Capillary liquid chromatography coupled with mass spectrometry for analysis of nanogram protein quantities on a wide-pore superficially porous particle column in top-down proteomics

Laurent Lignieres, Véronique Legros, Manel Khelil, Nicolas Senecaut, Matthew A Lauber, Jean-Michel Camadro, Guillaume Chevreux

Journal of Chromatography B 1214 (2023) 123566

Contents lists available at ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/jchromb

Capillary liquid chromatography coupled with mass spectrometry for analysis of nanogram protein quantities on a wide-pore superficially porous particle column in top-down proteomics

Laurent Lignieres [a,1], Véronique Legros [a,1], Manel Khelil [a], Nicolas Senecaut [a], Matthew A Lauber [b], Jean-Michel Camadro [a], Guillaume Chevreux [a,*]

[a] Université Paris Cité, CNRS, Institut Jacques Monod, F-75013 Paris, France
[b] Waters Corporation, 34, Maple Street, Milford, MA 01757-3696, United States

ARTICLE INFO

Keywords:
Top-down
Wide-pore
Superficially porous
Histones, monoclonal antibodies, ribosomes

ABSTRACT

In top-down proteomics experiments, intact protein ions are subjected to gas-phase fragmentation for MS analysis without prior digestion. This approach is used to characterize post-translational modifications and clipped forms of proteins, avoids several "inference" problems associated with bottom-up proteomics, and is well suited to the study of proteoforms. In the past decade, top-down proteomics has progressed rapidly, taking advantage of MS instrumentation improvements and the efforts of pioneering groups working to improve sample handling and data processing. The potential of this technology has been established through its successful use in a number of important biological studies. However, many challenges remain to be addressed like improving protein separation capabilities such that it might become possible to expand the dynamic range of whole proteome analysis, address co-elution and convoluted mass spectral data, and aid final data processing from peak identification to quantification. In this study, we investigated the use of a wide-pore silica-based superficially porous media with a high coverage phenyl bonding, commercially packed into customized capillary columns for the purpose of top-down proteomics. Protein samples of increasing complexity were tested, namely subunit digests of a monoclonal antibody, components of purified histones and proteins extracted from eukaryotic ribosomes. High quality mass spectra were obtained from only 100 ng of protein sample while using difluoroacetic acid as an ion pairing agent to improve peak shape and chromatographic resolution. A peak width at half height of about 15 s for a 45 min gradient time was observed on a complex mixture giving an estimated peak capacity close to 100. Most importantly, efficient separations were obtained for highly diverse proteins and there was no need to make method specific adjustments, suggesting this is a highly versatile and easy-to-use setup for top-down proteomics.

1. Introduction

Proteins are complex molecules that play many critical roles in the body. They represent the third component of the central dogma of molecular biology and directly exhibit the potential function of genes while doing most of the work required for a living organism *via* a wide array of functions. Among others, they can serve as catalytic enzymes, messengers, structural components or be involved in transport and storage. The proteome is the entire set of proteins that is produced or modified by living organisms, and proteomics aims to decipher a picture of a proteome at a given time in a given condition. Usually, the proteome is investigated differentially in order to elucidate key mechanisms and/or biological pathways explaining specific states or conditions of a living organism at the protein level.

Abbreviations: MS, Mass Spectrometry; TD, Top-Down; BU, Bottom-Up; PTMs, Post-Translational Modifications; ID, Internal Diameter; DOE, Design Of Experiment; RPLC, Reversed Phase Liquid Chromatography; FA, Formic Acid; DFA, DiFluoroacetic Acid; XIC, Extracted Ion Chromatogram; RSD, Relative Standard Deviation; HCD, Higher-energy C-trap Dissociation; FWHM, Full Width at Half Maximum; RES, Resolution; LC, Light Chain; HC, Heavy Chain; AGC, Automatic Gain Control; t_R, Retention Time; mAb, monoclonal Antibody.

* Corresponding author.
E-mail address: guillaume.chevreux@ijm.fr (G. Chevreux).

[1] These authors contributed equally.

https://doi.org/10.1016/j.jchromb.2022.123566
Received 6 October 2022; Received in revised form 1 December 2022; Accepted 2 December 2022
Available online 8 December 2022

Bottom-Up (BU) proteomics is based on a proteolytic digestion of proteins from a complex mixture prior to analysis by mass spectrometry (MS). Peptides resulting from the digestion are separated by liquid chromatography or capillary electromigration [1,2] coupled to MS for further sequencing and identification *via* a protein database search. Furthermore, intensities of the identified peptides can be used as a surrogate of the protein abundance in solution in order to compare protein expression of large biological systems [3,4]. This technology is now well established and has benefitted from dramatic improvement in MS instrumentation (resolution, mass accuracy, sensitivity and acquisition frequency). However, the beforehand proteolytic digestion results in a massive loss of information and precludes the analysis of a large diversity of proteoforms arising from combinatorial post-translational modifications (PTMs), splicing mechanisms, polymorphisms, protease degradation [5]. In this context, Top-Down (TD) proteomics has emerged as an alternative approach capable of identifying and quantifying these proteoforms. Proteins are analyzed without prior digestion and identifications rely on MS/MS spectra obtained by fragmentation of intact proteins in the mass spectrometer [6,7]. Although attractive and promising, this methodology remains difficult to implement due to three distinct challenges. First, the processing of TD data raises several specific issues. For instance, the deconvolution of monoisotopic *versus* average intact masses, the estimation of false discovery rates on identifications and the scoring of proteoform spectral matches as well as the multiplicity of variable modifications that translate into highly demanding computing resources. The field is very active and numerous software applications have been developed to provide reliable tools to the research community [8–13]. Secondly, TD proteomics inherently requires MS instrumentation that is able to resolve the isotopic patterns of intact proteins. This is needed in order to enhance the deconvolution of the large number of overlapping multicharged fragments that can originate from proteins. Fragmentation technologies orthogonal to collision induced dissociation (CID) are also essential, because they can increase sequence coverage and facilitate the analysis of labile PTMs. As a result, TD proteomics investigators continue to explore emerging technologies like ExD, IRMPD, UVPD [14–16]. Finally, the last challenge to overcome is achieving a highly effective intact protein separation at the front end of the analysis, which is directly tied to the observable dynamic range of protein expression and proteoforms. In particular, a more effective separation helps expand the intrascan dynamic range by reducing the spectral overlap of co-eluting proteins (and their complex charge state envelopes). Chromatographic separation of proteins in reversed phase liquid chromatography (RPLC) is not yet commonplace in the proteomics field. Nevertheless, there have been a few studies involving use of home-made columns. For several examples of work, investigators have found success with nano and capillary columns packed with 1000 Å pore size poly(styrene/divinylbenzene) particles (PLRP-S media) [17–19]. An alternative type of column was more recently proposed by Liang et al., which was a bridged hybrid monolithic column based on a bis(triethoxysilyl)ethylene matrix functionalized with C_8 groups [20]. This novel home-made column performed well in terms of protein separations and repeatability while exhibiting low backpressures. Wu et al. [21] demonstrated unprecedented separation efficiency of intact proteins using a prototype, pulled tip capillary densely packed with 0.47 μm silica particles. Ultimately, it should be noted that these studies have exclusively come from expert laboratories dedicated to the development of TD proteomics. It is our believe that the field of TD proteomics is held back by the lack of commercial, ready-to-use columns and robust operating.

In 2018, Bobaly *et al.* [22] described a new wide-pore silica-based superficially porous media with a high coverage phenyl bonding, 2.7 μm particles and an average pore size of 450 Å [23]. This media was proposed for 2.1- or 4.6-mm columns internal diameter (ID) and commercialized as BioResolve RP mAb Polyphenyl stationary phase for the reversed-phase analysis of intact or sub-unit digested monoclonal antibodies (mAbs) and antibody-drug conjugates (ADCs) in a pharmaceutical context. Excellent protein recovery, peak capacity as well as interesting selectivity was reported in those applications. Interestingly, custom made capillary columns (I.D. × L 0.3 mm × 150 mm) can be commercially ordered with this type of media. The aim of this study was to investigate if this type of media and column format can be a suitable alternative for protein separation in TD-proteomics applications. The first aspect of this work was to optimize protein separations for the best compromise between resolution, peak capacity and sensitivity within a 45-min long gradient. A second stage of the work was to explore the usability of this media for non-mAbs protein separation to expand its potential use for TD-proteomics application. First, the superior performance of this media to resolve hydrophobic compounds was demonstrated in comparison to a traditional C4 column operating under the same conditions. Then difluoroacetic acid (DFA) was investigated as an alternative to formic acid (FA) as a pairing ion agent to improve chromatographic performance without significant compromise to MS sensitivity. With this, we explored whether DFA would behave any differently under microflow electrospray conditions *versus* standard flow LC-MS [24,25]. Subsequently, intervariable dynamics and performance of the LC separation was investigated using a protein test mixture and a Design Of Experiment (DOE) approach that took into account the mobile phase flow rate (*i.e.,* flow rate), gradient slope and the temperature. A set of operating parameters was selected for further analysis of real-life samples of increasing complexity in order to appraise the potential of this capillary column chromatography for TD-proteomics. Results obtained from a purified IgG1, a bovine histone mixture and a protein extract from eukaryotic ribosomes are provided and discussed in terms of chromatographic separation and value from an MS perspective.

2. Material and methods

2.1. Material

Acetonitrile (A955-1), ammonium acetate (73594), guanidine-HCl (G4505), FA (A117-50), DFA (00922), histone from calf thymus (H9250) and SIGMAFAST™ Protease Inhibitor Tablets (S8820), magnesium acetate tetrahydrate (M5661), potassium acetate (P1190), HEPES buffer (54457), ammonium sulfate (A4418),) were purchased from Sigma Aldrich (St. Louis, MO, USA). Intact Protein Standard Mix (A33526) and TCEP-HCl (20490) were from ThermoFisher (Waltham, Massachusetts, USA). Difco Yeast Nitrogen Base (YNB) was from BD (Franklin Lakes, New Jersey, USA). NIST IgG1 mAb standard (RM 8671) was purchased from National Institute of Standards and Technology (NIST, Gaithersburg, MD, USA). Endoproteases FabRICATOR (IdeS) and FabALACTICA (IgdE) were obtained from Genovis (Lund, Sweden). All aqueous solutions were prepared using LC-MS grade water from ThermoFisher (Waltham, Massachusetts, USA).

2.2. Standard solubilization and dilution

The Intact Protein Standard Mix is a lyophilized mixture of 6 recombinant proteins (human IGF-1 LR3, human thioredoxin, *S. dysgalactiae* Protein G, Bovine Carbonic Anhydrase II, *Streptococcus* Protein AG, and *E. coli Exo* Klenow). For LC-MS experiments, the standard was dissolved in 200 μL of LC-MS grade water with 0.1 % FA to a final concentration of 0.4 μg/μL and aliquoted. Before analysis, aliquots were further diluted 4 times before injection of 1 μL corresponding to 100 ng of total protein material. Histones from calf thymus were supplied as a lyophilized powder that was resuspended in 50 mM pH 6.8 ammonium acetate buffer to reach a final concentration of 1 mg/mL. Aliquots were further diluted 10 times in mobile phase A before injection of a 1 μL volume corresponding to 100 ng of total protein material.

2.3. NIST mAb preparation

Enzymatic proteolysis with IdeS was achieved by a twofold dilution

of NIST mAb in a phosphate saline buffer 1X pH 7.4 followed by an incubation during 1 h at 37 °C with 1 unit IdeS protease per 1 μg of mAb. For IgdE proteolysis, NIST mAb was twofold diluted using 100 mM phosphate buffer pH 7.0 and digested by IgdE protease overnight at 37 °C with 1 unit per 1 μg of mAb. Samples were evaporated until dry by using a vacuum concentrator, then resuspended in Guanidine-HCl 6 M, 50 mM ammonium acetate adjusted to pH 5.6 using acetic acid. Samples were reduced by incubation for 30 min at 56 °C at a final concentration of TCEP 10 mM. A 1 μL injection volume corresponding to 100 ng of protein material was used for all experiments.

2.4. Yeast ribosomes preparation

Wild strain yeast *S. cerevisiae* S288C was grown in a synthetic medium made of YNB without amino acids, with ammonium sulfate as a nitrogen source, and glucose (0.5 % m/v) used as the sole source of carbon. 400 mL liquid cultures were inoculated at OD_{600nm} 0.03 using cells grown overnight in 10 mL of the same medium inoculated with a fresh colony from YPD plates. Cultures were incubated in an Infors orbital shaker operated at 200 rpm and 30 °C. When the cell density reached OD_{600nm} 4, cells were collected by centrifugation for 10 min at 4000 g. The cell pellet was resuspended in 5 mL of lysis buffer consisting of HEPES 20 mM pH 7.5, magnesium acetate 10 mM, potassium acetate 50 mM, and one tablet of protease inhibitor cocktail. Two volumes of acid-washed glass beads (0.45–0.5 mm diameter) were added to one volume of the cell suspension, and the cells were lysed by vortexing for three cycles of 5 min with 5 min on ice intervals between each cycle. The resulting cell homogenate was collected by centrifugation for 5 min at 3000 g, aliquoted in 1 mL fractions, and either rapidly frozen at −80 °C or immediately used for ribosome preparations. Total protein amount was measured using a Bradford micro-assay.

A 7 %–47 % sucrose gradient was prepared, by overlaying 2.6 mL each of sucrose solutions containing 7 %, 17 %, 27 %, 37 %, 47 % sucrose (m/m) respectively in Lysis Buffer (total gradient volume = 13.2 mL). These stacked volumes were allowed to sit overnight at 4 °C for equilibration of the phases. 1 mL of freshly prepared lysate, was loaded on top of the gradient, by replacing 1 mL of sucrose gradient on the top of each tube, then centrifuged for 2.5 h at 39,000 g in a SW41Ti rotor maintained at 4 °C (Beckman Coulter, Brea, CA, USA). The gradient was fractionated from the bottom of the centrifuge tube using a Pasteur pipette connected to a peristatic pump, and 200 μL fractions were collected in a microplate. Half of each fraction was transferred in a UV readable plate to record the distribution of ribosome across the fractionated gradient. The UV plate was read at 254 nm for RNA quantification.

Following the localization of the fractions containing the highest RNA content (and thus ribosomes), the total protein amount was measured using a Bradford micro-assay. The buffer composition of the fractionated ribosomes was exchanged to 25 mM ammonium acetate prior to LC-MS analyses, by loading a volume equivalent to 25 μg of protein sample onto a molecular weight cut-off centrifugal filter unit (0.5 mL, 10 kDa cutoff; Amicon, Merck Millipore Sigma) and running 8 cycles of concentration/dilution with the ammonium acetate solution. Samples were acidified with 0.1 % DFA (v/v) and injected directly without any additional precipitation step of nucleic acids.

2.5. LC-MS for top-down analysis

Samples were injected in microliter pickup mode onto a custom made BioResolve RP (450 Å, 2.7 μm, 0.3 mm × 150 mm) Polyphenyl Column (Waters, Milford, MA, USA) using an UltiMate 3000 RSLCnano chromatographic system fitted with a capillary flow meter from ThermoFisher Scientific (Waltham, MA, USA). Mobile phases A and B were respectively H_2O/ACN 98/2 and H_2O/ACN 20/80 acidified with 0.1 % (v/v) FA or DFA. The sample was eluted at a flow rate of 6 μL/min using the following slope change points: 0–5 min hold at 6.3 % B, 5 min slope until 31.3 % B, 40 min slope until 56.2 % B, 5 min wash at 100 % B and finally 10 min hold at 6.3 % B. The eluent was sprayed using the conventional ESI ion source of a Q-Exactive Plus mass spectrometer from ThermoFisher Scientific operating in the positive ion mode. MS spectra were acquired at a resolution of 70 K, with a mass range of *m/z* 500–2500 and an AGC target value of 1e6 for a maximum ion accumulation time of 500 ms. MS/MS data were acquired at similar resolution, in data-dependent top2 acquisition mode with an AGC target value of 5e5 for a maximum ion accumulation time of 250 ms. Parent ions were selected for fragmentation by Higher-energy C-trap Dissociation (HCD) with a dynamic exclusion time of 30 s.

Histone sample was analyzed in positive ion mode on an Orbitrap fusion instrument from ThermoFisher Scientific (Waltham, MA, USA). MS spectra were acquired within a mass range of *m/z* 500–2500, with an AGC target value of 4e5 for a maximum injection time of 250 ms at a resolution of 60 K and 4 microscans accumulation. MS/MS data were acquired at similar resolution, with a first *m/z* of 100, an AGC target value of 1.25e5 for a maximum injection time of 250 ms and 4 microscans accumulation. Parent ions were selected in the top speed 3 s mode for fragmentation by HCD with a dynamic exclusion time of 60 s.

2.6. Design of experiment (DOE)

DOE studies were conducted using Minitab 17. A central composite design (within surface response design) was selected with three continuous factors *i.e.,* the flow rate, the temperature and the gradient slope. The number of replicates was set at one, three points were set at the center of the cube, and alpha was selected as the default value ($\alpha = 1.682$). Two levels were chosen to define the cube points *i.e.,* flow rate from 4 to 6 μL/min, temperature from 40 to 60 °C and gradient slope defined as a gradient starting at 25 % ACN and reaching in 40 min a target factor varying from 45 to 65 % ACN. Finally, the study was designed with a total of 17 experiments, 8 cube points, 6 axial points and 3 center points (details available in Supplementary data 1) and using the Intact Protein Standard Mix. Raw data were processed with the Thermo Xcalibur v3.0.63 software. Peaks were integrated using the Genesis peak algorithm, a Boxcar smoothing value of 3 points and enabling valley detection in peak detection settings. For each peak, retention time (t_R), area and height were determined on the Total Ion Current (TIC) trace and used to calculate surrogate estimators of chromatographic performance and described below as "output variables":

The average area of integrated peaks was used to monitor sensitivity.

The relative t_R difference between two consecutive peaks $\Delta t_R = (t_{R2} - t_{R1})$ was used as a surrogate of selectivity because of difficulties to calculate retention factors accurately with the lack of clear unretained peak information using solely the TIC trace.

Full width at half maximum (FWHM) and resolution (RES) were calculated according to the following formula [26]:

$$FWHM = 0,939 \times \frac{Area}{Height}$$

$$RES = \frac{(t_{R2} - t_{R1})}{(FWHM_1 + FWHM_2)}$$

FWHM, peak area and Δt_R taken as an average from the six peaks given by the Intact Protein Standard Mix were finally used as the three output variables of interest to be optimized during the DOE study.

2.7. Data analysis

2.7.1. NISTmAb and histone samples

Protein mass spectra were deconvoluted using PMI Intact software (Protein Metrics, Cupertino, CA, USA) [27]. Theoretical masses were computed by PMI Intact with chemical element values deduced from MIDAs isotopic distribution exact center of mass [28]. GlycanMass tools provided by the Expasy website [29] and Unimod [30] were used to

calculate average masses for PTMs.

2.7.2. Ribosome preparation

The *S. cerevisiae* proteome from Uniprot was used as database (UP000002311 release 2022_01, 6062 entries) for all of the identification searches. Raw data were processed by Prosight 4.0 on the Proteome Discoverer 2.5 interface from ThermoFisher Scientific (Waltham, MA, USA). The ProSightPD High/High cRAWler node was used with default parameters and ProsightPD 4.0 Annotated Proteoform node was run with mass tolerances of 20 ppm and 2.2 Da for fragment and precursor masses, respectively. Subsequence search was used with a mass tolerance of 20 ppm for both fragment and precursor masses, then another ProsightPD 4.0 Annotated Proteoform node was used for a less stringent search with mass tolerance of 20 ppm and 100 Da for fragment and precursor masses, respectively. *N*-term methionine excision, *N*-term acetylation, acetylation (K), methylation/dimethylation (K, R), phosphorylation (S, T, Y) and oxidation (M) were set as variable modifications. A FDR (False Discovery Rate) of 1 % was applied for the identification.

Raw data were also converted into an mzML format using msConvertGUI [31] (ProteoWizard, version 3.0.19106-4a85cc56b). Then, mzML files were processed using OpenMS tools [32], TOPFD and TOPPIC (OpenMS-2.6.0-pre-HEAD-2020–10-23). TOPFD was used to deconvolute data with the following parameters: HCD fragmentation, maximum charge of 50, maximum mass of 100,000 Da, MS1 S/N ratio of 3.0, MS2 S/N ratio of 1.0, precursor windows size (m/z) of 3.0 and a m/z error of 0.02 Da. TOPPIC was used to analyze ms2.msalign files obtained with TOPFD. TOPPIC was set with the following parameters: Decoy Database, an Evalue of 0.01 was applied at both PrSM and proteoform levels, a mass error tolerance of 20 ppm and a PrSM cluster error tolerance of 1.2 Da were used. One single mass shift of +/− 500 Da was tolerated and variable PTMs as described previously were used for the search.

InteractiVenn [33] (https://wsww.interactivenn.net/) was used to generate a Venn diagram of the identified ribosomal proteins. StringDB [34] was used to generate a functional protein association network (version 11.5, 2021_08, https://string-db.org/) with the following parameters: Full String network and mapping of network edges based on confidence with an interaction score set at “high confidence” (0.700).

Peak capacity was calculated as follows according to both Wu [20] and Neue [35], assuming fully symmetric peaks and an effective separation time based on first and last detected protein signals:

$$Peakcapacity = 1 + \frac{separationtime}{1.7 \times FWHM}$$

3. Results and discussion

3.1. Optimization of protein separation

To begin this work, we performed an initial qualitative comparison of performance between the superficially porous phenyl media and a classical reverse phase C4 column with 300 Å pore size (see supplementary data 2). The intact protein standard mix was run on these two columns at three acetonitrile gradient slopes. All standard proteins exhibited more retention on the phenyl media *versus* its C4 comparator. Separation results were close in term of peak shape and FWHM for the first four eluted peaks while the last two most hydrophobic ones were significantly broader with the C4 column namely the Carbonic Anhydrase II and Exo Klenow proteins. This effect was all the more pronounced when the gradient slope was flattened with a sharp increase of tailing for the Exo Klenow peak and a decrease of its overall height. The superior performance of the phenyl-bonded superficially porous stationary phase for hydrophobic proteins is not surprising, as this sorbent was designed for efficient separations of both mAbs and antibody drug conjugates. In all, these results indicate that the superficially porous phenyl column should be able to better handle protein mixtures with diverse hydrophobicity when compared to a traditional C4 column.

A second part of this work was to investigate whether the use of an ion-pairing agent could be of benefit to TD-proteomics separations with the phenyl column chemistry. The use of TFA as an ion-pairing agent in the mobile phase has been very effectively employed, and it is widely used to improve the peak shape and width of eluted proteins [36]. However, in the case of chromatographic separations hyphenated to MS, FA is preferred because of the strong MS signal suppression induced by TFA. Recently, several papers reported the preferable use of DFA against TFA for its better compromise in terms of ion pairing to improve chromatographic performance over its negative signal suppression effect [25]. Herein, the use of DFA was evaluated and compared to FA in an experiment involving the injection of the intact protein standard mix in two conditions differing by the mobile phase composition only. Fig. 1 panel (a) shows the chromatogram obtained with FA while panel (b) shows the separation as obtained with DFA. The six standard proteins were eluted in the same order, while more retention was observed with the DFA condition – as would be expected from the increased hydrophobicity conferred by a more hydrophobic and effective ion pairing agent [37]. DFA improved also slightly the selectivity as can be observed from the elution of the six proteins spanning over 21.6 min *versus* 18 min in the FA condition. This effect translates directly into an increase of Δt_R between consecutive peaks (see Table 1). Peak sharpness was the most significantly improved parameter with on average a FWHM of 16 s in DFA *versus* 29 s in FA. Altogether, the improved peak widths and selectivity led to a 2-fold increase in resolution when using DFA. It is also worth noting that both peak width and shape were highly compromised in the FA based separation of Exo Klenow protein (~68 kDa) and that the positive effect of switching to DFA might be most convincing with high molecular mass proteins. Finally, the area of peaks from the total ion chromatograms was found to decrease on average by a factor of 3.5 in DFA *versus* FA while the height decreased only by a factor of 2 due to the balancing of effects from ion suppression *versus* improved peak widths. Considering that the data-dependent acquisition mode (DDA) in a TD-experiment is sensitive to a signal threshold for triggering MS/MS events, we could assume that the height of peaks is more relevant to estimate the capacity of recording fragmentation data for protein identification. In this way, we could assume only a 2-fold decrease in sensitivity for TD-experiments using DFA mobile phases while getting a 2-fold increase in protein resolution for an enhanced peak capacity that is by far the most limiting factor that currently needs improvement to maximize protein identifications. Finally, it should be noted that the DFA ion pairing effect was also observed with C4 columns but chromatographic results were still superior with the combined use of DFA with the superficially porous phenyl column (data not shown).

In the next experiments, a DOE-like study (see material & methods) was conducted in order to quickly optimize a couple other variables and to thereby obtain balanced performance in terms of protein separation, peak sharpness and time/method throughput. Three variable parameters were studied as input factors *i.e.,* temperature, flow rate and the eluent gradient slope. These are parameters that are known to be critical for chromatographic reverse phase separation [38]. In this experiment, DFA was kept at a constant mobile phase concentration of 0.1 % to limit the number of tests in the DOE study. FWHM, Δt_R and peak area were used as output factors (as surrogate indicators of separation efficiency, selectivity and sensitivity).

Fig. 2 shows the analysis of variance obtained from a set of 17 injections after fitting the data with a first order model. Results shows that both the flow rate and gradient steepness (slope) significantly affect peak width. This is expected as these parameters are tied to longitudinal diffusion. Average time distance between consecutive peaks was exclusively influenced by the gradient steepness with a fairly flat effect being seen for both flow rate and temperature. Finally, sensitivity, evaluated as the average area of the detected peaks, was shown to be mainly influenced by the flow rate, almost proportionally due to a

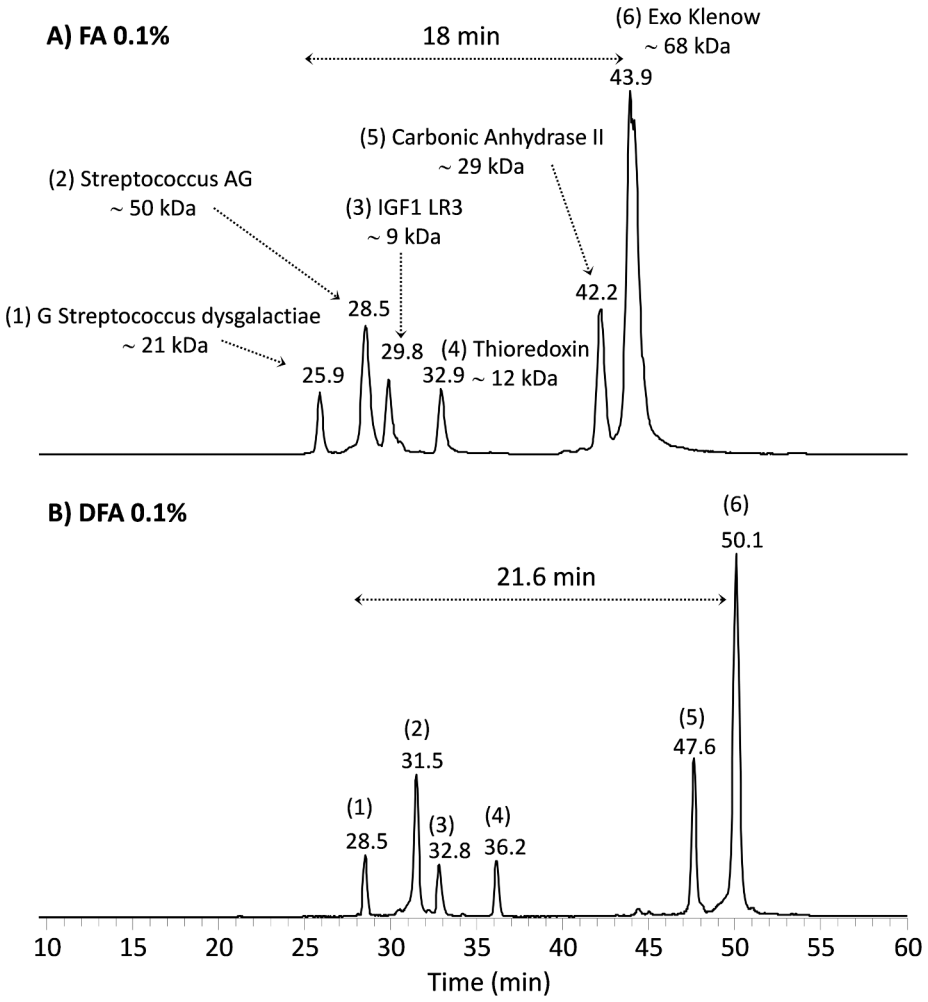


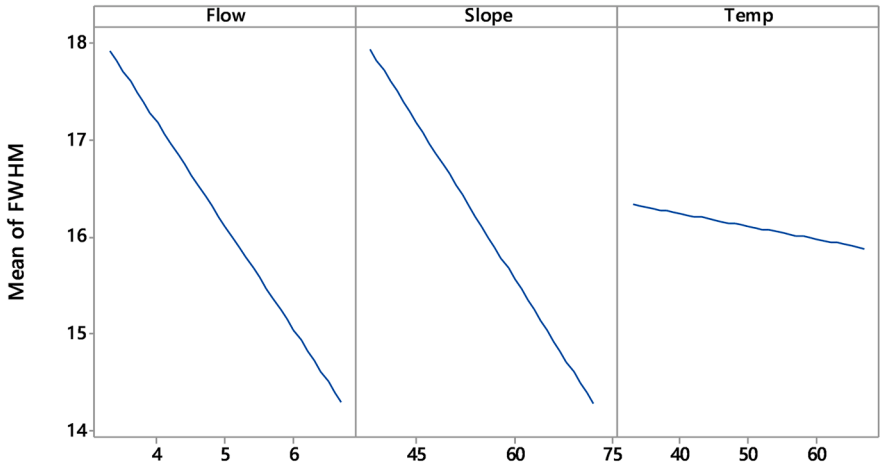
Fig. 1. TIC chromatograms from a 100 ng injection of protein standard mix, using either 0.1 % FA as a modifier (Panel A) or 0.1 % DFA (Panel B). TIC signal on the y-axis (not shown) was normalized to the largest peak displayed.

Table 1
Result summary of chromatographic performance parameters measured with FA and DFA acidic modifiers.

Protein	t_R(min) FA	Area	Height	Δt_R(min)	FWHM (s)	RES	t_R(min) DFA	Area	Height	Δt_R(min)	FWHM (s)	RES
(1) G Protein Streptococcus dysgalactiae, 21 kDa	25.8	7.E+09	3E+08		21	3.0	28.5	2.E+09	2.E+08		13	5.9
(2) Streptococcus AG, 50 kDa	28.5	2.E+10	6E+08	2.7	32	1.4	31.5	7.E+09	4.E+08	3.0	18	2.3
(3) IGF1 LR3, 9 kDa	29.8	9.E+09	3E+08	1.3	25	3.8	32.8	2.E+09	1.E+08	1.3	16	7.1
(4) Thioredoxin, 12 kDa	32.9	8.E+09	3E+08	3.1	24	11.0	36.1	2.E+09	2.E+08	3.4	12	24.7
(5) Carbonic Anhydrase II, 29 kDa	42.2	2.E+10	7E+08	9.3	27	1.3	47.6	6.E+09	4.E+08	11.5	15	3.7
(6) Exo Klenow, 68 kDa	43.8	8.E+10	2E+09	1.6	46		50.1	2.E+10	8.E+08	2.4	24	
Average	**33.8**	**2.E+10**	**6E+08**	**3.6**	**29**	**4.1**	**37.8**	**7.E+09**	**3.E+08**	**4.3**	**16**	**8.7**

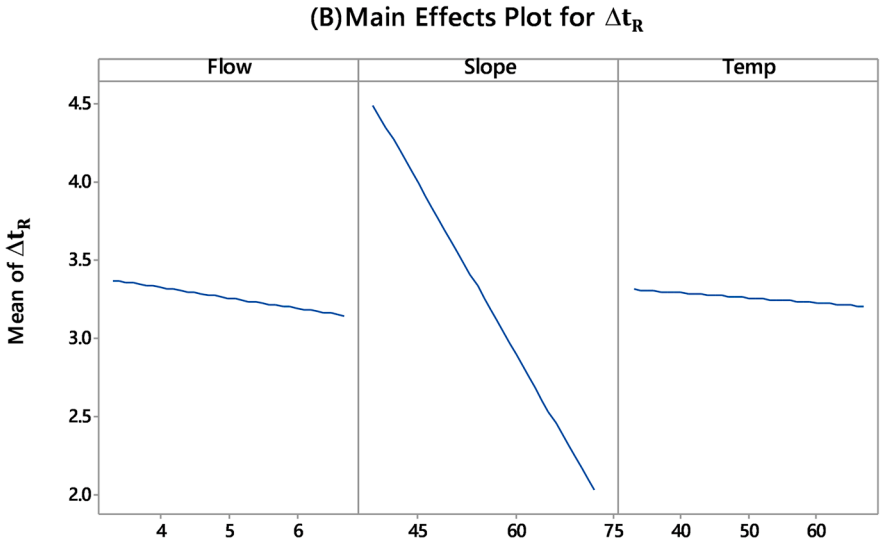
dilution effect observed with increased flow rates [39]. Interestingly, and even if it is not statistically significant (p-value of 0.14), temperature appeared here to have a slightly negative impact on detected peak area. Column temperature is usually seen to have a positive effect on protein recovery [22], so this contradictory outcome suggests an unexpected negative effect. There was no straightforward evidence of temperature induced protein hydrolysis as well, but this hypothesis cannot be fully ruled out. Another plausible explanation for this could be that there is a slight decrease in ionization at high temperature in relation to the electrospray operating conditions of the experiment.

Finally, to compromise between peak sharpness, protein selectivity and elution time, a gradient slope of 0.5 % ACN/min reaching a final value of 45 % ACN (56.2 % B) was preferred. In these conditions, the elution of the largest Exo Klenow protein (peak 6) occurred at 50.1 min, which is just at the start of the column wash with 100 % B. Regarding the flow rate, its increase will improve peak sharpness to the expense of



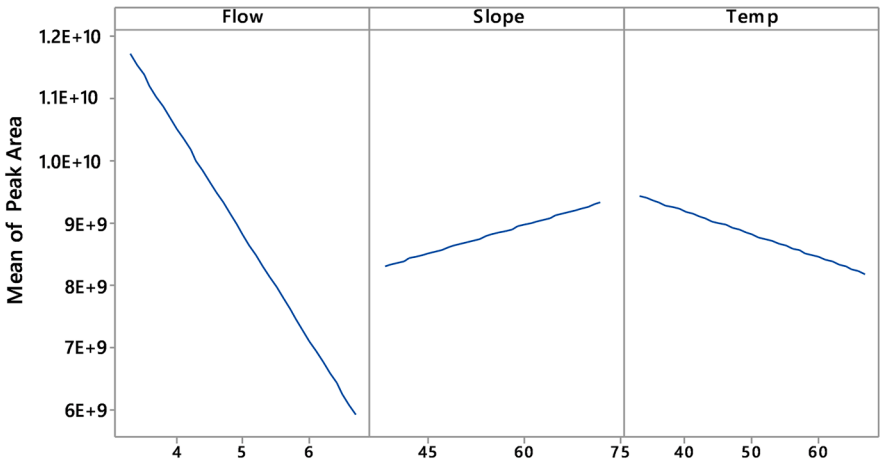
Analysis of Variance

Source	DF	Adj SS	P-Value
Linear	3	31.76	0.00
Flow	**1**	**15.65**	**0.00**
Slope	**1**	**15.88**	**0.00**
Temp	1	0.24	0.52
Error	13	7.08	
Lack-of-Fit	11	6.51	0.37
Pure Error	2	0.56	
Total	16	38.84	
R-sq =	81.78%		



Analysis of Variance

Source	DF	Adj SS	P-Value
Linear	3	7.29	0.00
Flow	1	0.06	0.37
Slope	**1**	**7.22**	**0.00**
Temp	1	0.01	0.66
Error	13	0.07	
Lack-of-Fit	11	0.08	0.00
Pure Error	2	0.00	
Total	16	8.2	
R-sq =	86.42%		



Analysis of Variance

Source	DF	Adj SS	P-Value
Linear	3	4.3E+19	0.00
Flow	**1**	**4.0E+19**	**0.00**
Slope	1	1.2E+18	0.22
Temp	1	1.8E+18	0.14
Error	13	9.8E+18	
Lack-of-Fit	11	9.4E+18	0.19
Pure Error	2	3.7E+17	
Total	16	5.3E+19	
R-sq =	81.49%		

DF: Degree of Freedom
Adj SS: Adjusted Sum of Squares

Fig. 2. Variance analysis of the DOE study with DFA modifier, data were fitted against a first order model, (A) effects plot for FWHM, (B) effects plot for Δt_R and (C) effects plot for peak area.

column operating pressure and MS sensitivity. A balanced flow rate of 6 µL/min was then chosen here for a very acceptable backpressure of about 105–97 bars along the separation. Even if temperature did not appear crucial for the intact protein mix analysis, we still made the choice to keep a relatively high temperature of 60 °C for further analysis as it is known to benefit to the elution of some difficult proteins like mAbs and to limit column backpressure. Finally, method performance was quickly investigated; consecutive dilutions of the protein standard allowed us to approximate a limit of detection at the low ng level. Linearity based on peak area was demonstrated for 5 to 500 ng mass loads of all peaks except for the Exo Klenow protein that exhibits the most intense signal and demonstrated saturation when the injected amount exceeded 100 ng. Repeatability of t_{Rs} was better than < 0.5 % RSD for all the detected peaks (See supplementary data 3). These performance data were obtained on previous generation instrumentations (both the chromatographic system and the mass spectrometer) and may be better when testing on state-of-the-art systems.

3.2. Analysis of the NIST mAb at intact and subunit levels

Antibody-based molecules continue to represent the most widely relied upon drug modality in the biopharmaceutical industry [40]. They

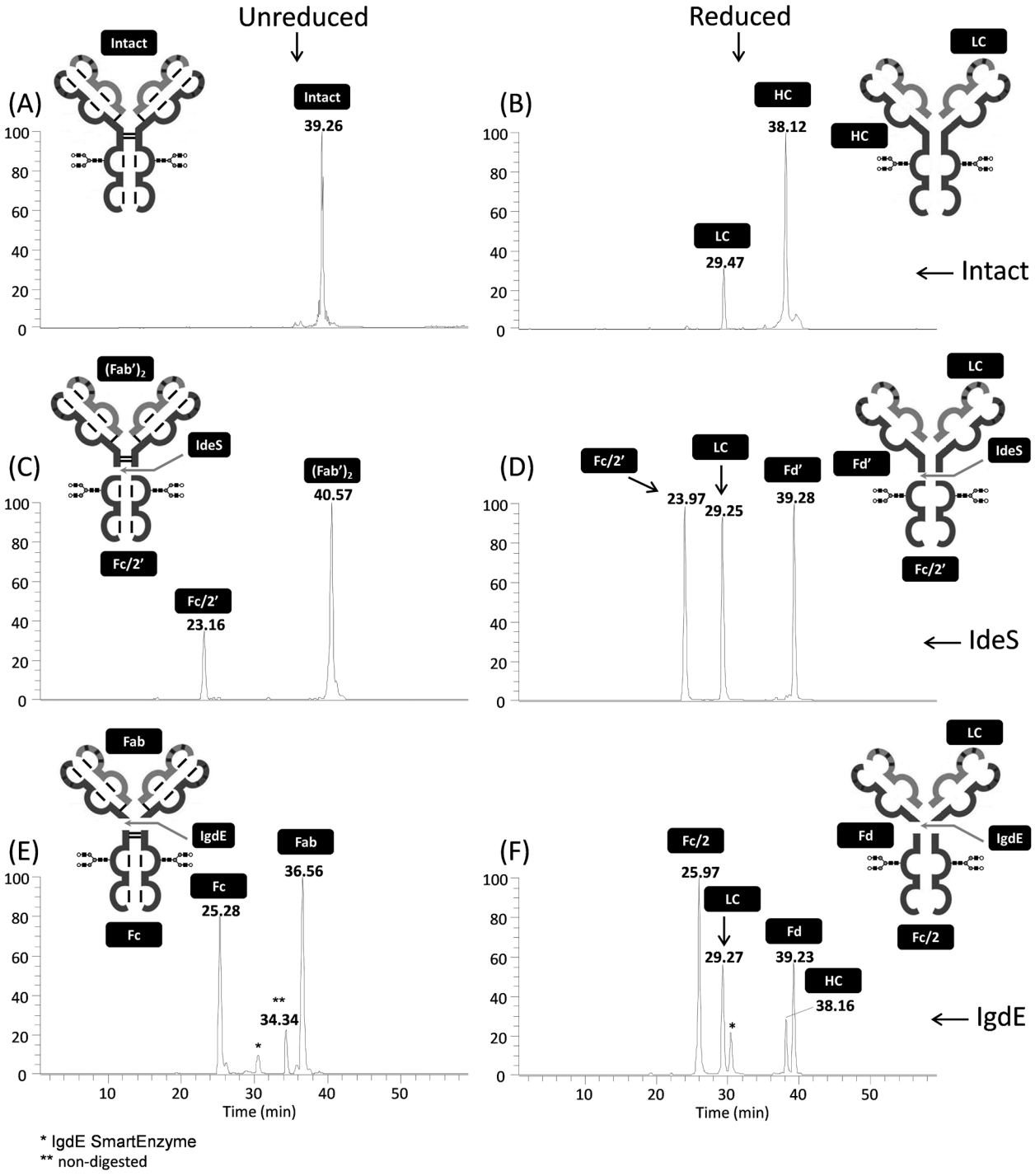


Fig. 3. TIC chromatograms from 100 ng injection of NISTmAb, in (A) intact condition, (B) IdeS condition, (C) IgdE condition, (D) reduced condition, (E) IdeS and reduced condition, (F) IgdE and reduced condition.

are, however, heterogeneous proteins due to multiple enzymatic and chemical PTMs that occur during their manufacturing process. Regulatory authorities require in-depth characterization of these PTMs and detailed control of any that might be deemed quality attributes. Only through this analytical work, can the consistency of a drug substance be soundly demonstrated.

Over the last 20 years, the usage of MS has increased steadily for the analysis of protein-based biotherapeutics [41] and it is now considered to be a fundamental technique for characterizing and monitoring quality attributes. Middle-down analysis is currently-one of the most promising MS-based workflows for gathering information about protein quality attributes [42]. It is based on specific enzymes (for instance immunoglobulin degrading enzymes, like IdeS) that enable minimal sample preparation with limited artefact introduction and the use of high resolution mass spectrometry to observe numerous PTMs in one single run [22,43,44]. In this context, we first tested the developed method with the analysis of the NIST IgG1 mAb standard (RM 8671) in both reduced and non-reduced conditions, at the intact level as well as after IdeS or IgdE digestion. Fig. 3 shows the corresponding TIC chromatograms generated with only 100 ng-injections of material. All chromatograms show an efficient separation and symmetrical, sharp peaks for all the generated species, namely intact, heavy chain (HC), light chain (LC), Fc/2′, $(Fab')_2$, Fd', Fc, Fab, Fc/2 and Fd. The largest peak width was observed on the intact mAb with a FWHM of 19 s. Others peaks were narrower and exhibited an average FWHM of 14 s. Of note, IgdE digestion was not complete as can be deduced in Fig. 3E by the presence of one (Fc + Fab) structure at 34.34 min. The incomplete IgdE digestion was also detected in the reduced version of the sample by the presence of one HC chain still visible at 38.16 min (see Fig. 3F). This partial digestion can be mitigated with an appropriate phosphate buffer concentration for digestion [45]. Apart from all these abundant subunits of interest, numerous minor peaks were also well separated in each condition. Fig. 4 provides a zoomed view of the non-reduced IdeS digestion. Several peaks below 1 % in relative abundance were reliably separated and assigned to their most likely structures according to mass matching. Peaks labelled from 2 to 4 correspond to one non-glycosylated and to partial reduced isoforms of the Fc/2. Peaks 5 to 7 represent clipped forms of the Fab subunit with identified cleavages at the upper hinge of the HC, the C-terminus of the LC and reduction of the hinge disulfide bridge. These results are in agreement with the knowledge of antibody fragmentation susceptibility [46] and they have already been reported in various studies [47–50]. Finally, peaks labelled as 9 and 10 were identified as singly reduced and oxidized Fab variants. In these separations, there is a peak labeled with an asterisk and this corresponds to the IdeS enzyme used during sample preparation. This artefact can be removed by slightly modifying the sample preparation using immobilized IdeS on agarose beads. In conclusion, the present method is fully compatible with the analysis of mAbs at both the intact and subunit level, and it has shown itself to be useful for characterizing and monitoring low levels of potential critical quality attributes with only 100 ng of injected material per run and only a few micrograms of mAb consumed during sample preparation.

3.3. Analysis of bovine histones

Histones are the major structural proteins of chromosomes; they associate with DNA in the nucleus and help condense it into chromatin. They are highly modified with PTMs that modulate chromatin structure which in turn plays an important role in control and adjustment of gene expression. Over the last decade, mass spectrometry has become a valuable tool to studying the heterogeneity of those highly modified proteins, especially analyzing them at the intact level, because it has provided unique information regarding the precise combination of multiple PTMs on the same protein [51,52]. We used here a bovine histone mixture from calf thymus to test the chromatographic performance of the developed method. This was an insight investigation since

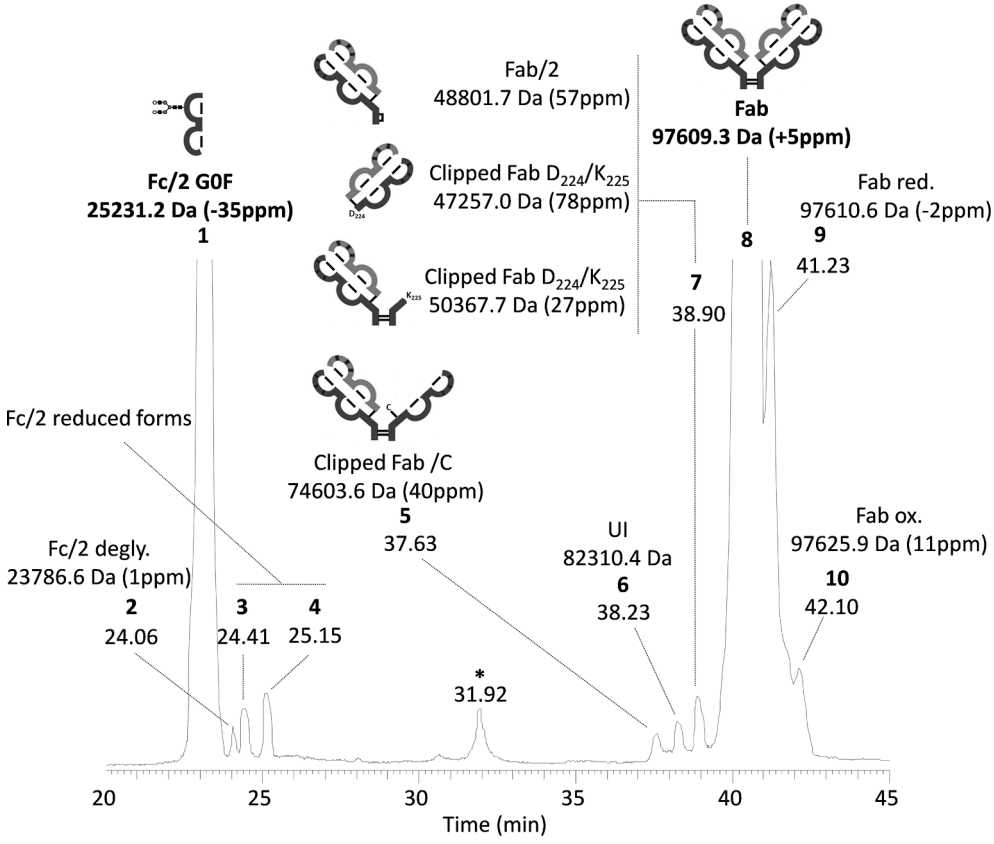


Fig. 4. Enlarged view of NISTmAb TIC chromatogram in IdeS non-reduced condition.

histones are thought to be challenging to work with highly basic proteins that can exhibit pronounced secondary interactions and peak broadening during RPLC. Fig. 5A shows the TIC profile obtained from a 100 ng-injection of material. Proteins were found to elute from 12 to 40 min; 13 peaks appeared to represent the majority of the TIC peaks areas. These peaks have been annotated with their protein assignments in Fig. 5A. Protein identification was performed by mass matching of histones theoretical sequences and sequencing data were used for confirmation. Table 2 summarizes the most relevant mass assignments made in this experiment while Fig. 5 panels B, C, D and E represent the deconvoluted mass spectra of H2B1, H4, H3 and H2A1 histone variants. Extracted ion chromatogram (XIC) signals from these 4 peaks were used to estimate an average LC peak FWHM of 16 s, which is reassuringly comparable to results obtained with the protein standard.

It is worthwhile to compare the results of this experiment with those published by Zhou *et al.* [39] wherein bovine histone RPLC-MS analysis was scaled to 1.0 mm ID chromatography with C_{18} bonded phases of fully *versus* superficially porous commercial columns. In their work, the authors obtained a reasonable separation for the most abundant histones from the five major families and even detected variants of lower

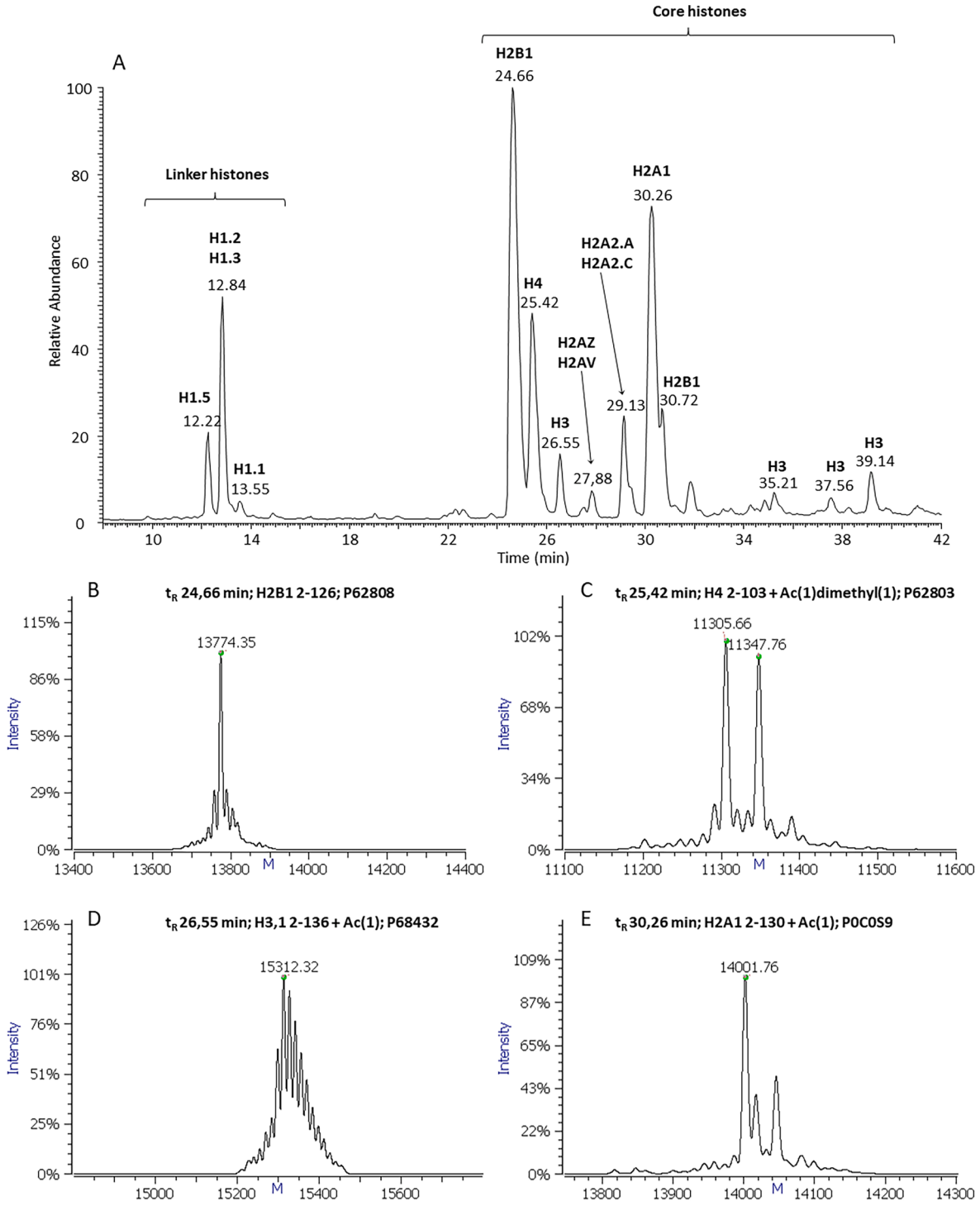


Fig. 5. TIC chromatogram from a 100 ng injection of bovine histone mixture from calf thymus (A), H2B1 deconvoluted mass spectrum at 24.6 m min (B), H4 deconvoluted mass spectrum at 25.42 min (C), H3 deconvoluted mass spectrum at 26.55 min (D) and H2A1 deconvoluted mass spectrum at 30.26 min (E).

Table 2
Masses and assignments of the main Bovine histone species detected by MS.

t_R (min)	histones	accession #	Sequence/PTMs	Theoretical masses (Da)	Experimental masses (Da)	Δm (Da)
12.22	H1.5	XP_010816761.1	2–226/Ac(1)	22604.83	22604.19	−0.64
12.84	H1.2	P02253	2–213/Ac(1)	21266.28	21265.94	−0.33
	H1.3	A7MAZ5	2–221/Ac(1)	22064.22	22063.56	−0.66
	H1.3	A7MAZ5	9–221/Ac(1)	21366.52	21367.58	1.06
13.55	H1.1	G3N131	2–218/Ac(1)	22011.06	22010.27	−0.78
24.66	H2B1	P62808	2–126	13774.81	13774.35	−0.46
25.42	H4	P62803	2–103/Ac(1)dimethyl(1)	11306.14	11305.66	−0.47
	H4	P62803	2–103/Ac(2)dimethyl(1)	11348.14	11347.76	−0.38
26.55	H3.1*	P68432	2–136/Ac(1)	15312.74	15312.32	−0.42
27.88	H2AV	Q32LA7	2–128	13377.35	13377.15	−0.2
	H2AZ	P0C0S4	2–128	13421.4	13421.06	−0.34
29.13	H2A2A	Q6FI13.3	2–130/Ac(1)	14006.15	14005.71	−0.44
	H2A2C	A1A4R1	2–129/Ac(1)	13899.04	13898.74	−0.3
30.26	H2A1	P0C0S9	2–130/Ac(1)	14002.14	14001.76	−0.38
30.72	H2B1	P62808	2–126/Ac(1)	13816.81	13816.46	−0.35
	H2B1N	Q32L48	2–126/Ac(1)methyl(1)	13847.85	13846.58	−1.27
35.21	H3.2*	P84227	2–136/Ac(1)	15298.68	15297.35	−1.33
37.56	H3	–	–	–	–	–
39.14	H3.1*	P68432	2–136/Ac(1)C(red)	15314.76	15314.69	−0.07

abundance by increasing protein injection up to 1.6 μg. They also tried to use different ion pairing agents but experienced significant peak broadening using DFA with columns made of superficially porous particles. Compared to these published results, we obtained here a very similar elution order of the detected species. Importantly, H4 and H2B1 variants were fully resolved in our experiments while those two species showed co-elution in previous work. Our scaling of chromatography down to 0.3 mm I.D. has also helped increase method sensitivity, with at least the same number of isoforms identified from a 10 to 20 reduced mass load injection. In that respect, this type of column and workflow could be of benefit when only a scarce amount of material can be extracted. Interestingly, bovine histone mixture was also studied previously by capillary zone electrophoresis as an alternative separation technology [53]. This technology requires only small amounts of proteins to be loaded in volumes as low as 25 nL and separation of 4 histones peaks was achieved over a 5 min window.

3.4. Chromatographic separation of yeast ribosomes

Ribosomes are ribonucleoprotein complexes found in the cytosol of eukaryotic and prokaryotic cells [54]. Their function is to synthesize proteins by decoding the information contained in mRNA transcripts. A wide diversity of ribosomes has been studied and it is known that each type can have very distinct ribosomal protein stoichiometries and a vast array of PTMs. By analogy with the "histone code" it is speculated that there may be a "ribosome code" composed of unique combinations of proteoforms and that each could hold potential functional roles in regulating translational outputs [55,56]. To this end, TD-proteomics could be used as an approach to help define the regulation mechanisms involving PTMs and ribosome activity. Keeping that in mind, we turned to testing the underlying method with a protein mixture composed of purified yeast ribosomes. Fig. 6A shows the TIC chromatogram resulting from the analysis of 1 μg of a ribosomal preparation. 25 major peaks labelled at a TIC threshold of 10 % were chromatographically resolved in a time range of 5–50 min. Of interest, protein derived signals were identified from 9.3 to 51.4 min with no significant detection in the column wash step. That is, proteins were detected across the entirety of the gradient with optimum separation. Fig. 6B shows an extracted ion-chromatogram of *m/z* values representing three distinct proteins that eluted at the start, middle and end of the gradient. Peak widths for these three species ranged from 11 to 17 s (FWHM). An effective peak capacity was also estimated according to an equation from Wu *et al.* [21] and by using the first and last real protein elution time as described by Neue [35] and Wang [57]. By taking an effective elution time window of 42.4 min and an estimated average FWHM of 15 s, a peak capacity close to 100 was approximated. Roth *et al.* [58] reported previously a protein separation at capillary level with extremely narrow peak widths of about 5 to 6 s but only with a very fast gradient of 3 min and a comparatively high flow rate *versus* what has been employed in this work. A very fast gradient can lead to poor peak capacity, as a result, Roth and co-workers also reported the separation of ~500 ng total protein from mouse heart homogenate with a longer 20 min gradient. From this experiment, they reported 59 unique identified proteins and a mean peak capacity of 106.

Usually, the acquisition of informative TD-mass spectra is quite slow to accumulate ion signal as compared to BU-proteomics. Top2 MSMS cycles are typically used with a sequencing speed close to 1 Hz in order to accumulate several microscans. This instrument limitation makes it more favorable to apply extended gradients. In our experiment, MS and MSMS data were next processed with Prosight and Toppic for protein identification. Fig. 6C shows the Venn diagram of the 89 unique identified proteins by combination of outputs from the two software. 64 proteins (about 72 %) were consistently identified by both while each of the two software applications singly identified 14 % of the remaining proteins *i.e.,* 12 and 13 unique proteins. Fig. 6D shows finally the STRING protein graph composed of 89 identified nodes with protein–-protein interaction links or connections among them. Red nodes show proteins with GO molecular function classified as structural constituents of the ribosome in accordance with expected identifications.

Altogether, these results show that significant amounts of information rich MS data can be acquired for top-down analysis with nanogram level injections of a rather complex protein mixture. Even if chromatographic resolution of proteins is not yet at the level of peptides (which tend to show 6 to 8 s peak widths on similar gradient time separations), we have demonstrated here the feasibility to elute proteins with symmetrical peak shape and FWHM widths only twice larger. The use of DFA contributes significantly to this improved resolution and it is shown here to not dramatically compromise MS sensitivity.

4. Conclusions

A wide-pore silica-based superficially porous phenyl bonded media, commercially available as 2.7 μm particles with an average pore diameter of 450 Å was evaluated for capillary scale TD-proteomics. The choice of a 0.3 mm ID column made it very easy to operate this chromatography with a conventional ESI ion source. The superior behavior of this media was observed for hydrophobic proteins when compared to traditional C4 columns used for protein separations. Then, the positive effect of using DFA as an ion pairing agent for TD-proteomics separations and its coupling with MS was demonstrated. Significant

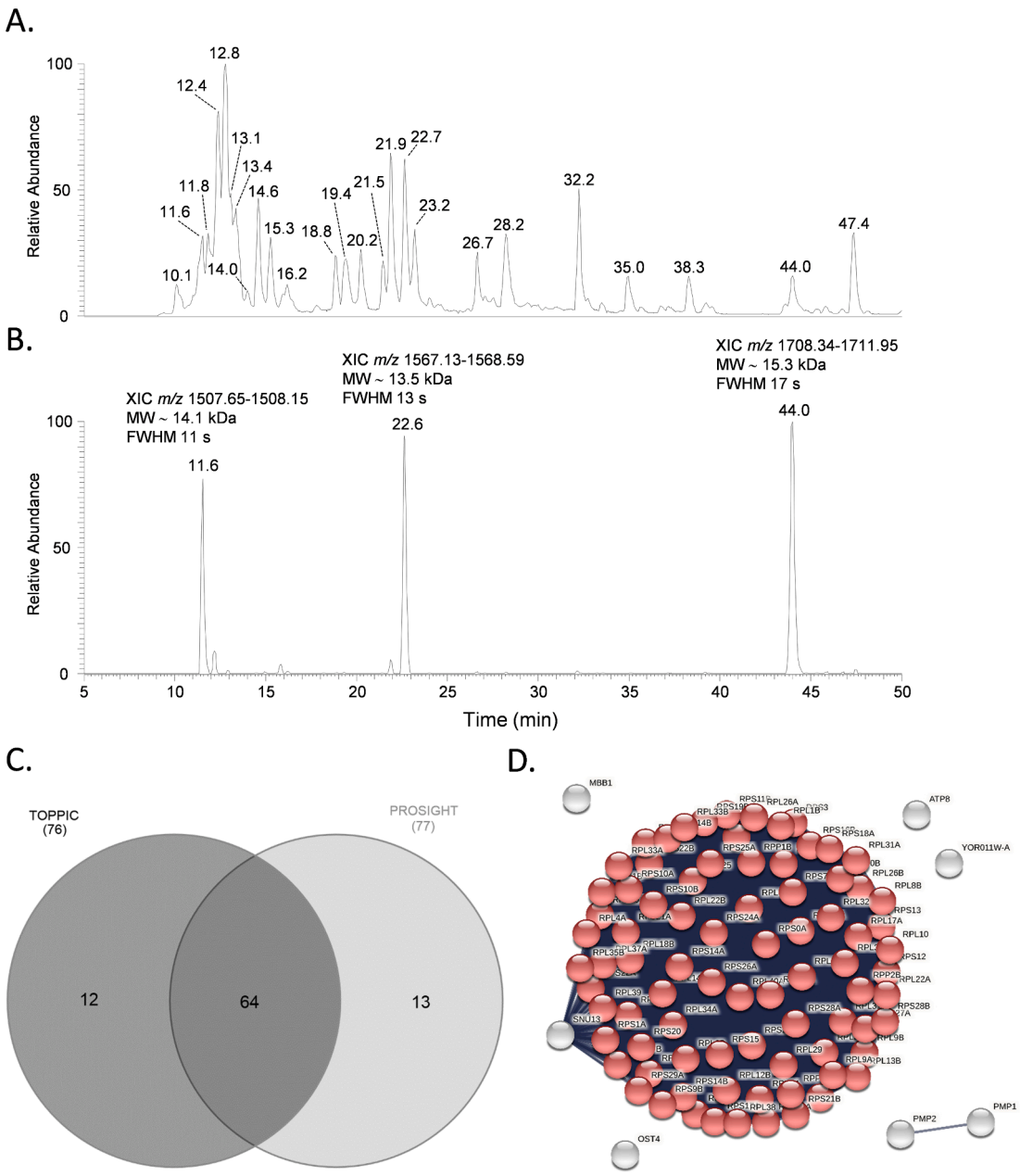


Fig. 6. (A) TIC chromatogram from a 1 μg injection of a ribosomal preparation from Saccharomyces Cerevisiae, (B) extracted ion-chromatogram of *m/z* signals from three distinct proteins eluted at the early, middle and end of the gradient (C) Venn diagram of identified proteins using either TOPPIC or PROSIGHT and (D) full STRING network of the 89 identified proteins, red nodes belong to the cytosolic ribosome cellular component from Gene Ontology enrichment. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

chromatographic benefits were obtained with only a 2–3-fold MS sensitivity loss, so DFA has been confirmed to be a balanced mobile phase compromise worth broader consideration for TD-proteomics. By combining DFA with the phenyl bonded superficially porous particle column, we have developed a method that proved itself to be robust and able to separate proteins with excellent peak shape and high peak capacity with only a 2-fold decrease in MS sensitivity. In comparison, TFA might show equivalent or slightly better chromatography but it will cause an up to 10-fold decrease in MS sensitivity *versus* FA. With the presented method, about 0.1–1 μg of total protein mixture can be routinely loaded onto the column, though the limit of detection will be close to 1 ng. Outstanding versatility was observed with no further optimization required to analyze heterogeneous samples such as mAb subunit digestions, histones and eukaryotic ribosomal proteins. This phenyl stationary phase can be requested for custom column packing and the use of standardized manufacturing processes should make for TD experiments. In our hands, operating pressures did not exceed 105 bars so there it can be broadly applied on all types of LC hardware so long as considerations are made for flow rate and dispersion optimization. In all, we believe this chromatographic approach will make for a reliable and readily accessible technique for other investigators looking for a starting point for TD-proteomics.

5. Funding sources

This study was supported by Université Paris Cité/CNRS and Region Ile de France.

6. Author statement

LL, VL, MK and NS made a substantial contribution to acquisition, data analysis and writing of the original draft. MAL and JMC both contributed to resources, scientific insights and critical review of both the original and revised draft. GC conceptualized and supervised the study, he contributed to draft the article and to its revision. All the authors commented on the paper.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Matthew Lauber is an employee of Waters Corporation, which has developed the wide-pore silica-based superficially porous media for commercial sale.

Data availability

The authors do not have permission to share data.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jchromb.2022.123566.

References

[1] S. Stepanova, V. Kasicka, Recent developments and applications of capillary and microchip electrophoresis in proteomics and peptidomics (2015-mid 2018), J. Sep. Sci. 42 (2019) 398–414.

[2] Z.B. Zhang, Y.Y. Qu, N.J. Dovichi, Capillary zone electrophoresis-mass spectrometry for bottom-up proteomics, Trends Anal. Chem. 108 (2018) 23–37.

[3] J.K. Eng, A.L. McCormack, J.R. Yates, An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database, J. Am. Soc. Mass Spectrom. 5 (1994) 976–989.

[4] J.R. Yates, C.I. Ruse, A. Nakorchevsky, Proteomics by mass spectrometry: approaches, advances, and applications, Annu. Rev. Biomed. Eng. 11 (2009) 49–79.

[5] R. Aebersold, J.N. Agar, I.J. Amster, M.S. Baker, C.R. Bertozzi, E.S. Boja, C.E. Costello, B.F. Cravatt, C. Fenselau, B.A. Garcia, Y. Ge, J. Gunawardena, R.C. Hendrickson, P.J. Hergenrother, C.G. Huber, A.R. Ivanov, O.N. Jensen, M.C. Jewett, N.L. Kelleher, L.L. Kiessling, N.J. Krogan, M.R. Larsen, J.A. Loo, R.R. Ogorzalek Loo, E. Lundberg, M.J. MacCoss, P. Mallick, V.K. Mootha, M. Mrksich, T. W. Muir, S.M. Patrie, J.J. Pesavento, S.J. Pitteri, H. Rodriguez, A. Saghatelian, W. Sandoval, H. Schluter, S. Sechi, S.A. Slavoff, L.M. Smith, M.P. Snyder, P.M. Thomas, M. Uhlen, J.E. Van Eyk, M. Vidal, D.R. Walt, F.M. White, E.R. Williams, T. Wohlschlager, V.H. Wysocki, N.A. Yates, N.L. Young, B. Zhang, How many human proteoforms are there? Nat. Chem. Biol. 14 (2018) 206–214.

[6] I. Ntai, K. Kim, R.T. Fellers, O.S. Skinner, A.D.t. Smith, B.P. Early, J.P. Savaryn, R. D. LeDuc, P.M. Thomas, N.L. Kelleher, Applying label-free quantitation to top down proteomics, Anal. Chem. 86 (2014) 4961–4968.

[7] L.V. Schaffer, R.J. Millikin, R.M. Miller, L.C. Anderson, R.T. Fellers, Y. Ge, N. L. Kelleher, R.D. LeDuc, X. Liu, S.H. Payne, L. Sun, P.M. Thomas, T. Tucholski, Z. Wang, S. Wu, Z. Wu, D. Yu, M.R. Shortreed, L.M. Smith, Identification and quantification of proteoforms by mass spectrometry, Proteomics 19 (2019) e1800361.

[8] A.J. Cesnik, M.R. Shortreed, L.V. Schaffer, R.A. Knoener, B.L. Frey, M. Scalf, S. K. Solntsev, Y. Dai, A.P. Gasch, L.M. Smith, Proteoform suite: software for constructing, quantifying, and visualizing proteoform families, J. Proteome Res. 17 (2018) 568–578.

[9] W. Cai, H. Guner, Z.R. Gregorich, A.J. Chen, S. Ayaz-Guner, Y. Peng, S.G. Valeja, X. Liu, Y. Ge, MASH suite pro: a comprehensive software tool for top-down proteomics, Mol. Cell. Proteomics 15 (2016) 703–714.

[10] Q. Kou, L. Xun, X. Liu, TopPIC: a software tool for top-down mass spectrometry-based proteoform identification and characterization, Bioinformatics 32 (2016) 3495–3497.

[11] Z. Wu, D.S. Roberts, J.A. Melby, K. Wenger, M. Wetzel, Y. Gu, S.G. Ramanathan, E. F. Bayne, X. Liu, R. Sun, I.M. Ong, S.J. McIlwain, Y. Ge, M.A.S.H. Explorer, A universal software environment for top-down proteomics, J. Proteome Res. 19 (2020) 3867–3876.

[12] R.T. Fellers, J.B. Greer, B.P. Early, X. Yu, R.D. LeDuc, N.L. Kelleher, P.M. Thomas, ProSight Lite: graphical software to analyze top-down mass spectrometry data, Proteomics 15 (2015) 1235–1238.

[13] C.J. DeHart, R.T. Fellers, L. Fornelli, N.L. Kelleher, P.M. Thomas, Bioinformatics analysis of top-down mass spectrometry data with ProSight Lite, Methods Mol. Biol. 1558 (2017) 381–394.

[14] Y. Qi, D.A. Volmer, Electron-based fragmentation methods in mass spectrometry: an overview, Mass Spectrom. Rev. 36 (2017) 4–15.

[15] L. Fornelli, K. Srzentic, T.K. Toby, P.F. Doubleday, R. Huguet, C. Mullen, R. D. Melani, H. Dos Santos Seckler, C.J. DeHart, C.R. Weisbrod, K.R. Durbin, J. B. Greer, B.P. Early, R.T. Fellers, V. Zabrouskov, P.M. Thomas, P.D. Compton, N. L. Kelleher, Thorough performance evaluation of 213 nm ultraviolet photodissociation for top-down proteomics, Mol. Cell. Proteomics 19 (2020) 405–420.

[16] J.F. Greisch, S.A.M. van der Laarse, A.J.R. Heck, Enhancing top-down analysis using chromophore-assisted infrared multiphoton dissociation from (phospho) peptides to protein assemblies, Anal. Chem. 92 (2020) 15506–15516.

[17] J.A. Melby, Y. Jin, Z. Lin, T. Tucholski, Z. Wu, Z.R. Gregorich, G.M. Diffee, Y. Ge, Top-down proteomics reveals myofilament proteoform heterogeneity among various rat skeletal muscle tissues, J. Proteome Res. 19 (2020) 446–454.

[18] D.P. Donnelly, C.M. Rawlins, C.J. DeHart, L. Fornelli, L.F. Schachner, Z. Lin, J. L. Lippens, K.C. Aluri, R. Sarin, B. Chen, C. Lantz, W. Jung, K.R. Johnson, A. Koller, J.J. Wolff, I.D.G. Campuzano, J.R. Auclair, A.R. Ivanov, J.P. Whitelegge, L. Pasa-Tolic, J. Chamot-Rooke, P.O. Danis, L.M. Smith, Y.O. Tsybin, J.A. Loo, Y. Ge, N. L. Kelleher, J.N. Agar, Best practices and benchmarks for intact protein analysis for top-down mass spectrometry, Nat. Methods 16 (2019) 587–594.

[19] H.D.S. Seckler, L. Fornelli, R.K. Mutharasan, C.S. Thaxton, R. Fellers, M. Daviglus, A. Sniderman, D. Rader, N.L. Kelleher, D.M. Lloyd-Jones, P.D. Compton, J. T. Wilkins, A. Targeted, Differential top-down proteomic methodology for comparison of ApoA-I proteoforms in individuals with high and low HDL efflux capacity, J. Proteome Res. 17 (2018) 2156–2164.

[20] Y. Liang, Y. Jin, Z. Wu, T. Tucholski, K.A. Brown, L. Zhang, Y. Zhang, Y. Ge, Bridged hybrid monolithic column coupled to high-resolution mass spectrometry for top-down proteomics, Anal. Chem. 91 (2019) 1743–1747.

[21] Z. Wu, B. Wei, X. Zhang, M.J. Wirth, Efficient separations of intact proteins using slip-flow with nano-liquid chromatography-mass spectrometry, Anal. Chem. 86 (2014) 1592–1598.

[22] B. Bobaly, M. Lauber, A. Beck, D. Guillarme, S. Fekete, Utility of a high coverage phenyl-bonding and wide-pore superficially porous particle for the analysis of monoclonal antibodies and related products, J. Chromatogr. A 1549 (2018) 63–76.

[23] M. Lauber, D. Walsh, Chromatographic Compositions, US11235308B2, Waters Technologies Corporation, Milford, MA (US), United State, 2022.

[24] H. Lardeux, B.L. Duivelshof, O. Colas, A. Beck, D.V. McCalley, D. Guillarme, V. D'Atri, Alternative mobile phase additives for the characterization of protein biopharmaceuticals in liquid chromatography - mass spectrometry, Anal. Chim. Acta 1156 (2021), 338347.

[25] J.M. Nguyen, J. Smith, S. Rzewuski, C. Legido-Quigley, M.A. Lauber, High sensitivity LC-MS profiling of antibody-drug conjugates with difluoroacetic acid ion pairing, MAbs 11 (2019) 1358–1366.

[26] N. Dyson, Chromatographic Integration Methods, The Royal Society of Chemistry, 1990.

[27] M. Bern, T. Caval, Y.J. Kil, W. Tang, C. Becker, E. Carlson, D. Kletter, K.I. Sen, N. Galy, D. Hagemans, V. Franc, A.J.R. Heck, Parsimonious charge deconvolution for native mass spectrometry, J. Proteome Res. 17 (2018) 1216–1226.

[28] G. Alves, A.Y. Ogurtsov, Y.K. Yu, Molecular isotopic distribution analysis (MIDAs) with adjustable mass accuracy, J. Am. Soc. Mass Spectrom. 25 (2014) 57–70.

[29] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31 (2003) 3784–3788.

[30] D.M. Creasy, J.S. Cottrell, Unimod: protein modifications for mass spectrometry, Proteomics 4 (2004) 1534–1536.

[31] M.C. Chambers, B. Maclean, R. Burke, D. Amodei, D.L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt, J. Egertson, K. Hoff, D. Kessner, N. Tasman, N. Shulman, B. Frewen, T.A. Baker, M.Y. Brusniak, C. Paulse, D. Creasy, L. Flashner, K. Kani, C. Moulding, S.L. Seymour, L.M. Nuwaysir, B. Lefebvre, F. Kuhlmann, J. Roark, P. Rainer, S. Detlev, T. Hemenway, A. Huhmer, J. Langridge, B. Connolly, T. Chadick, K. Holly, J. Eckels, E.W. Deutsch, R. L. Moritz, J.E. Katz, D.B. Agus, M. MacCoss, D.L. Tabb, P. Mallick, A cross-platform toolkit for mass spectrometry and proteomics, Nat. Biotechnol. 30 (2012) 918–920.

[32] H.L. Rost, T. Sachsenberg, S. Aiche, C. Bielow, H. Weisser, F. Aicheler, S. Andreotti, H.C. Ehrlich, P. Gutenbrunner, E. Kenar, X. Liang, S. Nahnsen, L. Nilse, J. Pfeuffer, G. Rosenberger, M. Rurik, U. Schmitt, J. Veit, M. Walzer, D. Wojnar, W.E. Wolski, O. Schilling, J.S. Choudhary, L. Malmstrom, R. Aebersold, K. Reinert, O. Kohlbacher, OpenMS: a flexible open-source software platform for mass spectrometry data analysis, Nat. Methods 13 (2016) 741–748.

[33] H. Heberle, G.V. Meirelles, F.R. da Silva, G.P. Telles, R. Minghim, InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams, BMC Bioinf. 16 (2015) 169.

[34] D. Szklarczyk, A.L. Gable, K.C. Nastou, D. Lyon, R. Kirsch, S. Pyysalo, N.T. Doncheva, M. Legeay, T. Fang, P. Bork, L.J. Jensen, C. von Mering, The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets, Nucleic Acids Res. 49 (2021) D605–D612.

[35] U.D. Neue, Peak capacity in unidimensional chromatography, J. Chromatogr. A 1184 (2008) 107–130.

[36] W.C. Mahoney, M.A. Hermodson, Separation of large denatured peptides by reverse phase high performance liquid chromatography. Trifluoroacetic acid as a peptide solvent, J. Biol. Chem. 255 (1980) 11199–11203.

[37] M. Shibue, C.T. Mant, R.S. Hodges, Effect of anionic ion-pairing reagent concentration (1–60 mM) on reversed-phase liquid chromatography elution behaviour of peptides, J. Chromatogr. A 1080 (2005) 58–67.

[38] N.S. Lakka, C. Kuppan, Principles of chromatography method development. Biochemical Analysis Tools - Methods for Bio-Molecules Studies, IntechOpen, London, United Kingdom, 2019.
[39] Y. Zhou, X. Zhang, L. Fornelli, P.D. Compton, N. Kelleher, M.J. Wirth, Chromatographic efficiency and selectivity in top-down proteomics of histones, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 1044–1045 (2017) 47–53.
[40] G. Walsh, Biopharmaceutical benchmarks 2018, Nat. Biotechnol. 36 (2018) 1136–1145.
[41] S. Rogstad, A. Faustino, A. Ruth, D. Keire, M. Boyne, J. Park, A retrospective evaluation of the use of mass spectrometry in FDA biologics license applications, J. Am. Soc. Mass Spectrom. 28 (2017) 786–794.
[42] D. Rathore, A. Faustino, J. Schiel, E. Pang, M. Boyne, S. Rogstad, The role of mass spectrometry in the characterization of biologic protein products, Expert Rev. Proteomics 15 (2018) 431–449.
[43] G.D. Pipes, P. Campbell, P.V. Bondarenko, B.A. Kerwin, M.J. Treuheit, H.S. Gadgil, Middle-down fragmentation for the identification and quantitation of site-specific methionine oxidation in an IgG1 molecule, J. Pharm. Sci. 99 (2010) 4469–4476.
[44] G. Chevreux, N. Tilly, N. Bihoreau, Fast analysis of recombinant monoclonal antibodies using IdeS proteolytic digestion and electrospray mass spectrometry, Anal. Biochem. 415 (2011) 212–214.
[45] V. Faid, Y. Leblanc, N. Bihoreau, G. Chevreux, Middle-up analysis of monoclonal antibodies after combined IgdE and IdeS hinge proteolysis: Investigation of free sulfhydryls, J. Pharm. Biomed. Anal. 149 (2017) 541–546.
[46] H. Liu, G. Gaza-Bulseco, E. Lundell, Assessment of antibody fragmentation by reversed-phase liquid chromatography and mass spectrometry, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 876 (2008) 13–23.
[47] J. Vlasak, R. Ionescu, Fragmentation of monoclonal antibodies, MAbs 3 (2011) 253–263.
[48] Y. Leblanc, V. Faid, M.A. Lauber, Q. Wang, N. Bihoreau, G. Chevreux, A generic method for intact and subunit level characterization of mAb charge variants by native mass spectrometry, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 1133 (2019), 121814.
[49] F. Fussl, K. Cook, K. Scheffler, A. Farrell, S. Mittermayr, J. Bones, Charge variant analysis of monoclonal antibodies using direct coupled pH gradient cation exchange chromatography to high-resolution native mass spectrometry, Anal. Chem. 90 (2018) 4669–4676.
[50] S. Wang, A.P. Liu, Y. Yan, T.J. Daly, N. Li, Characterization of product-related low molecular weight impurities in therapeutic monoclonal antibodies using hydrophilic interaction chromatography coupled with mass spectrometry, J. Pharm. Biomed. Anal. 154 (2018) 468–475.
[51] Y. Zheng, X. Huang, N.L. Kelleher, Epiproteomics: quantitative analysis of histone marks and codes by mass spectrometry, Curr. Opin. Chem. Biol. 33 (2016) 142–150.
[52] K. Contrepois, E. Ezan, C. Mann, F. Fenaille, Ultra-high performance liquid chromatography-mass spectrometry for the fast profiling of histone post-translational modifications, J. Proteome Res. 9 (2010) 5501–5509.
[53] D. Chen, Z. Yang, X. Shen, L. Sun, Capillary zone electrophoresis-tandem mass spectrometry as an alternative to liquid chromatography-tandem mass spectrometry for top-down proteomics of histones, Anal. Chem. 93 (2021) 4417–4424.
[54] M. van de Waterbeemd, K.L. Fort, D. Boll, M. Reinhardt-Szyba, A. Routh, A. Makarov, A.J. Heck, High-fidelity mass analysis unveils heterogeneity in intact ribosomal particles, Nat. Methods 14 (2017) 283–286.
[55] E. Emmott, M. Jovanovic, N. Slavov, Ribosome Stoichiometry: From Form to Function, Trends Biochem. Sci 44 (2019) 95–109.
[56] D. Simsek, M. Barna, An emerging role for the ribosome as a nexus for post-translational modifications, Curr. Opin. Cell Biol. 45 (2017) 92–101.
[57] X. Wang, D.R. Stoll, A.P. Schellinger, P.W. Carr, Peak capacity optimization of peptide separations in reversed-phase gradient elution chromatography: fixed column format, Anal. Chem. 78 (2006) 3406–3416.
[58] M.J. Roth, D.A. Plymire, A.N. Chang, J. Kim, E.M. Maresh, S.E. Larson, S.M. Patrie, Sensitive and reproducible intact mass analysis of complex protein mixtures with superficially porous capillary reversed-phase liquid chromatography mass spectrometry, Anal. Chem. 83 (2011) 9586–9592.