research articles Journal of proteome research

Human Plasma Proteome Analysis by Multidimensional Chromatography Prefractionation and Linear Ion Trap Mass Spectrometry Identification

Wen-Hai Jin,[†,‡] Jie Dai,[†] Su-Jun Li,[†] Qi-Chang Xia,[†] Han-Fa Zou,[*,‡] and Rong Zeng[*,†]

Research Center for Proteome Analysis, Key Lab of Proteomics, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, 200031, China, National Chromatographic R&A Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, 116011, China

Received December 15, 2004

A resurgence of interest in the human plasma proteome has occurred in recent years because it holds great promise of revolution in disease diagnosis and therapeutic monitoring. As one of the most powerful separation techniques, multidimensional liquid chromatography has attracted extensive attention, but most published works have focused on the fractionation of tryptic peptides. In this study, proteins from human plasma were prefractionated by online sequential strong cation exchange chromatography and reversed-phase chromatography. The resulting 30 samples were individually digested by trypsin, and analyzed by capillary reversed-phase liquid chromatography coupled with linear ion trap mass spectrometry. After meeting stringent criteria, a total of 1292 distinct proteins were successfully identified in our work, among which, some proteins known to be present in serum in <10 ng/mL were detected. Compared with other works in published literatures, this analysis offered a more full-scale list of the plasma proteome. Considering our strategy allows high throughput of protein identification in serum, the prefractionation of proteins before MS analysis is a simple and effective method to facilitate human plasma proteome research.

Keywords: proteomics • human plasma • mass spectrometry • two-dimensional liquid chromatography • protein fractionation

Introduction

Blood plasma is believed the most complex human-derived proteome, containing other tissue proteomes as subsets.[1,2] It has an extraordinary dynamic range in that more than 10 orders of magnitude separate albumin and the rarest proteins now measured clinically. It is known some proteins in plasma involved in coagulation, immune defense, small molecule transport, and protease inhibition have been functionally characterized and associated with disease processes.[3−12] Therefore, it is meaningful to identify proteins and find protein biomarkers in human plasma as much as possible.

The conventional proteomics technology platform most commonly used is based on 2-DE MS, which has been more successful in profiling proteins and their disease- or treatment-related quantitative changes in tissue homogenates than in plasma sample.[13] Anderson have reviewed the past decades of plasma proteome research works based on 2-DE:[2,14−20] the highest quantity of identified plasma proteins is only 319. Despite inherent disadvantages of 2-DE, the reason of these relatively poor results maybe lies in the following aspects: the relatively serious band broadening due to very high abundance of a few proteins in plasma, and the crowding of separated protein spots in the molecular mass range between 45 and 80 KD and in the isoelectric point range between 4.5 and 6.[13] Therefore, other high throughput methods in plasma protein identification should be developed. Anderson et al. adopted three individual methodologies to identify plasma proteins.[2] In their work, a total of 1175 distinct proteins were identified after combing results of three different methodologies with literature searching. Although this approach improved the quantity of identified proteins from human plasma remarkably, the time-consuming and labor intensive made this approach less attractive. At the same time, Adkins[21] et al. detected ~490 proteins in serum by an offline SCX chromatographic fractionation of trypsin digests, and online RPLC coupled with ion trap MS. However, this approach was based upon immunoglobulin depletion, as was potentially problematic for quantitative measurements due to the variable and selective losses of other proteins along with immunoglobulins.[22] Some other concurrent works were also involved that problem.[13,23,24] Recently, Shen et al.[22] used high efficiency nano-RPLC and SCXLC to obtain

* To whom correspondence should be addressed. Prof. Rong Zeng, Research Center for Proteome Analysis, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 YueYang Road, Shanghai, 200031. Tel: +86-21-54920170. Fax: +86-21-54920171. E-mail: zr@sibs.ac.cn. Prof. Hanfa Zou, National Chromatographic R&A Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116011, China. Tel: +86−411−83693409, Fax: +86−411−83693407. E-mail: zouhfa@mail.dlptt.ln.cn.

† Shanghai Institutes for Biological Sciences.

‡ Dalian Institute of Chemical Physics.

10.1021/pr049761h CCC: $30.25

Published on Web 03/18/2005

ultrahigh-efficiency separations in conjunction with tandem MS for characterization of the human plasma proteome, as a result, between 800 and 1682 human proteins were identified. This method was very simple and effective, besides the capillaries used were so long that the back pressure of the columns was far beyond 10 000 psi, which made this approach difficult to perform in ordinary lab.

In this work, proteins from human plasma were prefractionated using online SCXLC/RPLC. The resulting 30 samples were individually digested by trypsin, and analyzed by capillary reversed-phase liquid chromatography coupled with linear ion trap spectrometry. At last, a nonredundant list comprising 1292 distinct proteins was obtained, which contained numerous low abundance human plasma proteins, demonstrating this methodology a valid and effective technique in plasma proteome research.

Materials and methods

Chemicals. All the water used in the experiment was prepared using a Milli-Q system (Millipore, Bedford, MA); Dithiothreitol (DTT), ammonium bicarbonate and iodoacetamide (IAA) were all purchased from Bio-Rad (Hercules, CA); Urea and sodium chloride (NaCl), were obtained from Sigma (St. Louis, MO); Sodium di-hydrogen phosphate, 2-hydrate, was purchased from Fisher (Merck, Darmstadt, German); Trypsin was purchased from Promega (Madison, WI); Formic acid (FA), trifluoroacetic acid (TFA) and acetonitrile (ACN) were obtained from Aldrich (Milwaukee, WI). Phosphoric acid was obtained from Shanghai Chemical Plant (Shanghai, China). All the chemicals were of analytical grade except acetonitrile, which was of HPLC grade.

Human Plasma Preparation- The human blood plasma was obtained from one healthy female donor (ages 27, O type), provided by Shanghai Blood Center. Protease inhibitors were added when collected the blood. An initial protein concentration of ~100 mg/mL of plasma was determined using Bradford method. After that 500 μL sample was diluted to ~20 mg/mL with 50 mM phosphate buffer (pH 2.5, containing 5% ACN). The serum samples were filtered through 0.22 μm filters (Agilent technologies) by spinning at 16 000 $\times$ g at room temperature for 1.5 min. After that, the sample was exchanged to SCXLC buffer A, and the last protein concentration was 18.2 mg/mL.

Two-Dimensional Liquid Chromatography. Using the Biocad-vision workstation (Applied biosystems), a multistep chromatographic procedure was used to perform prefractionation of human plasma proteins. In detailed, the above human plasma samples of 1.1 mL (containing ~20 mg protein) was injected on a POROS HS20 4.6 $\times$ 100 mm SCXLC (PerSeptive Biosystems, Framingham, MA) column, which was tandem with POROS R2 10 2.1 $\times$ 100 mm RP column (PerSeptive Biosystems, Framingham, MA) through a six-port valve, thereafter the sample was separated using strong cation exchange liquid chromatography in the first dimension followed by reversed-phase HPLC in the second dimension. The flow rate was 1 mL/min all the time. The eluents for SCXLC column were A, 50 mM phosphate buffer at pH 2.5 and B, 2 M NaCl in 50 mM phosphate buffer at pH 2.5, both containing 5% ACN (v/v). Buffers for RP HPLC column were C, water with 0.1% trifluoroacetic acid, and D, 90% acetonitrile with 0.1% trifluoroacetic acid. Typical acetonitrile gradient elution for the second dimension was: 0% D to 30% D in 5 min, followed by 30% D to 100% D in 25 min, and an isocratic elution of 100% D in 10 min followed by 100% C flushing for another 10 min to rebalance RP HPLC column, Six step gradients in the first dimension were, 0% B, 10% B, 20% B, 40% B, 80% B, and 100% B corresponding to 0, 200, 400, 800, 1600, and 2000 mM NaCl buffers, each flushing the tandem columns for 40 min, followed by the second dimensional elution described above. It should be noted that in order to eliminate salts containing in the RP column during the SCX process, a 40 min desalting process using 100% buffer C to flush RP column was adopted before the second dimensional separation started. In our experiment, samples were collected into tubes every 8 min (from 0 to 40 min) in the second dimensional elution process in time-dependent mode, that was, 30 tubes of sample were collected at the end of chromatography separation.

Sample Preparation for MS Analysis. 30 tubes of sample obtained above were lyophilized to dryness and redissolved in 150 μL reducing solution (8 M Urea, 100 mM ammonium bicarbonate, pH 8.3), each mixed with 3 μL of 1 M DTT. The mixture was incubated at 37 °C for 2.5 h and then 20 μL of 1 M IAA was added and incubated for an additional 30 min at room temperature in darkness. After that, the protein mixtures in each fraction were exchanged into 100 mM ammonium bicarbonate buffer, pH 8.5 and each incubated with 20 mg trypsin at 37 °C for overnight to generate enough peptides. Undigested proteins and trypsin were removed by filtering the mixture through filters, cut off values of which were ≥10 k Da (Millipore technologies). The digested peptide mixtures were lyophilized, and then dissolved in 0.1% formic acid.

LC−MS/MS Analysis. A Surveyor liquid chromatography system (ThermoFinnigan, San Jose, CA), consisting of degasser, MS Pump and autosampler, equipped with a C 18 reversed phase column (RP, 150 μm $\times$ 120 mm, home packing), was used. The pump flow rate was split 1:120 to achieve a column flow rate of 1.5 μL/min. The samples were loaded onto the column with a mobile-phase gradient from 0% F (water with 0.1% formic acid) to 70% F (acetonitrile with 0.1% formic acid) in 150 min. The mass spectral data observed in this work were acquired on a LTQ linear ion trap mass spectrometer (Thermo, San Jose, CA) equipped with an electrospray interface operated in positive ion mode. The temperature of heated capillary was set at 170 °C. A voltage of 3.0 kV applied to the ESI needle resulted in a distinct signal. Normalized collision energy was 35.0. The number of ions stored in the ion trap was regulated by the automatic gain control. Voltages across the capillary and the quadrupole lenses were tuned by an automated procedure to maximize the signal for the ion of interest. The mass spectrometer was set that one full MS scan was followed by ten MS/MS scans on the 10 most intense ions from the MS spectrum with the following Dynamic Exclusion settings: repeat count 2, repeat duration 30 s, exclusion duration 90 s.

Data Analysis. To identify proteins in the above sample, spectra form each fraction was searched with the SEQUEST algorithm against the nonredundant human protein database from IPI. All output results were combined together using homemade software named BuildSummary to delete keratins and the redundant data. In detailed, peptide identifications should require the following criteria: $X_{corr} \geq 1.9$ with charge state 1+, $X_{corr} \geq 2.2$ with charge state 2+, or $X_{corr} \geq 3.75$ with charge state 3+. In addition, ΔCn cutoff values were ≥0.1 and the SP rank of the peptides ≤4.

Results

Two-Dimensional Liquid Chromatography. In this experiment, every step gradient of SCX was followed by a 50-min.

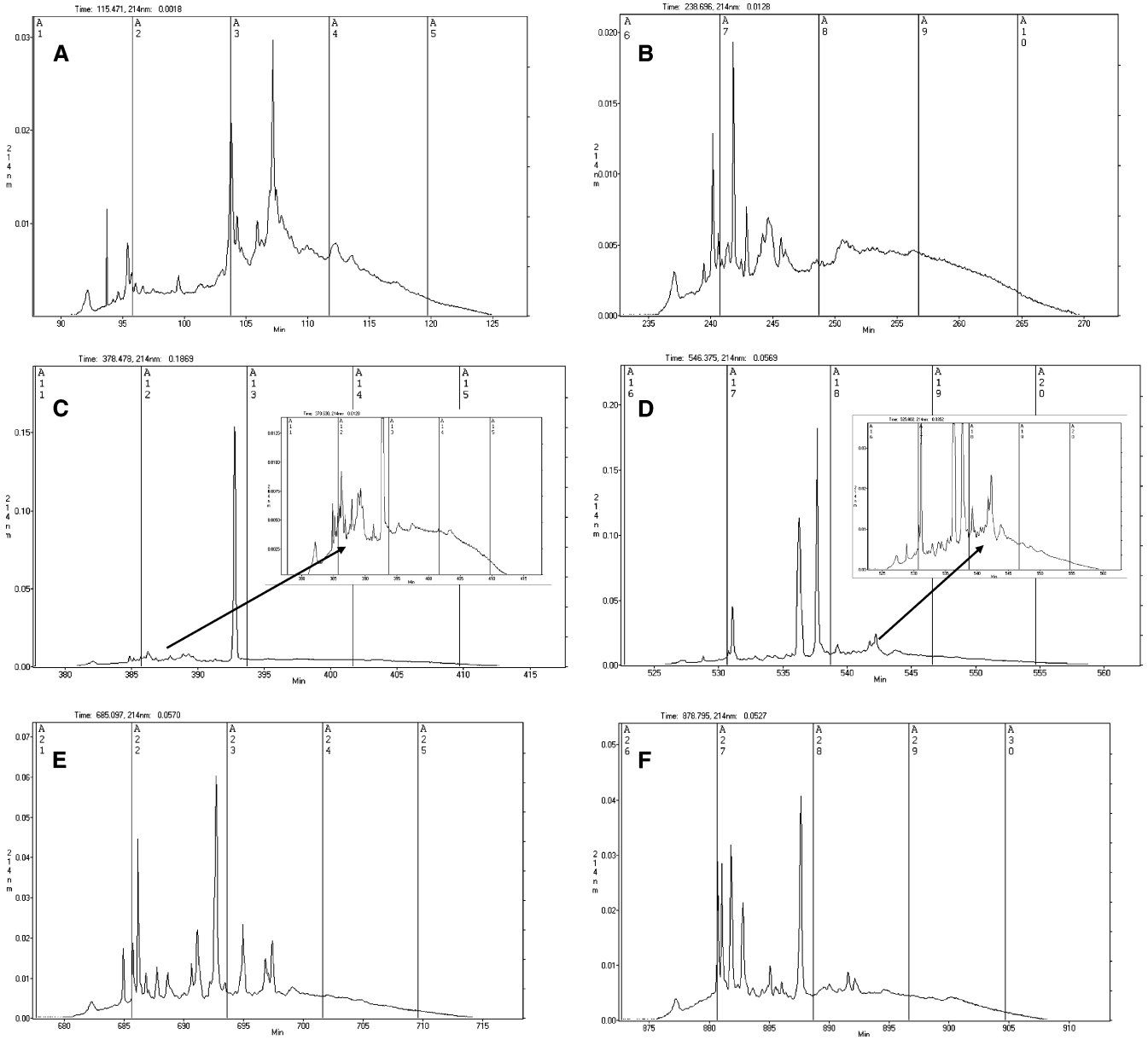

Figure 1. Separation of human plasma proteins on the tandem SCXLC/RPLC columns. RPLC Chromatograms of proteins eluted from SCX column by six step gradients of 0, 200, 400, 800, 1600, and 2000 mM NaCl are corresponding from (A) to (F). The 30 collected fractions are marked from A1 to A 30. Experimental conditions: SCX column, POROS HS 20 4.6 × 100 mm; RP column, POROS R2 10 2.1 × 100 mm; Buffers for SCX LC: A, 50 mM phosphate buffer at pH 2.5; B, 2 M NaCl in 50 mM phosphate buffer at pH 2.5, both containing 5% ACN (v/v). Buffers for RP HPLC: C, water with 0.1% trifluoroacetic acid, D, 90% acetonitrile with 0.1% trifluoroacetic acid. Acetonitrile gradient elution for the RP column: 0% D to 30% D in 5 min, followed by 30% D to 100% D in 25 min, and an isocratic elution of 100% D in 10 min followed by 100% C flushing for another 10 min. Flow rate: 1 mL/min, temperature: 25 °C, detection wavelength, 214 nm.

RP-HPLC run. Therefore, 6 step gradients of SCX were corresponding to 6 RP chromatograms (See experimental). Figure 1 shows the separation of human plasma proteins (fractionated by SCX) as detected by UV absorption at 214 nm on RP-HPLC column at corresponding separation time interval. From the figure it can be seen nearly 150 peaks were obtained in our experiment. Although the intensity of some peaks was very sharp, the resolutions of the separation, together with the symmetry of most peaks in the chromatograms were satisfied. To verify the results, the same separation was performed twice. At last, no obvious changes in peak shape and shifts in retention times were observed; demonstrating the reproducibility of the separation was very acceptable. In our work, 30 fractions, as annotated in the chromatograms, each containing about 8 mL of RP eluents, were lyophilized to dryness quickly, and stored in −80 °C refrigerator, which reduced the risk of protein degradation to the lowest level.

High-Throughput of the Human Plasma Proteins with LTQ Identification. About 30 000 scans of mass spectrum were obtained in one capillary LC−MS/MS run for one fraction analysis. At last, a total of 16 921 peptides corresponding to 750 thousands.out files were obtained in our experiment, After meeting the criteria mentioned in the experimental procedure, a very conservative method of selecting distinct proteins was also used in order to validate the identified proteins. As a result, any identified proteins assigned to the same cluster by

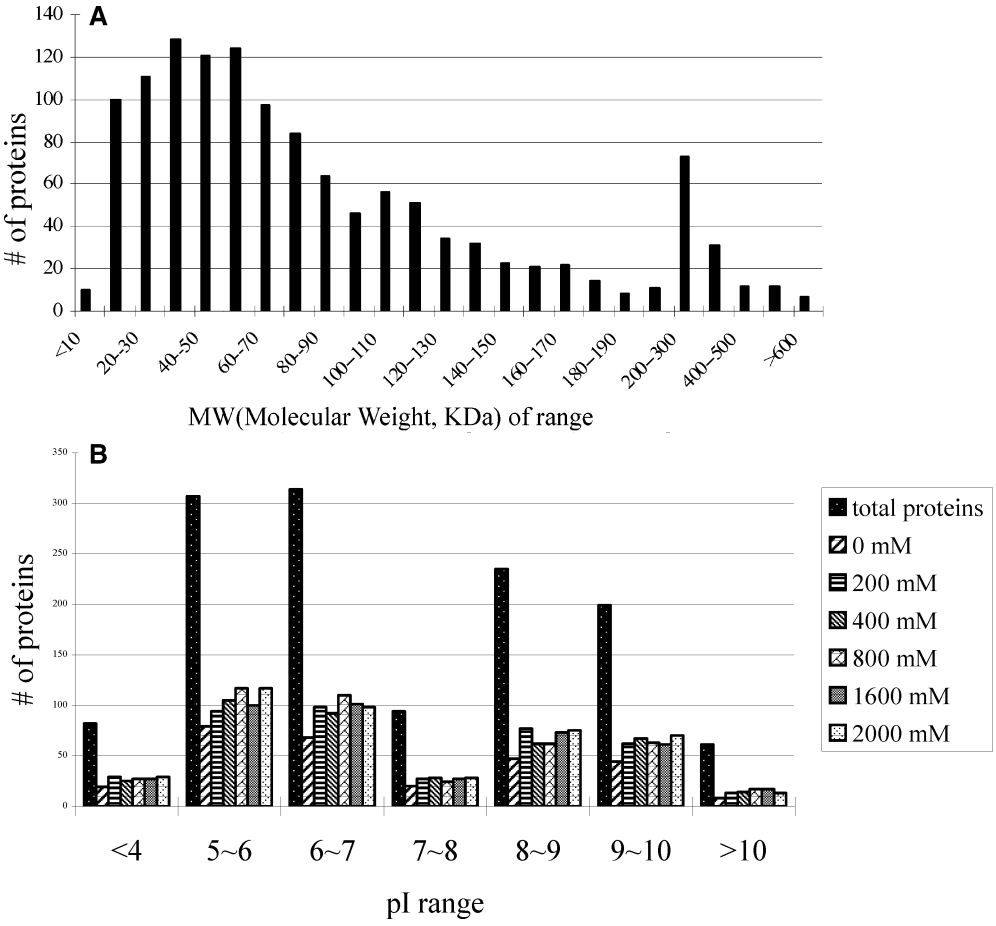

Figure 2. Physicochemical characteristics of the identified proteins. Mw distribution, (B) p*I* distribution. 0, 200, 400, 800, 1600, and 2000 mM represents the proteins identified in the 0, 200, 400, 800, 1600, and 2000 mM NaCl eluents from SCX column, respectively.

SEQUEST were reported as a single entry in the nonredundant through in-house software BuildSummary. At last, a total of 1292 distinct proteins were reported in this work, as listed in the Supporting Information Table. It seems more protein precursors are reported in our work. As a bottom-up strategy, The same peptide (s) may be assigned to one more proteins which treated as a group. To obtain the nonredundant result for functional annotation, protein with either most complete annotations in the database or highest molecular weight in one group were selected for the functional classification and the most coverage of the protein sequence.

The reporting proteins in our work comprise proteins from high abundance protein, e.g., HSA, IgG to the relatively low abundance proteins, such as interleukins family etc. At the same time, some known peptides and proteins, e.g., leptin, amylase and parathyroid hormone, which should be present under normal conditions, were not included in the table. We did detect these compounds, but either the X_{corr} or ΔCn of which cannot reach the cutoff values, indicating the false-negative resulted from the strict cutoff. The most important issue for the data quality in our work is to avoid false-positive using high score, which certainly led to the false-negative.

Physicochemical Characteristics of the Identified Proteins. The 1292 identified proteins were classified according to different physicochemical characteristics such as molecular weight (Mw) and isoelectric point (p*I*). The protein distribution patterns of each step gradient eluents of SCX were compared resulted from the above characteristics listed above (Figure 2). In the present work, for the total 1292 proteins, there were only 10 (0.8%) proteins with Mw < 10 kDa, and 407 (31.50%) proteins with Mw > 100 kDa. The smallest and largest Mw obtained were 4.53 kDa and 3816 kDa. Regarding the p*I* distribution, the 1292 proteins distributed cross a wide p*I* range (3.95−13.72). 321 (24.85%) proteins distributed among p*I* 5−7 intervals, but 82 (6.34%) proteins had p*I* < 4.0 and 61 (4.70%) proteins had p*I* > 10. Although no obvious regular patterns on proteins p*I* distribution were observed in different step gradient eluents from SCX, proteins with p*I* between 7 and 8 only took 7.28% of the total.

Gene Ontology Annotation. In his classic series entitled the plasma proteins, Putnam[25]defined true plasma proteins as those that carry out their functions in the circulation, thus excluding proteins that for example, serve as messengers between tissues or that leak in blood as a result of tissue damage. Elaborating on Putnam's classification from a functional viewpoint, Anderson et al. classified the protein content of plasma into the 8 groups:[1] 1. Protein secreted by solid tissues and that act in plasma; 2. immunoglobins; 3."long distance" receptor ligands; 4. "local receptor" ligands; 5. temporary passengers; 6. tissue leakage products; 7. aberrant secretions; 8. foreign proteins. But in fact, with more and more proteins were identified in human plasma, it was more difficult to define a distinct protein within these catalogues. Therefore, a sort based on Gene Ontology annotation was presented in our work.

On one hand, 865 (66.95%) were mapped to at least one annotation term within the GO molecular function category,

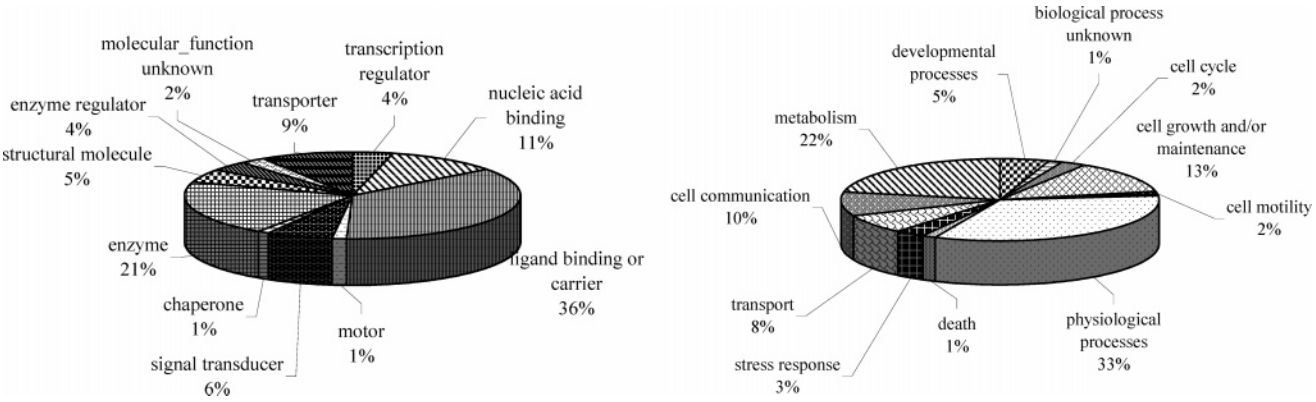

Figure 3. Classification of identified human plasma proteins based on GOA. (A) Molecular Function Ontology (B) Biological Process Ontology.

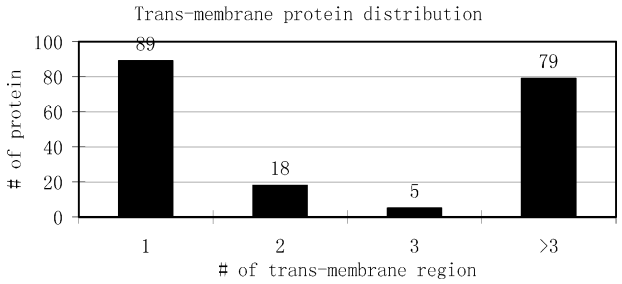

Figure 4. Trans-membrane distribution of proteins identified.

including 62 (4%) transcription regulator proteins, 177 (11%) nucleic acid binding proteins, 582 (36%) proteins with ligand binding or carrier capacity, 23 (1%) moto proteins, 104 (6%) signal transducer proteins, 16 (1%) chaperone proteins, 334 (21%) proteins with enzyme activity, 80 (5%) structural molecule proteins, 62 (4%) enzyme regulator proteins, 143 (9%) transporter proteins, and 40 (2%) proteins with molecular function unknown, as shown in Figure 3 A. On the other hand, 2162 GO numbers of 783 (60.60%) proteins with at least one annotation term in GO biological process category were classified into 11 sorts: developmental processes (109, 5.04%), cell cycle (49, 2.27%), cell growth and/or maintenance (281,13.00%), cell motility (33, 0.15%), physiological processes (724, 33.49%), death (28, 1.30%), stress response (63, 2.91%), transport (165, 7.63%), cell communication (206, 9.52%), metabolism (472, 21.83%), biological process unknown (32, 1.48%), as shown in Figure 3 B.

Trans-Membrane Protein Prediction. Trans-membrane domain of identified proteins was predicted by TMHMM. As shown in Figure 4, 191 (14.78%) of identified 1292 proteins had at least one trans-membrane region. Among them, 89 proteins with one trans-membrane region, 18 with two trans-membrane regions, and 84 proteins with at least three trans-membrane region. It should be noted in our experiment, a protein named ovarian cancer related tumor marker CA125, which is the currently the only tumor marker to have a validated role in the postoperative monitoring of ovarian cancer,[26] had no possible trans-membrane regions.

Discussion

Liquid Chromatography Based Works in Human Plasma Proteome Research. Most liquid chromatography based works in human plasma proteome research have focused on the fractionation of peptides mixture. In fact, prefractionation of proteins by multidimensional liquid chromatography has been proven to be powerful techniques in proteomics research,[27−31] as was used for example, in Liu's work[27] when online SCXLC/RPLC was used to analyze proteins form yeast ribosomal. In their works, the splitted eluents form 2D-HPLC was online interfaced with ESI-TOF−MS and the fraction collection was performed simultaneously. The components were detected by the electrospray and MALDI mass spectrometer. At last, about one hundred and sixty proteins were identified. Considering the yeast ribosomal proteome had a limited complexity, we extended this fractionation technique to plasma proteome which has much wider dynamic range than other systems. The objective of this study was to determine if this 2-D fractionation of proteins based on SCX-RPLC prior to trypsin-digestion and MS detection would allow for the identification of lower abundance proteins in plasma in the presence of the more common high abundant proteins, since depletion of which using affinity-based methods would resulted in the loss of physiological important low abundant proteins in human plasma. Due to the sequence homology, in our work, 1292 protein groups containing of 1971 protein accession numbers were obtained and after our de-redudance procedure by BiuldSummary (Manuscript submitted), 1292 representive proteins were finally defined. Compared to results in the literature and direct capillary LC−ESI−MS/MS analysis of plasma derived tryptic peptides mixture (data not shown), this method was an effective approach in human plasma protein identification, i.e., the proteins were fractionated before MS identification, which contributed greatly to the resolution of the system. Lescuyer et al.[32] reviewed recent progress in different chromatographic protein fractionation techniques and drew a conclusion that chromatographic technology was a powerful tool for enrichment of low abundance proteins because it might cluster the high abundance components to a single or more fractions. While in our work, 20 mg proteins were loaded on SCXLC column, thus the enrichment of the low abundance proteins of the sample would be much more outstanding not only because of the reason mentioned above, but also because elevating quantity of injected sample would lead to a increase to low abundance proteins introduced into the system. As chromatograms shown in Figure 1, there were very high peaks in fraction #12 and #17, indicating there was a great feasibility of clustering some high abundant proteins of human plasma in these several fractions. In fact, through the capillary RPLC−LTQ−MS/MS analysis, some known very high

abundant plasma proteins such as serum albumin precursor, and fibrinogen α/α-E chain precursor were identified in such fractions, seldom identified in other factions, which demonstrated Lescuyer's theory, thereby 2D LC fractioanation played a positive effect on the identification of low abundant proteins in human plasma in our work.

Although we have only identified very limited proteins with Mw < 10 KDa compared to Weissinger's work,[33] our strategy appeared to give an overall understanding of proteins in human plasma with little restrictions on the Mw and p*I*, For example, there were 417 (32.28%) proteins with Mw beyond traditional 2DE (Mw < 10 kDa, Mw > 100 kDa), and 260 (20.12%) with p*I* beyond routine 2DE (p*I* > 9), demonstrating that this methodology shows great potential for proteome research. But as shown in Figure 2, the experimental p*I* distribution of the identified proteins in each fraction seemed to have no relationships with their theoretical SCX chromatography behavior, for the solutes should have been eluted sequentially according to their p*I*. In fact, this contradiction observed in our experiment was reasonable, because the theoretical p*I*s of proteins in the databases are given for the full genomic length unmodified proteins, but proteins are often processed and post-translationally modified thus, those values do not apply for most proteins circulating in the bloodstream. In addition, the different homologous sequences, isoforms, splice variants would also influenced the experimental results. Furthermore, proteins could also bind to matrix through nonionic interactions. In fact, as illuminated by Lescuyer,[32] one drawback of IEC is this approach is restricted to soluble proteins since buffers contain neither urea nor detergents, but an advantage for the separation of plasma proteins since most proteins in it are soluble. The quantitative dynamic range currently accessible in the plasma proteome could be accessed by either printed papers or clinic references. At the high abundance end, HSA (Human Serum Albumin) has a normal concentration range of 35−50 mg/mL, while at the low abundance end, interleukin-12 has a normal range of 77 pg/mL,[22] both were successfully identified in our experiment. Although there is still controversy over whether the low abundance plasma proteins are more interesting or clinically meaningful than those of high abundance, it is of first importance to identify protein from plasma as much as possible. As in our work, 254 (19.66%) proteins were identified as hypothetical proteins, a high abundant Ras family (46, 3.56%), KIAA (71, 5.50%) family and zinc finger family (26, 2.01%) proteins were identified, which demonstrates this methodology potential of proteome analysis to human plasma. Furthermore, there were many proteins that might not ordinarily be expected to be abundant enough to appear in the "common component" list identified in our work, involving some interesting molecular functions, including extracellular calcium-sensing receptor precursor (senses changes in the extracellular concentration of calcium ions), tumor suppressor protein DCC precursor (implicated as a tumor suppressor gene), Fibrillin 2 precursor (structural component of connective tissue microfibrils that binds calcium), angiotensinogen precursor (which helps regulate volume and mineral balance of body fluids), FK506-binding protein 4 (component of inactivated mammalian steroid receptor), proto-oncogene tyrosine-protein kinase FER (involving in cell cycle control), cathepsin L precursor (suggested to be predictive for tumor growth and invasion processes). A number of these proteins have obvious relevance to important disease mechanisms, and thus are of potential diagnostic value. For example, extracellular calcium-sensing receptor precursor (CASR) has been demonstrated to have relevance with several human diseases: defects in CASR are not only the cause of familial hypocalciuric hypercalcemia but also the cause of neonatal severe primary hyperparathyroidism; In addition, abnormal of this protein are also the cause of autosomal dominant hypocalcemia and autosomal dominant hypoparathyroidism. In our experiment, the identification of this protein was based upon assignment of its tryptic peptide R·NSTHQNSLEAQKSSDTLTR·H with M 2+, X_{corr} of 2.3, ΔCn of 0.23 and RSp of 1. Ovarian cancer related tumor marker CA125 is of clinical value in the pre-operative diagnosis and monitoring of ovarian malignancies. Available data suggest that CA 125 is elevated in the majority of epithelial ovarian malignancies prior to clinical presentation, However, an increase in the blood levels of this tumor marker has recently been reported even in patients with congestive heart failure,[34] which made this protein more interesting. The identification of this protein in our work was base upon assignment of its tryptic peptide K.STVLSSVPTGATTEVSR.T with M 2+, X_{corr} of 2.5, ΔCn of 0.16 and RSp of 1 according to the peptide assignment criteria. The identification of CA125 in a healthy young woman indicates this methodology will be more attractive if combining with quantitative proteomics techniques. Though the numbers of "peptide hits" can indicate the rough quantitative information of a protein, the accurate quantitation still needs other method such as ICAT.

Although 1292 unique proteins were identified, it is observed that most of the proteins identified based only on single peptide (Supporting Information). The protein prefractionation increased the identification numbers of proteins, however, the proteins with at least two identified peptides did not increase significantly. Although the extremely high-abundant proteins were distributed in more fractions and low-abundant proteins can be detected in certain fractions, the peptides from high-abundant proteins still are bias to be detected in 1D-LC mass spectrometry. Therefore, for a very dynamic system, the protein fractionation should be combined with peptide fractionation to achieve more proteins identified, also more proteins with at least two identified peptides. Of course, it will be very time-consuming if both fractionations used.

In summary, we provided a strategy for human plasma proteome research: prefractionation of human plasma proteins using online SCXLC/RPLC followed by trypsin digestion and capillary RPLC−ESI−Linear Ion Trap-MS/MS identification, followed by bioinformatics annotation, which has provide as a high throughput, sensitive and effective analytical approach for proteomics of human plasma. Plenty of identified proteins could provide important links to damage or malfunction of the tissues in which they are specially expressed, providing leads for disease biomarkers.

Acknowledgment. This work was supported by National High-Technology Project (2002BA711A11) and Basic Research Foundation (2002CB713807).

Supporting Information Available: 1292 distinct proteins reported in this work (Supporting Information Table). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Anderson, N. L.; Anderson, N. G. *Mol. Cell Proteomics* **2002**, *1*, 845−867.
(2) Anderson, N. L.; Polanski, M.; Pieper, R.; Gatlin, T.; Tirumalai, R. S.; Conrads, T. P.; Veenstra, T. D.; Adkins, J. N.; Pounds, J. G.; Fagan, R.; Lobley, A. *Mol. Cell Proteomics* **2004**, *3*, 311−326.
(3) Pieper, R.; Gatlin, C. L.; Makusky, A. J.; Russo, P. S.; Schatz, C. R.; Miller, S. S.; Su, Q.; McGrath, A. M.; Estock, M. A.; Parmar, P. P.; Zhao, M.; Huang, S. T.; Zhou, J.; Wang, F.; Esquer-Blasco, R.; Anderson, N. L.; Taylor, J.; Steiner, S. *Proteomics* **2003**, *3*, 1345−1364.
(4) Stevens, E. V.; Liotta, L. A.; Kohn, E. C. *Int. J. Gynecol. Cancer* **2003**, *13 Suppl. 2*, 133−139.
(5) Steel, L. F.; Shumpert, D.; Trotter, M.; Seeholzer, S. H.; Evans, A. A.; London, W. T.; Dwek, R.; Block, T. M. *Proteomics* **2003**, *3*, 601−609.
(6) Ryu, J. W.; Kim, H. J.; Lee, Y. S.; Myong, N. H.; Hwang, C. H.; Lee, G. S.; Yom, H. C. *J Korean Med. Sci.* **2003**, *18*, 505−509.
(7) Rui, Z.; Jian-Guo, J.; Yuan-Peng, T.; Hai, P.; Bing-Gen, R. *Proteomics* **2003**, *3*, 433−439.
(8) Ransohoff, D. F. *Science* **2003**, *299*, 1679−1680.
(9) Poon, T. C.; Yip, T. T.; Chan, A. T.; Yip, C.; Yip, V.; Mok, T. S.; Lee, C. C.; Leung, T. W.; Ho, S. K.; Johnson, P. J. *Clin. Chem.* **2003**, *49*, 752−760.
(10) Stulik, J.; Hernychova, L.; Porkertova, S.; Pozler, O.; Tuckova, L.; Sanchez, D.; Bures, J. *Proteomics* **2003**, *3*, 951−956.
(11) Diamandis, E. P. *J. Natl. Cancer Inst.* **2004**, *96*, 353−356.
(12) Rosenblatt, K. P.; Bryant-Greenwood, P.; Killian, J. K.; Mehta, A.; Geho, D.; Espina, V.; Petricoin, E. F., 3rd; Liotta, L. A. *Annu. Rev. Med.* **2004**, *55*, 97−112.
(13) Pieper, R.; Su, Q.; Gatlin, C. L.; Huang, S. T.; Anderson, N. L.; Steiner, S. *Proteomics* **2003**, *3*, 422−432.
(14) Anderson, L.; Anderson, N. G. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5421−5425.
(15) Anderson, N. L.; Anderson, N. G. *Electrophoresis* **1991**, *12*, 883−906.
(16) Hughes, G. J.; Frutiger, S.; Paquet, N.; Ravier, F.; Pasquali, C.; Sanchez, J. C.; James, R.; Tissot, J. D.; Bjellqvist, B.; Hochstrasser, D. F. *Electrophoresis* **1992**, *13*, 707−714.
(17) Anderson, N. G.; Willis, D. D.; Holladay, D. W.; Caton, J. E.; Holleman, J. W.; Eveleigh, J. W.; Attrill, J. E.; Ball, F. L.; Anderson, N. L. *Anal.Biochem.* **1975**, *66*, 159−174.
(18) Tracy, R. P.; Currie, R. M.; Young, D. S. *Clin. Chem.* **1982**, *28*, 890−899.
(19) Harrison, H.; Miller, K.; Ober, C.; Refetoff, S.; Dick, M.; Elias, S. *Am. J. Hum. Genet.* **1991**, *48*, 362−369.
(20) Harrison, H. H.; Gordon, E. D.; Nichols, W. C.; Benson, M. D. *Am. J. Med. Genet.* **1991**, *39*, 442−452.
(21) Adkins, J. N.; Varnum, S. M.; Auberry, K. J.; Moore, R. J.; Angell, N. H.; Smith, R. D.; Springer, D. L.; Pounds, J. G. *Mol. Cell Proteomics* **2002**, *1*, 947−955.
(22) Shen, Y.; Jacobs, J. M.; Camp, D. G., 2nd; Fang, R.; Moore, R. J.; Smith, R. D.; Xiao, W.; Davis, R. W.; Tompkins, R. G. *Anal. Chem.* **2004**, *76*, 1134−1144.
(23) Tirumalai, R. S.; Chan, K. C.; Prieto, D. A.; Issaq, H. J.; Conrads, T. P.; Veenstra, T. D. *Mol. Cell Proteomics* **2003**, *2*, 1096−1103.
(24) Zhou, M.; Lucas, D. A.; Chan, K. C.; Issaq, H. J.; Petricoin, E. F., 3rd; Liotta, L. A.; Veenstra, T. D.; Conrads, T. P. *Electrophoresis* **2004**, *25*, 1289−98.
(25) Putnam, F. W., *The Plasma Proteins Structure, Function, and Genetic Control*; Academic Press: New York, 1975−1987; p 1−55.
(26) Schorge, J. O.; Drake, R. D.; Lee, H.; Skates, S. J.; Rajanbabu, R.; Miller, D. S.; Kim, J. H.; Cramer, D. W.; Berkowitz, R. S.; Mok, S. C. *Clin. Cancer Res.* **2004**, *10*, 3474−3478.
(27) Liu, H.; Berger, S. J.; Chakraborty, A. B.; Plumb, R. S.; Cohen, S. A. *J. Chromatogr. B Analyt Technol. Biomed. Life Sci.* **2002**, *782*, 267−289.
(28) Wagner, K.; Miliotis, T.; Marko-Varga, G.; Bischoff, R.; Unger, K. K. *Anal.Chem.* **2002**, *74*, 809−820.
(29) Kachman, M. T.; Wang, H.; Schwartz, D. R.; Cho, K. R.; Lubman, D. M. *Anal.Chem.* **2002**, *74*, 1779−1791.
(30) Watson, R.; Nadler, T. *Methods Mol. Biol.* **2004**, *244*, 425−445.
(31) Huber, C. G.; Walcher, W.; Timperio, A. M.; Troiani, S.; Porceddu, A.; Zolla, L. *Proteomics* **2004**.
(32) Lescuyer, P.; Hochstrasser, D. F.; Sanchez, J. C. *Electrophoresis* **2004**, *25*, 1125−1135.
(33) Weissinger, E. M.; Kaiser, T.; Meert, N.; De Smet, R.; Walden, M.; Mischak, H.; Vanholder, R. C. *Nephrol. Dial Transplant* **2004**, *19*, 3068−3077.
(34) Faggiano, P.; D'Aloia, A.; Bignotti, T.; Dei Cas, L. *Ital. Heart J.* **2003**, *4*, 497−499.

PR049761H