Development of Ultra-High Pressure Capillary Reverse-Phase Liquid Chromatography/Tandem Mass Spectrometry for High-Sensitive and High-Throughput Proteomics

Min-Sik Kim, Woo-Suk Choie, Yong Seung Shin,[†,‡] Myeong-Hee Yu,[†] and Sang-Won Lee[*]

Department of Chemistry and Center for Electro- & Photo-Responsive Molecules, Korea University, 1, 5-ka, Anam-dong, Seongbuk-ku, Seoul 136-701, Korea
[†]Functional Proteomics Center, Korea Institute of Science and Technology, Seoul 136-791, Korea
[‡]School of Life Sciences, Korea University, Seoul 136-701, Korea
Received September 16, 2004

Recently mass spectrometry and separation methods such as liquid chromatography have become major tools in the field of proteomics. In this report, we describe in detail our efforts to develop ultra-high pressure capillary reverse-phase liquid chromatography (cRPLC) and its online coupling to a mass spectrometer by a nano-electrospray (nanoESI) interface. The RPLC system is constructed in house to deliver LC solvents at the pressure up to 20,000 psig, which is four times higher than conventional RPLC systems. The high operation pressure allows the efficient use of packed micro-capillary columns (50, 75 and 150 μm i.d., up to 1.5 m long). We will discuss the effect of column diameter on the sensitivity of cRPLC/MS/MS experiments and the utility of the developed technique for proteome analysis by its application in the analysis of proteome samples having different levels of complexity.

Key Words : Proteomics, Liquid chromatography, Tandem mass spectrometry

Introduction

Methodologies for global profiling of gene expression at the mRNA level as a function of different cellular states have been well established.[1,2] However, many studies showed that the mRNA level has little correlation with the expression in cell proteins that are real "workers" in signal transduction and other cellular processes.[3] Therefore, the analysis of proteome, the whole mixture of the expressed proteins in cells, should provide more accurate and direct information on the biological system as a whole and delineate the information transfer pathways being mediated by proteins at the molecular level.

Human proteome consists of more than at least 40,000 proteins and the component proteins have vastly different expression level (up to 10^9)[4] with a wide spectrum of chemical modifications.[5] The enormous complexity and differential expression of real proteome sample require high throughput and sensitive analytical platforms. With the development of soft ionization techniques, such as MALDI and electrospray, to ionize peptides or proteins into the gas phase, mass spectrometry became a major tool in analyzing cell proteins. However, the highest dynamic range of the state-of-the-art mass spectrometric technology is only several thousands[6] and this is clearly not large enough to address the "real" proteome sample that has up to 10^9 concentration differences.

Currently, various separation methods, such as single dimensional reverse-phase liquid chromatography (RPLC)[7] or multi-dimensional separation (*e.g.* MudPIT),[8] are online coupled with mass spectrometry by means of electrospray ionization for identifying and characterizing complex proteome samples.[9,10] The improvement in separation efficiency should result in increased number of detected peptides from a single LC/MS/MS experiment. Small inner diameter (15-75 μm) and long (> 1 m) capillary columns are essential for enhanced sensitivity and high resolution separation of complex peptide mixtures, but require the use of high pressure of up to 10,000 psig for efficient operation.[11,12] Shen *et al.* demonstrated that a peak capacity of 1000 could be achieved with 83 cm long capillary columns (with i.d. as small as 15 μm) that are operated at 10,000 psig. Even with the employment of high pressure of 10,000 psig, however, the linear flow rate inside of the packed capillary columns can be measured to several tenths of nanoliters per minute, which entails to employ nanoelectrospray.[11]

Here we present in details our efforts in establishing ultrahigh pressure capillary RPLC system utilizing micro-capillary columns for high throughput and highly sensitive proteomic experiments. The RPLC system has the maximum operating pressure up to 20,000 psig, which is two times higher than previous experiments.[11] We also evaluate the performances of packed capillary columns having inner diameters as small as 50 μm. It was observed that the sensitivity is dramatically improved upon using smaller ID capillary columns as demonstrated by changing the inner diameters from 150 to 50 μm, which are packed with 3 μm silica based reversed phase packing materials. The employment of the high performance capillary column to homebuilt ultra-high pressure LC system resulted in

[*]Author to whom all correspondence should be addressed: Tel: +82-2-3290-3137; Fax: +82-2-3290-3121; e-mail: sw_lee@korea.ac.kr

significant increase in separation resolution. It is also demonstrated that the developed LC system can be effectively coupled to high-efficiency mass spectrometry to analyze complex polypeptide mixtures, resulting in vastly increased dynamic range of proteome analysis.

Experimental Section

Materials and Sample Preparation. BSA (Bovine serum albumin) was dissolved in digestion buffer, 50 mM ammonium bicarbonate, pH 8.0, and methanol. The final concentration during the denaturation process was maintained to be *ca*. 0.1 μg/μL in 40% methanol.[13] In-solution digestion was performed at 37 °C for 20 h using trypsin at 1 : 50 concentration. After digestion, the peptide mixtures were dried by a SpeedVac (ThermoSavant, NY, USA), and resuspended in solvent A (0.05% trifluoroacetic acid (TFA) and 0.2% acetic acid in water).

S.cerevisiae haploid strains Y2805 (*MATa pep::his3 prb1-D1.6R can1 his1-200 ura3-52*) and AF-2 (*HMLa or HMRa ho ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 ssd1*) were used. In shake-flask culture, yeast strains were pre-incubated overnight in YPD (yeast extract 10 g, peptone 20 g, dextrose 20 g/L) media, and then transferred to flask again containing 500 mL of YPD media. Yeast strains were harvested by centrifugation at mid-log phase (OD_{600} = 3.6), distributed in four baffled flasks and incubated in a shaking incubator (30 ^{o}C, 200 rpm). Harvested cell pellets were washed several times with cysotol buffer (20 mM hepes (pH 6.8), 250 mM sorbitol, 150 mM potassium acetate, 1 mM magnesium acetate., 1 mM dithiothreitol) and suspended in 2 volumes of the cysotol buffer. Subsequently an equal cell volume of glassbeads (Sigma, St. Louis, MO, USA) were added to each tube. Cells were subjected to lysis by 10 cycles of vortexing for 1 min followed by 1 min of incubation on ice. A separate homogenizer (ThermoSavant, NY, USA) was used as well for cell disruption. Whole lysates were centrifuged for postnuclear fraction (5,000 g, 5 min) and soluble cytosol fraction was obtained by centrifugation on an ultracentrifuge (100,000 g, 4 ^{o}C, 30 min). The concentration of proteins was determined by a bradford method. The soluble fraction of cell lysate is subjected to tryptic digestion by the similar method as described above.

Sequencing-grade immobilized trypsin was purchased from Promega (Madison, WI, USA). All other reagents were of analytical or HPLC grade.

Column Packing Apparatus and Method. Figure 1 depicts the experimental apparatus used to manufacture packed micro-capillary columns. An air-driven liquid pump (Model DSHF-302, Haskel, Inc., Burbank, CA, USA) was used to generate a liquid pressure up to 55,000 psig. Current packing apparatus holds a liquid pressure up to 25,000 psig. The liquid pressure from the Haskel pump was measured by an ultra-high pressure gauge (100,000 psig, Astraguage, Ivyland, PA, USA) before the pressurized liquid was delivered to the home-built slurry mixer device. The packing material (C18-bonded 3 μm particles, 300 Å pores, Phenomenex, Torrance, CA, USA) was uniformly mixed and dissolved in 2-propanol aqueous solution (50%, v/v) by a micro-magnetic stirring bar inside of the slurry mixer. The C18 particles were packed into an empty fused silica capillary (Polymicro Technologies, Phoenix, AZ, USA) via a stainless steel tee (Valco international, Schenkon, Switzerland). One port of the tee is connected to a short pre-packed column (750 μm i.d. × 10 cm), which is used to vent pressure after completion of column packing. The vent column is closed during packing and is open after finishing packing a capillary column to speed up the depressurization

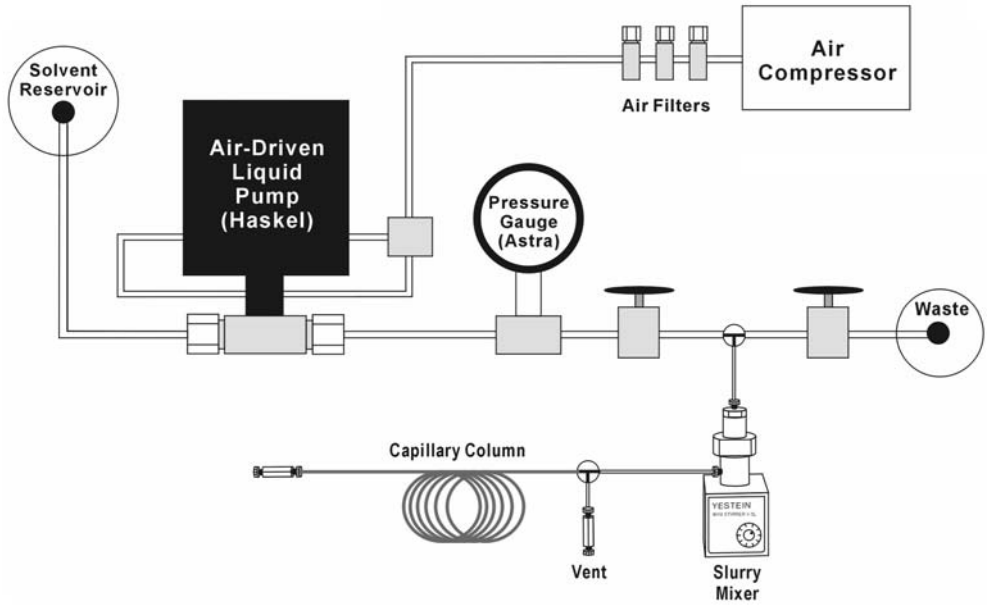

Figure 1. Schematics of ultrahigh pressure slurry packing apparatus used to manufacture capillary columns. Detailed procedures are described in the text.

while preventing packing material from backing up by sudden pressure decrease. Capillaries were cut into 1-1.5 m in length and connected to the tee, which is then connected to slurry mixer, by means of PEEK sleeve (400-μm i.d., Upchurch Scientific, Oak Harbar, WA, USA). The other end of the capillary was connected to a stainless steel union and a steel screen (2-μm pores, Valco international, Schenkon, Switzerland) was inserted to the union to prevent leaking of the packing material.

The slurry packing was initiated by applying liquid pressure of *ca*. 1,000 psig to the slurry mixer. Under an optical microscope (OAM 24SR, $\times$40 manification, Dongwon EnC, Sungnam, Korea), the packing process and condition are carefully inspected. As the column was being packed to several centimeters, the packing speed decreased. The liquid pressure was gradually increased up to 20,000 psig as packing proceeds to completion. Occasional tapping on the tee was necessary when the packing process was stopped. After the packing, the packed column was conditioned at 20,000 psig for 10 min in a sonic bath. The column was conditioned at 20,000 psig for additional 2 hours and completely depressurized overnight.

Ultra-high Pressure LC Systems. Figure 2 shows the schematics of ultra-high pressure two-column LC system. In this study, we have built two separate LC systems, a 10,000 and a 20,000 psig LC system respectively. Stable solvent flows at the ultra-high pressure were provided by two ISCO syringe pumps (Model 100DM for 10,000 psig system and Model 65D for 20,000 psig system, ISCO, Lincoln, NE, USA). Mobile phases (solvent A: 0.05% TFA and 0.2% acetic acid in water; solvent B: 90% acetonitrile, 0.1% TFA in water) were delivered by the two ISCO syringe pumps to a set of switching valves that allows development of solvent gradient, sample injection and column selection. The inset of Figure 2 shows the valve connection diagram. Two four port switching valves (Valco Instruments Co. Inc., Houston, TX, USA) and a six port switching valve (Valco Instruments Co.Inc., Houston, TX, USA) were used. A four port switching valve, termed solvent selector, is used for selecting solvent that enters to the solvent mixer and subsequently to packed capillary column. The internal volume of the solvent mixer is *ca*. 2.5 mL. Initially the mixer is filled with solvent A and, after the sample injection, solvent selector is set to solvent B to start gradient development. The mobile phases were mixed by a micro-magnetic stirring bar inside of the solvent mixer, and the solvent from the mixer was then directed to a six port switching valve (*i.e*. sample injector) and injects the sample loaded to the sample loop to one of the two capillary columns. Another six port switching valve, column selector, was used to select the column that will be used for experiments. As shown in the inset of Figure 2, one port of the column selector is directly connected to the solvent A pump and thereby allows one column being regenerated (or equilibrated) by solvent A while the other column is used for an LC/MS experiment. This 2-column configuration effectively eliminates the dead time (*ca.* 2 hrs) required for column washing by pure solvent B and column regeneration by pure solvent A between experiments since one capillary column was used for a separation while the other was being washed and regenerated. A fused-silica

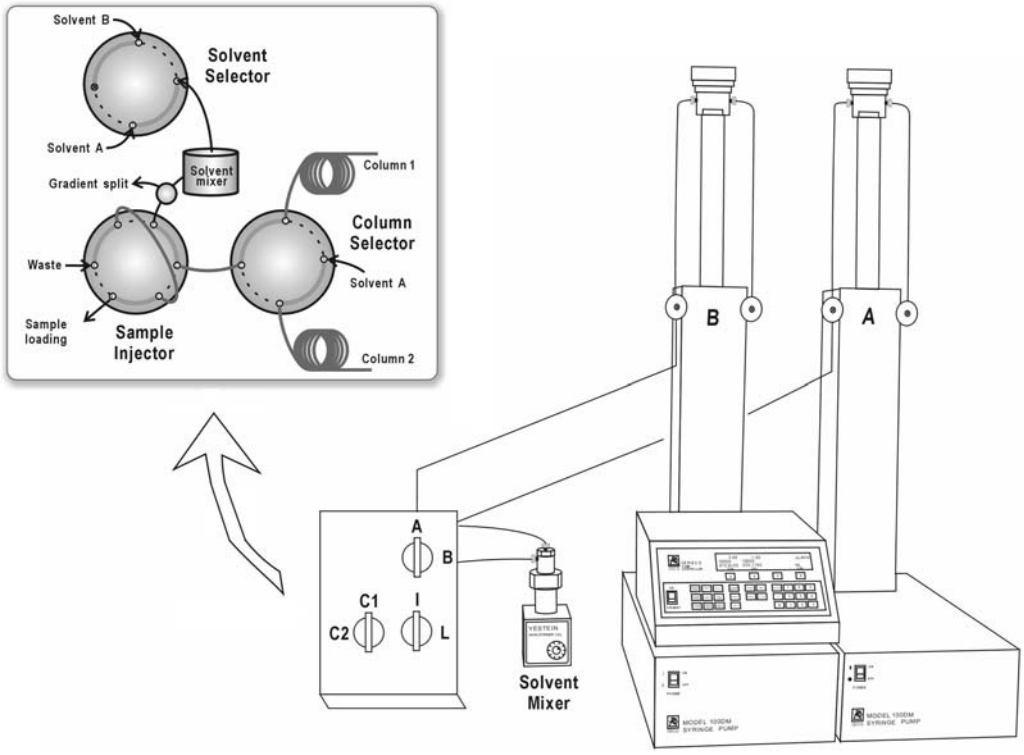

Figure 2. Schematics of the packed capillary LC system. Inset shows valve connection diagram for two-column system.

capillary flow restrictor (packed with C18 Jupiter, Phenomenex, Torrance, CA, USA) was connected between mixer and the sample injector valve to control the gradient development speed. Under the configuration of the current setup, an exponential gradient was generated over 200 min, encompassing a range of 0% to 75% solvent B.

cRPLC/ESI-MS. A quadrupole ion trap mass spectrometer (LCQ advantage, ThermoFinnigan, San Jose, CA, USA) and a quadrupole linear ion trap mass spectrometer (LTQ, ThermoFinnigan, San Jose, CA, USA) were used to for all experiments. Briefly, the mass spectrometers are equipped with home-built nanoESI interfaces which use an optical xyz-translational stage (460A series, Newport, Irvine, CA, USA) to precisely control the position of the spray emitter relative to the desolvation capillary. The interface is also designed to minimize the dead volume between the LC and the mass spectrometer by directly connecting the separation column and the spray emitter via a zero-volumn stainless steel union (Valco Instruments Co. Inc., Houston, TX, USA). An electrospray emitter was prepared by drawing a fused-silica capillary (30 μm i.d. and 150 μm o.d.) with a hung weight while applying a micro-torch flame (GT 3000S, Prince, Japan) until the capillary breaks. The electrospray emitter was connected to the packed capillary column by using a union. The temperature of the heated desolvation capillary was kept at 250 °C and electrospray voltage was 2.0 kV. No Neubulizer or counter-flowing drying gas was used.

Analysis of Mass Spectrometric Data. To identify the eluting peptides, the mass spectrometers were operated in data-dependent tandem MS mode where the four most abundant ions detected in a precursor MS scan were dynamically selected for subsequent MS/MS experiments (with collisional energy set to 35%), simultaneously incorporating dynamic exclusion option to prevent reacquisition of MS/MS spectra of the same peptides. The MS and MS/MS data from the LC/MS/MS experiment were analyzed using the SEQUEST program against a database which was constructed by combining the IPI yeast database (ftp://ftp.ebi.ac.uk/pub/databases/IPI/current) and a common contaminant database (http://www.ncbi.nih.gov). The database search is restricted to complete tryptic peptides. A home-built application written using Microsoft Visual Basic (VB) was used to filter the identified peptide sequences from the initial database search using criteria of Xcorr ≥ 1.8 for +1 charge state, ≥ 2.5 for +2 charge state, and ≥ 3.5 for +3.[14] The normalized difference between first and second match scores (ΔCn) higher than 0.08 was used in the filtering process as well. The VB application subsequently use each filtered peptide sequence to check "uniqueness" of the peptide by sequence matching of the peptide sequence against the protein database. If the peptide sequence is uniquely matched to a protein, the peptide is classified to a unique peptide. In order to minimize false positive identification, only proteins with 2 or more unique peptides were classified as positive protein identification. When a protein is identified by a single unique peptide, the corresponding MS/MS spectrum was manually inspected for positive identification.

Due to the sequence homologies of the proteins, many of filtered peptide sequences were matched to multiple proteins. In these cases, the VB application classifies proteins matched with the same peptide sequences into a protein group. The protein group with larger than 3 peptide matches was also considered as positive identification and only a single protein from the group that is either with largest sequence coverage or with representative annotation was reported. For all other cases, proteins were considered as positive identification only when the MS/MS spectra of matched peptides were manually inspected to be confident identification.

Results and Disscusion

Loading Capacity of Peptide Samples. In isocratic mode, increased sample volume results in degradation of separation resolution. However, in gradient mode of mobile phase composition, the introduction of relatively large sample volume has insignificant effect on the separation resolution, especially in the case where long and small ID capillary columns are employed. The initially long band of sample is compensated by the efficient and prolonged interaction of the analytes (in this case, peptides) with stationary phase. If sample is available in sufficient amount, it is desirable to increase sample loading quantity and this requires large loading capacity. Figure 3 shows chromatograms of LC/MS (LCQ advantage) experiments using 10, 100, and 1000 ng of tryptic BSA peptides. The column used was 1 m long and 50 μm-i.d capillary column. For comparison, the chromatograms were scaled to the data from the experiment of 100 ng injection. The chromatogram (Figure 3c) from 1000 ng

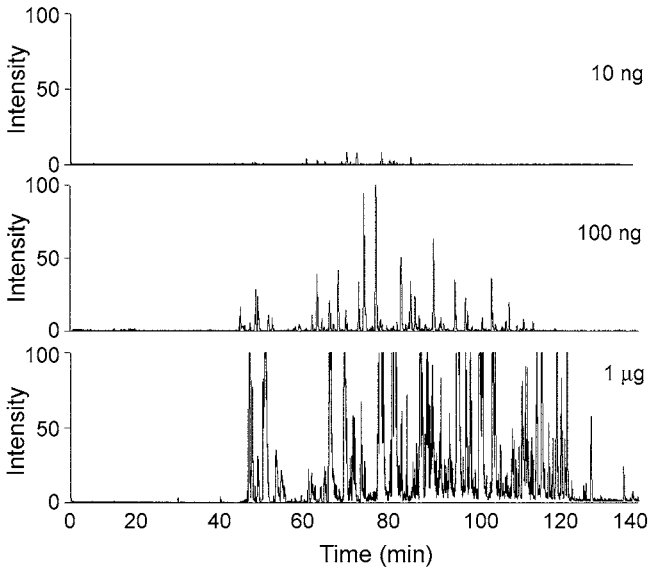

Figure 3. Comparison of chromatograms of LC/MS experiments using different amount of tryptic BSA peptides; (a) 10 ng, (b) 100 ng, (c) 1000 ng injection. The chromatograms were scaled to the data from the experiment of 100 ng injection.

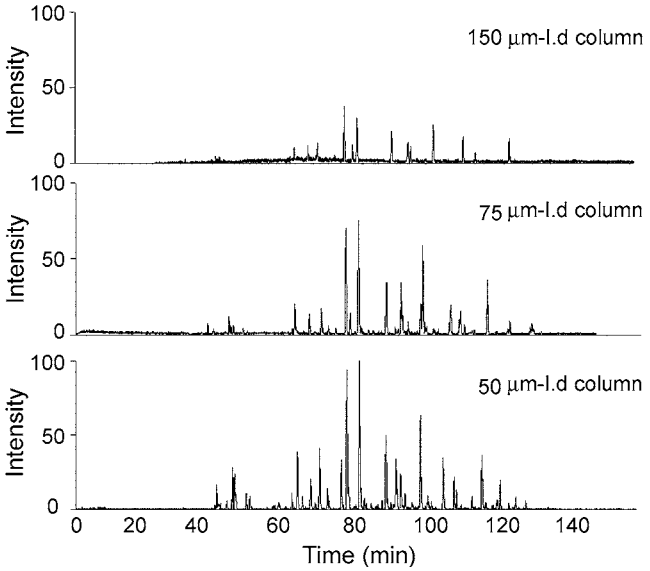

Figure 4. Comparison of chromatograms of LC/MS experiments using different inner diameters of capillary columns. The amount of injected samples was identical for the experiments. The chromatograms were scaled to the data from 50 μm capillary.

injection shows zoom-in views for 20% of the base peak intensity. Many additional peptides become detectable with increased amount of sample loaded. We observed no apparent peak broadening effect with increased amount of sample, suggesting sufficient sample loading capacity for proteomics application.

Dependence of Sensitivities on Inner Diameters of Capillary Columns. Improving resolution of LC separation enhances the detection dynamic range for extremely complex mixtures by resolving the low abundance species from high abundance components. Increased column length with decreased inner diameter is apparent choice to increase separation resolution and sensitivity. Figure 4 compares chromatograms of LC/MS (LTQ) experiments using different i.d. columns: (a) 150 μm for Figure 4a; (b) 75 μm for Figure 4b; and (c) 50 μm for Figure 4c. The columns are all 1 m in length. The amount of samples used for these experiments is 100 ng of tryptic BSA peptides and the operation pressure of LC was 10,000 psig. The sensitivities of LC/MS experiments were observed to dramatically increase with the decreased i.d. of capillary columns as evidenced by the increased peak intensities and the increased numbers of low lying peaks on the chromatograms while using the same amount of sample. However, the use of long and small i.d. capillary columns entails high back pressure to efficiently operate in conjunction with a mass spectrometer. The use of 1 m long and 50 μm-i.d. capillary column resulted in *ca*. 0.1 μL/min even at the operation pressure of 10,000 psig. While it is still possible to operate the nanoESI source at the flow rate lower than 0.1 μL/min, we experimentally observed the flow rate for practical nanoESI operation is around 0.1 μL/min or higher. Thereby, the use of even smaller i.d. and longer capillary columns than the 50 μm-i.d. column requires higher LC operation pressure, especially in conjunction with mass spectrometry.

Figure 5 shows a chromatogram obtained from an LC/MS (LTQ) experiment. The LC operation pressure for the experiment was 17,500 psig and the separation column was a 1.5 m long 50 μm-i.d. column. The inset of Figure 5 indicates that the full width at half maxima of the peak is ca. 8.5 s. The theoretical peak capacity of the LC/MS experiment is *ca*. 1,500, which is the highest for single dimensional peptide separation reported in literature.[11] While using long and small i.d.-capillary columns and maintaining good separation resolution, the linear flow rate through the capillary column was measured to be *ca*. 0.15 μL/min, which is practically well coupled to the nanoESI interface.

High Throughput and High Sensitive Proteomics Experiments by Ultra-high Pressure LC/MS/MS. Combining ultra-high pressure LC system and high efficient mass spectrometric analysis should provide a proteomic platform having both high sensitivity and high throughput. The ability of an LC system to separate complex proteome sample into more than 1,500 different fractions and the capability of MS system to dynamically select and/or exclude to avoid

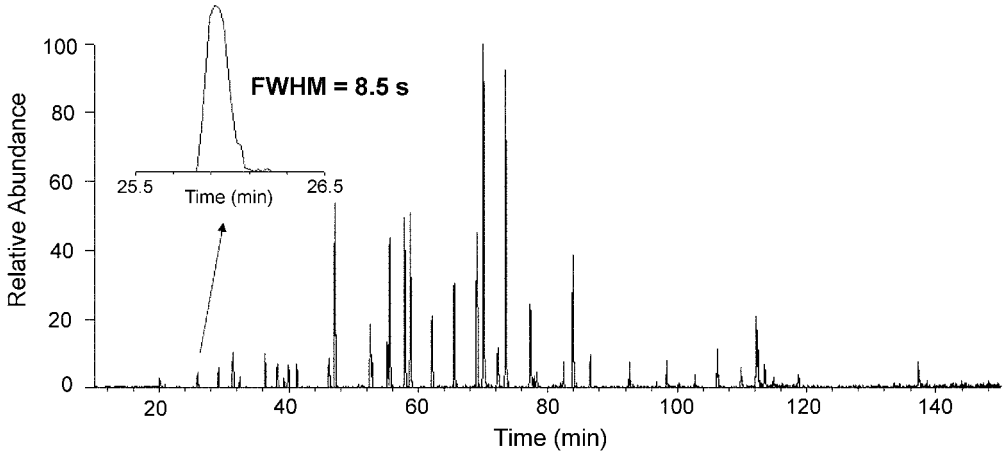

Figure 5. A chromatogram of LC/MS experiment using tryptic enolase peptides. The operation pressure was 17,500 psig. The inset indicates the peak width of a chromatogram peak is *ca*. 8.5 seconds, which corresponds to the total peak capacity of 1,500.

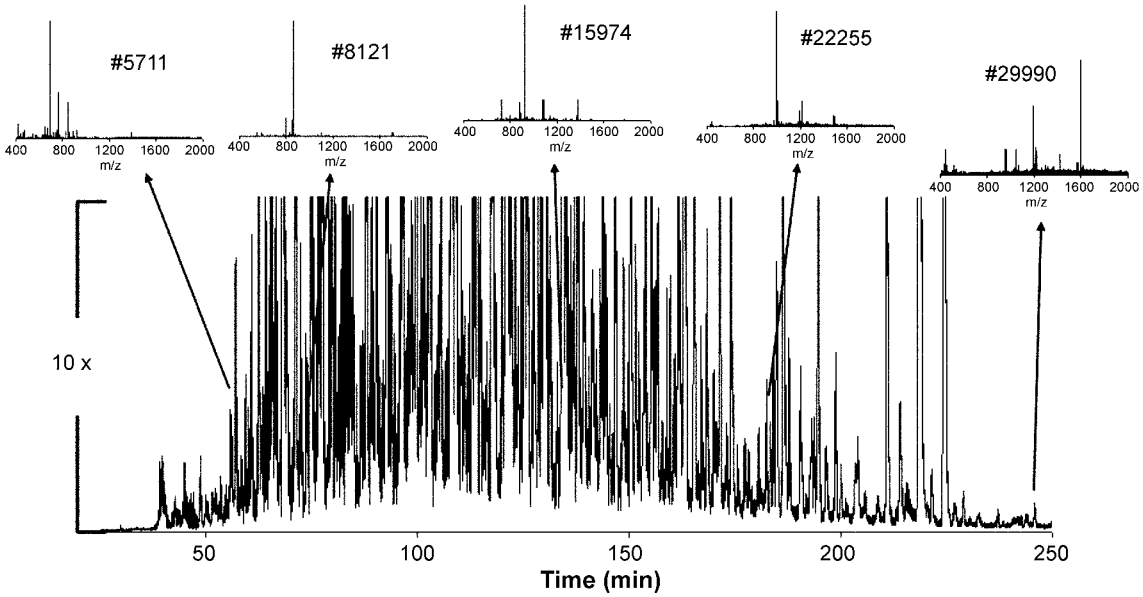

Figure 6. A chromatogram of LC/MS experiment using tryptic yeast peptides (20 μg injection). The insets display mass spectra measured from the experiment.

reacquisition of MS/MS spectra for the same peptides will significantly increase the dynamic range of the proteome analysis. Figure 6 shows a chromatogram for tryptic peptides of soluble fraction of yeast lysates. The chromatogram was obtained using a 150 cm long capillary column packed with 3-μm C18 particle (300 Å pore size) operating at 20,000 psig. The peptide mixture was directly injected to LC for analysis without any pre-treatment steps such as desalting. As evident from insets, which displays mass spectra corresponding to weak chromatographic peaks (less than 5% of base peak), many minor species were efficiently separated from their neighboring major components by this high-resolution separation and were detected with high signal-to-noise ratios. The database search followed by peptide filtering and group, which is described in detail in Experimental section, resulted in a total of 951 identified proteins. This is favorably comparable to any other previously reported numbers of identified yeast proteins. For example, Yates and co-workers previously identified a total of 1,465 yeast proteins using 420 μg soluble fraction, 440 μg lightly washed insoluble fraction and 490 μg heavily washed insoluble fraction, where each samples are independently subjected to 15 steps of SCX elutions for MudPIT experiments.[15] The total number of LC/MS/MS experiments was 45. We note that we have identified 951 proteins using the same search criteria but using only a limited amount of sample (*ca.* 20 μg) in a single LC/MS/MS experiment. It is also important to note that we used only soluble fraction of yeast lysates while the previous experiment used three different fractions of the cell lysates.

Conclusions

High performance packed capillary columns with i.d. as small as 50 μm and length as long as 1.5 m were manufactured in house using a home-built slurry packer. Details of the slurry column packer and the packing conditions were described. Long and small i.d. capillary columns were shown to be effectively employed to ultra-high pressure LC system having maximum operation pressure up to 20,000 psig. The high resolution separation established in this study can provide a peak capacity as high as 1,500. The combination of high resolution LC separation with high efficient mass spectrometric analysis significantly increased the dynamic range of proteome analyses. The combined LC/MS/MS technology identified over 900 proteins from a complex proteome sample from yeast lysates in a single experiment, demonstrating its utility for high throughput analyses of complex proteome samples having a broad spectrum of complexity.

Acknowledgement. This work was supported by the Korea Research Foundation Grant (KRF-2002-003-C00087).

References

1. Velculescu, V. E.; Zhang, L.; Zhou, W.; Vogelstein, J.; Basrai, M. A.; Bassett, D. E.; Hieter, P.; Vogelstein, B.; Kinzler, K. W. *Cell* **1997**, *88*, 243-251.
2. DeRisi, J. L.; Iyer, V. R.; Brown, P. O. *Science* **1997**, *278*, 680-686.
3. Gygi, S. P.; Rochon, Y.; Franza, B. R.; Aebersold, R. *Mol. Cell. Biology* **1999**, *19*, 1720-1730.
4. Anderson, N. L.; Anderson, N. G. *Mol. Cell. Proteomics* **2002**, *1*, 845-867.
5. O'Donovan, C.; Apweiler, R.; Bairoch, A. *Trends Biotechnol.* **2001**, *19*, 178-181.
6. Wolters, D. A.; Washburn, M. P.; Yates, J. R. *Anal. Chem*. **2001**, *73*, 5683-5690.
7. Shen, Y.; Toliæ, N.; Zhao, R.; Paša-Toliæ, L.; Li, L.; Berger, S. J.; Harkewicz, R.; Anderson, G. A.; Belov, M. E.; Smith, R. D. *Anal. Chem*. **2001**, *73*, 3011-3021.
8. Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.;

Morris, D. R.; Garvik, B. M.; Yates, J. R. *Nat. Biotechnol.* **1999**, *17*, 676-682.
9. Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. *J. Proteome Res.* **2003**, *2*, 43-50.
10. Shen, Y.; Smith, R. D. *Electrophoresis* **2002**, *23*, 3106-3124.
11. Shen, Y.; Zhao, R.; Berger, S. J.; Anderson, G. A.; Rodriguez, N.; Smith, R. D. *Anal. Chem.* **2002**, *74*, 4235-4249.
12. Banks, J. F. *Chromatogr.* **1996**, *743*, 99-104.
13. Russell, W. K.; Park, Z-. Y.; Russell, D. H. *Anal. Chem.* **2001**, *73*, 2682-2685.
14. Tabb, D. L.; McDonald, W. H.; Yates, J. R. *J. Proteome Res.* **2002**, *1*, 21-26.
15. Washburn, M. P.; Wolters, D.; Yates, J. R. *Nat. Biotechnol*. **2001**, *19*, 242-247.