2023;223. https://doi.org/10.1016/j.jpba.2022.115147
269. Shen X, Liang Z, Xu T, Yang Z, Wang Q, Chen D, et al. Investigating native capillary zone electrophoresis-mass spectrometry on a high-end quadrupole-time-of-flight mass spectrometer for the characterization of monoclonal antibodies. Int J Mass Spectrom. 2021;462:116541. https://doi.org/10.1016/j.ijms.2021.116541

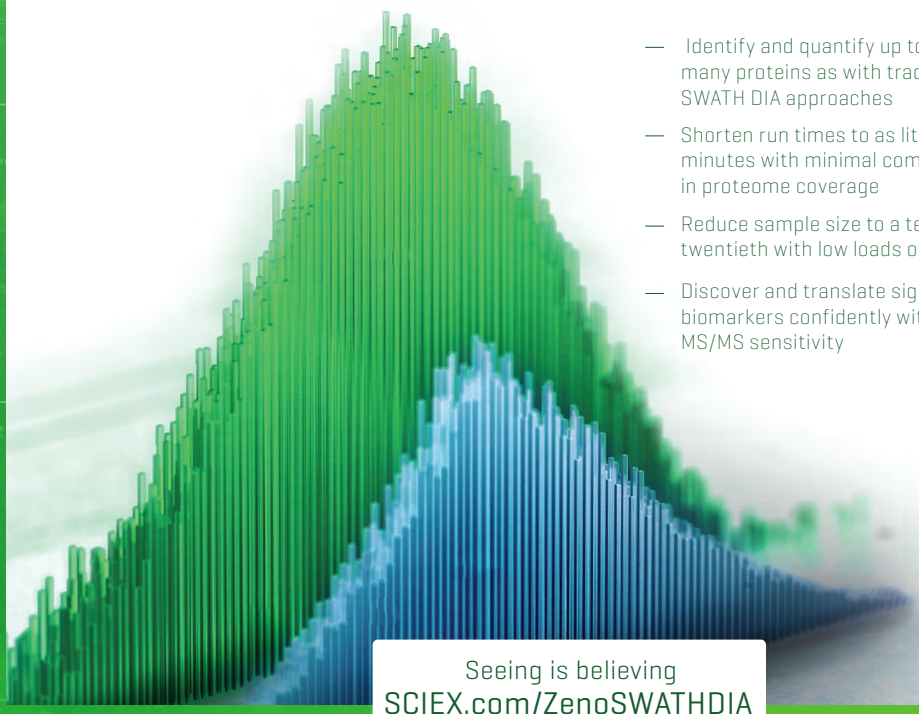
SCIEX
The Power of Precision

Are you ready for a step-change in data independent acquisition?

Introducing

Zeno SWATH
data independent acquisition (DIA)

Enhance speed, sensitivity and accuracy of your workflows using Zeno SWATH data independent acquisition (DIA) for MS-based proteomics.



Harnessing the power of the Zeno trap for Zeno SWATH DIA allows you to:

- Identify and quantify up to twice as many proteins as with traditional SWATH DIA approaches
- Shorten run times to as little as 5 minutes with minimal compromise in proteome coverage
- Reduce sample size to a tenth or twentieth with low loads of 10 ng
- Discover and translate significant biomarkers confidently with higher MS/MS sensitivity

Seeing is believing
SCIEX.com/ZenoSWATHDIA