Analyst

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/analyst

Integrated Platform with Combination of Online Digestion and ^{18}O Labeling for Proteome Quantification via Immobilized Trypsin Microreactor

Shen Zhang[1, 2], Huiming Yuan[1], Baofeng Zhao[1], Yuan Zhou[1], Hao Jiang[1, 2], Lihua Zhang[1*], Zhen Liang[1], Yukui Zhang[1]

1 National Chromatographic Research and Analysis Center, Key Laboratory of Separation Science for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Science, Dalian, 116023, China

2 University of Chinese Academy of Sciences, Beijing, China

* To whom correspondence should be addressed. Prof. Dr. Lihua Zhang, National Chromatographic Research and Analysis Center, Key Laboratory of Separation Science for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Science, Dalian, 116023, China. E-mail: lihuazhang@dicp.ac.cn; Tel. & fax: +0086 41184379720

Abstract

A novel automated integrated platform for quantitative proteome analysis was established, with combination of proteins on-line digestion and in-situ ^{18}O labeling by immobilized enzyme reactor (IMER), digests captured and desalted by a C18 trap column, and peptides analyzed by nanoRPLC-ESI-MS/MS. Bovine serum albumin (BSA) was used to evaluate the performance of the developed platform. Compared with traditional offline method, not only the digestion and labeling time was shortened from 36 h to just 1 h, but also the labeling efficiency was improved from 95% to 99%. Furthermore, the back-exchange from ^{18}O to ^{16}O could also be efficiently avoided by the use of IMER. The platform was further evaluated by the quantitative analysis of 100 ng ^{18}O and ^{16}O online labeled yeast sample with a mixing ratio of 1:1, and the results showed significantly improved sensitivity, reproducibility as well as quantitative accuracy than offline method. With these advantages, the integrated platform was finally applied to the quantitative profiling of 100 ng proteins extracted from two mouse hepatocarcinoma ascites syngeneic cell lines with high and low lymph node metastases rates, and 10 differentially expressed proteins were successfully found, most of which were related to tumorigenesis and tumor metastasis. All these results demonstrate the developed integrated platform can provide a new way for high efficiency ^{18}O labeling and quantitative analysis of trace amount of sample with high accuracy and high reproducibility.

1. Introduction

Accurate quantification of protein expression in biological systems is an increasingly important part of LC-MS-based proteomics research. Incorporation of differential stable isotopes for relative protein quantification has been widely used in the past decade.[1-4] Among them, trypsin-catalyzed stable isotope ^{18}O-labeling at C-terminal carboxyl groups of tryptic peptides, originally proposed by Fenselau and co-workers,[5] has been increasingly applied to quantitative proteomic studies in various biological systems for several advantages. First, the isotope tags are introduced in C-terminus enzymatically, so that any kind of protein sample can be labeled, and the secondary reactions inherent to chemical labeling are also avoided. Second, the light and heavy labeled peptide pairs coelute in high pressure liquid chromatography. Third, the method constitutes a universal strategy, which provides labeled peptides from all kinds of proteins, carrying any kind of post-translational modification.[6-8]

However, the limitations of this approach, including potentially incomplete labeling, back-exchange from ^{18}O to ^{16}O, and sample loss during labeling, have hindered its broad application in biological studies.[9-10] The issue of low efficiency for ^{18}O labeling has been efficaciously addressed by decoupling digestion and labeling, namely performing the trypsin-catalyzed labeling after digestion rather than during digestion, through which nearly all peptides can reach high labeling efficiency with optimized digestion and incubation conditions.[11] However, separation of these two reactions requires additional steps, thus, increasing the time for sample preparation as well as

the possibility of sample loss. Another potential limitation of the ^{18}O labeling technique is the possibility that back-exchange from ^{18}O to ^{16}O may occur because activated enzyme not removed can also catalyze ^{18}O to ^{16}O in C-terminal with $H_2{}^{16}O$ existence, leading to errors in relative quantitative measurements.[12] Many efforts have been devoted to solve the problem, including reducing and alkylating the enzyme,[13] heating and cooling the digestion to eliminate residual tryptic enzyme, and using filter to remove the enzyme.[14-18] However, these strategies may lead to precipitation or sample loss, making it less than ideal when trace amount of sample are investigated.

The most readily and effective method to avoid back-exchange is the use of particles with immobilized protease in labeling procedure, by the physical removal of the particles from the labeled peptides, which can simultaneously improve the ^{18}O labeling efficiency.[12, 19] Incorporation of immobilized trypsin into the initial digestion step can also minimize back exchange.[20] While the adoption of particles in these procedures brings extra centrifugation steps, and the nonspecific binding of peptides to the matrix of immobilized protease can also result in significant sample loss, making their applications towards trace amount of sample increasingly impractical.[16]

In this report, we described an automated integrated proteome quantitative analysis platform with simultaneously digestion and ^{18}O labeling that employed hydrophilic immobilized trypsin reactor in proteolytic digestion and labeling steps, to eliminate the back exchange, increase labeling efficiency, improve sensitivity and quantitative accuracy. This platform was successfully applied to the quantitative analysis of nanogram amount of Hca-F and Hca-P, two mouse hepatocarcinoma ascites syngeneic

cell lines with high and low lymph node metastases rates, showing a promising future in quantitative proteomics with its high accuracy, high sensitivity and easy manipulation.

2. Experimental

2.1 Materials and reagents

Bovine serum albumin (BSA, bovine serum), trypsin (bovine pancreas, TPCK treated), formic acid (FA), α-cyano-4-hydroxycinnamic acid (CHCA) and protease inhibitor cocktail were ordered from Sigma–Aldrich (St. Louis, MO, USA). Dithiothreitol (DTT) and iodoacetamide (IAA) were from Acros (Morris Plains, NJ, USA). Acetonitrile (ACN, HPLC grade) was bought from Merck (Darmstadt, Germany). Water was purified by a Milli-Q system (Millipore, Milford, MA, USA). $H_2{}^{18}O$ (97%) was obtained from Shanghai Research Institute of Chemical Industry (China). Hca-F and Hca-P were provided by Prof. Shujuan Shao from Dalian Medical University.

A precise syringe pump (Baoding Longer Pump Co., Baoding, China) was used to push the sample solution through the IMER. BCA protein assay kit was from Beyotime Institute of Biotechnology (Tianjin, China). Yeast Protein Extraction Reagent was ordered from Takara (Dalian, China). Fused-silica capillaries (250 μm i.d. × 375 μm o.d.) were obtained from Sino Sumtech (Handan, China). Acrylic polymer microspheres with amino groups (5 μm, 1000 Å) were bought from Shenzhen Nanomicro Technology (Shenzhen, China). Venusil XBP C18 particles (5

μm, 120 Å) were ordered from Bonna-Agela Technologies (Tianjin, China).

2.2 Sample preparation

BSA was separately dissolved in different buffers prepared with $H_2{}^{16}O$ or $H_2{}^{18}O$ (97%), respectively, with a concentration of 0.1 mg/mL, and then denatured by heating at 90°C for 10 min. After cooled to room temperature, the sample was reduced with 10 mM DTT at 56°C for 1.5 h, and alkylated by 25 mM IAA in the dark at room temperature for 40 min.

Yeast cells (strain ATCC 204508 / S288c) grown on YPD culture medium was cultured at 37°C for 24 h, and the mixture was centrifuged at the speed of 8000×g at 4°C for 2 min to precipitate cells. After washed with cold dH_2O, precipitate cells were suspended in Yeast Protein Extraction Reagent according to the standard procedure. Then the cell suspension was centrifuged at the speed of 12000×g at 4°C for 5 min. 8 M urea together with 1% (v/v) Protease Inhibitor Cocktail Set Ⅰ were added into the precipitates with the ratio of 4:1 (v/v), followed by ultrasonication for 180 s in ice-bath. Finally, the resulting mixture was centrifuged at 20000×g at 4°C for 30 min. The supernatant was collected and proteins were precipitated by the addition of cold acetone. After centrifugation, the pellets were lyophilized in a SpeedVac (Thermo Fisher Scientific, San Jose, CA, USA). The protein was re-dissolved in 50 mM NH_4HCO_3 (pH 7.8) prepared with $H_2{}^{16}O$ or $H_2{}^{18}O$ and the concentration was determined by BCA assay.

Approximately 2×10^6 Hca-F or Hca-P cells were inoculated subcutaneously and grown in the abdominal cavity of inbred Chinese 615 mice for 7 d. The cells collected

in ascites were then respectively washed three times with cold 1×PBS buffer, homogenized in 2 mL of lysis buffer (1% (v/v) protease inhibitor cocktail in 8 M urea) using a Tissue Tearor from Biospec Products (Bartlesville, OK, USA) at approximately 10000 rpm for 1 min, sonicated at 100 W for 100 s and centrifuged at 25000×g at 4°C for 40 min. The supernatant was collected and proteins were precipitated by the addition of cold acetone. After centrifugation, the pellets were lyophilized in a SpeedVac. The proteins from Hca-F were re-dissolved in 50 mM NH_4HCO_3 (pH 7.8) prepared with $H_2{}^{18}O$ and the proteins from Hca-P were re-dissolved in 50 mM NH_4HCO_3 (pH 7.8) prepared with $H_2{}^{16}O$. The protein concentration was determined by BCA assay.

The proteins extracted from Yeast, Hca-F and Hca-P were denatured by heating at 90°C for 10 min. After cooled to room temperature, the samples were reduced with 10 mM DTT at 56°C for 1.5 h, and alkylated by 25 mM IAA in the dark at room temperature for 40 min. Finally, two equal aliquot of 100 ng (0.05 mg/mL, 2μL) proteins from Yeast dissolved in $H_2{}^{16}O$ and $H_2{}^{18}O$ and two equal aliquot of 100 ng (0.05 mg/mL, 2μL) proteins from Hca-F and Hca-P were subject to online operation, and another two equal aliquot of 100 ng (1μg/mL, 100μL) proteins from Yeast dissolved in $H_2{}^{16}O$ and $H_2{}^{18}O$ were subject to offline operation.

2.3 The optimization of the condition of online digestion and labeling by IMER

The hydrophilic immobilized trypsin microreactor (250 μm i.d. × 10 cm) was prepared according to our previous procedure.[21] To optimize the labeling time, 100 ng BSA (0.05 mg/mL, 2μL) dissolved in 50 mM NH_4HCO_3 (pH 7.8) prepared with

$H_2{}^{18}O$ was loaded on the IMER at a flow rate of 1 μL/min, then the flow rate was stopped and the sample was incubated in the IMER for 0, 30, 60 and 90 min, respectively. The ^{18}O labeled peptides were collected and analyzed by MALDI-TOF MS. Furthermore, four aliquots of BSA was separately dissolved in different buffers prepared with $H_2{}^{18}O$, including 50 mM NH_4HCO_3 (pH 7.8), 50 mM NH_4Ac (pH 6.75), 100 mM NH_4HCO_3 (pH 8.5) and 100 mM NH_4Ac (pH 6.0), to optimize the labeling solvent.[16, 22-24] After digestion and labeling by IMER for 60 min, the peptides were also analyzed by MALDI-TOF MS. All of the above digestion and labeling steps were carried out at 37°C in the column oven.

The configuration of the integrated platform was shown in Figure 1, consisting of a syringe pump, two six-port/two-position switching valves, an IMER, a column oven and the one dimensional nanoRPLC-MS/MS analysis system. For the quantitative analysis of real samples, 100 ng yeast proteins dissolved in 50 mM NH_4HCO_3 (pH 7.8) prepared with $H_2{}^{16}O$ aforementioned were firstly loaded on IMER. After digested and labeled for 60 min, the peptides were flushed to a C18 trap column (150 μm i.d. × 3 cm) with 50 mM NH_4HCO_3 (pH 7.8) prepared with $H_2{}^{16}O$ for 10 min. Then the IMER was equilibrated with 50 mM NH_4HCO_3 (pH 7.8) prepared with $H_2{}^{18}O$ for 10 min. Another 100 ng yeast proteins dissolved in 50 mM NH_4HCO_3 (pH 7.8) prepared with $H_2{}^{18}O$ were digested and labeled on IMER. The ^{18}O labeled peptides were also flushed to the C18 trap column, but with 50 mM NH_4HCO_3 (pH 7.8) prepared with $H_2{}^{18}O$ (solid line mode in Figure 1). Finally, the ^{16}O and ^{18}O labeled peptides were eluted to the nanoRPLC column for separation (dashed line mode in Figure 1).

The quantitative proteome analysis of 100 ng Hca-F and 100 ng Hca-P proteins was performed as for yeast sample.

2.4 Offline digestion and labeling in solution

The offline digestion and labeling of BSA was performed using a previously described protocol with a slight modification.[22] First, the BSA was digested by trypsin with an enzyme to substrate ratio of 1:25 (m/m) in 50 mM NH_4HCO_3 (pH 7.5) at 37°C for 12 h. Then, two equal aliquots of trypsin digests from BSA were lyophilized in a vacuum centrifuge, and re-suspended in $H_2{}^{16}O$ and $H_2{}^{18}O$ -prepared reaction buffer containing NH_4HCO_3 (100 mM, pH 6.0) and trypsin (1:25, m/m), respectively. After incubation for 24 h at 37°C, the two aliquots were boiled for 10 min at 100°C and snap frozen at -80°C to deactivate residual trypsin. The offline digestion and labeling of yeast was performed as for BSA, except that the starting material of yeast before offline operation including digestion, desalting and labeling is 100 ng, for a fair comparison with online platform.

2.5 Mass spectrometry analysis.

MALDI-TOF MS was performed on Ultraflex III TOF/TOF (BrukerDaltonics, Bremen, Germany) with a SmartbeamNd-YAG laser at 355 nm, a repetition rate of 200 Hz, and an acceleration voltage of 20 kV. CHCA matrix solution (7 mg/mL) was prepared in ACN/H_2O/TFA (60:40:0.1, v/v/v). Equal amounts of the sample and CHCA were sequentially dropped onto the MALDI plate for MS analysis. Spectra were obtained in positive ionization mode using reflector detection.

The MS analysis of real sample was performed on an LTQ Orbitrap Velos mass spectrometer with a quaternary surveyor MS pump (Thermo Fisher, San Jose, CA, USA). The peptides mixture flushed onto the C18 trap column (150 μm i.d. × 3 cm) were then separated by a C18 capillary column (75 μm i.d. × 15 cm). The mobile phases were H_2O with 2% ACN and 0.1% formic acid (A) and ACN with 2% H_2O and 0.1% formic acid (B). The elution conditions contained 20 min of elution with 0% B for desalting. Then, the linear gradients was 0-10% buffer B for 10 min, 10-35% B for 60 min, 40-80% for 5 min then maintained at 80% B for 10 min, and the flow rate after splitting was about 300 nL/min. The LTQ-Orbitrap mass spectrometer instrument was operated in positive mode with a 2.0 kV applied spray voltage. The temperature of the ion transfer capillary was set at 200°C. One microscan was set for each MS and MS/MS scan. All MS and MS/MS spectra were acquired in the data dependent mode. A full scan MS acquired from m/z 300 to 1800 was followed by 15 data dependent MS/MS events. The dynamic exclusion function was set as follows: repeat count, 1; repeat duration, 40 s; exclusion duration, 40 s. The normalized collision energy for MS/MS scanning was 35%.

2.6 Data analysis

The acquired raw files were analyzed by MaxQuant (version 1.2.2.5). The Andromeda program, embedded in MaxQuant, was used to search the peak lists against Yeast database (downloaded from www.uniprot.org on Augest. 31 st, 2013, 7,786 entries) or IPI-mouse database (version 3.68, 56,729 entries). Common contaminants were added to this database. Cysteine carbamidomethylation was

searched as a fixed modification, whereas N-terminal acetylation and methionine oxidation were searched as variable modifications. Enzyme specificity was set as trypsin. Two miss cleavages were allowed and a minimum of six amino acids per identified peptide was required. For quantitative analysis, the ^{18}O was set as heavy label. A False Discovery Rate (FDR) of 0.01 for proteins and peptides was required. Proteins quantified in triplicate analyses and with at least a 2-fold changes were considered to be differentially expressed proteins.

3. Results and Discussion

3.1 Evaluation of the digestion and ^{18}O labeling efficiency by IMER

To improve the digestion and ^{18}O labeling efficiency, most of the reported off-line protocols employed different solvent in digestion and labeling steps.[22-24] In our platform, by confining the immobilized trypsin and proteins into the column, these two steps can be accomplished on the IMER in situ at the same time with the same solvent, beneficial to improving sensitivity and reproducibility.

To calculate ^{18}O labeling efficiency from MALDI data of BSA digests, several peptides within a wide range of m/z were chosen, and the efficiency of incorporation of one or two ^{18}O atoms in different incubation time were manually calculated according to Eq. (1) [23] and Eq. (2),

$$\mathrm{ratio}\left(\frac{^{16}O}{^{18}O}\right) = \frac{I_0}{I_4 - \frac{M_4}{M_0}I_0 + I_2\left(1-\frac{M_2}{M_0}\right) - \left(1-\frac{M_2}{M_0}\right)\frac{M_2}{M_0}I_0} \quad (1)$$

$$\mathrm{ratio}\left(\frac{^{18}O}{^{18}O+^{16}O}\right) = 1 - \frac{\mathrm{ratio}\left(\frac{^{16}O}{^{18}O}\right)}{\mathrm{ratio}\left(\frac{^{16}O}{^{18}O}\right)+1} \quad (2)$$

where I_0, I_2 and I_4 are the measured intensities of the first, third and fifth peak in the isotopic envelope, and M_0, M_2 and M_4 are the corresponding theoretical relative intensities of the isotopic envelope of the peptide. The theoretical intensities of M_i were calculated from its sequence, according to the MS-Isotope at ProteinProspector web site.[25]

To optimize the incubation time for on-column digestion and ^{18}O labeling of BSA, three replicated experiments were carried out to evaluate the sequence coverage, labeling efficiency and reproducibility under different incubation time. As shown in Table 1, when the incubation time was prolonged to 60 min, the sequence coverage could exceed 65% and the labeling efficiency could exceed 99%. With the incubation time further increased to 90 min, the sequence coverage and labeling efficiency kept stable, indicating that the digestion and labeling process had been stabilized after 60 min.

To further optimize the labeling solvent, four kinds of reported solvent which were compatible with subsequently nanoRPLC-MS/MS analysis were chosen to evaluate the labeling efficiency.[16, 22-24] As shown in Table2, when digested and labeled in 50 mM NH_4HCO_3 (pH 7.8), the highest labeling efficiency was obtained.

The typical MALDI-TOF MS spectra of three representative unlabeled peptides, labeled peptides obtained by traditional offline method and labeled peptides obtained by IMER under optimized condition (50 mM NH_4HCO_3, 60 min) are shown in Figure 2. For IMER based labeling, peptides with +4 Da dominated the spectra and other ^{18}O labeled variants (+0 or +2 Da) can hardly be observed, indicating high labeling

efficiency was achieved. Compared with traditional offline methods, the IMER based strategy for digestion and labeling not only can significantly improve the average labeling efficiency from 95% to 99% (calculated from 8 highest intense peptides in MALDI-TOF MS spectra), but also can shorten the required sample preparation time from more than 36 h to just 1 h. Furthermore, it can dramatically simplify the offline multi-step operation by just injecting the sample into the IMER and incubating for 60 min.

Such high labeling efficiency could be mainly attributed to the significantly accelerated digestion and labeling rate by tremendously increasing the possibility of contact between proteins/peptides and immobilized trypsin in the limited space of IMER, which was also accordance with the previous report, namely, the activity of immobilized enzyme was over 2000 BAEE-U with the temperature increased to 37°C, while the activity was only 14.3±0.8 BAEE-U for in solution trypsin (BAEE-U was defined as △A253 of 0.001 AU/min).[21]

3.2 Evaluation of the effectiveness of IMER in preventing back-exchange

When $H_2{}^{16}O$ is added to the ^{18}O labeled peptides, with the catalytic protease present, enzyme-catalyzed back exchange can occur. However, in our experiment, with the integrated platform shown in Figure 1, the trypsin immobilized particles were packed in a capillary, and further connected to the C18 trap column through a valve. During the analysis, the digested and labeled peptides could be easily separated from the immobilized trypsin by elution with corresponding buffer, so that the back exchange from ^{18}O to ^{16}O could be easily avoided.

To evaluate the effectiveness in preventing ^{18}O back-exchange of our strategy, we re-dissolved ^{18}O labeled peptides obtained by the IMER with $H_2{}^{16}O$, and kept them at room temperature for a week, followed by MALDI-TOF MS analysis. Figure S1 in Supplementary Information showed that before and after 1 week of storage, no obvious difference (99.40% vs. 99.37%) on the $^{18}O/ (^{16}O+^{18}O)$ abundance ratios of the 8 highest intense peptides was observed, confirming that almost no back-exchange occurred. Compared with other reported methods to prevent back exchange, such as heating and cooling the digestion to eliminate residual tryptic enzyme or using filter to remove the enzyme,[16, 17] by our integrated platform, the operation is not only easy but also automatic, beneficial to improve the accuracy and reproducibility of quantitative analysis.

3.3 Evalution of the integrated platform on quantitative analysis of 100 ng sample

Although ^{18}O labeling is more amenable for trace amount of sample compared with other chemical labeling strategies,[12] it still needs too many artificial operations on digestion, labeling and deactivating trypsin. The multi-step offline operations and unavoidable sample absorption on tube may lead sample loss,[26] making it still difficult for the identification of trace amount of sample. In our strategy, all these offline operations can be integrated in the platform, making the platform more accessible for the quantitative analysis of trace amount of sample.

To verify the feasibility of our strategy on the proteome quantitative analysis of trace amount of real sample, the integrated platform was further applied to analyze

100 ng of yeast sample labeled with $^{16}O/^{18}O$ at the ratio of 1:1. The LC-MS/MS spectrum was shown in Figure S2 to illustrate the performance of the platform. After database searching, 214, 226 and 237 proteins were successfully quantified in three technical replicates (including sample preparation and LC-MS/MS analysis), respectively, and 119 proteins were commonly quantified in all three technical replicates, all of which could be quantified with ratios ranging from 0.665 to 1.206, and the average RSD value[27, 28] between three technical replicates was 13.22% (Figure S3 in Supplementary Information). The results were further compared with those obtained by traditional offline in-solution method. As shown in Table 3, the results of total sample preparation time, commonly quantified protein number, the average relative errors of quantified proteins and the average RSD values between three technical replicates of online platform were all much better than that of offline method, suggesting the improved sensitivity, quantitative accuracy and reproducibility. The commonly quantified protein number of offline method is very low, but it is in the same magnitude compared with our previous result with similar offline operation,[26] which might be caused by the sample loss and poor reproducibility of the multi-step offline operation. For ^{16}O and ^{18}O online labeled yeast sample at mixing ratios of 1:5 and 1:10, the measured ratios were 5.12 and 9.81, respectively, with average relative errors (RE)[27, 28] of 9.9% and 11.8%, respectively (Figure S4 in Supplementary Information), which demonstrated the high quantitative accuracy in a wide dynamic range.

The obviously improved performance of our established platform for proteome

quantification could be attributed to two reasons. First, the less offline operations of our platform could reduce the possibility of sample loss and extra discrepancy. Second, the highly efficient and reproducible digestion could avoid the possibility of error introduced by nonrecurring of rapid digestion. As shown in Figure S5 in Supplementary Information, the percentage of miss-cleavage was only 10.25%, which is even better than that of in-solution digestion.[29] Furthermore, the ratio distribution of the peptides with miss-cleavage was also consistent with the peptides without miss-cleavage, indicating the miss-cleavage has little influence on the quantitative accuracy for reproducible miss-cleavage of the IMER.

3.4 Quantitative proteome profiling of Hca-F and Hca-P Cell Lines

Two mouse hepatocarcinoma ascites syngeneic cell lines, Hca-F with a highly lymphatic metastasis rate over 75%, and Hca-P with a low lymphatic metastasis rate less than 25%, were experimental subjects selected in many researches to gain insights into potential mechanisms that contribute to the Lymph node metastasis (LNM) associated tumor malignancy.[30, 31]

With high sensitivity, high reproducibility and good quantitative accuracy, the integrated platform was further applied to the quantitative proteome analysis of Hca-F and Hca-P cell lines. As shown in Figure 3, in three technical replicates, 329, 285, and 297 non-redundant proteins were successfully quantified from 736, 698, and 684 distinct peptides, respectively. The overlapping of quantified proteins in these three technical replicates also showed good reproducibility of the integrated platform.

If using the three-fold standard deviation in accuracy evaluation as the cutoff with

significance (99.7% confidence, 3σ), proteins with ratio <0.577 or >1.357 would be considered as significant changes.[32] Considering the variation of the MS-based quantification method, a more stringent cutoff (2-fold change), which is widely adopted by the other quantitative studies, is adopted as the selection of differential candidates in our experiment.[33, 34] In total, 199 unique proteins were reliably quantified in all three technical replicates, and 10 proteins of them displayed more than 2-fold expression differences, among which 5 proteins were up-regulated in Hca-F and 5 proteins were down-regulated in Hca-F (Supplementary Information Table1). Among these differently regulated proteins, 9 proteins were also consistently quantified in previous studies.[28] For example, Ahnak protein, which was reported as a tumor suppressor via modulation of TGFβ/Smad signaling pathway,[35] was found down-regulated in Hca-F in our study. The amount of transferrin receptor protein was reported to decrease according to the reduction of the proliferative ability of cancer cells,[36] and transferrin receptor protein 1 was found up-regulated in Hca-F in our result. Furthermore, Calreticulin was not found as a differently regulated protein in previous studies but has been reported to be related to the evolution and transfer of tumor,[37] which is deserved for further biological verification.

4. Conclusions

An automatic integrated online platform involving on-line digestion, ^{18}O labeling, peptide separation, identification and quantification was established, by which some inherent problems of ^{18}O labeling, such as limited labeling efficiency, back exchange and sample loss caused by tedious offline operation, could be significantly improved.

Through the applications in the quantitative analysis of nanograms of real samples, higher precision, accuracy, sensitivity and reproducibility could be achieved by our strategy compared with traditional offline methods. All these results demonstrated that such a platform might become a promising technique for the quantitative analysis of trace amount of sample.

Acknowledgement

The work was supported by the China State Key Basic Research Program Grant (2012CB910604), National Natural Science Foundation (21190043). The Creative Research Group Project by NSFC (21321064) and National key scientific instrument and equipment development projects (2012YQ120033-2). We also thank Prof. Shujuan Shao from Dalian Medical University for providing Hca-P and Hca-F samples.

References

1. R. Aebersold, M. Mann, *Nature*, 2003, **422**, 198-207.

2. S. E. Ong, M. Mann, *Nat. Chem. Biol*, 2005, **1**, 252-262.

3. Y. Y. Zhang, B. R. Fonslow, B. Shan, M. C. Baek, J. R. Yates 3rd, *Chem. Rev*, 2013, **113**, 2343-2394.

4. Q. Wu, H. Yuan, L. Zhang, Y. Zhang, *Analytica. Chimica. Acta*, 2012, **731**, 1-10.

5. X. Yao, A. Freas, J. Ramirez, P. A. Demirev, C. Fenselau, *Anal. Chem*, 2001, **73**, 2836-2842.

6. L. F. Dagley, C. A. White, Y. Liao, W. Shi, G. K. Smyth, J. M. Orian, A. Emili, A.

W. Purcell, *Proteomics*, 2014, **14**, 241-261.

7. M. Wang, Y. Shen, I. V. Turko, D. C. Nelson, Li. S, *Anal. Chem*, 2013, **85**, 11014-11019.

8. D. Lang, S. Anker, B. Kuropka, E. Krause, *Anal. Methods*, 2013, **5**, 1058-1061.

9. X. Ye, B. T. Luke, D. J. Johann Jr, A. Ono, D. A. Prieto, K. C. Chan, H. J. Issaq, T. D. Veenstra, J. Blonder, *Anal. Chem*, 2010, **82**, 5878-5886.

10. M. Miyagi, K. C. Rao, *Mass Spectrom. Rev*, 2007, **26**, 121-136.

11. W. J. Qian, M. E. Monroe, T. Liu, J. M. Jacobs, G. A. Anderson, Y. Shen, R. J. Moore, D. J. Aderson, R. Zhang, S. E. Calvano, S. F. Lowry, W. Xiao, L. L. Moldawer, R. W. Davis, R. G. Tompkins, D. G. Camp II, R. D. Smith, *Mol. Cell. Proteomics*, 2005, **4**, 700-709.

12. C. Fenselau, X. D. Yao, *J. Proteome Res*, 2009, **8**, 2140-2143.

13. A. Staes, H. Demol, J. V. Damme, L. Martens, J. Vandekerckhove, K. Gevaert, *J. Proteome Res*, 2004, **3**, 786-791.

14. H. F. Storms, R. V. D. Heijden, U. R. Tjaden, J. V. D. Greef, *Rapid Commun. Mass Spectrom*, 2006, **20**, 3491-3497.

15. T. Z. Kristiansen, H. C. Harsha, M. Gronborg, A. Maitra, A. Pandey, *J. Proteome Res*, 2008, **7**, 4670-4677.

16. B. O. Petritis, W. J. Qian, D. G. Camp II, R. D. Smith, *J. Proteome Res*, 2009, **8**, 2157-2163.

17. Y. Zhao, W. Jia, W. Sun, W. H. Jin, L. H. Guo, J. Y. Wei, W. T. Ying, Y. J. Zhang, Y. M. Xie, Y. Jiang, F. C. He, X. H. Qian, *J. Proteome Res*, 2010, **9**, 3319-3327.

18. E. Bonzon-Kulichenko, D. Pérez-Hernández, E. Núñez, P. Martínez-Acedo, P. Navarro, M. Trevisan-Herraz, M. D. C. Ramos, S. Sierra, S. Martínez-Martínez, M. Ruiz-Meana, E. Miró-Casas, D. García-Dorado, M. Ruiz-Meana, E. Miro-Casas, D. Garcia-Dorado, J. M. Redondo, J. S. Burgos, J. Vázquez, *Mol. Cell. Proteomics*, 2011, **10**, M110 003335.

19. W. J. Qin, Z. F. Song, C. Fan, W. J. Zhang, Y. Cai, Y. J. Zhang, X. H. Qian, *Anal. Chem*, 2012, **84**, 3138-3144.

20. J. R. Sevinsky, K. J. Brown, B. J. Cargile, J. L. Bundy, J. L. Stephenson Jr, *Anal. Chem*, 2007, **79**, 2158-2162.

21. H. Jiang, H. M. Yuan, Y. Liang, S. Xia, Q. Zhao, Q. Wu, L. H. Zhang, Z. Liang, Y. Zhang, *J. Chromatogr. A*, 2012, **1246**, 111-116.

22. S. J. Yang, A. Y. Nie, L. Zhang, G. Q. Yan, J. Yao, L. Q. Xie, H. J. Lu, P. Y. Yang, *Journal of Proteomics*, 2012, **75**, 5797-5805.

23. D. Lopez-Ferrer, A. Ramos-Fernandez, S. Martinez-Bartolome, P. Garcia-Ruiz, J. Vázquez, *Proteomics*, 2006, **6**, 4-6.

24. C. A. White, N. Oey, A. Emili, *J. Proteome Res*, 2009, **8**, 3653-3665.

25. MS-Isotope, http://prospector.ucsf.edu.

26. Q. Zhao, Y. Liang, H. M. Yuan, Z. G. Sui, Q. Wu, Z. Liang, L. H. Zhang, Y. K. Zhang, *Anal. Chem*, 2013, **85**, 8507-8512.

27. S. Zhang, Q. Wu, Y. C. Shan, Y. Zhou, L. H. Zhang, Y. K. Zhang, *J. Proteomics*, 2015, **114**, 152-160.

28. Y. Zhou, Y. C. Shan, Q. Wu, S. Zhang, L. H. Zhang, Y. K. Zhang, *Anal. Chem*,

2013, **85**, 10658-10663.

29. A. S. Hebert, A. L. Richards, D. J. Bailey, A. Ulbrich, E. E. Coughlin, M. S. Westphall, J. J. Coon, *Mol. Cell. Proteomics*, 2014, **13**, 339-347.

30. X. N. Cui, J. W. Tang, L. Hou, B. Song, L. Ban, *World J. Gastroenterol*, 2006, **12**, 6893-6897.

31. H. F. Li, M. Y. Ling, Y. Xie, H. Xie, *Oncol. Res*, 1998, **10**, 569-573.

32. Y. J. Weng, Y. Y. Qu, H. Jiang, Q. Wu, L. H. Zhang, H. M. Yuan, Y. Zhou, X. D. Zhang, Y. K. Zhang, *Anal. Chim. Acta*, 2014, **833**, 1-8.

33. C. X. Song, F. J. Wang, M. L. Ye, K. Chen, R, Chen, J. Zhu, Y. X. Tan, H. Y. Wang, D. Figeys, H. F. Zou, *Anal. Chem*. 2011, **83**, 7755–7762.

34. F. J. Wang, R. Chen, J. Zhu, D. G. Sun, C. X. Song, Y. F. Wu, M. L. Ye, L. M. Wang, H. F. Zou, *Anal. Chem*, 2010, **82**, 3007-3015.

35. I. H. Lee, M. Sohn, H. J. Lim, S. Yoon, H. Oh, S. Shin. J. H. Shin, S-H. Oh, J. Kim, D. K. Lee, D. Y. Noh, D. S. Bae, J. K. Seong, Y. S. Bae, *Oncogene*, 2014, **33**, 4675-4684.

36. M. Takao, T. Nobuyuki, E. Yoshinobu, A. Teruo, *International Journal of Oral and Maxillofacial Surgery*, 1994, **23**, 430-433.

37. M. Ni, A. S. Lee, *FEBS Lett*, 2007, **581**, 3641-3651.

Table 1. Sequence coverage and ^{18}O labeling efficiency of BSA digests obtained with different incubation time by IMER (n=3, n refers to the number of passages of the same sample through IMER).

Incubation time (min)	0	30	60	90
Sequence coverage (%)	54.33±1.15	59.67±0.58	65.33±2.52	66.00±1.73
Labeling efficiency (%)	88.07±0.61	95.97±0.38	99.37±0.21	99.40±0.20

Table 2. ^{18}O labeling efficiency (%) of two peptides in BSA digests in different labeling solvents (n=3, n refers to the number of passages of the same sample through IMER).

Peptide sequence	50 mM NH_4Ac (pH 6.75)	100 mM NH_4Ac (pH 6.0)	50 mM NH_4HCO_3 (pH 7.8)	100 mM NH_4HCO_3 (pH 8.5)
K.LGEYGFQN ALIVR.Y	91.86 ±0.72	97.31 ±0.05	99.69 ±0.03	99.75 ±0.15
R.RPCFSALTP DETYVPK.A	89.38 ±1.07	89.96 ±0.61	97.03 ±0.62	91.58 ±1.30

Table 3. Quantification results obtained by online platform and traditional offline in-solution method for the yeast sample.

	Total time	Commonly quantified protein number	Mean ratio	RE (%)	RSD (%)
Online	5 h	119	0.967	7.6	13.22
Offline	45 h	15	1.049	31.4	31.55

Figure 1. Schematic diagram of integrated platform for proteome quantification with on-line digestion and ^{18}O labeling by IMER.

Figure 2. Typical MALDI-TOF MS spectra of the unlabeled peptides (A), labeled peptides obtained by traditional offline method (B) and labeled peptides obtained by IMER (C). The amino acid sequences of peptides A, B and C depicted in all figure are LGEYGFQNALIVR, KVPQVSTPTLVEVSR and RPCFSALTPDETYVPK, respectively.

Figure 3. Quantified proteins in Hca-F and Hca-P cell lines in three technical replicates (A), and protein expression ratio (Hca-P/Hca-F) of proteins quantified in all three technical replicates with RSD < 50% (B).

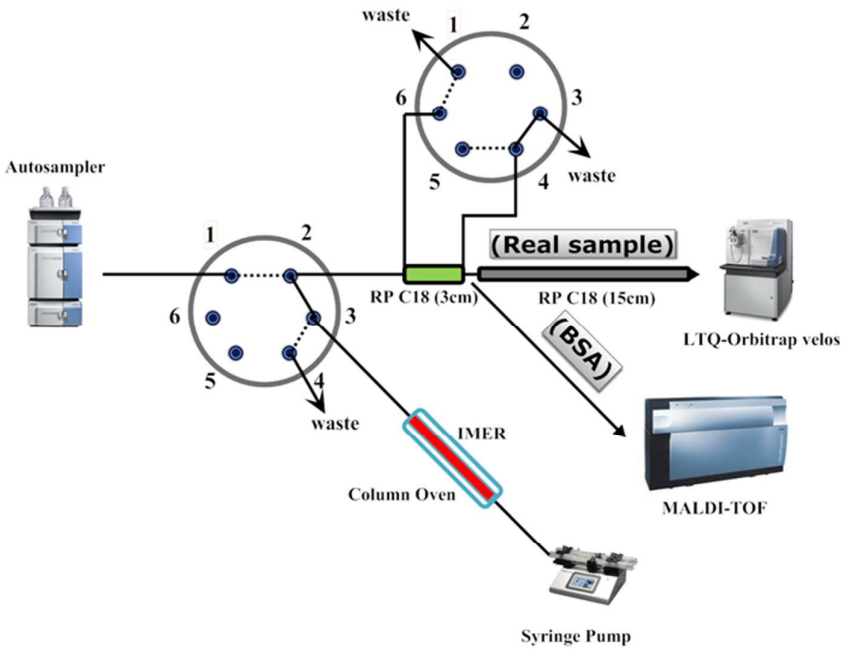

Figure 1. Schematic diagram of integrated platform for proteome quantification with on-line digestion and ^{18}O labeling by IMER.

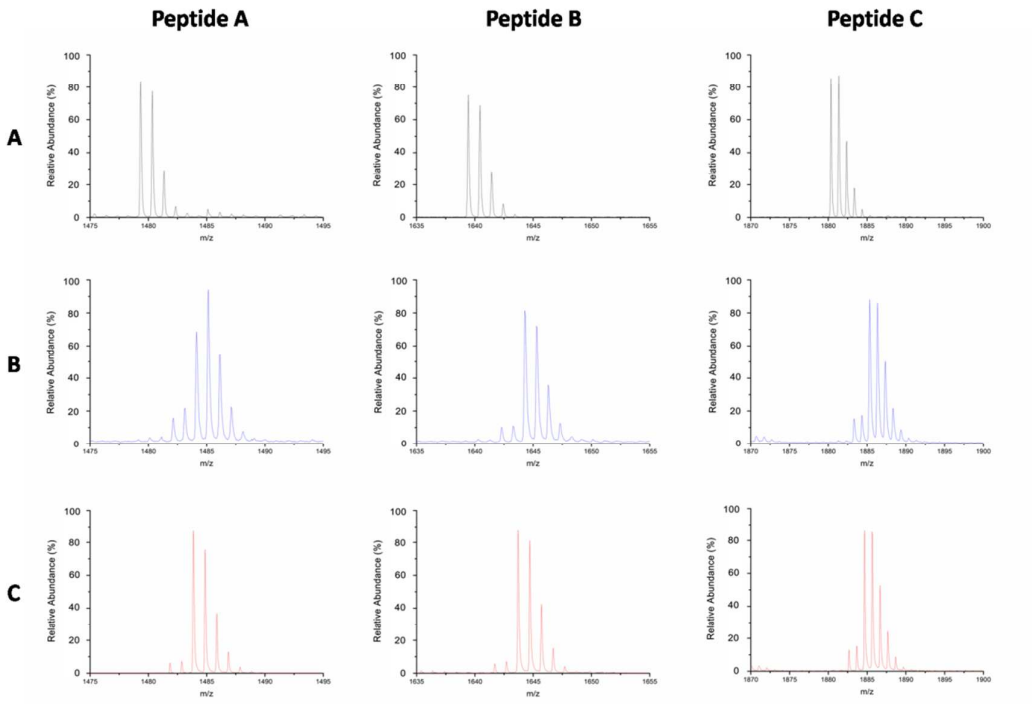

Figure 2. Typical MALDI-TOF MS spectra of the unlabeled peptides (A), labeled peptides obtained by traditional offline method (B) and labeled peptides obtained by IMER (C). The amino acid sequences of peptides A, B and C depicted in all figure are LGEYGFQNALIVR, KVPQVSTPTLVEVSR and RPCFSALTPDETYVPK, respectively.

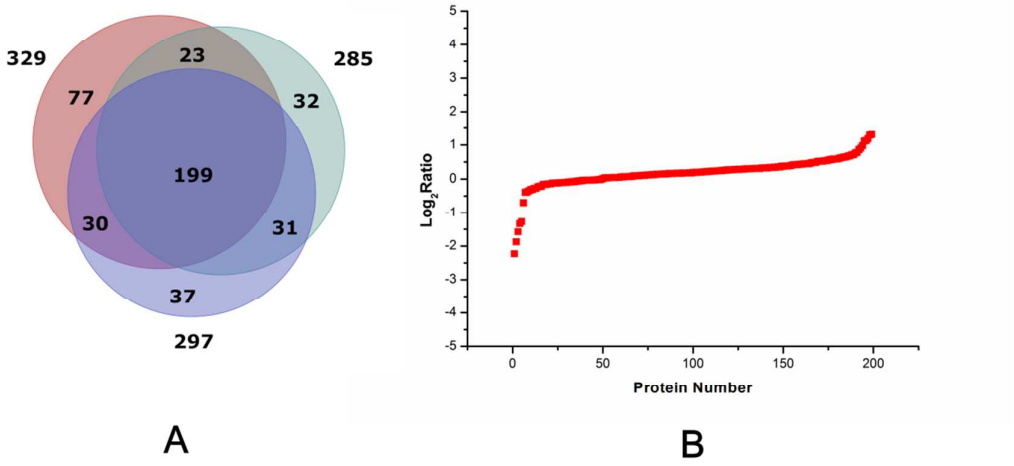

Figure 3. Quantified proteins in Hca-F and Hca-P cell lines in three technical replicates (A), and protein expression ratio (Hca-F/Hca-P) of proteins quantified in all three technical replicates with RSD < 50% (B).