Korean J. Chem. Eng., **30**(2), 417-421 (2013)
DOI: 10.1007/s11814-012-0134-9

A new flow path design for multidimensional protein identification technology using nano-liquid chromatography electrospray ionization mass spectrometry

Thangamani Rajesh*,‡, Hyung-Yeon Park,‡, Eunjung Song***, Changmin Sung***,****, Sung-Hee Park***,****, Jae-Hun Lee***,****, Dongwon Yoo***, Yun-Gon Kim*****, Jong-Min Jeon*, Byung-Gee Kim***,****, and Yung-Hun Yang*,†**

*Department of Microbial Engineering, College of Engineering, Konkuk University, 1, Hwayang-dong, Gwangjin-gu, Seoul 143-701, Korea
**Bio-MAX Institute, Seoul National University, Gwanak-gu, Seoul 151-742, Korea
***School of Chemical and Biological Engineering, Institute of Molecular Biology and Genetics, Seoul National University, Gwanak-gu, Seoul 151-742, Korea
****Interdisciplinary Program for Bioengineering, Seoul National University, Seoul 151-744, Korea
*****Chemical Engineering, Soongsil University, 511, Sangdo-dong, Seoul 156-743, Korea
(*Received 18 April 2012 • accepted 14 August 2012*)

Abstract−Multidimensional protein identification technology (MudPIT) is one of the most versatile methods for separating and identifying highly complex peptides or proteins. However, there are still inherent problems resulting from salt in eluent systems and instrumentation with MudPIT. We designed a novel and simple flow path using two-valve system and successfully performed a fully automated peptide analysis using MudPIT coupled with nano-liquid chromatography electrospray ionization mass spectrometry (nLC-ESI-MS). It enables to generate a remarkably stable nanospray during the MudPIT analysis and realize the fully automated MudPIT system. This column arrangement could be easily installed to avoid laborious loading steps and unstable ionization from discontinuous flow. Consequently, the new flow path design for MudPIT system guarantees the detection of more peptides and higher protein coverage in proteome analysis.

Key words: MudPIT, Nano LC-ESI-MS, SCX, C18, Peptide Identification

INTRODUCTION

The main focus of proteomics is to completely understand protein expression, dynamics, and connections among proteins (proteome) in cells under specific conditions [1]. The development of mass spectrometry (MS) and protein purification/separation techniques accelerated advances in the proteomic era and revolutionized proteomics [2]. MS is a tool to analyze molecules through ionization using electron energy and subsequent separation of ions [3]. It has been used for small molecules, but recently developed soft-ionization methods have spurred its application to biomolecules [4,5]. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) is the most widely used method for protein resolution and identification, although it is laborious, tedious, and time-consuming [1].

Robust and automated high throughput separation technologies are needed, and multidimensional protein identification technology (MudPIT) is the most versatile and distinguished method for rapid and large-scale proteome analysis using multidimensional liquid chromatography (LC) and tandem MS [6,7]. In the process of MudPIT, the denatured and reduced proteome is enzymatically digested to generate a peptide mixture, and this peptide mixture is applied to multidimensional LC composed of a strong cation exchange (SCX) and a reversed-phase (RP) chromatography column [8,9]. Whole peptide samples are loaded onto the SCX column, and eluted to the RP column using a solution containing low concentration of salt. Once the peptides are moved to the RP column, they are eluted with an acetonitrile (ACN) gradient and then sprayed into the MS. This process is repeated with a higher concentration salt solution until all samples are completely eluted. The combination of these two columns generates an alternative and efficient two-dimensional separation system like 2D-PAGE, and the two-stage separation increases efficiency of the MS analysis [8,9].

Although there have been several modifications to improve MudPIT by applying ultra-high-pressure, double vents and an autosampler, the flow configurations are still very complicated [8,9]. Compared with initial MudPIT experiments requiring a pressure vessel to load peptides, the loading steps are clearly simplified using an autosampler. However, the flow to the MS was not continuous in most cases, because of the washing step and the use of a vent tee for waste flow. Discontinuous flow stops ionization of the solution in the MS during the washing step, which normally lasts 5-10 min. As a result, when the sample is injected into the MS, the sample is lost because of inefficient ionization. Considering that a pre-run is applied before every MS analysis, it is easy to see that discontinuous flow to the MS greatly decreases the efficiency of the MS analysis.

However, when MudPIT is applied to nanoLC system, the high salt solution passes through the nano tip, decreasing the buildup of spray and bubbles produced in the process resulting in an unstable spray condition. This decreases ionization efficiency and thus, the analysis must be stopped to remove high salt around the nano tip

†To whom correspondence should be addressed.
E-mail: seokor@konkuk.ac.kr
‡Equal contribution by two authors.

[10-12]. In this study, we designed a flow path using a two-valve nano LC system to avoid problems associated with salt formation using nano tip in MudPIT. As a result, we successfully performed the fully automated process of peptide analysis by MudPIT with nLC-electrospray ionization (ESI)-MS.

MATERIALS AND METHODS

1. Materials

Bovine serum albumin (BSA), ammonium acetate, and ACN were purchased from Sigma (St. Louis, MO, USA) and C18 beads were obtained from Vydac (218TP51; Deerfield, IL, USA). The SCX column (5 μm, 4.0×125 mm, Waters, Milford, MA, USA) was cut with a lathe, and SCX beads were collected.

2. Peptide Preparation

The BSA sample (100 μg) was incubated in 10 mM dithiothreitol for 30 minutes at 56 ºC. After cooling to room temperature, the BSA sample was treated with iodoacetamide (final, 20 mM) for 30 minutes in the dark at room temperature for carboxyamidomethylation. Sequencing grade trypsin (100 ng; Promega, Madison, WI, USA) in 100 mM NH_4HCO_3 was added for in-solution digestion, and proteins were digested overnight at 37 ºC. Tryptic peptides were collected, dried in a vacuum drier, and stored at −20 ºC for further analysis [13].

3. Packing Method of Column

A fused silica capillary column (360 μm OD, 100 μm ID) was cut to 10 cm and washed with 100% methanol. Potassium silicate solution (KASIL) (150 μl) was placed in an Eppendorf tube and 50 μl of formamide was added, vortexed for 1 min, and centrifuged briefly. One end of the capillary column was dipped into mixture and wiped with a paper towel. The column was baked at 90-100 ºC for at least 3 hr and then washed with methanol for 3 min using a pressure vessel. The C18 beads were packed using a home-made pressure bomb to a 7 cm length for ACN gradient analysis, and the SCX and C18 beads were packed to 3 cm and 4 cm for MudPIT analysis, respectively.

4. Analytical Methods

A fused silica capillary column (360 μm OD, 100 μm ID) was pulled with a P-2000 laser puller (Sutter Instrument) to make nano spray tips [14]. Peptides were eluted with a 0-90% ACN (0.5% formic acid) gradient solution for 60 min at 200 nl/min and analyzed by nLC-MS/MS (Ultimate 3000, Dionex) using an LTQ-orbitrap MS (Thermo Finnigan; Thermo Fisher, Waltham, MA, USA). The peptides were eluted by MudPIT using ammonium acetate and an ACN gradient by 2 μl of double distilled water (DDW) (0 mM of ammonium acetate) from SCX to RP. Peptides were eluted with a 0-90% ACN (0.5% formic acid) gradient solution for 60 min at 200 nl/min and analyzed by MS. Then, 100, 200, and 500 mM ammonium acetate solution was injected in order instead of DDW (0 mM of ammonium acetate). Different concentrations of ammonium acetate were located in separate vials in the autosampler, and 2 μl was injected. Peptide ions were detected in full scan mode from 400-1,700 m/z followed by three data-dependent MS/MS scans (isolation width: 1.5 m/z, 35% normalized collision energy, dynamic exclusion for 5 min) in a completely automated fashion. Proteins were identified by searching MS/MS spectra against a nonredundant (NR) protein database using SEQUEST Sorcerer (Thermo Scientific) along with the following criteria to sort out proteins from the MS spectra: First, the cross-correlation score (Xcorr) should be >1.7 for +1 charged tryptic peptides, >2.5 for +2 charged peptides, or >3.0 for +3 charged peptides. Second, the delta correlation value (Cn) should be at least 0.15, regardless of the charge state [14].

RESULTS AND DISCUSSION

1. Peptide Analysis by Designing a nLC-ESI-MS Flow Path

In MudPIT analysis, the flow to the MS is not continuous due to washing step and usage of vent tee for waste flow. This kind of

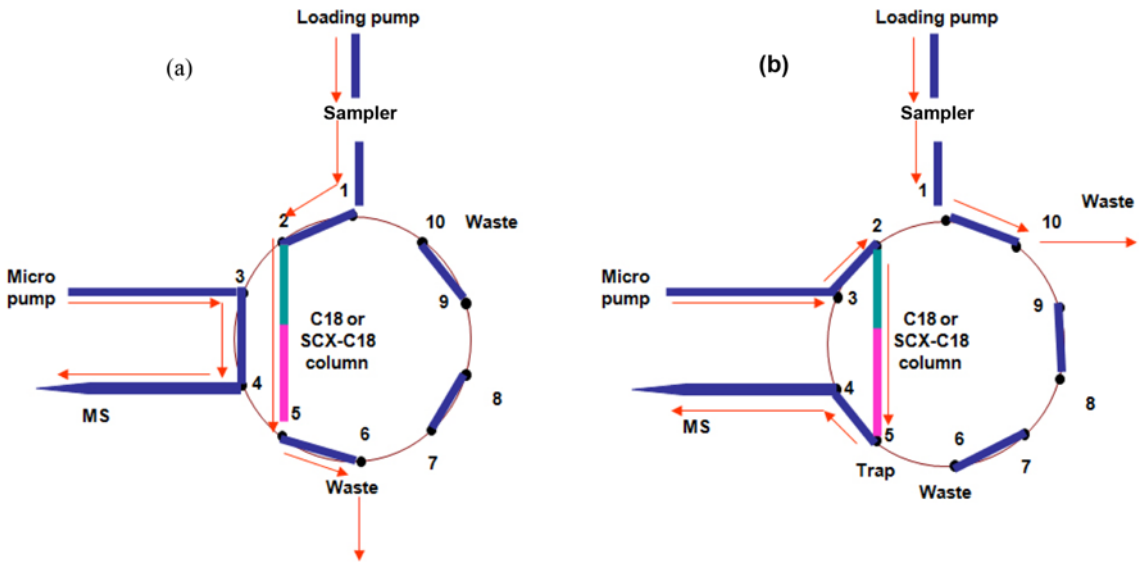

Fig. 1. Instrumental setup for fully automated multidimensional protein identification technology (MudPIT). The column part will be replaced by either C19 or SCX-C18 depending on analytical methods. The arrows represented the flow of buffer from micro pump and loading pump. Flow path for loading mode (a) and analytical mode (b) were changed from 1-2 position to 1-10 position for analysis.

discontinuous flow stops ionization of the solution in the MS during the washing step. This discontinuous flow to the MS greatly decreases the efficiency of the MS analysis. To solve the flow discontinuity, we designed a novel flow path with a nano LC. The column configuration is shown in Fig. 1. While the MS spray is continued through the whole analysis and loading process using a micro pump, the washing and separating steps are performed by a loading pump. Sample (2 μl) was loaded onto the C18 or SCX-C18 column from the autosampler at ports 1 or 2 connecting 3-4, 5-6, 7-8 and 9-10 (Fig. 1(a)). Then, the sample was washed by continuous flow from the loading pump (200 μl/min) for at least 3 min. After the valve was changed to the 1-10 position, connecting 2-3, 4-5, 6-7, and 8-9 (Fig. 1(b)), the column was connected to the MS without stopping flow. The ACN gradient could be applied with the micro pump, and peptides eluted as the ACN concentration increased.

Peptides were analyzed by two consecutive ACN gradient cycles within 80 min to examine the feasibility of this design. As seen from the total ion current chromatogram in Supplementary Fig. 1(a), peptides were well resolved and eluted using this new flow path design. Most peptides were eluted at the first gradient cycle and 42% of all peptide sequence coverage was achieved. Among the matched peptide peaks, b and y ions from peptide fragmentation were unambiguously detected by the collision induced dissociation method as shown for the "LVTDLTK" (Supplementary Fig. 1(b)) or "YLYEIAR" (Supplementary Fig. 1(c)) peptides. As a result, we analyzed the BSA tryptic peptides using this new flow path design, which simplified the whole peptide analytical process.

2. Peptide Analysis by ACN Gradient

Several analyses were examined with the C18 column before we applied MudPIT. We installed the home-made C18 column and performed four cycles of peptide analysis with the ACN gradient. The first cycle showed 39.0% coverage with 14 matched peptides, and this process was repeated with the same ACN gradient method (Table 1). The shapes of the total ion current chromatograms were similar (Supplementary Fig. 2(a)-(d)); however, most peptides were analyzed during the first cycle, suggesting that many peptides were eluted by ACN and that there was no other interaction between peptides with C18 except hydrophobic interactions. The overall coverage of peptides was 46.1% with 25 peptides (Table 1, Supplementary Fig. 2(e)).

3. Peptide Analysis by MudPIT

Sample (2 μl) was loaded onto the C18 or SCX-C18 column and injected from the autosampler at the ports 1 and 2 positions (Fig. 1(A)). Then, the valve was diverted to the 1-10 position so the ACN gradient could be applied by micro pump to elute the peptides (Fig. 1(B)). After one cycle was completed, the valve was reverted to the 1-2 position, and 2 μl of ammonium acetate (100 mM) was loaded to move the peptides bound to the SCX part to the C18 part of the column, and the sample was washed for 3 min with the loading pump. Once the valve was changed to the 1-10 position, the analysis was initiated, ACN was applied to the column, and more peptides were eluted. Similarly, 200 and 500 mM ammonium acetate was sequentially applied. Samples were eluted well following four cycles of the ACN and ammonium acetate gradient, showing a similar total ion current chromatogram (Supplementary Fig. 3(a)-(d)). As a result, 60.5% sequence coverage with 35 matched peptides was obtained (Table 1, Supplementary Fig. 3(e)). Compared with the C18 only method, the interaction of the peptide with SCX retained many peptides; thus, the second cycle also eluted many peptides like the first cycle of analysis (Table 1). The number of detected peptides is decreased dramatically from the third cycle, suggesting that 100 mM ammonium acetate eluted many of the peptides from the SCX column. A more detailed gradient condition such as 0, 20, 40, 60, 80, and 100 will be very useful to obtain higher peptide coverage.

4. Application of Simple MudPIT to Targeted Proteomics of *Streptomyces*

To examine whether simple MudPIT method works with a real system like complex proteome sample, we applied it to the screening of proteins interacting with RNA polymerase in *Streptomyces*. *Streptomyces coelicolor* J1980 whose carboxy terminus of the *S. coelicolor* β' subunit of RNA polymerase (RpoC) was modified to contain six histidine residues [15], was chosen to reveal what other subunit proteins interact with RpoC to make the whole RNA polymerase. *Streptomyces coelicolor* J1980 culture was harvested at 72 hrs of growth and the cells were disrupted by sonication [15]. Acquired proteome was subjected to nickel-NTA agarose affinity chromatography; then histidine tagged RpoC and possible proteins interacting with it were successfully captured. The proteins were digested with trypsin and the peptides were divided into halves to be analyzed by ACN gradient method or by MudPIT method.

MudPIT method gave a greater number of different peptides (289 hits) compared to ACN gradient method (281 hits) and MudPIT detected 12 proteins, whereas ACN gradient method detected only 8 proteins, which have more than two different peptides detected indicating that peptide mixture was separated evenly by the consecutive SCX and C18 chromatography (Table 2(b)).

Table 1. Analysis of bovine serum albumin trypsin-digested peptides using an acetonitrile gradient with C18 and multidimensional protein identification technology (MudPIT) with SCX-C18

Peptide analysis by ACN gradient with C18		Peptide analysis by MudPIT with SCX-C18	
Analysis cycle	Sequence coverage (%) (number of matched peptides)	Analysis cycle	Sequence coverage (%) (number of matched peptides)
1st	39.0 (14)	1st (0 mM)[a]	41.7 (11)
2nd	16.1 (6)	2nd (100 mM)	44.3 (12)
3rd	12.2 (5)	3rd (200 mM)	15.7 (7)
4th	10.1 (4)	4th (500 mM)	15.1 (6)
Overall	46.1 (25)	Overall	60.5 (35)

[a]Concentration of ammonium acetate

Table 2. Comparison of detected peptides with His-tagged RpoC from *Streptomyces coelicolor* J1980 using an acetonitrile gradient with C18 (a) and multidimensional protein identification technology (MudPIT) with SCX-C18 (b)

(a) Proteins detected using an acetonitrile gradient with C18

	Protein[a]	Score (Xc)[b]	SCO #	Gene	Peptide (Hits)
1	DNA-directed RNA polymerase beta chain	70.20	SCO4654	*rpoB*	61
2	Putative cell division trigger factor	50.19	SCO2620	-	71
3	Putative transferase	40.27	SCO2592	-	26
4	Putative carboxyl transferase	40.23	SCO5535	*accB*	36
5	DNA-directed RNA polymerase beta' chain (fragment)	30.26	SCO4655	*rpoC*	44
6	Glutamine synthetase I	30.14	SCO2198	*glnA*	19
7	Glucosamine-fructose-6-phosphate aminotransferase	20.15	SCO2789	*glmS2*	17
8	60 kD chaperonin cpn60	20.13	SCO4762	*groEL1*	7

(b) Proteins detected using multidimensional protein identification technology (MudPIT) with SCX-C18

	Protein	Score (Xc)	SCO #	Gene	Peptide (Hits)
1	60 kD chaperonin cpn60	90.23	SCO4762	*groEL1*	66
2	RNA polymerase alpha subunit	40.20	SCO4729	*rpoA*	30
3	50S ribosomal protein L2	30.28	SCO4705	*rplB*	29
4	DNA-directed RNA polymerase beta' chain (fragment)	30.24	SCO4655	*rpoC*	22
5	Glucosamine-fructose-6-phosphate aminotransferase	30.24	SCO2789	*glmS2*	18
6	Elongation factor TU-1	30.23	SCO4662	*tuf1*	54
7	DNA-directed RNA polymerase beta chain	30.18	SCO4654	*rpoB*	7
8	Putative cell division trigger factor	30.18	SCO2620	-	7
9	Hypothetical protein SC6D7.18c.	20.20	SCO1421	-	37
10	30S ribosomal protein S2	20.20	SCO5624	*rpsB*	4
11	Hypothetical protein	20.20	SCO2077	-	13
12	50S ribosomal protein L4	20.13	SCO4703	*rplD*	2

[a]Only proteins with two or more different peptides detected are listed in the table

[b]The preliminary scores obtained during the database search using SEQUEST Sorcerer. The score is calculated based on the number of ions from the MS/MS spectrum that matched the candidate peptide

It was reported that RNA polymerase is composed of multi-subunits such as α(*rpoA*), β(*rpoB*), β'(*rpoC*), ω in *E. coli* [16]; however, there are few reports on possible proteins that interact with RNA polymerases in prokaryotes, except subunits. In this study, both methods detected RpoC and RpoB (Table 2(a)-(b)). The α subunit (RpoA) unit was well detected with two different peptides by MudPIT, although α unit was also detected by one peptide (seven hits) with ACN gradient method (data not shown). Subunit ω was also detected by one peptide (seven hits) with MudPIT, not with ACN gradient method (data not shown). Other proteins such as 50S ribosomal protein L2 (SCO4705), 30S ribosomal protein S2 (SCO5624), 50S ribosomal protein L4 (SCO4703) were detected by MudPIT only and are highly expected to interact with RNA polymerase considering their roles with DNA at transcription and translation processes. Among detected proteins, 60 kD chaperonin Cpn60 (SCO4762) and putative cell division trigger factor (SCO2620) seem to be the newly found candidates to interact with RNA polymerase and will be interesting targets for further study. Overall, this simplified MudPIT method gives better peptide resolution and more peptides detected in both the model and the real system.

CONCLUSION

This study showed a simple application of MudPIT by designing an nLC-MS flow path and achieved continuous ionization into MS. This arrangement of columns avoids laborious loading steps and unstable ionization from discontinuous flow due to the venting splitter. This system is also easily installed and fully automated. As a result, we detected more peptides with greater protein coverage. This method will be easily adapted for peptide identification considering the high speed and high resolution of MS.

ACKNOWLEDGEMENTS

This paper was supported by Konkuk University in 2011.

SUPPORTING INFORMATION

Additional information as noted in the text. This information is available via the Internet at http://www.springer.com/chemistry/journal/11814.

REFERENCES

1. A. Gorg, W. Weiss and M. J. Dunn, *Proteomics*, **4**, 3665 (2004).
2. R. Aebersold and M. Mann, *Nature*, **422**, 198 (2003).
3. J. W. Hager, *Anal. Bioanal. Chem.*, **378**, 845 (2004).
4. K. Tanaka, K. Waki, Y. Ido, S. Akita, Y. Yoshida and T. Yoshida,

Rapid Commun. Mass Spectrom, **2**, 151 (1988).
5. J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science*, **246**, 64 (1989).
6. A. Ducret, I. Van Oostveen, J. K. Eng, J. R. Yates, 3rd and R. Aebersold, *Protein Sci*, **7**, 706 (1998).
7. A. J. Link, J. Eng, D. M. Schieltz, E. Carmack, G. J. Mize, D. R. Morris, B. M. Garvik and J. R. Yates, 3rd, *Nat. Biotechnol.*, **17**, 676 (1999).
8. D. A. Wolters, M. P. Washburn and J. R. Yates, 3rd, *Anal. Chem.*, **73**, 5683 (2001).
9. M. P. Washburn, D. Wolters and J. R. Yates, 3rd, *Nat. Biotechnol.*, **19**, 242 (2001).
10. A. Motoyama, J. D. Venable, C. I. Ruse and J. R. Yates, 3rd, *Anal. Chem.*, **78**, 5109 (2006).
11. A. W. Guzzetta and A. S. Chien, *J. Proteome. Res.*, **4**, 2412 (2005).
12. P. Taylor, P. A. Nielsen, M. B. Trelle, O. B. Horning, M. B. Andersen, O. Vorm, M. F. Moran and T. Kislinger, *J. Proteome. Res.*, **8**, 1610 (2009).
13. Y. H. Yang, K. Lee, K. S. Jang, Y. G. Kim, S. H. Park, C. S. Lee and B. G. Kim, *Anal. Biochem.*, **387**, 133 (2009).
14. Y. G. Kim, D. S. Shin, Y. H. Yang, G. C. Gil, C. G. Park, Y. Mimura, D. K. Cooper, P. M. Rudd, R. A. Dwek, Y. S. Lee and B. G. Kim, *Chem. Biol.*, **15**, 215 (2008).
15. M. J. Babcock, M. J. Buttner, C. H. Keler, B. R. Clarke, R. A. Morris, C. G. Lewis and M. E. Brawner, *Gene*, **196**, 31 (1997).
16. K. Severinov, R. Mooney, S. A. Darst and R. Landick, *J. Biol. Chem.*, **272**, 24137 (1997).