Journal of Chromatography B, 879 (2011) 610–614

Contents lists available at ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Increasing proteome coverage with offline RP HPLC coupled to online RP nanoLC–MS

Emine Gokce[a], Genna L. Andrews[a], Ralph A. Dean[b], David C. Muddiman[a,*]

[a] W.M. Keck FT-ICR Mass Spectrometry Laboratory, Department of Chemistry, North Carolina State University, Raleigh, NC 27695, USA
[b] Center for Integrated Fungal Research, North Carolina State University, Raleigh, NC 27695, USA

ARTICLE INFO

Article history:
Received 19 October 2010
Accepted 26 January 2011
Available online 4 February 2011

Keywords:
RP HPLC
Fractionation methods
Proteomics
Mass spectrometry

ABSTRACT

Fractionation prior to mass spectrometry is an indispensable step in proteomics. In this paper we report the success of performing offline reversed phase high pressure liquid chromatography (HPLC) fractionation on a C18 2.0 mm × 150 mm column at the peptide level with microliter per minute flow rates prior to online nano-flow reversed phase liquid chromatography mass spectrometry (nanoLC–MS) using the well-studied fungus *Saccharomyces cerevisiae*. A C18 75 μm × 150 mm column was used online and the online elution gradients for each fraction were adjusted in order to obtain well resolved separation. Comparing this method directly to only performing nanoLC–MS we observed a 61.6% increase in the number of identified proteins. At a 1% false discovery rate 1028 proteins were identified using two dimensions of RPLC versus 636 proteins identified in a single nano-flow separation. The majority of proteins identified by one dimension of nano-LC were present in the proteins identified in our two dimensional strategy. Although increasing analysis time, this non-orthogonal and facile pre-fractionation method affords a more comprehensive examination of the proteome.

1. Introduction

Mass spectrometry (MS) is the central technology for proteomic analyses. High sensitivity, rapid duty cycles, parts-per-million mass accuracy and high resolving power of hybrid mass spectrometers have made MS a valuable tool in proteomics. However, MS still has limitations in dynamic range (3–4 orders of magnitude), and detecting the entire proteome of a species is a target that has not yet been accomplished. The difficulties herein reside mostly in the wide dynamic range of proteins and the sheer complexity of the proteomes, which even the rapid duty cycles of the most recently developed mass spectrometers cannot handle. The development of improved MS platforms is still in progress and is essential to increase the proteome coverage.

The fungus *Saccharomyces cerevisiae* (*S. cerevisiae*) is commonly used as a model organism in technology and methodology evaluation studies and has to date the most extensively characterized proteome. Mann and coworkers [1] were the first in combining two dimensional polyacrylamide gel electrophoresis (2D-PAGE) fractionation [2] and a matrix assisted laser desorption ionization-time of flight mass spectrometer identifying 150 yeast proteins in their study in 1996. Twelve years later, again Mann and coworkers [3] identified 3639 proteins using 1D-PAGE [4,5] and 3987 yeast proteins using OFFGEL [6–10] both coupled to an online liquid chromatography (LC) linear trap quadrupole (LTQ)-Orbitrap mass spectrometer.

A variety of analytical strategies exists for the fractionation of complex mixtures prior to MS analysis and is either performed at the protein or peptide level, each with their own advantages and disadvantages. Aside from electrophoresis based fractionation methods (2D-PAGE and 1D-PAGE, OFFGEL and GelFree [11]), chromatography based methods have also been developed. Offline strong cation exchange (SCX) [12–14] prior to nanoLC–MS and multi dimensional protein identification technology (MudPIT) [15–17], combining SCX and reversed phase (RP) online, are both common methods used in proteomics.

RP-RP at two different pH values [18–20] has become an option in separation techniques as it was found to be identical in orthogonality to SCX-RP [21]. The orthogonality in RP-RP arises only from the difference in the pH values since the column used in both dimensions is C18 columns. In this study we performed peptide level offline RP high pressure liquid chromatography (HPLC) fractionation at micro-flow/min rates of a *S. cerevisiae* whole digest prior to RP nanoLC–MS (LC–LC–MS), using the same mobile phase and pH and compared this method to the performance of only nanoLC–MS analysis of the whole digest. The total number of proteins and protein groups identified from each analysis were compared.

* Corresponding author. Tel.: +1 919 513 0084; fax: +1 919 513 7993.
E-mail address: david_muddiman@ncsu.edu (D.C. Muddiman).

1570-0232/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2011.01.032

2. Experimental

2.1. Sample preparation

The *S. cerevisiae* strain Y15696 (BY4742; MaTα; his3D1; leu2D0; lys2D0; ura3D0; YIR034c::kanMX4), with a lys1 gene deletion purchased from EuroScarf (Frankfurt, Germany) was inoculated in liquid yeast peptone dextrose and harvested during log phase. A yeast cell pellet was collected after centrifugation of the culture at 5000 rpm for 10 min at 4 °C, washing with 50 mM Tris–HCl (Sigma–Aldrich, St. Louis, MO) buffer and followed by a second centrifugation step. By grinding the pellet with liquid nitrogen the cells were lysed and resuspended in 50 mM Tris–HCl. Cell debris was removed with a third centrifugation step at same conditions mentioned above.

Protein concentration was determined by a BCA and Bradford assay. To denature the protein, urea (Sigma–Aldrich, St. Louis, MO) was added to the yeast lysate to a final concentration of 8 M. Dithiothreitol (DTT) (Biorad, Hercules, CA) was added to a final of 5 mM and the sample was incubated for 30 min at 56 °C to reduce the protein disulfide bonds. Iodoacetamide (Sigma–Aldrich, St. Louis, MO) was added to a final of 20 mM and the sample was incubated for 30 min in the dark at room temperature for alkylation of the free thiols. DTT was added again to quench the alkylation reaction. In order to dilute the urea to 2 M, 50 mM Tris–HCl was added. Digestion was performed overnight with the addition of trypsin (Sigma–Aldrich, St. Louis, MO) at a 1:50 enzyme:protein ratio. Formic acid (FA) (Sigma–Aldrich, St. Louis, MO) was added to stop the digestion.

2.2. Offline RP HPLC

An aliquot of the yeast digest (63.5 μg), was reconstituted in 50 μl mobile phase A (98% HPLC-Grade H_2O (Burdick and Jackson, Muskegon, MI), 2% acetonitrile (ACN) (Burdick and Jackson, Muskegon, MI) and 0.2% FA). Peptide level fractionation was performed on a Shimadzu HPLC system (CBM-20A prominence communications bus module, DGU-20A_5 prominence degasser, two LC-20AD prominence pumps, CTO-20A prominence column oven at 40 °C, SPD-20A prominence ultraviolet–visible (UV–vis) detector) (Shimadzu, Columbia, MD) equipped with a 100 μl sample loop. The sample was loaded onto a 5 μm, 200 Å, Magic C18AQ, 2.0 mm × 150 mm column (Michrom Bioresources Inc., Auburn, CA). The pumps were operated at a combined flow rate of 200 μl/min. After 1 min equilibration time and a 10 min column wash at 5% mobile phase B (98% ACN, 2% H_2O and 0.2% FA) a 30 min gradient was performed from 5% to 50% B. The gradient was ramped up to 95% B over 1 min and held for 4 min to wash the column. The gradient was then ramped down to 5% B in 2 min and maintained for 3 min. A Gilson FC 203B fraction collector (Gilson Inc., Middleton, WI) was used to collect fractions every 3 min during the 30 min gradient for a total of 10 fractions. Protein concentrations of the fractions were calculated using the UV–vis spectrum and the Scope's method [22] (Table 1). The fractions were dried down and reconstituted in mobile phase A to a concentration of approximately 100 ng/μl.

2.3. NanoLC–MS

The nanoLC prior to MS/MS was performed on a nanoLC-1D system from Eksigent (Dublin, CA) at room temperature. A 75 μm i.d. IntegraFrit capillary (New Objective, Woburn, MA) was packed in house to 5 cm with Magic C18AQ packing material (Michrom BioReasources, Auburn, CA) and operated as a trap. A 75 μm i.d. PicoFrit capillary column (New Objective, Woburn, MA) was packed 15 cm with the same packing material. Separations were carried out using a continuous, vented column configuration as previously described by our group [23]. A 2 μl (200 ng) sample was injected into the 10 μl loop and loaded onto the trap column with approximately 10 column washes prior to analytical separation. For peptide separation on the analytical column, the flow rate was set to 350 nl/min. A 5 min column wash was performed at 2% B followed by a 1 h linear gradient. The gradient was ramped up to 90% B in 1 min and maintained for 10 min. Two minutes were required to establish 2% B and this was maintained for 2 min. Fraction 1 was subjected to a 1 h gradient of 5–9% B and the gradients for the following fractions were determined based on the elution end time of the previous fraction. All gradients consisted of a 4% increase in B over the 1 h time period. The gradient for each fraction is listed in Table 2. In addition, a whole yeast digest sample (200 ng) was analyzed by nanoLC–MS/MS and a 1 h gradient of 10–40% B was applied. Three technical replicates of the whole yeast sample were run.

Table 1
Calculated peptide concentrations per fraction.

Fraction #	Conc. [μg/ml]
1	3.23
2	3.23
3	4.84
4	7.26
5	8.87
6	9.68
7	8.87
8	7.26
9	4.84
10	3.23

MS analysis was performed using a hybrid LTQ-Orbitrap MS (Thermo Fisher Scientific, Bremen, Germany). The automatic gain control (AGC) limit for the Fourier Transform MS (FTMS) was set to 1×10^6 and the maximum injection time was 500 ms. For the ion trap the AGC limit was 8×10^3 and the maximum injection time was 80 ms. The resolving power was set to $30{,}000_{fwhm}$ at m/z 400 and 8 data dependent MS/MS events were performed for ions with charge states ≥+2. Singly charged ions were rejected for MS/MS. Dynamic exclusion was enabled, and ions selected for MS/MS interrogation were excluded for 180 s. The normalized collision energy was 35% and lock mass calibration using polydimethylcyclosiloxane present in ambient laboratory air (m/z 445.120025) was enabled. External calibration was also performed following manufacturer instructions and using manufacturer's calibration mix.

2.4. Data analysis

Data analysis was performed by converting the .raw files into .mgf files through MASCOT Distiller version 2.3.01 (Matrix Science Inc., Boston, MA) and searching the .mgf files against the target reverse yeast ORF database (orf_trans_all.fasta.gz) from the Stanford University web page: http://www.yeastgenome.org in MASCOT. Parameters used in MASCOT were ±5 ppm peptide ion tolerance, ±0.6 Da MS/MS fragment ion tolerance and 2 allowed missed cleav-

Table 2
One hour gradients of the fractions.

Fraction #	Gradient [%B]
1	5–9
2	8–12
3	10–14
4	13–17
5	15–19
6	18–22
7	21–25
8	25–29
9	29–33
10	33–37

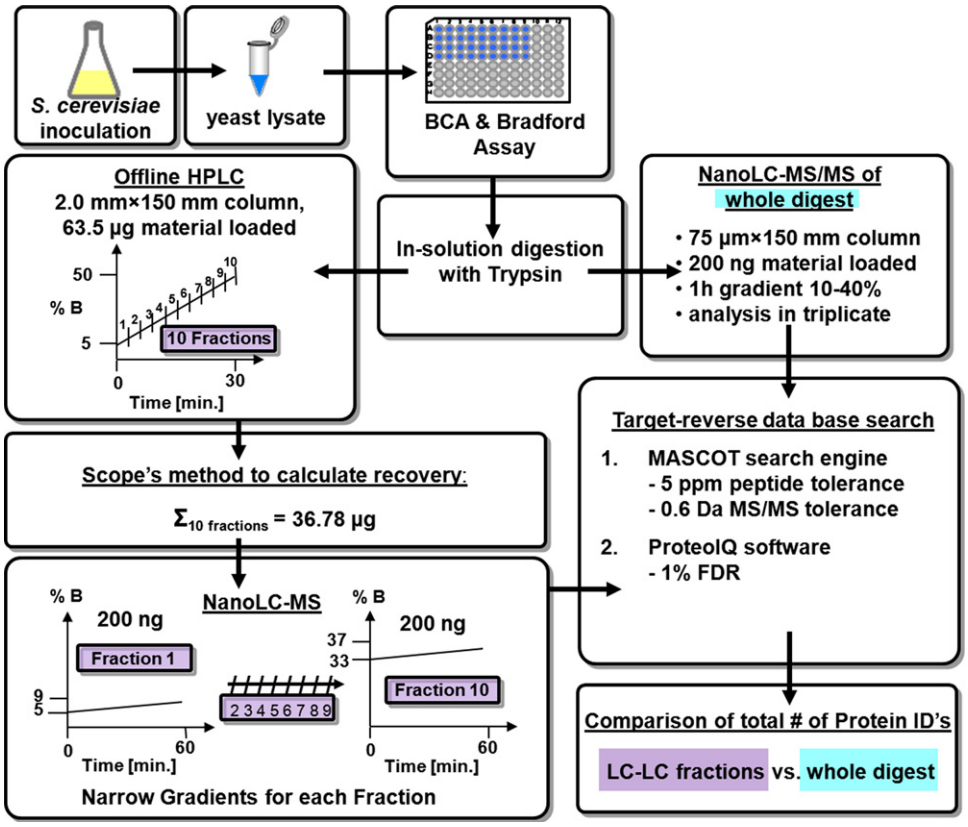

Fig. 1. Experimental workflow. *S. cerevisiae* was inoculated, harvested and lysed. After a tryptic digestion, 200 ng of the digest were analyzed in triplicate by nanoLC–MS. A sample containing 63.5 µg of yeast peptides was loaded onto the offline HPLC and 10 fractions were collected; narrower gradients were applied to the fractions on the nanoLC–MS. The data was searched in MASCOT and combined for comparative analysis in ProteoIQ.

ages. A fixed modification was set to be carbamidomethylation of cysteine and variable modifications were oxidation of methionine as well as deamidation of glutamine and asparagine. The output of the database search in the form of .dat files was analyzed in ProteoIQ version 2.1.01_SILAC_beta08 (BioInquire, Athens, GA) at 1% FDR [24].

3. Results and discussion

The experimental workflow is diagramed in Fig. 1. The *S. cerevisiae* strain was grown in liquid culture and harvested during log phase. BCA and Bradford assays were performed to estimate the concentration of the lysate and the sample was digested in solution using trypsin. An aliquot of the digest was run directly on the nanoLC–MS and another was collected as 10 fractions during offline RP HPLC prior to nanoLC–MS.

The *S. cerevisiae* sample fractionated offline was detected and approximately quantified by UV–vis at 205 nm (see Fig. 2A and Table 1). The total recovery was around 60% and sufficient to perform further nanoLC–MS analysis with the fractions. The total ion chromatogram (TIC) for the nanoLC–MS analysis, increasing only 4% B over a 60 min linear gradient, is shown in Fig. 2B. The TIC in Fig. 2C emphasizes the complexity of the *S. cerevisiae* whole sample digest applying a 60 min gradient for peptide elution of 10–40% B. By applying narrow gradients to the already fractionated *S. cerevisiae* digest the need for an additional level of separation is demonstrated through the chromatogram and more MS/MS spectra for that gradient range are obtained. Reducing the complexity of the sample in addition to increasing the total amount of material loaded yields the potential for more protein identifications, in particular, lower abundant species, which are frequently suppressed in data-dependent LC–MS experiments of complex samples. As an example, the protein: RPL31B SGDID:S000004398, Chr XII from 931754-931698, 931348-931064, reverse complement was identified with one peptide in the whole digest sample (21.05% sequence coverage), whereas in the LC–LC approach, with more opportunity for detection, the same protein was identified with 6 unique peptides (55.26% sequence coverage). See Supplemental Tables 1 and 2 for a full list of proteins identified by each technique as well as to observe other examples affording increased proteome coverage by the LC–LC approach.

The peptide overlaps between the fractions are described in Table 3. As expected early fractions have no peptide overlap with late fractions offering evidence that the separation was carried out successfully. The large overlap of 283 peptides between fractions 6 and 7 is due to the number of peptides eluting at the same time and thus bleeding out. Carryover is also attributable to peptides that are eluting at the time boundary as the fraction collector switches between fractions.

The Venn diagram in Fig. 3 presenting the results, shows that evaluating the 10 offline RP-LC fractions of the *S. cerevisiae* digest in narrow gradients by nanoLC–MS increased protein identification by 61.6% versus direct analysis of the whole digest by nanoLC–MS. In the LC–LC–MS experiment 1028 yeast proteins were identified at 1% FDR compared to 636 proteins resultant of nanoLC–MS analysis. The majority of proteins identified from the whole digest sample analyzed by nanoLC–MS were also identified within the total population of protein identifications from the LC–LC–MS experiment.

The analysis time for the fractionation must be considered as analyzing a sample in a timely manner is also very important. The time required for this fractionation is about an hour for the initial offline RP separation. Influenced by the number of

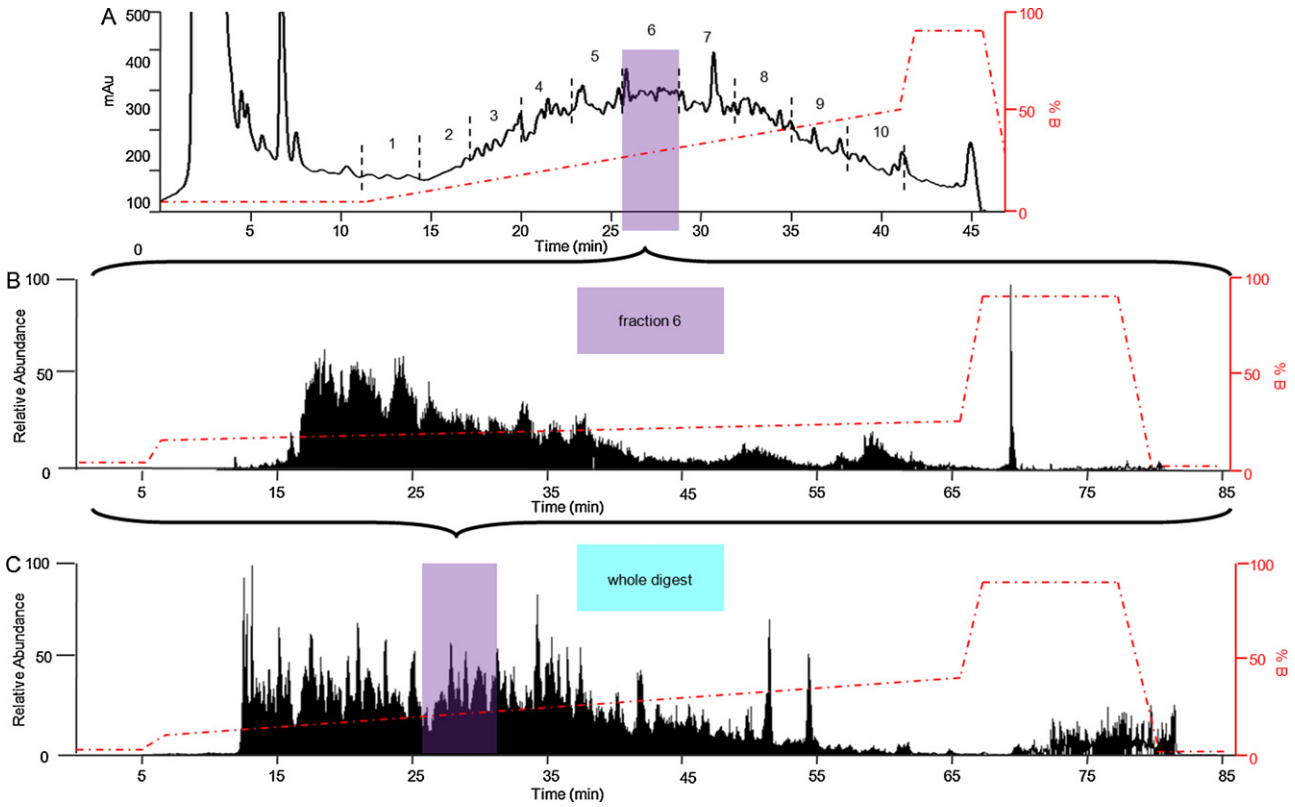

Fig. 2. Representative data comparing the complexity in chromatography. (A) UV–vis spectrum at 205 nm for the fractionation on the offline HPLC. Fractions 1–10 were collected during the time periods marked with dashed black lines. The gradient is implied with the dashed line in red. As an example fraction 6 was collected at 25–29% B and is highlighted in violet. (B) TIC of fraction 6. (C) TIC of whole digest. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Table 3
Peptide overlaps between fractions.

fraction#	1	2	3	4	5	6	7	8	9	10
1	247	29	6	10	7	1	0	1	0	0
2	29	257	25	10	12	4	2	1	0	0
3	6	25	461	43	30	17	7	1	1	1
4	10	10	43	631	137	96	42	10	4	3
5	7	12	30	137	1063	227	123	37	10	3
6	1	4	17	96	227	1227	283	108	28	15
7	0	2	7	42	123	283	1164	246	89	30
8	1	1	1	10	37	108	246	977	196	105
9	0	0	1	4	10	28	89	196	684	169
10	0	0	1	3	3	15	30	105	169	461

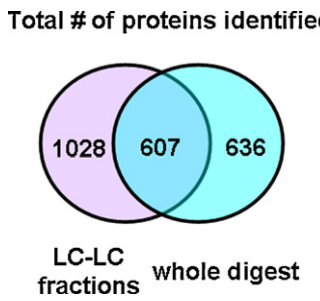

Fig. 3. Venn diagram for comparison of proteins identified with each method. The majority of proteins identified in the whole digest (636) are within the identified proteins from the LC–LC fractions (1028).

fractions collected, 10 fractions here, nanoLC–MS analysis of all fractions requires 10 h. Although LC–LC–MS analysis time compared to nanoLC–MS analysis time is a considerable increase, 1D separations are known to be insufficient for complex mixtures.

de Godoy et al. [3] analyzed 24 OFFGEL-fractions using 2 h gradients on the nano-LC and identified 3987 yeast proteins (83 proteins/h). In comparison, in our LC–LC–MS study 10 fractions yielded 1028 proteins using a non-orthogonal method and only 1 h gradients (102 proteins/h). The same database and the MASCOT search engine were used in both studies. The only difference in the bioinformatic platform was the analysis software; ProteoIQ employed for our LC–LC–MS analysis whereas de Godoy et al. used MaxQuant. Considering the rate of protein identification per hour, *vide supra*, it is demonstrated that LC–LC–MS has an acceptable analysis time in order to continue the pursuit of characterizing an entire proteome through MS analysis.

4. Conclusions

Fractionation prior to LC–MS is essential in order to overcome the dynamic range and complexity of biological samples. We have shown that offline HPLC fractionation at the peptide level coupled to online nanoLC–MS, a non-orthogonal method, is a simple and effective way of reducing the sample complexity and thus increasing the number of proteins identified. The increase of instrument time from 3 h (triplicate of whole digest) to 10 h (10 fractions), is a disadvantage (212 protein IDs/h versus 102 protein IDs/h). However, fractionation allows a greater amount of sample to be loaded onto the nano LC-column (2 μg versus 200 ng) and thus yields a 61.6% increase in number of proteins identified compared to employing nano-LC–MS without pre-fractionation.

Acknowledgements

The authors would like to thank Timothy Collier for his thoughtful inputs to the manuscript and assistance in preliminary analyses. We thank Yeon Yee Oh for inoculating and harvesting the yeast. We thank Professor Ron Orlando (University of Georgia) for suggesting this general idea to us. We also thank the National Science Foundation (Grant Number MCB-0918611), the W.M. Keck Foundation, and North Carolina State University for financial support. Genna L. Andrews acknowledges the financial support of the North Carolina State University Molecular Biotechnology Training Program and the National Institutes of Health (Grant 5T32GM00-8776-08).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jchromb.2011.01.032.

References

[1] A. Shevchenko, O.N. Jensen, A.V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, H. Boucherie, M. Mann, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 14440.
[2] J. Klose, U. Kobalz, Electrophoresis 16 (1995) 1034.
[3] L.M.F. de Godoy, J.V. Olsen, J. Cox, M.L. Nielsen, N.C. Hubner, F. Frohlich, T.C. Walther, M. Mann, Nature 455 (2008) 1251.
[4] C.L. Lee, H.H. Hsiao, C.W. Lin, S.P. Wu, S.Y. Huang, C.Y. Wu, A.H.J. Wang, K.H. Khoo, Proteomics 3 (2003) 2472.
[5] S.A. Beausoleil, M. Jedrychowski, D. Schwartz, J.E. Elias, J. Villen, J.X. Li, M.A. Cohn, L.C. Cantley, S.P. Gygi, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 12130.
[6] P.E. Michel, F. Reymond, I.L. Arnaud, J. Josserand, H.H. Girault, J.S. Rossier, Electrophoresis 24 (2003) 3.
[7] M. Heller, P.E. Michel, P. Morier, D. Crettaz, C. Wenz, J.D. Tissot, F. Reymond, J.S. Rossier, Electrophoresis 26 (2005) 1174.
[8] M. Heller, M.L. Ye, P.E. Michel, P. Morier, D. Stalder, M.A. Junger, R. Aebersold, F.R. Reymond, J.S. Rossier, J. Proteome Res. 4 (2005) 2273.
[9] P.E. Michel, D. Crettaz, P. Morier, M. Heller, D. Gallot, J.D. Tissot, F. Reymond, J.S. Rossier, Electrophoresis 27 (2006) 1169.
[10] P. Horth, C.A. Miller, T. Preckel, C. Wenz, Mol. Cell. Proteomics 5 (2006) 1968.
[11] J.C. Tran, A.A. Doucette, Anal. Chem. 81 (2009) 6201.
[12] J.M. Peng, J.E. Elias, C.C. Thoreen, L.J. Licklider, S.P. Gygi, J. Proteome Res. 2 (2003) 43.
[13] M. Vollmer, P. Horth, E. Nagele, Anal. Chem. 76 (2004) 5180.
[14] Y.F. Shen, J.M. Jacobs, D.G. Camp, R.H. Fang, R.J. Moore, R.D. Smith, W.Z. Xiao, R.W. Davis, R.G. Tompkins, Anal. Chem. 76 (2004) 1134.
[15] D.A. Wolters, M.P. Washburn, J.R. Yates, Anal. Chem. 73 (2001) 5683.
[16] M.P. Washburn, D. Wolters, J.R. Yates, Nat. Biotechnol. 19 (2001) 242.
[17] J. Dai, C.H. Shieh, Q.H. Sheng, H. Zhou, R. Zeng, Anal. Chem. 77 (2005) 5793.
[18] H. Toll, H. Oberacher, R. Swart, C.G. Huber, ×matogr. A 1079 (2005) 274.
[19] I. Francois, D. Cabooter, K. Sandra, F. Lynen, G. Desmet, P. Sandra, J. Sep. Sci. 32 (2009) 1137.
[20] L. Mondello, P. Donato, F. Cacciola, C. Fanali, P. Dugo, J. Sep. Sci. 33 (2010) 1454.
[21] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Anal. Chem. 77 (2005) 6426.
[22] R.K. Scopes, Anal. Biochem. 59 (1974) 277.
[23] G.L. Andrews, C.M. Shuford, J.C. Burnett, A.M. Hawkridge, D.C. Muddiman, J. Chromatogr. B 877 (2009) 948.
[24] Y. Benjamini, Y. Hochberg, J. R. Stat. Soc. Ser. B: Methodol. 57 (1995) 289.