Analyst

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/analyst

Journal Name

ARTICLE

A simple dual online ultra-high pressure liquid chromatography system (sDO-UHPLC) for high throughput proteome analysis.

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Hangyeore Lee,[a] Dong-Gi Mun,[a] Jingi Bae,[a] Hokeun Kim,[a] Se Yeon Oh,[a] Young Soo Park,[b] Jae-Hyuk Lee[c] and Sang-Won Lee[a]*

We report a new and simple design of a fully automated dual-online ultra-high pressure liquid chromatography system. The system employs only two nano-volume switching valves (a two-position four ports valve and a two-position ten ports valve) that direct solvent flows from two binary nano-pumps for parallel operation of two analytical columns and two solid phase extraction (SPE) columns. Despite the simple design, the sDO-UHPLC offers many advantageous features that include high duty cycle, back flushing sample injection for fast and narrow zone sample injection, online desalting, high separation resolution and high intra/inter-column reproducibility. This system was applied to analyze proteome samples not only in high throughput deep proteome profiling experiments but also in high throughput MRM experiments.

Introduction

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) is a powerful tool for in-depth proteome profiling and targeted MS experiments in many biological and clinical applications[1-3]. One of the ultimate goals of clinical proteomics is to find the biological marker molecules that can be used for detection of a specific disease or evaluation of its state. Initial phase of a biomarker research focuses on performing discovery experiments which are unbiased and semi-quantitative proteome profiling experiments[4, 5]. The efforts in this discovery phase would result in a list of candidate marker proteins that are differentially expressed between normal and disease states. These potential protein biomarkers then would be verified and validated on a much larger scale in terms of sample size[6, 7], thereby demanding high throughput analyses.

An important experimental requirement of MS-based biomarker research is high sensitivity of LC-MS/MS platforms for detecting low-abundant potential protein biomarkers present in biological samples (e.g. plasma or tissue)[8]. Advanced LC systems of high separation resolution such as ultra-high pressure liquid chromatography (UHPLC) can improve the detection sensitivity vastly not only for profiling experiments but also for verification/validation experiments. Köcher et al. demonstrated that the peak capacity is increased with increasing gradient time and analytical column length, and the increased peak capacity leads to improved detection of low abundant proteins[9]. A targeted MS approach such as multiple reaction monitoring (MRM) also requires high resolution separation for the sensitive detection of target peptides in the presence of complex sample matrix[10-14]. Shi et al. reported that the utilization of long gradient (i.e. 300 min) on a long packed capillary column (75 µm × 150 cm) can increase the sensitivity of LC-MRM experiments dramatically, achieving 3 times more multiplexing capacity than conventional LC-MRM and LOQ of 10 ng/mL without depleting abundant proteins and fractionating sample[15].

Another demanding factor of MS-based biomarker research is the experimental throughput of LC-MS/MS platforms. Despite dramatic improvement in sensitivity, the use of a long and narrow column with a long gradient in UHPLC system requires a long lead time for sufficient column equilibration before the next experiment. There have been several examples of utilizing multi-column LC systems employing valve-switching techniques to achieve a high duty cycle (i.e. the ratio of useful MS-acquisition time relative to the total analysis time) for LC-MS/MS experiments[16-20]. Livesay et al. reported a high-throughput LC system that utilized four columns[17]. This system achieved a near 100% duty cycle by doing a separation experiment on one column while performing sample loading and regeneration on the other columns. Wang et al. reported another HPLC system utilizing two columns connected in parallel[18]. This system employed two binary pumps to operate two pairs of analytical and trap columns for high throughput analysis by adding three switching valves. Recently, Orton et al. reported a dual-

[a] Department of Chemistry, Research Institute for Natural Sciences, Korea University, Seoul 136-701, South Korea.
E-mail: sw_lee@korea.ac.kr; Tel: +82-2-3290-3137; Fax: +82-2-3290-3121
[b] Department of Pathology, University of Ulsan College of Medicine, Asan Medical Center, Seoul 138-873, South Korea.
[c] Department of Pathology, Chonnam National University Medical School, Gwangju 501-746, South Korea.
† Electronic Supplementary Information (ESI) available. See
DOI: 10.1039/x0xx00000x

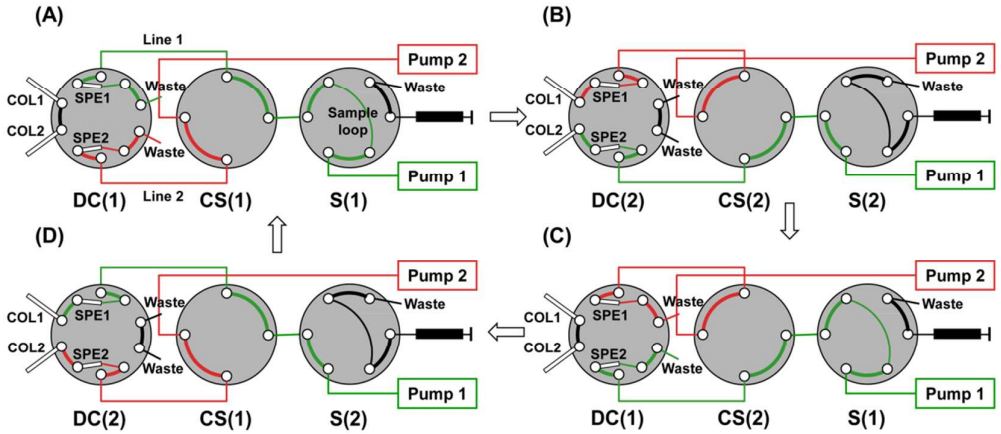

Fig 1. Schematic representation of the sDO-UHPLC system for the one cycle of two LC-MS/MS experiments. (A) Sample injection on SPE1 while SPE2 being equilibrated. (B) RPLC on SPE1/COL1 while SPE2/COL2 being equilibrated. (C) Sample injection on SPE2 while SPE1 being equilibrated. (D) RPLC on SPE2/COL2 while SPE1/COL1 being equilibrated. DC, CS, and S are dual column valve, column selection valve, and sample injection valve, respectively. The numbers in parentheses indicate the position of the three valves.

column interface for parallel chromatography[19]. Utilizing only one additional switching valve, the system successfully provided a near double LC-MS/MS uty cycle. The system, however, lacked SPE columns for online desalting.

In this study, we developed a simple dual online ultra-high pressure liquid chromatography (sDO-UHPLC) system. The system consists of two nano binary pumps, an autosampler and two additional switching valves. The automated time-control of the positions of the two valves allowed both parallel and alternative operation, in which two sets of analytical and SPE columns are being utilized for consecutive LC-MS/MS experiments with very small dead time in between. Application of the developed system to both targeted MS and deep proteome profiling experiments demonstrated increased experimental throughput.

Experimental

Chemicals and materials

LC/MS grade acetonitrile (A955-212, ACN) and water (W6-4) were purchased from Fisher Scientific (Fair Lawn, NJ). Enolase (E6126), formic acid (14265), and ammonium bicarbonate (A6141) were purchased from Sigma-Aldrich (St. Luis, MO). Sequence-grade modified porcine trypsin (90057) was purchased from Thermo Scientific (Rockford, IL). Fresh frozen tissue samples were collected from 2 patients with intestinal-type gastric cancers (M/44 and F/43) who underwent curative surgery at Asan Medical Center and Chonnam University Hwasun Hospital in Korea. Adjacent normal tissue was also collected from the same patients at least 2 cm from tumor border. The patient signed IRB-approved informed consents.

Peptide sample preparation

Tryptic enolase peptides and gastric tissue peptides were prepared as described in supplementary information†.

The gastric peptides were labeled using 4-plex iTRAQ reagent (4352135, AB Sciex, Foster City, CA) according to manufacturer instructions. The two adjacent normal peptide samples (100 µg each) were labeled with 114 and 116 iTRAQ reagents, respectively and the two cancer peptide samples (100 µg each) were labeled with 115 and 117 iTRAQ reagents, respectively. The four iTRAQ-labeled peptide samples were pooled and concentrated to 200 µL and the concentrate was immediately subjected to the basic pH reverse-phase fractionation where the initial 96 fractions were noncontiguously pooled into 24 fractions (supplementary information†)[21].

Simple dual online ultra-high pressure liquid chromatography (sDO-UHPLC) system

A schematic diagram of the sDO-UHPLC system is shown in Fig 1. The system consists of a sample injection valve (S, 6-ports/3-channels/2-positions) of a autosampler unit and two nanoflow binary pumps (Pump 1 and Pump 2, nanoACQUITY, Binary Solvent Manager, Waters, Milford, MA) and two switching valves: a column selection valve (CS, 4-ports/2-channels/2-positions, C72MX-4674, VICI, Houston, TX) and a dual column valve (DC, 10-ports/5-channels/2-positions, C72MX-4670, VICI, Houston, TX). The two additional valves were installed on universal actuators (EUHC, VICI, Houston, TX). Pump 1 is used to perform sample injection into the SPE column, online desalting and SPE/COL equilibration while Pump 2 is used to generate a gradient for LC separation. The numbers in parentheses in the figure indicate the position of the three valves

On the DC-valve, two capillary analytical columns (COL1 and COL2, 75 µm × 100 cm) and two SPE columns (SPE1 and SPE2, 150 µm × 3 cm) were installed. Both analytical and SPE columns were packed in house with C18 bonded particles (Jupiter, 300 Å, 3 µm, Phenomenex, Torrence, CA) as previously described[22]. By switching the DC valve, the COL/SPE pairs are in either "RPLC" position (i.e. COLs and SPEs being connected in series as in Fig. 1B and 1D) or

"INJECT" position (i.e. SPEs are being connected to waste lines as in Fig. 1A and 1C). The CS-valve directs the solvent flow from either of the two pumps to either SPE1/COL1 or SPE2/COL2 (i.e. from Pump 1 to SPE1/COL1 and from Pump 2 to SPE2/COL2 in CS(1) position as in Fig. 1D). The positions of the valves are automatically time-controlled by asserting (e.g. grounding) the digital I/O pins of the universal actuators using an electronic switch of the nanoACQUITY system running the MassLynx data system (Waters, Milford, MA). A signal relay (DS1E-M-DC5V, Panasonic Electric Works Co., Osaka, Japan) was used between the electronic switch and the digital I/O pins of the universal actuator to control the valve switching via an electronic switch of the LC system.

High throughput LC-MRM experiments

Scheduled MRM experiments utilizing the sDO-UHPLC system were performed on a triple quadrupole mass spectrometer (6490, Agilent, Santa Clara, CA). A specially designed two-position selection valve was used to make connections between the two columns of the sDO-UHPLC system and an electrospray emitter as described previously[20]. For the proof-of-concept demonstration of high-throughput MRM experiments, a 180 min gradient (1%-40% solvent B over 160 min, 40%-80% over 5 min, 80% for 10 min and 1% for 5 min, 300 nL/min) was used for the scheduled MRM experiments on the simple tryptic digests of enolase 1 protein from yeast (Saccharomyces cerevisiae) devoid of matrix. To evaluate the detection sensitivity and the retention time reproducibility, eight abundant peptides eluted at various gradient times were selected. Solvent A was 0.1% formic acid in water and solvent B was 0.1% formic acid in ACN. The temperature of the two analytical columns was kept at 60 ℃[23]. In order to compare the results from matrix free condition with ones in the presence of complex matrix, 10 pg, 100 pg, 1 ng, and 10 ng of tryptic enolase peptides were spiked into 1 μg of unlabeled peptides from normal gastric tissue, respectively, and analyzed under the same MRM conditions.

The capillary voltage of the triple quadrupole mass spectrometer was maintained at 3 kV. The drying gas flow rate and temperature were set to 6 L/min and 300 ℃, respectively. Q1 and Q3 were set to unit resolution. Skyline software (v1.1)[24] was used for generation of transition list and calculation of optimal collision energies for the each peptide. Three most abundant transitions per peptide were monitored with the time window of 4 min. The list of transitions and optimized collision energies are summarized in table S1†. The cycle time was set to 500 ms and the minimum and maximum dwell time were 81.71 ms and 165.64 ms, respectively.

Global proteome profiling experiments and data analyses

10 μg of labeled peptides were analyzed by the sDO-UHPLC system coupled to a bench top orbitrap mass spectrometer (Q-Exactive, Thermo Electron, San Jose, CA) as described in supplementary information†.

The precursor masses of the MS/MS scans were mass corrected and refined using PE-MMR (available at http://omics.pnl.gov/software/PEMMR.php)[25]. The refined MS/MS data (e.g. mgf files) were then searched against a composite database of uniprot human reference database (released May 2013; 90,219 entries) and contaminants (179 entries) using MS-GF+ (v9387, http://omics.pnl.gov/software/ms-gf)[26]. The database searches were performed with the following parameters: the mass tolerance for precursor ion = 10 ppm and the number of tryptic termini (NTT) = 1. Static modifications were iTRAQ to lysine and N-terminus and carbamidomethylation to cysteine. A variable modification for methionine oxidation was used. PSMs at the false discovery rate (FDR) of 0.01 were used for further analysis[27]. The resultant PSMs were used to infer proteins as described in supplementary information†.

Results and discussion

Previously, we developed a dual online SPE/capillary RPLC (DO-SPE/cRPLC) system that was equipped with six switching valves[16]. Using the system, one can perform high-resolution and highly reproducible peptide separation experiments in a high-throughput manner utilizing the two pairs of SPE and analytical columns. The throughput was demonstrated to increase two-folds by essentially removing the column equilibration time between successive LC-MS/MS experiments. In addition to the twice LC-MS/MS throughputs, the system offered fast "back-flushing" sample injection, online desalting, and high retention time reproducibility among experiments utilizing the two pairs of analytical columns and SPE trap columns. Despite the advantageous features, the system was equipped with six manual switching valves, requiring manual operation of the very complex LC procedures. The sDO-UHPLC simplifies and automates the complex LC operations while retaining all the advantageous features of the DO-SPE/cRPLC system.

Fully automated operation of the sDO-UHPLC

Fig. 1 demonstrates one cycle of two consecutive LC-MS/MS experiments utilizing the two analytical columns without the time loss in between for column equilibration (from Fig. 1A to Fig. 1D). The positions of the dual column valve (DC-valve) and column selection valve (CS-valve) were automatically time controlled by the LC control software as described in the Methods section. In this simple valve module, the DC-valve, which is 10-ports/2-positions switching valve, is equipped with two pairs of SPE and analytical columns in such a way that a pair of an SPE trap and an analytical column (i.e. SPE1 and COL1) is inter-connected by fluidic paths while they do not make fluid connection to the other SPE/COL pair (i.e. SPE2 and COL2). This allows independent and parallel utilization of the two SPE/COL pairs by using a single switching valve (i.e. DC-valve). At the same time, the SPE and analytical column of each pair can make connection or disconnection by simply switching the DC valve (i.e. switching between DC(1) and DC(2)).

In Fig. 1A, the sample in the sample loop of the S-valve (in S(1) position) was injected to SPE1 with the solvent A from the

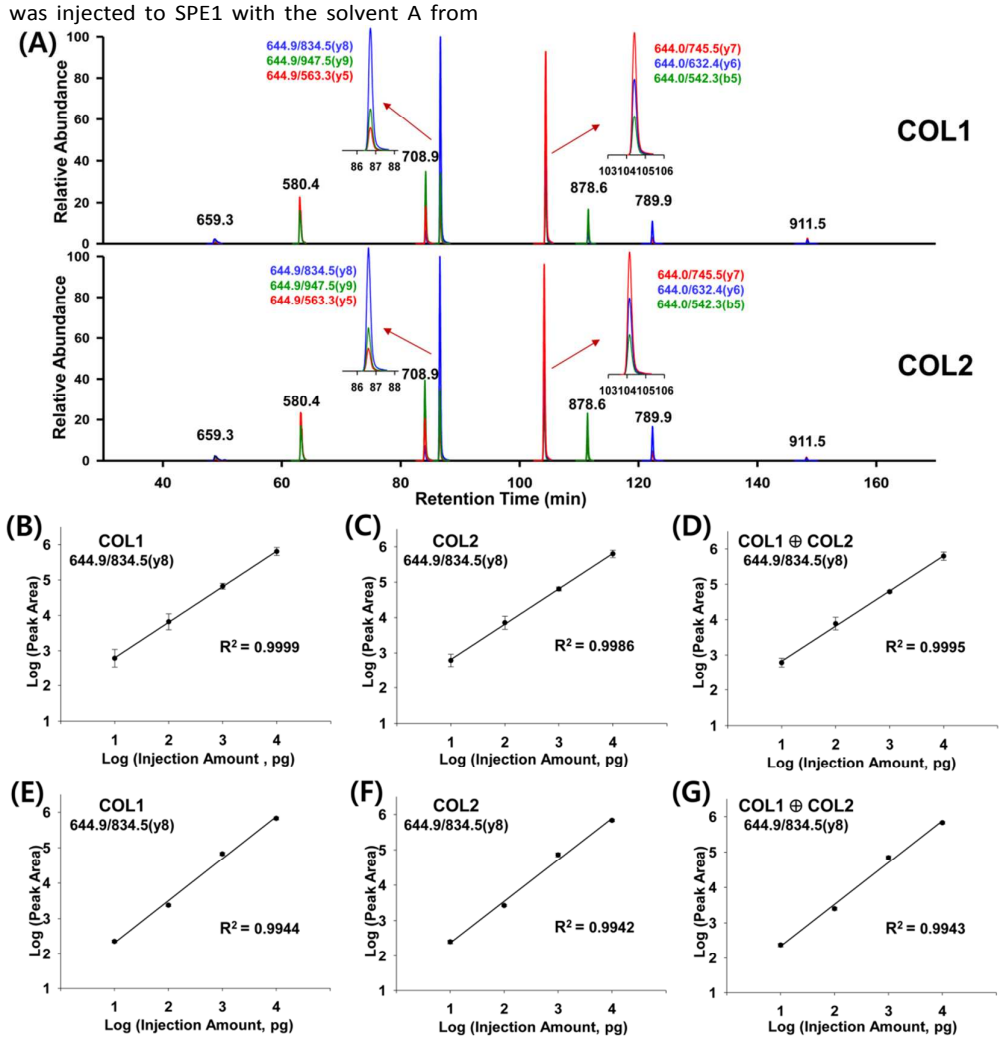

Fig. 2. (A) Extracted ion chromatograms of 24 transitions corresponding to eight enolase peptides in matrix free condition from LC-MRM experiment utilizing COL1 and COL2, alternatively. Insets show the details of two of eight target peptides: VNQIGTLSESIK (644.9) and NVNDVIAPAFVK (644.0). (B-D) Serial dilution curves using a transition (644.9/834.5, y8) of VNQIGTLSESIK from six consecutive experiments in matrix free condition: B using the transition data of COL1 only (3 LC-MRM experiments), C using the transition data of COL2 only (3 experiments), and D using the transition data of both COL1 and COL2 (6 experiments). (E-G) Serial dilution curves using a transition (644.9/834.5, y8) of VNQIGTLSESIK from six consecutive experiments using 1 µg of gastric tissue peptides as sample matrix: E using the transition data of COL1 only (3 LC-MRM experiments), F using the transition data of COL2 only (3 experiments), and G using the transition data of both COL1 and COL2 (6 experiments). The numbers on the curves are the coefficient of variation (CV).

Pump 1 while the CS-valve and the DC-valve were in CS(1) and DC(1) position, respectively. In this position, the sample injection on the SPE1 was made to be decoupled from COL1 so that the sample can be injected at a higher flow rate (i.e. 2 µL/min). After sample injection, an extra 2 µL of solvent A was flowed into the SPE1 and flowed out to the waste port for online desalting. During sample injection and desalting on the SPE1 column, solvent A from Pump 2 was flowed into SPE2 at 300 nL/min. For RPLC experiments on the injected sample on the SPE1 column, the DC-valve and the CS-valves were switched to the DC(2) and CS(2) positions, respectively, so that the RP solvent gradient from the Pump 2 flowed to the CS-valve, and to the SPE1 and COL1, respectively (Fig. 1B). The solvent gradient the SPE1 in the opposite direction of the loading before it reaches COL1. This back-flushing mode of injection provides fast sample loading and online clean-up while it also resulted in a narrow sample injection zone for improved separation resolution. During the RPLC experiments on SPE1 and COL1, the solvent A from Pump 1 flowed at a rate of 300 nL/min to the SPE2 and COL2 for equilibration. This allowed the same time for equilibrating both the SPE and analytical columns as the RPLC gradient time, which is ideal for reproducible separation.

Upon completion of the RPLC experiments on the SPE1 and COL1, the DC-valve was switched to the DC(1) position while

CS-valve remained in the CS(2) position (Fig. 1C) so that the next sample was injected to SPE2 by Pump 1. After sample injection and online desalting, the DC-valve and the CS-valves were switched to the DC(2) and CS(1) positions, respectively, for the next RPLC experiment on SPE2 and COL2 by Pump 2 (Fig. 1D), completing a cycle of two LC-MS/MS experiments. Note that both RPLC experiments on COL1 and COL2 shared the same binary nanoflow pump for gradient generation and solvent line, and had essentially identical gradient flow paths (i.e. by carefully making line 1 and line 2 in Fig. 1A and both SPEs the same lengths), which is critical in achieving high reproducibility in the gradient elution of peptides from the two columns. There was a small dead time between two consecutive LC-MS/MS experiments when sample injection and online desalting occurred. However, due to the fast injection flow rate (2 µL/min), the dead time (~3 min) for sample injection and online desalting is negligible, compared to the long gradient time (180 min) used for each LC-MS/MS experiment.

High-throughput and precise LC-MRM measurements using the sDO-UHPLC system

The efficiency of the sDO-UHPLC system for the high-throughput targeted MS analysis was first evaluated by LC-MRM experiments using "matrix free" tryptic enolase peptides with varying injection amounts: 10 pg, 100 pg, 1 ng, and 10 ng. A total of eight peptides were targeted and, for each peptide, three transitions were selected, corresponding to a total of 24 transitions (Fig. 2A). In order to check the reproducibility of LC-MRM experiments on the sDO-UHPLC system, six consecutive LC-MRM experiments were performed at each injection amounts by alternatively using the two analytical columns. The average CVs of intra- and inter- column reproducibility in retention time with respect to the 180min gradient was the same at 0.33 %. The high analytical precision of peptide separation allows accurate timing of transition in scheduled MRM experiments and potentially leads to more peptides monitored in an LC-MRM experiment.

Fig. 2B and C show the serial dilution curves obtained from three experiments using COL1 and COL2, respectively, and Fig. 2D shows the combined result from these six runs using both COL1 and COL2. The serial dilution curves of the LC-MRM experiments (measured for transition 644.9/834.5, y8) shows linear response with coefficient (R^2) of 0.9999 and 0.9986 for COL1 and COL2, respectively. The average CV of transition measurement using COL1 and COL2 were 3.77 % and 2.38 %, respectively. The serial dilution curve of the combined data also shows excellent linear correlation with R^2 of 0.9995 and average CV of 3.21 %, despite using two separate columns. The linearity and precision for other transitions are summarized in Fig. S1†.

In order to examine the performance metrics of LC-MRM experiments using sDO-UHPLC system in the presence of sample matrix, the same amounts of tryptic enolase peptides (i.e. 10 pg, 100 pg, 1 ng, and 10 ng) were spiked in 1 µg of unlabelled global peptides from the normal gastric tissue, respectively, and each samples were subjected to triplicate LC-MRM experiments under the same condition targeting the same eight peptides (Figure 2E-G). Even under the complex sample matrix, the retention time reproducibility (average CV of 0.67%), the reproducibility of transition measurements

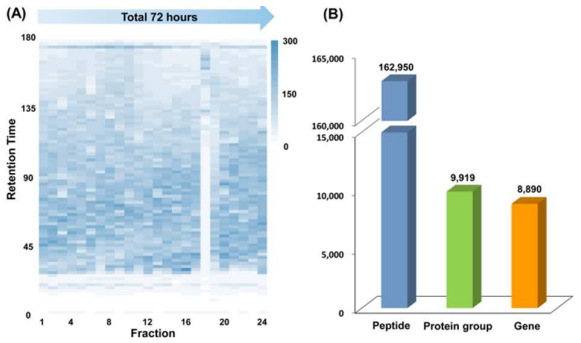

Fig. 3. (A) A heat map of the number of identified peptides in the 24 LC-MS/MS experiments. The x- and y-axes show the fraction number and a retention time. (B) The 24 LC-MS/MS dataset led to identification of 162,950 peptides (blue), 9,919 protein groups (green) and the identified protein groups were assembled to 8,890 protein coding genes (orange).

(average CVs of 0.81 %, 0.96 % and 1.09 % for COL1, COL2 and COL1⊕COL2, respectively), the measurement linear correlation (R^2 of 0.9944, 0.9942 and 0.9943 for COL1, COL2 and COL1⊕COL2, respectively) remained similar while the limit of quantitation for some transitions were higher, resulting in no measurements for them at 10 pg (Fig. S2†).

These results demonstrate that the sDO-UHPLC system significantly increased the duty cycle for the MRM experiment comparing to conventional single column LC system while it maintained the analytical precision and linear correlation of a single column LC system, even in the context of long gradient employing long analytical columns. To our knowledge, this is the first demonstration of utilizing a dual-column LC system for high-throughput LC-MRM experiment.

High-throughput global profiling of complex proteome using the sDO-UHPLC system

The increased experimental throughput of the sDO-UHPLC is especially valuable for deep and extensive proteome profiling experiments employing an extensive offline sample fractionation, in which a proteome sample is divided into many fractions to overcome undersampling of LC-MS/MS analysis platform[21]. In order to demonstrate the high-throughput and extensive proteome profiling, the sDO-UHPLC system was utilized to perform global profiling of a complex proteome sample from a tissue pair of a patient. As illustrated in Fig. 3A, twenty-four LC-MS/MS experiments on the twenty-four iTRAQ-labeled peptide fractions of a gastric tissue pair were completed in slightly over 72 hours, approaching the duty cycle of 100% (Fig. 3A, Fig. S3†) and resulted in a total of 162,950 non-redundant peptides of 9,919 protein groups, which were mapped to 8,890 genes (Fig. 3B). Assuming the use of the same time for the column equilibration as the time of RP gradient utilizing a single

column LC system, a similar global profiling experiment would require over 144 hours. This demonstrates the benefit of the sDO-UHPLC system for deep proteome profiling in high-throughput manner, facilitating clinical proteomics where a large number of clinical samples are analyzed[28].

Conclusions

We developed a simple fully automated dual-online ultrahigh pressure liquid chromatography system for throughput proteomics experiments. The system greatly simplifies the implementation of dual online UHPLC capability to most of the commercial UHPLC system, by employing only two additional switching valves. Two analytical RP capillary columns and two SPE columns were installed on a 10-ports switching valve in such a way that the two pairs of SPE and analytical columns operated independently and alternatively. For the MRM experiment, the sDO-UHPLC system increased the experimental throughput by about two fold, while achieving high reproducibility in retention time and high linear correlation despite using two separate analytical columns. In addition the system was successfully used for high throughput global profiling of a complex gastric tissue proteome. The developed system is significantly simpler than previous techniques and found to be easy and robust to operate. It thereby holds promise to become a base technique of various LC systems for high-throughput deep proteome profiling and targeted MS analysis. We are currently applying sDO-UHPLC system to quantify endogenous proteins in a large number of gastric tissue samples in biomarker-focused studies.

Acknowledgements

This work was supported in part by the Multi-omics Research Program through the National Research Foundation of Korea funded by the Korean Ministry of Science, ICT & Future Planning (NRF-2012M3A9B9036675). This work was also supported by a grant from the Ministry of Science, ICT, and Future Planning, Republic of Korea (Grant No. NRF-2014M3C7A1046047). Bio-specimens were provided by Chonnam University Hwasun Hospital (07SA2013010-001) and Asan Medical Cancer (2013-7), which are members of the National Biobank of Korea that is supported by the Ministry of Health, Welfare and Family Affairs. We thank Mr. Jack Reeves of Valco Instruments Co. Inc. for many years of invaluable helps and advices on developing our LC technologies until his retirement.

Notes and references

1. S. M. Hanash, S. J. Pitteri and V. M. Faca, *Nature*, 2008, **452**, 571-579.
2. P. Wang, J. R. Whiteaker and A. G. Paulovich, *Cancer biology & therapy*, 2009, **8**, 1083-1094.
3. Y. Zhang, B. R. Fonslow, B. Shan, M. C. Baek and J. R. Yates, 3rd, *Chemical reviews*, 2013, **113**, 2343-2394.
4. N. Rifai, M. A. Gillette and S. A. Carr, *Nature biotechnology*, 2006, **24**, 971-983.
5. S. Surinova, R. Schiess, R. Huttenhain, F. Cerciello, B. Wollscheid and R. Aebersold, *Journal of proteome research*, 2011, **10**, 5-16.
6. A. G. Paulovich, J. R. Whiteaker, A. N. Hoofnagle and P. Wang, *Proteomics. Clinical applications*, 2008, **2**, 1386-1402.
7. A. P. Drabovich, E. Martinez-Morillo and E. P. Diamandis, *Biochimica et biophysica acta*, 2015, **1854**, 677-686.
8. T. Shi, D. Su, T. Liu, K. Tang, D. G. Camp, 2nd, W. J. Qian and R. D. Smith, *Proteomics*, 2012, **12**, 1074-1092.
9. T. Kocher, R. Swart and K. Mechtler, *Analytical chemistry*, 2011, **83**, 2699-2704.
10. P. Picotti, B. Bodenmiller, L. N. Mueller, B. Domon and R. Aebersold, *Cell*, 2009, **138**, 795-806.
11. C. E. Parker, T. W. Pearson, N. L. Anderson and C. H. Borchers, *The Analyst*, 2010, **135**, 1830-1838.
12. N. Bisson, D. A. James, G. Ivosev, S. A. Tate, R. Bonner, L. Taylor and T. Pawson, *Nature biotechnology*, 2011, **29**, 653-658.
13. A. J. Percy, A. G. Chambers, J. Yang, D. B. Hardie and C. H. Borchers, *Biochimica et biophysica acta*, 2014, **1844**, 917-926.
14. T. Shi, X. Sun, Y. Gao, T. L. Fillmore, A. A. Schepmoes, R. Zhao, J. He, R. J. Moore, J. Kagan, K. D. Rodland, T. Liu, A. Y. Liu, R. D. Smith, K. Tang, D. G. Camp, 2nd and W. J. Qian, *Journal of proteome research*, 2013, **12**, 3353-3361.
15. T. Shi, T. L. Fillmore, Y. Gao, R. Zhao, J. He, A. A. Schepmoes, C. D. Nicora, C. Wu, J. L. Chambers, R. J. Moore, J. Kagan, S. Srivastava, A. Y. Liu, K. D. Rodland, T. Liu, D. G. Camp, 2nd, R. D. Smith and W. J. Qian, *Analytical chemistry*, 2013, **85**, 9196-9203.
16. H. K. Min, S. W. Hyung, J. W. Shin, H. S. Nam, S. H. Ahn, H. J. Jung and S. W. Lee, *Electrophoresis*, 2007, **28**, 1012-1021.
17. E. A. Livesay, K. Tang, B. K. Taylor, M. A. Buschbach, D. F. Hopkins, B. L. LaMarche, R. Zhao, Y. Shen, D. J. Orton, R. J. Moore, R. T. Kelly, H. R. Udseth and R. D. Smith, *Analytical chemistry*, 2008, **80**, 294-302.
18. H. Wang and S. M. Hanash, *Journal of proteome research*, 2008, **7**, 2743-2755.
19. D. J. Orton, M. J. Wall and A. A. Doucette, *Journal of proteome research*, 2013, **12**, 5963-5970.
20. H. Lee, J. H. Lee, H. Kim, S. J. Kim, J. Bae, H. K. Kim and S. W. Lee, *Journal of chromatography. A*, 2014, **1329**, 83-89.
21. Y. Wang, F. Yang, M. A. Gritsenko, Y. Wang, T. Clauss, T. Liu, Y. Shen, M. E. Monroe, D. Lopez-Ferrer, T. Reno, R. J. Moore, R. L. Klemke, D. G. Camp, 2nd and R. D. Smith, *Proteomics*, 2011, **11**, 2019-2026.
22. J. H. Lee, S. W. Hyung, D. G. Mun, H. J. Jung, H. Kim, H. Lee, S. J. Kim, K. S. Park, R. J. Moore, R. D. Smith and S. W. Lee, *Journal of proteome research*, 2012, **11**, 4373-4381.
23. S. W. Hyung, M. S. Kim, D. G. Mun, H. Lee and S. W. Lee, *The Analyst*, 2011, **136**, 2100-2105.
24. B. MacLean, D. M. Tomazela, N. Shulman, M. Chambers, G. L. Finney, B. Frewen, R. Kern, D. L. Tabb, D. C. Liebler and M. J. MacCoss, *Bioinformatics*, 2010, **26**, 966-968.

25. B. Shin, H. J. Jung, S. W. Hyung, H. Kim, D. Lee, C. Lee, M. H. Yu and S. W. Lee, *Molecular & cellular proteomics : MCP*, 2008, **7**, 1124-1134.
26. S. Kim and P. A. Pevzner, *Nature communications*, 2014, **5**, 5277.
27. K. Jeong, S. Kim and N. Bandeira, *BMC bioinformatics*, 2012, **13 Suppl 16**, S2.
28. B. Zhang, J. Wang, X. Wang, J. Zhu, Q. Liu, Z. Shi, M. C. Chambers, L. J. Zimmerman, K. F. Shaddox, S. Kim, S. R. Davies, S. Wang, P. Wang, C. R. Kinsinger, R. C. Rivers, H. Rodriguez, R. R. Townsend, M. J. Ellis, S. A. Carr, D. L. Tabb, R. J. Coffey, R. J. Slebos, D. C. Liebler and C. Nci, *Nature*, 2014, **513**, 382-387.