

Development of a two-dimensional protein–peptide separation protocol for comprehensive proteome measurements

George M. Janini, Thomas P. Conrads, Timothy D. Veenstra, Haleem J. Issaq*

Analytical Chemistry Laboratory, Separations Technology Group, SAIC-Frederick, Inc., National Cancer Institute at Frederick, P.O. Box B, Frederick, MD 21702, USA

Abstract

We have developed an effective two-dimensional fractionation protocol of complex proteome mixtures that extends the ability to conduct more comprehensive proteome measurements. A sample containing intact proteins extracted from *Saccharomyces cerevisiae* was fractionated by liquid phase isoelectric focusing, followed by tryptic digestion and solid-phase extraction (SPE) clean-up and reversed-phase liquid chromatography–electrospray ionization tandem mass spectrometry (LC–MS–MS) of the resultant peptides. The clean-up step is designed to desalt the fractions and rid them of urea and ampholytes prior to analysis by LC–MS–MS. Fifty milligrams of protein were separated into 20 fractions by liquid-phase isoelectric focusing, spanning a pH range of 3–10. The effectiveness of the removal of ampholytes was monitored by capillary zone electrophoresis and LC–MS–MS. The ability to analyze all of the 20 fractions without any noticeable decrease in the separation efficiency demonstrates the overall effectiveness of the SPE clean-up step. The results show that the separation strategy is effective for high throughput characterization of proteins from complex proteomic mixtures.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Proteomes; Proteins; Peptides

1. Introduction

The proteome of any cell, tissue, or biological fluid is a complex mixture of proteins that span a wide range of size, relative abundance, acidity/basicity, and hydrophobicity. For example, a human cell type may express up to 20 000 proteins at any time [1] with a predicted dynamic range of up to five orders of magnitude [2]. Though state-of-the-art mass spectrometry (MS) provides an invaluable analytical tool, the ability to detect and reliably

characterize large numbers of components depends greatly on the ability to resolve each species distinctly prior to MS detection. Many attempts for sample fractionation have been made using different electrophoretic, chromatographic and a combination of hyphenated techniques, both off-line and on-line. Chronologically, the most widely used method for fractionation of the complex protein mixtures prior to MS analysis is two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) [3–10], followed by enzymatic digestion of the separated protein spots. 2D-PAGE has unique advantages, but it also has limitations currently being redressed by the development of more conventional “non-gel”-based fractionation methods [9,10].

*Corresponding author. Tel.: +1-301-846-7186; fax: +1-301-846-6037.

E-mail address: issaqh@mail.ncifcrf.gov (H.J. Issaq).

One such separation technique, high-performance liquid chromatography–tandem mass spectrometry (LC–MS–MS), has been advanced as a viable alternative to 2D-PAGE. Indeed, a variety of innovative one-dimensional (1D) and 2D LC–MS schemes have been investigated in the last decade [11–20]. One advantage of LC methods over gel-based methods is that the sample remains in solution throughout the sample preparation and subsequent analysis; hence losses associated with poor recovery from gels are eliminated. While LC–MS–MS has proven to be of great utility in the analysis of protein mixtures, complex proteomes overwhelm separation capabilities of any single 1D separation technique. To overcome this limitation, coupled, yet orthogonal 2D separation methods are being developed and a variety of different separation techniques, such as LC–capillary electrophoresis (CE) and isoelectric focusing (IEF)–LC, are being combined for the fractionation and separation of complex proteomes [16–20].

Preparative scale liquid-phase IEF has been incorporated into several protein fractionation strategies [18,21–26]. Lubman and co-workers [18,21] developed a 2D method whereby the intact proteins were separated in the first dimension by IEF based on their isoelectric point (*pI*) using a Rotofor cell and in the second dimension based on their hydrophobicity by nonporous reversed-phase LC. Fractions were collected from the LC eluent and subjected to proteolytic digestion and matrix-assisted laser desorption–ionization MALDI-TOF–MS for protein identification. A major advantage of this strategy is that the proteins collected from the reversed-phase LC eluent remain in the liquid phase throughout the 2D process. Also the amount of a target protein is, potentially, 50 times higher than that obtainable from 2D-PAGE [21]. One disadvantage of this strategy is that the fractions collected from the Rotofor cell contain up to 8 M urea, which presents a problem for downstream LC–MS–MS analysis. The Rotofor fractions also contain 2.5% (w/v) of ampholytes, which are retained by the reversed-phase column and obstruct the effective detection of proteins and peptides in LC–MS-based analyses. In addition, many intact proteins are not soluble under the solvent conditions necessary for effective reversed-phase LC separations, and the loss of an intact protein leaves no remnant for its identification.

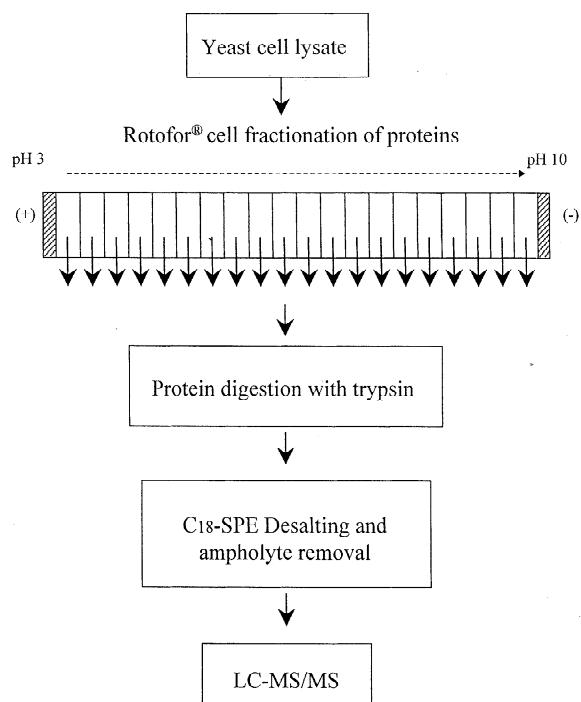


Fig. 1. Flow chart of the protein separation strategy.

In this work we present an effective sample fractionation strategy (Fig. 1) that incorporates liquid-phase prefractionation, followed by tryptic digestion, sample clean-up using solid-phase extraction (SPE) and LC–MS–MS analysis. Preparative liquid-phase IEF was used to fractionate the proteins extracted from a yeast cell lysate. Each fraction was subsequently digested with trypsin, desalted by SPE to remove urea and ampholytes, and directly analyzed using LC–MS–MS. The SPE cleanup of the Rotofor fractions prior to LC–MS–MS analysis increases the robustness of this analysis by eliminating the exposure of the reversed-phase column packing material to urea and ampholytes. In addition, the implementation of a tryptic digestion step prior to reversed-phase LC–MS–MS analysis increases the overall protein coverage since peptides are less likely to precipitate than intact proteins under these separation conditions. Even if some peptides are lost under reversed-phase LC separation conditions, it is highly likely that at least a few peptides from each protein will remain available and whose presence would allow identification of their protein of origin.

2. Experimental

2.1. Sample preparation

Saccharomyces cerevisiae (strain BJ5460) was grown in rich medium containing 1% (w/v) yeast extract, 2% (w/v) bacto-peptone and 2% (w/v) glucose. Cultures were inoculated at an $OD_{600} < 0.1$, incubated at 30 °C with shaking, and harvested at an $OD_{600} \sim 1.0$. Cell pellets were collected by centrifugation at 2000 g and stored at –20 °C until needed. Prior to lysis, cells were washed three times with 50 mM NH_4HCO_3 , pH 8.2, and resuspended to ~400 mg wet cells/ml in 50 mM NH_4HCO_3 , pH 8.2 (~4 g cells used for this study). Cells were lysed using a Mini Bead-Beater (Biospec Products, Bartlesville, OK, USA) operating at 6000 rpm for 1 min followed by 1 min incubation on ice. This procedure was repeated twice and the soluble cell extract was transferred to a new tube. The glass beads were rinsed with 50 mM NH_4HCO_3 , pH 8.2, and the wash was added to the reserved extract. Cell debris was removed from this combined extract by centrifugation for 30 min at 20000 g. The supernatant (18 ml) was collected and protein concentration (5 mg/ml) was determined using a bicinchoninic acid assay (Pierce, Rockford, IL, USA).

2.2. Protein prefractionation by liquid-phase isoelectric focusing

A preparative-scale isoelectric focusing device (Rotofor, Bio-Rad Laboratories, Hercules, CA, USA) was used for sample prefractionation as per manufac-

turer's instructions with a few modifications. Briefly, the focusing chamber was filled with distilled water and run for 5 min at 5 W constant power to remove residual ionic contaminants from the membrane core of the focusing chamber and the ion-exchange membranes that separate the chamber from the anolyte and catholyte compartments. Fifty ml of a 1 mg/ml protein solution were loaded in the focusing chamber. The final concentrations of the buffer constituents were 1 mg/ml total yeast protein, 2.5% Bio-Lyte 3-10 ampholytes (Bio-Rad), 2 M urea, and 2 mM dithiothreitol. The focusing chamber was cooled to 10 °C by a circulating refrigerated water bath and focusing was conducted at 12 W for 5 h. Twenty fractions containing 2.5 ml of protein solution each were collected. All of the fractions were clear except fractions 11–14 (pH range of 5–6), which were turbid. Establishment of the pH gradient in the focusing cell was confirmed by measuring the pH of each fraction using pH indicator paper (pHydrion). The proteins in the Rotofor fractions were profiled by SDS-PAGE using 4–15% Tris-glycine-SDS precast polyacrylamide slab gels (Bio-Rad), (Fig. 2). The total protein in each fraction was estimated based on visual inspection of the Coomassie-stained slab gels.

2.3. Enzymatic digestion of protein fractions

One ml of each fraction was transferred to a fresh microvial and NH_4HCO_3 was added to each to a final concentration of 100 mM. The proteins were denatured prior to tryptic digestion by boiling for 5 min. Sequencing grade-modified porcine trypsin

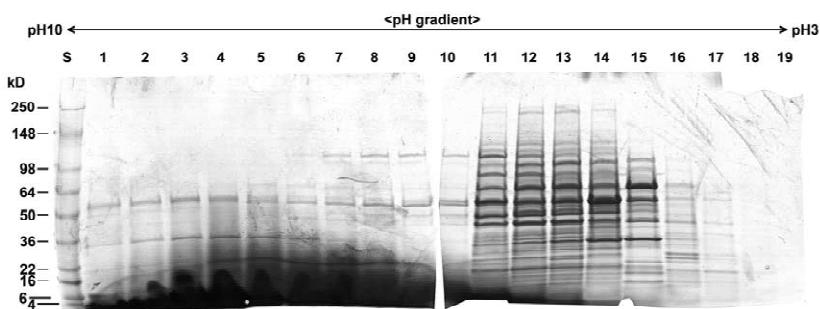


Fig. 2. SDS-PAGE image of the 20 Rotofor fractions after Coomassie blue staining. Lanes 1–19 correspond to the Rotofor fractions with lane 1 being the most basic. Lane S contains the molecular mass markers. Fifty μ l of each Rotofor fraction were mixed with 50 μ l of Laemmli sample buffer, and incubated at 95 °C for 5 min. Thirty μ l of these samples were loaded into each well.

(Promega, Madison, WI, USA) was added to each fraction to give an approximate protein–trypsin ratio of approximately 50:1 (w/w). The fractions were incubated at 37 °C for 20 h, and lyophilized to dryness. The lyophilized fractions were resuspended in 500 µl of 10 mM HCl. All fractions were clear after the digestion step including fractions 11–14, which were cloudy prior to digestion.

2.4. Ampholyte removal and peptide desalting

The resuspended fractions were desalted by solid-phase extraction using single-use bonded phase octadecyl (C₁₈) cartridges (Bond Elut 6 ml/500 mg, Varian, Harbor City, CA, USA). Each cartridge was preconditioned with 2 volumes (12 ml) of methanol, followed by washing with 2 volumes of 0.1% trifluoroacetic acid (TFA) in water. The sample was applied and the cartridge was washed again with 1 volume H₂O, 2 volumes 0.1% (w/v) NH₄OH, 1 volume H₂O, and 2 volumes 0.1% (v/v) TFA in water. The peptides were eluted from the cartridge with 4 ml of acetonitrile–H₂O–TFA (80:19.9:0.1, v/v/v) and collected in 4.5-ml glass vials. The volume of each sample was reduced to 1 ml by evaporation under argon atmosphere at 37 °C, and was further reduced to 100–500 µl by vacuum centrifugation.

2.5. Capillary LC–MS–MS analysis

The peptide fractions were separated using an in-house prepared 10-cm long, 75-µm I.D. capillary column (Polymicro Technologies, Phoenix, AZ, USA) packed with 5 µm Jupiter C₁₈ stationary phase (Phenomenex, Torrence, CA). The mobile phase consisted of 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The peptides were eluted using an Agilent 1100 microcapillary HPLC system operating at a flow-rate of 500 nl/min. One µl of sample was loaded onto the reversed-phase capillary LC column, and the peptides eluted using the following gradient: 95% A for 20 min followed by a linear gradient to 85% B over 60 min and then held at 85% B for 20 min. The column was re-equilibrated with 95% A prior to the next injection. Tandem MS analysis was performed using an ion trap mass spectrometer (LCQ-DECA XP, Finnigan

MAT, San Jose, CA, USA) equipped with a nanoelectrospray ionization (ESI) source operating under the following instrument conditions: spray voltage of 2.0 kV; capillary temperature of 180 °C; capillary voltage of 10 V; tube lens offset of 10 V. To identify the eluting peptides, the mass spectrometer was operated in a data-dependent tandem MS mode in which a full-scan mass spectrum was followed by three tandem MS scans. The molecular ions detected in each full-scan spectrum were dynamically selected for collision-induced dissociation based on their intensity in the preceding MS scan. The normalized collision energy was set to 38%. Peptides were identified using SEQUEST (ThermoFinnigan, Torrence, CA) and the *Saccharomyces cerevisiae* database included with the program. The results of peptide and subsequent protein identification will be presented in a future publication.

3. Results and discussions

The objective of this work is to develop an effective multi-dimensional fractionation and separation procedure that can be combined with MS to identify proteins within complex mixtures. The challenges in developing such a system is to maximize the solubility of the proteins and ultimately deliver the proteins to the mass spectrometer in peptide form, to capitalize on the instrument's ability to effectively identify peptides by tandem MS. We chose to perform the first-dimensional fractionation by preparative IEF since the proteins are kept in solution through this process, and a large quantity of sample can be initially loaded compared to other fractionating methods. While Rotofor has been criticized for poor focusing [27] related to instrumental factors such as cathodic and diffusional drift [23,24] there are also inherent protein properties, such as isoelectric microheterogeneity and protein–protein interactions, that can result in the same protein being present in different compartments. While there is a certain degree of overlap among neighboring compartments, as shown in Fig. 2, the intensities of equidistant bands seem to follow a Gaussian distribution pattern, where different proteins peak in different Rotofor compartments. This pattern is more evident at the low and high pH regions where the

proteins are less concentrated and more tightly focused at their respective *pI* values. Protein loss is another problem that has been reported in Rotofor fractionation [21]. This loss can be minimized, however, by adding solubilizing agents, such as urea, to the sample.

In this work, 50 mg of protein from the total cell lysate of *Saccharomyces cerevisiae* were loaded in the Rotofor focusing chamber, and proteins separated into 20 fractions, each at a different pH spanning the pH range from 3 to 10. Despite the addition of urea to the focusing buffer, some cloudiness was observed in fractions 11–14 (pH range 5–6) where the concentration of proteins was relatively high. Precipitation problems can be avoided by starting with a smaller sample; however, we elected to apply a large sample in order to improve the chances of identifying low abundance proteins. After tryptic digestion, however, all fractions including 11–14 were completely solubilized. While the Rotofor fractionation–LC–MS–MS procedure described above is similar to that adopted by other groups [18,21,22], the protocol described herein differs in subtle, yet key areas. In the previously described methods, both separation dimensions, IEF and reversed-phase LC, were performed on the intact proteins, which were, later, tryptically digested prior to MALDI-TOF analysis. In such an approach the hydrophobic proteins may stick to the C_{18} matrix by hydrophobic interaction. In addition, when using reversed-phase LC and SPE, small and highly charged proteins and peptides are lost in the initial sample application step, and highly hydrophobic proteins and peptides are not totally eluted from the column packing. Thus if the proteins are subjected to these procedures some proteins are completely lost. On the other hand, if the peptides from protein digests are subjected to these procedures, some peptides from any given protein may be lost, but the rest will be available for further separation and delivery to the LC–MS–MS system. In our approach, the intact proteins were digested prior to reversed-phased LC separation. By incorporating the digestion step at this point in the fractionation, the identification of the resultant peptides can be accomplished through easily automated LC–MS–MS.

After trypsin digestion, the fractions were lyophilized, and reconstituted in 500 μ l of 10 mM HCl.

Prior to reversed-phase LC–MS–MS analysis it is necessary to remove contaminants such as urea and ampholytes that negatively impact the separation and MS analysis of the peptides. The presence of urea will result in increasingly poor chromatographic separation and shorten the useful lifetime of a reversed-phase C_{18} capillary column. In addition, urea can clog the nanoelectrospray tip effectively preventing the column eluent from entering the mass spectrometer. Ampholytes, which are charged molecules, can bind to the reversed-phased matrix and decrease the available binding sites on the column for the peptides of interest. In addition, as they elute from the column into the mass spectrometer, the ampholytes will obscure low-molecular mass peptides and occupy space within the ion-trap, decreasing the effective number of peptide ions that can be detected at any point during MS analysis.

To remove both urea and ampholytes, a SPE clean-up step using a C_{18} matrix was incorporated after the tryptic digestion of the Rotofor fractions. The SPE clean-up step is used to buffer exchange and concentrate the peptides, but, more importantly, this step is necessary to remove the ampholytes and urea that, if present in the sample submitted to LC–MS–MS analysis, will foul the column and clog the nanospray tip and the ion transfer capillary.

The effectiveness of the SPE method to remove

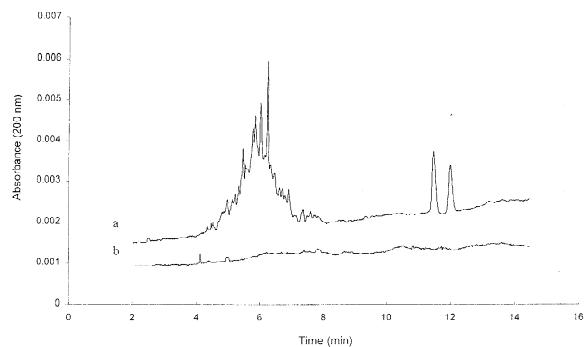


Fig. 3. Capillary zone electrophoresis profiles of Bio-Lyte 3-10 ampholytes (a) before and (b) after SPE treatment. Instrument, Beckman model P/ACE system 5500; column, 10% linear polyacrylamide-coated fused-silica; column dimensions, 37 cm (effective length 30 cm) \times 50- μ m I.D.; voltage, 12 kV; current, 25 μ A; buffer, 50 mM phosphoric acid adjusted to pH 2.5 with triethylamine; temperature, 22 °C; detection, UV at 200 nm; sample for trace a, 0.25%; injection, 5 s at 0.5 p.s.i.

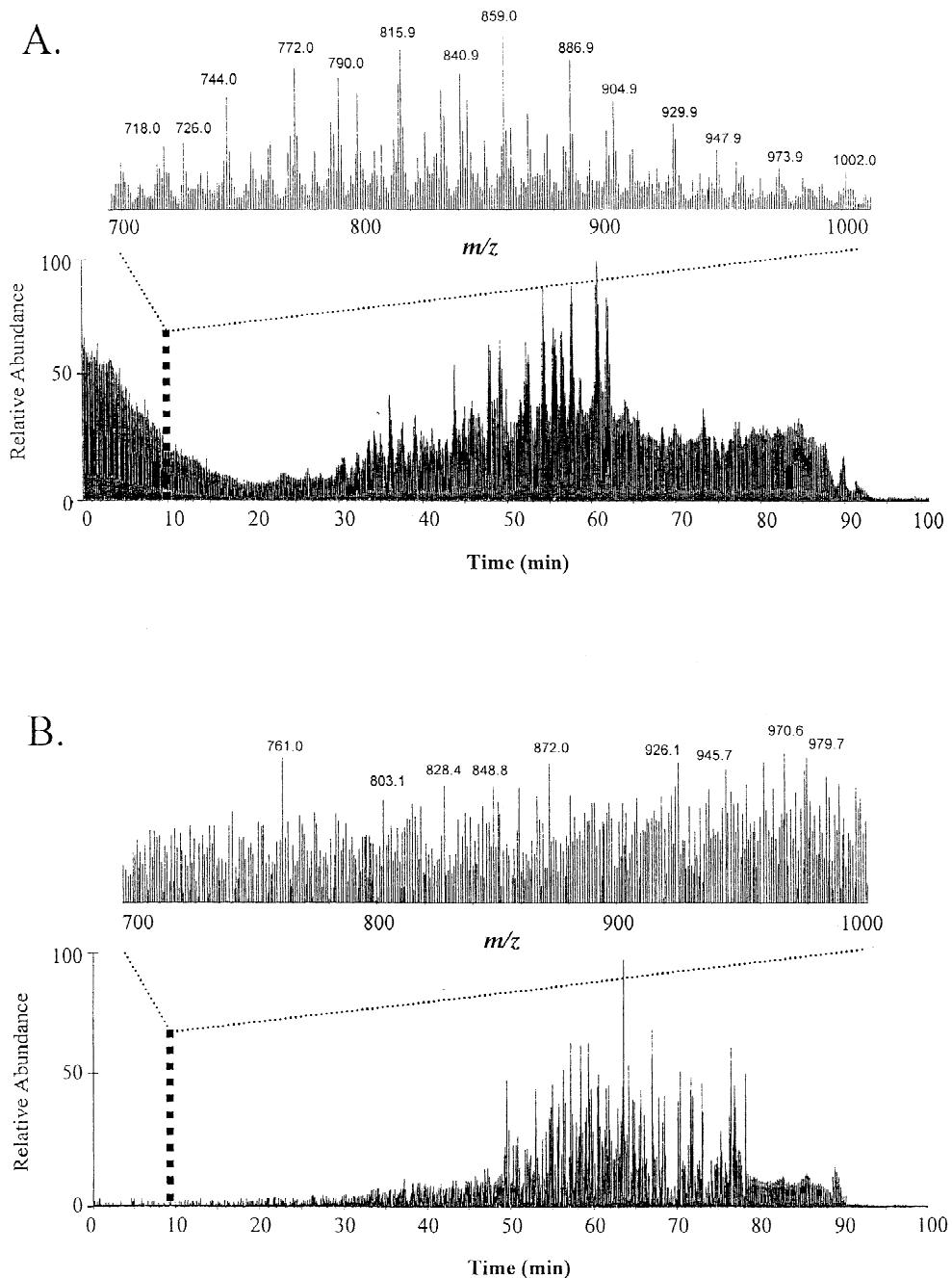


Fig. 4. Base-peak mass spectrometry chromatograms of Rotofor fraction number 4, (A) before and (B) after solid-phase extraction treatment. In the chromatogram of the sample prior to clean-up a significant contribution from ampholytes is observed, however, after clean-up these peaks are significantly reduced in intensity.

ampholytes was monitored by capillary zone electrophoresis (CZE) and MS. For CZE analysis, 1 ml of a 2.5% (w/v) solution of ampholytes was prepared and

10 μ l of this solution was diluted to 100 μ l with H₂O and analyzed by CZE. The remainder of the solution was applied to an SPE cartridge and treated

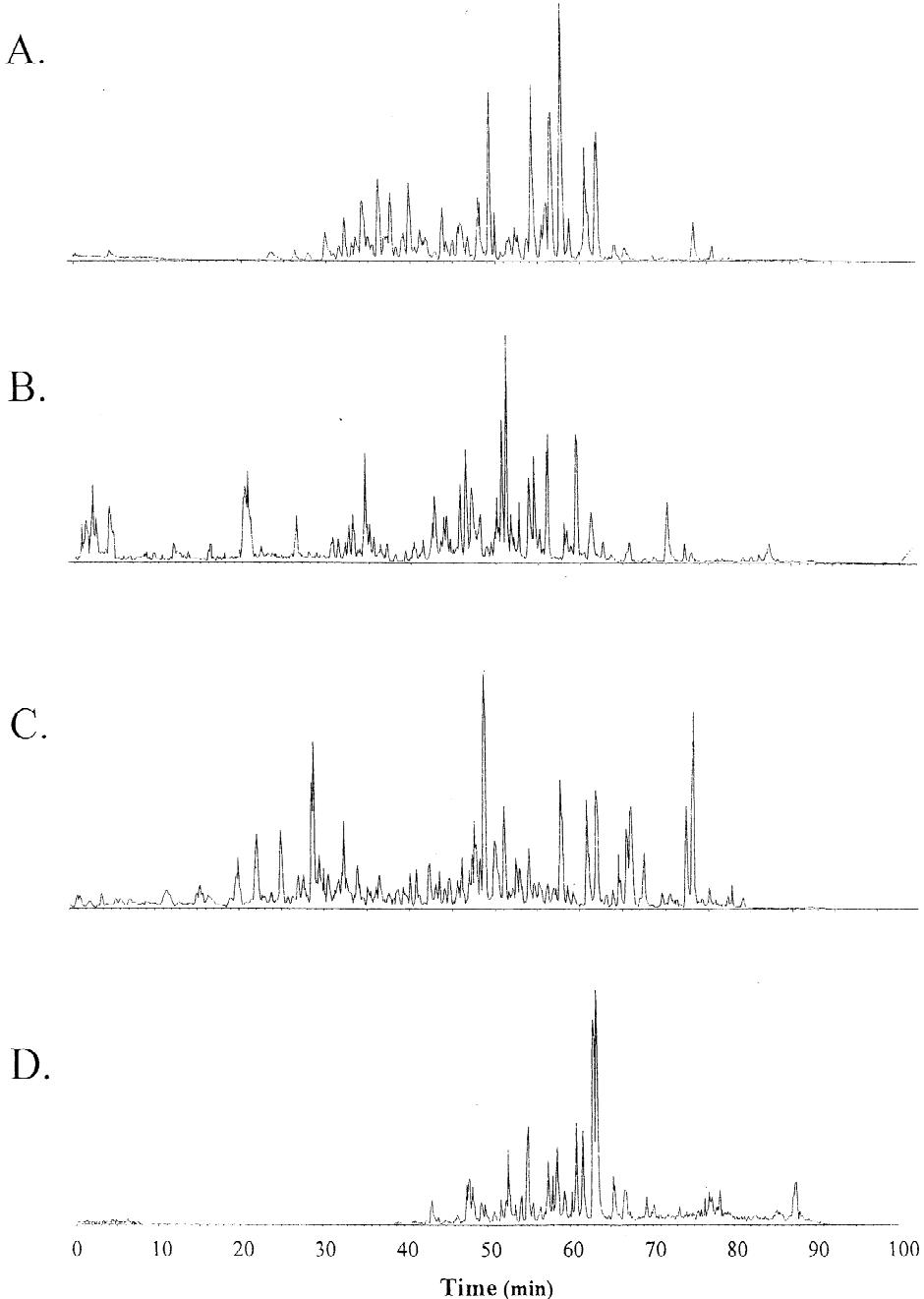


Fig. 5. Representative base-peak chromatograms from various Rotofor fractions. A total of over 50 fractions were analyzed using a single capillary column after SPE clean-up.

as described in the experimental section. Initially only acid wash was applied, but later, it was found that this was not sufficient to effectively rid the sample of ampholytes, especially from the basic fractions. The eluant from the SPE cartridge was concentrated down to 1 ml, and 10 μ l of it was diluted to 100 μ l and analyzed by CZE. The CZE profiles of Bio-Lyte 3–10 ampholytes before and after SPE clean-up are shown in Fig. 3. The presence of ampholytes is clearly seen in the sample prior to SPE clean-up. Signals related to the ampholytes are absent in Fig. 3b, demonstrating the effectiveness of the clean-up step.

The effectiveness of the SPE method to remove urea and ampholytes from a complex peptide mixture was also evaluated by analyzing a single Rotofor fraction by MS prior to and after SPE clean-up. The total ion chromatogram (TIC) of Rotofor fraction 4 (pH 9.0) prior to SPE clean-up is shown in Fig. 4A. A high background signal was observed early in the analysis. Within this region peaks arising from ampholytes were highly visible as shown in the inset of Fig. 4A. Furthermore, careful inspection of the data revealed the presence of peaks originating from the ampholytes in all regions of the TIC (data not shown). The same sample was reanalyzed after SPE clean-up (Fig. 4B). The TIC of this sample showed that, while the peptide peaks were unaffected, the background noise was greatly reduced, and no ampholyte-related peaks were detected. While the effectiveness of the SPE clean-up in removing urea was not monitored directly, the robustness of the LC–MS analysis provides an indirect measure of its efficacy. The base-peak chromatograms from yeast Rotofor fractions 3 (pH 9.5), 7 (pH 7.5), 12 (pH 6), and 17 (pH 4.5) analyzed by LC–MS are shown in Fig. 5. While the chromatograms differ due to the varying composition of each fraction, the overall resolution and efficiency of each is comparable. In addition, no clogging of the nanoelectrospray tip was observed during the analysis of all 20 Rotofor fractions, suggesting that minimal urea is likely to be present within any of the fractions. In total over 50 samples were run consecutively on the same column with no loss of overall resolution, and no signs of deposit build-up on the nanospray tip or the ion transfer capillary, both of which were observed prior to the SPE clean-up step.

4. Conclusions

An effective method to fractionate and process complex protein mixtures and analyze them by LC–MS at the peptide level was developed and tested. While the use of a liquid IEF (Rotofor) fractionation method provides many advantages such as increased sample loading, it does introduce contaminants such as urea and ampholytes that negatively impact the downstream LC–MS analysis. The procedure developed in this study incorporates a simple SPE clean-up procedure enabling the tryptically digested Rotofor fractions to be analyzed by reversed-phase LC–MS. The clean-up step prevents the ampholytes or urea from interfering with the MS analysis or impacting the chromatography. The introduction of the clean-up step between the orthogonal separation dimensions simplifies the overall procedure compared to previously published methods.

Acknowledgements

The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organization imply endorsement by the US Government. This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. NO1-CO-12400.

References

- [1] E. Celis, P. Gromov, Electrophoresis 10 (1999) 16–21.
- [2] J. Yates, Proteomics Workshop at HPCE 2001, Boston, MA, 2001.
- [3] P.H. O'Farrell, J. Biol. Chem. 250 (1975) 4007–4021.
- [4] A. Link, L. Hays, E. Carmack, J. Yates III, Electrophoresis 18 (1997) 1314–1334.
- [5] J. Harry, M.R. Wilkins, B.R. Herbert, N.H. Packer, A.A. Gooley, K.L. Williams, Electrophoresis 21 (2000) 1071–1081.
- [6] A. Shevchenko, O.N. Jensen, A.V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, A. Shevchenko, H. Boucherie, M. Mann, Proc. Natl. Acad. Sci. USA 93 (1996) 14440–14445.

- [7] S.P. Gygi, D.K.M. Han, A.-C. Gingras, N. Sonenberg, R. Aebersold, *Electrophoresis* 20 (1999) 310–319.
- [8] A. Shevchenko, M. Wilm, O. Vorn, M. Mann, *Anal. Chem.* 68 (1996) 850–858.
- [9] S.G. Gygi, G.L. Corthals, Y. Zhang, Y. Rochon, R. Aebersold, *Proc. Natl. Acad. Sci. USA* 97 (2000) 9390–9395.
- [10] R. Aebersold, D.R. Goodlett, *Chem. Rev.* 101 (2001) 269–295.
- [11] A. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J. Yates III, *Nat. Biotechnol.* 17 (1999) 676–682.
- [12] Y. Shen, R. Zhao, M.B. Belov, T.P. Conrads, G.A. Anderson, K. Tang, L. Pasa-Tolic, T.D. Veenstra, M.S. Lipton, H.R. Udseth, D.S. Smith, *Anal. Chem.* 73 (2001) 1766–1775.
- [13] J. Ding, P. Vouros, *Anal. Chem.* 69 (1997) 379–384.
- [14] A.H. Que, T. Konse, A.G. Baker, M.V. Novotny, *Anal. Chem.* 72 (2000) 2703–2710.
- [15] R.D. Smith, G.A. Anderson, M.S. Lipton, L. Pasa-Tolic, Y. Shen, T.P. Conrads, T.D. Veenstra, H.R. Udseth, *Proteomics* 2 (2002) 513–523.
- [16] M.P. Washburn, R. Ulaszek, C. Ceciu, D.M. Schieltz, J. Yates III, *Anal. Chem.* 74 (2002) 1650–1657.
- [17] H.J. Issaq, K.C. Chan, G.M. Janini, G.M. Muschik, *Electrophoresis* 20 (1999) 1533–1537.
- [18] M.T. Kachman, H. Wang, D.R. Schwartz, K.R. Cho, D.M. Lubman, *Anal. Chem.* 74 (2002) 1779–1791.
- [19] D.A. Wolters, M.P. Washburn, J. Yates III, *Anal. Chem.* 73 (2001) 5683–5690.
- [20] K. Wagner, T. Miliotis, G. Marko-Varga, R. Bischoff, K. Unger, *Anal. Chem.* 15 (2002) 809–820.
- [21] D.B. Wall, M.T. Kachman, S. Gong, R. Hindener, S. Parus, D.E. Misek, S.M. Hanash, D.M. Lubman, *Anal. Chem.* 72 (2000) 1099–1111.
- [22] J. Madoz-Gurpide, H. Wang, D.E. Misek, F. Brichory, S.M. Hanash, *Proteomics* 1 (2001) 1279–1287.
- [23] B. Herbert, P.G. Righetti, *Electrophoresis* 21 (2000) 3639–3648.
- [24] P.G. Righetti, A. Castagna, B. Herbert, *Anal. Chem.* 73 (2001) 320A–326A.
- [25] X. Zyo, D.W. Speicher, *Proteomics* 2 (2002) 58–68.
- [26] H.J. Issaq, T.P. Conrads, G.M. Janini, T.D. Veenstra, *Electrophoresis* 23 (2002) 3048–3061.
- [27] A. Ayala, J. Parrado, A. Machado, *Appl. Biochem. Biotechnol.* 69 (1998) 11–16.