Anal. Chem. **2009,** *81,* 8708–8714

Integrated Platform for Proteome Analysis with Combination of Protein and Peptide Separation via Online Digestion

Huiming Yuan,[†,‡] Lihua Zhang,[*,†] Chunyan Hou,[†,‡] Guijie Zhu,[†,‡] Dingyin Tao,[†,‡] Zhen Liang,[†] and Yukui Zhang[†]

Key Laboratory of Separation Science for Analytical Chemistry, National Chromatographic R. & A. Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China and Graduation School of Chinese Academy of Sciences, Beijing 100039, China

An integrated platform with the combination of protein and peptide separation was established via online protein digestion, by which proteins were first separated by a microcolumn packed with mixed weak anion and weak cation exchange (WAX/WCX) particles under a series of salt steps, online digested by a trypsin immobilized microenzymatic reactor (IMER), trapped and desalted by two parallel C8 precolumns, separated by microreversed-phase liquid chromatography (μRPLC) under a linear gradient of organic modifier concentration, and finally identified by electrospray ionization-MS/MS (ESI-MS/MS). To evaluate the performance of such a platform, a mixture of myoglobin, cytochrome c, bovine serum albumin (BSA), and α-casein, with mass ranging from 25 ng to 2 μg, was analyzed. Compared to the methods by off-line protein fractionation and shotgun based strategy, the analysis time, including sample preparation, digestion, desalting, separation, and detection, was shortened from ca. 30 to 5 h, and cytochrome c with abundance of 25 ng could be identified with improved sequence coverage. Furthermore, such an integrated platform was successfully applied into the analysis of proteins extracted from human lung cancer cells. Compared with the results obtained by the shotgun approach, the identified protein number was increased by 30%. All these results demonstrated that such an integrated approach would be an attractive alternative to commonly applied approaches for proteome research.

"Top-down" and "bottom-up" approaches are two main analytical strategies for proteome research.[1−3] By the "top-down" approach, intact proteins are usually separated and identified by mass spectrometry (MS).[1,2,4] Since the accurate mass of proteins could be obtained, this approach might be advantageous for identifying translational start and stop sites, mRNA splices variants and post-translational modifications (PTMs) of expressed gene products.[1,5] By the "bottom-up" approach, proteins are first digested into peptides and then separated by multidimensional chromatography, and finally identified by MS/MS.[6−11] Since the separation of peptides is much easier than that of proteins, the "bottom-up" strategy has recently become popular in proteome research.

However, both of the above-mentioned strategies have some shortcomings. By the "top-down" strategy, protein processing is challenging due to typical complications associated with intact protein purification. By the "bottom-up" strategy, the simultaneous separation of all peptides digested from the whole proteome brings great challenges not only to 2D-high performance liquid chromatography (2D-HPLC) separation but also to the identification by MS/MS. To solve this problem, off-line protein prefractionation was performed before peptide separation by HPLC, to decrease the complexity of samples.[12,13] However, most of the off-line approaches suffer from sample loss, time-consuming operation, and difficulty in automation. Therefore, a novel approach combining online protein separation, digestion, peptide separation, and protein identification might be a good solution.

To achieve rapid online digestion of proteins, one key product for the integration of protein processing and peptide analysis has been immobilized enzymatic reactors (IMER), in which proteases

* To whom correspondence should be addressed. Tel: 86-411-84379720. Fax: 86-411-84379779. E-mail: lihuazhang@dicp.ac.cn.

† Dalian Institute of Chemical Physics, Chinese Academy of Sciences.

‡ Graduation School of Chinese Academy of Sciences.

(1) Strader, M. B.; VerBerkmoes, N. C.; Tabb, D. L.; Connelly, H. M.; Barton, J. W.; Bruce, B. D.; Pelletier, D. A.; Davison, B. H.; Hettich, R. L.; Larimer, F. W.; Hurst, G. B. *J. Proteome Res.* **2004**, *3*, 965–978.

(2) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 806–812.

(3) Bogdanov, B.; Smith, R. D. *Mass. Spectrom. Rev.* **2005**, *24*, 168–200.

(4) Kelleher, N. L. *Anal. Chem.* **2004**, *76*, 196A–203A.

(5) Meng, F. Y.; Du, Y.; Miller, L. M.; Patrie, S. M.; Robinson, D. E.; Kelleher, N. L. *Anal. Chem.* **2004**, *76*, 2852–2858.

(6) Fujii, K.; Nakano, T.; Kawamura, T.; Usui, F.; Bando, Y.; Wang, R.; Nishimura, T. *J. Proteome Res.* **2004**, *3*, 712–718.

(7) Zhu, W. H.; Reich, C. I.; Olsen, G. J.; Giometti, C. S.; Yates, J. R. *J. Proteome Res.* **2004**, *3*, 538–548.

(8) Link, J. A.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R. *Nat. Biotechnol.* **1999**, *17*, 676–678.

(9) Boersema, P. J.; Divecha, N.; Heck, A. J. R.; Mohammed, S. *J. Proteome Res.* **2007**, *6*, 937–946.

(10) Motoyama, A.; Xu, T.; Ruse, C. I.; Wohlschlegel, J. A.; Yates, J. R. *Anal. Chem.* **2007**, *79*, 3623–3634.

(11) Dai, J.; Jin, W. H.; Sheng, Q. H.; Shieh, C. H.; Wu, J. R.; Zeng, R. *J. Proteome Res.* **2007**, *6*, 250–262.

(12) Martosella, J.; Zolotarjova, N.; Liu, H. B.; Nicol, G.; Boyes, B. E. *J. Proteome Res.* **2005**, *4*, 1522–1537.

(13) Gao, M. X.; Deng, C. H.; Yu, W. J.; Zhang, Y.; Yang, P. Y.; Zhang, X. M. *Proteomics* **2008**, *8*, 939–947.

10.1021/ac900310y CCC: $40.75
Published on Web 09/29/2009

are covalently bound to, trapped into, or physically adsorbed on different supports, such as membranes,[14] particles,[15−17] and monolithic materials.[18−20] Ye et al. integrated polymer monolith based IMER with capillary electrophoresis (CE) for the analysis of the denatured α-lactalbumin, and more than 20 peaks were detected with the column efficiency for the most peaks above 120 000 plates.[21] Cooper et al. coupled a miniaturized trypsin membrane reactor with transient capillary isotachophoresis/zone electrophoresis for the integrated analysis of proteins, which showed the significant advantages of combining analyte stacking, nanoscale electrophoretic separation, and nano-electrospray (ESI)-MS toward the characterization of low-abundance proteins.[22] Furthermore, Schriemer et al. utilized a particle based IMER to develop a novel integrated platform, in which proteins were separated by microreversed-phase liquid chromatography (μRPLC) and identified by MS/MS after online proteolytic digestion.[23,24] Although compared to off-line protein digestion, high throughput analysis was achieved by these integrated platforms, only one-dimensional separation of peptides or proteins was performed, resulting in insufficient resolving power and peak capacity for proteome study. Recently, Dovichi et al. presented a capillary electrophoresis (CE)-pepsin based microreactor-CE-ESI-MS/MS platform, which could be regarded as a proof-of-principle for a fully automated approach for online protein separation, digestion, peptide separation, and identification.[25] However, protein identification by ESI-MS/MS was limited by the too fast speed and too narrow peak width of CE separation.

In this paper, an integrated platform based on μHPLC separation that combines protein separation by microcolumn ion exchange chromatography (IEC) with mixed weak anion and weak cation exchange (WAX/WCX) particles, online digestion by an IMER, and peptide separation and protein identification by μRPLC-ESI-MS/MS was established. Through the analysis of a four-protein mixture and proteins extracted from human lung cancer cells, the platform demonstrated advantages such as high peak resolution, high protein identification confidence, high sequence coverage, good reproducibility, and ease of automation.

EXPERIMENTAL SECTION

Chemicals and Materials. Peeksil and PEEK tubes were purchased from Upchurch (Oak Harbor, Washington). Myoglobin (horse heart), cytochrome c (horse heart), trypsin (bovine pancreas), lysozyme (chicken egg white), and α-casein (bovine milk) were purchased from Sigma (St. Louis, MO). Bovine serum albumin (BSA) was obtained from Shanghai Milk Company (Shanghai, China). Acetonitrile (ACN, HPLC-grade) was bought from Merck (Darmstadt, Germany). Sodium cyanoborohydride was from Acros Organics (Geel, Belgium). Phenylmethanesulfonyl fluoride (PMSF) was ordered from Shanghai Sangon Biological Engineering Technology and Services (Shanghai, China). RPMI medium 1640 was ordered from Invitrogen Corporation (Carlsbad, CA).

XBP C18 packing material (5 μm, 200 Å) was ordered from Bona Inc. (Tianjin, China). WAX and WCX resins (5 μm, nonporous) were obtained from Sepax Technologies Inc. (Newark, NJ). Acrylic polymer particles with epoxy groups (5 μm, 1000 Å) were obtained from Shenzhen Nanomicro Technology Inc. (Shenzhen, China). C8 macrotrap precolumn (8 × 3 mm I.D.), C8 captrap precolumn (2 × 0.5 mm I.D.), Magic C18 column (50 × 0.3 mm I.D., 5 μm, 200 Å), and strong cation exchange (SCX) MicroBullet cartridge (28 × 1 mm I.D., 10 μm, 300 Å) were ordered from Michrom Bioresources Inc. (Auburn, CA).

Water was purified by a Milli-Q system (Millipore, Bedford, MA). Centrifugal filtration membrane (cutoff molecular weight, 3000 Da) was obtained from Millipore (Bedford, MA). All other chemicals and solvents were analytical grade.

Column Packing. WAX and WCX resins were mixed at the ratio of 3:1 (w/w), slurried in 100 mM sodium chloride, and packed into a Peeksil tube (100 × 0.3 mm I.D.) under a constant pressure of 4000 psi. With the same procedure, a microcolumn packed with WAX resin was prepared as well.

C18 packing material was slurried in bromoform and packed into Peeksil tubes (50 × 0.3 mm I.D. or 100 × 0.3 mm I.D.) under a constant pressure of 6000 psi, to prepare μRPLC columns.

Preparation of IMER. Acrylic polymer particles with epoxy groups (about 40 mg) were added into 25% (v/v) ammonium hydroxide solution and reacted at 40 °C for 3 h. Followed by centrifugation at 1000*g* for 5 min, the supernatant was discarded, and the residual particles were packed into a Peek tube (50 × 0.5 mm I.D.) with the same procedure for packing ion exchange microcolumns. Subsequently, 2.5% (w/v) glutaraldehyde solution was continuously flushed into the column for 1 h using a syringe pump, and then, 5 mg/mL trypsin solution, dissolved in 50 mM borate buffer (pH 8.2), was continuously pumped through the column for 2 h at room temperature. With both ends sealed, the column was left at room temperature overnight for trypsin immobilization. Finally, trypsin immobilized microreactor was treated with 25 mM sodium cyanoborohydride overnight and stored at 4 °C. Before use, IMER was flushed with 50% ACN (v/v) solution, containing 10 mM ammonium acetate, to remove residual unbound trypsin.

Sample Preparation. Standard proteins (cytochrome c, BSA, myoglobin, and α-casein) were, respectively, denatured in 50 mM NH_4HCO_3 buffer (pH 8.0) containing 8 M urea for 1 h at 37 °C in a water bath, followed by dilution with 50 mM NH_4HCO_3 until the concentration of each protein was 1 mg/mL. The denatured proteins were desalted by a C8 macrotrap precolumn, which was first activated with 98% (v/v) ACN containing 0.1% (v/v) TFA and equilibrated with 2% (v/v) ACN containing 0.1% (v/v) TFA at the flow rate of 0.5 mL/min. The samples were first loaded on the C8 precolumn and then washed with 2% (v/v) ACN containing 0.1% (v/v) TFA at 0.5 mL/min for desalting. Finally, proteins were

(14) Gao, J.; Xu, J.; Locascio, L. E.; Lee, C. S. *Anal. Chem.* **2001**, *73*, 2648–2655.
(15) Hsieh, Y. L. F.; Wang, H. Q.; Elicone, C.; Mark, J.; Martin, S. A.; Regnier, F. *Anal. Chem.* **1996**, *68*, 455–462.
(16) Dogruel, D.; Williams, P.; Nelson, R. W. *Anal. Chem.* **1995**, *67*, 4343–4348.
(17) Slysz, G. W.; Schriemer, D. C. *Rapid Commun. Mass. Spectrom.* **2003**, *17*, 1044–1050.
(18) Peterson, D. S.; Rohr, T.; Svec, F.; Fréchet, J.M. J. *Anal. Chem.* **2002**, *74*, 4081–4088.
(19) Xie, S.; Svec, F.; Fréchet, J. M. *J. Biotechnol. Bioeng.* **1999**, *62*, 30–35.
(20) Krenková, J.; Bilková, Z.; Foret, F. J. *Sep. Sci.* **2005**, *28*, 1675–1684.
(21) Ye, M. L.; Hu, S.; Schoenherr, R. M.; Dovichi, N. J. *Electrophoresis* **2004**, *25*, 1319–1326.
(22) Cooper, J. W.; Chen, J.; Li, Y.; Lee, C. S. *Anal. Chem.* **2003**, *75*, 1067–1074.
(23) Slysz, G. W.; Schriemer, D. C. *Anal. Chem.* **2005**, *77*, 1572–1579.
(24) Slysz, G. W.; Lewis, D. F.; Schriemer, D. C. *J. Proteome. Res.* **2006**, *5*, 1959–1966.
(25) Schoenherr, R. M.; Ye, M. L.; Vannatta, M.; Dovichi, N. J. *Anal. Chem.* **2007**, *79*, 2230–2238.

eluted with 1 mL of 80% (v/v) ACN and then dried at low temperature using a Speed vac Concentrator (Thermo-Fisher, San Jose, CA). The above-mentioned procedure was performed on a MAGIC MS4 dual solvent delivery system (Michrom, Auburn, CA).

Human lung cancer cell line H446 (kindly donated by Prof. Shujuan Shao, Dalian Medical University, Dalian, China) was cultured in RPMI medium 1640, supplemented with 10% (v/v) fetal calf serum, and incubated at 37 °C with 5% (v/v) CO_2. Cells were harvested, washed with ice-cold PBS 3 times, and then suspended in buffer composed of 9 M urea and 1 mM PMSF. The suspension was ultrasonicated for 30 s at 130 W and centrifuged at 12 000 rpm for 1 h. Then, the supernatant was collected, and the protein concentration was determined by the Bradford assay. The extracted proteins were reduced and alkylated with protocols mentioned elsewhere.[26] Small molecules, such as urea, DTT, IAA, and PMSF, were removed by filtration with a membrane with a cutoff molecular weight of 3000 Da. Proteins were collected and concentrated using a Speed Vac concentrator. After being desalted on a C8 precolumn with the procedures mentioned above, proteins were redissolved in 10 mM ammonium acetate buffer (pH 8.3) and divided equally for analysis by our new integrated platform and shotgun method.

Protein Digestion. Cytochrome c (100 μg/mL, 1 mL) and proteins extracted from human lung cancer cell line H446 (1 mg/mL, 1 mL) were in-solution digested by adding trypsin into pretreated samples, respectively, with the substrate-to-enzyme ratio of 50:1 and 20:1 (w/w), followed by incubation at 37 °C, respectively, for 12 and 20 h. Finally, 2 μL of formic acid was added to terminate the reaction.

Sample digestion by IMER was performed at 37 °C, by pumping samples into a trypsin immobilized microreactor at a constant flow rate.

Proteins Separation by μIEC. Microcolumn IEC experiments were performed on a MAGIC MS4 dual solvent delivery system. A five-protein mixture, including cytochrome c (1 mg/mL), myoglobin (1 mg/mL), BSA (1 mg/mL), lysozyme (0.6 mg/mL), and α-casein (0.4 mg/mL), with the injection volume of 1.5 μL, was separated respectively by WAX and WAX/WCX microcolumns, under the same conditions, and detected by a UV detector. Tris−HCl buffer (10 mM, pH 8.3) containing 10% (v/v) ACN (A) and A with additional 2000 mM sodium chloride (B) were used as the mobile phases. The applied linear gradient was 0% B (0 min) → 0% B (5 min) → 100% B (105 min), at the flow rate of 5 μL/min, and the eluates were detected by a UV detector at 214 nm. The above-mentioned experiments were repeated in triplicate.

Protein Analysis by an Integrated Platform. A four-protein mixture, cytochrome c (5 μg/mL), myoglobin (0.1 mg/mL), BSA (0.4 mg/mL), and α-casein (0.4 mg/mL), with the injection volume of 5 μL, was used to evaluate the performance of the integrated platform. For μIEC separation, 10 mM and 2000 mM ammonium acetate buffer (pH 8.3) were used to generate four salt steps, including 10, 200, 350, and 1000 mM ammonium acetate, at the flow rate of 3 μL/min. Each fraction was online digested by IMER. The digests were then captured in turn by two C8 parallel precolumns and analyzed by a Magic C18 microcolumn with online MS/MS detection. The mobile phases for peptide separation were 2% (v/v) ACN containing 0.1% (v/v) formic acid (C) and 98% (v/v) ACN containing 0.1% (v/v) formic acid (D). The gradient was set as follows: 0% D (0 min) → 0% D (10 min) → 40 %D (70 min) → 80% D (75 min), at the flow rate of 5 μL/min. After each μRPLC separation, the column was equilibrated with the initial mobile phase.

For the analysis of proteins extracted from human lung cancer cells, 30 μg of sample was injected onto a WAX/WCX microcolumn. For μIEC separation, 10 mM and 2000 mM ammonium acetate buffer, respectively, containing 2% (v/v) ACN (pH 8.3) were used to generate 11 salt steps, including 20, 60, 100, 140, 200, 300, 400, 500, 700, 1000, and 1600 mM ammonium acetate, with a flow rate of 1 μL/min. Each fraction was online digested by IMER. The digests were then captured in turn by two C8 precolumns and separated by a home packed XBP C18 microcolumn with online MS/MS detection. The mobile phases for peptide separation were the same as those for a four-protein mixture analysis, but the gradient was set as follows, 0% D (0 min) → 0% D (10 min) → 40% D (90 min) → 80% D (95 min) → 80% D (100 min), at the flow rate of 5 μL/min. After each μRPLC separation, the column was equilibrated with the initial mobile phase.

Protein Analysis by Traditional Methods. For comparison, the same four-protein mixture, including 25 ng of cytochrome c, 500 ng of myoglobin, 2 μg of BSA, and 2 μg α-casein, was further separated by two commonly used methods. For the off-line approach, the samples were separated by μIEC, off-line collected, in-solution digested (37 °C, 20 h) with a substrate-to-trypsin ratio of 20:1(w/w), and analyzed by μRPLC-ESI-MS/MS. All the experimental conditions for protein and peptide separation were the same as those applied for the integrated platform.

For the shotgun method, the four-protein mixture was in-solution digested (37 °C, 20 h) with a substrate-to-enzyme ratio of 20:1 (w/w), desalted, and further analyzed by 2D-SCX-RPLC-ESI/MS/MS. For peptide desalting, the procedure was the same as that for proteins, except that a C18 precolumn (10 × 4.6 mm I.D.), instead of a C8 column, was used. The collected solution was dried at low temperature using a Speed Vac Concentrator and redissolved in 2% (v/v) ACN containing 0.1% (v/v) formic acid. The solvents for SCX separation were 0.1% (v/v) formic acid (pH 3.5) and 2000 mM ammonium acetate (pH 3.5), respectively, with 2% (v/v) ACN. Seven salt steps, including 0, 40, 100, 200, 300, 500, and 1000 mM ammonium acetate, were used to elute peptides at the flow rate of 50 μL/min by 400 μL of each elution solvent, which were captured in turn by two C8 precolumns, and then separated by a XBP C18 column. The gradient was set as follows, 0% D (0 min) → 0% D (10 min) → 40% D (70 min) → 80% D (75 min) → 80% D (80 min) at the flow rate of 5 μL/min. After each μRPLC separation, the column was equilibrated with the initial mobile phase.

For shotgun method based analysis of proteins extracted from human lung cancer cells, sample was in-solution digested, desalted, and concentrated with the same procedure as described above. After being redissolved in 0.1% (v/v) formic acid containing 2% (v/v) ACN, 30 μg of sample was applied for 2D-SCX-RPLC analysis. The buffers for SCX separation were the same as those applied for analyzing a four-protein mixture. The same 11 salt steps as those for the integrated platform were applied at the flow rate of 100 μL/min by 400 μL of each elution solvent. The eluted peptides were captured in turn by two C8 precolumns and

(26) Dai, J.; Shieh, C. H.; Sheng, Q. H.; Zhou, H.; Zeng, R. *Anal. Chem.* **2005**, *77*, 5793–5799.

separated by a XBP C18 column under the same conditions as those for μRPLC applied in the integrated platform.

MS Identification. A Finnigan LCQ^{Duo} ion trap mass spectrometer (Thermo-Fisher, San Jose, CA) was hyphenated with μRPLC for protein identification. If not specially stated, the spray voltage of MS was 2.0 KV, and the temperature of ion transfer capillary was 150 °C. The MS/MS collision energy was set at 35%. During μRPLC-MS/MS analysis, the eluates were sprayed directly into the ESI source using a homemade interface without sheath or auxiliary gas. All MS and MS/MS spectra were acquired in the data-dependent mode, by which MS acquisition with the mass range of m/z 400−2000 was automatically switched to MS/MS acquisition with the control of Xcalibur software. The two most intense ions of the full MS scan were selected as the parent ions and subjected to MS/MS scan with an isolation width of m/z 2.0. The dynamic exclusion function was set as follows: repeat count, 2; repeat duration, 30 s; exclusion duration, 180 s.

Database Searching. The acquired MS/MS spectra were searched against protein database using Bioworks software (v3.1) with SEQUEST program (Thermo-Fisher, San Jose, CA). Trypsin was set as the enzyme for database searching. Peptides were searched using fully tryptic cleavage constraints and up to two missed internal sites were allowed for tryptic digestion. The mass tolerances were 2 Da for parent masses and 1 Da for fragment masses. For complex samples, cysteine residues were searched as a static modification of 57.0215 Da, and methionine residues as a variable modification of +15.9949 Da. Database searching of standard proteins was performed in horse.fasta database and bovine.fasta database, and the SEQUEST results were filtered by the cross-correlation score (X_{corr}). The peptides were considered positive if X_{corr} values were higher than 1.9, 2.2, and 3.75, respectively for singly, doubly, and triply charged peptides, and $\Delta Cn \geq 0.1$.

Database searching of proteins extracted from human lung cancer cell lines was performed in a combined database, which was a composite of the real human proteins (version 3.17, 60 234 entries) and reverse sequences of proteins, created by precisely reversing the order of the amino acid sequence for each protein. False positive rate (FPR) was calculated using the following equation, $FPR = 2 \times n(rev)/(n(rev) + n(real))$, where $n(real)$ is the number of peptides matched to "real" proteins, and $n(rev)$ is the number of peptides matched to "reverse" proteins, respectively.[27,28] The peptides were considered as positive identification if X_{corr} values were higher than 2.1, 2.65, and 3.75, respectively for singly, doubly and triply charged peptides, with $\Delta Cn \geq 0.25$. A FPR less than 5% was obtained for peptide identifications by the use of the above parameters.

RESULTS AND DISCUSSION

In this work, an integrated platform based on μHPLC separation that combines protein separation on a mixed WAX/WCX microcolumn, online digestion by an IMER, and peptide separation and protein identification by μRPLC-ESI-MS/MS, is to be established.

Protein Separation by Mixed Mode μIEC. To establish an integrated platform, the buffer for protein separation by μIEC and online protein digestion by IMER should be compatible. Therefore, in our experiments, buffers with pH value of 8.3 were chosen for protein separation. Since, in such buffers, both positive and negative charged proteins coexisted, a mixed mode μIEC, with a microcolumn packed with WAX/WCX materials (3:1, w/w), was applied for protein separation.

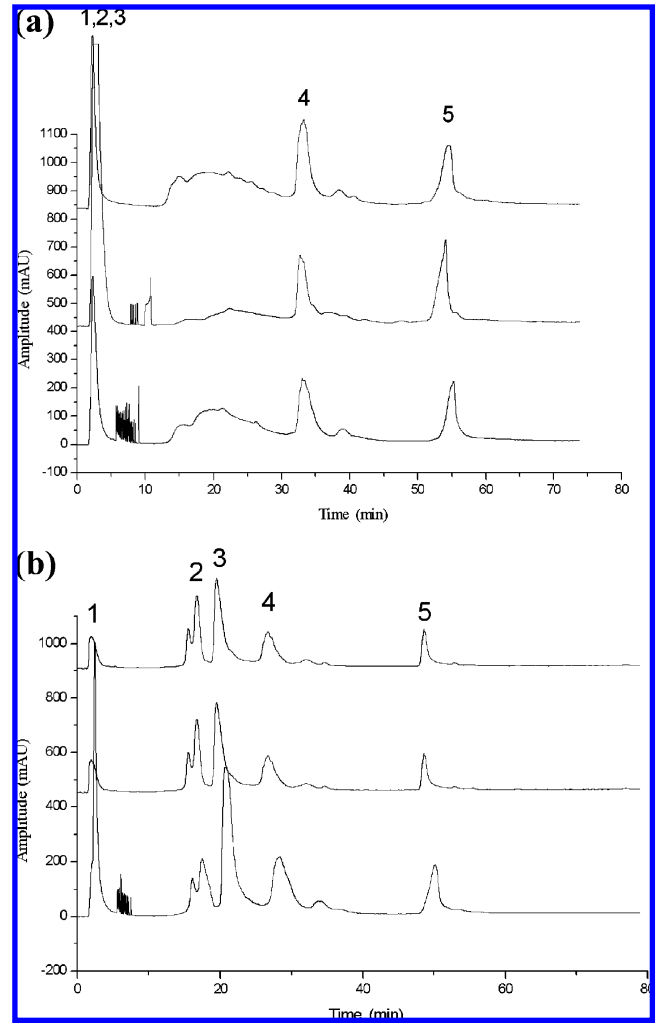

Figure 1. Separation of a five-protein mixture by WAX (a) and mixed WAX/WCX (b) microcolumns in triplicate runs with a UV detector. Experimental conditions were shown in the Experimental Section. Samples: 1, myoglobin; 2, cytochrome c; 3, lysozyme; 4, BSA; 5, α-casein.

To evaluate the performance of WAX/WCX microcolumns, a five-protein mixture was used to simulate a complex sample, among which myoglobin is near neutral, cytochrome c and lysozyme are positively charged, and BSA and α-casein are negatively charged. From Figure 1, it could be seen that, with a WAX microcolumn, BSA and α-casein were separated, while myoglobin, lysozyme, and cytochrome c were coeluted due to the extremely weak interaction with the positive charge on the resin surface. However, the baseline separation of all proteins, identified according to their respective retention time, was achieved with a WAX/WCX microcolumn, due to the improved interaction between samples and the stationary phase with both negative and positive charges. According to the results obtained in triplicate runs, it could be seen that, with separation buffers at pH 8.3, improved resolution of a mixture of acidic, neutral, and basic proteins could be achieved with good

(27) Peng, J. M.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. *J. Proteome Res.* **2003**, *2*, 43–50.
(28) Jiang, X. G.; Feng, S.; Tian, R. J.; Han, G. H.; Jiang, X. N.; Ye, M. L.; Zou, H. F. *Proteomics* **2007**, *7*, 528–539.

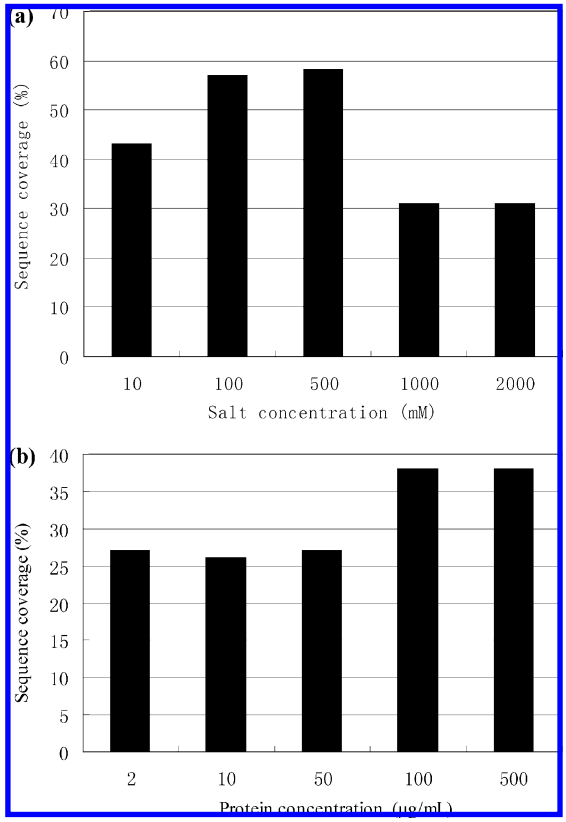

Figure 2. Effect of salt (a) and protein (b) concentration on sequence coverage of cytochrome c by IMER. Experimental conditions were shown in the Experimental Section. Sample concentration: 100 μg/mL (a); 2, 10, 50, 100, and 500 μg/mL (b).

reproducibility by a WAX/WCX microcolumn, compared to that obtained by a WAX microcolumn.

Online Protein Digestion by IMER. In this study, IMER was prepared with trypsin immobilized on acrylic polymer particles with epoxy groups. The enzymatic activity of such IMER was evaluated by the digestion of cytochrome c (0.1 mg/mL), which was pumped through the microreactor at the flow rate of 10 μL/min, and 1 mL of the digests were collected. With 2 μg of digests analyzed by μHPLC-ESI-MS/MS, as shown in Table S-1 in the Supporting Information, sequence coverages of 76% and 79% were obtained, respectively, for on-column and in-solution digestion, indicating that high efficient online protein digestion could be achieved by IMER.

Since proteins eluted by ammonium acetate in μIEC were subsequently online digested by IMER, the effect of ammonium acetate concentration, ranging from 10 mM to 2 M, on the enzymatic activity of IMER was studied by pumping cytochrome c dissolved in a different ammonium acetate buffer with the final concentration of 0.1 mg/mL through the IMER at the flow rate of 3 μL/min. A microliter of the digests from each sample were, respectively, collected, and 2 μg of each sample was further analyzed by μHPLC/MS/MS. From Figure 2a, it could be seen that high salt concentration, greater than or equal to 1 M, could result in decreased enzymatic activity. Although it was evident that salt concentrations of 1 and 2 M resulted in decreased protein sequence coverage clearly to 31% from ~56% (with a salt concentration of 100 and 500 mM), cytochrome c could still be identified with at least four peptides matching, which demonstrated that our developed IMER could endure high salt concentration and, thus, was compatible with μIEC separation.

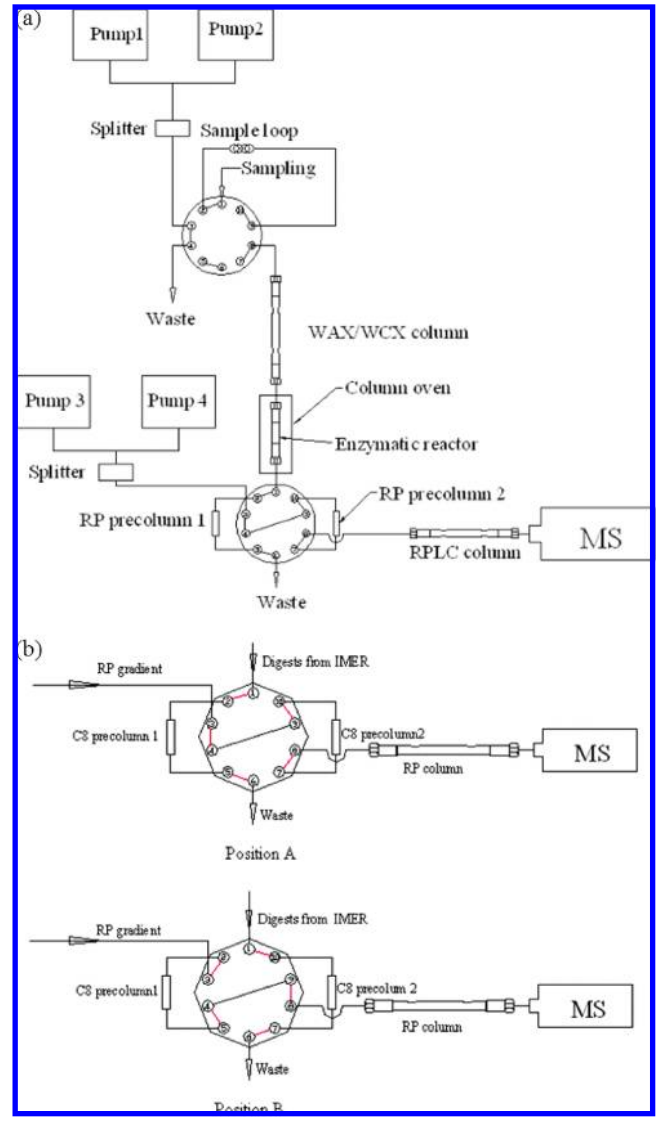

Figure 3. Schematic diagram of integrated platform for proteome analysis that combines online protein separation, digestion, peptide separation, and protein identification (a), and flowchart of two C8 precolumns for peptide trapping and separation (b).

The lifetime of IMER was investigated by the digestion of 50 μg/mL BSA in a consecutive 6 days, and 100 μL of the digests was collected for each day. With 250 ng of the digests analyzed by μRPLC with a UV detector, as shown in Figure S-1a in the Supporting Information, it could be seen that BSA could be digested sufficiently within a 6 day period. In addition, peptides collected in the sixth day were further analyzed by μRPLC-ESI/MS/MS, and the base peak chromatogram was shown in Figure S-1b in the Supporting Information. After database searching, six unique peptides generated from BSA were identified, corresponding to protein sequence coverage of ~16%. These results demonstrated that our developed IMER could be repeatedly used for at least 6 days.

In addition, the applicable protein concentration for IMER was studied by pumping cytochrome c (2 to 500 μg/mL), corresponding to the mass ranging from 2 to 500 μg, through the IMER, at the flow rate of 3 μL/min. With 40 ng of yielded peptides analyzed by

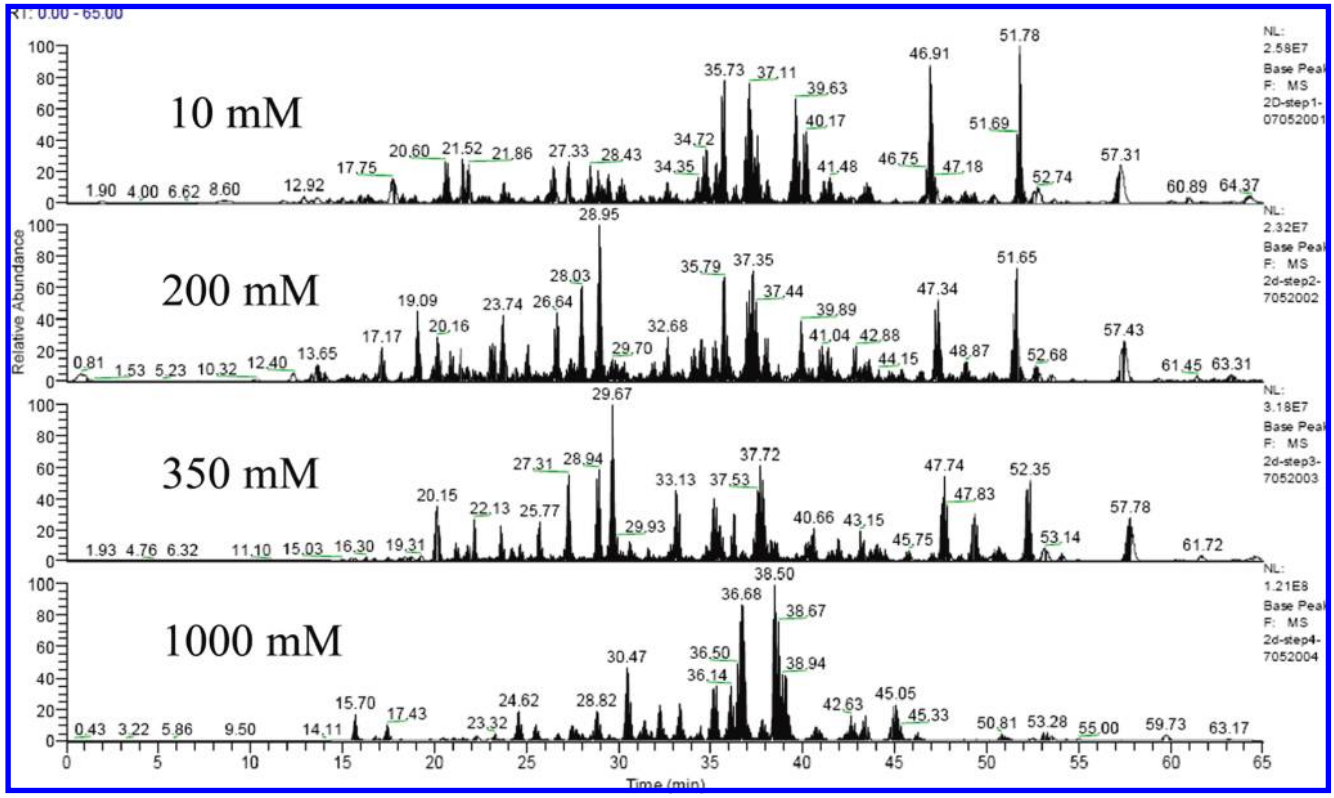

Figure 4. Base peak chromatograms of a four-protein mixture analyzed by the integrated platform, Experimental conditions were shown in the Experimental Section.

μRPLC-ESI-MS/MS, it was shown in Figure 2b that cytochrome c digested within a short residence time could be identified with sequence coverage over 25%, with at least four peptides matching. The enhanced enzymatic activity with trypsin immobilized on supports might be caused by the increased enzyme concentration in fixed space and by the decreased possibility for the autodigestion of enzyme.

Although all these results demonstrated that our developed acrylic polymer particles based IMER was competent in the integrated platform to couple proteins and peptide separation, further effort should be made to reduce protein residual and improve the digestion of low abundance proteins.

Construction of Integrated Platform. The schematic diagram of the integrated platform with the combination of online protein separation, digestion, peptide separation, and protein identification was shown in Figure 3a. During the operation, proteins were first injected onto a WAX/WCX column by switching valve 1 and then separated under a series of salt steps. Subsequently, each fraction eluted from the WAX/WCX microcolumn was online digested by an IMER. Finally, the digests were analyzed by μRPLC-MS/MS. In our experiments, to ensure the time synchronization of proteins and peptide separation, two C8 precolumns were used to in turn capture the protein digests by the control of valve 2, as shown in Figure 3b, to ensure the analysis time of peptides by μRPLC-MS/MS equal to that of proteins separated by a WAX/WCX microcolumn.

Evaluation of Integrated Platform. To evaluate the performance of our integrated platform, a four-protein mixture, including 25 ng of cytochrome c, 500 ng of myoglobin, 2 μg of BSA, and 2 μg α-casein, with p*I*s ranging from 4 to 9, was separated by mixed-mode μIEC into four fractions by four salt steps, including 10, 200, 350, and 1000 mM ammonium acetate, and each fraction was further online digested by IMER, followed by μRPLC-MS/MS analysis, as shown in Figure 4. By database searching, the sequence coverages for myoglobin, cytochrome-c, BSA, and α-casein were respectively 41 ± 3%, 41 ± 3%, 14 ± 2%, and 36 ± 4% in triplicate runs, as shown in Table 1. In addition, the RSDs of the retention time of identified peptides were found to be below 2.9% (as shown in Table S-2 in the Supporting Information). All these results demonstrate the good reproducibility of the established integrated platform.

To compare the performance of integrated platform with other commonly applied methods, the same four-protein mixture was also analyzed by off-line protein prefractionation by μIEC, in-solution digestion, and peptide analysis by μRPLC- MS/MS under

Table 1. Sequence Coverage[a] of Four Proteins Obtained by Three Different Methods in Triplicate Runs

	methods		
protein ID	integrated approach (5 h)[b]	off-line approach (~30 h)[b]	shotgun approach (~28 h)[b]
myoglobin (500 ng)	41% ± 3%	58% ± 9%	41% ± 3%
cytochrome c (25 ng)	41% ± 3%	0	13% ± 1%
bovine serum albumin (2000 ng)	14% ± 2%	19% ± 7%	38% ± 4%
α-casein (2000 ng)	36% ± 4%	36% ± 7%	18% ± 1%

[a] Peptides were considered positive if X_{corr} values were higher than 1.9, 2.2, and 3.75, respectively, for singly, doubly, and triply charged peptides, and ΔCn cutoff values were ≥0.1. Database: bovine.fasta and horse.fasta. [b] The total time for analysis of the four-protein mixture, including digestion, LC separation, and MS detection.

the same separation conditions. By database searching, the sequence coverages for myoglobin, cytochrome c, BSA, and α-casein were respectively 58 ± 9%, 0%, 19 ± 7%, and 36 ± 7% in triplicate runs, as shown in Table 1. It could be seen that although proteins with high abundance were identified with improved sequence coverage, the low abundance one, cytochrome c (25 ng), was not identified. In addition, the reproducibility of sequence coverages was worse than those obtained by integrated platform.

Furthermore, the four-protein mixture was also analyzed by the shotgun method. After database searching, the sequence coverages for myoglobin, cytochrome c, BSA, and α-casein were respectively 41 ± 3%, 13 ± 1%, 38 ± 4%, and 18 ± 1% in triplicate runs, as shown in Table 1. It could have also been that the sequence coverage of low abundance proteins was also worse than that obtained by integrated platform.

In addition, it could be seen that by the above-mentioned three methods, the analysis time for the integrated platform was the shortest.

Analysis of Proteins Extracted from a Human Lung Cancer Cell. The integrated platform was applied for the analysis of proteins (30 μg) extracted from human lung cancer cells, and the results were compared with those obtained by the shotgun method.

The base peak chromatograms obtained by the integrated approach and shotgun method were shown in Figure S-3 in the Supporting Information. By database searching, 1042 and 603 peptides were, respectively, identified, corresponding to 284 and 216 proteins (as shown in Tables S-3 and S-4 in the Supporting Information). Although the identified protein number was limited by the low detection sensitivity of the LCQ^{Duo} Mass spectrometer, obvious improvement on protein identification by integrated platform was achieved.

With the combination of proteins identified by both integrated platform and shotgun method, 104 proteins were identified by both approaches, accounting, respectively, for 37% and 49% of identified proteins, which demonstrated that more unique proteins could be identified by the integrated platform, which might be caused by three reasons. (1) The possibility of sample loss was reduced for the integrated platform, with online protein digestion achieved by IMER. (2) The efficiency of the trypsin digestion was increased by IMER due to the potentially improved trypsin−protein interaction. (3) The protein identification capacity was improved since peptides from fewer proteins were analyzed by μRPLC-MS/MS simultaneously.

A total of 284 proteins identified by the integrated platform were further analyzed, among which there were only 3 (1.05%) proteins with M_w less than 10 kDa, and 39 (13.7%) proteins with M_w more than 100 kDa (as shown in Figure 5a). The smallest and largest proteins were, respectively, of M_W 5.0 and 570 kDa. With the consideration of p*I*, 284 proteins were distributed cross a wide p*I* range from 3 to ~12. As shown in Figure 5b, 74 (26.1%) and 13 (4.57%) proteins were, respectively, distributed between p*I* 3 and 5 and over 10. In addition, 13 proteins (4.57%) were distributed between p*I* 7 and 8 (as shown in Figure 5b). All these results demonstrated that there was no discrimination for protein identification by integrated platform.

In addition, by the integrated platform, the total sample analysis time, including sample preparation, digestion, separation, and detection was shortened to 24 h, about half of that taken by the shotgun method (44 h). Although it would take some time and effort to prepare an IMER, it could be directly coupled with RPLC-ESI-MS/MS for online digestion and repeatedly used for for at least 6 days.

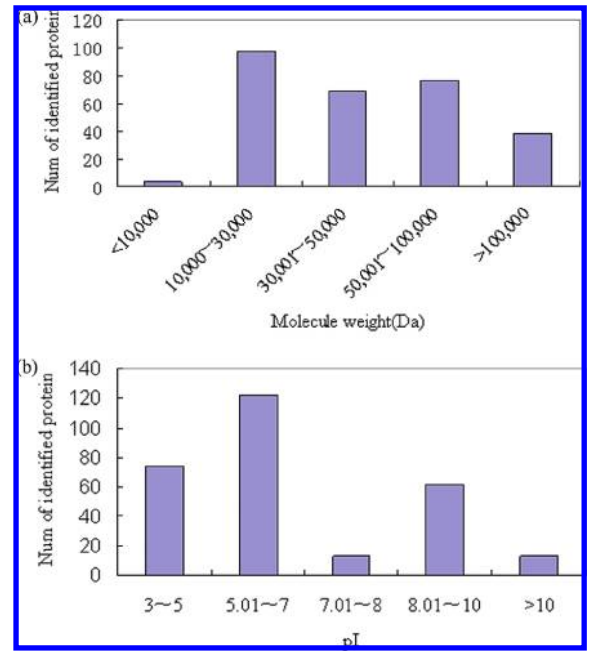

Figure 5. Molecule weight (a) and isoelectric point distribution (b) of proteins identified by the integrated platform.

CONCLUSION

A novel integrated platform based on μHPLC separation that combines protein separation on a mixed WAX/WCX microcolumn, online digestion by an IMER, and peptide separation and protein identification by μRPLC-ESI-MS/MS was established. By comparison with an off-line protein prefractionation method and shotgun strategy, the advantages of improved sequence coverage for low abundance proteins, good reproducibility, high throughput, and ease for automation were demonstrated in the analysis of a four-protein mixture. Furthermore, such a platform was successfully applied into the analysis of proteins extracted from human lung cancer cells. By comparison with the shotgun method, more proteins were identified within almost half the time. All these results demonstrated that such an integrated platform would be an attractive alternative to traditional approaches in proteome research.

ACKNOWLEDGMENT

The financial support from National Natural Science Foundation (20935004), National Basic Research Project of China (2007CB914100), National Key Technology R. & D. Program (2006BAK03A07), and Knowledge Innovation Project of Chinese Academy of Sciences (KJCX2.YW.H09) is gratefully acknowledged.

NOTE ADDED AFTER ASAP PUBLICATION

This manuscript was originally published ASAP on September 29, 2009. A text correction of EEI to ESI was made, and the corrected manuscript was published ASAP on October 5, 2009.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review February 9, 2009. Accepted September 10, 2009.

AC900310Y