Anal Chem. Author manuscript; available in PMC 2020 October 09.

Published in final edited form as:
Anal Chem. 2020 October 06; 92(19): 12774–12777. doi:10.1021/acs.analchem.0c03395.

Development of an Online 2D Ultrahigh-Pressure Nano-LC System for High-pH and Low-pH Reversed Phase Separation in Top-Down Proteomics

Zhe Wang, **Dahang Yu**, **Kellye A. Cupp-Sutton**
Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019, United States

Xiaowen Liu,
School of Informatics and Computing, Indiana University-Purdue University Indianapolis, Indianapolis, Indiana 46202, United States

Kenneth Smith,
Arthritis & Clinical Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma 73104, United States

Si Wu
Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019, United States

Abstract

The development of novel high-resolution separation techniques is crucial for advancing the complex sample analysis necessary for high-throughput top-down proteomics. Recently, our group developed an offline 2D high-pH RPLC/low-pH RPLC separation method and demonstrated good orthogonality between these two RPLC formats. Specifically, ultrahigh-pressure long capillary column RPLC separation has been applied as the second dimensional low-pH RPLC separation for the improvement of separation resolution. To further improve the throughput and sensitivity of the offline approach, we developed an online 2D ultrahigh-pressure nano-LC system for high-pH and low-pH RPLC separations in top-down proteomics. An online microtrap column with a dilution setup was used to collect eluted proteins from the first dimension high-pH separation and inject the fractions for ultrahigh-pressure long capillary column low-pH RPLC separation in the second dimension. This automatic platform enables the characterization of 1000+ intact proteoforms from 5 μg of intact *E. coli* cell lysate in 10 online-collected fractions. Here, we have demonstrated that

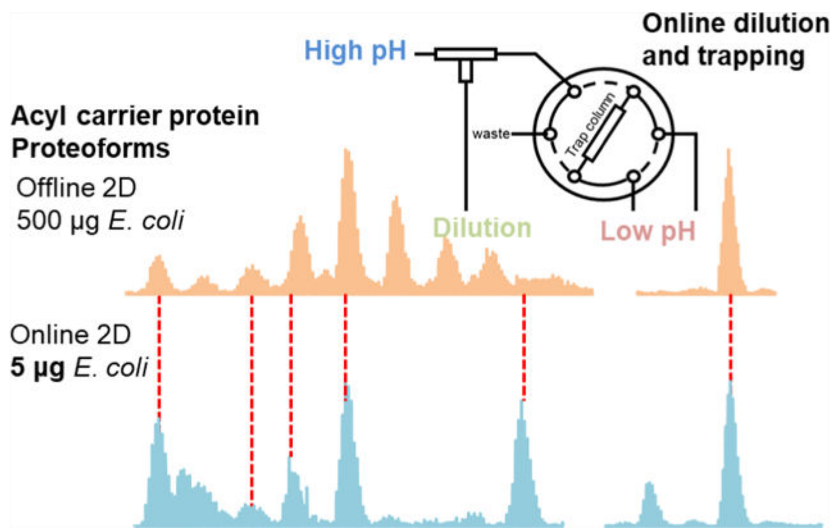
Corresponding Author: Si Wu – Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019, United States; Phone: (405) 325-6931; si.wu@ou.edu; Fax: (405) 325-6111.

Supporting Information
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.0c03395.
Supplementary experimental details; stepwise operation of online 2D nano-LC system and the gradient setup; validation of the dilution factor for 2DLC analysis; online 2D chromatograms of two 2-fraction runs; LC–MS base peak chromatograms of the online 2D-LC analysis with 10 online collected fractions; online 2D pH RP/RPLC–MS analysis of *E. coli proteins*; proteoforms of the apo-acyl carrier protein identified using offline 2D and online 2D pH RP/RPLC–MS; 5-minute segment from 1DLC and online 2DLC; proteoform elution patterns in high pH fractions; and 2D LC gradient schedule (PDF)
Lists of identified proteoforms and unique proteins (XLSX)

The authors declare no competing financial interest.

our online 2D pH RP/RPLC system coupled with top-down proteomics holds the potential for deep proteome characterization of mass-limited samples because it allows the identification of hundreds of intact proteoforms from complex biological samples at low microgram sample amounts.

Graphical Abstract



Top-down proteomics analyzes intact proteoforms resulting from genetic variation, alternative RNA splicing, and post-translational modifications.[1] In contrast to bottom-up proteomics, which is robust and sensitive but provides limited information on proteoforms, top-down proteomics can provide a complete description of the primary structure of a protein and its modifications.[2] However, there are significant challenges posed to top-down proteomics by the complexity of protein samples.[3] Proteomes present a wide dynamic range of thousands of proteoforms, which makes the separation, accurate mass spectrometric detection, and data analysis extremely complicated. One of the major efforts to address these challenges is the improvement of separation resolution prior to MS analysis.

Various separation techniques have been developed to be directly coupled to MS for high-throughput top-down proteomics including RPLC[4–6] and capillary electrophoresis (CE).[7–10] Multidimensional separation platforms have also emerged to improve the separation resolution of intact proteins from complex biological samples based on their orthogonal separation mechanisms.[7,11–20] Recently, our group reported the development of a 2D-LC platform that coupled high-pH RPLC to low-pH RPLC for top-down proteomics. Fractions from the first dimension high-pH RPLC were further separated using an ultrahigh-pressure long capillary column for high-resolution low-pH RPLC-MS analysis. The offline platform was applied to characterize intact proteoforms from *E. coli* lysate[14] and from human *HeLa* lysate.[21] Orthogonal separation by altering the pH conditions of the mobile phases of RPLC was achieved, and the use of RPLC in both dimensions offered higher resolution and better sensitivity than 1D-LC techniques. However, additional sample handling steps in offline 2D-

LC makes the techniques lower-throughput, introduces degradation products, and results in poor sample recovery.

Online 2D-LC has been developed to overcome some limitations of offline 2D-LC.[4,11,12,22,23] Recently, Baghdady et al. reported a proof-of-principle online 2D system that couples high-pH RPLC and low-pH RPLC for intact protein separation.[22] In Baghdady's work, microflow RPLC columns were used in both dimensions and two sample storage loops were used for online coupling. In the second dimension low pH RPLC separation, a very steep gradient (<1 min) was utilized which has been reported to enable the injection of sample volumes up to 70% of the column dead volume without band broadening for peptides and proteins.[22,24] However, the separation resolution of the second dimension RPLC was relatively low, with a peak capacity of 19. This low resolution in the second dimension makes this approach less practical for improving proteoform identification without significantly improved first dimension resolution since online MS detection largely relies on decreased protein complexity eluted from the second dimension RPLC.

Here we report an automatic microtrap column-based online 2D-LC platform using high-pH RPLC and low-pH RPLC separation. Fractions in the first dimension high-pH RPLC were collected into the online microtrap column with dilution. An ultrahigh-pressure long capillary column (i.e., 75 μm i.d., 100 cm length) was used in the second dimension, low-pH RPLC separation, which demonstrated a peak capacity of ~200 in our recent report.[5] A nanoscale UPLC column was also used in the first dimension, high-pH RPLC to improve the sensitivity of the system. To operate these columns, ultra-high-pressure nanoflow pumps are usually required. However, commercially available nanoflow UPLC systems are often expensive and hard to maintain. To provide an alternative, affordable solution, we customized a high-pressure normal flow LC system with two Thermo Accela LC pumps (operational pressure 10 000+ psi, operational flow rate 10–1000 μL/min), two high-pressure six-port switching valves and splitting columns,[5] and one Thermo Accela autosampler. The schematic diagram of the online 2D system is shown in Figure 1 (experimental details in the Supporting Information, Figure S1). In the first dimension, we utilized a custom-packed 10 cm C4 capillary column (Waters BEH300, 3.5 μm, 300 Å, 75 μm i.d.) for high-pH RPLC separation (pH = 10). The eluate from the first dimension was diluted online (1:10) and collected using a microtrap SPE column (Waters BEH300, 3.5 μm, 300 Å, 20 mm × 150 μm). In the second dimension, we utilized a 100 cm C5 column (Jupiter particles, 300 Å, 5 μm, 75 μm i.d.) for low-pH RPLC separations (pH = 2) that were directly coupled to an LTQ Orbitrap Velos Pro mass spectrometer for top-down MS analysis.

To evaluate the trapping efficiency of the microtrap column (e.g., if the organic content was sufficiently diluted and pH was low enough for the second dimension RPLC separation), 5 μg of *E. coli* protein sample was injected onto the first dimension column and then directly eluted using 100% high-pH mobile phase B. Both flow-through and eluate were collected on the trap column. After trapping, the sample was separated using the second dimension low-pH RPLC column for MS detection. The protein elution profile, protein intensity, and protein identification from this run were compared to a 1D low-pH RPLC separation of the same amount of *E. coli* protein sample (Figure S2). These results showed that proteins can be efficiently eluted from the high-pH RPLC column and trapped using the online microtrap for the second dimension low-pH RPLC separation. We further evaluated the reproducibility of the online 2D system with duplicated 2-fraction runs (1st-fraction is flow-through and second fraction is eluate from 0% to 100% buffer B) (Figure S3). The two 2D-LC runs had similar elution profiles in both flow-through and elution experiments.

We then optimized the first dimension fraction collection steps using intact *E. coli* proteins. An initial test of 11 fractions (Figure S4) was evaluated with the following elution steps: 0% for the first fraction, 0% to 10% for the second fraction, and a 10% increase in the concentration of the elution buffer for each subsequent fraction. Analysis of intact *E. coli* cell lysate suggested that the majority of proteins were eluted in the range of 20% to 70% of mobile phase B of the first dimension high-pH RPLC separation with most mass features detected in the fractions from 40% to 70% of mobile phase B. Based on these observations, we combined 0–40% mobile phase B into one online fraction and 70–100% into another online fraction. The optimized 10-fraction gradient setup is shown in Figure S1. The base peak chromatograms (BPCs) of the second dimension low-pH RPLC–MS separation for each fraction are shown in Figure 2 (example fraction 4 in Figure S5) and demonstrate different elution profiles in each fraction. Still, the BPCs of fractions collected after 60% MPB present similar profiles. In the last three fractions, there were only 87 unique proteoforms identified. More precise optimization of gradient setup for different types of samples (e.g., combining the last three fractions into one fraction) may be needed to better utilize the instrument time.

We previously reported the identification of the apo-acyl carrier protein, which is a unique protein working as a coenzyme in fatty acid and polyketide biosynthesis. Due to improved separation using the offline 2D pH RP/RPLC platform, low-abundance proteoforms of the apo-acyl protein were characterized with the phosphopantetheine modification and its unique oxidative forms of the thiol group on this modification. In this study, we were able to characterize most of the low-abundance proteoforms of the apo-acyl protein using the online comprehensive 2D-LC platform (Figure S6). Different proteoform patterns in online and offline analysis may be related to different batches of sample preparation. Also, some differences may be from the offline sample handling process. To further evaluate the improvement of the online 2D-LC platform, we evaluated the 5 min 1D-LC gradient segment and the same retention time segment in each online 2D-LC collected fraction (Figure S7). Nineteen mass features were detected using 1D-LC, and 57 mass features were detected using online 2D-LC in the same 5 min segment across all 10 online fractions. In total, 1 507 nonredundant level 1–3 proteoforms in 308 unique proteins were identified in 5 μg of intact *E. coli* cell lysate with our online 2D separation, which is comparable with our previous reported results using 500 μg of intact *E. coli* cell lysate using the offline 2D separation.[14]

Despite the success of offline 2D high-pH/low-pH RPLC separation in bottom-up proteomics, there has been limited success for online 2D high-pH/low-pH RPLC separation for bottom-up analyses. One major reason is that high-pH RPLC and low-pH RPLC are only partially orthogonal for peptide separations. Therefore, fraction concatenation is required to more efficiently utilize the separation space.[25] In top-down proteomics, we observe good orthogonality between high-pH RPLC separation and low-pH RPLC separation, as we discussed in previous offline studies.[14,21] The improved orthogonality of the top-down 2D platform allows efficient utilization of separation space without fraction concatenation, which makes the online coupling of high-pH RPLC and low-pH RPLC feasible. We further evaluated the orthogonality between high-pH RPLC separation and low-pH RPLC separation using the online platform. For each fraction, a heatmap was generated using the relative number of uniquely identified proteoforms in each bin (10 min windows) (Figure S8). We noticed that most proteoform identification occurred from 110–160 min; however, we also noticed that the detected proteoforms are actually evenly distributed in many fractions (Figure 2). Based on the results, the 2D LC gradient can be further optimized to improve the ID coverage, and the improvement of MS instrumentation can also improve proteoform identification. In addition, the 2D LC gradient can be further optimized to achieve more comprehensive proteoform identification of different types of sample. Overall, our results suggest good orthogonality of the online 2D platform, which is comparable with the results from our previously reported offline 2D platform.[14,21]

While offline fractionation is still the most commonly applied method, online 2D-LC platforms offer the advantage of less sample loss and higher sensitivity, which enables the analysis of small quantity samples. In addition, since the online 2D-LC system is automated and no sample processing between the two dimensions is involved, artificial modifications such as degradation can potentially be reduced. Here, we report the development of an online 2D ultrahigh-pressure nano-LC system for high-pH and low-pH reversed phase separation for the top-down analysis of complex biological samples. With the online 2D system, small sample quantities can be analyzed to achieve comparable proteoform identification results compared with the equivalent offline 2D-LC platform using much higher starting sample amounts. Good orthogonality between first dimension high-pH RPLC and second dimension low-pH RPLC were demonstrated using the online 2D system, which is consistent with our previous reports using the offline 2D system. The performance of the proposed online 2D system can also be further optimized with the incorporation of commercially available nanoflow UPLC pumps; and the incorporation of dual trap columns reduces the MS dead time in between runs through the parallel trapping and elution processes.[26] Additionally, high-performance and high frequency MS instrumentation can be adapted for improved proteoform identification using shorter second dimension gradients and columns with smaller particle sizes. Also, first dimension RPLC separation conditions such as buffer composition and column temperature can be optimized to improve the elution of hydrophobic proteins, which can increase column life and avoid back pressure problems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was partly supported by grants from NIH NIAID R01 AI41625 and NIH NIGMS R01 GM118470.

REFERENCES

(1). Smith LM; Kelleher NL; The Consortium for Top Down Proteomics. Nat. Methods 2013, 10, 186–187. [PubMed: 23443629]

(2). Chait BT Science 2006, 314, 65–66. [PubMed: 17023639]

(3). Chen B; Brown KA; Lin Z; Ge Y Anal. Chem 2018, 90, 110–127. [PubMed: 29161012]

(4). Gargano AFG; Shaw JB; Zhou M; Wilkins CS; Fillmore TL; Moore RJ; Somsen GW; Pasa-Tolic LJ Proteome Res. 2018, 17, 3791–3800.

(5). Wang Z; Liu X; Muther J; James JA; Smith K; Wu S Sci. Rep 2019, 9, 2345. [PubMed: 30787393]

(6). Shen Y; Tolic N; Piehowski PD; Shukla AK; Kim S; Zhao R; Qu Y; Robinson E; Smith RD; Pasa-Tolic LJ Chromatogr A 2017, 1498, 99–110.

(7). McCool EN; Lubeckyj RA; Shen X; Chen D; Kou Q; Liu X; Sun L Anal. Chem 2018, 90, 5529–5533. [PubMed: 29620868]

(8). Zhang ZB; Zhu GJ; Peuchen EH; Dovichi NJ Microchim. Acta 2017, 184, 921–925.

(9). Choi SB; Zamarbide M; Manzini MC; Nemes PJ Am. Soc. Mass Spectrom 2017, 28, 597–607.

(10). Wang Y; Fonslow BR; Wong CC; Nakorchevsky A; Yates JR 3rd Anal. Chem 2012, 84, 8505–8513. [PubMed: 23004022]

(11). Kwiatkowski M; Krosser D; Wurlitzer M; Steffen P; Barcaru A; Krisp C; Horvatovich P; Bischoff R; Schluter H Anal. Chem 2018, 90, 9951–9958. [PubMed: 30014690]

(12). Stoll DR; Harmes DC; Staples GO; Potter OG; Dammann CT; Guillarme D; Beck A Anal. Chem 2018, 90, 5923–5929. [PubMed: 29614857]

(13). Berkecz R; Tomosi F; Kormoczi T; Szegedi V; Horvath J; Janaky TJ Pharm. Biomed. Anal 2018, 149, 308–317.

(14). Wang Z; Ma H; Smith K; Wu S Int. J. Mass Spectrom 2018, 427, 43–51. [PubMed: 31097918]

(15). Zhu Z; Chen H; Ren J; Lu JJ; Gu C; Lynch KB; Wu S; Wang Z; Cao C; Liu S Talanta 2018, 179, 588–593. [PubMed: 29310280]

(16). Chen D; Shen X; Sun L Analyst 2017, 142, 2118–2127. [PubMed: 28513658]

(17). Tran JC; Zamdborg L; Ahlf DR; Lee JE; Catherman AD; Durbin KR; Tipton JD; Vellaichamy A; Kellie JF; Li M; Wu C; Sweet SM; Early BP; Siuti N; LeDuc RD; Compton PD; Thomas PM; Kelleher NL Nature 2011, 480, 254–258. [PubMed: 22037311]

(18). Valeja SG; Xiu L; Gregorich ZR; Guner H; Jin S; Ge Y Anal. Chem 2015, 87, 5363–5371. [PubMed: 25867201]

(19). Xiu L; Valeja SG; Alpert AJ; Jin S; Ge Y Anal. Chem 2014, 86, 7899–7906. [PubMed: 24968279]

(20). Tian Z; Tolic N; Zhao R; Moore RJ; Hengel SM; Robinson EW; Stenoien DL; Wu S; Smith RD; Pasa-Tolic L Genome Biol. 2012, 13, R86. [PubMed: 23034525]

(21). Yu D; Wang Z; Cupp-Sutton K; Liu X; Wu SJ Am. Soc. Mass Spectrom 2019, 30, 2502.

(22). Baghdady YZ; Schug KA Anal. Chem 2019, 91, 11085–11091. [PubMed: 31366196]

(23). Cao JL; Wang SS; Hu H; He CW; Wan JB; Su HX; Wang YT; Li PJ Chromatogr A 2018, 1536, 216–227.

(24). Sarrut M; D'Attoma A; Heinisch SJ Chromatogr A 2015, 1421, 48–59.

(25). Wang Y; Yang F; Gritsenko MA; Wang Y; Clauss T; Liu T; Shen Y; Monroe ME; Lopez-Ferrer D; Reno T; Moore RJ; Klemke RL; Camp DG 2nd; Smith RD Proteomics 2011, 11, 2019–2026. [PubMed: 21500348]

(26). Cao JL; Ma LJ; Wang SP; Deng Y; Wang YT; Li P; Wan JB Anal. Chim. Acta 2019, 1079, 237–251. [PubMed: 31387717]

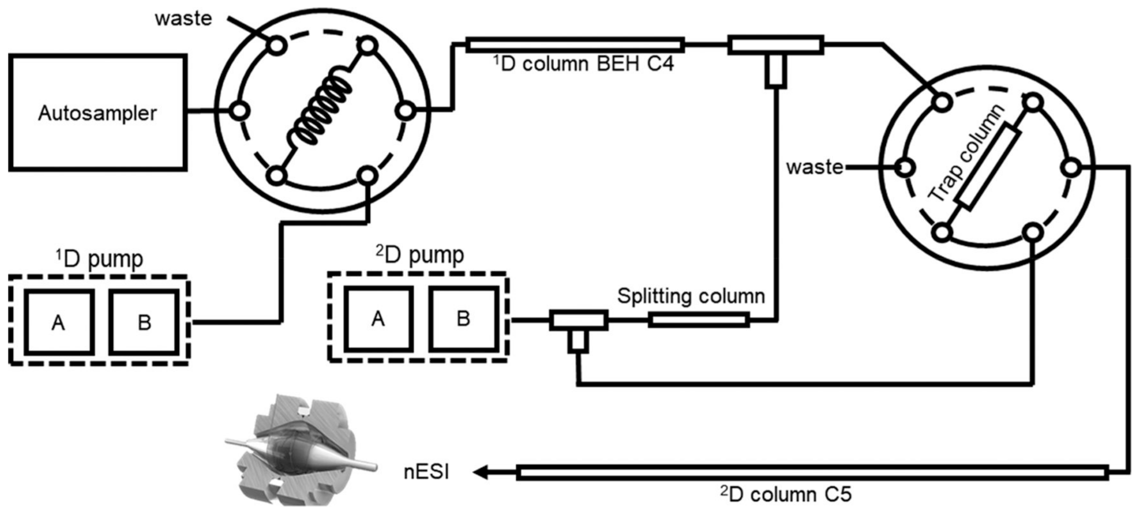


Figure 1.
Schematic diagram of the online comprehensive 2D-LC platform. A 25 μL sample loop was installed on a valve in the autosampler. The ^{1}D pump (high-pH mobile phase) was used to pump the mobile phase for sample loading and first dimension separation. The ^{2}D pump (low-pH mobile phase) was used to provide mobile phase to dilute the elution from the first dimension separation and provide the gradient for the second dimension separation. A splitting column was used to regulate the pressure and flow rate of the SPE sample loading and first dimension separation.

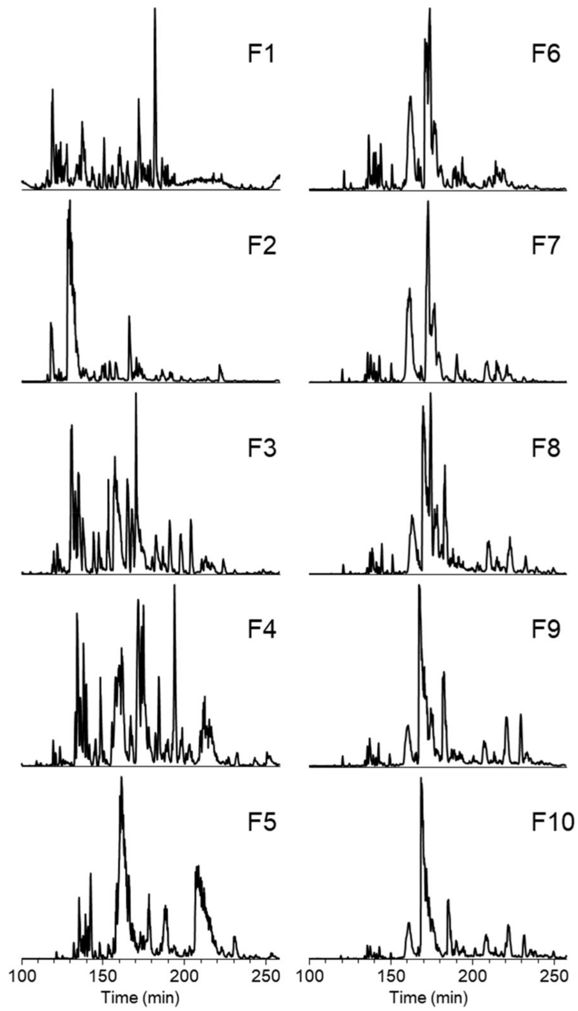


Figure 2.

Optimized online 2D LC analysis for *E. coli* protein separation with 10-fraction elution steps listed in Figure S1. For each fraction, LC–MS base peak chromatograms (m/z 600–1100) are shown.