



CZE-MS peptide mapping: To desalt or not to desalt?

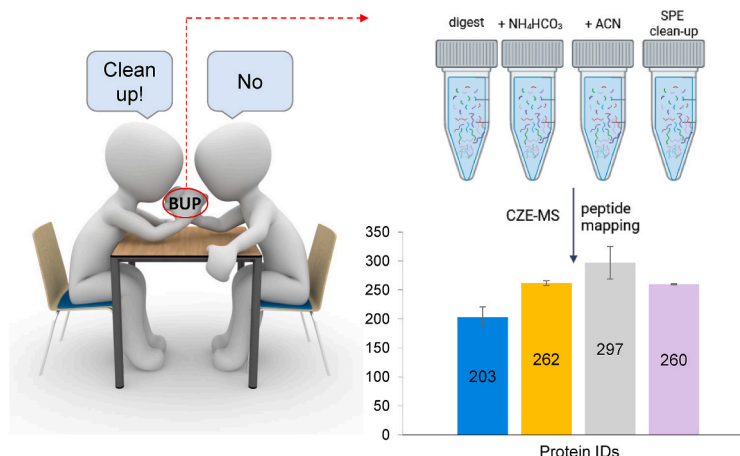
Cynthia Nagy, Melinda Andrasi, Ruben Szabo, Attila Gaspar*

Department of Inorganic and Analytical Chemistry, University of Debrecen, Egyetem ter 1, Debrecen, 4032, Hungary

HIGHLIGHTS

- Peptide maps of desalted digests were compared to those of non-desalted digests.
- Benefits in sequence coverage can be gained compared to desalted samples.
- Desalting offered no considerable benefits in sensitivity compared to non-desalted samples.
- Sample clean-up in BUP sample preparation prior to CZE-MS can be omitted.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Desalting
Bottom-up proteomics
Capillary zone electrophoresis
Mass spectrometry
Peptide mapping

ABSTRACT

Background: In “shotgun” approaches involving high-performance liquid chromatography or capillary zone electrophoresis (CZE), matrix removal prior to sample analysis is considered as an indispensable tool. Despite the fact that CZE offers a high tolerance towards salts, most publications reported on the use of desalting. There seems to be no clear consensus on the utilization of desalting in the CZE-MS community, most probably due to the absence of works addressing the comparison of desalted and non-desalted digests. Our aim was to fill this research gap using protein samples of varying complexity in different sample matrices.

Results: First, standard protein digests were analyzed to build the knowledge on the effect of sample clean-up by solid-phase extraction (SPE) pipette tips and the possible stacking phenomena induced by different sample matrices. Desalting led to a somewhat altered peptide profile, the procedure affected mostly the hydrophilic peptides (although not to a devastating extent). Nevertheless, desalting samples allowed remarkable stacking efficiency owing to their low-conductivity sample background, enabling a so-called field-amplified sample stacking phenomenon. Non-desalted samples also produced a stacking event, the mechanism of which is based on transient-isotachopheresis due to the presence of high-mobility ions in the digestion buffer itself. Adding either extra ammonium ions or acetonitrile into the non-desalted digests enhanced the stacking efficiency. A complex sample (yeast cell lysate) was also analyzed with the optimal conditions, which yielded similar tendencies.

* Corresponding author.

E-mail address: gaspar@science.unideb.hu (A. Gaspar).

Significance: Based on these results, we propose that sample clean-up in the bottom-up sample preparation process prior to CZE-MS analysis can be omitted. The preclusion of desalting can even enhance detection sensitivity, separation efficiency or sequence coverage.

1. Introduction

Proteomics is a field of study addressing the extensive characterization of the proteome. Mass spectrometers (MS) have a leading role in determining a wealth of information about protein samples, including quantitative and qualitative aspects as well as the identification and localization of post-translational modification, etc [1,2]. A typical bottom-up proteomic (BUP) workflow consists of the hydrolysis of proteins into peptides, the separation of these peptides by high performance liquid chromatography (HPLC) or capillary electrophoresis (CE) and their tandem MS detection [3]. The use of bioinformatics tools enables the identification of peptides and ultimately proteins in our sample.

Among the myriad of BUP-related publications, reverse-phased liquid chromatography (RPLC) still seems to be dominating the stage. However, the significance of capillary (zone) electrophoresis (CZE) is also increasing [4,5], mostly due to the orthogonal separation mode it offers, which is best demonstrated by works comparing the performance of RPLC and CZE for a given sample [6–14]. The results obtained show that the two methods complement each other to some extent regarding peptide and protein identifications. Several groups have harnessed this complementarity by utilizing both chromatographic and electrophoretic dimensions either as SPE - multistep elution - CZE workflows [7,15–19] or “true” LC-CZE-MS/MS systems [19–26], which typically enhanced proteome coverage.

Historically, the general discontent with CZE-MS derives from the fact that its sensitivity lags behind that of HPLC-MS, mostly because of the low injection volumes (few nL) and the dilution of the CE effluent by the sheath liquid. The relatively poor concentration sensitivity of CZE inhibits the acquisition of sufficient MS/MS spectral data. There has been considerable improvement in CZE-MS performance, which can be attributed to the evolution of interfacial designs and the utilization of on-line sample enrichment techniques (where analytes present in a large injected volume are transferred to a narrow, concentrated zone). The use of on-line preconcentration strategies is quite convenient since the process requires no additional modification of instrumental setup – analytes accumulate on the boundaries of a discontinuous liquid system (i.e. solutions having different physicochemical properties) due to changes in their velocities. Electrophoretic preconcentration comes in many forms, including field-amplified sample stacking (FASS) [8,15,27], transient isotachopheresis (t-ITP) [28–30], dynamic pH-junction (DPJ) [19,21,31] or pH barrage [32,33]. Analyte enrichment is generally achieved by the proper pairing of BGE and sample (matrix) compositions. This pairing is based on differences in conductivity (FASS), pH (dynamic pH junction) or salt content (t-ITP) between the BGE and the sample solution, however, this implies that samples need to be processed following digestion to fit the requirements of efficient stacking conditions [34]. Most BUP workflows using CE-MS resort to a preconcentration technique or a combination of those (Table S1).

The enzymatic cleavage of proteins is usually preceded by a series of sample preparation steps in order to render the protein accessible by proteases. In general, these steps include denaturation (e.g., heat, urea, thiourea, guanidine-HCl, Rapigest™), reduction (e.g., dithiothreitol (DTT), tris(2-carboxyethyl)phosphine hydrochloride (TCEP)) and alkylation of cysteine residues (e.g., iodoacetamide (IAA), chloroacetamide). The optimum pH for digestion is ensured by dilution with an appropriate buffer (e.g., ammonium bicarbonate (AB), Tris-HCl). After proteolysis, the reaction is quenched by acidification (e.g., formic acid (FA), trifluoroacetic acid (TFA)). Sample cleanup is commonly carried out using C18 solid phase extraction (SPE). SPE offers the

elimination of salts from the sample as well as enrichment prior to analysis. However, the use of SPE can lead to sample loss. Similar to RPLC separations, small hydrophilic peptides might not be retained, while larger hydrophobic peptides cannot be easily eluted from the packing, leading to an incomplete set of cleavage products.

A thorough literature search on BUP-related CZE-MS works was conducted in order to explore the prevalence of utilizing desalting (matrix removal) and preconcentration techniques. It was found that regardless of potential sample loss, most publications (38 out of 53 listed in Table S1) reported on the use of desalting. Desalting certainly offers greater freedom in the choice of subsequent stacking technique. On the contrary, in cases where there was no post-digestion sample clean-up, predominantly t-ITP stacking was employed. This sparks the question whether desalting is indeed necessary. It is well-known that during the CZE analysis of peptides some sort of stacking occurs nearly automatically when large sample volumes are injected, whatever the sample matrix. However, to the best of our knowledge, there has been no comparative CZE-MS investigation of desalted and non-desalted samples with respect to the stacking they induce.

In this work we studied the effect of sample cleanup with SPE in protein digests of varying complexity. In order to explore whether there is anything to be gained from matrix removal, peptide maps of desalted digests were compared to those of non-desalted digests. The possibility of further enhancing stacking was also explored in the case of non-desalted samples.

2. Materials and methods

2.1. Reagents and solutions

Analytical grade reagents were used. Urea, DTT, IAA, NH₄HCO₃ (AB), porcine pancreas trypsin (Type IX-S, lyophilized powder) (all Sigma products, St. Louis, MO, USA) stocks solutions were prepared in double-deionized water (Elix-3, Millipore, Darmstadt, Germany). Human serum albumin (HSA), myoglobin from equine heart (MYG), lysozyme from chicken egg white (LYS) and β -casein from bovine milk (β -CAS) (Sigma) were used as model proteins. Yeast cell lysate was purchased from Promega (Madison, WI, USA). Formic acid solutions, isopropanol (IPA), methanol (MeOH) and acetonitrile (ACN) were purchased from VWR (Radnor, PA, USA). Non-stimulated human tear samples were collected using a sterile glass capillary from a healthy female volunteer.

2.2. Instrumentation, software

The analysis of protein digests was conducted in 90 cm \times 50 μ m id. bare fused silica capillaries (Polymicro, Phoenix, AZ, USA) with a 7100 model CE instruments (Agilent, Waldbronn, Germany) using MS detection (maXis II UHR ESI-QTOF MS, Bruker, Bremen, Germany). CZE-MS hyphenation was achieved with a CE-ESI Sprayer interface (G1607B, Agilent).

Sample introduction was carried out hydrodynamically (HD), optimal injection parameters: 50 mbar, 112 s. In most cases the BGE consisted of 1 M formic acid (pH 1.9). Specific injection parameters and BGE compositions are given at the figure captions. Peptides were separated using 21 kV voltage. Capillary was preconditioned before each measurement with BGE (6 min, 4 bar); the postconditioning included successive washing steps with MeOH (2 min, 4 bar), de-ionized water (2 min, 4 bar) and BGE (6 min, 4 bar) and finally the application of voltage (20 s, 15 kV).

A 1260 Infinity II isocratic pump (Agilent) was used for delivering the sheath liquid (SL) (IPA:water = 1:1 + 0.1 % formic acid) at a flow rate of 4 $\mu\text{L}/\text{min}$. The CE instrument and the pump were operated by OpenLAB CDS Chemstation software (Agilent), the MS instrument was controlled by otofControl version 4.1 (build: 3.5, Bruker). Peptide maps and mass spectra were processed by Compass DataAnalysis version 4.4 (build: 200.55.2969, Bruker). Parameters applied for MS acquisition: positive ionization mode; nebulizer pressure: 0.5 bar; dry gas temperature: 200 $^{\circ}\text{C}$; dry gas flow rate: 4 L/min; capillary voltage: 4500 V; end plate offset: 500 V; MS spectra rate: 3 Hz; MS/MS spectra rate: 1–4 Hz; mass range: 50–2200 m/z . Collision induced dissociation (CID) was applied for producing fragment ions. MS/MS spectra were acquired in data-dependent mode. From each MS scan 6 (HSA, model protein mix) or 8 (tear samples, yeast cell lysate) precursor ions were selected. Active exclusion was applied (after the acquisition of 1 MS/MS spectrum the precursor was excluded from fragmentation for 0.5 min, except if the precursor intensity showed 5x increase); singly charged ions were excluded. Na-formate was injected after each run, which enabled internal m/z calibration. The generated peak lists were exported in MGF format. Peptide identification was performed with Byonic software (ver.: 3.9.6., Protein Metrics, Cupertino, CA, USA). Byonic runs were carried out with the following settings: fully specific digestion; maximum number of missed cleavages: 2; precursor mass tolerance: 15 ppm; fragment mass tolerance: 40 ppm; carbamidomethylation (+57.0215 Da) at Cys as fixed modification, deamidation (+0.9840 Da) at Asn and Gln, oxidation (+15.9949 Da) at Met, Trp and Tyr, phosphorylation (+79.9663) at Ser, Thr and Tyr as variable modifications.

2.3. Sample preparation

Four protein samples were investigated: (i) HSA, (ii) standard protein mixture containing HSA, MYG, LYS and β -CAS, (iii) yeast cell lysate and (iv) human tear. The pretreatment and digestion were executed as follows: \sim 2.5 mg protein was dissolved in 100 μL 25 mM AB and immediately 300 μL 8 M urea solution (30 min, room temperature) was added. The mixture was allowed to stand at room temperature for 30 min to denature the protein. In order to reduce the disulfide bonds, 40 μL 100 mM DTT was used (1 h, 37 $^{\circ}\text{C}$). The recombination of disulfide bonds was prevented by adding 40 μL 200 mM IAA (alkylating agent) to the solution (45 min, room temperature in dark). Finally, the mixture was diluted with 2 mL 25 mM AB to reduce urea concentration $<$ 1 M. For proteolysis, 50 μL 1 mg/mL freshly prepared trypsin solution was pipetted into the pretreated sample. Digestion was performed overnight (16 h, 37 $^{\circ}\text{C}$). Enzyme activity was quenched by the addition of 1 % formic acid (280 μL).

Digestion of the yeast cell lysate was performed as follows: to 20 μL lysate 30–30 μL 8 M urea and 25 mM AB were added. For reduction and alkylation 8–8 μL 100 mM DTT and 200 mM IAM were used. After diluting the mixture with 149 μL AB solution trypsin was added (5 μL , $c = 1$ mg/mL). Digestion was terminated by adding 27 μL 1 % formic acid to the solution.

For the digestion of human tears, the following quantities were used: 0.9 mg urea, 5 μL tear sample; 0.3–0.3 μL 100 mM DTT and 200 mM IAM, 9.5 μL 25 mM AB, 0.3 μL trypsin solution (1 mg/mL), 1.7 μL 1 % formic acid.

Desalting was carried out with 10 μL C18 ZipTips (Sigma) according to the manufacturer's protocol. 0.1 % formic acid was used for wetting the resin and for desalting. ACN:water (70:30) + 0.1 % formic acid ensured the elution of adsorbed peptides.

Four types of samples were analyzed: (i) non-desalted, (ii), non-desalted + 18 mM AB, (iii) non-desalted + 50 % ACN and (iv) desalted.

3. Results and discussion

3.1. Analysis of non-desalted digests

For FASS, the fundamental requirement is that the conductivity of the sample solution should be considerably lower than that of the BGE, for which desalting is actually a suitable solution. The digested samples typically contain a generous amount of organic solvent [15,27,35,36], which further decreases the conductivity of the sample solution and thereby increases the FASS effect. However, the essence of t-ITP is that the target analytes arrange between a leading (LE) and terminating electrolyte (TE) in order of decreasing mobility. It is quite customary to dissolve or dilute the samples after digestion with ammonium acetate (AmAc), whereby the high mobility NH_4^+ ions present in high concentration behave as leading ions [28,29,37,38]. Usually, protein samples are prepared in AB in order to set an optimal pH for digestion, therefore, a certain level of preconcentration can be expected because of the presence of NH_4^+ in high abundance. In order to quickly assess the validity of this assumption, as a starting point, we analyzed HSA digests in two different sample backgrounds: digests were 2-fold diluted either with (i) BGE, which is 1 M formic acid or (ii) deionized water, using relatively high sample loading (\sim 107 nL, 6 % of total capillary length) (Fig. S-1.). The CZE-MS peptide map belonging to the sample matrix containing a considerable amount of formic acid showed excessive peak broadening and low signal intensities. This was in line with our expectations, since the basic requirement of stacking – discontinuity in terms of the physicochemical properties of the liquids introduced into the capillary – was not satisfied.

Without further optimization to the sample solution we analyzed the non-desalted HSA digest with increasing injection volumes to explore the extent of preconcentration (Fig. 1). Increasing the sample volume leads to enhanced peak heights with some level of peak broadening and overlapping, which are extensive only at very high injection times (50 mbar \times 170s–163 nL).

The co-migration of peptides can cause ion-suppression phenomena and the undersampling of peaks for fragmentation [37,39], leading ultimately to limited sequence coverage. Therefore, besides the enrichment of components, efforts are also focused on enhancing analytical resolution by extending the migration window [36,40,41].

For a more detailed evaluation, three peaks (largely differing in their migration times and hence physico-chemical properties) were selected in the electropherograms (Fig. 1) whose height, width (FWHM) and theoretical plate number data were compared (Fig. 2). Extending the injection times results in a considerable increase in signal height, especially for the fastest component with only a modest contribution to peak broadening. The later migrating component 3, however, shows smaller levels of preconcentration, the increase of peak width is considerable. Most probably these slower peptides have mobilities similar to or lower than that of the terminating ion, thus they might be excluded from the focusing step. Based on peak intensity, FWHM and SC % values, an injection volume of \sim 107 nL (50 mbar, 112 s) proved to be most optimal and thus was chosen for further experiments. SC% values compare well with those found in recently published works, however, in the literature mostly a sequence homolog - bovine serum albumin (BSA) - is used as model protein for method optimization. For BSA various results were obtained, mostly depending on the stacking technique of choice and the sample concentration: \sim 23.4 %; 48.6 % and 56.4 % for CZE methods employing no preconcentration, DPJ and t-ITP, respectively (c : 0.033 mg/mL) [40]. Zhu et al. reported on 70 % SC for 0.05 mg/mL BSA digest utilizing DPJ, whereas the conventional injection of 1 mg/mL digest yielded 66 % coverage [42]. Our results obtained by CE-MS are consistent with those attainable by HPLC-MS, where \sim 60 % coverage was found for HSA [43].

In addition to t-ITP, other mechanisms also contribute to analyte stacking, e.g., FASS and t-ITP often go hand in hand. FASS is also known to be less effective for components having very low velocity. Therefore,

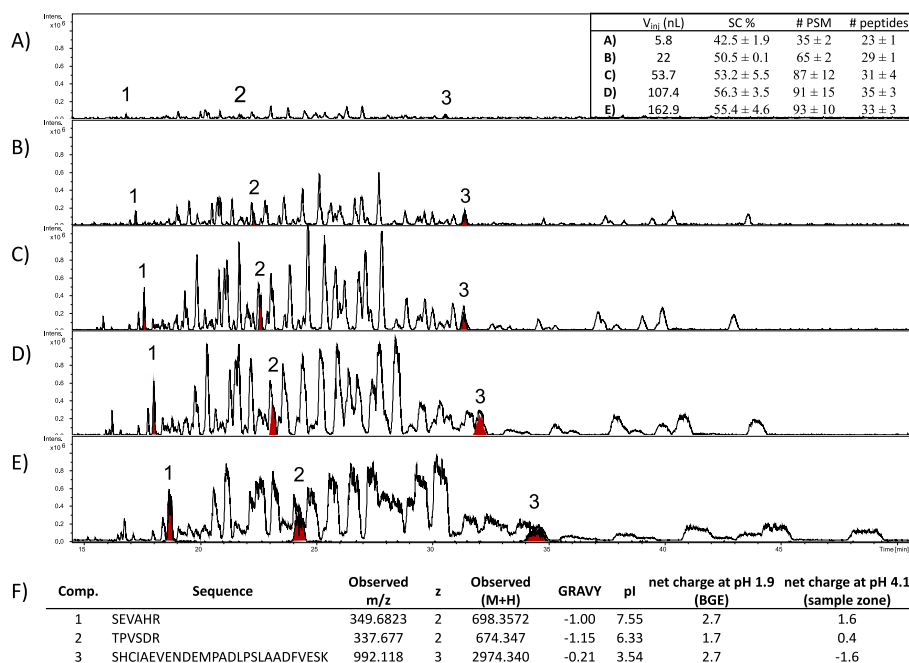


Fig. 1. CZE-MS peptides maps of a non-desalted HSA tryptic digest ($c = 0.089$ mg/mL) with increasing sample loading. HD injection using 50 mbar (A) 6 s; (B) 23 s; (C) 56 s; (D) 112 s and (E) 170 s corresponding to injection volumes of 5.8; 22; 53.7; 107.4 and 162.9 nL, respectively. CZE conditions: fused silica capillary (90 cm × 50 μm id.); BGE: 1 M formic acid (pH 1.9); U: 21 kV. MS acquisition parameters: positive ionization mode; nebulizer pressure: 0.5 bar; dry gas temperature: 200 °C; dry gas flow rate: 4 L/min; capillary voltage: 4500 V; end plate offset: 500 V; MS spectra rate: 3 Hz; MS/MS spectra rate: 1–4 Hz; mass range: 50–2200 m/z . Sheath liquid: IPA:water = 1:1 + 0.1 % formic acid; flow rate: 4 μL/min. (F) The amino acid sequences of the peptides denoted with labels: 1: SEVAHR; 2: TPVSDR and 3: SHCIAEVENDEMPADLPSLAADFVESK and their physico-chemical properties. Top right corner: sequence coverage values; number of peptide-spectrum-matches and identified peptides with increasing sample loading ($n = 3$). Electropherograms are set to the same scale.

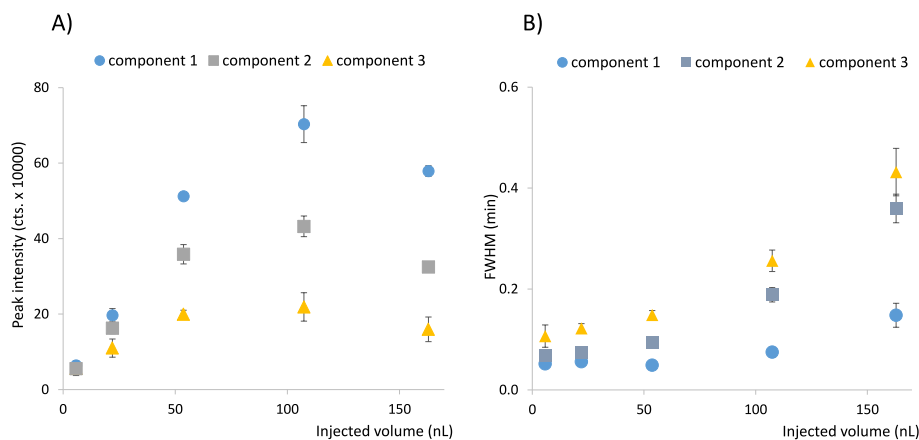


Fig. 2. The dependence of (A) peak intensity and (B) peak width (FWHM) on the volume injected in the case of the 3 selected components shown in Fig.1 ($n = 3$).

adjustments needed to be made either to the BGE or the sample matrix. These adjustments were specific to the type of stacking to be accentuated. Generally, in FASS-type of stacking, the conductivity difference of the BGE and sample solution can be modified (e.g., higher percentage of ACN in the sample), if the optimization of t-ITP is the objective, either leading or terminating ions can be added to the BGE/sample in order to prolong the ITP state.

The TE in our setup was suspected to be formic acid, which was actually the BGE. Although H^+ ions have an extremely high mobility in water, they can be slowed down by the buffering counterions of the LE [44]. In our case, the LE counterion was HCO_3^- , which in theory would migrate towards to anode, retarding the movement of H^+ ions. However, it is important to note that the pH of the sample was around 4.1, which is well below the pKa of carbonic acid, which is around 6.3. This means

that in these conditions the bicarbonate ions tend to protonate and approximately only 0.5 % of molecules is in an ionized state. Therefore, its role as counterion is questionable. On the other hand, because we usually terminate the enzymatic digestion with the addition of formic acid, the formate ions present in the sample could participate as buffering counterions. The pKa of formic acid is ~ 3.7 so presumably at around pH 4.1 there was considerable amount of formate ions that could neutralize the H^+ . The BGE concentration was varied in the 0.1–1 M range (Fig. S-2). In order to avoid excessive Joule-heating and consequent zone-dispersion, formic acid concentration was not raised above 1 M. Decreasing formic acid concentration clearly had a deleterious effect on overall peak profile, peptides seem to have blended into one another and peak intensities were progressively lower (slightly) as the BGE concentration decreased. The reason for this is most probably a

combination of factors. Firstly, peptides tend to adsorb to the capillary surface, which is known to be more pronounced when low ionic strength BGEs are applied. The increased pH might also be responsible for the higher tendency of peptide adsorption. Furthermore, with the decrease of TE concentration, ITP performance can be weaker, as well. Since peptides peaks were generally wider with decreasing BGE concentration, a given component got selected for fragmentation multiple times, which explains the gradual increase in peptide-spectrum matches (#PSM). Therefore, there was a redundancy of identical MS/MS spectra for higher-abundance peptides. Interestingly, SC% values did not show the decline one would expect from the appearance of the electropherograms. Nevertheless, for further experiments, 1 M formic acid was continued to be used.

Increasing the LE concentration was also investigated (Fig. 3). The addition of extra AB to the digest extended the duration of the t-ITP state, which allowed more time for the peptides to “pile up” in the stacking zone. This had a favourable impact on peak heights as well as resolution. The advantage of extra AB was more evident if the sample was highly diluted (Fig. 3C and D). It is important to note that the addition of AB increased not only the concentration of the leading ion (NH_4^+) but also the pH of the sample. From the original pH = 4.1 environment, the pH shifted above 7, which suggests the possibility of a dynamic pH-junction phenomenon, as well. Presumably, the high stacking efficiency can be attributed to the joint effect of t-ITP and DPJ.

3.2. Comparison of peptide patterns obtained from non-desalted and desalted digests

Having confirmed the possibility to utilize preconcentration in samples with absolutely no post-digestion sample treatment, desalted samples were also analyzed for comparison. The elution of adsorbed peptides from the SPE packing was carried out with an ACN-water mixture. This sample matrix had a significantly lower conductivity than the BGE containing 1 M formic acid, thus being directly amenable for FASS-type preconcentration, theoretically. In addition, the effect of adding ACN to non-desalted samples was also studied (Fig. 4. A). Peak intensities are lower in the case of the non-desalted, original digest; desalted and ACN-containing non-desalted digests both yielded increased intensities (by a factor of ~2). At this point it was clear that FASS and t-ITP were the main mechanism responsible for the stacking

phenomena in desalted and non-desalted original digests, respectively. The incorporation of ACN into the non-desalted sample could induce a so-called pseudo ITP state [45]. Since the digest contained a large amount of AB in this case, as well, the leading ion remains NH_4^+ . However, the role of the terminating ion was taken over by the organic solvent. ACN basically served the same purpose of creating a low-conductivity zone, where the movement of analyte ions was accelerated due to the high field-strength.

Apart from peak intensities, slight differences arose also in peak profiles. The peptide map of the desalted digest lacks a couple of peaks compared to those of non-desalted digests. Although SPE is inevitable in some cases, there are often concerns over peptide loss as small, strongly hydrophilic peptides are hardly retained, while long hydrophobic peptides may adsorb too firmly. Indeed, some of the peptides not identified in the case of desalted digests were small, hydrophilic peptides, as indicated by the Grand Average of Hydropathy (GRAVY) score (Fig. 4. B). It is worth noting, however, that the peptides not identified in desalted digests were present only in a relatively low concentration in the non-desalted digest and were not (that) critical for maximizing sequence coverage (SC%) since some of these peptides contained a missed cleavage close to the actual cleavage site (e.g.: ETCFAEEGKK or RMPCAEDYLSVVLNQLCVLHEK; missed cleavages indicated in bold) and were identified in their fully hydrolyzed form (e.g. MPCAEDYLSVVLNQLCVLHEK) in the desalted digests, as well. Furthermore, these “missing” peptides were not all completely lost as a result of SPE, only their full recovery was hindered. Extracted ion electropherograms (EIEs) of the peptides in question were generated, which indicate the decrease of intensities in the case of desalted digests (Fig. S-3.). Reduced intensities led either to their exclusion from MS/MS events or poor, uninformative fragment mass spectra.

The results presented are not the best representation if the goal is to compare purely the efficiency of the different preconcentration strategies (FASS vs t-ITP) since desalting alters not only the sample matrix but also the composition of the peptide mixture. In order to explore differences is stacking efficiencies, samples were desalted, lyophilized and reconstituted in solvents/solvent mixtures that correspond to their original composition: (I) 18 mM AB solution + 0.1 (v/v)% formic acid, (II) ACN:water = 1:1 containing 18 mM AB and (III) ACN:water = 7:3 containing 0.1 (v/v)% formic acid (elution solvent mixture during desalting) (Fig. 4. C). The tendency is similar to that of Fig. 4. A, the

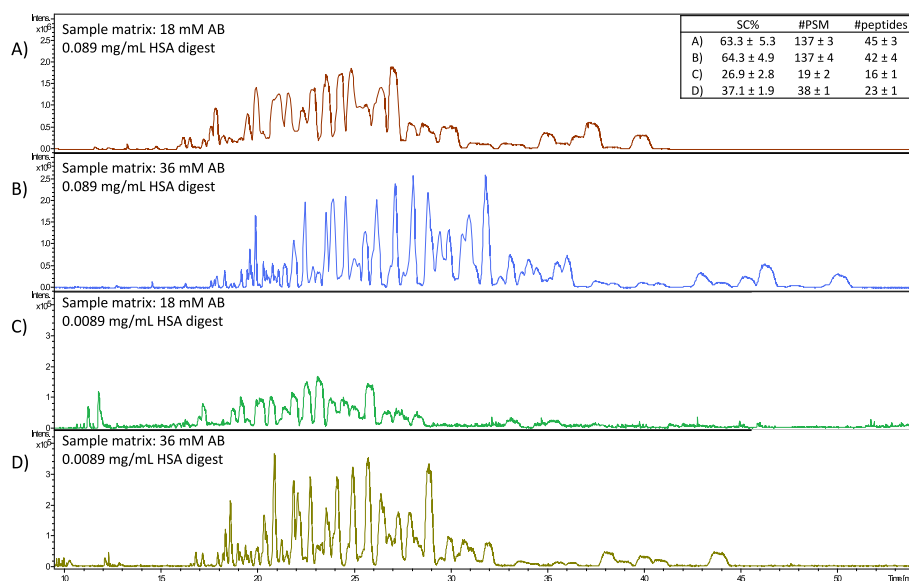


Fig. 3. Peptide maps of HSA digests (A, C) containing 18 mM AB and (B, D) containing 36 mM AB. Sample introduction: 50 mbar × 170 s (injected volume: 162.9 nL). Peptide concentration of HSA digests: (A, B) 0.089 mg/mL and (C, D) 0.0089 mg/mL. All other parameters are same as in Fig. 1. Electropherograms A-B and C-D are set to the same scale. SC%, number of PSMs and identified peptides for each case is indicated in the top right corner (n = 3).

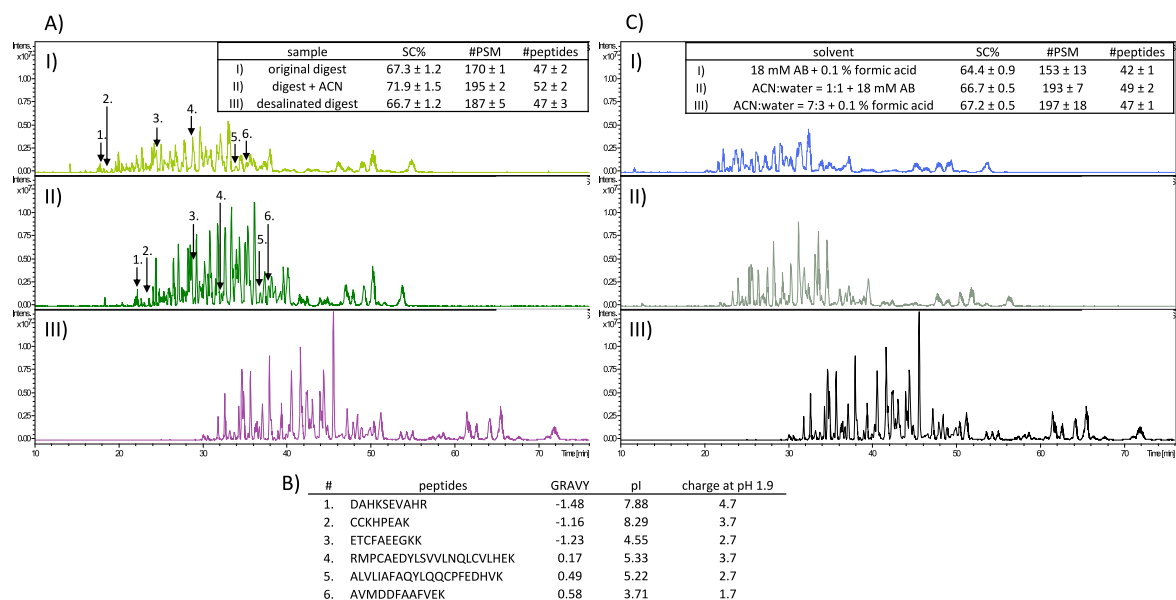


Fig. 4. CZE-MS analyses of HSA digests in different sample backgrounds. (A) Comparison of the peptide maps obtained by analyzing non-desalted samples having I) original matrix (digestion milieu) and II) added ACN (50 %) with III) desalted digests. (B) Peptides (1–6) not identified in the case of desalted digests. These peptides are labelled in the electropherograms corresponding to the non-desalted digests. (C) Comparison a peptide maps obtained by analyzing desalted, lyophilized HSA digests reconstituted in I) 18 mM NH_4HCO_3 + 0.1 (v/v) % formic acid, II) ACN:water = 1:1 + 18 mM NH_4HCO_3 and III) ACN:water = 7:3 + 0.1 (v/v) % formic acid. Sample introduction: 50 mbar, 112 s; all other CZE and MS parameters are the same as in Fig. 1. Electropherograms are set to the same scale. The concentration of HSA digest was the same in all cases ($c = 0.45$ mg/mL). All analyses were carried out in triplicate.

electropherograms obtained for samples containing a generous amount of ACN suggest higher stacking efficiencies (increased peak intensities, narrower peaks). Compared to traditional t-ITP, pseudo ITP allows a more relaxed approach, since it has no strict requirements on pH, concentration and counter-ions [46]. This is logical, considering the inherent low conductivity of ACN, whereas the mobility of H^+ (and hence, conductivity) is largely dependent on buffering.

In order to get deeper insights into sample matrix behaviour in cases when more proteins are included in the sample, a standard protein mixture was analyzed that contains albumin, myoglobin, β -casein and lysozyme. Again, three different types of samples were analyzed: (I) original digest, (II) digest containing 50 % ACN and (III) desalted digest. Judging by the base peak electropherograms, intensities doubled for digests containing ACN (Fig. S-4) but SC% values were only moderately higher (if at all) (Tables S-3). Altogether 69 common peptides were identified, whose intensities and FWHM data were investigated in the three sample types (Fig. 5). Relative values were illustrated, taking the original digest as a reference point. Generally, the samples including ACN showed increased intensities, with an average gain of 6 % and 4 % for non-desalted and desalted samples, respectively. In some cases a very sharp increase (or decrease) could be seen in peak intensities, the list of these few peptides and their physico-chemical properties can be found in the supplementary material (Table S-4, Table S-5). Peptides showing increased intensities encompass a diverse range of physico-chemical properties so there seems to be no clear bias towards specific peptide characteristics. Interestingly, there are some peptides whose intensities are decreased in the case of both samples containing ACN. This is logical for desalted digests, such decrease in intensities can be attributed to incomplete peptide recovery during the SPE process. As for peak widths, interestingly, the non-desalted digest did not show a consistent decrease relative to the original digest (in fact, an average increase by 2.5 % was found). On the other hand, the desalted digest undoubtedly offered narrower signal, demonstrating a 25 % average decrease in peak widths.

3.3. Analysis of yeast cell lysate

To demonstrate tangible results on the “performance” of the different

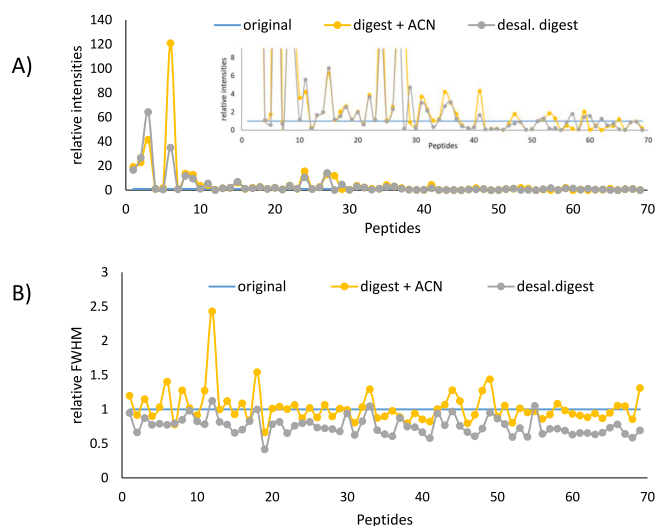


Fig. 5. The effect of sample matrix on (A) peak intensity and (B) FWHM values. Comparison of original digest, the digest containing ACN and desalted digest. The 69 shared peptides were numbered from 1 to 69, these peptide identifiers are on axis x. Relative intensities and FWHM values were used for illustrating the differences compared to the original digest. In part A) inset shows a zoom in on relative intensities. All experimental conditions are same as in Fig. 4.

post-digestion sample treatment approaches, a yeast cell lysate was also analyzed and the number of identified proteins was compared. In this case, four sample types were investigated: (1) original digest, (2) digest containing extra AB (+18 mM), (3) digest containing 50 % ACN and (4) desalted digest (Fig. 6, Fig. S-5); where mainly the following stacking mechanisms could be expected: (1) t-ITP, (2) combined t-ITP and DPJ, (3) pseudo t-ITP and (4) FASS. The conventional, low-volume injection ($V_{inj} \sim 6$ nL) yielded poor results (14 protein IDs), however, larger sample loadings utilizing (pseudo) t-ITP or FASS preconcentrations offer

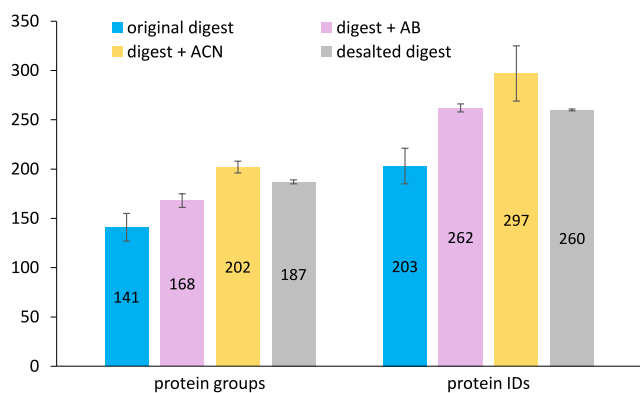


Fig. 6. Number of protein groups and protein identifications in a yeast cell lysate for the four sample types: (1) original digest, (2) digest with added ammonium bicarbonate (+18 mM), (3) digest with added ACN (50 %) and (4) desalted digest. Peptide concentration was 0.36 mg/mL in all cases. Measurements performed in triplicate. Sample introduction: 50 mbar, 112 s, all other CZE and MS parameters are same as in Fig. 1.

a ~15-fold increase in the number of identified proteins. Depending on what was used for diluting the non-desalted digest, increasing numbers of proteins could be identified. Sample #3 yielded the best results, compared to sample #1 (where the stacking is based mainly on t-ITP) ~40 % increase could be observed regarding identified protein groups. Desalted digests proved to be better in this case than non-desalted digests containing no ACN, peptide loss was not significant.

Altogether 175 proteins are shared by all four samples. The effect of sample matrix on the sequence coverages was also investigated (Fig. 7). As could be expected, samples #2–4 outperformed sample #1, especially the samples containing a generous amount of ACN. SC% values tend to be higher, in some cases even by a factor of 6. Sequence coverage values for the 175 shared proteins were in a relatively good alignment when desalted and ACN-containing non-desalted digests were

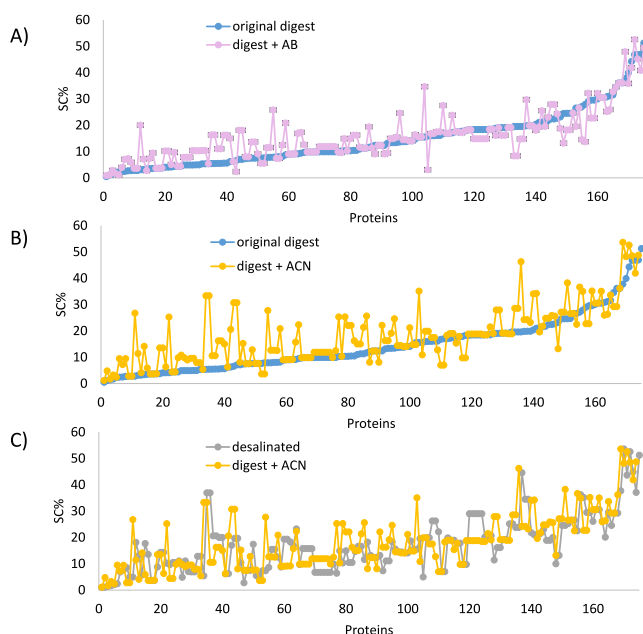


Fig. 7. The effect of sample matrix on SC% values in the case of digested yeast cell lysate. Comparison of (A) original digest and the digest containing extra AB, (B) original digest and the digest containing ACN and (C) desalted digest and digest containing ACN. The 175 shared proteins were numbered from 1 to 175, these protein identifiers are on axis x ($n = 3$).

compared, which suggests that ACN was a key ingredient in achieving efficient stacking regardless of whether sample went through SPE or not. The commercially available yeast protein extract is actually a purified standard often used for method development. To test whether our observations thus far had any validity in samples with biological matrices, human tear samples undergoing no pre-digestion purification were also analyzed. Although the overall depth of characterization was rather low, the tendencies coincided with those obtained for standard proteins and the cells (Tables S–6).

The latest version of the Swiss-Prot database contains 6727 curated and reviewed entries for *Saccharomyces cerevisiae* (viewed: 2023.12.08). With our CZE-MS/MS only a fraction of the whole yeast proteome could be identified and the results certainly lag behind those reported previously in both RPLC-MS [47–49] and CE-MS [36,50] studies. CE-MS allowed the identification of ~1500–3500 proteins depending on whether prefractionation was performed or not [36,50], while with RPLC-MS systems up to 4000 protein identifications could be achieved [49]. All these works emphasize the importance of long analytical runs for improved peptide separation. The migration window (~75 min) for the yeast digest in our study is comparable to those found in the CE-MS literature (~90 min). The substandard number of protein IDs in the present work can, therefore, be mainly attributed to the lower efficiency of (i) ion transfer from solution into gas phase and (ii) MS/MS sampling. We believe that our greatest setback in achieving a more comprehensive proteome characterization is that lack of a nanoESI source. In addition, the studies mentioned above all used Orbitrap MS systems, which have been shown to outperform QTOF analyzers [51].

4. Conclusions

The fundamental concept of this work was to investigate the importance of desalting. Sample clean-up prior to analysis is generally considered to be a necessity, since the presence of salts or other contaminants in the sample can interfere with both the separation and the MS detection.

In the majority of cases SPE desalting is utilized, mostly to eliminate the large amount of chaotropic agents; other additives (e.g., DTT, IAA) are usually present only in a low concentration, therefore their interference with the analysis is not suspected to be overwhelming. About a quarter of the papers studied (Tables S–1) excluded sample clean-up before analysis. Most of these cases utilize Rapigest™ (an anionic surfactant designed to be MS-compatible) instead of urea (or SDS) for the solubilization and denaturation of proteins. It can be cleaved by the addition of acid, therefore subsequent matrix removal is not necessary [2,29,30,35,37,38]. Only very few examples can be found that show similarity to our present work in the sense that no desalting was applied despite the use of urea [13,28].

Desalted samples were also analyzed and results were compared to those acquired from non-desalted digests. During the C18 ZipTip procedure, the adsorbed peptides were eluted with ACN-water mixture, the conductivity of which is considerably lower than that of the BGE, generating a FASS-type preconcentration at the beginning of the CZE run. The enrichment efficiencies of FASS (desalted digest) and t-ITP (non-desalted digest) were comparable, however, desalting led to peptide loss due to the negligible retention of strongly hydrophilic peptides on the SPE packing. Preconcentration efficiency could be further enhanced by adding ACN to the non-desalted digests. The ACN-salt mixtures have already been shown to promote the t-ITP effect, with ACN acting as terminating electrolyte [45].

Based on these results, we propose that sample clean-up in the bottom-up sample preparation process prior to CZE-MS analysis can be omitted. In addition to being an extra, time-consuming procedure, desalting offered no added considerable benefits in terms of detection sensitivity, separation efficiency or sequence coverage and actually led to slightly poorer results when compared to non-desalted samples. Although we propose the exclusion of desalting for routine applications,

in cases where deep bottom-up proteome profiling is required, resorting to prefractionation might be necessary, which inevitably entails the removal of salts before the actual CE analysis.

CRedit authorship contribution statement

Cynthia Nagy: Investigation, Methodology, Writing – original draft, Writing – review & editing. **Melinda Andradi:** Investigation, Methodology. **Ruben Szabo:** Data curation, Investigation. **Attila Gaspar:** Conceptualization, Methodology, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors acknowledge the financial support provided to this project by the National Research, Development and Innovation Office, Hungary (K142134) and the New National Excellence Program of the Ministry for Innovation and Technology (ÚNKP-22-3-II). The technical assistance of Nikolett Boros is highly appreciated.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.aca.2023.342162>.

References

- N. Said, R. Gahoual, L. Kuhn, A. Beck, Y.N. François, E. Leize-Wagner, Structural characterization of antibody drug conjugate by a combination of intact, middle-up and bottom-up techniques using sheathless capillary electrophoresis – tandem mass spectrometry as nanoESI infusion platform and separation, *Method. Anal. Chim. Acta* 918 (2016) 50–59.
- J. Giorgetti, A. Beck, E. Leize-Wagner, Y.N. François, Combination of intact, middle-up and bottom-up levels to characterize 7 therapeutic monoclonal antibodies by capillary electrophoresis – mass spectrometry, *J. Pharm. Biomed. Anal.* 182 (2020), 113107.
- M.P. Washburn, D. Wolters, J.R. Yates, Large-scale analysis of the yeast proteome by multidimensional protein identification technology, *Nat. Biotechnol.* 19 (2001) 242–247.
- K.A. Cupp-Sutton, M. Fang, S. Wu, Separation methods in single-cell proteomics: RPLC or CE? *Int. J. Mass Spectrom.* 481 (2022), 116920.
- V. Kasička, Peptide mapping of proteins by capillary electromigration methods, *J. Separ. Sci.* 45 (2022) 4245–4279.
- K. Faserl, B. Sarg, L. Kremser, H. Lindner, Optimization and evaluation of a sheathless capillary electrophoresis–electrospray ionization mass spectrometry platform for peptide analysis: comparison to liquid chromatography–electrospray ionization mass spectrometry, *Anal. Chem.* 83 (2011) 7297–7305.
- Y. Wang, B.R. Fonslow, C.C. Wong, A. Nakorchevsky, J.R. Yates, Improving the comprehensiveness and sensitivity of sheathless capillary electrophoresis–tandem mass spectrometry for proteomic analysis, *Anal. Chem.* 84 (2012) 8505–8513.
- G. Zhu, L. Sun, X. Yan, N.J. Dovichi, Single-shot proteomics using capillary zone electrophoresis–electrospray ionization–tandem mass spectrometry with production of more than 1250 *Escherichia coli* peptide identifications in a 50 min separation, *Anal. Chem.* 85 (2013) 2569–2573.
- K.R. Ludwig, L. Sun, G. Zhu, N.J. Dovichi, A.B. Hummon, Over 2300 phosphorylated peptide identifications with single-shot capillary zone electrophoresis–tandem mass spectrometry in a 100 min separation, *Anal. Chem.* 87 (2015) 9532–9537.
- L. Sun, G. Zhu, Z. Zhang, S. Mou, N.J. Dovichi, Third-generation electrokinetically pumped sheath-flow nanospray interface with improved stability and sensitivity for automated capillary zone electrophoresis–mass spectrometry analysis of complex proteome digests, *J. Proteome Res.* 14 (2015) 2312–2321.
- M. Ibrahim, R. Gahoual, L. Enkler, H.D. Becker, J. Chicher, P. Hammann, Y. N. François, L. Kuhn, E. Leize-Wagner, Improvement of mitochondria extract from *Saccharomyces cerevisiae* characterization in shotgun proteomics using sheathless capillary electrophoresis coupled to tandem mass spectrometry, *J. Chromatogr. Sci.* 54 (2016) 653–663.
- H. Chen, P. Shi, F. Fan, M. Tu, Z. Xu, X. Xu, M. Du, Complementation of UPLC-Q-TOF-MS and CEESI-Q-TOF-MS on identification and determination of peptides from bovine lactoferrin, *J. Chromatogr. B* 1084 (2018) 150–157.
- K.R. Johnson, M. Greguš, J.C. Kostas, A.R. Ivanov, Capillary electrophoresis coupled to electrospray ionization tandem mass spectrometry for ultra-sensitive proteomic analysis of limited samples, *Anal. Chem.* 94 (2022) 704–713.
- M.J. Gou, M.C. Kose, J. Crommen, C. Nix, G. Cobraiville, J. Caers, M. Fillet, Contribution of capillary zone electrophoresis hyphenated with drift tube ion mobility mass spectrometry as a complementary tool to microfluidic reversed phase liquid chromatography for antigen discovery, *Int. J. Mol. Sci.* 23 (2022), 13350.
- X. Yan, D.C. Essaka, L. Sun, G. Zhu, N.J. Dovichi, Bottom-up proteome analysis of *E. coli* using capillary zone electrophoresis–tandem mass spectrometry with an electrokinetic sheath-flow electrospray interface, *Proteomics* 13 (2013) 2546–2551.
- Z. Zhang, X. Yan, L. Sun, G. Zhu, N.J. Dovichi, Detachable strong cation exchange monolith, integrated with capillary zone electrophoresis and coupled with pH gradient elution, produces improved sensitivity and numbers of peptide identifications during bottom-up analysis of complex proteomes, *Anal. Chem.* 87 (2015) 4572–4577.
- Z. Zhang, T. Albanetti, T. Linkous, C.J. Larkin, R. Schoner, J.B. McGivney, N. J. Dovichi, Comprehensive analysis of host cell impurities in monoclonal antibodies with improved sensitivity by capillary zone electrophoresis mass spectrometry, *Electrophoresis* 38 (2016) 401–407.
- Z. Zhang, G. Zhu, E.H. Peuchen, N.J. Dovichi, Preparation of linear polyacrylamide coating and strong cationic exchange hybrid monolith in a single capillary, and its application as an automated platform for bottom-up proteomics by capillary electrophoresis–mass spectrometry, *Microchim. Acta* 184 (2017) 921–925.
- Z. Yang, X. Shen, D. Chen, L. Sun, Microscale reversed-phase liquid chromatography/capillary zone electrophoresis–tandem mass spectrometry for deep and highly sensitive bottom-up proteomics: identification of 7500 proteins with five micrograms of an MCF7 proteome digest, *Anal. Chem.* 90 (2018) 10479–10486.
- J.S. Mellors, W.A. Black, A.G. Chambers, J.A. Starkey, N.A. Lacher, J.M. Ramsey, Hybrid capillary/microfluidic system for comprehensive online liquid chromatography–capillary electrophoresis–electrospray ionization–mass spectrometry, *Anal. Chem.* 85 (2013) 4100–4106.
- D. Chen, X. Shen, L. Sun, Capillary zone electrophoresis–mass spectrometry with microliter-scale loading capacity, 140 min separation window and high peak capacity for bottom-up proteomics, *Analyst* 142 (2017) 2118–2127.
- D. Chen, X. Shen, L. Sun, Strong cation exchange–reversed phase liquid chromatography–capillary zone electrophoresis–tandem mass spectrometry platform with high peak capacity for deep bottom-up proteomics, *Anal. Chim. Acta* 1012 (2018) 1–9.
- Z. Yang, X. Shen, D. Chen, L. Sun, Improved nanoflow RPLC–CZE–MS/MS system with high peak capacity and sensitivity for nanogram bottom-up proteomics, *J. Proteome Res.* 18 (2019) 4046–4054.
- D. Chen, K.R. Ludwig, O.V. Krokhn, V. Spicer, Z. Yang, X. Shen, A.B. Hummon, L. Sun, Capillary zone electrophoresis–tandem mass spectrometry for large-scale phosphoproteomics with the production of over 11,000 phosphopeptides from the colon carcinoma HCT116 cell line, *Anal. Chem.* 91 (2019) 2201–2208.
- R. Kumar, R.L. Shah, A.S. Rathore, Harnessing the power of electrophoresis and chromatography: offline coupling of reverse phase liquid chromatography–capillary zone electrophoresis–tandem mass spectrometry for peptide mapping for monoclonal antibodies, *J. Chromatogr. A* 1620 (2020), 460954.
- R. Kumar, R.L. Shah, S. Ahmad, A.S. Rathore, 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* 42 (2021) 735–741.
- L. Sun, G. Zhu, S. Mou, Y. Zhao, M.M. Champion, N.J. Dovichi, Capillary zone electrophoresis–electrospray ionization–tandem mass spectrometry for quantitative parallel reaction monitoring of peptide abundance and single-shot proteomic analysis of a human cell line, *J. Chromatogr. A* 1359 (2014) 303–308.
- M. Larsson, E.S.M. Lutz, Transient isotachopheresis for sensitivity enhancement in capillary electrophoresis–mass spectrometry for peptide analysis, *Electrophoresis* 21 (2000) 2859–2865.
- J.M. Busnel, B. Schoenmaker, R. Ramautar, A. Carrasco-Pancorbo, C. Ratnayake, J. S. Feitelson, J.D. Chapman, A.M. Deelder, O.A. Mayboroda, High capacity capillary electrophoresis–electrospray ionization mass spectrometry: coupling a porous sheathless interface with transient–isotachopheresis, *Anal. Chem.* 82 (2010) 9476–9483.
- C. Lew, J.L. Gallegos-Perez, B. Fonslow, M. Lies, A. Guttman, Rapid level-3 characterization of therapeutic antibodies by capillary electrophoresis electrospray ionization mass spectrometry, *J. Chromatogr. Sci.* 53 (2015) 443–449.
- Z. Zhang, A.S. Hebert, M.S. Westphall, J.J. Coon, N.J. Dovichi, Single-shot capillary zone electrophoresis–tandem mass spectrometry produces over 4400 phosphopeptide identifications from a 220 ng sample, *J. Proteome Res.* 18 (2019) 3166–3173.
- J. Cheng, L. Wang, C.M. Rive, R.A. Holt, G.B. Morin, D.D. Chen, Complementary methods for de novo monoclonal antibody sequencing to achieve complete sequence coverage, *J. Proteome Res.* 19 (2020) 2700–2707.
- L. Wang, J. Cheng, J.E. McNutt, G.B. Morin, D.D. Chen, Dynamic pH barrage junction focusing of amino acids, peptides, and digested monoclonal antibodies in

- capillary electrophoresis–mass spectrometry, *Electrophoresis* 41 (2020) 1832–1842.
- [34] J.P. Quirino, On-line electrophoretic, electrochromatographic, and chromatographic sample concentration in CE-MS, in: *Capillary Electrophoresis–Mass Spectrometry (CE-MS): Principles and Applications*, first ed., John Wiley & Sons, 2016, pp. 103–127.
- [35] C. Lombard-Banek, S.A. Moody, M.C. Manzini, P. Nemes, Microsampling capillary electrophoresis mass spectrometry enables single-cell proteomics in complex tissues: developing cell clones in live *Xenopus laevis* and zebrafish embryos, *Anal. Chem.* 91 (2019) 4797–4805.
- [36] L. Sun, A.S. Hebert, X. Yan, Y. Zhao, M.S. Westphall, M.J. Rush, G. Zhu, M. M. Champion, J.J. Coon, N.J. Dovichi, Over 10000 peptide identifications from the HeLa proteome by using single-shot capillary zone electrophoresis combined with tandem mass spectrometry, *Angew. Chem. Int. Ed.* 53 (2014) 13931–13933.
- [37] A.A. Heemskerck, J.M. Busnel, B. Schoenmaker, R.J. Derks, O. Klychnikov, P. J. Hensbergen, A.M. Deelder, O.A. Mayboroda, Ultra-low flow electrospray ionization-mass spectrometry for improved ionization efficiency in phosphoproteomics, *Anal. Chem.* 84 (2012) 4552–4559.
- [38] R. Gahoual, J.M. Busnel, A. Beck, Y.N. François, E. Leize-Wagner, Full antibody primary structure and microvariant characterization in a single injection using transient isotachopheresis and sheathless capillary electrophoresis–tandem mass spectrometry, *Anal. Chem.* 86 (2014) 9074–9081.
- [39] S. Garza, M. Moini, Analysis of complex protein mixtures with improved sequence coverage using (CE–MS/MS)ⁿ, *Anal. Chem.* 78 (2006) 7309–7316.
- [40] M.J. Gou, G. Nys, G. Cobraiville, A. Demelonne, A.C. Servais, M. Fillet, Hyphenation of capillary zone electrophoresis with mass spectrometry for proteomic analysis: optimization and comparison of two coupling interfaces, *J. Chromatogr. A* 1618 (2020), 460873.
- [41] Z. Zhang, L. Sun, G. Zhu, O.F. Cox, P.W. Huber, N.J. Dovichi, Nearly 1000 protein identifications from 50 ng of *Xenopus laevis* zygote homogenate using online sample preparation on a strong cation exchange monolith based microreactor coupled with capillary zone electrophoresis, *Anal. Chem.* 88 (2015) 877–882.
- [42] G. Zhu, L. Sun, X. Yan, N.J. Dovichi, Bottom-up proteomics of *Escherichia coli* using dynamic pH junction preconcentration and capillary zone electrophoresis–electrospray ionization–tandem mass spectrometry, *Anal. Chem.* 86 (2014) 6331–6336.
- [43] M. Naldi, M. Baldassarre, M. Domenicali, F.A. Giannone, M. Bossi, J. Montomoli, T. D. Sandahl, E. Glavind, H. Vilstrup, P. Caraceni, C. Bertucci, Mass spectrometry characterization of circulating human serumalbumin microheterogeneity in patients with alcoholic hepatitis, *J. Pharm. Biomed. Anal.* 122 (2016) 141–147.
- [44] T. Melzer, B. Wimmer, S. Bock, T.N. Posch, C. Huhn, Challenges and applications of isotachopheresis coupled to mass spectrometry: a review, *Electrophoresis* 41 (2020) 1045–1059.
- [45] Z.K. Shibabi, Transient pseudo-isotachopheresis for sample concentration in capillary electrophoresis, *Electrophoresis* 23 (2002) 1612–1617.
- [46] Z.K. Shibabi, Stacking in capillary zone electrophoresis, *J. Chromatogr. A* 902 (2000) 107–117.
- [47] N. Selevsek, C.-Y. Chang, L.C. Gillet, P. Navarro, O.M. Bernhardt, L. Reiter, L.-Y. Cheng, O. Vitek, R. Aebersold, Reproducible and consistent quantification of the *Saccharomyces cerevisiae* proteome by SWATH-mass spectrometry, *Mol. Cell. Proteomics* 14 (2015) 739–749.
- [48] S.S. Thakur, T. Geiger, B. Chatterjee, P. Bandilla, F. Fröhlich, J. Cox, M. Mann, Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation, *Mol. Cell. Proteomics* 10 (2011), M110.003699.
- [49] A.L. Richards, A.S. Hebert, A. Ulbrich, D.J. Bailey, E.E. Coughlin, M.S. Westphall, J. J. Coon, One hour proteome analysis in yeast, *Nat. Protoc.* 10 (2015) 701–714.
- [50] K. Faserl, L. Kremser, M. Müller, D. Teis, H.H. Lindner, Quantitative proteomics using ultralow flow capillary electrophoresis–mass spectrometry, *Anal. Chem.* 87 (2015) 4633–4640.
- [51] P.J. Reddy, S. Ray, G.J. Sathe, A. Gajbhiye, T.S.K. Prasad, S. Rapole, D. Panda, S. Srivastava, A comprehensive proteomic analysis of totarol induced alterations in *Bacillus subtilis* by multipronged quantitative proteomics, *J. Proteomics* 114 (2015) 247–262.