Anal. Chem. **2010,** *82,* 3007–3015

A Fully Automated System with Online Sample Loading, Isotope Dimethyl Labeling and Multidimensional Separation for High-Throughput Quantitative Proteome Analysis

Fangjun Wang,[†] Rui Chen,[†] Jun Zhu,[†] Deguang Sun,[‡] Chunxia Song,[†] Yifeng Wu,[‡] Mingliang Ye,[†] Liming Wang,[‡] and Hanfa Zou*[,†]

CAS Key Lab of Separation Sciences for Analytical Chemistry, National Chromatographic Research and Analysis Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China, and The Second Affiliated Hospital of Dalian Medical University, Dalian 116027, China

Multidimensional separation is often applied for large-scale qualitative and quantitative proteome analysis. A fully automated system with integration of a reversed phase-strong cation exchange (RP-SCX) biphasic trap column into vented sample injection system was developed to realize online sample loading, isotope dimethyl labeling and online multidimensional separation of the proteome samples. Comparing to conventionally manual isotope labeling and off-line fractionation technologies, this system is fully automated and time-saving, which is benefit for improving the quantification reproducibility and accuracy. As phosphate SCX monolith was integrated into the biphasic trap column, high sample injection flow rate and high-resolution stepwise fractionation could be easily achieved. ~ 1000 proteins could be quantified in ~30 h proteome analysis, and the proteome coverage of quantitative analysis can be further greatly improved by prolong the multidimensional separation time. This system was applied to analyze the different protein expression level of HCC and normal human liver tissues. After three times replicated analysis, finally 94 up-regulated and 249 down-regulated (HCC/Normal) proteins were successfully obtained. These significantly regulated proteins are widely validated by both gene and proteins expression studies previously. Such as some enzymes involved in urea cycle, methylation cycle and fatty acids catabolism in liver were all observed down-regulated.

Mass spectrometry (MS) has been widely applied in protein quantification of various biological samples.[1−3] And it is emerging as a powerful tool for elucidation of different physiological and pathological processes in biological systems as well as discovery of useful protein biomarkers for early detection of serious diseases such as cancers.[4−8] Up to now, a number of technologies were applied to qualitative and quantitative proteome analysis of different biological samples, such as cell lines, tissues, and serum. Two-dimensional electrophoresis (2DE) is the most frequently used technology for protein quantifications of clinical samples.[9−11] As 2DE is unable to quantify very acidic or basic proteins, extremely large or small proteins and membrane proteins, MS is a good alternative without these limits for high throughput proteome analysis.[1]

In order to accurately quantify proteins by MS in shotgun proteome technology, two strategies are usually applied. The first one is label-free approach, which obtains the relative quantity of each peptide among samples by comparing the corresponding peak intensity in parallel nanoflow liquid chromatography coupled with tandem mass spectrometry (μLC-MS/MS) analyses.[12−14] The advantage of label-free approach is that no chemical labeling is required and several samples can be compared simultaneously. However, the poor reproducibility of μLC-MS/MS analysis might compromise the accuracy of quantification. The other one is stable isotope labeling, which usually codes different samples with different isotope reagents at first, and then the samples are mixed and analyzed by μLC-MS/MS. The relative quantity of each peptide among different samples is obtained by comparing the

* To whom correspondence should be addressed. Phone: +86-411-84379610. Fax: +86-411-84379620. E-mail: hanfazou@dicp.ac.cn.

[†] Chinese Academy of Sciences.

[‡] The Second Affiliated Hospital of Dalian Medical University.

(1) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198–207.
(2) Ong, S. E.; Mann, M. *Nat. Chem. Biol.* **2005**, *1*, 252–262.
(3) Abu-Farha, M.; Elisma, F.; Zhou, H. J.; Tian, R. J.; Zhou, H.; Asmer, M. S.; Figeys, D. *Anal. Chem.* **2009**, *81*, 4585–4599.
(4) Chen, R.; Pan, S.; Brentnall, T. A.; Aebersold, R. *Mol. Cell. Proteomics* **2005**, *4*, 523–533.
(5) Kuramitsu, Y.; Nakamura, K. *Proteomics* **2006**, *6*, 5650–5661.
(6) Bertucci, F.; Birnbaum, D.; Goncalves, A. *Mol. Cell. Proteomics* **2006**, *5*, 1772–1786.
(7) Zheng, J. J.; Gao, X.; Beretta, L.; He, F. C. *Proteomics* **2006**, *6*, 1716–1718.
(8) Seow, T. K.; Liang, R.; Leow, C. K.; Chung, M. C. M. *Proteomics* **2001**, *1*, 1249–1263.
(9) Seow, T. K.; Ong, S. E.; Liang, R.; Ren, E. C.; Chan, L.; Ou, K.; Chung, M. C. M. *Electrophoresis* **2000**, *21*, 1787–1813.
(10) Liang, C.; Leow, C. K.; Neo, J. C. H.; Tan, G. S.; Lo, S. L.; Lim, J. W. E.; Seow, T. K.; Lai, P. B. S.; Chung, M. C. M. *Proteomics* **2005**, *5*, 2258–2271.
(11) Sun, W.; Xing, B. C.; Sun, Y.; Du, X. J.; Lu, M.; Hao, C. Y.; Lu, Z. A.; Mi, W.; Wu, S. F.; Wei, H. D.; Gao, X.; Zhu, Y. P.; Jiang, Y.; Qian, X. H.; He, F. C. *Mol. Cell. Proteomics* **2007**, *6*, 1798–1808.
(12) Wang, W.; Zhou, H.; Lin, H.; Roy, S.; Shaler, T. A.; Hill, L. R.; Norton, S.; Kumar, P.; Anderle, M.; Becker, C. H. *Anal. Chem.* **2003**, *75*, 4818–4826.
(13) Radulovic, D.; Jelveh, S.; Ryu, S.; Hamilton, T. G.; Foss, E.; Mao, Y.; Emili, A. *Mol. Cell. Proteomics* **2004**, *3*, 984–997.
(14) Wang, F. J.; Ye, M. L.; Dong, J.; Tian, R. J.; Hu, L. H.; Han, G. H.; Jiang, X. N.; Wu, R. A.; Zou, H. F. *J. Sep. Sci.* **2008**, *31*, 2589–2597.

10.1021/ac100075y
Published on Web 03/15/2010

intensity of specific isotopic peaks in the same μLC-MS/MS analysis.[15−18] Stable isotope labeling with amino acids in cell culture (SILAC) is the most popular strategy among all labeling methods, which could compare the expression levels of thousands of proteins in a single labeling experiment using data analysis software MSQuant or MaxQuant.[19−21] Recently, the proteins in living mouse could be isotope labeled by dosing a diet containing either the natural or the $^{13}C_6$ substituted version of lysine, and the SILAC-mouse approach exhibited a versatile tool for determining protein functions under complex in vivo conditions of model mouse.[22] However, SILAC needs expensive reagents and it is impossible to obtain SILAC clinical samples of human, such as serum, urine, and tissues.

Stable isotope dimethyl labeling at peptide level is another popular strategy for protein quantification, which globally labels the N-terminus and ε-amino group of lysine with dimethyl groups using inexpensive reagents formaldehyde and sodium cyanoborohydride.[23−25] Heck and co-workers optimized this strategy and standardized three types of labeling protocols, in-solution, online with LC-MS, and on-column using solid-phase extraction (SPE) columns.[26−28] Standard isotope labeling procedures usually need many manually handling steps such as sample desalting, adding labeling regents, incubation, quenching labeling, sample redesalting, and so on, which may result in sample loss and poor reproducibility. Moreover, the manually handling procedure is time-consuming and labor-intensive. To circumvent these limitations, Raijmakers et al. established an online sequential isotope dimethyl labeling methods, which allowed sample loading, isotope dimethyl labeling and 1-dimensional (1D) LC-MS/MS analysis in fully automated manner. They applied this method to determine the different protein expression between bovine liver and spleen 20 S core proteasome complexes and discovered that the tissue-specific immunoproteasome assembly is more diverse than previously assumed.[29]

Hepatocellular carcinoma (HCC) is the most common primary cancer of the liver, especially in Africa, Southeast Asia, and China.[7,8] Stable isotope labeling quantification strategy was also applied to study the HCC samples for identifying differentially regulated proteins. Chaerkady et al quantified over 600 proteins among which 59 and 92 proteins were found up- and down-regulated, respectively, in HCC as compared to adjacent normal tissue by using isobaric tagging for relative and absolute quantification (iTRAQ) technique.[30] Chen et al quantified 2335 proteins and found 91 and 61 proteins were up- and down-regulated in HCC cell line HCCLM6 as compared to MHCC97L. And their results provide valuable insights into the HCC metastasis mechanisms as well as potential biomarkers for detection of HCC metastasis.[31] However, all of these isotope-labeling approaches were manually performed and off-line SCX chromatography or SDS-PAGE was used for peptides or proteins prefractionation. It is known that online multidimensional separation displays advantages such as higher sensitivity, minimal loss of sample, no vial contamination, and no sample dilution effect.[32,33] Especially in μLC separation, online multidimensional separation exhibits an efficient way for prefractionation of limited amount of biological samples.

In our previous work, we have developed a SCX-RP biphasic column with a segment of 10 cm-long phosphate SCX monolith and another 65 cm-long C12 hydrophobic monolith within a single 100 μm i.d. capillary column for shotgun proteome analysis.[32] In this study, we constructed a RP-SCX biphasic column with 7 cm-long segment packed with C18 particles and 7 cm-long segment of SCX monolith within a single 200 μm i.d. capillary column. This biphasic column was used as a trap column in vented sample injection system for sample loading and labeling.[33] At first, the tryptic digests of the protein samples extracted from the HCC and normal liver tissues were loaded onto RP segment of the biphasic trap column and labeled with different types of isotope dimethyl regents in sequence. Then, all the enriched peptides were transferred to SCX segment of the biphasic column, followed with online multidimensional separation by stepwise salt eluting of the peptides to 15 cm-long RP separation column and binary gradient μLC-MS/MS analysis. Over 1000 proteins could be quantified when the online multidimensional separation with about 29 h analysis time was applied. When the analysis time was increased to 63 h, over 1700 proteins could be quantified. After three replicated runs, finally 94 significantly up-regulated and 249 significantly down-regulated proteins were successfully observed. About 40% of the differentially expressed proteins in this study have been previously observed in the literature by cDNA miroarray, 2DE, and MS analyses. Therefore, this sequential isotope labeling and SCX fractionation technology with RP-SCX biphasic

(15) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17*, 994–999.
(16) Han, D. K.; Eng, J.; Zhou, H. L.; Aebersold, R. *Nat. Biotechnol.* **2001**, *19*, 946–951.
(17) Zhang, R. J.; Sioma, C. S.; Wang, S. H.; Regnier, F. E. *Anal. Chem.* **2001**, *73*, 5142–5149.
(18) Regnier, F. E.; Riggs, L.; Zhang, R. J.; Xiong, L.; Liu, P. R.; Chakraborty, A.; Seeley, E.; Sioma, C.; Thompson, R. A. *J. Mass Spectrom.* **2002**, *37*, 133–145.
(19) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. *Mol. Cell. Proteomics* **2002**, *1*, 376–386.
(20) Graumann, J.; Hubner, N. C.; Kim, J. B.; Ko, K.; Moser, M.; Kumar, C.; Cox, J.; Scholer, H.; Mann, M. *Mol. Cell. Proteomics* **2008**, *7*, 672–683.
(21) Cox, J.; Matic, I.; Hilger, M.; Nagaraj, N.; Selbach, M.; Olsen, J. V.; Mann, M. *Nat. Protocols* **2009**, *4*, 698–705.
(22) Kruger, M.; Moser, M.; Ussar, S.; Thievessen, I.; Luber, C. A.; Forner, F.; Schmidt, S.; Zanivan, S.; Fassler, R.; Mann, M. *Cell* **2008**, *134*, 353–364.
(23) Hsu, J. L.; Huang, S. Y.; Chow, N. H.; Chen, S. H. *Anal. Chem.* **2003**, *75*, 6843–6852.
(24) Huang, S. Y.; Tsai, M. L.; Wu, C. J.; Hsu, J. L.; Ho, S. H.; Chen, S. H. *Proteomics* **2006**, *6*, 1722–1734.
(25) Ji, C. J.; Li, L. J.; Gebre, M.; Pasdar, M.; Li, L. *J. Proteome Res.* **2005**, *4*, 1419–1426.
(26) Boersema, P. J.; Aye, T. T.; van Veen, T. A. B.; Heck, A. J. R.; Mohammed, S. *Proteomics* **2008**, *8*, 4624–4632.
(27) Boersema, P. J.; Raijmakers, R.; Lemeer, S.; Mohammed, S.; Heck, A. J. R. *Nat. Protocols* **2009**, *4*, 484–494.
(28) Boersema, P.; Foong, L. Y.; Ding, V.; Lemeer, S.; Mohammed, S.; Raijmakers, R.; van Breukelen, B.; Boekhorst, J.; Choo, A. B. H.; Heck, A. J. R. *Mol. Cell. Proteomics* **2009**, S47–S47.
(29) Raijmakers, R.; Berkers, C. R.; de Jong, A.; Ovaa, H.; Heck, A. J. R.; Mohammed, S. *Mol. Cell. Proteomics* **2008**, *7*, 1755–1762.
(30) Chaerkady, R.; Harsha, H. C.; Nalli, A.; Gucek, M.; Vivekanandan, P.; Akhtar, J.; Cole, R. N.; Simmers, J.; Schulick, R. D.; Singh, S.; Torbenson, M.; Pandey, A.; Thuluvath, P. J. *J. Proteome Res.* **2008**, *7*, 4289–4298.
(31) Chen, N.; Sun, W.; Deng, X. Y.; Hao, Y. W.; Chen, X. L.; Xing, B. C.; Jia, W.; Ma, J.; Wei, H. D.; Zhu, Y. P.; Qian, X. H.; Jiang, Y.; He, F. C. *Proteomics* **2008**, *8*, 5108–5118.
(32) Wang, F. J.; Dong, J.; Ye, M. L.; Jiang, X. G.; Wu, R.; Zou, H. F. *J. Proteome Res.* **2008**, *7*, 306–310.
(33) Wang, F. J.; Dong, J.; Jiang, X. G.; Ye, M. L.; Zou, H. F. *Anal. Chem.* **2007**, *79*, 6599–6606.

trap column exhibits a promising way for comparative protein expression analysis.

EXPERIMENTAL SECTION

Materials. Daisogel ODS-AQ (5 μm, 12 nm pore) was purchased from DAISO Chemical CO., Ltd. (Osaka, Japan). PEEK tubing, sleeves, microtee, and microcross were obtained from Upchurch Scientific (Oak Harbor, WA). Fused silica capillaries with 50 and 200 μm i.d. were purchased from Yongnian Optical Fiber Factory (Hebei, China), and with 75 μm i.d. from Polymicro Technologies (Phoenix, AZ). All the water used in experiments was purified using a Mill-Q system from Millipore Company (Bedford, MA). Ethylene glycol methacrylate phosphate (EGMP), γ-methacryloxypropyltrimethoxysilane (γ-MAPS), methylene bis-acrylamide, formaldehyde, sodium cyanoborohydride, dithiothreitol (DTT), iodoacetamide, and trypsin were obtained from Sigma (St. Louis, MO). Azobisisobutyronitrile (AIBN) was obtained from Shanghai Fourth Reagent Plant (Shanghai, China). Formic acid (FA) was obtained from Fluka (Buches, Germany). Acetonitrile (ACN, HPLC grade) was from Merck (Darmstadt, Germany).

Sample Preparation. The HCC and normal human liver tissues were obtained from the Second Affiliated Hospital of Dalian Medical University (Dalian, China). The utilization of human tissues complied with guideline of Ethics Committee of the Hospital. The normal liver tissues have been verified by histopathological examination, which excluded the presence of invading or microscopic metastatic cancer cells. The protein extraction procedures were the same as our previous works.[34,35] Briefly, human liver tissue was placed in an ice-cold homogenization buffer consisting of 8 M urea, 4% CHAPS w/v, 65 mM DTT, a mixture of protease inhibitor (Complete Mini protease inhibitor cocktail tablets, 1 tablet for 10 mL homogenization buffer) and 100 mM NH_4HCO_3 at pH 7.8. After homogenized by a Potter-Elvejhem homogenizer with a Teflon piston, sonicated for 100 W $\times$ 30 s and centrifuged at 25 000*g* for 1 h, the protein concentration was determined by Bradford assay. In our study, the same amount of protein extracts from five HCC or normal liver tissues were mixed together, respectively. Then, the proteins in the HCC or normal samples were precipitated by chloroform/methanol precipitation as described by Wessel et al.[36] After washing with methanol, the pellets were resuspended in 1 mL denaturing buffer containing 50 mM Tris/HCl (pH 8.1) and 8 M urea and the protein concentration was determined again by Bradford assay. The protein samples were reduced by DTT at 37 °C for 2 h and alkylated by iodoacetamide in dark at room temperature for 40 min. Then the solutions were diluted to 1 M urea with 50 mM Tris/HCl (pH 8.1). Finally, trypsin was added with weight ratio of trypsin to protein at 1/25 and incubated at 37 °C overnight. Then, the tryptic digests were purified with a homemade C18 solid phase cartridge and exchanged into buffer A (0.1% formic acid water solution). Finally, the samples were stored at −30 °C before usage.

Column Preparation. To prepare the RP-SCX biphasic column, the 200 μm i.d. capillary column was pretreated with

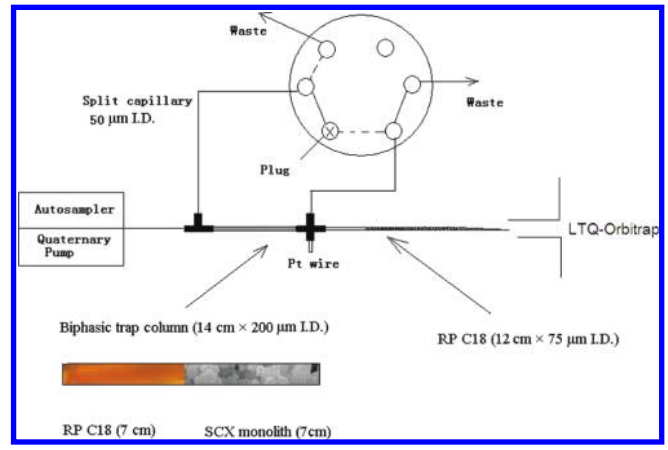

Figure 1. Schematic diagram of the system with automated sample injection and isotope dimethyl labeling. Solid line: sample injection and sequential isotope dimethyl labeling onto the RP-SCX biphasic trap column; dashed line: stepwise salt eluting of sample to analytical column and gradient RPLC analysis.

γ-MAPS as described elsewhere.[33] 25 cm-long pretreated capillary column was filled with 10 cm-long SCX monolithic polymerization mixture containing 80 μL ethylene glycol methacrylate phosphate, 60 mg methylene bis-acrylamide, 270 μL dimethylsulfoxide, 200 μL dodecanol, 50 μL *N*,*N*′-dimethylformamide, and 2 mg AIBN. Then, the capillary column was sealed at both ends with rubber stoppers, submerged into a water bath and allowed to react for 12 h at 60 °C. The resultant monolithic capillary column was washed with methanol using a HPLC pump to remove unreacted monomers and porogens and dried with nitrogen gas.[33] Finally, the SCX monolithic segment was cut at 7 cm length, followed with packing C18 material into the empty capillary to reach a length of about 7 cm adjacent the SCX monolith by a homemade pneumatic pressure cell at constant nitrogen gas pressure of about 580 psi with a slurry packing method.

To prepare the separation column, one end of a 75 μm i.d. capillary was first manually pulled to a fine point of ~3 μm with a flame torch, and then the C18 particles were packed until the packing section reached the length of 15 cm by the same method as described above.

RP Gradient μLC Separation. The HPLC system (Thermo, San Jose, CA) consisted of a degasser and a quaternary Surveyor MS pump. 0.1% FA aqueous solution (buffer A) and ACN with 0.1% formic acid (buffer B) were used for gradient separation. In all gradient separation, the flow rate after splitting was adjusted to ~300 nL/min (dashed line mode in Figure 1). Two types of RP separation gradient were used in our study. The first one was 95-min gradient, which was developed from 0 to 10% buffer B for 5 min, from 10 to 35% for 90 min, and from 35 to 80% for 5 min. After flushing 80% buffer B for 10 min, the separation system was equilibrated by buffer A for 15 min. The second one was 155 min gradient, which was the same as the 95 min gradient except for from 10 to 35% buffer B for 150 min.

Mass Spectrometry Analysis. The MS analysis was performed on LTQ-Orbitrap mass spectrometer (Thermo, San Jose, CA) at a resolution of 60 000. The temperature of the ion transfer capillary was set at 200 °C. The spray voltage was set at 1.8 kV and the normalized collision energy was set at 35.0%. One microscan was set for each MS and MS/MS scan. All MS and

(34) Han, G. H.; Ye, M. L.; Zhou, H. J.; Jiang, X. N.; Feng, S.; Jiang, X. G.; Tian, R. J.; Wan, D. F.; Zou, H. F.; Gu, J. R. *Proteomics* **2008**, *8*, 1346–1361.
(35) Chen, R.; Jiang, X. N.; Sun, D. G.; Han, G. H.; Wang, F. J.; Ye, M. L.; Wang, L. M.; Zou, H. F. *J.Proteome Res.* **2009**, *8*, 651–661.
(36) Wessel, D.; Flügge, U. I. *Anal. Biochem.* **1984**, *138*, 141–143.

MS/MS spectra were acquired in the data dependent mode. Survey full scan MS was acquired from m/z 400 to 2000, and seven most intense ions were selected for MS/MS scan by collision induced dissociation. The target ion setting was 5e5 for the Orbitrap, with a maximum fill-time of 500 ms. MS/MS scans were acquired in the LTQ with a target ion setting of 3e4 and a maximum fill-time of 100 ms. The dynamic exclusion function was set as follows: repeat count 2, repeat duration 30 s, and exclusion duration 90 s.

Sequential isotope Labeling and Online Multidimensional Separation. The configuration for automated sample injection using biphasic trap column was consisted with an autosampler, a six-port/two-position switching valve, a microtee and a microcross as shown in Figure 1. During sample or labeling reagent injection, the switching valve was switched to close the splitting flow and the flow through from the trap column was switched to waste (solid line mode in Figure 1). At first, 20 μL tryptic digest of the first sample was injected at a flow rate of 6 μL/min for 15 min. Second, 50 μL light isotope labeling reagent containing 0.04% formaldehyde, 6 mM cyanoborohydride, and 50 mM sodium phosphate (pH 7.5) was injected at a flow rate of 6 μL/min for 20 min. Third, 20 μL tryptic digest of the second sample was injected at a flow rate of 6 μL/min for 15 min. Fourth, 50 μL heavy isotope labeling reagent containing 0.04% deuterated formaldehyde, 6 mM cyanoborohydride, and 50 mM sodium phosphate (pH 7.5) was injected at a flow rate of 6 μL/min for 20 min. In all the injection procedures, 0.1% FA aqueous solvent (buffer A) was used as injection solvent, and 8 min additional equilibrium by buffer A was already included in each procedure. Then, the switching valve was switched to open the splitting flow (dashed line mode in Figure 1).

For one-dimensional separation, the peptides enriched onto the RP segment of the biphasic trap column was transferred to SCX segment with elution by 80% buffer B for 10 min, followed to separation column with elution by 100% buffer C (1000 mM NH_4Ac, pH 2.8) for 20 min. After re-equilibrated for 15 min with buffer A, a 95 min RP gradient μLC-MS/MS analysis on separation column was conducted.

For 29 h online multidimensional separation, a 95 min RP gradient μLC-MS/MS was applied at first (0 mM fraction). Therefore, all of the peptides enriched onto the RP segment of the biphasic trap column were transferred to the SCX monolith segment, and some unretained peptides were separated by the 15 cm long separation column and detected by MS. Then, a series stepwise elution (generated by buffer A and buffer C) with salt concentrations of 50, 100, 150, 200, 250, 300, 350, 400, 500, and 1000 mM NH_4Ac was used to gradually elute peptides from SCX segment onto the C18 separation column. Each salt step lasts 10 min except last one for 20 min. After whole system was re-equilibrated for 15 min with buffer A, the 95 min binary RP gradient μLC-MS/MS analysis as described above was applied to separate peptides prior to MS detection in each cycle.

To increase the proteome coverage, a 63 h online multidimensional separation was also applied in our study. The stepwise salt elution was developed as 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 500, and 1000 mM. And the 155 min separation gradient μLC-MS/MS analysis as described above was applied in each cycle. The other procedures were the same as the 29 h online multidimensional separation.

Protein Identification and Quantification. All the MS/MS spectra in one acquired raw file were converted to single *.mgf file using DTASupercharge (v2.0a7).[37] Then the *.mgf file was searched against the International Protein Index (IPI) human database (v3.52, 73928 entries) using Mascot Version 2.1 (Matrix Science). To evaluate the false discovery rate (FDR), reversed sequences were appended to the database. Cysteine residues were searched as static modification of +57.0215 Da, methionine residues as variable modification of +15.9949 Da. And both light and heavy dimethylation of peptide amino termini and lysine residues were set as variable modification of +28.0313 Da and +32.0564 Da, respectively. Peptides were searched using fully tryptic cleavage constraints and up to two missed cleavages sites were allowed for tryptic digestion. The mass tolerances were 10 ppm for parent masses and 0.8 Da for fragment masses. Peptides with Mascot score ≥25 (rank 1, $P \leq 0.05$) were used for protein quantification.

Protein quantification was performed using a dimethyl-adapted version of MSQuant (v2.0a81). Peptide ratios were obtained by calculating the extracted ion chromatograms (XIC) of the light and heavy forms of the peptide using the monoisotopic peaks only and protein ratios were calculated from the average of the all quantified peptides.[31,37,38] Then, all the MSQuant outputs of the same online multidimensional separation were imported in StatQuant (v1.2.2) and the quantified proteins were normalized against the log2 of the median ratio of all peptides quantified.[38,39]

RESULTS AND DISCUSSION

Labeling Efficiency and Quantification Accuracy. Raijmakers et al first realized online sequential isotope dimethyl labeling and 1D RP gradient separation.[29] However, the separation capability of 1D separation is always limited when complex biological samples are analyzed. In order to increase the proteome coverage in large-scale protein quantification, especially to quantify the proteins with relatively low abundance, it is necessary to increase the separation capability of the LC system. As online multidimensional separation exhibits powerful separation capability as well as advantages such as higher sensitivity, minimal loss of sample, and so on, it is a good choice to couple online isotope labeling with online multidimensional separation to increase the quantification proteome coverage. However, as the pH value of the online isotope dimethyl labeling regents is optimized at 7.5, it is improper to conduct peptides labeling onto SCX trap column.[27] Therefore, a RP-SCX biphasic trap column was developed and applied in the vented sample injection system as shown in Figure 1. The RP segment of the biphasic column is used for isotope dimethyl labeling similar to the work of Raijmakers et al.[29] The labeled peptides were transferred to the SCX segment by ACN elution. Then, online multidimensional separation can be conducted just as our previous works.[32,33] As the phosphate SCX monolith has

(37) Mortensen, P.; Gouw, J. W.; Olsen, J. V.; Ong, S.-E.; Rigbolt, K. T. G.; Bunkenborg, J.; Cox, J. r.; Foster, L. J.; Heck, A. J. R.; Blagoev, B.; Andersen, J. S.; Mann, M. *J. Proteome Res.* **2010**, *9*, 393–403.
(38) Lemeer, S.; Jopling, C.; Gouw, J.; Mohammed, S.; Heck, A. J. R.; Slijper, M.; den Hertog, J. *Mol. Cell. Proteomics* **2008**, *7*, 2176–2187.
(39) van Breukelen, B.; van den Toorn, H. W. P.; Drugan, M. M.; Heck, A. J. R. *Bioinform.* **2009**, *25*, 1472–1473.

Table 1. Number of Quantified Proteins and Peptides by Each Analysis

	control-1	control-2	control-3	29 h-1	29 h-2	29 h-3	63 h
pro. num.	954	823	1023	1181	1096	1119	1767
pep. num.	4425	3855	5303	5473	4826	4982	12129
FDR (pro.)	2.20%	1.58%	1.86%	1.61%	0.73%	1.61%	1.19%
FDR (pep.)	0.54%	0.36%	0.38%	0.38%	0.16%	0.38%	0.20%

superior permeability and sample loading capacity, high-resolution online salt stepwise fractionation can be realized to increase the separation capability of the μLC system efficiently.

To evaluate the performance of the system for automatically sequential labeling and online multidimensional μLC-MS/MS analysis, tryptic digests of the same amount (15 μg) of extracted protein from normal human liver tissues were injected in sequence as testing sample and labeled with light and heavy dimethyl groups, respectively. Then, a 29 h online multidimensional separation with 11 fractions was conducted as described above. After a database search by Mascot, finally 30738 peptides (corresponding to 1115 proteins) with Mascot score $\geq$25 (rank 1, $P \leq 0.05$) were obtained at FDR 1.16% and 29 483 peptides (95.92%) were successfully labeled with dimethyl groups. In all the labeled peptides, 320 peptides only have dimethylated lysine, 15 668 peptides only have dimethylated N termini, and 13 495 peptides have both dimethylated lysine and N termini. Among the 15 668 peptides that only have the dimethylated N termini, 11 821 peptides do not have lysine residue, and over 85% ((11 821 + 13 495)/29 483) of the labeled peptides with the dimethylated N termini and lysine residue were completely labeled. The numbers of identified peptides with light and heavy dimethyl labels are 15 135 and 14 348, respectively, and the ratio is 1.05. Therefore, the injected peptides in the two labeling steps are nearly all exclusively labeled with different types of isotope dimethyl groups. Then, these identified peptides were paired with the MS spectra using MSQuant and the quantification results were combined and normalized by StatQuant as described above. Finally, 4425 peptides corresponding to 954 proteins were successfully quantified as summarized in Table 1 (control-1), and the identification FDR were 2.20 and 0.54% for proteins and peptides, respectively. 597 proteins (62.6%) were quantified by at least two peptides. The typical isotope isoforms of peptides R.MFLSFPTTK.T and K.TCNVLVALEQQSPDIAQCVHLDR.N in MS spectra were given in Figure 2 A and B. Obviously, the peptides were reliably labeled with dimethyl groups and could be quantified by comparing the intensity of the isotope isoforms. The log2 ratios of the quantified proteins were shown in Figure 3 A. Though most of the log2 ratios were around 0 (ratio 1), some of the values were widely scattered between -6 and 6 and the average value of the log2 ratios is 0.220 (theoretically 0), which is not acceptable. Among the 179 proteins that have changed their quantity two folds (theoretically 1), 93 proteins (52.0%) were quantified by at least two peptides. Therefore, if we just used the filtering criterion that proteins must quantified by at least two peptides, 52.0% false quantification can pass the filter criterion but accurate quantification of many proteins by 1 peptide is filtered. Therefore, it is necessary to establish a more proper filtering criterion in isotope labeling quantitative proteome analysis.

Then, two replicated online isotope dimethyl labeling and multidimensional separation as described above were conducted

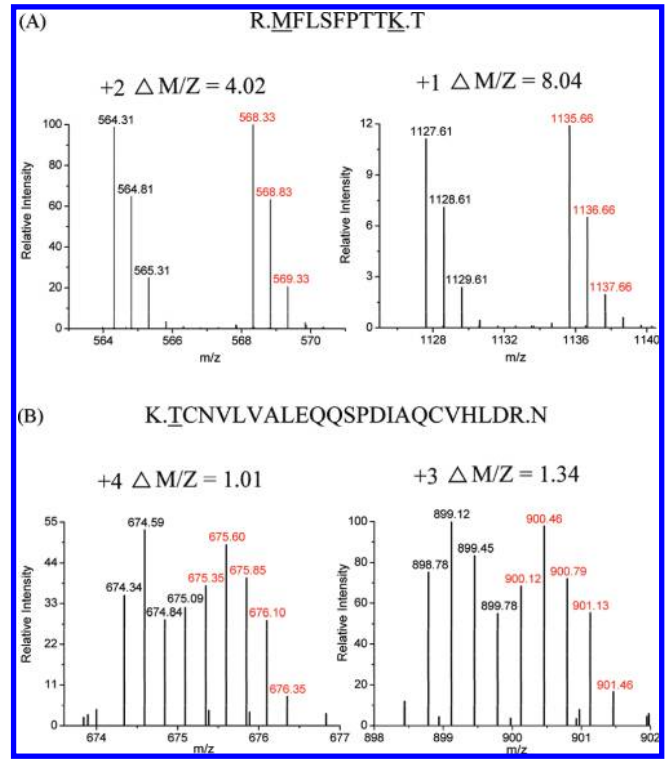

Figure 2. The peaks for isotope isoforms of peptides R.MFLSFPTTK.T (singly and doubly charged) (A) and K.TCNVLVALEQQSPDIAQCVHLDR.N (triply and quadruply charged) (B) in MS spectra.

and finally 3855 and 5303 peptides and 823 and 1023 proteins were successfully quantified, respectively, as summarized in Table 1 (control-2 and control-3). It was observed that totally 588 proteins were quantified in all of the three replicated analyses and the relative quantity of 394 proteins had relative standard deviation (RSD) < 50%. The log2 ratios (average value of the two analyses) of the 394 proteins are shown in Figure 3 B, and it can be seen that there is no any protein with log2 ratio >1 and only 4 protein with log2 ratio < -1. Therefore, only 1.0% (4) of the quantified proteins in replicated analyses has changed their quantity more than two folds (theoretically 1). And the average value of the log2 ratios is -0.05, which is much better than the result obtained by just one time analysis. In the other 194 proteins that quantified in all of the three replicated analyses but have RSD > 50%, only two proteins (1.0%) have simultaneously changed relative quantity more than two folds. Therefore, these results indicate the quantification accuracy is significantly improved by three times replicated analyses. Similar results were also obtained in online sequential isotope dimethyl labeling combined with one-dimensional analysis using the strategy developed by Raijmakers et al[29] (data not shown). Then three replicated analyses were conducted in the quantitative proteome analysis of the HCC and normal liver tissues lately. And the proteins quantified in all of the three analyses with RSD < 50% were considered as reliable quantifica-

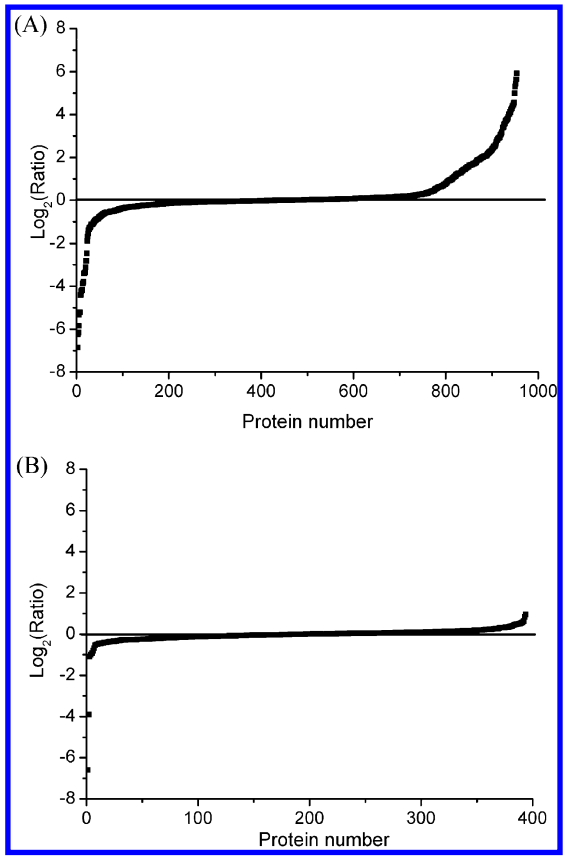

Figure 3. Protein expression ratios (normal/normal) in log2 obtained by the sequential light/heavy isotope labeling coupled with 29 h online multidimensional separation (A); three times replicated quantification analyses with RSD < 50% (B).

tion. The proteins with average log2 ratio >1 or < −1 were considered as significantly up-regulated or down-regulated. Furthermore, the proteins exhibited simultaneously up-regulated (log2 ratio >1) and down-regulated (log2 ratio <1) in all of the three replicated analyses were also considered as significantly regulated proteins even if their RSD > 50%.

Application to Differential Analysis of HCC and Normal Liver Tissues. Tryptic digests of the same amount of proteins (15 μg) extracted from HCC and normal human liver tissues were injected in sequence and labeled with light and heavy isotope dimethyl groups, respectively. Then, the 29 h online multidimensional separation coupled with MS/MS detection was conducted as described above and the μLC-MS/MS base peak chromatograms for the 11 fractions were shown in Figure 4. It can be seen that the peptides are symmetrically distributed among the fractions. After database search and peptides quantification, finally 1181 proteins could be quantified as summarized in Table 1 (29 h-1). Then, the whole procedures for sequential isotope dimethyl labeling and online multidimensional μLC-MS/MS analysis were repeated two times, and 1096 and 1119 proteins were successfully quantified as summarized in Table 1 (29 h-2 and 29 h-3). Therefore, the average numbers of quantified peptides and proteins are 5094 (RSD = 6.6%, n = 3) and 1132 (RSD = 3.9%, n = 3). And in these three replicated analyses, 797 proteins were quantified in all of the three analyses (as shown in the Supporting Information (SI) S1). The relative quantity of 483 proteins have RSDs < 50%. Among the proteins with RSD > 50%, 71 proteins exhibits simultaneously up-regulated (ratio >2) or down-regulated (ratio <0.5) in the replicated analyses, and we think these proteins are also reliably quantified as described above. Therefore, totally 554 proteins are reliably quantified and 94 proteins exhibit up-regulated and 249 proteins exhibit down-regulated (as shown in SI S2).

Different expression of 218 genes were detected with cDNA microarrays in HCC tumor comparing to nontumor liver samples by Neo et al.[40] Among the 554 reliably quantified proteins in our study, the genes of 21 proteins have also been studied by Neo et al as summarized in Table 2. Among the 21 proteins, T-complex protein 1 subunit zeta (CCT6A), 40S ribosomal protein S7 (RPS7), isoform long of 14-3-3 protein beta/alpha (YWHAB), ATP-citrate synthase (ACLY), and annexin A2 isoform 1 (ANXA2) are considered as up-regulated; 3-ketoacyl-CoA thiolase peroxisomal (ACAA1), cytochrome P450 2C8 (CYP2C8), and alcohol dehydrogenase 1B (ADH1B) are determined to be down-regulated in our isotope dimethyl labeling quantification, which are consistent with the status of genes expression. The only exception is programmed cell death protein 5 (PDCD5), which is down-regulated in protein quantification (ratio 0.46), but its gene expression shows up-regulated. Though the relative quantity of the other 12 proteins are considered as nonchanged in our study, many of them also exhibit up-regulation trends as summarized in Table 2. Therefore, most of the gene regulation level (~70%) is consistent with our protein quantification results.

Chaerkady et al quantified over 600 proteins (HCC/Normal) by using iTRAQ technique.[30] Among these proteins, 263 proteins were also quantified in our experiment as shown in SI S3. Eleven proteins are up-regulated and 60 proteins are down-regulated simultaneously in the both quantification results (see SI S3). Among these proteins, extracellular superoxide dismutase [Cu−Zn] (SOD3), Serotransferrin (TF), and isoform long of 14-3-3 protein beta/alpha (YWHAB) are also reported as up-regulated proteins in other studies.[10,40,41] Argininosuccinate lyase (ASL), argininosuccinate synthase (ASS1), ornithine carbamoyl transferase (OTC), and carbamoyl phosphate synthase (CPS1) are all enzymes involved in urea cycle, which is an essential metabolic pathway of liver for detoxification of ammonia. And they are all significantly down-regulated in the HCC tissues. Aspartate amino transferase (GOT), which is involved in intermediary metabolism, associate with urea cycle, is also down-regulated. Chaerkady et al also further validated the down-regulation of CPS1 and OTC in HCC samples by Western blot and gene expression studies.[30]

The other proteins that quantified in this study are also widely validated by the previous studies,[42−46] as the results shown in SI S2. Such as three interesting metabolic enzymes, methionine adenosyltransferase (MAT1A), glycine *N*-methyltransferase (GNMT), and betaine-homocysteine *S*-methyltransferase (BHMT) that are involved in the methylation cycle in the liver are all significantly

(40) Neo, S. Y.; Leow, C. K.; Vega, V. B.; Long, P. M.; Islam, A. F. M.; Lai, P. B. S.; Liu, E. T.; Ren, E. C. *Hepatology* **2004**, *39*, 944–953.
(41) Liang, R.; Neo, J. C. H.; Lo, S. L.; Tan, G. S.; Seow, T. K.; Chung, M. C. M. *J. Chromatogr., B* **2002**, *771*, 303–328.
(42) Yokoyama, Y.; Kuramitsu, Y.; Takashima, M.; Lizuka, N.; Toda, T.; Terai, S.; Sakaida, I.; Oka, M.; Nakamura, K.; Okita, K. *Proteomics* **2004**, *4*, 2111–2116.
(43) Li, C.; Tan, Y. X.; Zhou, H.; Ding, S. J.; Li, S. J.; Ma, D. J.; Man, X. B.; Hong, Y.; Zhang, L.; Li, L.; Xia, Q. C.; Wu, J. R.; Wang, H. Y.; Zeng, R. *Proteomics* **2005**, *5*, 1125–1139.

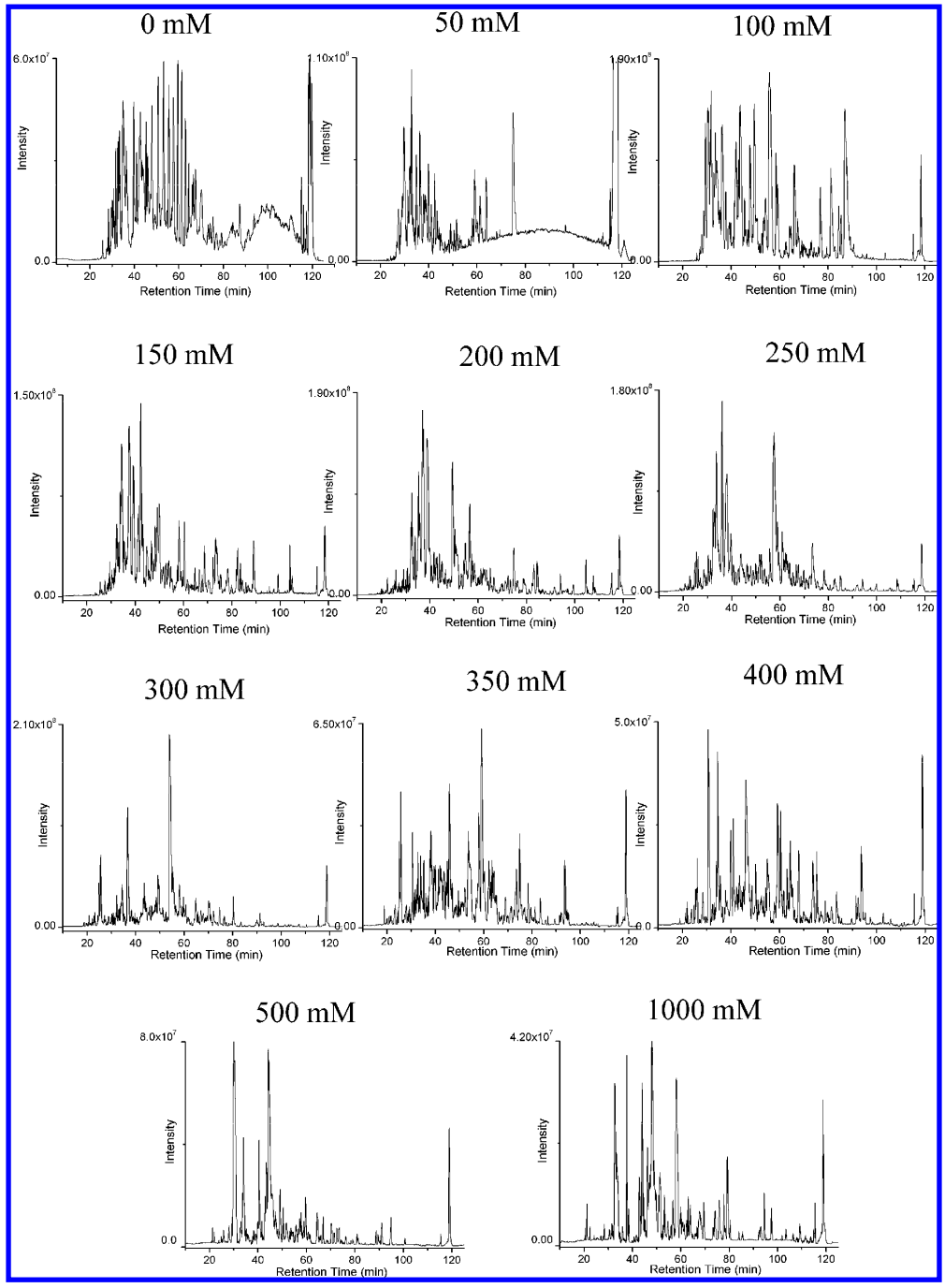

Figure 4. Base peak chromatograms of an 11-cycle online multidimensional analysis of isotope dimethyl labeled samples (HCC liver sample labeled with light dimethyl groups and normal liver sample labeled with heavy dimethyl groups). The three buffer solutions used for the chromatography were 0.1% formic acid aqueous solution (buffer A), 0.1% formic acid ACN (buffer B), and 1000 mM ammonium acetate at pH 2.7 (buffer C). Cycle 1 (0 mM) consisted of a 5 min gradient from 0 to 10% buffer B, a 90 min gradient from 10 to 35% buffer B, and a 5 min gradient from 35 to 80% buffer B, after a 10 min hold at 80% buffer B, the system was equilibrated with 100% buffer A for 15 min. Each of the next nine cycles was 150 min with the follow procedures: 10 min of X% buffer C, 15 min of 100% buffer A, then the separation gradient was just the same as cycle 1. The 10 min buffer C in cycle 2-10 was as follows: cycle 2, 5% (50 mM); cycle 3, 10% (100 mM); cycle 4, 15% (150 mM); cycle 5, 20% (200 mM); cycle 6, 25% (250 mM); cycle 7, 30% (300 mM); cycle 8, 35% (350 mM); cycle 9, 40% (400 mM); cycle 10, 50% (500 mM). Cycle 11 (1000 mM) consisted of a 20 min 100% buffer C wash followed by a 15 min 100% buffer A wash, and the separation gradient was also the same as cycle 1.

down-regulated, which is consistent with the 2DE results obtained by Liang et al.[10] And the expression level of these three enzymes has a major influence on the level of *S*-adenosylmethionine (AdoMet), a vital intermediate metabolite required for the proper functioning of the liver. It is known that HCC usually exhibits an accumulation of fatty acids and other lipids in cells. Mitochondrial

(44) Lim, S. O.; Park, S. J.; Kim, W.; Park, S. G.; Kim, H. J.; Kim, Y. I.; Sohn, T. S.; Noh, J. H.; Jung, G. *Biochem. Biophys. Res. Commun.* **2002**, *291*, 1031–1037.

(45) Kim, J.; Kim, S. H.; Lee, S. U.; Ha, G. H.; Kang, D. G.; Ha, N. Y.; Ahn, J. S.; Cho, H. Y.; Kang, S. J.; Lee, Y. J.; Hong, S. C.; Ha, W. S.; Bae, J. M.; Lee, C. W.; Kim, J. W. *Electrophoresis* **2002**, *23*, 4142–4156.

Table 2. List of the Quantified Proteins with the Genes Expression Level Studied (HCC/Normal) by Neo et al.[40a]

no.	IPI no.	gene symbol and protein name	ratio	ref 40
1	IPI:IPI00027626.3	CCT6A T-complex protein 1 subunit zeta	10.11	↑
2	IPI:IPI00013415.1	RPS7 40S ribosomal protein S7	7.92	↑
3	IPI:IPI00216318.5	YWHAB isoform long of 14-3-3 protein beta/alpha	4.69	↑
4	IPI:IPI00021290.5	ACLY ATP-citrate synthase	4.42	↑
5	IPI:IPI00418169.3	ANXA2 annexin A2 isoform 1	2.54	↑
6	IPI:IPI00444262.3	NCL cDNA FLJ45706 fis, clone FEBRA2028457, highly similar to nucleolin	1.58	↑
7	IPI:IPI00003865.1	HSPA8 isoform 1 of heat shock cognate 71 kDa protein	1.44	↑
8	IPI:IPI00010740.1	SFPQ isoform long of splicing factor, proline- and glutamine-rich	1.39	↑
9	IPI:IPI00026182.5	CAPZA2 F-actin-capping protein subunit alpha-2	1.17	↑
10	IPI:IPI00003362.2	HSPA5 HSPA5 protein	1.16	↑
11	IPI:IPI00215884.4	SFRS1 isoform ASF-1 of Splicing factor, arginine/serine-rich 1	1.15	↑
12	IPI:IPI00022774.3	VCP transitional endoplasmic reticulum ATPase	1.05	↑
13	IPI:IPI00419585.9	PPIA peptidyl-prolyl cis−trans isomerase A	0.84	↑
14	IPI:IPI00296337.2	PRKDC isoform 1 of DNA-dependent protein kinase catalytic subunit	0.84	↑
15	IPI:IPI00030131.3	TMPO isoform beta of lamina-associated polypeptide 2, isoforms beta/gamma	0.66	↑
16	IPI:IPI00010720.1	CCT5 T-complex protein 1 subunit epsilon	0.65	↑
17	IPI:IPI00290770.3	CCT3 chaperonin containing TCP1, subunit 3 isoform b	0.63	↑
18	IPI:IPI00023640.3	PDCD5 programmed cell death protein 5	0.46	↑
19	IPI:IPI00012828.3	ACAA1 3-ketoacyl-CoA thiolase, peroxisomal	0.14	↓
20	IPI:IPI00473031.6	ADH1B alcohol dehydrogenase 1B	0.03	↓
21	IPI:IPI00290301.1	CYP2C8 cytochrome P450 2C8	0.10	↓

[a] ↑, gene expression up-regulated; ↓, gene expression down-regulated.

acyl-CoA dehydrogenase (ACADSB), involved in the β-oxidation pathway of fatty acids catabolism to acetyl-CoA, is also observed down-regulated in this study, which was observed previously by Kim et al.[45] Recently, Lee et al[46] reported that cytochrome b5 (CYB5A), which is one of the phase I ncatabolic enzymes of xenobiotics in normal liver, is down-regulated in HCC, and the same result was also obtained in this study.

Though many technologies have been developed to study the protein expression level of clinical samples, strategy with high throughput and reproducibility is still needed for large-scale proteins quantification. In this study, a system with automatically sequential isotope dimethyl labeling coupled with online multidimensional separation was developed for protein quantification by using a RP-SCX biphasic trap column. In this system, all procedures including sample injection, sequential sample labeling, online stepwise fractionation, and μLC-MS/MS analysis are automatically operated, which is benefit for improving the reproducibility and accuracy of quantification. Comparing to online labeling methods previously reported,[29] this system can realize online fractionation conveniently, which is a powerful strategy to increase the proteome coverage in protein quantification. Furthermore, the proteome coverage can be efficiently increased along with the increasing of the number of stepwise fractions and the time of RP separation gradient. In our experiments, we also increased the number of salt step fractions and extended the reversed phase binary gradient to develop a 63 h online multidimensional μLC-MS/MS analysis as described in experimental section. At last, the relative quantification information of 12 129 peptides corresponding to 1767 proteins was obtained as summarized in Table 1 (63 h). Therefore, as the analysis time for online multidimensional μLC-MS/MS analysis increased 117.2%, the number of quantified peptides and proteins increased 138.1 and 56.1%, respectively. It can be seen that increasing the number of stepwise fractions and gradient separation time is a very effective methods to increase the proteome coverage in large-scale protein quantification. Another advantage of this system is that a 7 cm-long phosphate SCX monolith is integrated into the biphasic trap column. This type of phosphate SCX monolith has much higher permeability and sample loading capacity comparing to commercially available SCX materials, which is benefit for improving the resolution in SCX fractionation.[32,33] Though 14 cm-long biphasic trap column was utilized, sample loading and isotope labeling could be all operated at back pressure less than 1000 psi. By comparing with the quantification of about 1400 proteins by applying 25 off-line SCX fractionation for tryptic digest of about 180 μg proteins extracted from zebrafish embryos,[38] or about 600 proteins by using 39 off-line SCX fractionation for tryptic digest of about 160 μg proteins from human livers, about 1000 proteins were quantified by applying 11 online SCX fractionation for tryptic digest of about 30 μg proteins from human liver by our system, which exhibits better proteome coverage and detection sensitivity.

CONCLUSIONS

A system with automatically sequential isotope dimethyl labeling coupled with online multidimensional separation system was developed and applied in large-scale proteins quantification of HCC sample. The quantification accuracy of the system was validated by using the same amounts of testing sample. And it was observed that three replicated analyses was an efficient way to decrease the false quantification. As this system has powerful separation capacity in online multidimensional separation, ~1000 proteins can be quantified in ~30 h. And the proteome coverage can be also efficiently increased by increasing the number of stepwise fractions on the first dimensional trap column and the separation time on the second dimensional analytical column. Finally, 554 proteins in HCC tissues were reliably quantified by comparing to normal tissues and 343 proteins were significantly up or down-regulated. The differently regulated proteins are also widely validated previously in other studies, including both gene expression and protein quantification analyses.

(46) Lee, N. P.; Chen, L.; Lin, M. C.; Tsang, F. H.; Yeung, C.; Poon, R. T.; Peng, J. R.; Leng, X. S.; Beretta, L.; Sun, S.; Day, P. J.; Luk, J. M. *J. Proteome Res.* **2009**, *8*, 1293–1303.

ACKNOWLEDGMENT

Financial support is gratefully acknowledged from the National Natural Sciences Foundation of China (Grants 20735004 and 20975101), the China State Key Basic Research Program (Grants 2005CB522701 and 2007CB914102), the China High Technology Research Program (Grants 2006AA02A309 and 2008ZX10002-017), the Knowledge Innovation program of CAS (Grants KJCX2.YW.HO9 and KSCX2-YW-R-079), and the Knowledge Innovation program of DICP to H.Z. and the China High Technology Research Program (Grants 2008ZX1002-020) and the National Natural Sciences Foundation of China (Grants 20605022 and 90713017) to M.Y. We are grateful to Dr. Albert J. R. Heck and Shabaz Mohammed for their generous help in protein quantification by MSQuant.

SUPPORTING INFORMATION AVAILABLE

Three tables show additional information. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review January 12, 2010. Accepted March 4, 2010.

AC100075Y