Journal of Chromatography B, 905 (2012) 96–104

Contents lists available at SciVerse ScienceDirect

Journal of Chromatography B

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

Protein pre-fractionation with a mixed-bed ion exchange column in 3D LC–MS/MS proteome analysis

Luofu Zhang[a,b], Ling Yao[a,b], Yan Zhang[a,b], Ting Xue[a], Guangchen Dai[a], Keying Chen[a], Xiaofang Hu[a,*], Lisa X. Xu[a,*]

[a] School of Biomedical Engineering and Med-X Research Institute, Shanghai Jiao Tong University, Shanghai 200030, China
[b] School of Life Science and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

ARTICLE INFO

Article history:
Received 9 January 2012
Accepted 6 August 2012
Available online 14 August 2012

Keywords:
3D LC–MS/MS
Mixed-bed IEX chromatography
Protein pre-fractionation

ABSTRACT

The fractionation of complex samples at the protein level prior to shotgun proteomics analysis is an efficient means to more comprehensive analysis of samples. A mixed-bed ion-exchange (IEX) column, packed with both weak anion exchange (WAX) and weak cation exchange (WCX) materials, was used for the first dimensional separation of complex samples at the protein level using volatile solvents. The peptides from digestion of each fraction were then identified by 2D SCX-RP-LC–MS/MS. We applied this 3D strategy to mouse mammary tumor 4T1 cell lysate and identified a total of 3084 proteins in a typical experiment. The moderate separation performance of the mixed-bed IEX column facilitated the in-depth identification of the proteins in the complex sample. There were some acceptable inter-fraction overlaps. Nearly half (45.8%) of the proteins were only identified in single fractions, while 82.3% were identified in no more than 3 fractions. The identified proteins covered a broad range of p*I*, size and grand average hydrophobicity (GRAVY) values. Detailed analysis of proteins identified in each fraction elucidated the separation characteristics of mixed-bed IEX. Retention on mixed-bed IEX was associated, but not restricted to the extreme p*I* values (p*I* < 5, p*I* > 10) and to the percentage of charged residues of both signs. In conclusion, we have exploited the mixed-bed IEX column to establish an efficient and comprehensive identification method for complex samples.

1. Introduction

In proteomics research, the efficient fractionation of complex samples is critical for the comprehensive identification of proteins [1–3]. In the strategy of shotgun proteomics, complementary modes of chromatography are used to distribute peptides into subsets of a size manageable by mass spectrometry. This is an important way to improve the identification depth of complex samples [2,4,5]. Most commonly, strong cation-exchange (SCX) and reversed-phase (RP) chromatography are used to separate peptides by electrostatic and hydrophobic differences in MudPIT (multi-dimensional protein identification technology) strategy, respectively [2,5]. For more extensive identification of proteins in complex samples, some researchers have extended peptide fractionations from 2D (two-dimensional) LC to 3D [6,7] or even 4D LC [8]. However, it should be noted that with these strategies, peptides produced from proteins of high abundance are distributed throughout the fractions even after multiple dimensions of separation, shielding the effective detection of peptides from proteins of low abundance. Thus the addition of LC separation at the peptide level has limited its improvement of the identification depth of complex samples [9].

Pre-fractionation of complex samples at the protein level, followed by fractionation of peptides with the usual 2D LC–MS/MS method, is a powerful way to improve identification [9–11]. Proteins of high abundance can be collected in discrete fractions, avoiding the shielding effect of their peptides on peptides from other fractions. Moreover, compared to peptides, proteins are more diverse in physicochemical characteristics, such as sequence, size, shape-conformation, electrostatic state, p*I* and polarity. This provides the basis for better selectivity. Among various chromatographic modes, ion-exchange (IEX) chromatography has been widely used in protein fractionation for its high resolution and compatibility with biochemical samples [12–16]. However, anion-exchange (AX) or cation-exchange (CX) columns are each biased in adsorbing and separating proteins toward one end of the p*I* range. Proteins in the rest of the range may not be well-retained and separated [10,17,18].

* Corresponding authors at: School of Biomedical Engineering and Med-X Research Institute, Shanghai Jiao Tong University, 1954 Hua Shan Rd., Shanghai 200030, China. Tel.: +86 21 62933234; fax: +86 21 62932302.
E-mail addresses: xfhu_sjtu@163.com (X. Hu), lisaxu@sjtu.edu.cn (L.X. Xu).

1570-0232/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jchromb.2012.08.008

To address this problem, El Rassi et al. designed mixed-bed IEX columns, with a 1:1 mixture of SAX (strong anion-exchange) and SCX material, to achieve better adsorption and separation of protein samples [19]. Using a standard mixture of 7 proteins with large p*I* diversity, the mixed-bed IEX column outperformed traditional IEX columns under similar chromatographic conditions by achieving better selectivity and wider p*I* separation. In recent years, several groups have used mixed-bed IEX columns of WAX (weak anion-exchange) and WCX (weak cation-exchange) materials to separate complex samples at the protein level, followed by separation and detection of the resulting peptides by 1D RPLC–MS/MS. In Lee's work [20], 1733 and 1355 proteins were identified from *Escherichia coli* lysate and HeLa cell mitochondrial lysate, respectively. In Le's work [21], a total of 293 proteins were found in both bovine colostral and mature milk, among which 176 proteins had not been detected before with other methodologies. These works showed the potential of mixed-bed IEX columns as an effective approach to fractionation of complex samples, facilitating the detection of more proteins.

The present work aims to establish an effective 3D LC–MS/MS methodology for in-depth identification of complex samples by using mixed-bed IEX chromatography as the first dimension to fractionate complex samples at the protein level, followed by the usual online 2D SCX-RP-LC–MS/MS sequence for peptide separation and detection.

2. Materials and methods

2.1. Cell lines and reagents

The 4T1 (mouse mammary tumor) cell line was obtained from American Type Culture Collection (ATCC). Dithiothreitol (DTT), iodoacetamide (IAA), Tris-base, ammonium bicarbonate (NH_4HCO_3), ammonium acetate (NH_4OAc), formic acid (HCOOH), sodium dodecyl sulfate (SDS), acrylamide (Acr), bis-acrylamide (Bis), N,N,N′,N′-tetramethylethylenediamine (TEMED) and ammonium persulfate (AP) were purchased from Sigma–Aldrich (St. Louis, MO). Protease inhibitor cocktail was obtained from Roche (Mannheim, Germany). Sequencing grade modified trypsin was obtained from Promega (Madison, WI). Acetone, ethanol and acetonitrile were purchased from TEDIA (Fairfield, OH). The cell culture cupboard and the Snakeskin Pleated Dialysis Tubing (3500 MWCO) were purchased from Thermo (San Jose, CA). The centrifuge equipments were Allegra X-22R and Avanti J-26 XPI Centrifuge from Beckman Coulter Inc. (Fullerton, CA). Amicon 10 kDa centrifugal filter units were obtained from Millipore (Bedford, MA). All of the water used in the experiment was prepared using a Millipore Elix water purification system.

2.2. Cell culture and protein sample preparation

4T1 cells were cultured in RPMI 1640 medium (Gibco, Grand Island, NY) supplemented with 10% fetal bovine serum, penicillin (40 kU/L) and streptomycin (40 mg/L) at 37 °C in 5% CO_2. When cells were grown to approximately 90% confluence in 10 cm culture dishes, 4T1 cells were collected and washed 3 times with 10 mM PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.4 mM KH_2PO_4, pH 7.4) to remove culture medium. The collected cells were lysed in hypotonic buffer (69 mM NaCl, 5 mM Na_2HPO_4, 0.7 mM KH_2PO_4, 1.5 mM $MgCl_2$, 10 mM KCl, protease inhibitor cocktail (1:100, v/v)) at the ratio of 1:200 (1 dish of cells:200 μL lysis buffer), fractured by freezing and thawing 4 times in liquid nitrogen and 37 °C, and then sonicated at 190 W for 5 min. The cell lysate was centrifuged at 60,200 × *g* for 30 min, and the supernatant was collected and dialyzed in 10 mM ammonium acetate solution (pH 7.0, 1:1000, v/v) overnight at 4 °C for buffer changes before protein pre-fractionation. Protein concentration of dialyzed cell lysate was measured with the Bradford assay (Shanghai Sangon Biological Engineering Technology Inc., Shanghai, China). The cell lysate was then added with Protein Inhibitor Cocktail (1:100, v/v) before being stored at −80 °C.

2.3. Proteins pre-fractionation on the mixed-bed IEX column

Ion exchange HPLC fractionation was performed with an Agilent 1200 HPLC system. 2.5 mg of protein sample of cell lysate (250 μL) was injected into a mixed-bed ion exchange column (PolyCAT A™ and PolyWAX LP™ 1:1, column dimension 4.6 mm i.d. × 200 mm, particle size 5 μm, pore size 1000 Å, PolyLC Inc., Columbia, MD). Mobile phases A and B were 10 mM and 1 M ammonium acetate (pH 7.0), respectively. The gradient started with 0% phase B, up to 12% B in 10 min, 60% B in 25 min, and then 100% B in 38 min and maintained at 100% B until the absorbance returned to the baseline. UV absorbance was monitored at 280 nm. The column was eluted at 1 mL/min flow rate and column effluent was collected into 16 × 4 min fractions from 2nd minute to 66th minute. The column was re-equilibrated with phase A for 40 min before the next run. The fractions were then lyophilized at 25 °C to remove the volatile ammonium acetate. The above-mentioned protein level fractionation procedure of 2.5 mg of protein sample was repeated four times, three for chromatography repeatability validation and one for Bradford protein assay of each fraction. Another set of cell lysate, containing 0.8 mg of protein sample of 4T1 cell lysate, was also separated and 16 fractions were collected for SDS-PAGE analysis.

2.4. SDS-PAGE analysis of fractions from mixed-bed IEX chromatography

0.8 mg of protein sample of 4T1 cell lysate was loaded onto the mixed-bed IEX column and fractionated with the above-mentioned process. Ammonium acetate in collected fractions was removed by lyophilization. 16 fractions of dried protein samples were then analyzed on the 12% SDS-PAGE gel in parallel. The gels were stained by Coomassie Blue, followed by extensive decolorization and image capturing (Image Quant 150, GE healthcare, Piscataway, NJ).

2.5. Protein concentration assay of fractions from mixed-bed IEX chromatography

A 2.5 mg of cell lysate sample was fractionated with mixed-bed IEX chromatography, lyophilized and re-dissolved in re-suspension buffer (8 M urea, 65 mM DTT, 40 mM Tris–HCl, pH 7.4), with gentle shake for 30 min. The protein concentration of each fraction was then measured with the Bradford protein assay.

2.6. In-solution digestion

The procedure of protein in-solution digestion was described in [22] with small adaptation. The lyophilized proteins were dissolved and reduced in re-suspension buffer containing 8 M urea, 65 mM DTT, 40 mM Tris–HCl (pH 7.4) and protease inhibitor cocktail (1:100, v/v) at 37 °C for 2.5 h. Alkylation was performed with 325 mM IAA for 1 h. After the proteins were precipitated overnight by pre-cooled acetone (1:6, v/v, sample:acetone) and centrifuged at 22,000 × *g* for 30 min, the pellets were washed with 70% ethanol followed by air-drying. The protein pellets were further digested with trypsin (0.1 μg/μL) at a ratio of 1:50 (w/w, trypsin:protein) in 50 mM ammonium bicarbonate (pH 7.8) for 20 h. The peptides were ultrafiltered with 10 kDa centrifugal filter units (Millipore) and lyophilized for further analysis.

2.7. Online 2D SCX-RP-LC–MS/MS

Online two-dimensional LC–MS/MS experiments were performed on an MDLC system (Michrom BioResources Inc., Auburn, CA) coupled with a Thermo Finnigan 2D linear ion trap mass spectrometer (LTQ^{XL}, Thermo Inc., San Jose, CA). Each peptide sample was re-dissolved in 60 μL 2% (v/v) acetonitrile with 0.05% (v/v) formic acid and then 50 μL of this peptide solution was injected and loaded onto a Strong Cation Exchange Trap column (MicroTrap^{TM}, 1 mm × 8 mm, Michrom BioResources Inc.) with the autosampler of the MDLC system. Eight ammonium formate steps of 0 mM, 10 mM, 20 mM, 50 mM, 100 mM, 200 mM, 500 mM, and 800 mM were used to wash the SCX column in sequence at a flow rate of 10 μL/min. Each eluted peptide fraction was further desalted and concentrated with a Peptide Captrap column (CapTrap^{TM}, 0.5 mm × 2 mm, Michrom BioResources Inc.) with 2% (v/v) acetonitrile with 0.1% (v/v) formic acid at a flow rate of 10 μL/min for 10 min.

The desalted peptide fractions were further separated on a C18 capillary column (0.1 mm i.d. × 150 mm, 3 μm, 200 Å; Michrom BioResources Inc.) with a continuous gradient. The flow rate was 370 nL per minute. Mobile phase A was 0.1% (v/v) formic acid in water and phase B was 0.1% (v/v) formic acid in acetonitrile. The 128-min elution gradient started at 2% B for 10 min, reached 35% B in 120 min, then 80% B in the next 2 min, held at 80% B for 2 min and returned to 2% B for the next run. LTQ mass spectrometer (Thermo) with a nano-ESI source (ADVANCE; Michrom BioResources Inc.) were used to acquire ESI/MS/MS spectra in data-dependent acquisition mode. The MS scan utilized an *m/z* range of 400–1800. The temperature of the heated capillary was set at 200 °C and spray voltage was 1.2 kV. The mass spectrometer was set as one full MS scan followed by ten MS/MS scans on the ten most intense ions from the MS spectrum with the following dynamic exclusion settings: repeat count = 2, repeat duration = 15 s, exclusion duration = 30 s.

2.8. Direct analysis of cell lysate by online 2D SCX-RP-LC–MS/MS

Cultured 4T1 cells were collected, washed and then lysed in lysis buffer (1 dish of cells:1 mL lysis buffer) containing 8 M urea, 20 mM Tris–HCl (pH 7.4) and protease inhibitor cocktail (1:100, v/v) with gentle mixing for 1 h. After centrifugation at 60,200 × *g* for 30 min, the supernant was collected and protein concentration was measured with the Bradford assay. 4T1 cell lysate (500 μg of protein included) was added with 2 μL of 1 M DTT, and incubated at 37 °C for 2.5 h. Then 10 μL of 1 M IAA was added and incubated with the mixture for 1 h at room temperature in darkness. The protein sample was acetone-precipitated and peptides were prepared as Section 2.6 described through digestion, ultrafiltration and lyophization. The peptides were then directly loaded onto the online 2D LC–MS/MS system, in which peptides were separated on SCX and RP columns and detected via ESI-MS/MS as described in Section 2.7.

2.9. Peptide and protein identification

All data files were created by searching MS/MS spectra against the Mouse International Protein Index protein sequence database (IPI.MOUSE.v3.72.fasta), using the TurboSEQUEST program in the BioWorks 3.3 software suite, with a precursor-ion mass tolerance of 2.0 amu and fragment-ion mass tolerance of 0.8 amu. Trypsin was set as the protease with two missed cleavage sites allowed. Carbamidomethylation (+57.02150 Da) was searched as a fixed modification on cysteine, representing alkylation with IAA, while oxidized methionine (+15.99492 Da) was searched as the variable modification. The searched peptides and proteins were validated by PeptideProphet and ProteinProphet in the Trans-Proteomic Pipeline (TPP, v4.2) using default parameters. Proteins with ProteinProphet P value greater than 0.9 and with no less than one unique peptide were considered as true identifications. A randomized database of the IPI.MOUSE.v3.72.fasta was used as a decoy database to calculate the false discovery rate (FDR) of protein identification. The FDR was calculated by the ratio of the number of matches to the randomized database to the combined number of matches to the IPI.MOUSE.v3.72.fasta and its randomized derivative. FDR for Prophet 0.9 was less than 1%. Proteins containing the same peptides were grouped, and only one protein with the highest probability in each group was retained.

2.10. Data analysis

After spectra data were treated by SEQUEST and TPP, 16 lists of protein identification were generated corresponding to all peptide fractions run via MudPIT. Each protein in the lists was identified by at least one unique peptide. With a homemade protein merging software, total protein identifications in all fractions were compiled by pooling all identities and deleting redundant ones among fractions.

The inter-fraction protein overlap rate equals the total non-redundant proteins identified in two fractions divided by the overlapping proteins identified in the same two fractions, per the following formula:

$$\text{Overlap rate} = \frac{\text{Overlapping proteins between fractions A and B}}{\text{Proteins of fractions A+B} - \text{Overlapping proteins between fractions A and B}}$$

Physico-chemical properties, such as p*I*, molecular weight and GRAVY value, of proteins identified were calculated by online proteomics tools. Among these characteristics, molecular weight and p*I* were calculated on Expasy Proteomics Server (http://expasy.org/tools/pi_tool.html) and the GRAVY value was assessed by online GRAVY Calculator (http://www.gravy-calculator.de/index.php?page=file).

3. Results

3.1. 3D LC–MS/MS strategy

In our present work, we applied the mixed-bed IEX column for the first dimension of LC separation of intact proteins, with the combination of the widely used online 2D SCX-RP-LC–MS/MS method for subsequent peptide separation and identification. The workflow of this 3D strategy is displayed in Fig. 1. The lysate of a biological sample (here, a 4T1 cell line) was prepared and dialyzed. The dialyzed sample was loaded onto the mixed-bed IEX column and distributed into 16 fractions with a two-segment linear gradient of increasing ammonium acetate concentration. After lyophilization, the proteins in the fractions were re-dissolved, reduced, alkylated and trypsin-digested in parallel to prepare peptides. In the subsequent analysis, every peptide fraction was further separated into 8 sub-fractions by SCX chromatography and analyzed by RP-LC–MS/MS. The MS/MS spectra acquired were searched against the IPI mouse database using the SEQUEST algorithm and peptide and protein identifications were validated with the Prophet tools. Finally, protein identifications in all fractions were pooled by deleting redundant identifications in different fractions, resulting in the non-redundant list of total protein identifications.

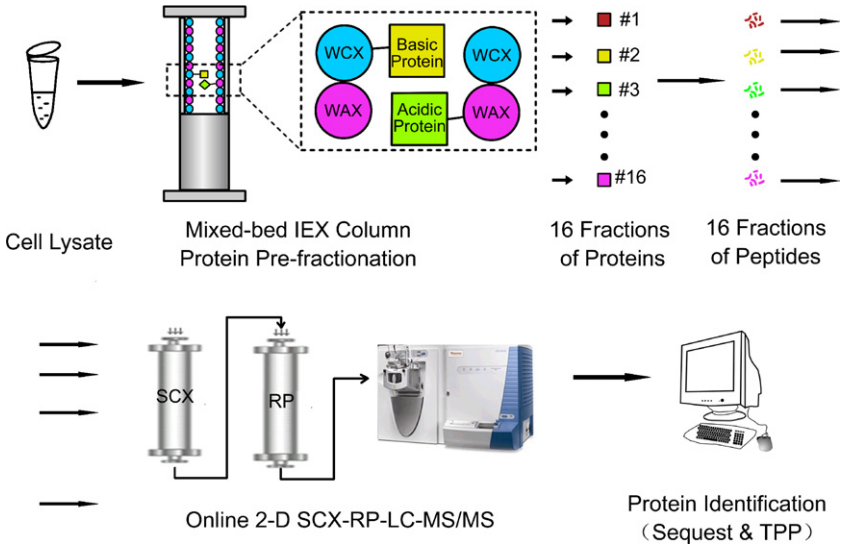

Fig. 1. Scheme of the 3D LC–MS/MS protein identification system. 4T1 cells lysate, the model of complex sample, was fractionated at protein level using a mixed-bed IEX column. The resulting protein fractions were digested individually. The peptides were then separated and identified via online 2D SCX-RPLC–MS/MS. MS/MS spectra generated were searched against the IPI.MOUSE database using the SEQUEST algorithm to identify peptides and proteins, which were validated by PeptideProphet and ProteinProphet in the TPP.

3.2. Performance of mixed-bed IEX chromatography to fractionate cell lysate

To establish an efficient 3D LC–MS/MS method, the variables needed to be optimized. NaCl or $NaClO_4$ gradients have been used to fractionate proteins on mixed-bed IEX columns. The fractionated samples must then be desalted by precipitation or dialysis before protein digestion. This increases the amount of work involved and possible loss of sample [20,21]. To improve the sample recovery rate, we replaced NaCl and $NaClO_4$ with ammonium acetate, a volatile salt that could be removed by lyophilization, minimizing the additional work amount. Moreover, we optimized other HPLC parameters such as salt concentrations and gradient to elute all proteins in sample with even distribution along a 50-min gradient. After testing and comparison, 1 M ammonium acetate of phase B was proved to be more suitable than 800 mM of that, which was the advised parameter of the column but not powerful enough to elute all proteins in our complex sample. On the other hand, the advised gradient of "0–10 min, 0–12% B; 10–30 min, 12–60% B; 30–40 min, 60–100% B" was gentle in the 10–30 min period but steep in the 30–40 min period, and thus we adjusted the gradient into "0–10 min, 0–12% B; 10–25 min, 12–60% B; 25–38 min, 60–100% B; 38–70 min, 100% B".

Fig. 2 shows a representative chromatogram of 4T1 cell lysate proteins eluted from the mixed-bed IEX column. Dozens of peaks (more than 40 here) were evenly distributed throughout the empirically optimized salt gradient, a satisfactory separation. While a very small group of proteins were not retained, resulting in the peak in the void volume fraction 1, most components were retained and fractionated with the gradient. With time intervals of 4 min, the eluted proteins were collected into 16 fractions until the UV absorbance signal returned to baseline. To ascertain if all proteins had been eluted, the protein concentrations in a set of 16 fractions were measured by the Bradford assay after lyophilization and re-suspension. The protein recovery of fraction 1–12 was satisfactory (~98%) (Table S1).

To evaluate the repeatability of the mixed-bed IEX chromatography, three replicates of 4T1 cell lysate were run on the column via the above method. The UV absorbance trace of the three runs superimposed well (Fig. 2B), an indication of acceptable repeatability with complex samples.

To briefly assess the separation performance of mixed-bed IEX chromatography, protein fractions were analyzed by SDS-PAGE. 4T1 cell lysate (0.8 mg protein) was fractionated and the 16 resulting fractions were individually dried, re-suspended and loaded onto SDS-PAGE gels. From the gel profiles (Fig. 3), lanes 1–12 exhibited large differences in distribution of bands, implying good fractionation of the complex sample. The samples of lanes 13–16 were collected with isocratic 1 M ammonium acetate. These lanes showed light bands and similar distribution to lane 12, suggesting that most proteins had been fractionated effectively in fractions 1–12.

3.3. Application of mixed-bed IEX chromatography in 3D LC–MS/MS

The 16 fractions from the mixed-bed column were digested individually and subjected to SCX-RPLC–MS/MS. 8 sub-fractions were collected from the SCX column by step salt elution and further separated by a 128-min gradient on an RP capillary. Using MS data processing with bioinformatics tools, we identified proteins in each fraction (Table 1). The protein identification data files of all fractions were pooled by depleting redundant proteins (same IPI number) to generate a total list of non-redundant proteins. In one typical 3D LC–MS/MS experiment, with the criteria of unique peptides ≥ 1, in the first 12 fractions we identified 3084 proteins, with 2612 proteins identified via at least 2 unique peptides. Few new proteins were found in fractions 13–16 (Fig. S1), so only data for fractions 1–12 were analyzed in subsequent research.

To demonstrate the efficacy of mixed-bed IEX pre-fractionation, 4T1 cell lysate was also digested and analyzed directly by 2D SCX-RPLC–MS/MS. In a typical experiment, 1292 proteins were identify with unique peptides ≥ 1, and among them 1073 ones were identified with at least 2 unique peptides. Proteins identified via 2D and 3D strategy were compared, as shown in Fig. 4. The 3084 identified proteins of 3D strategy covered 1112 proteins (86%) that were found by 2D strategy, and included another 1972 proteins at the same time, indicating that the pre-fractionation of proteins by the mixed-bed IEX column was critical for in-depth identification. The 180 proteins only identified in 2D LC–MS/MS had an average spectra count of 10.2, which was much lower than that of the total 1292

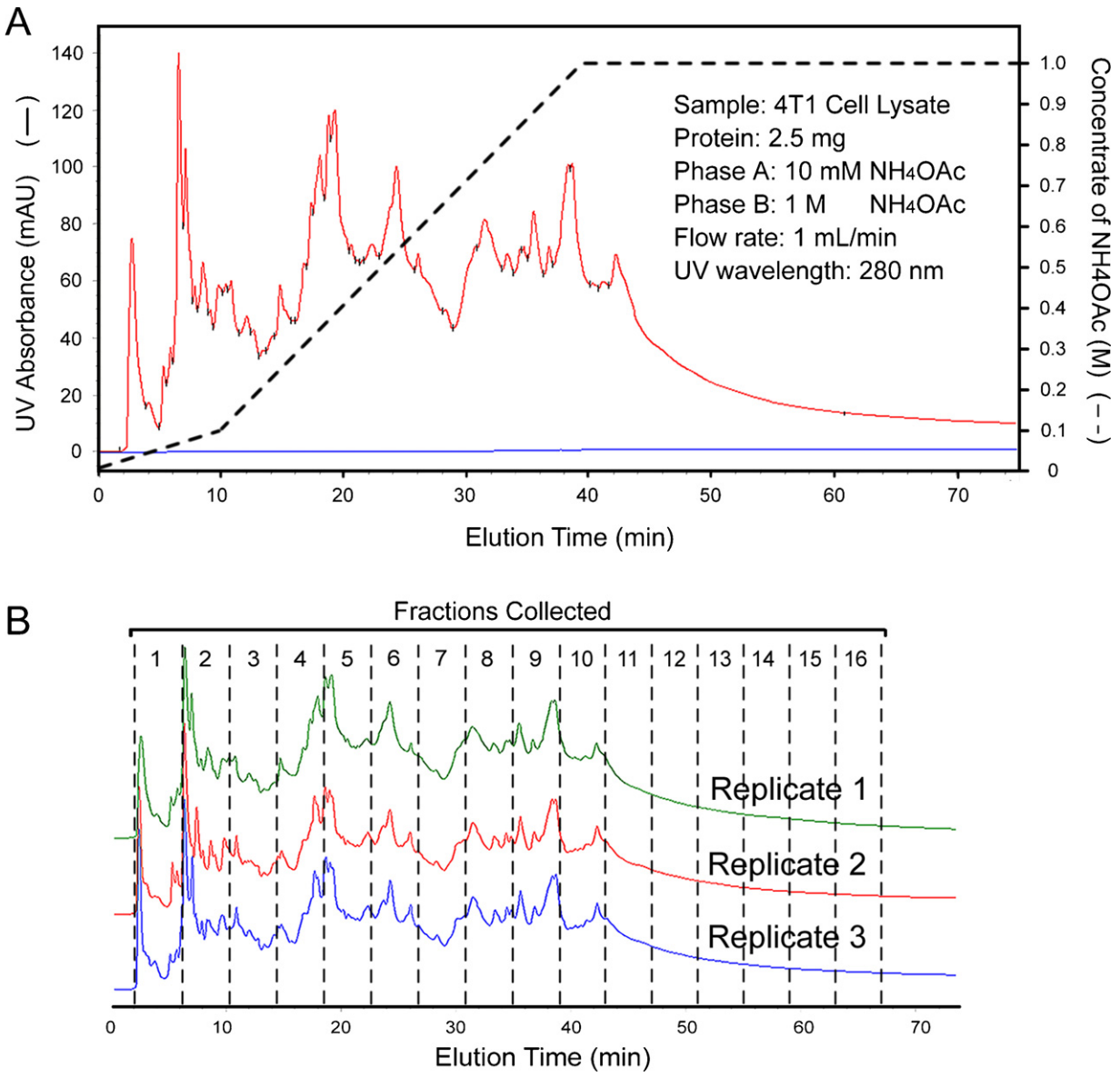

Fig. 2. Protein fractionation using the mixed-bed IEX column. (A) A representative chromatogram, with an elution gradient of salt concentration as shown. (B) The repeatability of the mixed-bed IEX chromatography, showing three replicates. Eluates were collected at 4-min intervals, with 16 resulting fractions in all. Mobile phase A: 10 mM ammonium acetate, pH 7.0; phase B: 1 M ammonium acetate, pH 7.0. Flow rate: 1 mL/min. Detection: 280 nm.

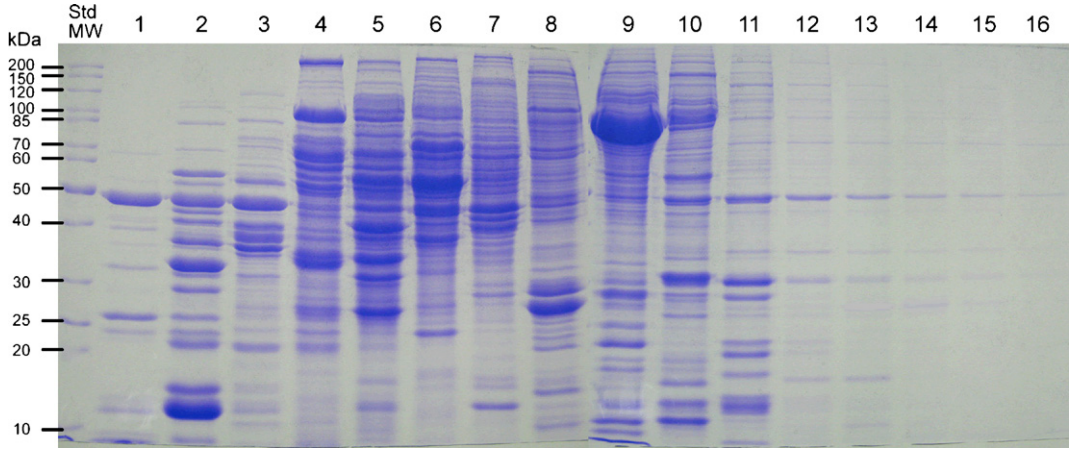

Fig. 3. SDS-PAGE analysis of the resulting protein fractions from mixed-bed IEX chromatography. Lanes 1–16 represent SDS-PAGE profiles of the 16 protein fractions. Lanes of MW Stds correspond to protein size markers ranging from 10 to 200 kDa.

Table 1
Protein identifications per fraction in a typical 3D LC–MS/MS analysis.

Fraction number		1	2	3	4	5	6	7	8	9	10	11	12	Total
Protein Identification	Unique peptides ≥ 1	119	220	370	433	612	717	802	918	813	728	685	657	3084
	Unique peptides ≥ 2	98	187	297	358	501	605	643	739	659	604	522	476	2612
	Unique peptides ≥ 3	58	143	210	258	365	442	449	542	474	433	354	291	1932

proteins of 44.8. Thus, it could be inferred that they were proteins with too low abundance to be stably detected in MS scans.

To assess the degree of fractionation, the protein overlap rate in every two fractions was calculated with the formula described in Section 2 and a heatmap was generated (Fig. 5A). There were some overlaps between adjacent fractions. As the distance between the two fractions in question increased, the overlap rate between them decreased correspondingly. With fractions 1–9 (0–38 min), eluting during the salt gradient, adjacent fractions had overlap rates of 20–35% and an average of 29%. Fractions 10–12, eluted isocratically after 38 min, had overlap rates of 35–50% and an average of 44%. In order to evaluate the fractionation efficiency on the larger scale of the complete elution program, all 3084 proteins were grouped through the number of fractions in which they were detected (Fig. 5B). 1414 proteins (45.9%) were identified only in one fraction, 768 proteins (24.9%) in two fractions, and 2502 proteins (82.3%) were identified in no more than three fractions. These results seemed to demonstrate moderate separation power of the mixed-bed IEX column and that the majority of proteins in the lysate were satisfactorily enriched.

To further assess the properties of the proteins identified through the present method, the 3084 proteins were classified according to several physicochemical characteristics, such as p*I*, molecular weight (MW), and grand average hydrophobicity (GRAVY). As shown in Fig. 6, proteins identified via the 3D LC method varied widely in these properties. The lowest and highest p*I* values were 3.7 and 12.55, respectively. Of the 3084 identified proteins, 2340 proteins (75.9%) were distributed within the p*I* 5–9 interval, 339 proteins (11.0%) had p*I* < 5 and 405 proteins (13.1%) had p*I* > 9. The proteins concentrated mainly within the p*I* ranges of 4–7 and 7.5–10, with few in the interval 7.0–7.5. This compares favorably with the distribution reported for mouse brain [23]. The molecular weight of identified 3084 proteins ranges from 6 kDa to 363 kDa. There were 319 proteins (10.3%) with MW < 20 kDa, 2143 proteins (69.5%) within 20–100 kDa and 622 proteins (20.2%) with MW > 100 kDa. This is in accordance with the SDS-PAGE results as well as the distribution for mouse brain [23]. The most positive GRAVY score was 0.898 and the most negative score was −1.957. Of these 3084 proteins, most proteins (2732 identities, 88.6%) were hydrophilic, with GRAVY value between −1 and 0. With GRAVY value ≤ −1, 223 proteins (7.3%) were highly hydrophilic, and 129 proteins (4.2%) hydrophobic, with GRAVY > 0. These data suggested that our method permits the successful separation and identification of proteins of all kinds in a complex cell lysate.

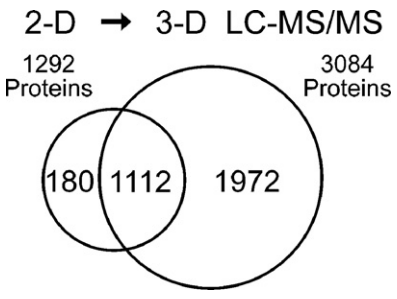

Fig. 4. Venn diagram of the total identified proteins using traditional 2D LC–MS/MS and presented 3D LC–MS/MS.

3.4. The effect of charge characteristics on retention

For deeper understanding of the separation mechanism of the mixed-bed IEX column, we correlated protein retention with their electrostatic properties. The retention time was represented by the fraction numbers. First, retention was assessed as a function of p*I* value. In each fraction, the percentage of proteins within various p*I* range was displayed in Fig. 7A. Neutral proteins, within p*I* range 6–9, were eluted continuously during the gradient, but clearly constituted a steadily decreasing percentage of the total as the gradient progressed. At the same time, extremely acidic or basic proteins (p*I* < 5 or >10) generally eluted at the end of the gradient, with their percentages of the total protein quantity and their absolute numbers increasing at the same time (Fig. 7B and C).

Next, protein retention was plotted as a function of the percentage of charged residues (positively charged: Lys, Arg, His; negatively charged: Asp, Glu) in the sequence, regardless of the p*I* value (Fig. 8). The error bars represented the deviations in percentages of the charged residues of all proteins from the average values in each fraction. Fig. 8A shows that the average total percentage of charged residues rose steadily with elution time, from 25% in fraction 1 to 31% in fraction 12. Fig. 8B shows that this trend pertained to both acidic and basic residues; in general, both of them increased in total residue percentage with later retention time. For neutral proteins, with p*I* 6–8, the same increasing trend was evident for both acidic and basic residues as the gradient progressed (Fig. 8C and D). Although the range of percentage of charged residues of hundreds of proteins in each fraction was broad, leading to big error bars, a general trend could be seen that the percentage of charged residues, both positively and negatively charged ones, could interact with column materials, contributing to the retention of protein. However, the range of observed percentage of charged residues in every fraction was broad, which meant that other factors, such as protein shape-conformation and charge distribution influenced the protein retention on the mixed-bed IEX column [24,25], as well as the scale of both positive and negative charges.

The amount of charged residues is one of the most important factors that affect retention in IEX chromatography. For acidic proteins and basic proteins, such as proteins with p*I* < 5 or p*I* > 10, they usually have strong charge bias and carry more net anion or cation respectively, tending to be eluted later by higher salt concentration generally. For neutral proteins, despite equilibrated numbers of positively charged residues and negatively charged residues, some of them are charged less while others be more. Therefore, neutral proteins appear along the whole eluting gradient, with increasing average charged residues proportion of both kinds.

4. Discussion

Identification of proteins in depth is facilitated by the fractionation of complex samples at the protein level prior to digestion and the usual 2D SCX-RP-LC–MS/MS analysis of the peptides. In the present study, mixed-bed IEX chromatography was assessed for the protein-level fractionation step. More than 3000 proteins were identified from a 4T1 cell lysate using this 3D strategy, and the proteins exhibited a wide range of physicochemical characteristics.

To establish a highly efficient workflow, we first used ammonium acetate gradient to separate proteins on the mixed-bed IEX

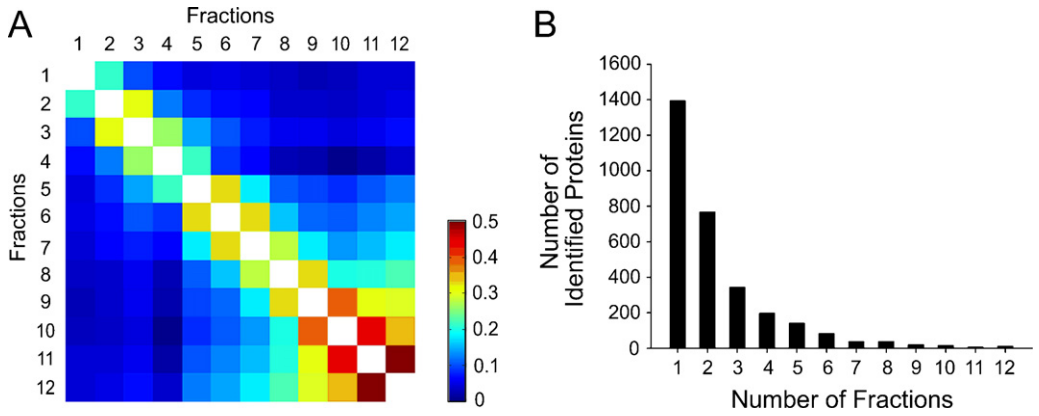

Fig. 5. Separation degree of protein pre-fractionation on the mixed-bed IEX column. (A) Heatmap of inter-fraction protein overlap rate. (B) Number of proteins identified vs. the number of fractions in which these proteins were detected.

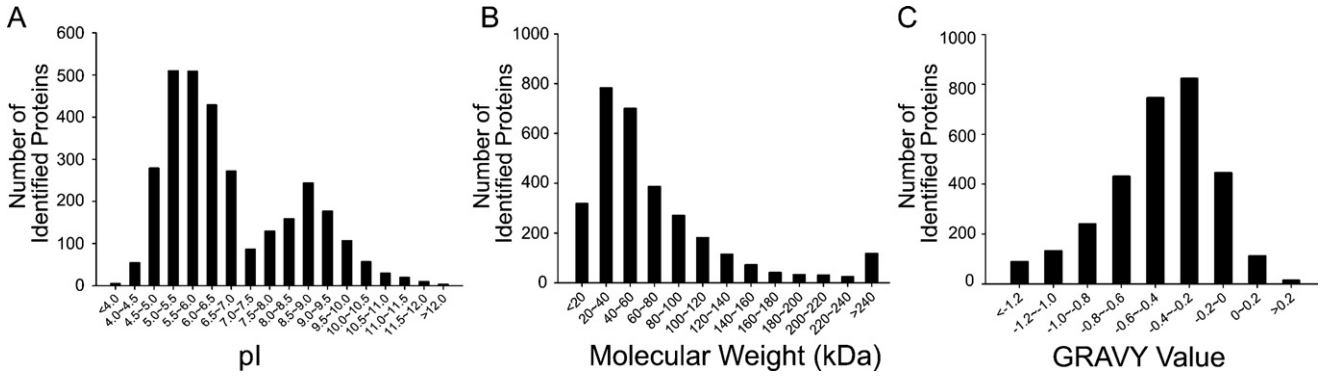

Fig. 6. Distribution of physicochemical properties of the 3084 proteins identified: (A) number of proteins with specific p*I* values; (B) molecular weight; (C) GRAVY value.

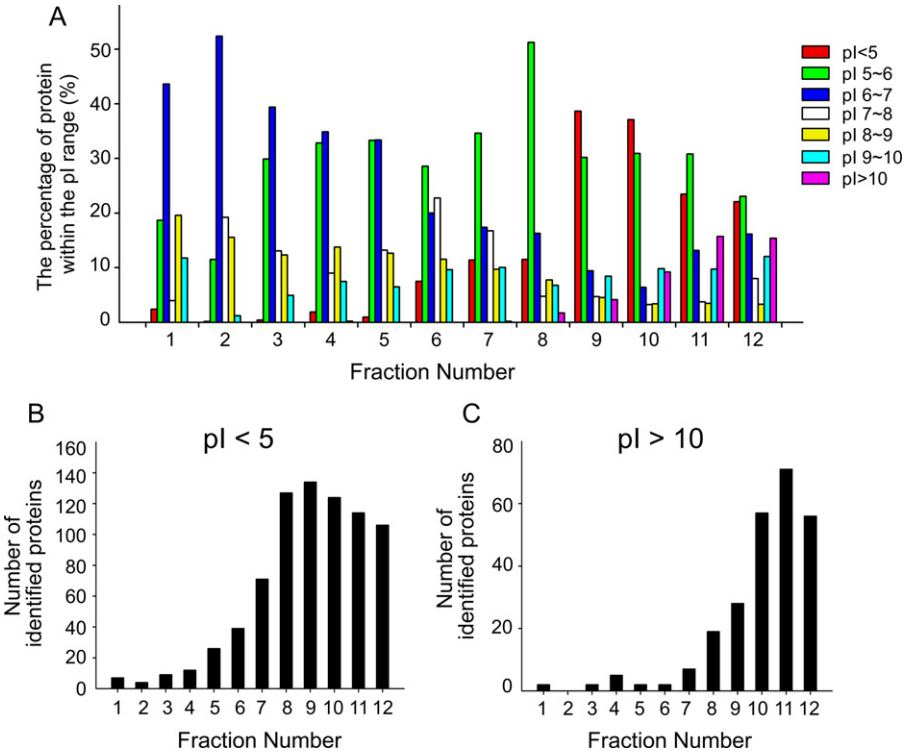

Fig. 7. Analysis of protein p*I* and retention time. The distribution of the proteins within the various p*I* ranges in 12 fractions (A). The height of the bars was proportionate to the percentage of intensity of the proteins within the various p*I* ranges represented regarding to the total intensity of the proteins in each fraction. Numbers of identified proteins with p*I* < 5 (B) and p*I* > 10 (C) in 12 fractions.

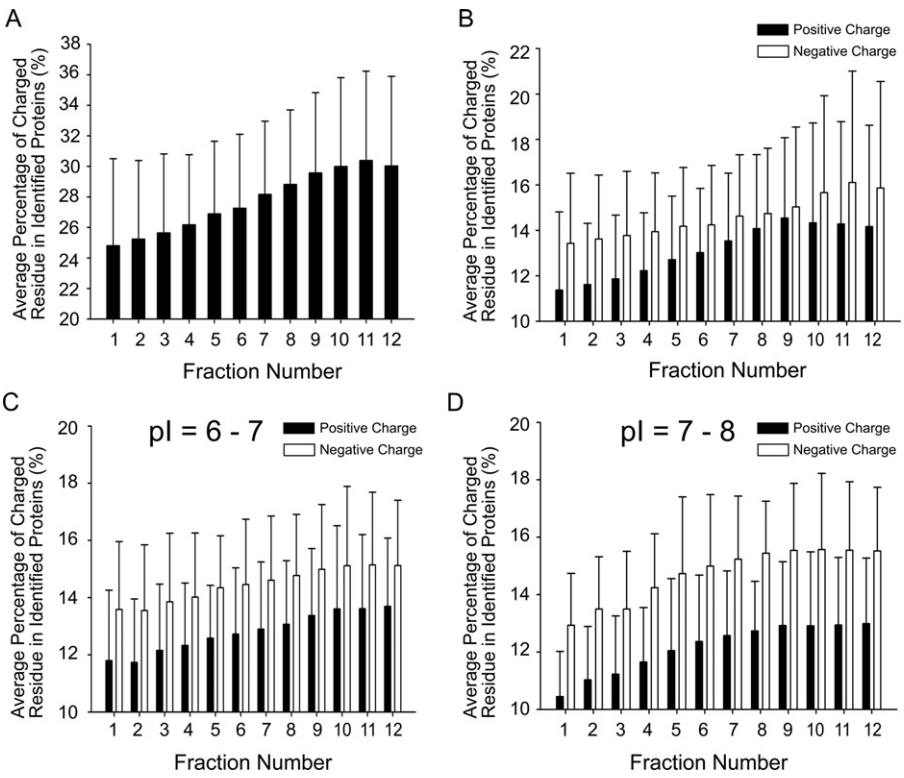

Fig. 8. Effect of the percentage of charged residues in protein sequences on retention. (A) Average percentage of charged residues in protein sequences of all proteins in 12 fractions. (B) Average percentage of both positively charged residues (Lys, Arg, His) and negatively charged residues (Asp, Glu) of all proteins in 12 fractions. Average percentage of both positively and negatively charged residues in 12 fractions for proteins within p*I* range 6–7 (C) and those within p*I* range 7–8 (D). The error bar in each fraction represents the standard deviations of the percentage of the charged residues of all counted proteins with the average value of percentage.

column, instead of the routine-used NaCl or $NaClO_4$ [20,21]. Ammonium acetate is a volatile salt and thus could be easily removed by lyophilization. With the elution system of ammonium acetate gradient, complex protein samples (4T1 cell lysate) were efficiently separated (Fig. 2). UV detection showed that nearly 50 peaks distributed evenly during the 50-min gradient, indicating ammonium acetate gradient was suitable to fractionate proteins on the mixed-bed IEX column. After fractionation, the collected samples were lyophilized and re-solubilized to determine protein concentration, yielding as high as ~98% recovery rate. Compared to NaCl or $NaClO_4$ elution, the utilization of ammonium acetate avoided sample loss as well as additional laboratory work, which are caused by desalting through dialysis or organic solvent precipitation [26,27]. Moreover, after lyophilization, the protein pellets could be treated by standard in solution digestion procedure, which included effective protein reduction and alkylation under urea circumstances prior to trypsin digestion, leading toward more complete digestion and high qualified peptides for MS detection. Thus we established a more compatible mobile system for protein separation on mixed-bed IEX column and efficient connecting steps between protein fractionation and digestion, generating a smooth workflow of 3D LC–MS/MS.

By extensive data analysis of proteins identified in each fraction, we noticed two important features of mixed-bed IEX column contributed by the combination of WAX and WCX. First, after sample loading, mixed-bed column can absorb both positively charged groups and negatively charged groups on protein surface, enabling majority proteins in complex sample binding to the column with only 2.5% of the total peak area being found in the void volume. This indicated that majority proteins could be well retained on the mixed-bed IEX column. Second, since the negatively and positively charged residues bind with WAX and WCX materials respectively, the mixed-bed IEX column could take full advantage of charged groups of either sign. This enabled satisfactory separation of all neutral, acidic and basic proteins, especially proteins with extreme p*I* value. The property distribution range of 3084 identified proteins provided an overall exhibition of the wide scope of protein identification, indicating the applicability of the mixed-bed IEX column in separation complex samples with diverse properties.

Our 3D LC–MS/MS strategy provides the basis for its wide application in proteomic analysis of complex samples, such as cell lysate, tissue lysate, biological fluid etc. Researchers can adopt our 3D LC–MS/MS system with following improvement in the consideration of some special use. In cases where more protein identifications are desired, a more gradual gradient could be implemented and a larger number of fractions collected. Loading a larger amount of sample would also increase the amount of low abundant proteins present, to facilitate the downstream identification of their peptides. The experimental design and equipment with more powerful separation capacity could benefit the deeper identification. Peptides could get a higher resolution on a longer chromatography column and 2D LC system with more optimized gradient and more sub-fractions collected [23].

5. Conclusion

In this study, a 3D LC–MS/MS protein identification strategy was developed based on the mixed-bed IEX protein pre-fractionation and the classic 2D SCX-RP-LC–MS/MS for further peptide separation and detection. The method was applied to complex proteome sample of mammalian cell lysate with more than 3000 proteins identified in a typical experiment. The mixed-bed IEX pre-fractionation reduced the complexity of protein sample and helped for covering proteins with a wide range. Detailed data analysis revealed that the protein retention on mixed-bed IEX was

associated, but not restricted to the extreme of p*I* and the scale of charged residues of both signs.

Acknowledgments

We thank Instrumental Analysis Center of Shanghai Jiao Tong University for the support in mass spectrometry analysis, PolyLC Inc. for providing the mixed-bed IEX column. Thank Dr. A. Alpert for instructions of language and discussion.

This work was supported by Ministry of Science and Technology of China (973 Program, 2010CB834300), National Natural Science Foundation of China (81071792).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jchromb.2012.08.008.

References

[1] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J.R. Yates 3rd, Nat. Biotechnol. 17 (1999) 676.
[2] M.P. Washburn, D. Wolters, J.R. Yates 3rd, Nat. Biotechnol. 19 (2001) 242.
[3] D.A. Wolters, M.P. Washburn, J.R. Yates 3rd, Anal. Chem. 73 (2001) 5683.
[4] M.P. Washburn, R. Ulaszek, C. Deciu, D.M. Schieltz, J.R. Yates 3rd, Anal. Chem. 74 (2002) 1650.
[5] J. Peng, J.E. Elias, C.C. Thoreen, L.J. Licklider, S.P. Gygi, J. Proteome Res. 2 (2003) 43.
[6] W.H. McDonald, R. Ohi, D.T. Miyamoto, T.J. Mitchison, J.R. Yates 3rd, Int. J. Mass Spectrom. 219 (2002) 245.
[7] J. Wei, J. Sun, W. Yu, A. Jones, P. Oeller, M. Keller, G. Woodnutt, J.M. Short, J. Proteome Res. 4 (2005) 801.
[8] A.W. Guzzetta, A.S. Chien, J. Proteome Res. 4 (2005) 2412.
[9] J. Zhang, X. Xu, M. Gao, P. Yang, X. Zhang, Proteomics 7 (2007) 500.
[10] B.L. Hood, M. Zhou, K.C. Chan, D.A. Lucas, G.J. Kim, H.J. Issaq, T.D. Veenstra, T.P. Conrads, J. Proteome Res. 4 (2005) 1561.
[11] J.M. Jacobs, H.M. Mottaz, L.R. Yu, D.J. Anderson, R.J. Moore, W.N. Chen, K.J. Auberry, E.F. Strittmatter, M.E. Monroe, B.D. Thrall, D.G. Camp 2nd, R.D. Smith, J. Proteome Res. 3 (2004) 68.
[12] A. Butt, M.D. Davison, G.J. Smith, J.A. Young, S.J. Gaskell, S.G. Oliver, R.J. Beynon, Proteomics 1 (2001) 42.
[13] P.G. Righetti, A. Castagna, B. Herbert, F. Reymond, J.S. Rossier, Proteomics 3 (2003) 1397.
[14] T. Linke, A.C. Ross, E.H. Harrison, J. Chromatogr. A 1043 (2004) 65.
[15] P. Lescuyer, D.F. Hochstrasser, J.C. Sanchez, Electrophoresis 25 (2004) 1125.
[16] W.H. Jin, J. Dai, S.J. Li, Q.C. Xia, H.F. Zou, R. Zeng, J. Proteome Res. 4 (2005) 613.
[17] P.C. Havugimana, P. Wong, A. Emili, J. Chromatogr. B: Analyt. Technol. Biomed. Life Sci. 847 (2007) 54.
[18] Z.B. Ning, Q.R. Li, J. Dai, R.X. Li, C.H. Shieh, R. Zeng, J. Proteome Res. 7 (2008) 4525.
[19] Z. El Rassi, C. Horvath, J. Chromatogr. 359 (1986) 255.
[20] S. Lee, Y. Chen, H. Luo, A.A. Wu, M. Wilde, P.T. Schumacker, Y. Zhao, J. Proteome Res. 9 (2010) 5705.
[21] A. Le, L.D. Barton, J.T. Sanders, Q. Zhang, J. Proteome Res. 10 (2011) 692.
[22] Y.B. Wu, J. Dai, X.L. Yang, S.J. Li, S.L. Zhao, J.S. Tang, G.Y. Zheng, Y.X. Li, J.R. Wu, R. Zeng, Mol. Cell. Proteomics 8 (2009) 2809.
[23] H. Wang, W.J. Qian, M.H. Chin, V.A. Petyuk, R.C. Barry, T. Liu, M.A. Gritsenko, H.M. Mottaz, R.J. Moore, D.G. Camp Ii, A.H. Khan, D.J. Smith, R.D. Smith, J. Proteome Res. 5 (2006) 361.
[24] F.E. Regnier, Science 238 (1987) 319.
[25] W. Xu, H. Zhou, F.E. Regnier, Anal. Chem. 75 (2003) 1931.
[26] L. Jiang, L. He, M. Fountoulakis, J. Chromatogr. A 1023 (2004) 317.
[27] J.R. McDonald, M. Ong, C. Shen, Z. Parandoosh, B. Sosnowski, S. Bussell, L.L. Houston, Protein Expr. Purif. 8 (1996) 97.