Characterization of Complex Proteoform Mixtures by Online Nanoflow Ion-Exchange Chromatography - Native Mass Spectrometry

Ziran Zhai,*[a,b] Despoina Mavridou,[a,b] Matteo Damian,[a] Francesco G. Mutti,[a] Peter J. Schoenmakers,[a,b] Andrea F.G. Gargano*[a,b]

[a] Van’t Hoff Institute for Molecular Sciences, Analytical Chemistry Group and Biocatalysis Group, University of Amsterdam, Amsterdam, The Netherlands.
[b] Centre for Analytical Sciences Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands.

*Corresponding author: Andrea F.G. Gargano; Ziran Zhai
E-mail: a.gargano@uva.nl; z.zhai@uva.nl

https://doi.org/10.26434/chemrxiv-2024-59vrz ORCID: https://orcid.org/0000-0003-3857-3843 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0
Abstract

The characterization of proteins and complexes in biological systems is essential to establish their critical properties and to understand their unique functions in a plethora of bioprocesses. However, it is highly difficult to analyze low levels of intact proteins in their native states (especially those exceeding 30 kDa) with liquid chromatography (LC) - mass spectrometry (MS). Herein, we describe for the first time the use of nanoflow ion-exchange chromatography directly coupled with native MS to resolve mixtures of intact proteins. Reference proteins and protein complexes with molecular weight between 10 and 150 kDa and a model cell lysate were separated using a salt-mediated pH gradient method using volatile additives. The method allowed for low detection limits (0.22 pmol of monoclonal antibodies) while proteins presented non-denatured MS (low number of charges and limited charge state distributions) and the oligomeric state of the complexes analyzed was mostly kept. Excellent chromatographic separations including the resolution of different proteoforms of large proteins (> 140 kDa) and a peak capacity of 82 in a 30-minute gradient, were obtained. The proposed setup and workflows show great potential for analyzing diverse proteoforms in native top-down proteomics, opening unprecedented opportunities for clinical studies and other sample-limited applications.

Keywords: ion exchange, intact proteins, mass spectrometry, nanoflow, non-denaturing analysis.
There is a great diversity in critical functions associated with proteoforms, a term used to describe protein products arising from homologous genes as a result of sequence variations, alternative splicing, and post-translational modifications. Therefore, the accurate, and comprehensive identification of intact proteins and complexes is crucial for fundamental studies of biological processes (e.g. cancer biology). However, the standard approach to characterize proteins, which is bottom-up proteomics, cannot directly identify proteoforms as the presence of proteins is inferred from peptides. Top-down proteomics and intact protein mass spectrometry circumvent these limitations, allowing the direct identification and analysis of intact proteoforms. This approach has shown great potential in characterizing protein mixtures, probing cellular heterogeneity, and unveiling underlying proteoforms biological functions. However, the most common top-down methods use denaturing LC-MS approaches, under conditions in which proteins are unfolded and non-covalent protein complexes are lost. Typically, reversed-phase LC (RPLC) approaches are used, allowing for medium-throughput analysis. However, when analyzing complex protein mixtures such as cell lysates, (i) molecular weight (MW) fractionation prior analysis needs to be performed (e.g. size-exclusion chromatography), and (ii) only the limited proteoforms identification above 30-60 kDa can be obtained.

An alternative to denaturing top-down proteomics is native MS. In native MS, water-based solutions using volatile buffers at neutral pH are employed, operating under conditions where proteins are non-denatured and labile protein complexes can be preserved. To date, the majority of the studies that apply native MS use purified samples and direct infusion with nano-spray sources. As a result, the current native MS analysis has limited applications in the analysis of non-purified complex mixtures. Therefore, there is a pressing need for novel strategies to directly analyze complex mixtures of intact proteins and complexes with a wide range of molecular weights by native LC-MS approaches. Also, it is imperative to achieve high MS sensitivity, given the low abundance of many target proteins in biological samples. Hyphenating native separations to native MS allows the measurement of complex samples, resolving proteoforms according to specific mechanisms (and therefore aiding identification) and increasing the dynamic range of the measurement. Cation-exchange chromatography (CEX) in particular is a non-denaturing separation mode that is considered the benchmark method for separating charge variants of proteins. Minor modifications of proteins often result in changes of iso-electric points (pI), resulting in alterations of their surface-charge distributions and, therefore, influencing the retention in CEX. To obtain highly sensitive detection and detailed characterization of native proteins, CEX can be coupled to high-resolution MS. The use of volatile salts (e.g. ammonium acetate) facilitates direct coupling of CEX to MS. This approach has been successfully applied to the characterization of biotechnological protein products available in relatively large amounts (e.g. injecting several tens of µg of purified biopharmaceuticals) but is not sensitive enough for broad application to biological systems (e.g. top-down proteomics).

In this work, we have developed a method to characterize intact proteins and complexes in their native states, covering a broad MW range with a small sample consumption (down to 33 ng injection for reference monoclonal antibodies). In our approach, we directly coupled nanoflow (250 or 500 nL min⁻¹) strong cation exchange chromatography (SCX) to nano-electrospray-ionization (ESI) under native MS (nMS) conditions. As a complex sample, a cell lysate from a culture of E. coli used to produce recombinant proteins was used. The cell lysis was performed under non-denaturing conditions, buffer-exchanged and concentrated using spin filters and analyzed with
our nanoSCX-nMS approach. Proteins were separated on packed capillary SCX columns and eluted according to their pI values. High contents of volatile salts (up to 250 mM) were exploited in combination with a pH gradient (from pH 5.0 to 8.5) to ensure elution. Such a combined elution strategy is referred as salt-mediated pH gradient.\textsuperscript{[15]} To perform nanoESI-MS, a nano-emitter coated with a hydrophobic material was used.\textsuperscript{[16]} The use of this emitter type was crucial to avoid clogging of the flow path by the salts. The low flow promoted desolvation/ionization efficiency allowing for sensitive detection of low-abundant proteins and complexes. We successfully applied our method to analyze an \textit{E. coli} cell lysate and observed hundreds of proteins with masses up to 150 kDa. We believe that the proposed nanoSCX-nMS is a promising approach for characterizing proteoforms and provides a universal strategy to overcome detection limitations in native top-down proteomics. In particular our method, with respect to state-of-the-art RPLC-MS methods,\textsuperscript{[8b]} it enables (i) the analysis of proteins under non-denaturing conditions, allowing the study of protein complexes and (ii) the identification of complicated protein mixtures covering different MW ranges without the need for MW fractionation.\textsuperscript{[17]} With respect to native methods such as native capillary zone electrophoresis,\textsuperscript{[18]} it permits to achieve high-performance separations for proteins with a wide range of pI values as well as to concentrate the sample before the analysis, thereby increasing the sensitivity of the measurement.

Our research was inspired by the work carried out at analytical scale in the characterization of charge variants of monoclonal antibodies (mAbs) by SCX-MS.\textsuperscript{[14a]} In particular, a SCX column (100 mm × 4.6 mm I.D.) was used to develop an analytical flow (0.4 mL min\textsuperscript{-1}) UV-based separation method. Proteins with different properties, i.e. ribonuclease A (RNase-A), carbonic anhydrase (CA), myoglobin (Myo), bovine serum albumin (BSA), pembrolizumab (Pem), cetuximab (Cet), and trastuzumab (Tra) (for details, see supporting information Table S1), were employed to test different elution conditions and evaluate the performance of the methods. Compared with conventional pH-gradient and salt-gradient approaches, the salt-mediated pH-gradient provides greater separation potential for complex samples.
To develop our method, firstly we performed a set of experiments to optimize the salt-mediated pH-gradient method over a broad pH range, aiming to extend its application to proteins with pI values of 5 and above. To establish optimum conditions, gradients ranging from pH 5.0 to a value between 8.0 and 10.0 were employed, incorporating initial salt concentrations of either 20 mM or 50 mM, and utilizing the elution mobile phase concentrations spanning from 140 mM and 250 mM. Figure S1 indicates that in our experiments a higher final pH (above 9.0) reduces resolution for mixtures of proteins and mAbs, especially for analytes with high pI values (Myo and RNase A). This presumably relates to the non-linear pH gradient that results in an abrupt change of pH above 7, quickly changing the in-solution charge state of the proteins. In contrast, increasing the concentration of salts (pH 8.5) allowed the model...
proteins Myo and RNase-A to be completely separated (see Figure S2). In addition, the slightly higher initial concentration of AmAc (50 mM) yielded a better separation. By monitoring the pH profile (see Figure S3) we succeeded in creating a more-linear pH change zone (over 18 min) that started from 50 mM AmAc at pH 5.0 (adjusted with acetic acid) to 250 mM AmAc at pH 8.5 (adjusted with ammonia). This method was employed for all subsequent measurements.

Next, the analytical column was emptied to collect the packing particles for preparing capillary SCX columns and trap columns. Fused-silica capillaries were end-sealed before use with frits of about 4 mm in length. Next, small amounts of packing materials were packed into the capillaries using a high-ionic-strength solvent (50 mM PBS plus 500 mM Na₂SO₄, pH 7.0) at a pressure of 12 MPa (see Table S2 for details). The homemade capillary SCX columns (150 mm × 75 or 100 µm i.d.) and trap columns (40 mm × 150 µm I.D.) were prepared following identical procedures and then flushed with 20 mM AmAc after packing to displace the non-volatile salts. The nanoSCX-nMS analytical system was then constructed as described in Figure S4. The nano-spray interface in the nanoSCX-nMS system entails non-denaturing conditions by circumventing the desolvation stage with heated nitrogen gas typically present in capillary and analytical flow setups thus limiting protein denaturation. The improved desolvation efficiency of the nano-spray also leads to an enhancement of the sensitivity.[19]

To investigate whether our method was sufficiently sensitive, we analyzed small amounts of BSA (0.1 µg to 1.0 µg injected on column). The MS intensity (Figure S5) of the extracted ion chromatograms over three charge states varied from 2.4E+7 (for 0.1 µg) to 7.0E+8 (for 1 µg). Mixtures of proteins and mAbs (0.1 mg mL⁻¹) were used to evaluate the separation performance of the homemade capillary SCX columns. BSA, CA, and RNase-A (analytes with different pI values and MWs) were well separated and eluted in the order of increasing pI (Figure 1a). Three mAbs (Pem, Cet, Tra) (pI between 7.6 and 9.1) were also separated (Figure 1b), underlining the potential of the method for analyzing a wide diversity of complex biological samples. To further increase the sensitivity of the method, a trap column (150 µm i.d. × 40 mm) was employed to pre-concentrate the low-concentration, high-volume samples. Such focusing is vital to avoid undesired peak broadening and loss of resolution in the nanoflow regime when injecting large volumes (in relation to the separation column volume). Figure 1c shows a comparison of a trap injection of 10 µL of mAbs mixture with a concentration of 0.01 mg mL⁻¹ and a trap-less injection of 1 µL of the same mixture with a concentration of 0.1 mg mL⁻¹. Clearly, using a focusing trap creates possibilities to analyze amount-limited complex samples without sacrificing the resolution.

From the low number of charge states of proteins and mAbs in MS spectra (shown in Figures S6 and S7), we conclude that the proteins remained in their native or non-denatured states during the entire analysis process.[20] To further evaluate the denaturation induced by our method we analyzed model oligomeric proteins: alcohol dehydrogenase (ADH, mainly expressed as tetramer) and L-asparaginase (having a distribution from monomer to octamer). The mild conditions applied retained the complexes during the analysis (Figure S8).
Next, the *E. coli* cell lysate was analyzed and used to evaluate the performance of the trap-based nanoSCX-nMS method. The sample was found to be a complex mixture of proteins with a broad range of MW as seen by denaturing SDS-PAGE (Figure 2a). The approximate concentration measured by Bradford assay was about 9.5 mg mL\(^{-1}\) of which 5 µL were injected (about 50 µg on column, Figure S9). In the total ion chromatogram (TIC, Figure 2b), numerous peaks were distributed across a broad time range (10 min to 37 min) and high MS intensities (up to 8.0E+8) were observed. Several proteins with MWs above 30 kDa were detected, as shown in Figure 2c (peaks 1 through 6). The spread of intensities and retention times indicate a broad range of proteins present in the cell lysates, both in terms of abundance and pI values. The analysis of the average peak width of the 6 proteoforms extracted in Figure 2c (with MW from 30 till 140 kDa revealed) resulted in a peak capacity of about 82 with a gradient of 30 min (Table S4). These proteins retained their native structures as judged by their mass spectra (Figure 3a through 3f). Notably, our method can resolve large proteoforms (around 141 kDa) differing in mass by only 54 Da (Figure 3e and 3f, peaks 5 and 6). Moreover, the protein charge state distributions approach the maximum m/z reachable with the tested mass spectrometer (8000 m/z). Results on fifteen other representative peaks with MW ranging from around 10 kDa to 100 kDa are shown in Figure S10 and Table S3. About 140 proteoforms between 10 and 140 kDa could be observed applying deconvolution and are plotted in Figure 3g and Figure S11.
In conclusion, we developed a novel approach for online nanoflow strong cation exchange chromatography with native MS detection. The method allowed rapid and accurate analysis of intact proteins and complexes under native conditions. The high sensitivity of the method, resulting from a higher desolvation efficiency and a trap-and-elute setup, suggests a great potential for the characterization of low-abundant species. Maintaining the native states of proteins leads to simpler mass spectra with reduced numbers of charge states respect to denaturing MS, which is beneficial for the characterization of complex protein mixtures of with broad MW ranges. In addition, the throughput can be improved as sample-processing steps such as molecular fractionation could be avoided. The nanoSCX-nMS method provides a streamlined and efficacious approach for analyzing protein complexes by native top-down proteomics, with far-reaching applications to help elucidate the roles of proteins in biological processes.

Acknowledgements

The authors gratefully acknowledge the financial support from the Chinese Scholarship Council. They also wish to thank Winfried Roseboom (University of Amsterdam).

References


https://doi.org/10.26434/chemrxiv-2024-59vrz ORCID: https://orcid.org/0000-0003-3857-3843 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0